



Caractérisation des variations climatiques du dernier millénaire à l'aide de nouvelles méthodes statistiques : reconstructions paléoclimatiques et intégration dans un modèle de circulation globale du climat

Simon Michel

► To cite this version:

Simon Michel. Caractérisation des variations climatiques du dernier millénaire à l'aide de nouvelles méthodes statistiques : reconstructions paléoclimatiques et intégration dans un modèle de circulation globale du climat. Climatologie. Université de Bordeaux, 2020. Français. NNT : 2020BORD0176 . tel-04246131

HAL Id: tel-04246131

<https://theses.hal.science/tel-04246131>

Submitted on 17 Oct 2023

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

En vue de l'obtention du grade de
DOCTEUR DE L'UNIVERSITÉ DE BORDEAUX EN SÉDIMENTOLOGIE
MARINE ET PALÉOCLIMATS

ED 304 : ÉCOLE DOCTORALE SCIENCES ET ENVIRONNEMENTS

présentée par

Simon MICHEL

**CARACTÉRISATION DES VARIATIONS CLIMATIQUES DU
DERNIER MILLÉNAIRE À L'AIDE DE NOUVELLES
MÉTHODES STATISTIQUES :
RECONSTRUCTIONS PALÉOCLIMATIQUES ET
INTÉGRATION DANS UN MODÈLE DE CIRCULATION
GLOBALE DU CLIMAT**

Travaux réalisés au laboratoire Environnements et Paléoenvironnements Océaniques et
Continentaux sous la direction de Didier SWINGEDOUW

Soutenue le 13 Novembre 2020 devant le jury composé de :

Xavier CROSTA	DR CNRS	EPOC, Pessac	Président
Pascal YIOU	DR CEA	LSCE/IPSL, Gif-sur-Yvette	Rapporteur
Christophe CASSOU	DR CNRS	CERFACS, Toulouse	Rapporteur
Pascale BRACONNOT	DR CEA	LSCE/IPSL, Gif-sur-Yvette	Examinatrice
Hugues GOOSSE	DR FNRS	SST/ELI/ELIC, Louvain	Examineur
Joël GUIOT	DR CNRS	CEREGE, Aix-en-Provence	Examineur
Didier SWINGEDOUW	CR1 CNRS	EPOC, Pessac	Directeur de thèse

Remerciements

Mes premiers remerciements vont évidemment à mon directeur de thèse Didier Swingedouw. Outre la grande qualité de son accompagnement scientifique, il fut un directeur de thèse très présent, attentionné, motivant, patient et ne m’a jamais pris de haut. Si cette thèse a pu parfois prendre du retard et tourner en rond ce n’est en aucun cas de sa faute et il aura même été là pour me re-motiver, me redonner de l’élan et m’encourager dans les moments où j’en avais besoin. Mes manières de voir le monde et de penser les choses ont considérablement évolué pendant ces trois ans, et je le dois en grande partie à l’esprit critique et scientifique qui m’a été inculqué par Didier. Au-delà de la science, Didier est humainement quelqu’un d’admirable, ouvert, drôle et toujours de bonne humeur, si bien que je le considère aujourd’hui comme un ami. Merci encore pour ces trois années.

Je tiens également à remercier Juliette Mignot, Myriam Khodri et Pablo Ortega, des personnes avec qui j’ai pu collaborer pour mes productions scientifiques et qui ont chacun passé un temps considérable pour m’apporter leur expertise et des conseils très avisés. J’appuie notamment mes remerciements à Juliette, que je peux considérer comme ma co-directrice de thèse officieuse, puisqu’elle aura su réserver une part considérable de son temps pour m’aider et me conseiller tout le long de ma thèse.

Je souhaite également remercier Joël Guiot, Hugues Goosse et Xavier Crosta pour avoir accepté d’être membres des deux comités du suivi de ma thèse. Ils se sont montrés patients, intéressés et encourageants, et c’était précieux. En particulier, le deuxième comité de thèse m’a permis de prendre confiance en mes résultats et de me donner un nouvel élan pour la fin de ma thèse.

Je remercie également l’ensemble du personnel scientifique et non scientifique de l’OASU, constitué de personnes intéressantes, ouvertes, agréables et sympathiques, ce qui est idéal et nécessaire pour réaliser une thèse dans de bonnes conditions. Bien que je ne puisse passer en revue tout le monde, j’adresse des remerciements à des personnes comme Frédérique Eynaud, Phillippe Martinez, Jacques Giraudeau, Marion Devillers ou Giovanni Sgubin, avec qui j’ai pu passer des pauses cafés aux longues discussions passionnantes pendant ces trois années. J’ai également une pensée toute particulière pour mon *ex-dream team* de bureau à savoir Kelly, Caly et Stan, qui sont devenus d’excellents amis et qui sont des personnes que j’apprécie énormément. J’ai passé de super moments avec vous : se faire des bouffes au *Carpe’*, des discussions sans fin, les soirées qu’on a pu passer ensemble, et tout le reste. Je remercie également les membres de mon *ex-ex-team* de bureau, Eleanore et Brieuc, lorsque j’étais en première année, qui ont su me mettre à l’aise et m’intégrer alors que la tâche était loin d’être facile. J’adresse également des remerciements à mes partenaires de *foodtruck* préférés, à savoir Audrey, Polo, et Léa, qui sont également devenus des amis et des personnes que j’apprécie beaucoup. Je remercie également Maëva, arrivée dans mon bureau en septembre, agréable à vivre et sympathique, à qui j’adresse également le maximum d’encouragements pour la suite de sa thèse. Je remercie également Jian, arrivé récemment, qui m’a laissé collaborer avec lui pour sa superbe première publication au sein du laboratoire. Je le remercie par ailleurs pour m’avoir payé un excellentissime restaurant typique chinois pour célébrer sa publication. Je remercie le reste des doctorants du laboratoire qui sont tous des personnes que j’apprécie. Je remercie de manière générale l’ensemble des acteurs du

laboratoire EPOC pour son dynamisme et son professionnalisme. J'adresse également des remerciements à Pierre-Antoine Dessandier, ancien d'EPOC, qui m'a pris sous son aile aux EGU 2018 et 2019 et m'a fait passer deux super semaines de colloques scientifiques, que j'appréhendais tant en début de thèse.

J'adresse aussi des remerciements au personnel de l'école doctorale Sciences et environnements qui sont des personnes très professionnelles et pleinement dévouées à la réussite de chaque doctorant.

J'adresse des remerciements particuliers à mon pote Paulin avec qui j'ai pu régulièrement décompresser après le travail : "rdv victoire 18h". Je le remercie aussi pour les 5 années pré-doctorat pour lesquelles on aura systématiquement été dans la même promotion. En plus de Paulin, je remercie également Julien, Camille, et le/la 5^{ème} joueur/se de tarot variable pour ces nombreuses parties qu'on a pu se faire pendant ces trois ans, en espérant pouvoir s'en refaire une prochainement.

Je remercie bien évidemment tous mes *bros* de Niort pour les moments que j'ai passé et passe avec eux, les amis pour la vie, que j'ai pu plus ou moins voir ces dernières années, mais qui seront toujours dans mon cœur : Roch, Pipeu, Pant'air, Julo, Ju, Nouma, Filex, Max, Dodo, Nono, Masgnaux, Péter, Nanar, Rara. Je remercie toutes les personnes qui entourent ce petit monde que j'apprécie énormément mais que je ne pourrais malheureusement pas tous citer : Angélique, Chacha, Camille F, Camille, Lolo, Pépé, Mamar, Roro, Anouk, Pierre, Théo, Léa, Vandier, Jules, et tous les autres.

J'ai une pensée particulière pour Manon, Antoine et Popo qui sont des amis très chers que je n'ai malheureusement pas pu beaucoup voir ces trois dernières années. Je les remercie cependant de se rendre disponible chaque fois que je suis de passage à Paris et de me contacter chaque fois qu'ils passent à Bordeaux.

Mes derniers et plus intimes remerciements vont à ma famille que j'ai une chance incommensurable d'avoir et que j'ai la chance de voir régulièrement. Je remercie tout d'abord ma soeur, Zoé, et mes frères, Elie et Sacha, qui sont des personnes admirables et pour qui je souhaite le meilleur dans les années qui viennent. Je remercie en particulier mes parents, qui ont été et seront toujours là pour moi. En plus d'être présent dans les moments les plus difficiles de ma thèse ils m'ont également hébergé pendant toute la période de confinement et après, ce qui sinon aurait rendu la fin de ma thèse extrêmement compliqué. Je remercie également mon beau frère, Polo, pour les moments que j'ai pu passer ces trois dernières années avec lui, comme à Rennes l'été dernier par exemple. Je remercie par ailleurs mes nièces, Mona et Prune, pour m'apporter un peu de simplicité et de joie enfantine dans ma vie. Pour finir, je remercie mon chat pour sa compagnie pendant toute la rédaction de ma thèse, et qui est de loin l'être avec qui j'ai passé le plus de temps ces derniers mois.

Résumé

La variabilité climatique joue un rôle considérable dans de nombreux domaines de nos vies. L'augmentation globale des températures, forcée par les émissions croissantes de gaz à effet de serre d'origine anthropique, a connu une période de hiatus entre 1998 et 2012. Les causes de cette séquence font l'objet de controverses scientifiques concernant les rôles relatifs de la variabilité intrinsèque au système climatique, et des forçages externes. Afin de résoudre ce type de débat, il est nécessaire de mieux comprendre les modes de variabilité décennaux, issus de l'organisation à grande échelle du système climatique. Cependant, 150 années d'observations instrumentales ne permettent pas d'estimer de manière rigoureuse leurs caractéristiques et donc de comprendre leur dynamique.

Cette thèse cherche à parfaire la compréhension de la variabilité climatique multi-décennale en produisant des reconstructions sur le dernier millénaire d'indices ou de grilles climatiques. Pour cela, on se propose d'introduire des méthodes statistiques modernes, dites de *machine learning*, appliquées à des données paléoclimatiques issues d'archives naturelles (cernes d'arbres, carottes de glace...). La base de données PAGES 2K récemment développée propose plus de 700 de ces enregistrements sur le dernier millénaire et sera abondamment utilisée pour construire les modèles statistiques proposés.

Des métriques mathématiques objectives permettent de montrer que l'application d'une méthode non-linéaire telle que la forêt aléatoire produit en général des résultats plus robustes que les méthodes linéaires usuelles. On utilise donc la forêt aléatoire pour reconstruire des variations du mode principal des températures de surface océanique (SST) de l'Atlantique Nord, l'*Atlantic Multidecadal Variability* (AMV), notamment liée à la circulation océanique de retournement dans ce bassin (AMOC). Cette reconstruction suggère que des changements dans cette circulation, survenus à la fin du XII^e siècle, furent probablement le catalyseur d'une entrée précoce dans le petit âge de glace, une période relativement froide du dernier millénaire. Les importantes activités volcaniques des XIII^e, XV^e et XIX^e siècles sont cependant identifiées comme les causes principales de la prolongation du petit âge de glace dont le paroxysme se situe au XIX^e siècle. De plus, des changements récents dans la variabilité de l'AMV semblent, d'après la théorie des systèmes dynamiques, être un signal d'approche d'un point de bascule de l'AMOC.

Une reconstruction des SST sur une grille a également été produite par forêt aléatoire sur tout le millénaire. Cette reconstruction montre que les phases positives de l'AMV auraient tendance à conduire, environ une décennie après, à une phase négative du mode principal des SST du Pacifique, l'*Interdecadal Pacific Variability* (IPV). Comme les observations montrent que le hiatus de 1998-2012 correspond à une phase négative de l'IPV, ces résultats soutiennent l'hypothèse que l'AMV est l'un des moteurs principaux de telles périodes de hiatus. Dans la lignée des hypothèses de ralentissement de l'AMOC déduites du spectre de l'AMV reconstruite, cette grille suggère que celle-ci atteint aujourd'hui un niveau bas sans précédent, depuis au moins l'an 850. Enfin, cette même grille de reconstruction a été assimilée par méthode de rappel en SST au sein du modèle de circulation globale IPSL-CM5A2-LR. Bien que la simulation reproduit convenablement les températures et paléo-températures mesurées, la méthode de rappel semble moins efficace que des méthodes d'assimilation plus complexes tel que le filtre de Kalman. De plus, l'AMOC simulée dans cette réanalyse ne correspond pas à son estimation à partir de méthodes statistiques, sans doute du fait que la salinité est déterminante pour forcer la convection océanique et l'AMOC. Ainsi, assimiler la salinité semble nécessaire pour reproduire correctement les variations d'AMOC sur le dernier millénaire à l'aide d'un modèle.

Abstract

Climate variability strongly impacts our lives in many ways. The observed growth in global temperatures, forced by anthropogenic greenhouse gas emissions, has experienced a period of hiatus between 1998 and 2012. Causes at the origin of such a sequence are subject to intensive scientific controversies, notably concerning the relative roles of climate variability intrinsic to the climate system and external radiative forcings. To address these questions, it is necessary to better understand decadal modes of variability resulting from the global scale organization of the climate system. However, the only 150 years of instrumental observations are not enough to rigorously characterise their nature, variability, and dynamics.

This thesis aims at improving our knowledge of multidecadal climate variability by producing reconstructions of climate indices and grids over the last millennium. For doing so, up-to-date statistical methods are used, namely machine learning technics, which are applied to paleoclimate data coming from natural archives (tree rings, ice cores...). The recently developed PAGES 2k database provides more than 700 of these records, and will thus be abundantly used for training the proposed statistical models.

Objective mathematical metrics show that a nonlinear technique, namely the random forest, generally produces more robust results than the usual linear technics. We thus use the random forest method to reconstruct variations of the preferential mode of North Atlantic sea surface temperatures (SST) de l'Atlantique Nord, namely the Atlantic Multidecadal Variability (AMV), which is notably related to the Atlantic Meridional overturning circulation (AMOC). This reconstruction suggests that changes in this circulation, which occurred at the end of the 12th century, has probably been the catalyzer of early onset of the little ice age, a relatively cold period of the last millennium. The strong volcanic activities from the 13th, 15th, and 19th centuries have nevertheless been identified as the main causes of the little ice age extension, which reached its climax during the 19th century. Moreover, according to dynamical systems theory, recent changes in the AMV variability constitute an early warning signal of an incoming tipping point in the AMOC.

A gridded SST reconstruction over the last millennium has also been produced using the random forest method. This reconstruction shows that positive phases of AMV are likely to drive, after about a decade, a negative phase of the preferential mode of Pacific SST, namely the Interdecadal Pacific Variability. Since observation show that the 1998-2012 hiatus period exactly matches with a long term negative phase of the AMV, these results suggest that the AMV is one of the main drivers of such hiatus periods. In line with the assumption of an AMOC weakening deduced from spectral properties of the reconstructed AMV this grid also suggest that the AMOC has reached an unprecedentedly low level of intensity, since at least 850 C.E. Finally, the same gridded reconstruction has been assimilated using SST nudging method within the IPSL-CM5A2-LR general circulation model. While the reanalysis fairly reproduces the observed temperatures and paleo-temperatures, the nudging method seems to be less efficient than more complex data assimilation technics, such as Kalman filtering. Moreover, the simulated AMOC in this reanalyses does not correspond to the one obtained from statistical reconstructions, which might be because salinity largely influences ocean surface layers densities, and thus to force oceanic convection and the AMOC. Assimilating surface salinity data thus appears to be essential to correctly reproduce the last millennium AMOC variability using a general circulation model.

Liste des sigles et acronymes

- **AMOC** : *Atlantic Meridional Overturning Circulation* - Circulation méridienne de retournement de l'Atlantique (acronyme)
- **AMV** : *Atlantic Multidecadal Variability* - Variabilité multi-décennale de l'Atlantique (sigle)
- **CART** : *Classification And Regression Trees* - Arbres de classification et de régression (acronyme)
- **CCSM** : *Community Climate System Model* - Nom d'un modèle de circulation globale hors version (sigle)
- **CE** : *Coefficient of Efficiency* - Coefficient d'efficacité (sigle)
- **CESM** : *Community Earth System Model* - Nom d'un modèle de circulation globale hors version (sigle)
- **ClimIndRec** : *Climate Index Reconstruction* - Nom d'un programme informatique de reconstruction (acronyme)
- **CMIP** : *Coupled Model Intercomparison Project* - Projet d'inter-comparaison de modèles couplés (acronyme)
- **CNRM** : Centre National de Recherche en Météorologie (sigle)
- **EKF** : *Ensemble Kalman Filter* - Filtre de Kalman d'ensemble (sigle)
- **Enet** : *Elastic net regression* - Régression *elastic net* (acronyme)
- **ENSO** : *El Niño Southern Oscillation* - El Niño-oscillation australe (acronyme)
- **EOF** : *Empirical Orthogonal Function* - Fonction orthogonale empirique (sigle)
- **GCM** : *General Circulation Model* - Modèle de circulation générale (sigle)
- **GIEC** : Groupe d'Experts Intergouvernemental sur l'Évolution du Climat (acronyme)
- **GISS** : *Goddard Institute for Space Studies* - Nom d'une laboratoire de recherche américaine (acronyme)
- **Had** : *Hadley Center for Climate Prediction and Research* - Centre de recherche britannique (acronyme)
- **INCA** : INtéracTion avec la Chimie et les Aérosols - Nom du modèle de chimie troposphérique d'IPSL-CM5 (acronyme)
- **IPSL** : Institut Pierre-Simon-Laplace (sigle)
- **IPSL-CM** : **IPSL Climate Model** - Nom du modèle de climat hors version de l'IPSL (sigle)

- **IPV** : *Interdecadal Pacific Variability* - Variabilité inter-décennale du Pacifique (sigle)
- **ITCZ** : *InterTropical Convergence Zone* - Zone de convergence intertropicale (sigle)
- **KF** : *Kalman Filter* - Filtre de Kalman (sigle)
- **LIA** : *Little Ice Age* - Petit âge de glace (sigle)
- **LIM** : *Louvain-la-neuve sea Ice Model* - Nom du modèle glace de mer de NEMO (acronyme)
- **LMLE** : *Last Millennium Large Ensemble* - Large ensemble du dernier millénaire (sigle)
- **LMD** : Laboratoire de Météorologique Dynamique (sigle)
- **LMDz** : Nom du modèle de circulation globale atmosphérique hors version d'IPSL-CM5 (sigle)
- **LV** : *Latent Variable* - Variable latente (sigle)
- **MCA** : *Medieval Climate Anomaly* - Anomalie climatique médiévale (sigle)
- **MOC** : *Meridional Overturning Circulation* - Circulation méridienne de retournement (acronyme)
- **NAO** : *North Atlantic Oscillation* - Oscillation Atlantique Nord (sigle)
- **NASST** : *North Atlantic Sea Surface Temperatures* - Températures de la surface océanique de l'Atlantique Nord (sigle)
- **NCAR** : *National Center for Atmospheric Research* - Nom d'un centre de recherche américain (sigle)
- **NEMO** : *Nucleus for European Modelling of the Ocean* - Nom du modèle de circulation globale océanique d'IPSL-CM5 (acronyme)
- **NOAA** : *National Oceanic and Atmospheric Administration* - Nom d'une agence d'étude américaine (acronyme)
- **OMM** : Organisation Météorologique Mondiale (sigle)
- **OPA** : Océan PARallélisé - Nom du modèle de dynamique océanique de NEMO (sigle)
- **ORCA** : Nom de la configuration de grille hors version de NEMO (acronyme)
- **ORCHIDEE** : *Organic Carbon and Hydrology in Dynamic Ecosystems* - Nom du modèle de végétation continentale d'IPSL-CM5 (acronyme)
- **PAGES** : *PAst Global ChangES* - Nom d'un consortium scientifique (acronyme)

- **PC** : *Principal Component* - Composante Principale (sigle)
- **PCA** : *Principal Component Anaysis* - Analyse en composantes principales (sigle)
- **PCR** : *Principal Component Regression* - Régression en composantes principales (sigle)
- **PF** : *Particle Filter* - Filtre à particules (sigle)
- **PISCES** : *Pelagic Interactions Scheme for Carbon and Ecosystems Studies* - Nom du modèle de biogéochimie marine de NEMO (acronyme)
- **PLS** : *Partial Least Squares* - Moindres carrés partiels (sigle)
- **PPE** : *PseudoProxy Experiment* - Expérience *pseudo-proxy* (sigle)
- **RAPID** : *RAPID climate change* - Nom d'un projet de recherche international (acronyme)
- **RE** : *Reduction of Error* - Réduction d'erreur (sigle)
- **REPROBUS** : *REactive Processes Ruling the Ozone in the Stratosphere* - Nom du modèle de chimie stratosphérique d'IPSL-CM5 (acronyme)
- **RF** : *Random Forest* - Forêt aléatoire (sigle)
- **SAT** : *Surface Air Temperatures* - Température atmosphérique de surface (sigle)
- **SLP** : *Sea Level Pressure* - Pression atmosphérique de surface (sigle)
- **SPG** : *SubPolar Gyre* - Gyre subpolaire (sigle)
- **SROCC** : *Special Report on the Ocean and Cryosphere* - Nom d'un rapport spécial du GIEC (acronyme)
- **SSS** : *Sea Surface Salinity* - Salinité de la surface océanique (sigle)
- **SST** : *Sea Surface Temperatures* - Températures de la surface océanique (sigle)
- **T2m** : *Tempeature at 2 meters* - Température atmosphérique à 2 mètres d'altitude (sigle)
- **TCL** : *Théorème Central Limite* (sigle)
- **TSI** : *Total Solar Irradiance* - Irradiance solaire totale (sigle)
- **VA** : *Variational Approach* - Approche variationnelle (sigle)
- **ZB** : *Zeta Bytes* - Nom d'une unité de mesure de volume de données (sigle)

Table des figures

1.1	Schéma du système climatique et échelles de temps sur lesquelles ses composantes évoluent	1
1.2	Illustration schématique de l'effet de la rotondité de la Terre sur la quantité d'énergie reçue par unité de surface selon la latitude	2
1.3	Illustration schématique de l'effet de l'obliquité de la Terre sur la répartition saisonnière de l'énergie solaire	2
1.4	Représentation schématique de la position de l'ITCZ selon la saison .	3
1.5	Représentation schématique de la circulation atmosphérique globale	4
1.6	Schéma très simplifié de la circulation thermohaline	6
1.7	Représentation schématique des composantes de l' <i>Atlantic Meridional Overturning Circulation</i> (AMOC) selon une coupe à 26°N	7
1.8	Représentation du bilan radiatif terrestre	8
1.9	Peinture de William Turner - <i>Dido building Carthage</i> (1816)	10
1.10	Peinture de William Ascroft - <i>Sans nom. Une de ses nombreuses productions entre 1883 et 1884</i>	10
1.11	Couche de cendres volcanique visible observée dans une carotte de glace prélevée au Groenland (à droite de la carotte)	12
1.12	Observation des anomalies de SST de l'Atlantique Nord relatives à la période 1870-2019	14
1.13	Indices AMV calculés sur la période 1870-2019	17
1.14	Représentation schématique des effets de la NAO sur les régions adjacentes de l'Atlantique Nord	18
1.15	Représentation schématique et simplifiée de la cellule de Walker du Pacifique lorsque le phénomène ENSO est en phase neutre	19
1.16	Effets des variations de l'IPV sur la SST du Pacifique	20
1.17	Importance relative de chaque source d'incertitudes dans la prévision décennale de la moyenne globale de la température atmosphérique .	21
1.18	Importance relative de chaque source d'incertitudes dans la prévision décennale de la moyenne globale et d'une moyenne régionale de la température atmosphérique	22
1.19	Moyenne globale des anomalies de température observées entre 1850 et 2017	23
1.20	Estimations de la concentration de CO ₂ atmosphérique moyenne . .	26
1.21	Reconstructions des forçages volcaniques et solaires sur le dernier millénaire	28
1.22	Caractéristiques spatio-temporelles des enregistrements paléoclimatiques de la base de données PAGES 2k (PAGES 2k Consortium 2017)	29
2.1	Terminologie d'un arbre computationnel quelconque	59
2.2	Schéma de la phase de reconstruction d'un modèle de forêt aléatoire	61
3.1	Scheme summarizing the main features of ClimIndRec	69

3.2	Scheme of the initial data	73
3.3	Scheme of a K-fold cross-validation procedure to select the optimal control parameter of a specific learning method $\mathcal{M}$	75
3.4	Scheme of the whole procedure for scores calculation for a given method $\mathcal{M}$	75
3.5	Boxplot of CE scores obtained for the four methods and different groups of proxy records by reconstructing the NAO index of the period 1000–1970 with $R = 50$ training/testing samples	78
3.6	Boxplot of CE scores for different reconstruction periods	79
3.7	Comparison of different NAO reconstructions from this study and Ortega et al. 2015	80
3.8	Comparison of reconstructions from this study with the original Jones et al. (1997) NAO index (Purple line) over their common period	81
3.9	P-values obtained from Shapiro-Wilk normality tests on the residuals from each reconstruction of Fig 3.7	82
3.10	Map of the 46 proxy records used for the reconstruction of the NAO index from Jones et al. (1997) on the time window 1000-1972 using the RF method	82
3.11	Superposed epoch analysis of the NAO response from two years (N-1) before to five years after (N+4) to the largest volcanic eruptions used by Ortega et al. (2015) (10 eruptions) and the 11 largest from Sigl et al. (2015).	84
4.1	Climatic impacts of the AMV over the historical era	99
4.2	Nested reconstruction of the AMV and related proxies	101
4.3	Pseudo-Proxy validation	102
4.4	Comparison with NASST and early warning signal test	104
4.5	Extended Data Figure 1 - Estimated forced component of historical NASST	106
4.6	Extended Data Figure 2 - Score by level of inputs for the 496 reconstructions	107
4.7	Extended Data Figure 3 - NASST and AMV response to the 10 largest volcanic eruptions of the last millennium	108
5.1	CE score by time-frame of reconstruction	124
5.2	Proxy records' types and temporal availability	125
5.3	Score, uncertainties and standard deviations for 4 oceanic basins Densities by grid points of 3 different metrics for North Atlantic, South Atlantic, Indian and Pacific basins : (a) Median CE score; (b) Standard deviation; (c) Ratio between regression standard error and standard deviation.	125
5.4	Global SST anomalies, solar minima and large volcanic eruptions	126
5.5	Spatial response of SST to large volcanic eruptions	127
5.6	Spatial response of SST to grand solar minima	129
5.7	Optimal global-scale contrasting periods for SST over the last millennium	130
5.8	Definition of the areas over which climate indices are calculated	131
5.9	Analysis of the Pacific	132
5.10	PCA of the North Atlantic	133

5.11	Time series and regression maps (in °C per unitless PC value) for two candidate indices for AMOC	134
5.12	Cross-correlations between AMV_{TS} and IPV_H (black) and $PC2_{NA}$. . .	136
5.13	Study of the link between AMV and IPV	137
6.1	Schéma de la structure du modèle IPSL-CM5 dans sa "configuration pleine", c'est-à-dire lorsque tous les modèles qui le composent sont actifs (Dufresne et al. 2013)	146
6.2	Validation de la réanalyse	148
6.3	Étude de la cohérence entre deux réanalyses distinctes	149
6.4	Cartes de corrélation entre les observations instrumentales et les réanalyses <i>Rap.</i> et <i>KFH16</i>	150
6.5	Comparaison des moyennes globales issues des deux réanalyses . . .	151
6.6	Corrélation entre les enregistrements paléoclimatiques existants et les réanalyses étudiées	152
6.7	Comparaison d'estimations des variations de l'AMOC sur le dernier millénaire	153

Liste des tableaux

3.1	Table of correlations between five NAO reconstructions	81
3.2	Proxy records not in PAGES 2k	87
3.3	Ten largest volcanic eruptions from Ortega et al. 2015	88
3.4	Ten largest volcanic eruptions from Sigl et al. 2015	89
3.5	Variance de la NAO reconstruite	90
3.6	R packages used in ClimIndRec	91
4.1	Extended Data Table 2 - List of the 37 models used for the estimation of the NASSTs forced component	109
4.2	Extended Data Table 3 - 10 last millennium largest volcanic eruptions (Sigl et al. 2015)	110
5.1	Eleven largest volcanic eruptions of the last millennium and location (Sigl et al. 2015)	127
5.2	Largest solar minima of the last millennium (Lean 2018)	128

Table des matières

1	Introduction générale	1
1.1	Contexte et pré-requis	1
1.1.1	Définitions générales	1
1.1.2	Fonctionnement général du système océan-atmosphère	2
1.1.3	Forçages radiatifs externes et variabilité climatique interne . .	8
1.1.4	Outils d'étude du climat	11
1.2	Modes de variabilité et implication pour la prévision décennale . . .	13
1.2.1	Modes de variabilité de l'Atlantique Nord	14
1.2.2	Modes de variabilité du Pacifique	19
1.2.3	Intérêt pour la prévision décennale	21
1.3	Le dernier millénaire comme outil de compréhension de la variabilité décennale	25
1.3.1	Une période idéale pour comprendre la variabilité naturelle et ses impacts	25
1.3.2	Deux séquences climatiques bien distinctes	27
1.3.3	Une base de données sans précédent dédiée à la reconstruc- tion paléoclimatique	29
1.3.4	<i>Machine learning</i> : de nouvelles opportunités pour la recons- truction millénaire	30
1.3.5	Assimilation de données dans un contexte paléoclimatique .	31
1.4	Problématiques et objectifs de ces travaux de thèse	32
1.4.1	Problématiques soulevées	32
1.4.2	Organisation de la thèse	33
2	Description des méthodes et outils utiles à la compréhension des travaux	35
2.1	Analyse de séries temporelles	35
2.1.1	Moyenne, variance, covariance et corrélation empiriques . . .	35
2.1.2	Stationnarité, autocovariance, autocorrélation et corrélation croi- sée	36
2.1.3	Analyses spectrales	37
2.2	Méthodes d'analyses statistiques pour les données climatiques	38
2.2.1	Analyse en composantes principales et fonctions orthogonales empiriques	38
2.2.2	Régression linéaire	40
2.3	Calculs de significativité statistique	42
2.3.1	Test de Student de significativité d'une moyenne	44
2.3.2	Test de Student de comparaison de deux moyennes	46
2.3.3	Test de Student corrigé de significativité d'une corrélation . .	48
2.3.4	Test du coefficient de pente d'un modèle de régression linéaire simple	49
2.3.5	Méthodes de Monte-Carlo et exemple d'application	51

2.4	Méthodes de <i>machine learning</i>	52
2.4.1	Contexte théorique	52
2.4.2	PCR	54
2.4.3	Moindres carrés partiels	55
2.4.4	Régression Elastic Net	56
2.4.5	Régression par forêt aléatoire	58
2.4.6	Optimisation de paramètres de contrôle par validation croisée	61
2.4.7	Évaluation d'un modèle par <i>hold-out</i>	62
3	Développement d'un outil de reconstruction robuste et automatique	65
3.1	Reconstruction par apprentissage statistique	65
3.2	Article publié dans <i>Geoscientific Model Development</i>	67
3.2.1	Introduction	67
3.2.2	Data, notations and methodologies	70
3.2.3	Results	77
3.2.4	Discussion and conclusion	83
3.2.5	Supplementary informations	87
3.3	Devenir de ClimIndRec : perspectives et améliorations possibles	93
4	Reconstruction des fluctuations de l'AMV : implications pour son évolution passée et future	95
4.1	L'AMV comme mode de variabilité interne : une stratégie de reconstruction particulière	95
4.2	Article soumis au journal <i>Nature Geoscience</i>	97
4.2.1	Robust and validated AMV reconstruction	99
4.2.2	An early internally-driven onset of LIA	103
4.2.3	Extended data figures and tables	106
4.2.4	Methods	111
4.3	Implications des résultats obtenus	114
5	Reconstruction d'une grille de SST basée sur une méthode de forêt aléatoire	117
5.1	Objectifs et motivations de ce chapitre	117
5.2	Article soumis au journal <i>Climate Dynamics</i>	118
5.2.1	Introduction	119
5.2.2	Method : Random forest-based gridded SST reconstruction applied to PAGES2K	122
5.2.3	Results	124
5.2.4	Discussion	138
5.3	Intérêts apportés par les résultats d'une telle étude	141
6	Rappel vers la SST reconstruite sur 1000 ans dans le modèle IPSL-CM5A2	145
6.1	Méthode d'assimilation de données : application simple dans un contexte paléoclimatique	145
6.2	Données et méthode	146
6.2.1	Le modèle IPSL-CM5A2-VLR comme base de simulation	146
6.2.2	Application de la méthode de rappel	147
6.3	Validation et comparaison avec une étude précédente	148
6.4	Comparaisons de différentes estimations millénaires de l'AMOC	152
6.5	Perspectives de ce chapitre	153

7	Conclusions, critiques et perspectives	157
7.1	Résumé des résultats présentés	157
7.2	Regard critique et réflexions sur les travaux présentés	162
7.3	Perspectives	166

« *Le grand ouvrier de la nature est le temps [...]* »

Georges-Louis Leclerc de Buffon (1756),
Histoire naturelle, vol. 6, p. 60

Chapitre 1

Introduction générale

1.1 Contexte et pré-requis

1.1.1 Définitions générales

Le climat se définit comme la description statistique en terme de moyenne et de variance de variables météorologiques appropriées (température, précipitation, pression atmosphérique...) sur une période allant du mois aux millions d'années (Groupe d'Experts Intergouvernemental sur l'Evolution du Climat, GIEC 2014). L'Organisation Météorologique Mondiale (OMM) propose par convention que la période classique sur laquelle on décrit le climat soit de 30 ans. On peut relever le caractère anthropocentré de cette notion de climat étant donné que la période de référence correspond environ à celle d'une génération humaine, et que les valeurs des variables considérées pour le définir sont celles observées en surface. Bien qu'arbitraire, cette définition permet de déterminer statistiquement ce que l'on considère comme changements climatiques (Paillard 2008).

Le système climatique (Fig. 1.1), dont l'évolution implique entre autres les variations du climat comme ci-avant définit, se caractérise comme un ensemble de composantes, ainsi que leurs interactions : l'atmosphère, l'hydrosphère (océans, rivières, lacs...), la biosphère (humains, animaux, végétaux...), la cryosphère (banquise, glaciers...) et la lithosphère (croûte terrestre). Ces différentes composantes possèdent chacune une variabilité intrinsèque sur des échelles de temps spécifiques (Fig. 1.1) et

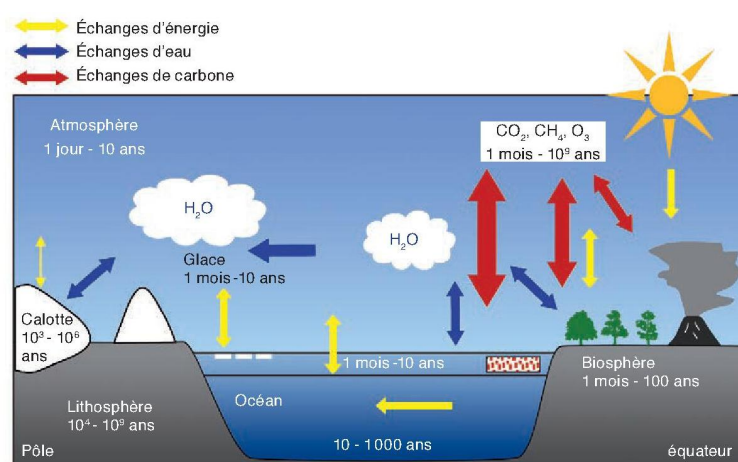


FIGURE 1.1 | Schéma du système climatique et échelles de temps sur lesquelles ses composantes évoluent.

échantent en permanence de l'énergie, de l'eau et du carbone. L'énergie que ces différentes composantes échantent provient principalement du rayonnement solaire. Ainsi, si ce flux constant d'énergie venait à s'arrêter, le système climatique deviendrait quasi-immobile.

L'océan et l'atmosphère sont les deux principales composantes du système climatique. L'atmosphère est le fluide dont les propriétés météorologiques de surface impactent directement la vie terrestre. L'océan, quant à lui, régule les conditions atmosphériques de par des échanges radiatifs, mécaniques et de matière continus avec l'atmosphère.

1.1.2 Fonctionnement général du système océan-atmosphère

Cette section ne se veut pas exhaustive en matière de description du système océan-atmosphère. Elle vise la description des mouvements des deux fluides dans leurs grandes lignes et à grande échelle, et permet ainsi de définir des notions générales, utiles à la compréhension des travaux qui seront présentés par la suite.

De par la rotondité et l'inclinaison de la Terre, l'énergie provenant du rayonnement électromagnétique solaire est répartie de manière inéquitable selon la latitude (Fig. 1.2) et selon les saisons (Fig. 1.3). Cette distribution différentielle de l'énergie dans l'espace et dans le temps met en mouvement l'atmosphère et l'océan via les gradients de densité. Ces mouvements induisent un transport de chaleur méridien vers les pôles, limitant ainsi les effets de la distribution hétérogène de l'énergie solaire. Le système couplé composé de l'océan et l'atmosphère joue un rôle crucial dans la redistribution de cette énergie. Ceci se fait soit par les vents, soit par les courants marins qui transportent les eaux équatoriales vers les pôles, en réponse au gradient de température méridien, limitant ainsi l'effet de la sphéricité de la Terre.

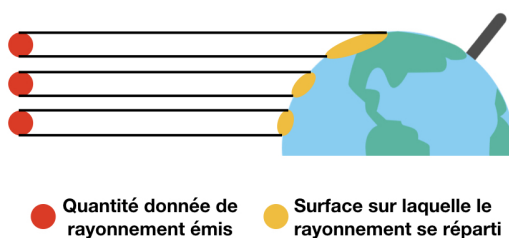


FIGURE 1.2 | Illustration schématique de l'effet de la rotondité de la Terre sur la quantité d'énergie reçue par unité de surface selon la latitude. Les points rouges représentent chacun une même quantité de rayonnement solaire émis. Les aires jaunes désignent la surface sur laquelle une quantité d'énergie solaire se répartit à la surface du globe.

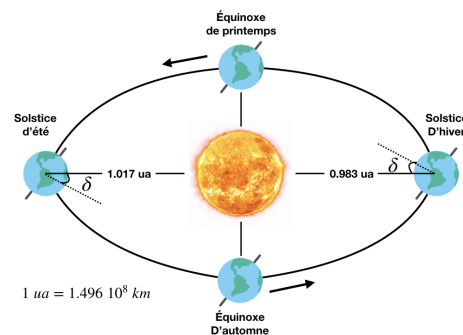


FIGURE 1.3 | Illustration schématique de l'effet de l'obliquité de la Terre sur la répartition saisonnière de l'énergie solaire. *u.a.* désigne l'unité astronomique, qui correspond à la distance moyenne entre la Terre et le soleil. δ correspond à l'angle de déclinaison solaire, c'est-à-dire, l'angle qu'il fait avec l'équateur géographique, qui varie environ entre -23.5° et 23.5° . Source : schéma inspiré de la figure de Wald (2007).

Dans le cas de l'atmosphère, les gradients de densité entre les masses d'air sont déterminés par leur température et leur pression atmosphérique. La pression atmosphérique en un point d'altitude donné se traduit comme l'intensité de la force exercée par la colonne d'air se situant au dessus de ce point. Cette pression peut donc être altérée par les gaz à concentrations variables qui constituent une masse d'air, comme la vapeur d'eau par exemple. Ainsi, si l'humidité (et donc la concentration en vapeur d'eau) d'une masse d'air augmente, alors sa masse, et donc sa densité, diminuent en conséquence. Dans le cas de l'océan, la densité des masses d'eau est principalement caractérisée par leur température et leur salinité (Rahmstorf 2002).

Au delà des gradients de densité, qui mettent en mouvement l'océan et l'atmosphère, les deux fluides sont également soumis à la force de Coriolis. Il s'agit d'une force inertielle qui s'applique perpendiculairement à la tangente de la trajectoire de tout corps en mouvement dans un référentiel en rotation. Pour le cas du sphéroïde qu'est la Terre, ceci implique que cette force est nulle à l'équateur et s'accroît selon la latitude. Elle s'applique vers la gauche du mouvement dans l'hémisphère sud et vers la droite du mouvement dans l'hémisphère nord. La compensation de cette force avec celle des gradients de pressions horizontaux et verticaux, résulte en une organisation sous forme de grands tourbillons de l'océan et de l'atmosphère : anticyclones, dépressions, gyres...

Les régions inter-tropicales sont gouvernées par une circulation atmosphérique particulière. En conséquence de la force de Coriolis, les régions océaniques subtropicales sont dominées par de grands anticyclones quasi-permanents. On observe ainsi des vents réguliers amenés par ces anticyclones en direction de l'équateur (alizés), et ce, dans chacun des hémisphères. Du fait du sens de rotation inverse des anticyclones entre les hémisphères par effet de Coriolis, ces vents proviennent du Nord-Est dans la partie nord de la bande tropicale, et proviennent du Sud-Est dans sa partie sud. Lorsque ces deux flux d'air se rencontrent, ces derniers sont fortement chargés en humidité et en énergie, du fait de l'abondante évaporation de l'océan dans les tropiques. Ceci provoque une instabilité dans la colonne d'air, enclenchant ainsi le transport en altitude de ces masses d'air relativement peu denses, en réponse au gradient de pression vertical et d'un effet de Coriolis faible. Ceci forme alors une grande ceinture nuageuse tropicale composée de cumulonimbus, signature de l'évaporation massive de l'océan dans les régions associées. On appelle cette région la zone de convergence intertropicale (*Inter Tropical Convergence Zone*, ITCZ). La position de l'ITCZ varie saisonnièrement (1.4), selon l'endroit où le maximum de température océanique de surface est atteint. Cette grande structure connaît également des variations sur des échelles de temps plus longues : à l'échelle

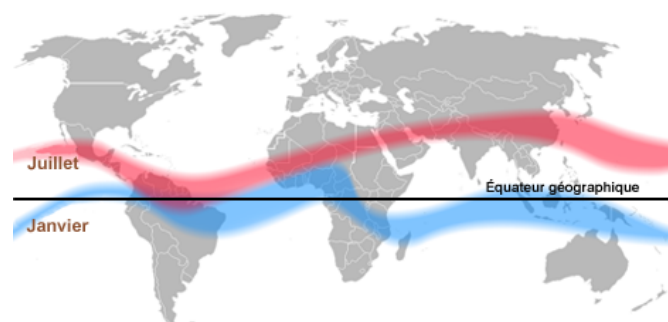


FIGURE 1.4 | Représentation schématique de la position de l'ITCZ selon la saison. Sa position moyenne approximative lors de l'été boréal est indiquée en rouge et sa position moyenne approximative lors l'été austral en bleu. La ligne noire représente l'équateur géographique. Source : figure de produite par Halldin (2006) - Wikimedia Commons [Domaine public].

multi-décennale et pluri-centennale selon l'activité des courants océaniques méridiens (Knight et al. 2005, Persechino et al. 2013), ou encore à des échelles géologiques, comme par exemple au cours d'un processus de déglaciation (Jacobel et al. 2016).

Ainsi, les variations de l'ITCZ participent à la distribution des moussons et des sécheresses dans les régions tropicales (Knight et al. 2005, Persechino et al. 2013, Gadgil 2018). Une fois que les masses d'air ascendantes de l'ITCZ atteignent le haut de la tropopause à la jonction avec la stratosphère (à environ 15 km d'altitude à l'équateur), elles ont déjà perdu une grande partie de leur humidité et ne peuvent continuer de gagner en altitude, le gradient de pression vertical avec les masses d'air de la stratosphère ne le permettant pas. Ces masses d'air sont alors contraintes de se déplacer vers les hautes latitudes en conséquence du flux constant de nouvelles masses d'air ascendantes provenant de plus basses altitudes. À mesure que ces masses d'air se déplacent vers les hautes latitudes, leur humidité et donc les précipitations qui leur sont associées décroissent fortement. C'est ainsi que se forment les deux ceintures désertiques à environ 30°S et 30°N. Une fois arrivées à de tels degrés de latitude, les masses d'air sèches descendent en altitude jusqu'à atteindre la surface. Une partie de ces vents rencontre alors les anticyclones subtropicaux qui les redistribuent partiellement en direction de l'équateur sous forme d'alizés. Ces mouvements de masses d'air méridiens et verticaux portent le nom de cellules de convection atmosphérique. Ces cellules tropicales particulières portent le nom de cellule de Hadley (Fig. 1.5). Du fait de leur position tropicale et donc de leur forte exposition à l'irradiance solaire, les cellules de Hadley sont celles qui redistribuent le plus d'énergie et dont les circulations sont les plus intenses.

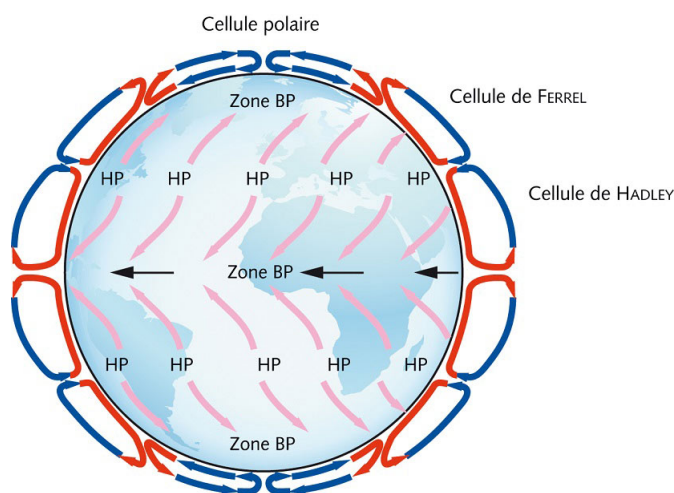


FIGURE 1.5 | Représentation schématique de la circulation atmosphérique globale. Principales cellules de convection atmosphérique méridienne et position latitudinale des centres de basse pression (BP) et de haute pression (HP). *Source : figure de Moreau (2013).*

Tout comme pour les cellules de Hadley, les cellules polaires de chaque hémisphère sont directement impliquées par les gradients thermiques de l'océan et de l'atmosphère, notamment dûs à l'inégale distribution du rayonnement électromagnétique solaire. Au niveau des centres de dépression subpolaires, environ situés au 60^{ème} parallèle de chaque hémisphère, les masses d'air sèches et denses provenant des pôles rencontrent et se mélangent aux masses d'airs relativement chaudes et humides venant des plus basses latitudes. La basse pression de ces masses d'air

provoque alors un processus de convection atmosphérique. Ces vents suivent ainsi un mouvement vertical jusqu'à la tropopause (à environ 8 km d'altitude dans ces régions) et sont renvoyés à de plus hautes latitudes du fait du flux vertical constant de nouvelles masses d'air. Une fois que les masses d'air arrivent près des pôles, elles se sont préalablement déchargées en humidité sous forme de précipitations. De par

leur refroidissement elles descendent alors selon un mouvement vertical et sont ensuite renvoyées en surface à de plus basses latitudes, en direction des centres de dépression subpolaires.

À l'inverse des cellules polaires et des cellules de Hadley, les cellules de Ferrel sont des cellules de circulation indirectes sur le plan thermique. Il s'agit du résultat dynamique qui découle de l'activité des deux autres types de cellules. Au niveau des 60^{ème} parallèle, l'air ascendant qui constitue la cellule polaire n'est pas entièrement renvoyé vers les pôles. Une partie est renvoyée vers les basses latitudes, en divergeant vers l'ouest par effet de Coriolis. Lorsque ces masses d'air rencontrent l'air chaud et sec de la cellule de Hadley au niveau des 30^{ème} parallèles, elles descendent avec elles et renforcent les centres de hautes pression subtropicaux. Ces cellules permettent de partiellement redistribuer l'air chaud et humide venant des tropiques et l'air froid et sec venant des hautes latitudes. Elles se trouvent donc en moyennes latitudes, du 30^{ème} au 60^{ème} parallèle dans chaque hémisphère.

Aux jonctions de cellules de convection de différents types, on observe de grandes structures de vents régulières que l'on appelle courants-jets (*jet streams*). On appelle front subtropical (respectivement polaire) l'intersection entre une cellule de Ferrel et une cellule de Hadley (respectivement polaire). Ces fronts se caractérisent comme des zones de rencontre entre masses d'air aux propriétés très différentes en matière de température et de pression. Leur fort contraste thermique en altitude force une partie des masses d'air à ne plus suivre le chemin tracé par les cellules de convection, qui s'écoule alors de manière horizontale. En réponse à la rotation de la Terre (et donc par effet de Coriolis), ces courants-jets soufflent d'ouest en est dans les deux hémisphères. Les courants-jets subtropicaux sont déterminants pour la redistribution des orages et des températures hivernales dans chacun des hémisphères. Néanmoins, ils sont particulièrement puissants en hiver (relatif à chaque hémisphère), puisque le contraste thermique entre les masses d'air tropical et celles de plus hautes latitudes s'intensifie. Leurs mouvements d'ouest en est ne sont pas parfaits, ces derniers pouvant être déviés selon l'intensité des centres de haute pression (anticyclones subtropicaux) et des centres de basse pression (dépressions subpolaires). Pour ce qui est des courants-jets subpolaires, ils sont beaucoup plus rapides, du fait d'un fort contraste thermique entre les régions polaires et les régions de moyennes latitudes.

Au-delà de ces grandes structures relativement stables (Eliassen 1951), les courants atmosphériques, en surface comme en altitude, possèdent une variabilité chaotique (Hasselmann 1976). C'est pour cette raison que les prévisions météorologiques, qui ne sont autres que des prévisions de l'état de l'atmosphère, ne sont en général données qu'à l'horizon de deux semaines. D'un point de vue stochastique, Hasselmann (1976) propose qu'un signal atmosphérique est un processus sans mémoire (ou bruit blanc), autrement dit, qu'il s'agit d'un signal chaotique, dont l'évolution à un temps donné ne dépend pas de ses états précédents. Frankignoul and Hasselmann (1977) suggèrent qu'à l'inverse de l'atmosphère, l'océan possède une inertie et une capacité thermique bien plus importante de par sa forte viscosité et sa forte densité, l'amenant ainsi à évoluer sur des échelles de temps plus longues : multi-décennales (> 10 ans), pluri-centennales (> 100 ans). Il a donc été proposé de considérer les signaux océaniques comme des processus markoviens du premier ordre

(ou marche aléatoire), forcés par le bruit blanc stochastique de l'atmosphère (Hasselmann 1976). En effet, les mouvements des courants atmosphériques impliquent localement des mouvements d'advection océanique à la jonction entre les deux fluides (surface). Ainsi, les mouvements atmosphériques, très turbulents, pouvant dans un même lieu changer de direction d'une heure à l'autre, constituent une source de variabilité haute fréquence pour l'océan. Néanmoins, ces variations hautes fréquences sont trop rapides pour mettre en mouvement l'océan au delà de la turbulence induite. Ceci implique, en général, une prévisibilité à long-terme plus élevée pour l'océan que pour l'atmosphère. Néanmoins, il faut noter qu'au delà des hypothèses proposées par Hasselmann (1976) et Frankignoul and Hasselmann (1977), on sait maintenant que les interactions entre l'océan et l'atmosphère peuvent également faire que l'on observe des évolutions de processus atmosphériques à basse fréquence (Smith et al. 2020).

Comme évoqué au début de cette section, la salinité et la température d'une masse d'eau déterminent en grande partie sa densité. C'est pourquoi on nomme fréquemment circulation thermohaline (Fig. 1.6), la circulation océanique globale (du grec *thermo* pour température et *halin* pour la salinité) (Rahmstorf 2006). Grossièrement, cette circulation se divise en deux branches : l'une en surface, et l'autre en profondeur (entre 1000 et 6000 mètres selon les régions). Cette circulation a de très forts impacts sur le climat de différentes régions puisque sa branche de surface advecte des eaux équatoriales chaudes. Ces eaux vont alors réchauffer l'atmosphère, notamment par le biais des flux de chaleurs, ce qui joue un rôle direct sur la météorologie et la climatologie de nombreuses régions. Puisqu'il s'agit de mouvements océaniques, cette circulation se fait à très basse fréquence (Frankignoul and Hasselmann 1977). On estime en effet qu'il faut environ 1000 à 1500 ans pour qu'une molécule d'eau parcoure l'ensemble de la boucle formée par cette circulation (Rahmstorf 2002, Fig. 1.6)

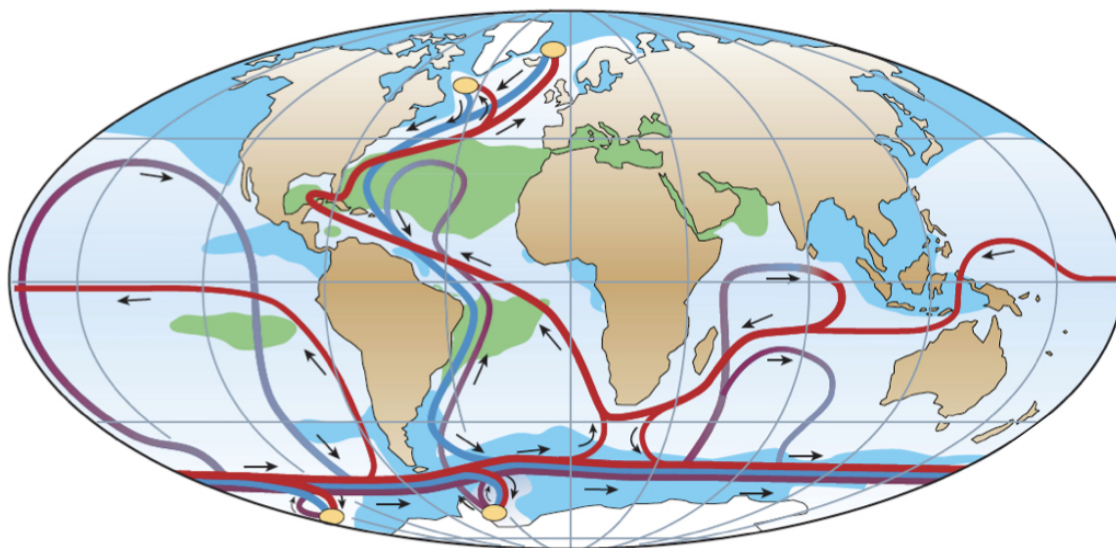


FIGURE 1.6 | Schéma très simplifié de la circulation thermohaline. La branche rouge est celle de surface, et la bleue est celle en profondeur. Les flèches noires indiquent le sens de circulation des courants. Les points jaunes indiquent les principaux points de convection océanique. Les surfaces bleues ciel sont celles pour lesquelles on observe une salinité inférieure à 34‰ entre 1990 et 1996. Les surfaces vertes sont celles pour lesquelles on observe une salinité supérieure à 36‰ entre 1990 et 1996. Source : figure de Rahmstorf (2006), données de Broecker (1999).

Généralement, l'océan se divise en plusieurs couches aux densités différentes, les couches les plus denses se trouvant en profondeur, et les moins denses en surface. On appelle stratification de l'océan cette organisation sous forme de couches aux densités suffisamment différentes pour que celles-ci communiquent très peu, voire pas du tout, selon les régions. Néanmoins, certains phénomènes, à l'origine de la circulation thermohaline, permettent à ces différentes strates de communiquer.

En particulier, la turbulence des courants atmosphériques et les marées jouent un rôle essentiel dans la stabilité de la circulation thermohaline, notamment parce qu'elles permettent à des eaux profondes de remonter en surface (Wunsch 2000, Wunsch 2002). Ce phénomène a généralement lieu dans des régions équatoriales ou dans des régions côtières. Les processus physiques à l'origine de ces remontées d'eaux varient en fonction de la région. En règle générale, il s'agit d'un phénomène de pompage d'Ekman. Ce phénomène est dû au fait qu'il existe une zone de transition thermique très rapide entre les couches océaniques, spécifique à chaque région, que l'on appelle thermocline. En effet, puisque l'océan possède une certaine transparence, les rayons solaires ne le traversent que sur quelques dizaines de mètres, ne réchauffant que les couches les plus hautes. Cette profondeur est plus importante dans les régions tropicales, où les rayons solaires sont plus nombreux par unité de surface et l'angle d'incidence du rayonnement solaire y est plus faible. Apparaît alors un fort contraste thermique entre les eaux de surface directement chauffées par le rayonnement solaire et affectées par les mouvements de l'atmosphère, et celles plus en profondeur. Lorsque des vents réguliers ou saisonniers (alizés, moussons...), font se déplacer perpendiculairement les eaux de surface (transport d'Ekman), ceci crée en ce lieu un vide à la surface de l'océan. On a alors une thermocline plus proche de la surface océanique, ce qui peut engendrer un phénomène de remontées d'eaux froides en surface, comblant ainsi le vide créé par les courants atmosphériques. Ce phénomène montre en particulier comment la variabilité atmosphérique peut impacter la circulation de l'océan à basse fréquence.

La convection océanique est le second et principal phénomène permettant aux différentes strates océaniques de communiquer et donc de former les deux branches de la circulation thermohaline. Ce phénomène a lieu lorsque les couches de surface deviennent plus denses que les couches plus profondes, ces dernières vont alors se mélanger et former des eaux profondes. Il existe plusieurs régions de convection océanique, mais on en distingue trois principales qui participent majoritairement à la circulation thermohaline. La première se situe dans la mer de Weddell (océan Austral), au sud de l'Amérique du Sud. Les deux autres régions principales où se forment les eaux profondes se trouvent à la jonction entre l'océan Arctique et l'océan Atlantique. On trouve la première d'entre elles dans la région

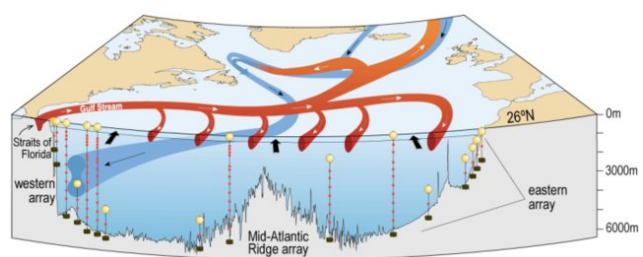


FIGURE 1.7 | Représentation schématique des composantes de l'Atlantic Meridional Overturning Circulation (AMOC) selon une coupe à 26°N. Les flèches noires indiquent le transport d'eaux équatoriales vers les pôles (transport d'Ekman). Le branchement rouge est la branche de surface en direction du Nord, la branche bleue est celle qui se situe en profondeur en direction du Sud. Les points jaunes illustrent le tableau d'amarrage permettant de mesurer les courants (Cunningham et al. 2007). *Source : figure de Smeed et al. (2014).*

du gyre subpolaire (*Subpolar Gyre*, SPG) de l'Atlantique, située au sud du Groenland, comprenant les mers du Labrador et d'Irminger, et la seconde dans les mers Nordiques, au nord des pays scandinaves (Fig. 1.6, 1.7). Les eaux convectives, qui plongent sur plusieurs kilomètres, repartent ensuite vers le sud sous forme d'un large courant d'eaux profondes. Ceci implique donc que l'Atlantique Nord est l'un des acteurs clefs de cette circulation de grande échelle (Fig. 1.6), et fait donc l'objet d'un intérêt scientifique particulier. Il est difficile de discerner les mouvements dus aux gradients de densité et ceux dus à d'autres phénomènes, tels que les mouvements de l'atmosphère (Wunsch 2000). On étudie généralement la circulation océanique par le biais de la circulation méridienne de retournement (*Meridional Overturning Circulation*, MOC). On l'étudie en particulier dans l'Atlantique (*Atlantic MOC*, AMOC, Fig. 1.7), à laquelle le *Gulf Stream* participe en surface. On considère ainsi l'AMOC comme un excellent indicateur de l'état de la circulation océanique engendrée par les gradients de densité (Wunsch 2002, Rahmstorf 2006).

1.1.3 Forçages radiatifs externes et variabilité climatique interne

L'énergie redistribuée au sein du système climatique provient en grande partie du flux électromagnétique solaire (99.97%), composé de rayonnements aux différentes longueurs d'ondes (infrarouge, ultraviolets, lumière visible...). Le reste de l'énergie (0.03 %) provient de la chaleur interne de la Terre. Chaque composante du système climatique possède une capacité d'absorption et de réfraction du rayonnement incident qui lui est spécifique. Le rayonnement électromagnétique ré-émis par les composantes de la Terre est fonction de sa température (loi de Planck), et se situe majoritairement dans le spectre infrarouge. On appelle alors bilan radiatif terrestre la différence entre l'énergie reçue et perdue par le système climatique (Fig. 1.8). C'est au travers de celui-ci que des éléments externes au système climatique peuvent modifier ses variations.



FIGURE 1.8 | Représentation du bilan radiatif terrestre. Données calculées à partir d'observations diverses, dont des observations satellite allant de mars 2000 à mars 2004. Source : figure francisée de Trenberth et al. (2009).

Les premières mesures directes de variables météorologiques remontent à la fin du XIX^e siècle, ce qui, non sans coïncidence, a lieu au cours de la révolution industrielle. Marquée par la découverte et le début de l'exploitation de nombreux gisements d'énergies fossiles, cette période constitue un tournant dans l'histoire de l'humanité, puisqu'elle a vu sa population totale multipliée par sept en environ 200 ans (donnée basée sur les estimations de Biraben 1979 et Bucci and Raurich 2017). Aujourd'hui, on sait que la combustion de ces énergies fossiles, indispensable au fonctionnement des sociétés humaines contemporaines, participe à modifier significativement la composition chimique de l'atmosphère terrestre. Ceci se fait notamment par l'apport de molécules comme le dioxyde de carbone (CO₂), le protoxyde d'azote (N₂O) ou le méthane (CH₄), responsables du renforcement de l'effet de serre terrestre. Ce constat rend donc très probable le caractère anthropique de l'augmentation globale des températures.

L'activité humaine est également à la base d'émissions d'autres types de gaz, contenant des aérosols. Il s'agit de très fines particules qui restent en suspension dans l'air, c'est-à-dire que leur vitesse de chute est faible. Une fois injectées dans l'atmosphère, ces particules diffusent une fraction des rayonnements incidents et en ré-émettent une partie vers l'espace (effet dit "parasol"), jouant ainsi un rôle négatif sur la température atmosphérique de surface moyenne. Ils peuvent également absorber une partie du rayonnement et réchauffer les couches de l'atmosphère où ils se trouvent. Ces particules favorisent d'autre part l'effet d'albédo des nuages de par des modifications de leurs propriétés microphysiques et optiques, ce qui joue également un rôle négatif sur les températures atmosphériques de surface (Booth et al. 2012, Boucher et al. 2013, Smith et al. 2016). Toutefois, ces particules injectées dans la troposphère inférieure finissent par en ressortir au bout de quelques jours, soit par la pluie ou la neige, soit par dépôts directs d'humidité. Ainsi, à moins d'être émises de manière continue, ces particules ont un effet long-terme sur les températures bien moins marqué que celui des gaz à effet de serre.

D'autres facteurs, externes au système climatique, jouent un rôle sur le bilan radiatif de la Terre. Aux échelles de temps décennales et multi-décennales, les deux principaux sont des variations de l'irradiance solaire totale (*Total solar irradiance*, TSI, e.g. Swingedouw et al. 2011) et les éruptions volcaniques (Mignot et al. 2011, Booth et al. 2012, Khodri et al. 2017). Les éruptions volcaniques libèrent des gaz riches en sulfates, dans une quantité qui dépend à la fois de la nature du volcan et de la puissance de son éruption. Pour certaines d'entre elles, les éruptions dites stratosphériques, qui sont en général celles de plus grande ampleur, une quantité abondante de ces gaz en partie composés de dioxyde de soufre (SO₂) est envoyée directement dans la stratosphère, et y forment des particules aérosols via des processus complexes (Swingedouw et al. 2017). La vitesse des vents atteint plusieurs centaines de km.h⁻¹ dans cette couche de l'atmosphère, si bien qu'il ne faut que quelques mois après l'éruption pour que le nuage volcanique couvre l'ensemble du globe. À l'inverse des émissions de gaz aérosols anthropiques, les particules formées par les nuages volcaniques peuvent rester plusieurs années dans l'atmosphère du fait de l'altitude à laquelle elles se trouvent. L'effet de refroidissement global impliqué par ce type d'éruption est donc de plus long-terme (en général 2 à 3 ans, Robock 2000, Khodri et al. 2017). Pour illustrer les effets d'une éruption stratosphérique massive, l'année "sans été" (1816) est souvent prise en exemple. Cet événement de l'histoire est survenu à la suite de l'éruption du volcan Tambora (Indonésie, île de *Sumbawa*) en 1815,

estimée comme la 6^{ème} plus puissante éruption des 2500 dernières années (Sigl et al. 2015), et a engendré un bilan humain et matériel catastrophique sur l'île (Brönnimann and Krämer 2016). En terme d'impact global de l'éruption, l'année 1816 est marquée par un été boréal particulièrement froid, au cours duquel des chutes de neiges en Europe (Trigo et al. 2009) et des pluies incessantes en Asie (Brönnimann and Krämer 2016) sont documentées dans des archives historiques. Les agriculteurs de l'époque ont alors connu des récoltes désastreuses, amenant émeutes, famines et migrations au travers de l'Europe, l'Asie et l'Amérique du Nord (Brönnimann and Krämer 2016).

Puisque l'on ne dispose pas de mesures météorologiques chiffrées et précises pour cette époque, il est difficile de comprendre la signature climatique exacte de cet événement. En revanche, des archives historiques telles que les peintures des artistes anglais William Turner et William Ascroft (Fig. 1.9, 1.10) ont permis de documenter les répercussions mondiales des éruptions de Tambora en 1815 et du Krakatoa en 1883 en terme de composition chimique de l'atmosphère. Dans leurs tableaux respectivement parus en 1816 et 1883, on aperçoit un ciel aux teintes rougeâtres très prononcées. L'analyse chromatique de plusieurs dizaines de tableaux anciens a permis de montrer que ces couleurs sont dues à une importante quantité de poussières volcanique dans l'atmosphère les années suivant de grosses éruptions, et de valider l'ampleur mondiale de tels phénomènes sans même disposer de données météorologiques et climatiques chiffrées (Zerefos et al. 2007).



FIGURE 1.9 | Peinture de William Turner - *Dido building Carthage (1816)*. Source : Wikimedia commons [Domaine public].



FIGURE 1.10 | Peinture de William Ascroft - *Sans nom. Une de ses nombreuses productions entre 1883 et 1884*. Source : Wikimedia commons [Domaine public].

Plus récemment ont eu lieu d'autres éruptions, bien moins importantes que celle de Tambora (1815), comme celles des monts Agung (1963), El chichòn (1982) et Pinatubo (1991). On estime que le trop peu d'observations de ces événements et leur relative faiblesse ne permettent pas d'estimer de manière statistiquement robuste leur empreinte sur les variations climatiques, notamment en terme de rapport signal/bruit (Swingedouw et al. 2017). L'imprévisibilité de l'ampleur, de l'occurrence et de la localisation des éruptions volcaniques rend la tâche particulièrement difficile pour faire des projections climatiques fiables, mais une compréhension approfondie de leur impact sur le climat peut s'avérer fructueuse en matière de planification si un événement de type Tambora venait à survenir dans les années à venir (Hermanson et al. 2020).

Les événements naturels et externes au système climatique jouant un rôle sur le bilan radiatifs de la Terre, comme les variations de la TSI et les éruptions volcaniques

stratosphérique, sont appelés forçages climatiques naturels. Ce terme s'oppose aux forçages climatiques anthropiques, découlant quant à eux de l'activité humaine. Indépendamment de ces différents phénomènes, le climat continue de fluctuer selon les grandes structures océaniques et atmosphériques décrites dans la section précédente, et ce, même sans modification de son bilan radiatif. On parle de variabilité climatique interne. Le terme de variabilité climatique interne est né selon l'idée qu'au sein des modèles numériques de climat, le climat continue de fluctuer même si aucun forçage radiatif ne lui est imposé. Dans la pratique, le bilan radiatif terrestre évolue de manière continue, si bien qu'il n'est pas possible de directement observer la variabilité climatique dite interne, au sens de sa définition initiale. Bien que l'existence de cette variabilité ne fasse aucun doute, de nombreuses questions résident dans sa capacité à influencer sur le climat, en particulier dans le contexte de changement climatique actuel.

1.1.4 Outils d'étude du climat

En tant que système physique, le climat évolue de manière continue dans le temps et l'espace. Pour décrire son comportement, on dispose de grilles de mesures de variables météorologiques discrétisées dans l'espace et dans le temps. C'est ce qu'on appelle les observations *in situ*. Celles-ci sont des mesures de l'évolution réelle du climat. À mesure que les techniques s'améliorent, on dispose de plus en plus de données, sur des grilles toujours plus fines, en particulier depuis la mise en place de satellites dans les années 1980, qui enregistrent continuellement des données à de très hautes résolutions. Il s'agit des uniques données directement mesurées que nous ayons de la réalité, et ces observations ne prennent en compte qu'une part infime de l'histoire climatique de la Terre. Elles ne permettent donc qu'une compréhension très partielle des processus physiques à l'oeuvre dans les variations du climat, et ce, sur des échelles de temps très restreintes.

Pour étudier le climat de manière plus dynamique, on a recours à l'utilisation de modèles numériques qui simulent son comportement à partir des lois physiques et des cycles biogéochimiques connus. Il s'agit de modèles couplant différents sous-modèles : modèle d'océan, modèle d'atmosphère... On les appelle modèles de circulation générale (ou globale) du climat (*General circulation models*, GCM). Comme pour tout modèle physique, il s'agit de représenter au mieux la réalité en la schématisant de manière simplifiée. Ils sont implémentés de manière à résoudre, au sein de grilles discrétisées et sur une succession de pas de temps, les équations découlant des lois de la physique et des cycles biogéochimiques connus. Ils fonctionnent tel un système dynamique déterministe, c'est-à-dire que pour un état initial donné, la simulation d'un même modèle ne dépendra alors que des conditions aux limites qui lui sont imposées. Les conditions aux limites ici sont les forçages externes, que l'on peut préalablement prescrire à un modèle, afin de simuler l'effet d'émissions anthropiques de CO₂, ou de simuler une éruption volcanique stratosphérique par exemple.

Il existe une quarantaine de ces modèles de circulation globale à travers le monde, développés au sein de grands centres de recherche. On peut par exemple citer les deux modèles Français, développés à l'Institut Pierre-Simon Laplace (IPSL) et au Centre National de Recherche en Météorologie (CNRM). Les schémas numériques des modèles existants sont différents, si bien que pour des états initiaux similaires,

ils ne vont pas simuler le même climat. Ceci s'illustre notamment par l'incertitude des projections associées aux différents modèles numériques présentée par l'étude de Hawkins and Sutton (2009), qui sera discutée dans la section 1.2.3. Ainsi, une approche souvent utilisée pour étudier des phénomènes physiques à l'aide de modèles est de considérer un ensemble de simulations de ces derniers. Ces modèles de circulation sont ceux dont les projections sont présentées dans les rapports de synthèse du GIEC. De ce fait, chaque centre de recherche est chargé de proposer une nouvelle version de son modèle de climat, à intervalle régulier, de manière à être à disposition pour aider à la rédaction des rapports du GIEC (tous les 7-8 ans). Chaque génération propose des modèles basés sur les nouvelles connaissances accumulées depuis la génération précédente. Aujourd'hui (2020), les modèles de la 6^{ème} phase du *Coupled Model Intercomparison Project* (CMIP6) constituent la toute nouvelle génération de modèles de circulation globale. Puisque celle-ci est récente, les modèles et sorties de modèles utilisés dans cette thèse sont issus de la génération CMIP5 (modèles parus en 2013).

Si les modèles de circulation globale se veulent représentatifs des réalités de la physique et des cycles biogéochimiques, ils ne sont pas voués à directement donner des observations de la réalité. Il ne permettent donc pas de directement résoudre le problème du manque d'observations *in situ*. Une manière indirecte d'étudier des variations réelles du climat ayant eu lieu avant qu'on ne le mesure, est de regarder les variations de marqueurs biochimiques ou géochimiques, mesurés à partir d'archives naturelles. Par exemple, les écarts entre les anneaux de croissance de cernes d'arbres de moyennes latitudes permettent de bien retracer la température ou les précipitations estivales (période de croissance) dans une région donnée (Esper et al. 2007). On peut également retracer la température océanique ou l'intensité de courants à partir de limons présents dans des carottes de sédiment marin (Thornalley et al. 2009, Moffa-Sánchez et al. 2014, Mjell et al. 2016, Moffa-Sánchez and Hall 2017). Les carottes de glace prélevées dans des régions de forte accumulation de neiges, en général en haute latitude et/ou altitude, sont également un excellent moyen d'étudier des climats anciens. Celles-ci ont notamment la particularité de pouvoir retracer l'évolution passée des forçages externes. Par exemple, on retrouve des traits noirs, visibles à l'oeil nu dans des carottes de glace polaires. Ces marques correspondent à du dépôt de cendres volcaniques, qui sont du fait de leur présence, issues d'éruptions stratosphériques pour qu'elles puissent atteindre de telles latitudes (Sigl et al. 2015, Fig. 1.11).

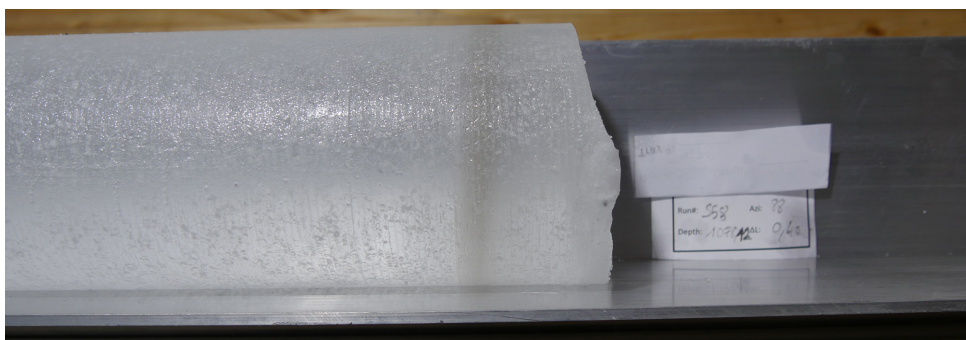


FIGURE 1.11 | Couche de cendres volcanique visible observée dans une carotte de glace prélevée au Groenland (à droite de la carotte). Cette couche a été attribuée à l'éruption du volcan Saksunarvatn (Islande), dont l'éruption a eu lieu il y a environ 10400 ans. Source : Image proposée par Anders Svensson [[Lien vers le site eastgrip.org](https://www.eastgrip.org)].

La glace a également l'avantage d'emprisonner des micro-bulles d'air. Ainsi, selon les couches dans lesquelles elles se trouvent, on peut donner une estimation de la composition chimique de l'atmosphère lorsqu'on les prend en nombre suffisant. On peut ainsi retracer l'évolution de la concentration de gaz à effet de serre tel que le CO₂ avec ce type d'approche. Enfin, les carottes de glace peuvent également permettre de retracer les variations de la TSI, à l'aide d'un isotope du Bérilyum (¹⁰Be). Il existe d'autres types d'archives naturelles permettant de retracer le climat du passé. On peut citer : des pollens présents dans du sédiment lacustre, des coraux, des spéléothèmes ou encore des archives historiques (PAGES 2k Consortium 2017). Pour une grande partie d'entre eux, y compris les carottes de glace, on peut notamment retracer les paléo-températures grâce au $\delta^{18}\text{O}$ (Jouzel et al. 2003). Il s'agit d'un rapport entre deux isotopes stables (non radioactifs) de l'oxygène : le ¹⁸O, et le ¹⁶O (qui est le plus répandu, 99,8%). On appelle enregistrements paléoclimatiques ces mesures indirectes de variations climatiques passées. Chacune de ces données n'est qu'une reconstruction locale ou régionale d'une variable climatique donnée et elles possèdent chacune des incertitudes de mesure et de méthodologie. On appelle alors reconstruction paléoclimatique tout recours à la combinaison de ces données pour produire des estimations de variations climatiques passées à plus grandes échelles.

On peut également étudier le climat avec plusieurs des approches décrites précédemment, notamment puisqu'elles ont chacune leurs avantages et leurs inconvénients. Les données climatiques *in situ* peuvent permettre de valider la physique des modèles si ces derniers parviennent à bien les reproduire. Il en va de même pour les données paléoclimatiques, mais cette fois-ci pour valider la physique des modèles sur des temps plus longs que la période de mesures instrumentales. On peut également utiliser les modèles dans le but inverse. C'est-à-dire que si des processus physiques inconnus sont à l'origine de variations observées sur les observations *in situ* ou les enregistrements paléoclimatiques, alors les modèles reproduisant bien ces observations peuvent potentiellement renseigner sur les processus qui en sont à l'origine. Enfin, il est intéressant de confronter données *in situ* et données paléoclimatiques, notamment dans un contexte de changement climatique. En effet, les enregistrements ou les reconstructions paléoclimatiques peuvent permettre d'estimer si des variations climatiques observées ont un caractère exceptionnel ou sans précédent, et donc être attribuées ou non à une réponse aux forçages radiatifs anthropiques.

1.2 Modes de variabilité et implication pour la prévision décennale

On distingue sur les observations climatiques, et selon différentes échelles de temps, des structures météorologiques cohérentes et récurrentes. On a ainsi remarqué qu'en deux lieux très espacés (plusieurs milliers de kilomètres), on peut avoir pour une échelle de temps donnée, des variations climatiques phasées. On parle de téléconnexions. L'abondance de ces téléconnexions signifie que le climat s'organise autour de grandes structures, selon différentes échelles spatiales et temporelles. On appelle modes de variabilité climatique ces grandes structures du climat. De par leur cohérence et leur aspect récurrent, la compréhension de ces modes est cruciale pour étudier convenablement le climat, puisque ces derniers sont associés à des fluctuations et des changements climatiques très caractéristiques.

1.2.1 Modes de variabilité de l'Atlantique Nord

L'océan Atlantique Nord est marqué par la succession prononcée de décennies anormalement froides ou anormalement chaudes de ses températures de surface (*sea surface temperatures*, SST), indépendamment de la tendance globale de réchauffement (Kushnir 1994, Kerr 2000). Ces fluctuations de la SST de l'Atlantique Nord (*North Atlantic SST*, NASST, moyenne spatiale entre les latitudes 0° et 60°N) peuvent de ce fait accentuer ou ralentir l'effet du réchauffement global anthropique dans les régions avoisinantes (Ting et al. 2009, Mann et al. 2014).

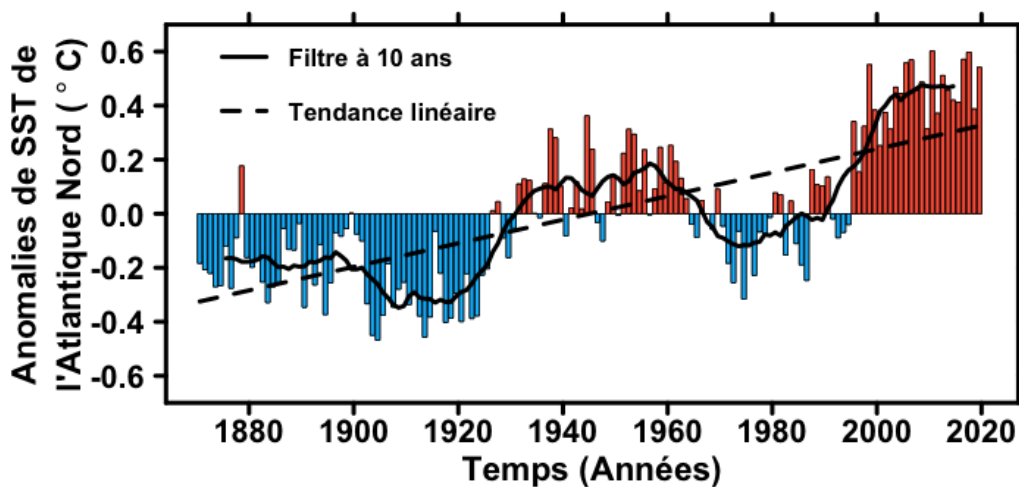


FIGURE 1.12 | Observation des anomalies de SST de l'Atlantique Nord relatives à la période 1870-2019. Les anomalies négatives sont présentées sous forme de barres bleues et les anomalies positives sous forme de barres rouges. La courbe noire pleine est la série des données filtrées à 11 ans par moyenne mobile. La droite noire pointillée indique la tendance linéaire estimée par régression (section 2.2.1). Source : jeu de données HadISST (Rayner et al. 2003).

On appelle Variabilité Multi-décennale de l'Atlantique (*Atlantic Multidecadal Variability*, AMV, Fig. 1.12), le mode de variabilité sous-jacent à ces variations de température. D'une part, on sait que les phases positives et négatives de ce phénomène ont d'importants impacts sur les conditions météorologiques estivales de l'Amérique du Nord, l'Amérique Centrale et l'Europe (Sutton and Hodson 2005, Sutton and Dong 2012, O'Reilly et al. 2017, Azuz-Adeath et al. 2019). D'autre part, les variations de l'AMV régulent le débit de certaines rivières d'Amérique du Nord (Enfield et al. 2001, Kelly and Gore 2008) et d'Europe (Boé and Habets 2014), notamment par les variations de températures qui influent sur le volume, et donc la fonte, des glaciers de haute montagne. L'AMV contrôle également la distribution des anomalies inter-hémisphériques de SST dans l'Atlantique Nord, ce qui joue un rôle sur la localisation des régions de convection atmosphérique tropicale de l'Atlantique, et donc sur le positionnement de l'ITCZ (Knight et al. 2006). En conséquence, ces changements dans le positionnement latitudinal de l'ITCZ impliquent des modifications dans la distribution des moussons et des sécheresses des pays tropicaux, notamment dans la région sahélienne, le Nord-Est du Brésil, et l'Asie centrale (Knight et al. 2006, Zhang and Delworth 2006, Defrance et al. 2017). C'est ainsi que la dernière phase négative

de l'AMV, associée à un décalage vers le Sud de l'ITCZ (Knight et al. 2006, Persechino et al. 2013), ayant commencé à la fin des années 1960 (Fig. 1.12), est probablement responsable des sécheresses et famines répétées en Afrique sahélienne de 1967 à 1972 (Defrance et al. 2017). La phase positive de l'AMV qui suit, ayant commencé à la fin des années 1990 (Fig. 1.12), est associée à un réchauffement anormal de la SST de l'Atlantique Nord, et notamment dans sa partie tropicale. Plusieurs études soutiennent que cela a engendré un accroissement et une intensification des ouragans tropicaux de l'Atlantique Nord (Emanuel 2005, Trenberth and Shea 2006, Zhang and Delworth 2006, Enfield and Cid-Serrano 2010), responsables de dégâts humains et matériels considérables dans les Antilles, les Caraïbes et la côte Atlantique tropicale de l'Amérique du Nord, dans les années 2000 notamment (*e.g.* Katrina, 2005; Rita, 2005; Wilma, 2005; Ike, 2008). Plus récemment, on observe que l'AMV semble terminer sa phase positive et devrait entrer dans une période négative d'au moins deux décennies (Frajka-Williams et al. 2017, Robson et al. 2018).

Les processus à l'origine des variations multi-décennales de l'AMV font l'objet d'une controverse scientifique très animée. Les désaccords viennent du fait que la période sur laquelle l'AMV est observée (depuis 1870 environ), est trop courte pour établir de manière précise ses moteurs et ses propriétés spectrales. En effet, bien que les mesures existantes semblent montrer que l'AMV a une périodicité de 40-80 ans (Kerr 2000), cette estimation ne peut être considérée comme statistiquement robuste avec seulement 150 ans de données. À l'inverse, les simulations numériques des modèles climatiques récents produisent en général une périodicité de l'AMV dans la branche 10-30 ans (Lin et al. 2019). Ceci souligne donc l'importance de disposer d'une période d'observation plus longue pour comprendre les moteurs de l'AMV.

Bien qu'au début des années 2000 il semblait se dessiner un consensus sur la nature interne de l'AMV (Enfield et al. 2001, Knight et al. 2005, Knight et al. 2006, Trenberth and Shea 2006, Zhang and Delworth 2006), des études parues quelques années plus tard viennent remettre en question cette idée. Une partie de la communauté des sciences du climat soutient en effet depuis quelques temps l'hypothèse selon laquelle la teneur en aérosols de l'atmosphère, qu'elle soit d'origine naturelle ou anthropique, peut très bien expliquer les variations observées de l'AMV (Booth et al. 2012, Undorf et al. 2018, Wanatabe and Tatebe 2019). Dans cette veine, Smith et al. (2016) ont montré que les émissions d'aérosols étaient maximales en Amérique du Nord et en Europe dans les années 1960 (phase d'AMV négative émergente), et ont commencé à graduellement décroître dans les années 1980 (phase d'AMV positive émergente) suite à des politiques de régulation de la pollution atmosphérique. Puisque ces régions bordent de part et d'autre l'Atlantique Nord, ceci soutient l'idée selon laquelle les dernières périodes négative et positive de l'AMV pourraient être en grande partie expliquées par les émissions de gaz aérosols, et en particulier celles d'origine anthropique.

Néanmoins, les récentes observations d'une phases négative émergente de l'AMV (Frajka-Williams et al. 2017, Robson et al. 2018) montrent que les processus de variabilité climatique interne pourraient bien être à l'origine de ce refroidissement relatif récent de l'Atlantique Nord. De plus, des études ont montré que des enregistrements paléoclimatiques marins et terrestres, situés dans des régions affectées par l'AMV, possèdent de fortes fluctuations multi-décennales, mais sur des périodes bien plus longues que celle couverte par les données provenant de mesures instrumentales

(Gray et al. 2004, Mjell et al. 2016). On suggère ainsi que des processus associés à la circulation océanique, et en particulier à l'AMOC, sont à l'origine de cette variabilité multi-décennale (McCarthy et al. 2015, Muir and Fedorov 2015, Yan et al. 2018, Zhang et al. 2019a). Cependant, ce n'est que depuis 2004, après la mise en place du projet *Rapid Climate Change* (RAPID) (Cunningham et al. 2007), que des observations directes de l'AMOC sont disponibles. Ce projet de recherche international a pour but de mesurer, en continu et sur toute la profondeur de l'océan, le transport de masses d'eau (exprimé en Sverdrup, $1 \text{ Sv} \simeq 100 \text{ (m}^3/\text{s)}$) méridien, selon une coupe longitudinale faite à 26.5°N le long de l'Atlantique. Bien qu'au sein des modèles on discerne un lien statistique robuste entre les variations de l'AMV et de l'AMOC (Knight et al. 2005, Persechini et al. 2013), l'explication physique de ce lien se donne selon différentes théories (Knight et al. 2005, Persechini et al. 2013, Zhang et al. 2019a). Une théorie très en vogue implique un mécanisme océanique faisant évoluer l'AMV sur des périodes de 20-30 ans et fut décrit à partir d'observations du niveau marin et de simulations de modèle physique par Frankcombe et al. (2008) et Frankcombe and Dijkstra (2009). Ils proposent qu'une intensification de l'AMOC signifie que le transport d'eaux chaudes de l'Équateur vers l'Arctique s'accroît. L'accumulation d'eaux de surface anormalement chaudes en Atlantique pendant plusieurs années (cf. AMV positive, Fig. 1.12) conduit à une perte anormale de densité des eaux de surface dans les régions de convection. La re-stratification partielle de la colonne d'eau entraîne alors un ralentissement de l'AMOC, et à terme un refroidissement anormal de l'Atlantique Nord (cf. AMV négative, Fig. 1.12). La boucle du processus se termine lorsque les eaux de surfaces anormalement froides qui s'accumulent dans les régions nordiques finissent par anormalement gagner en densité et font l'objet d'une convection plus intense, ce qui est synonyme d'une intensification de l'AMOC. En se basant sur cette hypothèse, l'AMV pourrait constituer une observation indirecte du comportement de l'AMOC sur une bien plus longue période que celle couvrant les observations directes de cette dernière. Néanmoins, on ne peut directement associer les deux puisque la SST ne dépend pas que des variations de l'AMOC, mais aussi des flux de chaleur associés aux forçages stochastiques de l'atmosphère (Clement et al. 2015), ou de la présence de particules aérosols dans la haute troposphère et la stratosphère (Booth et al. 2012, Undorf et al. 2018, Wanatabe and Tatebe 2019).

Du fait que l'origine des variations de l'AMV ne font pas consensus, calculer des indices de ce mode soulève à nouveau des questions. En effet, les indices AMV actuellement développés sont en général construits de manière à retirer de la manière la plus robuste possible l'effet des différents forçages radiatifs externes sur les variations de la NASST, afin de ne garder que sa composante de variabilité interne. Ceci se base donc sur l'hypothèse forte et débattue selon laquelle les forçages radiatifs externes ne sont que dans une moindre mesure à l'origine des variations multi-décennales de l'AMV. Une manière précautionneuse et non polémique de définir l'AMV est donc de la considérer comme la composante de variabilité multi-décennale de la NASST, quelle que soit son origine.

Les méthodes de calcul d'indices de l'AMV basées sur une suppression des forçages externes permettent de bien isoler la composante multi-décennale de la NASST. La première fut proposée par Enfield et al. (2001), puis abondamment utilisée au début des années 2000 (Knight et al. 2006, Zhang and Delworth 2006). Celle-ci consiste

à estimer une tendance par régression linéaire (section 2.2.2) de la série de NASST qui lui est ensuite retirée (Fig. 1.13).

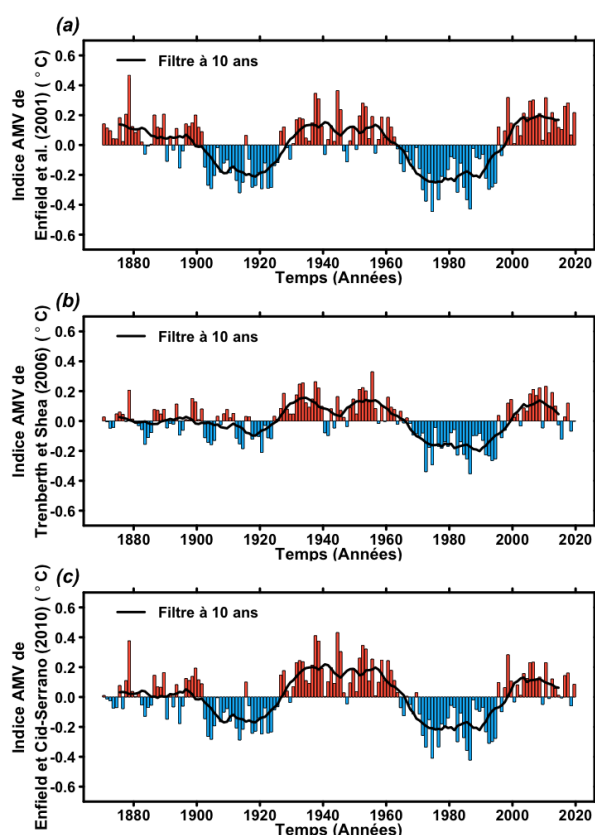


FIGURE 1.13 | Indices AMV calculés sur la période 1870-2019. a) Indice obtenu par suppression de la tendance linéaire (Enfield et al. 2001). b) Indice obtenu par suppression de la tendance quadratique (Enfield and Cid-Serrano 2010). c) Indice obtenu par la différence entre la moyenne des anomalies de la SST de l'Atlantique Nord et la moyenne globale des anomalies de la SST (Trenberth and Shea 2006). Source : jeu de données HadISST (Rayner et al. 2003).

est ensuite retirée de la NASST observée par régression (section 2.2.2). L'estimation de la composante forcée est soit calculée comme la première composante principale (section 2.2.1) des différentes NASST des simulations historiques (Ting et al. 2009), soit par une moyenne d'ensemble de celles-ci (Steinman et al. 2015). Il existe encore d'autres méthodes de calcul d'indice de l'AMV dont une description détaillée est proposée par Frankignoul et al. (2017). Cette abondance d'indices amène notamment des études de l'AMV à se baser sur un ensemble d'indices plutôt qu'un seul (Qasmi et al. 2017, Frankignoul et al. 2017). Ceci met par ailleurs en évidence le besoin grandissant d'une meilleure connaissance des processus à l'origine de l'AMV pour en donner une définition adéquate.

En terme de circulation atmosphérique, l'Oscillation Nord Atlantique (*North Atlantic Oscillation*, NAO) est le mode de variabilité préférentiel du bassin de l'Atlantique Nord à grande échelle (Hurrell 1995). La fluctuation des échanges de masses d'air entre le centre de haute pression de l'anticyclone des Açores et le centre de basse pression de la dépression d'Islande entraîne une modification du gradient de

Néanmoins, des études ont souligné que cette méthode de calcul d'indice se base sur des hypothèses de linéarité de l'effet des forçages peu réalistes (Trenberth and Shea 2006, Enfield and Cid-Serrano 2010, Marini and Frankignoul 2013, Tandon and Kushner 2015). En partant sur les mêmes base que Enfield et al. (2001), Enfield and Cid-Serrano (2010) ont ajouté un terme quadratique dans l'estimation de la tendance de la NASST sur la période historique pour retirer son signal forcé (Fig. 1.13). Trenberth and Shea (2006) ont quant à eux proposé de calculer un indice AMV comme la différence entre la NASST et la SST globale, en soutenant le fait que cette dernière permet d'obtenir une estimation de l'effet global des forçages externes (Fig. 1.13). L'effet des forçages pourrait néanmoins être plus complexe dans leur évolution spatio-temporelle (Ting et al. 2009). Pour en tenir compte, les méthodes plus récentes s'appuient sur les simulations numériques d'un ensemble de modèles climatiques pour estimer la composante forcée de l'Atlantique Nord, qui est en-

pression atmosphérique de surface (*Sea Level Pressure, SLP*) entre le front polaire et le front tropical de l'Atlantique Nord.

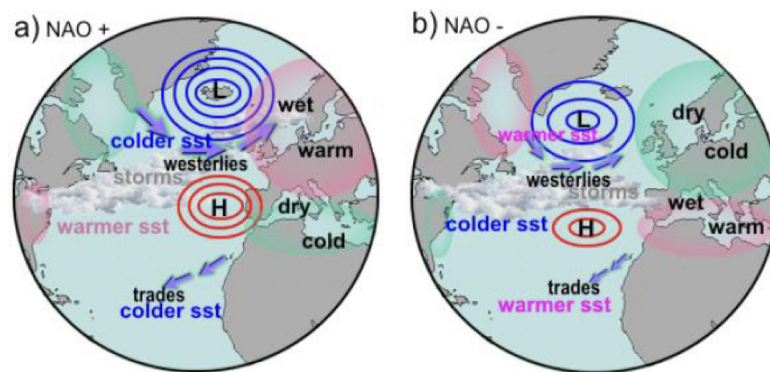


FIGURE 1.14 | Représentation schématisée des effets de la NAO sur les régions adjacentes de l'Atlantique Nord. Les effets sont exprimés en terme de température et de précipitations hivernales lors de sa phase positive (a) et de sa phase négative (b). H est le point de haute pression de l'anticyclone des Açores et L est le point de basse pression de la dépression Islandaise. Source : figure de Bojariu (2003).

Lorsque les deux centres d'action s'intensifient simultanément, le courant-jet de l'Atlantique venant de l'Ouest est dévié vers le Nord, il s'agit d'une phase de NAO positive. À l'inverse, le courant-jet est dévié vers le Sud lorsque les deux centres d'action perdent simultanément en intensité. Le courant-jet chargé en nuages, amène ainsi en phase positive de NAO des hivers anormalement froids et secs en Europe du Sud, Maghreb et au nord-est de l'Amérique du Nord, et des hivers anormalement chauds et humides en Europe du Nord et au sud-est de l'Amérique Nord (Cassou et al. 2004, Cattiaux et al. 2010, Casado et al. 2013, Fig. 1.14). On retrouve les structures météorologiques inverses lors d'une phase négative de NAO (Fig. 1.14). De par ses impacts de grande échelle, il a été montré que la NAO impacte de manière significative les environnements marins et terrestres (Mysterud et al. 2001, Drinkwater et al. 2003), mais aussi les sociétés humaines. Par exemple, il a été récemment montré que la NAO joue sur le niveau marin de certaines régions qui bordent l'Atlantique Nord (Kenigson et al. 2018). De plus, de par sa signature sur les structures météorologiques hivernales de l'Atlantique Nord et des régions avoisinantes, la NAO impacte les marchés de l'énergie européens et nord-américains sous forme de modifications de la consommation et de la production (d'énergies renouvelables uniquement) (Ely et al. 2013, Ravestein et al. 2018).

A l'inverse de l'AMV, la NAO possède une variabilité inter-annuelle beaucoup plus forte puisqu'il s'agit d'un mode de variabilité atmosphérique, tel que l'avaient proposé Frankignoul and Hasselmann (1977) pour un tel type de phénomène avant même que Hurrell (1995) ait officiellement identifié le mode. Néanmoins, la NAO n'est pas considérée comme un mode chaotique et imprévisible (Dunstone et al. 2016). Notamment, une récente étude de Smith et al. (2020) montre que la NAO aurait une prévisibilité décennale bien plus forte qu'on ne le pensait auparavant. On suppose depuis longtemps que la basse fréquence des fluctuations observées de la NAO pourrait être due à son interaction avec l'océan. En effet, on suggère que le forçage stochastique de la NAO aux hautes latitudes modifie les flux de chaleurs de l'océan, ce qui altère le gradient de densité vertical et horizontal de l'Atlantique sub-polaire, et joue donc un rôle sur la convection océanique et la formation d'eaux profondes (Delworth and Mann 2000, Eden and Jung 2001, Delworth et al. 2016). C'est

de cette manière indirecte, via des modifications de l'AMOC, que l'on suspecte un effet de la NAO sur l'AMV, ce qui prend 10 à 20 ans pour se mettre en place (Robson et al. 2012). L'effet de l'AMV sur la NAO est beaucoup plus en phase dans le temps. La signature spatiale de l'AMV, caractérisée par un tripôle de SST, joue également sur le gradient de pression de grande échelle (Cassou et al. 2004). Selon son signe, l'AMV peut modifier la NAO de par des changements dans le transfert d'énergie via les flux de chaleurs associés à l'évaporation de l'océan (Czaja and Frankignoul 2002, Wen et al. 2005, Peings and Magnusdottir 2014). L'AMV joue également un rôle sur la formation de glace de mer Arctique (Miles et al. 2014), qui à son tour modifie la pression atmosphérique dans la région, et modifie donc les variations de la NAO (Koenigk et al. 2016, Caian et al. 2018).

Il convient de préciser que l'AMV et la NAO sont considérés comme des modes de variabilité de par la cohérence des structures climatiques qui leur sont associées. L'AMOC est quant à elle considérée comme une grande structure du système climatique, qui joue un rôle sur les variations des modes de variabilité tels que la NAO et l'AMV.

1.2.2 Modes de variabilité du Pacifique

Le Pacifique couvre près d'un tiers de la surface du globe et s'étend sur plus de 150° de longitude à l'équateur. Il est pleinement ouvert sur l'océan austral et est également ouvert sur l'océan Arctique via le détroit de Béring. Ce gigantesque océan est dominé par un phénomène de grande échelle connu sous le nom de *El Niño Southern Oscillation* (ENSO). *El Niño* (le fils) est le nom d'origine religieuse qui fut initialement donné par les pêcheurs péruviens qui avaient noté un courant chaud récurrent, se dirigeant vers le sud et longeant la côte. Ils notèrent aussi le phénomène inverse, à savoir que ce courant pouvait être anormalement froid, qu'ils appelèrent *La Niña* (la fille) par opposition au phénomène *El Niño*. Aujourd'hui on se réfère à ces deux termes pour une anomalie qui se répand du Pacifique central équatorial jusqu'à la côte sud-américaine équatoriale. Cette anomalie de température n'est en fait que la signature climatique la plus visible d'un phénomène de plus grande échelle. Comme expliqué dans la section 1.1.2, la rotation de la Terre fait que les vents vont majoritairement d'est en ouest au sein de la bande tropicale. La tendance générale au sein du Pacifique est donc que des eaux chaudes, advectées par les vents réguliers issus de la circulation des cellules de Hadley, vont s'accumuler à l'Ouest des tropiques. La région dans laquelle les eaux chaudes de

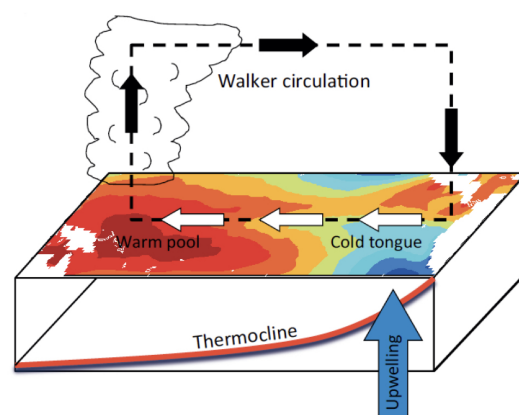


FIGURE 1.15 | Représentation schématique et simplifiée de la cellule de Walker du Pacifique lorsque le phénomène ENSO est en phase neutre. À droite du schéma se trouve le continent sud-américain, et à gauche se trouve le continent océanien. Sur l'océan les couleurs rouges indiquent des anomalies de températures positive et les couleurs bleues indiquent des anomalies négatives. Plus les couleurs sont foncées, plus les anomalies sont grandes en valeur absolue. Les anomalies représentées ici sont uniquement schématiques et ne découlent pas de données réelles. Les courants atmosphériques en altitude sont indiqués par des flèches noires et les courants atmosphériques de surface sont indiqués par des flèches blanches. Source : figure de Christensen et al. (2013).

surface s'accumulent porte le nom de *warm pool* ("piscine chaude"). De par la température élevée des eaux de surface qui s'y trouvent, de grands mouvements de convection atmosphérique y ont lieu. Une partie des cumulonimbus qui s'y forment se déchargent sur l'Asie sous forme de moussons (Tanaka et al. 2004), et une autre s'écoule le long de l'équateur en suivant les vents d'ouest de la haute troposphère, et redescendent à l'est du Pacifique équatorial. Cette cellule de convection zonale se nomme la cellule de Walker, et ses mécanismes furent introduits par Bjerknes (1969) (Fig. 1.15). Elle existe également le long des océans Indien et Atlantique, qui peuvent communiquer les uns avec les autres via des modulations de leurs circulations. Notamment, les gradients thermiques inter-bassin jouent un rôle important sur l'intensité de ces cellules (McGregor et al. 2014). Par contraste avec les eaux chaudes qui s'accumulent à l'ouest du Pacifique équatorial, les masses d'eaux se trouvant le plus à l'est (la région *El Niño*) y sont plus froides. La thermocline est donc plus proche de la surface dans cette région, et donc plus encline à déclencher un phénomène de remontées d'eaux froides. Ceci a notamment lieu lorsque la cellule de Walker du Pacifique s'intensifie et que les vents d'est associés deviennent plus puissants, il s'agit alors du phénomène *La Niña*. À l'inverse, lorsque la circulation de Walker est anormalement faible en intensité, la thermocline gagne en profondeur, les eaux profondes ne remontent plus, ce qui crée une anomalie positive dans le Pacifique équatorial Est, synonyme d'un événement *El Niño*. Cette région du Pacifique peut également se trouver dans une phase dite neutre, c'est-à-dire qu'on ne discerne pas particulièrement de phase *El Niño* ou *La Niña*. Derrière le phénomène ENSO, existe une basse fréquence, notamment due à des réactions en chaîne associées aux phases *El Niño* et *La Niña* (Fig 1.16).

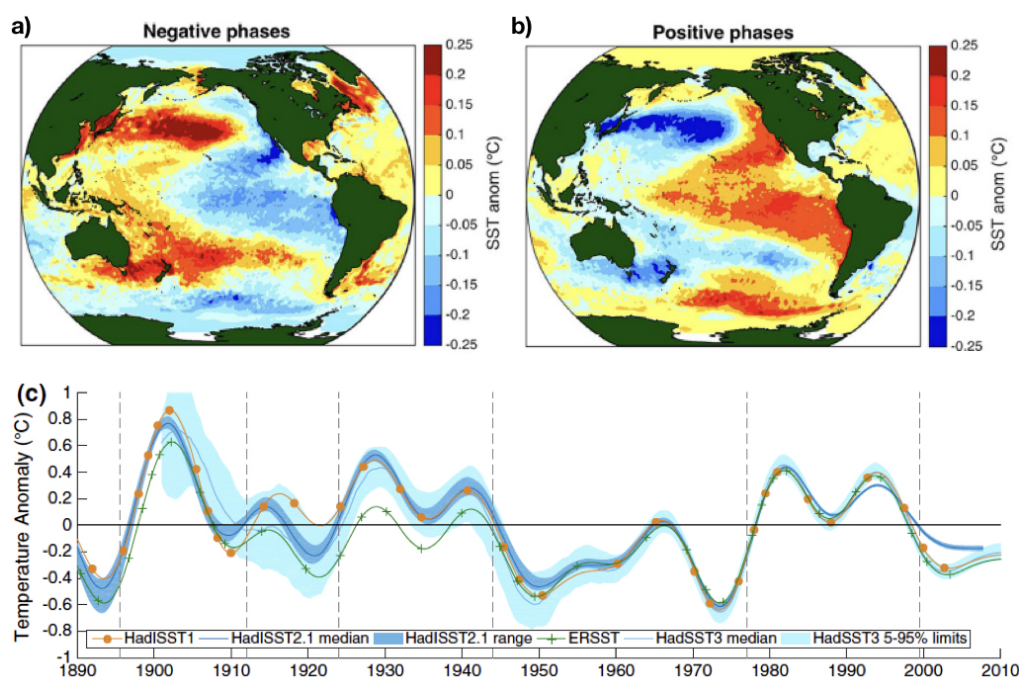


FIGURE 1.16 | Effets des variations de l'IPV sur la SST du Pacifique. a) Carte composite (moyenne sur le temps) lors d'une phase d'IPV positive sur les données de HadISST (Rayner et al. 2003) (voir c). b) Carte composite lors d'une phase d'IPV négative sur les données de HadISST (Rayner et al. 2003) (voir c). c) Indice IPV calculé à partir de la différence entre la moyenne des anomalies du Pacifique équatorial Est et la moyenne des anomalies du Pacifique Sud et Nord (Henley et al. 2015). À partir des jeux de données HadISST (Rayner et al. 2003) et ERSSTv3b (Smith et al. 2008) Source : figures de Henley et al. (2015).

En effet, si le fait que la cellule de Walker s'intensifie implique une remontée d'eaux froides dans le Pacifique équatorial Est, alors le gradient thermique ouest-est en est d'autant plus renforcé, ce qui implique des vents d'est tropicaux plus puissants, signifiant que la cellule de Walker s'intensifie davantage et s'étend sur l'intégralité du Pacifique équatorial. À l'inverse, la phase *El Niño* implique un gradient thermique moins important et donc des vents plus faibles et une désintensification de la cellule de Walker : sa taille se rétrécit et la région de convection atmosphérique se décale vers l'est. Dans les deux cas, la réaction en chaîne fait que l'on observe des périodes de une ou plusieurs décennies avec une majorité d'évènements *La Niña*, et d'autres avec une majorité d'évènements *El Niño*. On appelle le phénomène à basse fréquence correspondant la Variabilité Inter-décennale du Pacifique (*Interdecadal Pacific Variability*, IPV) (Parker et al. 2007, Henley et al. 2015) (Fig. 1.16). Dans chacune des phases, les eaux de la *warm pool* sont redistribuées vers de plus hautes latitudes. Ceci fait que l'empreinte spatiale de l'IPV est très caractéristique : en phase positive on retrouve des anomalies négatives dans le nord, l'ouest et le sud du Pacifique, et des anomalies positives du centre des tropiques jusqu'à l'est (Fig. 1.16). On observe la configuration inverse lors d'une phase d'IPV négative (Fig. 1.16).

1.2.3 Intérêt pour la prévision décennale

Hawkins and Sutton (2009) ont réalisé une étude dans laquelle ils quantifient les différentes sources d'incertitudes dans la prévision du climat à venir, selon les modèles de circulation globale de la génération CMIP3 (section 1.1.4).

Ils identifient trois sources majeures d'incertitudes dans la prévision décennale de cet ensemble de modèles (Fig. 1.17). La première est associée à la différence entre la physique et les schémas numériques des différents modèles ; la seconde est associée aux différents scénarios d'émission de gaz à effet de serre imposés à ces modèles ; la dernière est associée à la variabilité interne du climat. Cette étude souligne en particulier l'importance de réduire l'incertitude associée à la variabilité interne, notamment en approfondissant les connaissances des mécanismes qui la composent. Ils montrent qu'au delà de 30-40 ans de prévision de la moyenne globale, la variabilité interne joue un rôle moins important en proportion que les autres sources d'incertitudes. En revanche, ces auteurs soulignent que la variabilité climatique interne compte pour une très large part de l'incertitude dans les projections climatiques à court terme (de dix ans à quelques décennies), et ce, d'autant plus pour la prévision de moyennes régionales (Hawkins and Sutton 2009, Fig. 1.18).

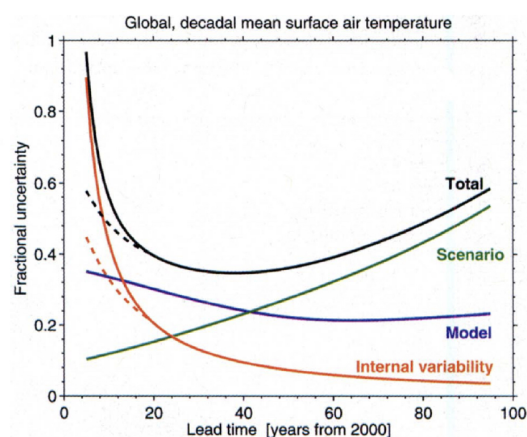


FIGURE 1.17 | Importance relative de chaque source d'incertitudes dans la prévision décennale de la moyenne globale de la température atmosphérique. Les modèles utilisés sont ceux de la génération CMIP3 (voir section 1.1.4). Chaque ligne représente l'incertitude fractionnelle associée à l'une des sources pour la prévision décennale de la moyenne globale, relative à l'année 2000 (voir Hawkins and Sutton 2009 pour la méthode). Les lignes pointillées montrent la réduction d'incertitude atteignable de par l'initialisation des modèles au travers de l'assimilation de données océaniques observées. Source : figure de Hawkins and Sutton (2009)

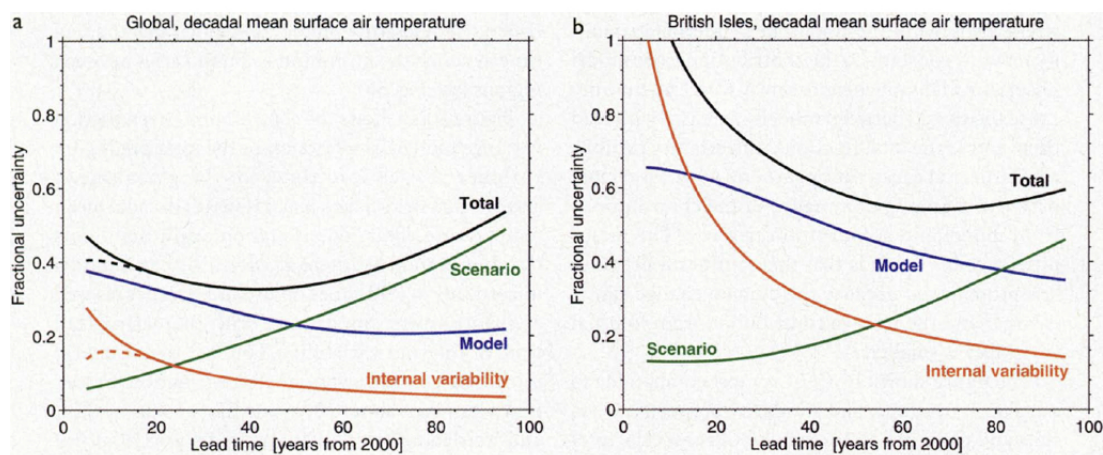


FIGURE 1.18 | Importance relative de chaque source d'incertitudes dans la prévision décennale de la moyenne globale et d'une moyenne régionale de la température atmosphérique. a) Même figure que Fig. 1.17, exceptée que les incertitudes sont calculées relativement à la moyenne de 1970-2000 (voir Hawkins and Sutton (2009) pour la méthode). b) Identique à a) mais appliqué à la moyenne régionale simulée sur les îles britanniques. *Source : figures de Hawkins and Sutton (2009).*

Les auteurs montrent que réduire l'incertitude associée à la variabilité climatique interne est d'une importance primordiale, en soutenant que la prévision décennale à l'échelle régionale est essentielle pour prendre des décisions politiques adéquates en terme de planification (Fig. 1.18, Hawkins and Sutton 2009). Des études postérieures ont également montré l'importance de réduire cette incertitude pour améliorer la prévision décennale à l'échelle continentale (Deser et al. 2012, Deser et al. 2020). Une meilleure compréhension de la variabilité climatique interne peut donc constituer un tremplin pour la modélisation du climat et sa prévision, aux échelles décennales et multi-décennales (Hawkins and Sutton 2009, Cassou et al. 2018). Ceci permettrait d'une part d'implémenter des modèles numériques plus réalistes et d'autre part de sélectionner au sein de ces modèles, un sous-ensemble de ceux qui reproduisent au mieux les comportements connus des modes de variabilité climatique interne. On parle d'approche de sélection par contraintes émergentes (Hall et al. 2019).

Alors que les émissions anthropiques de gaz à effet de serre n'ont cessé de croître de manière exponentielle depuis la fin du XIX^e siècle, on distingue dans l'évolution de la température globale une période de hiatus entre 1998 et 2012 (Kaufmann et al. 2011, Kosaka and Xie 2013, England et al. 2014, Medhaug et al. 2017, Fig. 1.19). Cette période est marquée par un signal négatif très fort dans la région *El Niño*, accompagné de signal positif dans le Pacifique Ouest, Nord et Sud. Autrement dit, cette période de hiatus est dominée par une phase négative d'IPV très caractéristique (Kaufmann et al. 2011, Kosaka and Xie 2013, section 1.2.2). L'étude des moteurs pouvant amener une telle phase négative de l'IPV a ainsi fait l'objet d'intenses recherches au cours des dernières années. McGregor and Timmermann (2011) ont montré dans une étude basée sur un unique modèle numérique, qu'une éruption volcanique suffisamment importante aurait une probabilité plus forte d'entraîner des anomalies négatives dans la région *El Niño*, ce qui pourrait entraîner une phase négative d'IPV les années suivantes. Smith et al. (2016) ont, dans cette lignée, émis l'hypothèse que le début de la phase négative de l'IPV en 1998 pourrait être une réponse retardée dans le temps à l'éruption du volcan Pinatubo (Philippines, île de *Luçon*) en 1991. Néanmoins, ces hypothèses peuvent être nuancées puisque d'autres études (Ohba et al. 2013, Khodri et al. 2017), basées sur des ensembles de modèles numériques, ont

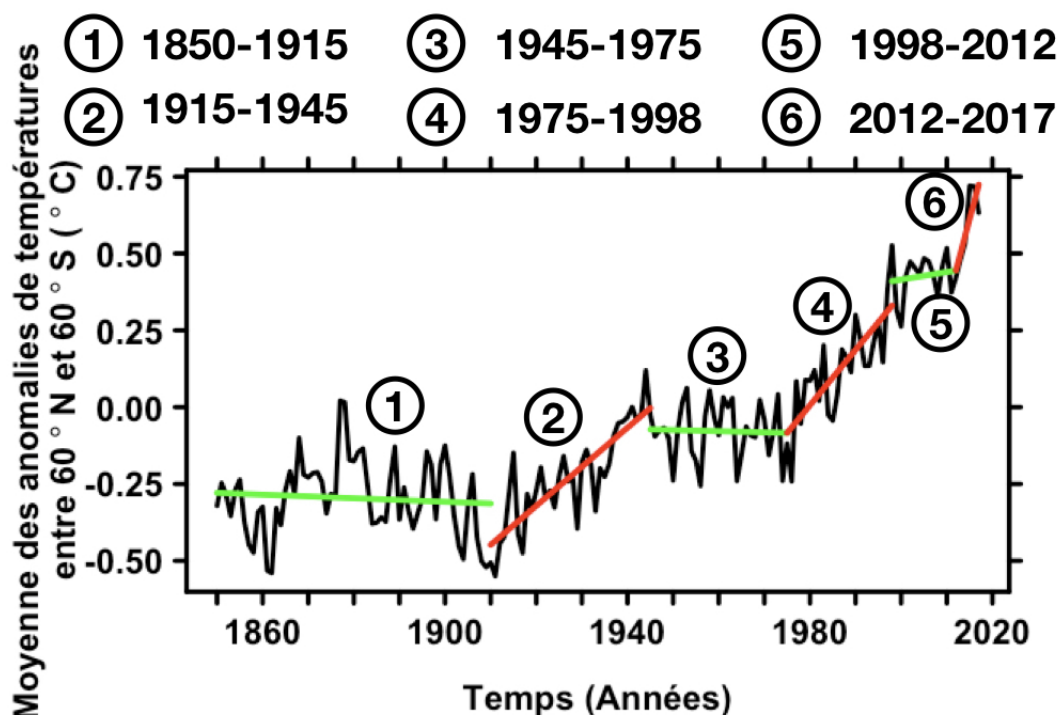


FIGURE 1.19 | Moyenne globale des anomalies de température observées entre 1850 et 2017. Les lignes rouges représentent la tendance linéaire pour des périodes au cours desquelles celle-ci est significativement positive au seuil de significativité de 90% (section 2.3.5). Les lignes vertes représentent la tendance linéaire pour des périodes au cours desquelles celle-ci n'est pas significativement positive au seuil de significativité de 90% (section 2.3.5). Source : jeu de données HadCRUT4 (Morice et al. 2012).

montré des résultats qui contredisent ceux de McGregor and Timmermann (2011). Ces autres études trouvent en effet que des éruptions stratosphériques riches en aérosols sulfatés sont plutôt enclines à désintensifier la cellule de Walker du Pacifique, et donc à réduire la remontée d'eaux froides équatoriale, et ainsi stimuler une phase d'*El Niño*. Kosaka and Xie (2013) ont quant à eux montré par simulations numériques, qu'une telle phase négative d'IPV peut être causée par de la variabilité climatique interne uniquement, notamment par un processus de réaction en chaîne causé par un gradient de SST ouest-est suffisamment important au sein du Pacifique équatorial. En réponse à un tel gradient thermique, la cellule de Walker du Pacifique s'intensifie, et provoque une advection vers l'ouest plus importante et une remontée d'eaux profondes froides accrue dans le Pacifique Est, ce qui va renforcer le gradient thermique. Puisque la cellule de Walker s'intensifie, les eaux anormalement chaudes de la *warm pool* sont rapidement distribuées au nord et au sud du Pacifique, de part et d'autre de l'équateur, ce qui est typique d'une phase négative d'IPV. L'étude de England et al. (2014) vient soutenir les travaux de Kosaka and Xie (2013), notamment en notant des alizés particulièrement puissants (sans précédent sur la période observationnelle) dans le Pacifique équatorial, signe d'une circulation plus intense au sein de la cellule de Walker. À cet égard, l'étude réalisée par McGregor et al. (2014) montre qu'un tel gradient thermique au sein du Pacifique pourrait avoir été forcé par la partie tropicale de l'Atlantique Nord anormalement chaude dans les années 1990 (cf. AMV positive). Cette réponse négative de l'IPV 10 ans après une entrée en phase positive de l'AMV est cohérente avec ce que l'on observe sur l'ensemble de la période instrumentale (Marini and Frankignoul 2013). Ainsi, une telle phase négative de l'IPV pourrait être le fruit de changements préalables dans les

variations de l'AMV, et donc possiblement dans celles de l'AMOC.

Le rôle de l'AMOC dans la redistribution globale de la chaleur et dans l'évolution du climat ne fait plus aucun doute depuis des décennies (Ganachaud and Wunsch 2000, Wunsch 2000, Wunsch 2002, Knight et al. 2005, Persechini et al. 2013, Buckley and Marsahll 2016, Zhang et al. 2019a). Cependant, du fait du peu d'observations, une reproduction réaliste de sa structure de grande échelle et de la distribution de chaleur associée constitue un élément crucial dans l'amélioration des prévisions par modèles numériques.

L'AMOC est par ailleurs reconnue pour être un élément de bascule du système climatique, capable de modifier profondément son état (Collins et al. 2019). À cet égard, des simulations de modèles numériques ont montré que la fonte rapide de la banquise arctique et de la calotte groenlandaise pourrait avoir un effet irréversible de réduction de l'AMOC, et donc du transport méridien de chaleur par l'océan (Swingedouw et al. 2007, Bakker et al. 2016). Ceci s'explique par l'apport d'eaux douces dans les couches océaniques de surface suite à la fonte de la glace, ce qui va les désaliniser et donc réduire leur densité, et ainsi ralentir la convection (Swingedouw et al. 2007, Bakker et al. 2016). Il existe donc de fortes suspicions quant à un ralentissement à venir de l'AMOC, qui a même été qualifié de *très probable* avant 2100, même avec des scénarios optimistes en matière d'émission de gaz à effets de serre, selon le dernier rapport SROCC (*Special Report on the Ocean and Cryosphere*, Collins et al. 2019) du GIEC. À ce sujet, il a même été envisagé un arrêt quasi-complet de cette circulation, bien que ceci ait été qualifié de *très improbable* avant 2100 selon ce même rapport, même sous les scénarios les plus pessimistes en matière d'émission de gaz à effet de serre (Collins et al. 2019).

En matière d'observations, Smeed et al. (2018) indiquent que l'AMOC était significativement plus intense entre 2004 et 2007 qu'entre 2008 et 2017. Néanmoins, ces 14 années de mesures ne permettent pas de déterminer de manière robuste s'il s'agit d'une tendance de long terme puisque les données instrumentales ne permettent pas l'étude des variations multi-décennales de l'AMOC. Il n'est donc pas possible de discerner si la tendance observée est comparable en terme d'amplitude à des variations précédentes. Ainsi, si la tendance récente est qualifiable de "sans précédent" sur une période plus longue, on pourrait très certainement attribuer les récentes variations de l'AMOC aux forçages radiatifs anthropiques. Dans le cas contraire, cela supposerait que les effets escomptés par les modèles numériques de l'activité humaine sur l'AMOC ne sont pas encore visibles.

Pour répondre à cette problématique, des études ont tenté d'estimer le comportement naturel de l'AMOC sur de plus longues périodes de temps. Pour ce faire, celles-ci se basent soit sur des modèles numériques (Zhang 2008, Swingedouw et al. 2013), soit sur des enregistrements ou des reconstructions paléoclimatiques (Rahmstorf et al. 2015, Thornalley et al. 2018), soit sur la signature de l'AMOC sur la SST de l'Atlantique Nord (Caesar et al. 2018). Selon Rahmstorf et al. (2015) et Thornalley et al. (2018), la tendance négative de l'AMOC pourrait avoir commencé dès la fin du XIX^e siècle, et serait de nos jours sans précédent depuis au moins deux millénaires. Cette conclusion soutient donc l'idée selon laquelle l'activité humaine aurait déjà grandement perturbé l'AMOC, qui demeurerait dans un état relativement stable depuis des siècles. En s'appuyant sur un indice de circulation basé sur la SST dans la

régions du SPG de l'Atlantique Nord, qui remonte donc en 1870, Caesar et al. (2018) ont montré que l'on discerne une forte tendance de ralentissement de l'AMOC depuis le début des années 1960. Bien que les preuves d'un ralentissement anthropique de l'AMOC s'accumulent, sa prévisibilité, notamment par le biais de l'AMV, représente un enjeu crucial dans la prévision décennale. En effet, Sgubin et al. (2017) ont montré qu'à l'horizon 2100, il se pourrait que l'on observe un ralentissement brutal de la convection dans la région du SPG de l'Atlantique Nord, entraînant un ralentissement partiel de l'AMOC et un refroidissement de l'Atlantique Nord (cf. AMV négative). Les auteurs ont montré que si un tel événement arrivait selon les scénarios optimistes en matière d'émission de gaz à effet de serre, il se pourrait que l'on observe d'importants refroidissements locaux dans les régions qui bordent l'Atlantique Nord, notamment en Europe.

Les intérêts politiques, sociétaux et économiques de la prévision décennale sont multiples et reposent sur la compréhension des processus de variabilité interne qui continuent de diriger les fluctuations du climat à ces échelles de temps, sur un plan global comme régional (Hawkins and Sutton 2009). Mais les 150 ans d'observation de la SST et les 14 ans d'observation de l'AMOC constituent un échantillon bien trop faible pour comprendre la variabilité climatique multi-décennale. Le recours aux observations paléoclimatiques permet donc de disposer de mesures indirectes des variations climatiques passées. Cependant, chacune de ces données possède des incertitudes de mesure ainsi que des incertitudes temporelles et statistiques. La reconstruction de signaux climatiques globaux, comme ceux associés à de grands modes de variabilité par combinaison appropriée de données paléoclimatiques, peut permettre de réduire ces différentes incertitudes associées. À cet égard, la période du dernier millénaire fait l'objet d'un intérêt scientifique particulier pour mieux comprendre la variabilité climatique interne et naturelle aux échelles décennales à pluricentennales.

1.3 Le dernier millénaire comme outil de compréhension de la variabilité décennale

1.3.1 Une période idéale pour comprendre la variabilité naturelle et ses impacts

Une récente étude de glaciologie (Rubino et al. 2019) a donné une nouvelle estimation de la concentration de CO₂ atmosphérique du dernier millénaire, à partir de micro-bulles de gaz emprisonnées dans un grand nombre de carottes de glaces (section 1.1.4). Celles-ci ont été prélevées dans la région du *Law Dome* (Antarctique de l'Est), connue pour être une région de forte accumulation de neige, et donc idéale pour étudier le climat du passé avec une fréquence annuelle (section 1.1.4). Cette concentration est donnée en parties par millions (ppm), ce qui correspond au nombre estimé de molécules de CO₂ par million de molécules contenues dans l'atmosphère (Fig. 1.20).

L'étude de Rubino et al. (2019) suggère qu'avant 1800, la concentration de CO₂ dans l'atmosphère, a varié de 13.2 ppm entre son minimum (271.2 ppm, Rubino et al. 2019) et son maximum (284.2 ppm, Rubino et al. 2019) (Fig. 1.20). En comparaison, la différence entre le maximum de cette même période pré-1800 et le niveau

de concentration atteint en 2019 (411.4 ppm, Tans and Keeling 2020) est d'un peu plus de 126 ppm (Fig. 1.20, Rubino et al. 2019). C'est dans cette faible variation relative de la concentration de CO₂ atmosphérique que se trouve l'intérêt d'étudier les variations climatiques du dernier millénaire pour comprendre des phénomènes climatiques internes, ainsi que leurs réponses aux forçages naturels. En effet, un faible forçage radiatif de CO₂ implique une forte contribution de la variabilité interne et des forçages naturels aux évolutions du climat sur cette période.

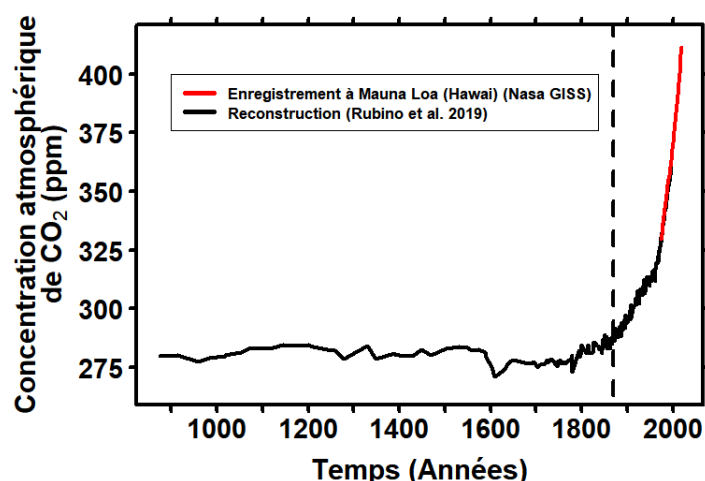


FIGURE 1.20 | Estimations de la concentration de CO₂ atmosphérique moyenne. Reconstruction millénaire (noir) et observations instrumentales (rouge), données en parties par millions (ppm). Source : données de Rubino et al. (2019) (noir) et Tans and Keeling (2020) (rouge).

Puisqu'il s'agit d'une période proche dans le temps, les différences par rapport à l'ère industrielle en matière de conditions orbitales (positions relatives de la Terre et du Soleil) et géologiques (positions relatives des océans et continents) sont très faibles. On peut donc supposer que les processus internes qui sont à l'oeuvre de nos jours sont sensiblement similaires à ceux de cette période.

Rappelons que nous vivons actuellement un contexte de changement climatique significatif, qui ne semble que commencer, et qui devrait rapidement bouleverser les modes de vies que nous avons connus ces dernières décennies. Il existe de ce fait un intérêt particulier à observer comment les variations climatiques ont pu, et donc pourraient à l'avenir, impacter les sociétés humaines. Bien que les sociétés humaines du passé étaient, semble-t-il, moins résilientes aux changements environnementaux qu'aujourd'hui, il faut toutefois souligner que le réchauffement observé depuis 1870 est deux à trois fois supérieur à la différence entre le minimum et le maximum de températures observées sur les 2000 années précédentes (Abram et al. 2016). Différentes études ont ainsi montré que de grands bouleversements sociétaux ont pu être causés par la variabilité climatique naturelle. Par exemple, Singh et al. (2018a) ont montré, au travers de reconstructions de précipitations et de simulations de modèles, que la grande sécheresse de 1876-1878 et les famines qu'elle a causées en Asie, au Brésil et en Afrique furent les conséquences d'un décalage anormal de l'ITCZ vers le nord, en partie dû à une possible phase positive de l'AMV, associée à une configuration océanique globale particulière (section 1.2.1, Singh et al. 2018a). En matière d'impacts d'éruptions volcaniques, bien que l'on s'éloigne légèrement du dernier millénaire, Sigl et al. (2015) ont montré que les deux éruptions massives ayant eu lieu au VI^e siècle (540 et 574) ont été les catalyseurs des nombreuses famines et des vagues épidémiques de la peste de Justinien qui était à son paroxysme à cette époque. Les auteurs soulignent que ces crises ont provoqué un déclin socio-économique et démographique dans de nombreuses régions, notamment dans l'empire Byzantin, considérablement impacté à l'époque (Sigl et al. 2015).

1.3.2 Deux séquences climatiques bien distinctes

Les forçages anthropiques au cours de la période du dernier millénaire sont faibles par rapport aux dernières décennies (Rubino et al. 2019, Fig. 1.20). Il existe cependant un large consensus sur le fait qu'on peut distinguer, sur de nombreuses archives naturelles, deux périodes contrastées en matière de températures : la relativement chaude anomalie climatique médiévale (*Medieval Climate Anomaly*, MCA), et le relativement froid petit âge de glace (*Little Ice Age*, LIA) (Bradley and Jones 1993). Il n'existe en revanche pas de consensus sur l'évolution spatio-temporelle exacte des variations climatiques au cours de ces périodes. On considère en général que le MCA se situait au début du millénaire (environ 900-1300), suivi du LIA (environ 1400-1800) (Guiot 2012). Ces dernières ont été identifiées pour la première fois par Lamb (1965), en étudiant des archives naturelles de l'Europe et de l'Atlantique. LaMarche (1974) a à son tour identifié deux périodes similaires, mais cette fois-ci dans des archives nord-américaines, dans la région de la Californie. Ainsi, une telle téléconnexion sur des échelles de temps centennales de part et d'autre de l'Atlantique Nord a soulevé la question d'une réorganisation du climat à l'échelle globale, comme le suggérait déjà Lamb (1965). Par la suite, à l'aide d'un ensemble de simulations de modèles numériques, et en comparant les variations simulées avec les enregistrements paléoclimatiques existants, Goosse et al. (2005) ont montré que le MCA et le LIA étaient des événements qui s'étendent au moins à l'ensemble de l'hémisphère Nord.

À mesure que de nouvelles mesures paléoclimatiques sont prélevées, on distingue d'autres régions dans lesquelles des changements majeurs ont eu lieu au cours de la transition entre le MCA et le LIA. On peut citer de manière non exhaustive des changements dans : les régimes de mousson de l'Asie de l'Est (Sinha et al. 2007) et ceux de la région sahélienne (Shanahan et al. 2009, Carré et al. 2019); la couverture de glace de mer Arctique (Massé et al. 2008, Kinnard et al. 2011, Jomelli et al. 2016, Halloran et al. 2020); les températures du Moyen-Orient (Kaniewski et al. 2011), de certaines régions de l'Océanie (Griffiths et al. 2016, McGowan et al. 2018), ou encore de l'Amérique du Sud (Luterbacher et al. 2011, Apaéstegui et al. 2014). Cette profusion de preuves que le climat a significativement changé d'état entre le MCA et le LIA soulève un intérêt encore plus fort pour l'étude de ces périodes. Néanmoins, une étude récente a montré qu'au cours du MCA comme du LIA, les maxima et minima de températures sont atteints à des moments relativement hétérogènes selon les différentes régions du globe (Neukom et al. 2019). Si bien que même si l'on peut distinguer des différences significatives de températures aux échelles pluri-centennales, la synchronisation spatiale des températures au cours du MCA et du LIA est incomparable à celle observée au cours de l'ère industrielle (Neukom et al. 2019).

La proximité dans le temps de cette période fait également qu'on dispose de reconstructions des forçages naturels robustes, notamment grâce à l'étude de carottes de glace (section 1.1.4, Fig 1.20, 1.21, Vieira et al. 2011, Sigl et al. 2015, Lean 2018, Rubino et al. 2019). Ainsi la reconstruction paléoclimatique peut être un parfait outil pour comprendre les contributions relatives de la variabilité climatique interne et des forçages radiatifs dans la mise en place de séquences climatiques telles que

le MCA et le LIA. À ce sujet, il n'existe pas réellement de consensus sur les processus ayant engendré leur mise en place. Bien qu'il semblerait qu'au vu de certains enregistrements paléoclimatiques les variations de la TSI aient varié dans de

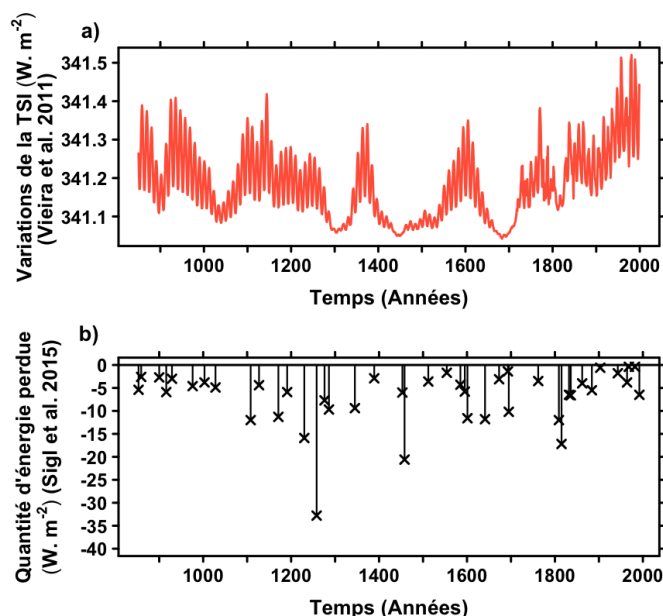


FIGURE 1.21 | Reconstructions des forçages volcaniques et solaires sur le dernier millénaire. a) Reconstruction de la TSI en W.m^{-2} sur le dernier millénaire. Source : données de Vieira et al. (2011). b) Reconstruction des éruptions volcaniques majeures du dernier millénaire, estimée en perte moyenne de W.m^{-2} . Source : données de Sigl et al. (2015).

évaluée comme la troisième plus importante du dernier millénaire (Sigl et al. 2015, Fig. 1.21). Brönnimann et al. (2019) indiquent par ailleurs que deux autres éruptions très puissantes ayant eu lieu au cours de cette période ont pu potentialiser celle de Tambora. Ces éruptions ont eu lieu en 1808 (ou 1809 selon les estimations) et en 1831, mais leurs localisations respectives sont à l'heure inconnues. Quant à eux, Miller et al. (2012), en se basant sur des simulations de modèles et des données paléoclimatiques, pointent le regroupement de très grosses éruptions ayant eu lieu au cours du XIII^e siècle et le rôle qu'il a pu jouer sur la persistance d'un volume cryosphérique Arctique élevé. Cette période comprend en effet la plus grosse éruption des 1000 dernières années (Sigl et al. 2015, Fig. 1.21). Il s'agit de l'éruption du volcan Samalas (Indonésie, île de Lombok), ayant eu lieu en 1257 ou 1258 selon les estimations, identifiée par Lavigne et al. (2013). Cette éruption précède par ailleurs la quatrième plus importante du dernier millénaire, ayant eu lieu en 1230, mais dont la localisation est pour le moment inconnue (Sigl et al. 2015, Fig. 1.21). Venant s'ajouter aux effets des éruptions, d'autres études ont montré par des simulations de modèles que des processus de variabilité climatique interne associés à des changements dans la circulation océanique et dans ses échanges avec l'atmosphère, peuvent également expliquer la transition entre les deux périodes (Schleussner et al. 2015, Moreno-Chamarro et al. 2016). Goosse et al. (2012) ont à ce sujet montré par des réanalyses obtenues à l'aide d'une méthode de filtre à particules que de tels changements de régimes pouvaient s'expliquer par de très faibles variations du bilan radiatif entre le MCA et le LIA (e.g. 0.25 W.m^{-2}), notamment aux hautes latitudes.

moindres proportions et auraient donc eu peu d'influences sur les variations climatiques (Schurer et al. 2014), de nombreuses études suggèrent que les éruptions volcaniques pourraient être responsables de la transition entre les deux périodes (Crowley et al. 2008, Miller et al. 2012, Brönnimann et al. 2019). Si l'on regarde plus en détail, Brönnimann et al. (2019) montrent, à l'aide de reconstructions paléoclimatiques et de modèles, que la fin du LIA, ayant eu lieu la première moitié du XIX^e siècle, et considérée comme sa phase la plus froide (PAGES 2k Consortium 2017), a très largement été influencée par une forte activité volcanique. En effet, on y trouve l'éruption de Tambora (1815),

D'autres études, basées sur des reconstructions de la couverture cryosphérique Arctique, ont mis en évidence des avancées importantes de glaciers et de glace de mer au XII^e siècle (Kinnard et al. 2011, Jomelli et al. 2016), c'est-à-dire bien avant la fin de la période conventionnellement définie du MCA (~900-1300), mais aussi bien avant le regroupement de puissantes éruptions volcaniques du XIII^e siècle. Pour finir, l'étude de Trouet et al. (2009) suggère que la NAO est passée d'un état positif permanent avant 1400, à un état plutôt neutre ensuite. Les auteurs soutiennent en effet que cet état permanent de la NAO a amené des conditions hivernales relativement chaudes et humides, et donc *a priori* moins hostiles, au cours du MCA. Ils soutiennent ensuite que s'en sont suivies des conditions plus variables à l'échelle inter-annuelle, avec notamment la présence de plusieurs hivers froids et secs en Europe au cours du LIA. Toutefois, le protocole de Trouet et al. (2009) ne se base que sur deux enregistrements, ce qui entraîne une grande incertitude comme l'a montré une étude ayant reproduit ce protocole en *pseudo-proxy*, au sein de modèles de circulation globale (Lehner et al. 2012).

1.3.3 Une base de données sans précédent dédiée à la reconstruction paléoclimatique

Le dernier millénaire est une période couverte par un nombre très satisfaisant de données paléoclimatiques, à haute résolution (annuelle), constituant ainsi une période d'étude adéquate pour la variabilité décennale induite par des processus climatiques internes et des forçages radiatifs naturels. À cet égard, le consortium scientifique PAGES 2k (*Past Global Changes 2k*) dirige la mise en place et la maintenance d'une base de données paléoclimatiques qui couvrent, au moins en partie, les deux derniers millénaires et dont la variable reconstruite est la température. La dernière version de cette base de données est disponible depuis 2017 (PAGES 2k Consortium 2017) et comprend près de 700 enregistrements paléoclimatiques dont les fiabilités respectives ont été évaluées et vérifiées (Fig. 1.22). On peut voir qu'au fur et à mesure qu'on s'éloigne dans le temps, le nombre de données

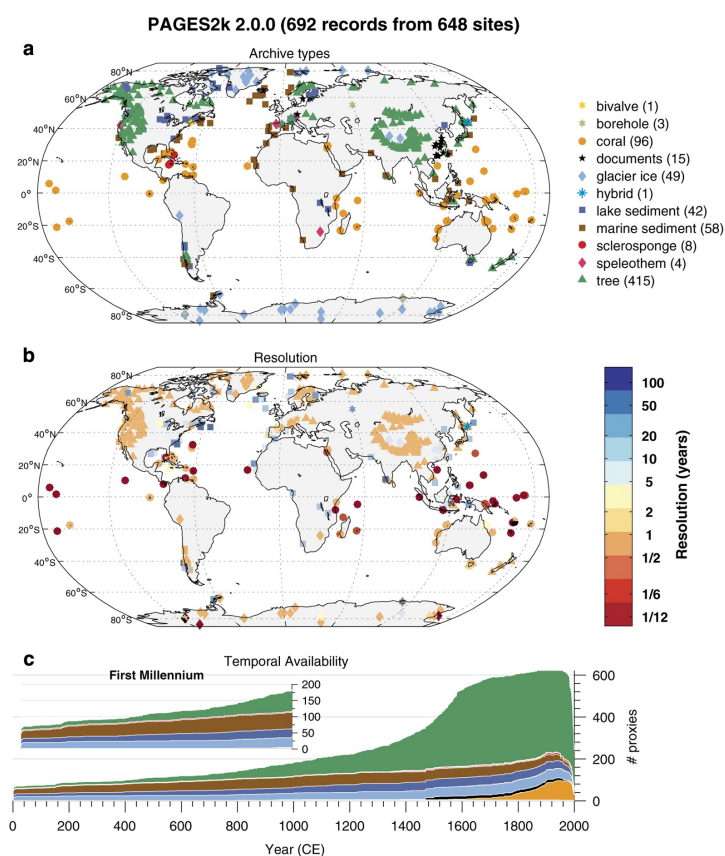


FIGURE 1.22 | Caractéristiques spatio-temporelles des enregistrements paléoclimatiques de la base de données PAGES 2k (PAGES 2k Consortium 2017). a) Localisation, type d'archive. b) Localisation, type d'archive et résolution. c) Période couverte par les enregistrements, par type d'archive naturelle. Source de PAGES 2k Consortium (2017).

On peut voir qu'au fur et à mesure qu'on s'éloigne dans le temps, le nombre de données

disponibles décroît sensiblement. À partir de l'an 1000, il ne reste qu'un peu moins de 180 enregistrements disponibles, et à l'an 0, il n'en reste qu'un peu plus de 70.

Puisque cette base de données se veut apte à la reconstruction paléoclimatique sur les deux derniers millénaires, elle met à disposition un très grand nombre de données à résolution annuelle (>500 , Fig. 1.22b), permettant d'avoir de nombreuses mesures en commun avec les observations sur la période historique, offrant ainsi la possibilité de calibrer des modèles statistiques robustes pour faire des reconstructions. Dans l'ensemble de la thèse, cette base de données servira de socle et va constituer plus de 90% des données utilisées dans les travaux présentés.

1.3.4 *Machine learning* : de nouvelles opportunités pour la reconstruction millénaire

L'augmentation du nombre de données disponibles et des capacités de calculs n'est pas spécifique à la paléoclimatologie. Selon le rapport de 2018 de l'*International Data Corporation* (Reinsel et al. 2018), le volume de données stockées dans le monde est passé de environ 1 zettabytes (ZB, 1 ZB= 10^{21} bytes) en 2010, à 33 ZB en 2018. Les auteurs du rapport estiment ensuite qu'au vu des progrès technologiques, notamment en terme de capacité de calcul et de stockage de données, le volume de données stockées dans le monde devrait atteindre 175 ZB d'ici à 2025. Cette abondance de données, dont la croissance en est tout juste à ses débuts, a fait naître un intérêt grandissant pour la mise en place de programmes informatiques capables de traiter efficacement de grands nombres de données et d'en faire des modèles prédictifs. On appelle ce type d'algorithme "méthodes d'apprentissage statistique", plus fréquemment appelées méthodes de *machine learning*, qui constituent une des branches de l'Intelligence Artificielle (Vapnik 2000).

Le *machine learning* est une stratégie d'approche moderne qui consiste à "apprendre" d'un ensemble de données observées, dont certaines que l'on cherche à modéliser (variables à expliquer) à l'aide d'autres, dont on sait qu'on pourra disposer de nouvelles valeurs (variables explicatives). Le but est en général d'anticiper ce qui va se passer au vu de ce qui a été précédemment appris sur les données, et au vu de nouvelles valeurs de données de variables explicatives.

Les intérêts pour l'application de ces méthodes dans de nombreux champs de la science sont grandissants, notamment ceux dans lesquels de nombreuses données sont disponibles : génomique (Chen and Ishwaran 2012), microbiologie (Qu et al. 2019), démographie (Luo et al. 2015), cosmologie (Ntampaka et al. 2019), mais aussi en étude du changement climatique (Huntingford et al. 2019). Les domaines d'application en dehors du monde de la recherche sont également multiples et façonent de plus en plus le monde qui nous entoure. On peut par exemple citer : la prévision de cours boursiers, l'estimation de risque en actuariat, l'estimation de la solvabilité d'emprunteurs de crédits bancaires, le webmarketing, la vente de billets de transport, l'optimisation de la production agricole ou encore la mise en place de plans d'urbanisme. Dans cette thèse, puisque l'on appliquera ces méthodes dans un contexte paléo, il s'agira de reconstructions et non de prévisions.

Comme indiqué dans la section précédente, les enregistrements paléoclimatiques sont particulièrement intéressants pour le type de reconstruction que l'on cherche à

produire. En effet, puisque l'on commence à disposer d'un nombre convenable d'années d'observations de variables climatiques (>100 ans), plus les enregistrements paléoclimatiques ont des valeurs en commun avec les observations, plus on peut établir des liens et des modèles robustes. Les méthodes de *machine learning* gagnant en robustesse à mesure qu'elles disposent de données, la mise en place de la base de données PAGES 2k (PAGES 2k Consortium 2017) est donc l'occasion parfaite pour introduire ces approches dans un contexte de reconstruction paléoclimatique pluri-centennale. Les notations mathématiques associées à ces techniques seront développées dans le chapitre suivant.

1.3.5 Assimilation de données dans un contexte paléoclimatique

Même lorsque l'on dispose de données instrumentales qui ne sont que des observations de la réalité, on ne dispose jamais de l'état de l'intégralité du système climatique. Pour répondre à ce problème, les méthodes d'assimilation de données sont des approches, souvent statistiques, permettant de faire tendre la trajectoire d'un modèle physique, comme par exemple un modèle de circulation globale, vers des observations, et ainsi de disposer d'un état complet du système tel que simulé par le modèle au vu des observations. Une approche simple d'assimilation de données est la relaxation newtonienne, également connue sous le nom de méthode de rappel, ou méthode de *nudging* (Stauffer and Seaman 1990), qui n'est pas une approche basée sur des méthodes statistiques.

La méthode de rappel provient de la météorologie. Elle permet notamment de répondre à un problème lié entre les différences d'échelle des modèles météorologiques mondiaux et régionaux qui les composent (Liu et al. 2012). La non linéarité des paramétrisations physiques et des formulations des dynamiques des deux types de modèles, fait que l'imbrication de ces derniers peut impliquer des variations chaotiques uniquement liées à ce changement d'échelle (Alexandru et al. 2009). Le principe est d'ajouter un terme de relaxation dans les équations d'un modèle basé sur la différence entre l'état du modèle et une ou plusieurs grilles de données de référence, pour chaque point de grille et chaque pas de temps (Stauffer and Seaman 1990). Il est possible d'appliquer cette technique sur une variété de variables climatiques : la SST (Servonnat et al. 2015, Ortega et al. 2017, Estella-Perez et al. 2020), la salinité de surface de l'océan (*Sea Surface Salinity*, SSS) (Servonnat et al. 2015, Estella-Perez et al. 2020), les vents (Zhang et al. 2014), ou encore la température atmosphérique de surface (Zhang et al. 2014). Typiquement, en assimilant par une méthode de rappel un nombre réduit de variables, on obtient l'évolution des variables non-assimilées selon la physique des modèles et les variables assimilées.

En climatologie, on utilise la méthode de rappel dans différents buts. Une des principales utilisations est de produire des réanalyses de la période historique, selon différentes variables météorologiques. En particulier, les mesures instrumentales de température et de salinité de surface et de sous-surface sont utilisées pour produire des reconstructions de la circulation océanique sur la période récente (Swingedouw et al. 2013, Servonnat et al. 2015, Ortega et al. 2017, Estella-Perez et al. 2020). On utilise également le rappel de données historiques pour initialiser un modèle numérique dans un but de projections (Mignot et al. 2016).

Il existe d'autres méthodes, plus complexes et plus coûteuses, mais reconnues pour être plus efficaces qu'une approche par méthode de rappel. On peut par exemple

citer les approches variationnelles (*Variational Approaches*, VA), les filtres de Kalman (*Kalman Filters*, KF), les filtres de Kalman d'ensemble (*Ensemble Kalman Filters*, EKF) ou les filtres à particules (*Particle Filters*, PF). Ces méthodes sont également abondamment utilisées pour assimiler les données satellitaires en météorologie (Houtekamer and Zhang 2016), pour laquelle il existe des données d'observation nombreuses et à très haute résolution spatio-temporelle.

Pour appliquer de telles méthodes en paléoclimatologie, il faut donc disposer de la reconstruction dans l'espace et dans le temps d'au moins une variable à assimiler. Ainsi, si les reconstructions de champs de variables climatiques s'améliorent à mesure que de nouvelles données sont disponibles (section 1.3.4), il en va de même pour l'application des méthodes d'assimilation de données dans un contexte de paléoclimatologie. En effet, pour qu'une méthode d'assimilation tende à reproduire la réalité, il faut évidemment disposer d'observations (ou de reconstructions) qui en soient de bonnes estimations. Ainsi, des reconstructions millénaires évaluées comme robustes, pourraient offrir, une fois assimilées au sein d'un modèle de circulation globale, une vision d'ensemble des variations climatiques sur cette période.

1.4 Problématiques et objectifs de ces travaux de thèse

1.4.1 Problématiques soulevées

- **Comment automatiser la reconstruction d'indices climatiques avec une approche objective?** Il existe de nombreuses reconstructions dans la littérature scientifique. Cependant, la trop grande variété des approches utilisées fait que ces dernières sont souvent contradictoires et peu comparables en matières de métriques mathématiques. Pour répondre à cette problématique, cette thèse a pour but d'introduire des méthodes de *machine learning* à la reconstruction paléoclimatique. L'objectif sera de proposer une méthodologie qui se veut transparente et qui permet d'évaluer des reconstructions faites avec des approches bien distinctes. On s'intéressera pour cela à la période du dernier millénaire, nous permettant ainsi de se baser sur un grand nombre d'enregistrements (section 1.3.3). On cherchera en particulier, à produire la reconstruction d'indices de modes de variabilité tels que la NAO et l'AMV, mais aussi de grilles de températures de surface. On estimera ainsi, en se basant sur des reconstructions paléoclimatiques précédentes, dans quelle mesure l'apport de nouvelles méthodes peut permettre d'obtenir des reconstructions significativement plus robustes. Une fois cette question étayée, on pourra alors se pencher sur des problématiques associées aux variations climatiques du dernier millénaire.
- **Quel rôle pour la variabilité interne et les forçages climatiques naturels dans la mise en place du MCA et LIA, ainsi que pour la transition entre les deux périodes?** En reconstruisant des indices forcés et non forcés de la SST de l'Atlantique Nord, on s'intéressera à la possible implication de cet océan pour la mise en place des séquences du LIA et du MCA. Pour cela on identifiera la réponse climatique de ces modes lors des grands minima solaires (Lean 2018), et lors d'éruptions volcaniques stratosphériques (Sigl et al. 2015). On cherchera alors à bien identifier les périodes sur lesquelles s'étendent le MCA et le LIA ainsi que les différents processus qui en sont à l'origine.

- **Peut-on estimer les variations passées de l'AMOC à l'aide de reconstruction d'indices basés sur de la SST?** Puisque l'AMOC fait partie des nombreux facteurs pouvant jouer sur les variations de la SST, on se demandera si des reconstructions d'indices de l'AMV peuvent permettre d'estimer des variations passées de l'AMOC. Dans un deuxième temps, on cherchera à extraire d'une grille de reconstruction de SST, les indices les plus fiables possibles pour estimer les variations passées de l'AMOC.
- **Y a-t-il un effet de l'océan Atlantique Nord sur les variations du Pacifique? Quelles implications pour la période de hiatus observée entre 1998 et 2012?** Bien que l'effet de l'AMV sur l'IPV sur une échelle de 10 ans est significatif sur la période historique (Marini and Frankignoul 2013), peut-on réellement à partir de 150 ans de mesures estimer que ceci a toujours été vrai? A partir d'une reconstruction de grille de SST nous étudierons la co-évolution de ces deux modes au cours du dernier millénaire, et nous déterminerons quel type de configuration climatique en matière d'état moyen, permet à l'un de ces derniers d'influer ou non sur les variations de l'autre.
- **Que dire des résultats de modèles en terme de variabilité décennale et multi-décennale?** On cherchera à partir de nos reconstructions, à valider ou réfuter des résultats obtenus par des simulations de modèles d'autres études, notamment en matière de propriétés spectrales (fréquence, périodicité) des modes, mais aussi des relations qu'ils ont les uns avec les autres. Ceci se veut alors être un tremplin pour des études futures pour évaluer quels modèles sont les plus enclins à simuler convenablement la variabilité climatique interne et naturelle (Hall et al. 2019), ce qui est essentiel pour l'amélioration des prévisions climatiques, aux échelles régionales, continentales et globales (Hawkins and Sutton 2009, Deser et al. 2012, Cassou et al. 2018, Deser et al. 2020).
- **Une assimilation de données par rappel appliquée à une reconstruction robuste de SST permet-elle de faire aussi bien que des approches par KF?** Nous utiliserons une grille de SST produite à partir des nouvelles méthodes statistiques introduites. On cherchera alors à lui appliquer une méthode de rappel au sein d'un modèle de circulation globale afin d'obtenir une vue d'ensemble de la variabilité climatique du dernier millénaire. On comparera alors la reconstruction obtenue avec une autre, obtenue par KF (Hakim et al. 2016). En plus de ne pas utiliser la même méthode d'assimilation, cette reconstruction diffère d'autant plus de la nôtre, puisqu'elle utilise un autre modèle de circulation globale (CCSM4), et une autre grille de reconstruction de température (Hakim et al. 2016). On s'intéressera notamment à déterminer si les deux méthodes parviennent à bien reproduire localement les variations données par les enregistrements paléoclimatiques de la base de données PAGES 2k (PAGES 2k Consortium 2017).

1.4.2 Organisation de la thèse

Pour tenter de répondre aux différentes problématiques soulevées ci-avant, nous proposons d'organiser les travaux de cette thèse comme suit. Le chapitre 2 expose l'ensemble des méthodes et des approches ayant permis de produire les reconstructions présentées dans cette thèse. La reconstruction ne donnant qu'une base pour

potentiellement obtenir des résultats scientifiques, les méthodes permettant d'analyser les reconstructions obtenues seront également présentées. Le chapitre 3 fera connaître un outil d'optimisation et de comparaison de reconstructions, développé au cours de cette thèse. Alors que cet outil utilise des concepts peu abordables pour les non-experts, il présente l'avantage d'être facile d'utilisation et en accès libre sur internet. Une application proposée à titre d'exemple à la reconstruction d'un indice NAO dans ce même chapitre. Le chapitre 4 reprendra l'outil présenté dans le chapitre 3 pour produire une reconstruction robuste de l'AMV en tant que mode de variabilité interne de la SST de l'Atlantique Nord. Une nouvelle interprétation du rôle qu'ont pu jouer la circulation océanique et les forçages externes sur la mise en place du MCA et du LIA sera alors proposée. De plus, un test de point de bascule sera présenté, dans le but de savoir si l'Atlantique Nord pourrait à l'avenir connaître un refroidissement abrupt, notamment associé à un hypothétiques ralentissement de l'AMOC (Rahmstorf et al. 2015, Sgubin et al. 2017, Caesar et al. 2018). On constatera après comparaison des méthodes de *machine learning* dans les chapitres 3 et 4, que l'application d'approches non linéaires telle que la méthode de forêt aléatoire (section 2.4), surclasse les méthodes linéaires usuelles, en particulier la très fréquemment utilisée régression en composantes principales (Luterdacher et al. 2001, Cook et al. 2002, Ortega et al. 2015, Wang et al. 2017, Cook et al. 2019, section 2.4). Ainsi, la méthode de forêt aléatoire sera étendue à la reconstruction millénaire d'une grille quasi-intégrale de SST de résolution $1^\circ \times 1^\circ$. On explorera alors une grande variété de phénomènes d'un point de vue dynamique, tels que l'effet des forçages sur la SST ou la relation entre l'Atlantique Nord et le Pacifique. Il sera également extrait à partir d'une méthode d'analyse en composantes principales, une série temporelle dont la covariance avec le reste de la SST semble indiquer qu'il s'agirait d'une très bonne estimation de l'AMOC au cours du dernier millénaire. Nous explorerons en particulier les variations récentes de cet indice AMOC au vu de celles reconstruites, afin de valider ou réfuter l'hypothèse d'un ralentissement sans précédent proposée par Rahmstorf et al. (2015), Thornalley et al. (2018) et Caesar et al. (2018). Enfin, chapitre 6 est une tentative d'appliquer une méthode d'assimilation de données simple (méthode de rappel) dans le but comparer la réanalyse obtenue avec celle obtenue à l'aide d'une méthode alternative : le KF (Hakim et al. 2016). Cette méthode, plus complexe que la méthode de rappel, étant très coûteuse, le but sera ici de comparer la fiabilité de méthode à la complexité algorithmique moindre par rapport au KF, telle que la méthode de rappel. Pour finir, le chapitre 7 présente un résumé des travaux présentés auxquels nous apporterons certaines critiques objectives, puis nous décrirons les perspectives amenées par les différents résultats obtenus.

Chapitre 2

Description des méthodes et outils utiles à la compréhension des travaux

Chacun des chapitres 3 à 6 constitue une étude scientifique distincte. Ces études utiliseront des concepts communs mais se baseront également chacune sur des concepts spécifiques dont certains seront explicités. Ce chapitre décrit l'ensemble des concepts et méthodes utilisés pour les travaux présentés dans les chapitres 3 à 6 dès lors qu'ils n'y sont pas déjà explicités.

2.1 Analyse de séries temporelles

On appelle série temporelle, ou série chronologique, une famille de variables aléatoires $(X_t)_{t \in T}$ définie dans l'espace de temps T , indexée par t qui représente le temps. Selon les cas, t peut représenter n'importe quel pas de temps, tel que la seconde, la minute, le mois ou l'année. On utilise le terme de série temporelle pour les signaux climatiques et paléoclimatiques puisqu'il s'agit de mesures météorologiques discrétisées dans le temps et l'espace. Ces séries sont en réalité des estimations de l'évolution de processus continus dans l'espace-temps. On peut caractériser ces signaux par une fonction f du temps $t \in T$, avec $f : T \rightarrow \mathbb{R}$.

Puisque les mesures climatiques et paléoclimatiques, ainsi que les sorties de modèles, sont des séries finies, on considèrera dans cette thèse que $T \subset \mathbb{Z}$, $\#(T) < \infty$, où $\#(\bullet)$ est l'opérateur cardinal. C'est à dire que toute série $(X_t)_{t \in T}$ est définie dans un ensemble fini et dénombrable. Puisque T est fini et dénombrable, on considèrera dans cette section que $n = \#(T)$.

2.1.1 Moyenne, variance, covariance et corrélation empiriques

On note $\bar{X}_n$, la moyenne empirique de la série temporelle $X = (X_t)_{t \in T}$ la fonction de $\mathbb{R}^n \rightarrow \mathbb{R}$ dont la valeur est donnée par :

$$\bar{X}_n = \frac{1}{n} \sum_{t \in T} X_t \quad (2.1)$$

La variance empirique et l'écart type empirique de la série $X = (X_t)_{t \in T}$, respectivement notés s_n^2 et s_n , fonctions de $\mathbb{R}^n \rightarrow \mathbb{R}_+$, sont donnés par :

$$s_n^2 = \frac{1}{n} \sum_{t \in T} (X_t - \bar{X}_n)^2 \quad (2.2)$$

$$s_n = \sqrt{s_n^2} \quad (2.3)$$

On introduit une seconde série temporelle $Y = (Y_t)_{t \in T}$ définie sur le même espace de temps T que $X = (X_t)_{t \in T}$. On appelle covariance empirique de X et Y la fonction de $\mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$, qui est donnée par :

$$\hat{Cov}_n(X, Y) = \frac{1}{n} \sum_{t \in T} (X_t - \bar{X}_n)(Y_t - \bar{Y}_n) \quad (2.4)$$

Finalement, on appelle coefficient de corrélation linéaire empirique de $X = (X_t)_{t \in T}$ et $Y = (Y_t)_{t \in T}$, la fonction de $\mathbb{R}^n \times \mathbb{R}^n \rightarrow [-1, 1]$, donnée par :

$$\hat{Cor}_n(X, Y) = \frac{\hat{Cov}_n(X, Y)}{s_n^{(X)} s_n^{(Y)}} \quad (2.5)$$

Où $s_n^{(X)}$ et $s_n^{(Y)}$ sont les écarts-type empiriques respectifs de X et Y

2.1.2 Stationnarité, autocovariance, autocorrélation et corrélation croisée

On dit que la série temporelle $X = (X_t)_{t \in T}$ est stationnaire, si et seulement si :

$$\left\{ \begin{array}{l} \bullet \mathbb{E}[X] = \mu \in \mathbb{R}, \forall t \in T \\ \bullet X_t \text{ est de carré intégrable, i.e. } X_t^2 < \infty, \forall t \in T \\ \bullet Cov(X_s, X_{s+t}) = Cov(X_{s-1}, X_{s-1+t}) = \dots = Cov(X_0, X_t), \forall t \in T, s \in T_+, \\ \text{tels que } t + s \in T \end{array} \right. \quad (2.6)$$

On appelle fonction d'autocovariance de la série temporelle stationnaire $(X_t)_{t \in T}$, la fonction γ de $\mathbb{Z} \rightarrow \mathbb{R}$, définie $\forall t \in T \forall h \in \mathbb{Z}$, tels que $t + h \in T$, par :

$$\gamma(h) = Cov(X_t, X_{t+h}) = Cov(X_0, X_h), \text{ par stationnarité} \quad (2.7)$$

On appelle ainsi la fonction d'autocorrélation ρ , la fonction de $\mathbb{Z} \rightarrow [-1, 1]$ définie $\forall t \in T, \forall h \in \mathbb{Z}$, tels que $t + h \in T$, par :

$$\rho(h) = \frac{Cov(X_t, X_{t+h})}{Var(X_t)} \quad (2.8)$$

$$\rho(h) = \frac{Cov(X_t, X_{t+h})}{Cov(X_t, X_t)} \quad (2.9)$$

$$\rho(h) = \frac{Cov(X_0, X_h)}{Cov(X_0, X_0)}, \text{ par stationnarité} \quad (2.10)$$

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)}, \text{ d'après (2.7)} \quad (2.11)$$

L'étude de l'évolution de la corrélation sur T de $(X_t)_{t \in T}$ et $(Y_t)_{t \in T}$ est donnée par la fonction de corrélation croisée R_{XY} définie $\forall h \in \mathbb{Z}$ tel que $t + h \in T, \forall t \in T$

$$R_{XY}(h) = \frac{\text{Cov}(X_t, Y_{t+h})}{\sqrt{\text{Var}(X)}\sqrt{\text{Var}(Y)}} \quad (2.12)$$

2.1.3 Analyses spectrales

On considère $T = \llbracket 0, \dots, n \rrbracket$, un espace de temps de taille n , où $\llbracket \bullet \rrbracket$ est un intervalle de nombres entiers. Pour tout série temporelle $X = (X_t)_{t \in T}$, on peut calculer son spectre de puissance sur T obtenue à partir de sa transformée de Fourier discrète. La transformée de Fourier discrète de X pour la fréquence f est donnée par :

$$\hat{s}(f) = \frac{1}{\sqrt{n}} \sum_{t=0}^n X_t e^{-i \frac{2\pi f t}{n}} \quad (2.13)$$

On calcule alors la transformée de Fourier pour chaque fréquence $f_i = \frac{i}{n}, \forall i \in \llbracket 1, \lfloor \frac{n}{2} \rfloor \rrbracket$, correspondant aux périodes $p_i = \frac{n}{i}, \forall i \in \llbracket 1, \lfloor \frac{n}{2} \rfloor \rrbracket$, où $\lfloor \bullet \rfloor$ est l'opérateur de partie entière inférieure. Le spectre de puissance de X est donné, $\forall f \in \{f_1, \dots, f_{\lfloor \frac{n}{2} \rfloor}\}$, par :

$$S(f) = \|\hat{s}(f)\|^2 \quad (2.14)$$

Où $\|\bullet\|$ est ici la norme complexe. L'analyse par spectre de puissance connaît des limites dans le cadre géophysique, car les signaux étudiés sont parfois non-stationnaires sur une période donnée.

Pour analyser l'évolution des propriétés spectrales d'une série temporelle $(X_t)_{t \in T}$ dans le temps et selon différentes fréquences, on utilise les transformées en ondelettes. Il s'agit de donner une représentation du signal permettant de faire apparaître simultanément des informations temporelles (localisation dans le temps, durée) et fréquentielles, pouvant ainsi aider à la caractérisation physique de la source du signal. L'analyse en ondelettes se distingue de la transformée de Fourier car elle permet d'analyser des processus non-stationnaires. Les propriétés de la transformée en ondelettes est ainsi particulièrement efficace dans l'analyse des signaux géophysiques (Grinsted et al. 2004). On calcule dans un premier temps un jeu de fonctions $\phi_{\tau s}$, différentes dérivées de translation et de rotation d'une ondelette mère Ψ :

$$\phi_{\tau s}(t) = \frac{1}{\sqrt{s}} \Psi\left(\frac{t - \tau}{s}\right) \quad (2.15)$$

Où $s \in \mathbb{R}_+$ et $\tau \in \mathbb{R}$. s et τ sont respectivement appelés les paramètres d'expansion et de translation. Les fonctions $\phi_{\tau s}$ portent le nom d'ondelettes. Pour construire une famille d'ondelettes, plusieurs valeurs pour τ et s sont choisies arbitrairement. L'ondelette mère, Ψ , doit être à support compact et de moyenne nulle, on parle de conditions d'admissibilité :

$$\int |\Psi(t)|^2 dt < \infty \quad (2.16)$$

$$\int \Psi(t) dt = 0 \quad (2.17)$$

La convolution de $f(t)$ avec cet ensemble d'ondelettes est la transformée continue en ondelettes de f :

$$W(\tau, s) = \int f(t) \bar{\Psi}_{\tau, s}(t) dt \quad (2.18)$$

Où $\bar{\Psi}$ est le complexe conjugué de Ψ .

Une des ondelettes mères remplissant les conditions d'admissibilité (2.16) et (2.17) la plus fréquemment utilisée dans l'études de signaux géophysiques est l'ondelette de Morlet. Elle est donnée par :

$$\Psi(t) = \pi^{-\frac{1}{4}} e^{-i\pi t} e^{-\frac{t^2}{2}} \quad (2.19)$$

2.2 Méthodes d'analyses statistiques pour les données climatiques

2.2.1 Analyse en composantes principales et fonctions orthogonales empiriques

Les variations du climat résultent d'un nombre considérable d'interactions complexes. Ainsi les variables climatiques observées ou de sortie de modèles sont relativement volumineuses selon leur résolution spatiale et temporelle. Le but d'une analyse en composantes principales (*i.e.* les composantes principales, *Principal Component analysis*, PCA), proposée par Hotelling (1933), est de résumer une grande part de la variabilité initiale d'un jeu de données avec de nombreuses variables, en un nombre très restreint de variables (*Principal Components*, PC). Preisendorfer (1988) propose une application de la PCA aux données météorologiques, qu'il appelle également analyse en fonctions orthogonales empiriques (*Empirical Orthogonal functions*, EOF). Le terme d'EOF renvoie aux vecteurs propres de la base orthogonale construite au cours de l'analyse.

On considère la grille de surface d'une variable météorologique quelconque et discrétisée, c'est à dire que celle-ci se décompose dans un champ spatial fini et dénombrable à 2 dimensions et un espace temporel fini et dénombrable à une dimension. On appelle F le champ de mesure associé à cette grille. Il s'agit d'une fonction du temps t , de la latitude θ et de la longitude ϕ . F est donc définie sur les pas de temps t_i , $i = 1, \dots, n$, pour les points de grilles aux latitudes θ_l , $l = 1, \dots, p_1$ et aux longitudes ϕ_k , $k = 1, \dots, p_2$. Les valeurs de $F_{i,j,k}$ sont données par :

$$F_{i,j,k} = F(t_i, \theta_l, \phi_k) \quad (2.20)$$

Pour appliquer une PCA sur un champs de valeurs de surface, il faut la transformer en une matrice à deux dimensions $X(t, s)$ (Preisendorfer 1988), où s est la position spatiale dans la grille latitude/longitude. On notera s_j , $j = 1, \dots, p$, $p = p_1 p_2$, les paires θ_l, ϕ_k , $\forall l = 1, \dots, p_1$, $\forall k = 1, \dots, p_2$.

On a donc :

$$X = \begin{pmatrix} x_{1,1} & \cdots & x_{1,p} \\ \vdots & \ddots & \vdots \\ x_{n,1} & \cdots & x_{n,p} \end{pmatrix} \quad (2.21)$$

On peut donc écrire :

$$X = (X^j)_{1 \leq j \leq p} = (X_i)_{1 \leq i \leq n}^T \quad (2.22)$$

Où $\bullet^T$ est l'opérateur de transposée d'une matrice.

Ainsi, si l'on note :

$$\bar{x}_{.j} = \frac{1}{n} \sum_{i=1}^n x_{ij} \quad (2.23)$$

On peut définir la moyenne de la variable climatique X sur l'ensemble de la période $T = (t_i)_{1 \leq i \leq n}$ par :

$$\bar{x} = (\bar{x}_{.j})_{1 \leq j \leq p} \quad (2.24)$$

Le champs des anomalies est alors donné par :

$$X^{(a)} = X - \mathbb{1}_n \bar{x} = (I - \frac{1}{n} \mathbb{1}_n \mathbb{1}_n^T) X \quad (2.25)$$

Où $\mathbb{1}_n = (1, \dots, 1)^T$ et $\#(\mathbb{1}_n) = n$.

Ainsi, $\forall 1 \leq t \leq n$, $X_t^{(a)}$ est une série de la répartition spatiale des anomalies au pas de temps t , et $X^{(a)j}$ est la série temporelle qui décrit les anomalies du point de grille j sur T .

Il s'agit ici de rapporter la matrice $X^{(a)}$ à une base orthogonale de plus petite dimension, qui explique le plus possible de variabilité initiale de $X^{(a)}$ dans ses premiers axes. Pour cela, on calcule la matrice de covariance des points de grilles, notée S :

$$S = \frac{1}{n} X^{(a)T} X^{(a)} \quad (2.26)$$

On cherche à trouver les vecteurs $v_1, \dots, v_p$, formant une base orthogonale de $\mathbb{R}^p$, qui sont tels que la variance globale du jeu de données est concentrée dans les premiers vecteurs de cette base.

On procède à la diagonalisation de S . On peut alors facilement trouver le premier vecteur propre v_1 car on ne dispose pas encore de contrainte d'orthogonalité. On maximise la variance de $X^{(a)}v_1$, qui correspond à la représentation des données de $X^{(a)}$ sur le premier axe de la nouvelle base :

$$\text{Var}(X^{(a)}v_1) = \frac{1}{n} \|X^{(a)}v_1\|^2 = \frac{1}{n} (X^{(a)}v_1)^T (X^{(a)}v_1) = v_1^T S v_1 \quad (2.27)$$

Ceci revient à résoudre :

$$v_1 = \arg \max_{\substack{v^T v = 1 \\ v \in \mathbb{R}^p}} (v^T S v) \quad (2.28)$$

Le vecteur v_1 est le premier vecteur propre de S et est tel que $\text{Var}(X^{(a)}v_1) = v_1^T S v_1 = \lambda_1$. On appelle λ_1 la première valeur propre, associée au premier vecteur propre de S .

On trouve $v_2 \in \mathbb{R}^p$ de la même manière en rajoutant la contrainte d'orthogonalité avec v_1 :

$$v_2 = \arg \max_{\substack{v^T v = 1 \\ v \in \mathbb{R}^p \\ v^T v_1 = 0}} (v^T S v) \quad (2.29)$$

En répétant l'opération jusqu'à p en ajoutant une contrainte d'orthogonalité à chaque étape, on obtient les p vecteurs propres de S ainsi que les p valeurs propres associées : $\lambda_k = \text{Var}(Xv_k) = v_k^T S v_k, \forall 1 \leq k \leq p$, qui sont telles que $\lambda_1 < \lambda_2 < \dots < \lambda_p$, et qui vérifient :

$$\sum_{j=1}^p \lambda_j = \sum_{j=1}^p \text{Var}(X^j) \quad (2.30)$$

On peut ainsi associer des poids aux vecteurs propres à partir des valeurs propres de S . Ces poids, $\omega_1, \dots, \omega_p$, représentent la part de variance initiale expliquée par chacun des vecteurs propres de S :

$$\omega_j = \frac{\lambda_j}{\sum_{k=1}^p \lambda_k} \quad (2.31)$$

Au final, on a :

$$SV = V\Lambda \quad (2.32)$$

Où $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_p) \in \mathbb{R}^{p \times p}$.

Les PC $u_1, \dots, u_p$ correspondent aux coordonnées de X dans la nouvelle base définie par les colonnes de V :

$$U = XV \quad (2.33)$$

On appelle U , la matrice des PC. La matrice V est la matrice des EOF. Les colonnes de U traduisent des parts indépendantes de la variabilité temporelle initiale de $X^{(a)}$. Les colonnes de V traduisent des parts indépendantes de la variance de $X^{(a)}$.

2.2.2 Régression linéaire

Le but d'une régression linéaire est d'exprimer par l'équation d'une droite une variable à expliquer Y comme la combinaison linéaire d'un ensemble de p variables explicatives $X^1, \dots, X^p$. Le modèle de régression linéaire est le suivant.

$$Y = X\beta + \epsilon \quad (2.34)$$

Où $\beta = \begin{pmatrix} \beta_0 \\ \cdots \\ \beta_p \end{pmatrix} \in \mathbb{R}^{p+1}$. Puisque le modèle linéaire connaît un coefficient de pente (β_0 , la matrice $X \in \mathbb{R}^{n \times (p+1)}$ s'écrit $X = (\mathbb{1}_n, X^1, \dots, X^p)$. $\epsilon \sim \mathcal{N}(0_n, I_n \sigma^2)$ est un vecteur de bruits blancs gaussiens de même variance σ^2 , où 0_n est le vecteur nul de taille n et $I_n = \text{diag}(n)$ la matrice identité d'ordre n .

Bien que la régression linéaire est une approche de modélisation simple à appliquer et efficace, celle-ci repose sur des hypothèses fortes nécessaires au calcul des coefficients de régression. Ces hypothèses sont les suivantes :

- Les variables explicatives $X^1, \dots, X^p$ sont non colinéaires c'est à dire qu'aucune d'entre elles ne peut s'écrire comme une combinaison linéaire des $(p-1)$ autres. En effet, l'estimateur de β fait appel à l'inversion de la matrice $X^T X$. Or, cette dernière n'est pas inversible si certains de ses vecteurs sont colinéaires, ce qui autrement dit signifie que S n'est pas diagonalisable. Notons que dans le cas d'une régression linéaire simple, c'est à dire à une seule variable explicative, on a $X = (\mathbb{1}_n, X^1)$. L'unique cas dans lequel le modèle de régression n'est pas applicable selon cette hypothèse est donc le cas où X^1 est un vecteur constant, ce qui n'a dans la pratique aucun intérêt.
- Les résidus du modèle $\epsilon_1, \dots, \epsilon_p$ vérifient les hypothèses de Gauss-Markov. C'est-à-dire qu'elles doivent être distribuées selon une loi gaussienne, leurs espérances sont nulles et leurs variances sont égales à σ^2 (homoscédasticité des résidus). De plus, elles doivent être telles que $\text{Cov}(\epsilon_i, \epsilon_j) = 0, \forall i \neq j$.
- Les variables explicatives sont exogènes et non aléatoires, ce qui de par l'homoscédasticité des résidus revient à dire que $\forall 1 \leq i \leq n, \forall 1 \leq j \leq p, \epsilon_i | X_i^j \sim \mathcal{N}(0, \sigma^2)$.

L'estimation de β est donnée par l'estimateur des moindres carrés, à savoir :

$$\hat{\beta} = (X^T X)^{-1} X^T Y \quad (2.35)$$

Ce qui, dans le cas d'un modèle de régression linéaire simple revient à :

$$\hat{\beta}_1 = \frac{\text{Cov}(X, Y)}{\text{Var}(Y)} \quad (2.36)$$

$$\hat{\beta}_0 = \bar{Y}_n - \hat{\beta}_1 \bar{X}_n \quad (2.37)$$

La droite du modèle de régression est donc donnée par l'équation :

$$\hat{Y} = X \hat{\beta} \quad (2.38)$$

La qualité du modèle de régression linéaire, c'est à dire la propension de la droite estimée à se rapprocher des vrais valeurs de Y , est donnée par la statistique du R^2 . Il s'agit d'un rapport de la somme des carrés des écarts à la moyenne des valeurs prédites par la droite $\hat{Y}$ et la somme des carrés des écarts à la moyenne des vraies valeurs Y . Le R^2 de régression, définie dans $[-1, 1]$, est donné par :

$$R^2 = \frac{\sum_{i=1}^n (\hat{Y}_i - \bar{Y}_n)^2}{\sum_{i=1}^n (Y_i - \bar{Y}_n)^2} \quad (2.39)$$

Dans cette thèse, le modèle linéaire dans sa version multiple servira pour l'application de certaines méthodes de *machine learning* (sections 2.4, 3, 4, 5). On utilisera le modèle linéaire dans sa version simple pour produire des cartes de régression point de grille par point de grille, entre le champ d'une variable météorologique donnée (SST, SLP...), et un indice climatique donné. Le lien entre l'indice et la variable météorologique s'exprime alors comme l'estimation du coefficient de régression de pente $\hat{\beta}_1$. Ceci permet alors de renseigner sur le signe et l'intensité du lien entre un indice climatique et une variable météorologique en un point de grille donné. Ce lien spatial entre un indice et un champ de données peut également s'exprimer de manière équivalente en matière de signe par des cartes de corrélation. Dans les deux cas, le calcul de la significativité de la statistique sera détaillé dans la section suivante qui lui est dédiée.

2.3 Calculs de significativité statistique

Les tests statistiques permettent d'estimer la significativité d'une statistique calculée à partir d'un échantillon d'observations donné, que l'on notera ici s . Le but est de tester une hypothèse nulle H_0 contre une hypothèse dite alternative H_1 . Dans la pratique, on construit l'hypothèse nulle de sorte que sous celle-ci, le caractère testé d'une statistique donnée n'est pas significatif. On peut alors définir les risques de première et seconde espèces α et β , qui sont donnés par :

$$\alpha = \mathbb{P}[H_1|H_0] \quad (2.40)$$

$$\beta = \mathbb{P}[H_0|H_1] \quad (2.41)$$

Où, $\mathbb{P}[x|y]$ est la probabilité de x conditionnelle à y .

A l'aide des distributions de probabilité de la statistique s sous H_0 , on peut calculer la probabilité d'observer la statistique s alors que H_0 est vraie, on parle de p-valeur :

$$\text{pval} = \mathbb{P}[s|H_0] \quad (2.42)$$

Ainsi on peut dire qu'un test donné, pour un risque de première espèce α , est significatif au seuil $1 - \alpha$, dès lors que $\text{pval} < 1 - \alpha$.

Le plupart des tests présentés ci-après sont des tests de Student. Ceci requiert donc qu'on évoque avant tout la définition de la loi de Student, qui est essentielle pour comprendre les sections 2.3.1 à 2.3.4.

Définition : Soit $Z \sim \mathcal{N}(0, 1)$, une variable aléatoire de loi normale centrée réduite. Soit alors $U \sim \chi_n^2$ une variable aléatoire distribuée selon une loi du χ^2 à n degrés de liberté et indépendante de Z . Alors, on peut dire que la variable aléatoire $\mathcal{T} = \frac{Z}{\sqrt{U/n}}$ suit une loi de Student à n degrés de liberté, i.e. $\mathcal{T} \sim St(n)$.

De même, nous ferons fréquemment appel au concept d'estimateur sans biais de la variance théorique. Introduisons préalablement la notion de biais d'un estimateur.

Définition : Soit $\hat{\theta}$ un estimateur donné d'une valeur théorique θ . On peut alors définir le biais de l'estimateur $\hat{\theta}$, noté $\mathcal{B}(\hat{\theta})$, donné par :

$$\mathcal{B}(\hat{\theta}) = \mathbb{E}[\hat{\theta}] - \theta \quad (2.43)$$

On dit alors que $\hat{\theta}$ est un estimateur sans biais de θ si $\mathcal{B}(\hat{\theta}) = 0$

Pour définir les lois sous-jacentes aux tests statistiques présentés ci-après, nous devons introduire les estimateurs sans biais de l'espérance μ et de la variance σ^2 d'une série donnée $X = (X_t)_{t \in T}$. Les estimateurs empiriques naturels (que l'on appelle estimateurs du maximum de vraisemblance) de ces derniers, sont respectivement donnés par :

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i \quad (2.44)$$

$$\begin{aligned} S_n^2 &= \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2 = \frac{1}{n} (X_i^2) - \frac{2}{n} \bar{X}_n \sum_{i=1}^n (X_i) - \bar{X}_n^2 \\ &= \frac{1}{n} \sum_{i=1}^n (X_i^2) - \left(\frac{1}{n} \sum_{i=1}^n X_i \right)^2 \end{aligned} \quad (2.45)$$

En utilisant les propriétés de la linéarité de l'espérance, d'après (2.43), on peut montrer que $\bar{X}_n$ est un estimateur sans biais de μ :

$$\begin{aligned} \mathcal{B}(\bar{X}_n) &= \mathbb{E}[\bar{X}_n] - \mu \\ &= \mathbb{E}\left[\frac{1}{n} \sum_{i=1}^n X_i\right] - \mu = \frac{1}{n} \sum_{i=1}^n \mathbb{E}[X_i] - \mu \\ &= \frac{n\mu}{n} - \mu = \mu - \mu = 0 \end{aligned} \quad (2.46)$$

Le biais de S_n^2 , quant à lui, est défini par :

$$\mathcal{B}(S_n^2) = \mathbb{E}[S_n^2] - \sigma^2 \quad (2.47)$$

Notons que $\mathbb{E}(S_n^2) = \mathbb{E}\left[\frac{1}{n} \sum_{i=1}^n X_i^2\right] - (\mathbb{E}[\bar{X}_n^2])$ d'après (2.45) et de par linéarité de l'espérance. Le calcul de $\mathbb{E}[\bar{X}_n^2]$ donne :

$$\begin{aligned} \mathbb{E}[\bar{X}_n^2] &= \frac{1}{n^2} \mathbb{E}\left[\left(\sum_{i=1}^n X_i\right)^2\right] \\ &= \frac{1}{n^2} \mathbb{E}\left[\sum_{i=1}^n (X_i^2) + \sum_{i=1}^n \sum_{i \neq j} X_i X_j\right] \\ &= \frac{1}{n^2} (n\mathbb{E}[X^2] + n(n-1)\mathbb{E}[X]^2) \\ &= \frac{1}{n} \mathbb{E}[X^2] + \frac{n-1}{n} (\mathbb{E}[X])^2 \end{aligned} \quad (2.48)$$

On a donc :

$$\begin{aligned}
 \mathbb{E}[S_n^2] &= \frac{1}{n} \mathbb{E}\left[\sum_{i=1}^n X_i^2\right] - \frac{1}{n} \mathbb{E}[X^2] - \frac{n-1}{n} (\mathbb{E}[X])^2 \\
 &= \frac{n-1}{n} \mathbb{E}[X^2] - \frac{n-1}{n} (\mathbb{E}[X])^2 \\
 &= \frac{n-1}{n} (\mathbb{E}[X^2] - (\mathbb{E}[X])^2) \\
 &= \frac{n-1}{n} \sigma^2, \text{ d'après la définition de la variance}
 \end{aligned} \tag{2.49}$$

Et finalement :

$$\mathcal{B}[S_n^2] = \frac{n-1}{n} \sigma^2 - \sigma^2 = -\frac{1}{n} \sigma^2 \tag{2.50}$$

La variance empirique du maximum de vraisemblance, S_n^2 , est donc un estimateur biaisé de σ^2 , dont le biais vaut $-\frac{1}{n} \sigma^2$ d'après (2.50). En revanche, on peut noter que $s_n^2 = \frac{n}{n-1} S_n^2$ est un estimateur sans biais de σ^2 puisque d'après (2.49) :

$$\mathbb{E}\left[\frac{n}{n-1} S_n^2\right] = \sigma^2 \tag{2.51}$$

$$i.e. \mathbb{E}[s_n^2] = \sigma^2 \tag{2.52}$$

Et donc :

$$\mathcal{B}(s_n^2) = \mathbb{E}[s_n^2] - \sigma^2 = \sigma^2 - \sigma^2 = 0 \tag{2.53}$$

Puisque la variance théorique d'un signal géophysique sur une période de temps donnée n'est jamais connue dans la pratique, on estimera toujours σ^2 par son estimateur sans biais s^2 dans la suite de cette thèse. Tout ceci a été introduit car ceci a d'importantes conséquences dans la réalisation des tests présentés ci-dessous, notamment en terme de loi de probabilité sous-jacente sous H_0 .

2.3.1 Test de Student de significativité d'une moyenne

Soit une série temporelle $X = (X_t)_{t \in \mathbb{Z}}$ de moyenne empirique sur T $\bar{X}_n = \frac{1}{n} \sum_{t=1}^n X_t$ et de variance empirique estimée sur T (sans biais) $s_n^2 = \frac{1}{n-1} \sum_{t=1}^n (X_t - \bar{X}_n)^2$. Dans ce contexte, il est alors possible d'appliquer 3 types de test de significativité de Student pour la moyenne : bilatéral et unilatéral à gauche et à droite. Dans cette thèse, on s'intéressera pour ces tests au cas particuliers où la valeur testée est 0. En effet, puisque l'on travaillera en terme d'anomalies (écarts à la moyenne sur une période), on cherchera à savoir si une série temporelle d'anomalies est significativement positive ou négative. On appelle tests de Student l'ensemble des tests pour lesquels la statistique calculée suit une loi de Student. On notera que pour $n > 30$, ce qui, dans cette thèse, sera toujours le cas pour les tests de cette section et des sections 2.3.2 et 2.3.4, la loi de Student converge vers une loi gaussienne centrée réduite, ce qui simplifiera les procédures de test présentées dans cette sous-section et dans les sections 2.3.2 et 2.3.4.

Le premier type de test est le test bilatéral de $\bar{X}_n$, qui consiste à savoir si la moyenne de la série X est significativement différente de 0, quelque soit le signe. La procédure de test s'écrit alors :

$$\begin{cases} H_0 : \mu = 0 \\ H_1 : \mu \neq 0 \end{cases} \quad (2.54)$$

Où $\mu = \mathbb{E}[X]$ est l'espérance (ou moyenne théorique) de la série X . On note également σ^2 la variance théorique de la série X .

On a alors :

$$\mathbb{E}[\bar{X}_n] = \frac{1}{n} \sum_{i=1}^n \mathbb{E}[X_i] = \mu \quad (2.55)$$

$$\text{Var}(\bar{X}_n) = \frac{1}{n^2} \sum_{i=1}^n \text{Var}(X_i) = \frac{\sigma^2}{n} \quad (2.56)$$

Ainsi, puisque $\frac{\sigma_n^2}{n} \xrightarrow[n \rightarrow +\infty]{} 0$, d'après la loi forte des grands nombres, $\bar{X}_n$ est convergente. Comme on considère $n > 30$, $\bar{X}_n$ étant une somme de variables aléatoires, alors le théorème central limite (TCL) ainsi que (2.55) et (2.56), nous disent que :

$$Z = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \sim \mathcal{N}(0,1) \quad (2.57)$$

Précisons que Z suit une loi gaussienne lorsque σ^2 est connue (très rare dans la pratique). Dès lors que σ^2 est inconnue, son estimateur s^2 est une variable aléatoire et on ne peut pas dire que la loi de Z est une gaussienne. On peut noter que $\mathcal{T} = \frac{Z}{\sqrt{U/(n-1)}}$, avec $U = \frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \bar{X}_n)^2$, donne l'expression de Z où la variance théorique est substituée par la variance empirique :

$$\mathcal{T} = \frac{\sqrt{n}(\bar{X}_n - \mu)}{s} \quad (2.58)$$

D'après le théorème de Cochran, on a alors que $U \sim \chi_{n-1}^2$, où χ_{n-1}^2 est la loi du χ^2 à $n - 1$ degrés de liberté. En tant que quotient d'une loi gaussienne centrée réduite et d'une loi de χ_{n-1}^2 , on peut donc dire que $\mathcal{T} \sim St(n - 1)$, où $St(n - 1)$ est la loi de Student à $n - 1$ degrés de liberté.

On note t^{stat} la statistique de test obtenue sur une série temporelle observée, dont la valeur sous l'hypothèse H_0 ($\mu = 0$) est :

$$t^{stat} = \frac{\sqrt{n}\bar{X}_n}{s_n} \quad (2.59)$$

Si H_0 est vraie, alors t^{stat} est une réalisation de $\mathcal{T} \sim St(n - 1)$. Puisque $n > 30$, ceci revient, par approximation normale, à considérer que t^{stat} est une réalisation de $\mathcal{T} \sim \mathcal{N}(0,1)$. H_0 est rejetée si t^{stat} est trop improbable d'être obtenue sous l'hypothèse que cette dernière est vraie. La zone de rejet de H_0 au risque $\alpha \in]0, 1[$ W_α , telle

que $\mathbb{P}_{H_0}[W_\alpha] = \alpha$, peut donc s'écrire :

$$W_\alpha = \{|t^{stat}| > u_{1-\frac{\alpha}{2}}\} \quad (2.60)$$

Où $u_{1-\frac{\alpha}{2}}$ est le quantile d'ordre $1 - \frac{\alpha}{2}$ de la loi normale centrée réduite. Le quantile d'une loi pour une valeur donnée s'obtient par l'inverse de la fonction de répartition de la loi normale. Les logiciels de statistiques permettent de calculer ces quantiles, qui sont estimés par des algorithmes d'approximation.

Pour le cas de tests unilatéraux, on cherche à savoir si la série temporelle X est significativement positive, ou significativement négative. Dans ce cas, on a les hypothèses alternatives respectives $H_1 : \mu > 0$ et $H_1 : \mu < 0$. La statistique de test est la même que pour les tests bilatéraux. En revanche, pour que la région de rejet de H_0 satisfasse $\mathbb{P}_{H_0}[W_\alpha] = \alpha$, on a respectivement :

$$W_\alpha = \{t^{stat} > u_{1-\alpha}\} \quad (2.61)$$

et,

$$W_\alpha = \{t^{stat} < u_\alpha\} \quad (2.62)$$

Pour les trois procédures, H_0 est rejetée si $t^{stat} \in W_\alpha$, c'est-à-dire que la moyenne est considérée comme significativement différente de 0 dans le cas bilatéral, et significativement positive ou négative dans le cas unilatéral, au risque $\alpha \in]0, 1[$. Dans le cas contraire, si $t^{stat} \notin W_\alpha$, H_0 n'est pas rejetée, et la moyenne de la série temporelle observée n'est pas considérée comme significativement différente de 0 dans le cas bilatéral, et n'est pas significativement positive ou négative dans le cas unilatéral, au risque $\alpha \in]0, 1[$.

2.3.2 Test de Student de comparaison de deux moyennes

Ce test est très similaire à celui présenté ci-avant. Comme pour le test de significativité d'une moyenne, on traitera initialement le cas d'un test bilatéral, puis nous en déduirons les procédures des tests unilatéraux. Le but du test bilatéral de comparaisons de moyennes est de déterminer si les moyennes des séries sont significativement différentes. Dans cette thèse, on sera par moment confronté à la comparaison de moyennes de séries définies sur deux périodes différentes. On note donc ici que $X = (X_t)_{t \in T_X}$ est définie sur l'espace de temps T_X et que $Y = (Y_t)_{t \in T_Y}$ est définie sur l'espace de temps T_Y . On note alors les tailles des deux échantillon $n_X = \#(T_X)$ et $n_Y = \#(T_Y)$. Ici on notera que $\mu_X = \mathbb{E}[X]$ et $\mu_Y = \mathbb{E}[Y]$ sont les espérances respectives des deux séries, et que $\bar{X}_n$ et $\bar{Y}_n$ en sont les réalisations empiriques. On note alors σ_X^2 et σ_Y^2 les variances respectives des deux séries, et s_X^2 et s_Y^2 en sont les réalisations.

Les hypothèses peuvent donc se poser comme suit :

$$\begin{cases} H_0 : \mu_X = \mu_Y \\ H_1 : \mu_X \neq \mu_Y \end{cases} \quad (2.63)$$

On remarque que si l'on note $\mu = \mu_X - \mu_Y$, l'espérance de la variable aléatoire $X - Y$, on peut alors noter les hypothèses sous la forme :

$$\begin{cases} H_0 : "\mu = 0" \\ H_1 : "\mu \neq 0" \end{cases} \quad (2.64)$$

Il s'agit donc d'appliquer exactement le même test que précédemment, mais appliqué à la moyenne de $X - Y$. Il suffit donc de déterminer les paramètres de la gaussienne $\bar{X}_n - \bar{Y}_n$, qui permettront d'appliquer le TCL. On a :

$$\mathbb{E}[\bar{X}_n - \bar{Y}_n] = \mathbb{E}[\bar{X}_n] - \mathbb{E}[\bar{Y}_n] = \mu_X - \mu_Y \quad (2.65)$$

$$Var(\bar{X}_n - \bar{Y}_n) = Var(\bar{X}_n) + Var(\bar{Y}_n) = \frac{1}{n_X^2} n_X \sigma_X^2 + \frac{1}{n_Y^2} n_Y \sigma_Y^2 = \frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y} \quad (2.66)$$

Puisque dans cette thèse on sera toujours dans le cas où $n_X > 30$ et $n_Y > 30$, on peut directement remplacer les variances par leurs estimateurs empiriques et utiliser l'approximation normale pour la statistique de test. Le TCL nous dit alors que :

$$Z = \frac{(\bar{X}_n - \bar{Y}_n) - (\mu_X - \mu_Y)}{\sqrt{\frac{s_X^2}{n_X} + \frac{s_Y^2}{n_Y}}} \sim \mathcal{N}(0, 1) \quad (2.67)$$

Or, sous H_0 , on a $\mu_X = \mu_Y$. On peut donc écrire, sous H_0 :

$$Z = \frac{\bar{X}_n - \bar{Y}_n}{\sqrt{\frac{s_X^2}{n_X} + \frac{s_Y^2}{n_Y}}} \sim \mathcal{N}(0, 1) \quad (2.68)$$

Si l'on note t^{stat} la réalisation de Z , on peut définir la zone de rejet W_α telle que :

$$W_\alpha = \{|t^{stat}| > u_{1-\frac{\alpha}{2}}\} \quad (2.69)$$

Où $u_{1-\frac{\alpha}{2}}$ est le quantile d'ordre $1 - \frac{\alpha}{2}$ de la loi normale centrée réduite.

Pour le cas de tests unilatéraux, on cherche à savoir si la série temporelle X est significativement supérieure à Y , ou significativement inférieure. Dans ce cas, on a les hypothèses alternatives respectives $H_1 : "\mu_X > \mu_Y"$ et $H_1 : "\mu_X < \mu_Y"$. La statistique de test est la même que pour les tests bilatéraux. Pour que la région de rejet de H_0 satisfasse $\mathbb{P}_{H_0}[W_\alpha] = \alpha$, on a respectivement :

$$W_\alpha = \{t^{stat} > u_{1-\alpha}\} \quad (2.70)$$

et,

$$W_\alpha = \{t^{stat} < u_\alpha\} \quad (2.71)$$

Pour les trois procédures, H_0 est rejetée si $t^{stat} \in W_\alpha$, c'est-à-dire que les moyennes sont considérées comme significativement différentes au risque $\alpha \in]0, 1[$ dans le cas bilatéral. Pour les tests unilatéraux, on conclut alors que μ_X est significativement supérieur (ou inférieur selon le test unilatéral) à μ_Y , au risque $\alpha \in]0, 1[$. Dans le cas

contraire, si $t^{stat} \notin W_\alpha$, H_0 n'est pas rejetée, et les moyennes des séries temporelles observées ne sont pas considérées comme significativement différentes au risque $\alpha \in]0, 1[$ dans le cas bilatéral. Pour les tests unilatéraux, on conclut alors que μ_X n'est pas significativement supérieur (ou inférieur selon le test unilatéral) à μ_Y , au risque $\alpha \in]0, 1[$.

2.3.3 Test de Student corrigé de significativité d'une corrélation

Les données paléoclimatiques sont définies sur des pas de temps qui peuvent différer et n'ont pas toujours les mêmes résolutions, que ce soit entre elles (Fig. 1.22b), ou par rapport aux données instrumentales. Ceci ne permet donc pas de calculer directement une corrélation. On peut ramener deux séries temporelles à des résolutions similaires avec plus ou moins de rigueur. Une première approche consiste à moyenner les observations de la série à la résolution la plus haute sur les pas de temps de la série à la résolution la plus basse (on parle de filtre). Une seconde approche consiste à interpoler la série à la résolution la plus basse sur les pas de temps de la série à la résolution la plus haute. Dans les deux cas, ceci biaise significativement l'estimation de la corrélation entre la série, notamment en la surestimant, ce qui peut amener à faussement décréter qu'une corrélation entre deux séries est significative (Bretherton et al. 1999). Il est cependant possible de ré-évaluer la statistique de test obtenue à partir de la corrélation, notamment en corrigeant le degré de liberté permettant de la calculer. On se place ici dans le cadre où r est la corrélation empirique calculée entre la série $(X_t)_{t \in T}$ et la série $(Y_t)_{t \in T}$, toutes deux définies sur le même espace de temps T tel que $\#(T) = n$. La procédure présentée ci-après a été proposée par McCarthy et al. (2015), en reprenant les travaux de Bretherton et al. (1999), dans le but d'étudier les relations entre l'AMOC et le niveau marin de différentes régions.

On se place tout d'abord dans le cas d'un test bilatéral. Les hypothèses d'un test de significativité de la corrélation, qu'il soit corrigé ou non, s'écrivent alors comme suit :

$$\begin{cases} H_0 : "r = 0" \\ H_1 : "r \neq 0" \end{cases} \quad (2.72)$$

La statistique classique du test de Student de significativité de la corrélation est donnée par :

$$t^{stat} = \sqrt{n-2} \frac{r}{\sqrt{(1-r^2)}} \quad (2.73)$$

Dans le cadre d'un test de corrélation classique, la statistique de test suit une loi $St(n-2)$.

Pour corriger le degré de liberté, on va se servir des fonctions d'auto-corrélation des deux séries (section 2.1.2), que l'on note respectivement ρ_X et ρ_Y , ce qui va permettre de s'affranchir de tout filtre ou interpolation appliquée à l'une des séries. On note ce dernier n_{eff} , qui est donné par :

$$n_{eff} = \left\lfloor n \times \frac{1 - \rho_X(1)\rho_Y(1)}{1 + \rho_X(1)\rho_Y(1)} \right\rfloor + 1 \quad (2.74)$$

Où $\lfloor \bullet \rfloor$ est l'opérateur de partie entière. La statistique de Student corrigée est donc donnée par :

$$t^{stat} = \sqrt{n_{eff}} \frac{r}{\sqrt{(1-r^2)}} \quad (2.75)$$

On a donc sous H_0 $t \sim St(n_{eff})$. À l'inverse des sections 2.3.1 et 2.3.2, $n > 30$ n'implique pas nécessairement que $n_{eff} > 30$. On peut donc déterminer la zone de rejet de H_0 au risque $\alpha \in]0, 1[$, notée W_α , donnée par :

$$W_\alpha = \begin{cases} \{|t^{stat}| \leq t_{1-\frac{\alpha}{2}, n_{eff}}\}, & \text{si } n_{eff} \leq 30 \\ \{|t^{stat}| \leq u_{1-\frac{\alpha}{2}}\}, & \text{si } n_{eff} > 30 \end{cases} \quad (2.76)$$

Où, $t_{1-\frac{\alpha}{2}, n_{eff}}$ est le quantile d'ordre $1 - \frac{\alpha}{2}$ de la loi de Student à n_{eff} degrés de liberté, et $u_{1-\frac{\alpha}{2}}$ est le quantile d'ordre $1 - \frac{\alpha}{2}$ de la loi normale centrée réduite.

Pour le cas de tests unilatéraux, on cherche à savoir si r est significativement positif ou significativement négatif. Dans ce cas, on a les hypothèses alternatives respectives $H_1 : "r > 0"$ et $H_1 : "r < 0"$. La statistique de test est la même que pour les tests bilatéraux. En revanche, pour que la région de rejet de H_0 satisfasse $\mathbb{P}_{H_0}[W_\alpha] = \alpha$, on a respectivement :

$$W_\alpha = \begin{cases} \{t^{stat} > t_{1-\alpha, n_{eff}}\}, & \text{si } n_{eff} \leq 30 \\ \{t^{stat} > u_{1-\alpha}\}, & \text{si } n_{eff} > 30 \end{cases} \quad (2.77)$$

et,

$$W_\alpha = \begin{cases} \{t^{stat} < t_{\alpha, n_{eff}}\}, & \text{si } n_{eff} \leq 30 \\ \{t^{stat} < u_\alpha\}, & \text{si } n_{eff} > 30 \end{cases} \quad (2.78)$$

Ainsi, dans le cas de test unilatéral comme bilatéral, si $t^{stat} \in W_\alpha$, alors on rejette H_0 et on considère que la corrélation est significative (ou positive/négative dans la cas des tests unilatéraux), pour un seuil de significativité α . Inversement, si $t^{stat} \notin W_\alpha$, alors on ne rejette pas H_0 et on considère que la corrélation n'est pas significative (ou positive/négative dans la cas des tests unilatéraux), pour un seuil de significativité α .

On appliquera de nombreuses fois ce test corrigé dans cette thèse. Entre autres, on présentera des cartes de corrélation avec tests entre un indice climatique quelconque (Y), et le champ d'une variable météorologique donnée $X_1, \dots, X_q$ définie sur q points de grille.

2.3.4 Test du coefficient de pente d'un modèle de régression linéaire simple

Dans la même veine que pour les cartes de corrélations, on étudiera également à l'aide d'un modèle de régression linéaire simple (à une variable, section 2.2.2) le lien entre un indice climatique quelconque (Y), et le champ d'une variable météorologique donnée $X_1, \dots, X_q$ défini sur q points de grille. En effet, de manière équivalente à la corrélation en matière de signe, le coefficient de pente d'un tel modèle de régression peut renseigner sur l'intensité et le signe d'un lien entre deux variables

et ainsi être présenté sous forme de cartes. On note X une série quelconque parmi $X_1, \dots, X_q$. Notons ici que $X = (X_t)_{t \in T}$ et $Y = (Y_t)_{t \in T}$ sont définis sur le même espace de temps T . On peut alors poser le modèle de régression linéaire simple $\forall 1 \leq i \leq n$:

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i, \epsilon_i \sim \mathcal{N}(0, \sigma^2) \quad (2.79)$$

L'idée du test ici est de déterminer si $\hat{\beta}_1$ permet de dire que β_1 est significativement différent de 0 (dans le cas bilatéral). On peut alors poser les hypothèses suivantes :

$$\begin{cases} H_0 : \beta_1 = 0 \\ H_1 : \beta_1 \neq 0 \end{cases} \quad (2.80)$$

Si l'on note s_{XY} , la covariance empirique de X et Y , on peut dire d'après (2.36) que :

$$\hat{\beta}_1 = \frac{s_{XY}}{s_X^2} = \frac{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)(Y_i - \bar{Y}_n)}{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2} \quad (2.81)$$

Pour rappel, β_1 et β_0 sont les coefficients de régression théoriques, et ne sont pas aléatoires. De même, les $(X_i)_{1 \leq i \leq n}$ sont, par hypothèse, exogènes, et sont également non-aléatoires (voir 4^{me} hypothèse du modèle linéaire section 2.2.2). Ainsi, on peut dire, $\forall 1 \leq i \leq n$, que $\mathbb{E}[Y_i] = \beta_0 + \beta_1 X_i$. On peut alors montrer que $\text{Var}(\epsilon_i) = \text{Var}(Y_i) - \text{Var}(\beta_0 + \beta_1 X_i) = \text{Var}(Y_i) = \sigma^2$ (par hypothèse, section 2.2.2). Puisque les ϵ_i sont gaussiens, alors les Y_i sont gaussiens, tels que, $\forall 1 \leq i \leq n$, $Y_i \sim \mathcal{N}(\beta_0 + \beta_1 X_i, \sigma^2)$. On peut donc en déduire que $\hat{\beta}_1$ est une variable aléatoire gaussienne puisque les Y_i et $\bar{Y}_n$ sont des variables aléatoires gaussiennes. D'après ce qui a été dit dans ce paragraphe, on peut par calcul montrer que :

$$\mathbb{E}[\hat{\beta}_1] = \beta_1 \quad (2.82)$$

$$\text{Var}(\hat{\beta}_1) = \frac{\sigma^2}{\sum_{i=1}^n (X_i - \bar{X}_n)^2} \quad (2.83)$$

D'après (2.82) et (2.83), le TCL nous dit que :

$$\frac{\hat{\beta}_1 - \beta_1}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^n (X_i - \bar{X}_n)^2}}} \sim \mathcal{N}(0, 1) \quad (2.84)$$

Puisque la variance résiduelle σ^2 est toujours inconnue, on l'estime par son estimateur sans biais :

$$s^2 = \frac{1}{n-2} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 = \frac{1}{n-2} \sum_{i=1}^n \hat{\epsilon}_i^2 \quad (2.85)$$

D'après le théorème de Cochran, on peut montrer que :

$$\frac{(n-2)s^2}{\sigma^2} \sim \chi_{n-2}^2 \quad (2.86)$$

Donc, d'après (2.86) et la définition de la loi de Student, on peut dire que :

$$\frac{\frac{\hat{\beta}_1 - \beta_1}{\sqrt{\sigma^2 \sum_{i=1}^n (X_i - \bar{X}_n)^2}}}{\sqrt{\frac{(n-2)s^2}{\sigma^2}}}{n-2} = \frac{\hat{\beta}_1 - \beta_1}{s / \sum_{i=1}^n (X_i - \bar{X}_n)^2} \sim St(n-2) \quad (2.87)$$

Puisque on aura toujours $n > 30$, l'approximation normale nous permet de dire :

$$\frac{\hat{\beta}_1 - \beta_1}{s / \sum_{i=1}^n (X_i - \bar{X}_n)^2} \sim \mathcal{N}(0, 1) \quad (2.88)$$

Si l'on note t^{stat} la réalisation de Z , on peut définir la zone de rejet W_α telle que :

$$W_\alpha = \{|t^{stat}| > u_{1-\frac{\alpha}{2}}\} \quad (2.89)$$

Où $u_{1-\frac{\alpha}{2}}$ est le quantile d'ordre $1 - \frac{\alpha}{2}$ de la loi normale centrée réduite.

Comme pour les sections précédentes, les tests unilatéraux consistent à utiliser les $u_{1-\alpha}$ et u_α pour la définition de W_α (voir sections 2.3.1 à 2.3.3).

Pour les trois procédures, H_0 est rejetée si $t^{stat} \in W_\alpha$, c'est-à-dire que le coefficient de pente du modèle est considérée comme significativement différent de 0 dans le cas bilatéral, et significativement positif ou négatif dans le cas unilatéral, au risque $\alpha \in]0, 1[$. Dans le cas contraire, si $t^{stat} \notin W_\alpha$, H_0 n'est pas rejetée, et le coefficient de pente du modèle n'est pas considéré comme significativement différent de 0 dans le cas bilatéral, et n'est pas significativement positif ou négatif dans le cas unilatéral, au risque $\alpha \in]0, 1[$.

2.3.5 Méthodes de Monte-Carlo et exemple d'application

Les tests présentés ci-avant ne sont applicables que dans des contextes particuliers, permettant ainsi de poser une loi de probabilité bien connue sous H_0 , que suivra la statistique de test calculée sous cette même hypothèse. Fréquemment, lorsque l'on étudie des données réelles, on peut être amené à ne pas connaître la loi sous H_0 de la statistique de test calculée. C'est dans le but de pallier à ce problème que les approches de Monte-Carlo existent. En effet, dans cette thèse, on va parfois être amené à calculer des statistiques tellement dépendantes du contexte qu'aucun test usuel ne peut être appliqué. L'approche de Monte-Carlo a pour but, par le biais d'un grand nombre de réalisations, de simuler la loi de probabilité inconnue qui se cache derrière une statistique calculée, et ainsi d'estimer si cette dernière a un caractère exceptionnel ou non.

Prenons un exemple que nous serons amenés à rencontrer. Pour étudier la réponse d'un signal climatique donné X l'année suivant de très grosses éruptions volcaniques, on peut regarder l'écart moyen entre la valeur de l'indice lors des années de grosses éruptions et lors des années suivant ces éruptions. Admettons alors que l'on note $v^{(10)}$ la valeur de l'écart moyen de l'indice entre les 10 années d'éruption et les années suivantes correspondantes. Supposons que l'on cherche à savoir si la

réponse de X aux très grosses éruptions est positive, on peut alors se proposer de tester les hypothèses suivantes :

$$\begin{cases} H_0 : "v^{(10)} = 0" = "Pas d'effet des très grosses éruptions sur l'indice" \\ H_1 : "v^{(10)} > 0" = "Réponse positive de l'indice après de très grosses éruptions" \end{cases} \quad (2.90)$$

On va alors chercher à estimer la loi de probabilité sous H_0 . Pour cela, on va réaliser un grand nombre (noté M) de tirages aléatoires de 10 années quelconques au sein de la série X , et regarder l'écart moyen entre ces années tirées aléatoirement et les années suivantes. Pour chacun des M tirages aléatoires de 10 années, puisque des années de très grosses éruptions seront rarement tirées dans un nombre important (étant donné qu'elles sont rares), l'écart moyen correspond à une réalisation de $v^{(10)}$ sous H_0 . On obtient alors M réalisations de $v^{(10)}$ sous H_0 que l'on ordonne. On les note alors $v_{(1)}^{(10)}, \dots, v_{(M)}^{(10)}$, tels que $v_{(1)}^{(10)} < v_{(2)}^{(10)} < \dots < v_{(M)}^{(10)}$. On peut alors, à partir de cette densité de probabilité estimée de $v^{(10)}$, définir une région de rejet de H_0 W_α , pour $\alpha \in]0, 1[$ telle que :

$$W_\alpha = \{v^{(10)} > v_{(\lfloor (1-\alpha) \times M \rfloor)}^{(10)}\} \quad (2.91)$$

Ainsi, si $v^{(10)} \in W_\alpha$, ceci veut dire que la moyenne des écarts des réponses à 1 an aux 10 volcans est supérieure à $\lfloor (1 - \alpha) \times M \rfloor$ des moyennes de 10 écarts entre deux années successives quelconques de la série parmi les M calculées. On considérera alors au risque α que les très grosses éruptions volcaniques impliquent significativement une réponse positive à un 1 an de la série X . Notons que l'on peut reproduire exactement la même procédure de test pour la deuxième année après l'éruption, mais aussi la troisième et ainsi de suite. Quoi qu'il en soit, dès lors que la manière dont se calcule la statistique de test est bien reproduite dans l'estimation de sa distribution de probabilité, l'approche de Monte-Carlo est applicable.

Puisque pour ce type de procédure chaque notation va dépendre de la statistique calculée et des hypothèses testées, les notations ci-avant ne se veulent pas être des notations mathématiques formelles et ont uniquement pour vocation de faciliter la compréhension de ce type d'approche.

2.4 Méthodes de *machine learning*

Dans cette thèse, on s'intéressera à la reconstruction de variables météorologiques, ou bien d'indices climatiques calculés à partir de variables météorologiques. Les algorithmes présentés par la suite seront donc appliqués dans leurs versions dites de régression, puisque nous travaillerons uniquement sur des variables quantitatives. Les prédicteurs sont donc dans notre cas les enregistrements paléoclimatiques.

2.4.1 Contexte théorique

On se placera ici dans un contexte de régression à une seule variable à expliquer, à savoir un indice climatique. Afin de reprendre les notations de la section 2.2, on

note $Y = (Y_t)_{t \in T}$ la série d'observations d'un indice climatique défini sur un espace de temps T quelconque.

Soit alors $X = (X_t^j)_{t \in T, 1 \leq j \leq p}$ une matrice donnée de p enregistrements paléoclimatiques observés sur la même période que $Y : T$. On note alors $X'_t = (X_t'^j)_{t \in T', 1 \leq j \leq p}$, une matrice des mêmes p enregistrements observés sur une période T' telle que $T \cap T' = \emptyset$, avec $\#(T') = m$. Notons qu'ici, pour éviter que certaines méthodes de régression ne donnent trop de poids aux données paléoclimatiques à forte variance, les données de X seront centrées et réduites et les données de X' seront centrées et réduites sur les valeurs initiales de X . Ceci s'applique de manière très fréquente en reconstruction paléoclimatique (Luterbacher et al. 2001, Cook et al. 2002, Luterbacher et al. 2002, Gray et al. 2004, Li et al. 2011, Ortega et al. 2015, Wang et al. 2017).

On cherchera ici à étudier et à modéliser des séries dont les pas de temps sont annuels. On peut donc, $\forall i \in T$, voir le vecteur $(Y_i, X_i^1, \dots, X_i^p)$ comme un ensemble de mesures de l'état du système climatique à l'année i . Il s'agira donc de reconstruire Y sur T' à partir de ses relations avec les enregistrements paléoclimatiques sur T , pré-établies à l'aide d'un algorithme de *machine learning*. Ce que l'on va appeler un algorithme de *machine learning* est une procédure permettant d'ajuster une méthode statistique de la manière la plus robuste possible aux données de X et de Y (on parle d'optimisation de la méthode), d'évaluer la qualité de ce modèle optimisé, et enfin, d'appliquer ce modèle entraîné et optimisé sur de nouvelles données paléoclimatiques de X' . Ainsi, les notations présentées ici cherchent à être très générales, puisque par la suite T' peut très bien être une sous-période de la période instrumentale sur laquelle on cherche à évaluer un modèle statistique, mais elle peut également être une période pré-industrielle sur laquelle on applique une reconstruction avec un modèle statistique optimisé.

Il existe une très grande variété de méthodes de régression que l'on peut appliquer dans un contexte de *machine learning*. Il existe à cet égard un théorème (théorème du *No free lunch*, Wolpert 1996) qui implique qu'il est strictement impossible qu'il existe une méthode statistique d'apprentissage meilleure que toutes les autres sur tous les jeux de données, d'où l'abondance des méthodes existantes. Le théorème du *No free lunch* est très puissant puisqu'au delà de l'implication précédemment évoquée, ce dernier implique de manière générale que pour toute méthode d'apprentissage A et B, on peut toujours trouver un jeu de données sur lequel A est meilleure que B, et inversement. Dans cette thèse, nous explorerons en particulier 4 méthodes, dont 3 n'ont pas été utilisées dans un contexte de reconstruction paléoclimatique sur le dernier millénaire. Ces méthodes, ainsi que les approches permettant de les optimiser et de les comparer au vu d'un jeu de données donné, sont expliquées dans le détail dans la suite de cette section. Les trois nouvelles méthodes ont été choisies pour différentes raisons qui seront également explicitées dans les sections correspondantes (2.4.2 à 2.4.5).

Les méthodes de régression seront présentées en deux parties : une première qui explique la phase de modélisation, qui revient à établir des liens entre Y et les données paléoclimatiques X selon une approche statistique donnée, et une seconde, que l'on appellera la phase de reconstruction, qui consiste à donner une estimation de Y sur T' à partir des relations établies lors de la première phase et des données paléoclimatiques X' .

2.4.2 PCR

Modélisation

Comme évoqué précédemment, la PCR est l'une des méthodes de reconstruction d'indices climatiques les plus répandues (Luterbacher et al. 2001, Cook et al. 2002, Luterbacher et al. 2002, Gray et al. 2004, Li et al. 2011, Ortega et al. 2015, Wang et al. 2017). Comme l'indique son nom, la PCR consiste à établir un modèle de régression linéaire de Y par une sous-matrice de la matrice U , où U est la matrice des PC de la matrice d'enregistrements paléoclimatiques X (cf. section 2.2.1 pour la construction de U). La différence avec la section 2.2.1 est ici que la matrice X n'est pas un champ de données météorologiques, mais l'analyse en PC s'applique de la même manière dès lors que les lignes de X représentent le temps et que ses colonnes représentent les différents enregistrements.

Comme expliqué section 2.2.1, les PC sont construites de manière à ce qu'elles soient orthogonales les unes aux autres (corrélations nulles deux-à-deux, section 2.2.1) et que la fraction de variance expliquée par chacune d'entre-elles est décroissante. La réduction de dimension étant l'objectif principal d'une analyse PCA, on va ici chercher à restreindre le nombre de PC que l'on utilise pour le modèle de régression linéaire, mais uniquement celles à fortes variances expliquées (les premières), qui permettent le mieux de reconstruire Y . Le choix du nombre $q \leq p$ de variables utilisées pour la régression linéaire de Y se fera par validation croisée, une méthode décrite dans la section 2.4.6 et mathématiquement détaillée dans le chapitre 3. C'est ce choix du paramètre q , inhérent à la méthode, que l'on appelle l'optimisation du modèle de PCR.

Après optimisation du modèle, on note la matrice $U = (U^j)_{j \leq 1 \leq q}$, contenant les q premières PC de X , que l'on utilise comme régresseurs linéaires de Y , ce qui, selon la section 2.2.2, donne le modèle suivant :

$$Y = \beta_0 + \beta_1 U^1 + \cdots + \beta_q U^q + \epsilon \quad (2.92)$$

La phase de modélisation sur T' se termine par l'estimation des coefficients β . Toujours selon la section 2.2.2, l'estimateur des moindres carrés, qui minimise $\|\hat{\epsilon}\| = \|Y - \hat{Y}\|$, est donné par :

$$\hat{\beta} = (U^T U)^{-1} U^T Y \quad (2.93)$$

Où, $U = [\mathbb{1}_n, U]$ (section 2.2.2).

Reconstruction

Avant d'appliquer les coefficients de régression estimés, il va s'agir d'obtenir les coordonnées de X' dans la base orthogonale composée des colonnes de V , la matrice des EOF de X (section 2.2.1). Pour cela, on projette X' sur la matrice V . La matrice de projection de X' dans la base des EOF de X est notée U' :

$$U' = X' V \in \mathbb{R}^{m \times p} \quad (2.94)$$

La reconstruction s'obtient alors en appliquant les coefficients de régression estimés précédemment sur les q premières PC projetées de U' . On obtient alors $\hat{Y}$, l'estimation de Y sur T' :

$$\hat{Y} = \hat{\beta}_0 + U^1 \hat{\beta}_1 + U'^2 \hat{\beta}_2 + \dots + U'^p \hat{\beta}_p \quad (2.95)$$

2.4.3 Moindres carrés partiels

La PCA cherche à maximiser la variance initiale de X lorsque sont construites les EOF. Cependant, ces EOF sont construites sans prendre en compte les variations de Y , ce qui peut amener à construire des régresseurs expliquant beaucoup de variance initiale de X , mais qui ne permettent pas forcément de bien reconstruire Y . Une méthode alternative et proche de la PCA, qui permet de pallier à ce problème, est la méthode de régression par moindres carrés partiels (*Partial Least Squares*, PLS), introduite par Wold et al. (1984).

Modélisation

L'idée de cette méthode est également de régresser Y par des vecteurs orthogonaux, issus d'une transformation de X . La différence est qu'ici il s'agit de maximiser la covariance entre y la projection de X sur les vecteurs de la base orthogonale.

Puisque la base orthogonale n'est pas construite de la même manière que pour la PCR, on appellera variables latentes (*Latent variables*, LVs) les équivalents des PC pour la PCR.

La première LV, notée ξ^1 , est une combinaison linéaire de $X^1, \dots, X^p$. C'est à dire qu'il existe $v_1 \in \mathbb{R}^p$ tel que $Xi^1 = Xv_1$, et qui vérifie :

$$v_1 = \arg \max_{\substack{v \in \mathbb{R}^p \\ \|v\|=1}} Cov(Y, Xv) \quad (2.96)$$

De manière analogue à la PCR, on construit les LVs suivantes en ajoutant les contraintes d'orthogonalité successives :

$$v_2 = \arg \max_{\substack{v \in \mathbb{R}^p \\ \|v\|=1 \\ v^T v_1 = 0}} Cov(Y, Xv) \quad (2.97)$$

$\vdots$

$$v_p = \arg \max_{\substack{v \in \mathbb{R}^p \\ \|v\|=1 \\ v^T v_1 = 0 \\ \dots \\ v^T v_{p-1} = 0}} Cov(Y, Xv) \quad (2.98)$$

On note $\Xi = (\xi^j)_{1 \leq j \leq p}$ la matrice des LVs. Les $v_1, \dots, v_p$ sont donc les vecteurs qui forment la base orthogonale sur laquelle est projetée X , ce sont donc les vecteurs équivalents aux EOF dans le cadre de la PCR. Il s'agit bien de EOF au sens de la définition, mais pour ne pas porter à confusion, nous les appellerons *loadings*, comme ils sont régulièrement appelés (Wold et al. 1984).

Le désavantage de cette méthode est que le calcul des *loadings* est beaucoup plus coûteux car ils s'obtiennent à l'aide d'algorithmes récursifs et non pas simplement par l'inversion d'une matrice. L'algorithme PLS1 (Höskuldsson 1988) est celui qui sera utilisé au cours de cette thèse. Son pseudo-code est le suivant :

```

1: procedure PLS1
2:    $X_0 \leftarrow X$ 
3:   for  $h = 1, \dots, r$ 
4:      $v_h \leftarrow \frac{X_{h-1}^T Y}{\|X_{h-1}^T Y\|^2}$ 
5:      $\tilde{\zeta}_h \leftarrow X_{h-1} v_h$ 
6:      $X_h = X_{h-1} - \frac{\tilde{\zeta}_h \tilde{\zeta}_h^T}{\|\tilde{\zeta}_h\|^2} X_{h-1}$  (Phase de déflation)
7: end procedure

```

Comme pour la PCR, le modèle sera optimisé (section 2.4.5) sur un nombre $r < p$ de variables latentes utilisées pour le modèle linéaire.

Après optimisation, on va alors estimer par moindres carrés les coefficients de régression du modèle linéaire :

$$Y = \beta_0 + \beta_1 \tilde{\zeta}^1 + \dots + \beta_r \tilde{\zeta}^r \quad (2.99)$$

Avec l'estimateur donné par :

$$\hat{\beta} = (\Psi^T \Psi)^{-1} \Psi^T Y \quad (2.100)$$

Où $\Psi = [\mathbb{1}_n, \Xi]$ (section 2.2.2).

Reconstruction

La reconstruction par la méthode PLS est donnée de la même manière que pour la PCR. On calcule d'abord la matrice des LVs étendues sur T' par projection de X' sur la base orthogonale V :

$$\Xi' = X'V \in \mathbb{R}^{m \times p} \quad (2.101)$$

La reconstruction s'obtient en appliquant les coefficients de régression estimés précédemment aux nouvelles données projetées dans la base orthogonale des LVs :

$$\hat{Y} = \Psi' \hat{\beta} \in \mathbb{R}^m \quad (2.102)$$

2.4.4 Régression Elastic Net

Sans utiliser de transformation orthogonale des enregistrements initiaux comme pour les méthodes PCR et PLS, le modèle de reconstruction le plus simple que l'on peut poser est le modèle de régression linéaire multiple (section 2.2.2), appliqué directement aux enregistrements. L'estimateur des moindres carrés est dans ce cas la solution du problème d'optimisation convexe de la fonction de coût L , donnée

par :

$$L(\beta) = \|Y - \sum_{j=1}^p \beta_j X^j\|^2 \quad (2.103)$$

Cependant, il a été montré que malgré l'abondante utilisation du modèle linéaire, l'estimateur des moindres carrés donne très souvent des résultats relativement peu fiables, du fait qu'il se base sur un ensemble d'hypothèses telles que l'homoscédasticité et la normalité des résidus, trop peu souvent vérifiées (Poole and O'Farrell 1971). Notons que c'est également un problème pour les méthodes PCR et PLS, puisque chacune d'entre elles fait appel, *in fine*, à un estimateur des moindres carrés. C'est dans le but de pallier à ce problème que furent proposés les modèles de régression dits « pénalisés » ou « régularisés ». Ici nous présenterons la méthode d'*Elastic Net* (Enet, Zou and Hastie 2005), qui utilise les concepts de deux autres types de régressions pénalisées : la méthode du Lasso (Tibshirani 1996) et la régression Ridge (Hoerl and Kennard 1970). Ces méthodes furent développées dans le but d'éviter la forte variabilité de l'estimateur des moindres carrés quand le nombre de variables explicatives augmente. La méthode Enet a l'avantage qu'une fois ses paramètres optimisés, elle peut découler en un modèle Lasso ou en un modèle Ridge, ou bien alors en une combinaison de ces deux approches. Les méthodes de pénalisation ont pour but d'ajouter un terme à l'équation de la fonction de coût. Dans le cas du Lasso et de la régression Ridge, le terme de pénalisation agit comme une contrainte de seuil, basée sur la norme l_k de β : $\|\beta\|_k^k = \sum_{j=1}^p |\beta_j|^k$. Avec $k = 1$ pour le Lasso, et $k = 2$ pour la régression Ridge. Les fonctions de coût des trois approches sont données par :

$$L^{Ridge}(\beta) = \|Y - \sum_{j=1}^p \beta_j X^j\|^2 + \lambda_2 \sum_{j=1}^p \beta_j^2 \quad (2.104)$$

$$L^{Lasso}(\beta) = \|Y - \sum_{j=1}^p \beta_j X^j\|^2 + \lambda_1 \sum_{j=1}^p |\beta_j| \quad (2.105)$$

$$L^{Enet}(\beta) = \|Y - \sum_{j=1}^p \beta_j X^j\|^2 + \lambda_1 \sum_{j=1}^p |\beta_j| + \lambda_2 \sum_{j=1}^p \beta_j^2 \quad (2.106)$$

Avec $\lambda_1, \lambda_2 > 0$.

Si l'on note $w = (w_j)_{1 \leq j \leq p} = (\text{sgn}(\beta_j))_{1 \leq j \leq p}$, où $\text{sgn}(\bullet)$ est la fonction signe, les fonctions de coûts peuvent alors s'écrire :

$$L^{Ridge}(\beta) = \|Y - \sum_{j=1}^p \beta_j X^j\|^2 + \lambda_2 \beta^T \beta \quad (2.107)$$

$$L^{Lasso}(\beta) = \|Y - \sum_{j=1}^p \beta_j X^j\|^2 + \lambda_1 w^T \beta \quad (2.108)$$

$$L^{Enet}(\beta) = \|Y - \sum_{j=1}^p \beta_j X^j\|^2 + \lambda_1 w^T \beta + \lambda_2 \beta^T \beta \quad (2.109)$$

Les estimateurs obtenus par la résolution des problèmes convexes donnés par les fonctions de coût (2.107 à 2.109) sont donnés par :

$$\hat{\beta}^{Lasso} = (X^T X)^{-1} (X^T Y - \frac{\lambda_1}{2} w) \quad (2.110)$$

$$\hat{\beta}^{Ridge} = (X^T X + \lambda_2 I)^{-1} X^T Y \quad (2.111)$$

$$\hat{\beta}^{Enet} = (X^T X + \lambda_2 I)^{-1} (X^T Y - \frac{\lambda_1}{2} w) \quad (2.112)$$

Zou and Hastie (2005) ont montré que l'estimateur $\hat{\beta}^{Enet}$ peut également s'écrire :

$$\hat{\beta}^{Enet} = (X^T X + (1 - \alpha) I)^{-1} (X^T Y - \frac{\alpha \lambda}{2} w) \quad (2.113)$$

Où $\alpha \in [0, 1]$. Ainsi on remarque que si $\alpha = 1$, $\hat{\beta}^{Enet}$ est un estimateur de régression Ridge, tandis que si $\alpha = 0$, il s'agit d'un estimateur de Lasso. Sinon, il s'agira effectivement d'un estimateur Enet. Ceci a beaucoup d'implications puisque l'optimisation de α permet en réalité de sélectionner le meilleur modèle parmi les trois. Ainsi le paramètre α permet de contrôler le poids relatifs des pénalisations associées aux normes l_1 et l_2 . Le paramètre λ va, quant à lui, permettre de contrôler l'intensité globale de pénalisation du modèle. Ici, la méthode Enet requiert l'optimisation de deux paramètres, ce qui rend l'optimisation de cette méthode beaucoup plus coûteuse que la PLS et la PCR comme nous le verrons dans la section 2.4.6 et le chapitre 3. Ces dernières ne requièrent en effet l'optimisation que d'un paramètre de contrôle.

2.4.5 Régression par forêt aléatoire

La méthode de forêt aléatoire (*Random forest*, RF) est une méthode d'agrégation de prédicteurs non linéaires, formellement introduite par Breiman (2001). Pour l'anecdote, l'auteur de l'article, Leo Breiman, décède malheureusement peu après la publication de celui-ci (en 2005), qui était alors cité un peu plus de 200 fois dans la littérature scientifique. Or cette méthode est relativement coûteuse. Mais avec les progrès accélérés du stockage des données et des capacités de calcul, ce coût s'est amorti et la méthode a connu un succès fulgurant : l'article de Breiman (2001) est désormais cité plus de 60000 fois dans des articles scientifiques (statistiques Google Scholar en juillet 2020). Ceci montre à quel point cette méthode est abondamment utilisée dans de nombreux champs scientifiques, témoignant également de la qualité des résultats qu'elle permet d'obtenir. Cette méthode s'inscrit dans la catégorie des méthodes d'apprentissage d'ensemble. Les méthodes d'ensemble consistent à considérer un ensemble de prédictions données par différents modèles statistiques et de les agréger, selon une combinaison donnée. À l'inverse des algorithmes précédemment évoqués, la RF peut également s'appliquer dans un cadre de classification, où l'on cherche à prédire (ou reconstruire) des variables catégorielles. Il existe d'autres méthodes d'ensemble, très en vogue ces dernières années, dont on peut citer en particulier les méthodes du gradient boosté (*Gradient boosting*, Friedman 1999) et d'arbres extrêmement randomisés (*Extremely randomized trees*, Geurts et al. 2006). Celles-ci ne sont pas explorées dans cette thèse mais il est clair que leur application dans un cadre de reconstruction paléoclimatique pourrait être d'un très grand intérêt.

La formalisation mathématique de la RF est relativement modeste au vu de son coût en matière de temps et de ressources de calcul par rapport à d'autres méthodes, comme celle présentée ci-avant. En effet, il s'agit simplement d'une moyenne de différentes prédictions données par des arbres de régression (Breiman 2001). On appelle en général CART (*Classification And Regression Trees*) ces arbres de décision, introduits par Breiman et al. (1984), qui sont des régresseurs non linéaires, c'est-à-dire qu'on ne peut les écrire comme une combinaison linéaire des enregistrements (ou de versions modifiées de ces derniers comme pour la PCR ou la PLS par exemple).

Avant de formaliser la RF, nous devons préalablement définir ce qu'est un arbre de régression. On considère l'ensemble d'apprentissage défini sur la période historique, constitué de l'indice climatique et des p enregistrements utilisés, que l'on note $\{(Y_i, X_i)_{1 \leq i \leq n}\}$, qu'on appelle n -uplet. On a donc que $X_i = (X_i^j)_{1 \leq j \leq p}$ est le vecteur d'observation d'enregistrements paléoclimatiques de l'année i , et Y_i l'observation de l'indice climatique pour la même année. Tout comme pour la PLS, l'arbre de régression est un algorithme récursif (2.4.3), c'est-à-dire qu'il se constitue de plusieurs étapes, dont chacune va dépendre des précédentes. De cette manière, l'arbre se construit de sa racine jusqu'à ses feuilles. Notons que dans la théorie des arbres (qu'ils soient décisionnels ou non), on appelle racine la base de l'arbre, qui représente en général le départ d'un algorithme récursif. On appelle nœud chaque étape intermédiaire d'un arbre. On appellera nœud-fils d'un nœud donné, tout nœud issu de ce nœud. Enfin, on appellera feuille tout nœud qui à la fin de l'algorithme, ne possède aucun nœud-fils (Fig. 2.1).

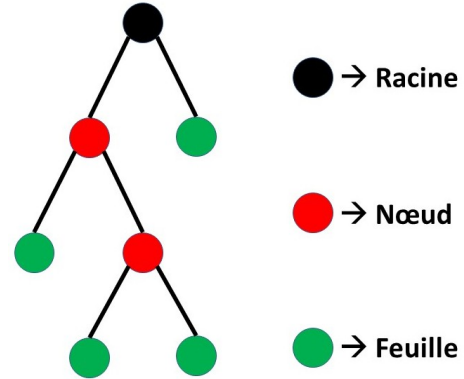


FIGURE 2.1 | Terminologie d'un arbre computationnel quelconque.

La première étape de l'algorithme consiste à considérer l'ensemble des années d'observation et de déterminer, parmi les p enregistrements, lequel permet de séparer en deux nœud-fils les données de l'indice climatique, telle que la somme de ses variances au sein des deux nœuds-fils est minimale. On appelle coupe le fait de séparer un nœud de l'arbre de décision en deux nœud-fils. On peut noter toute coupe de la manière suivante :

$$\{X^j < d\} \cup \{X^j > d\} \quad (2.114)$$

Où d est la valeur-seuil à partir de laquelle toute année d'observation dont la $j^{\text{ème}}$ variable y est supérieure sont placées dans le nœud-fils droit. De la même manière, toute année d'observation dont la $j^{\text{ème}}$ variable y est inférieure à d sont placées dans le nœud-fils gauche. Trouver la coupe optimale de la racine de l'arbre revient alors

à résoudre le problème d'optimisation convexe à deux variables suivant :

$$(j, d) = \arg \min_{\substack{d \in \mathbb{R} \\ 1 \leq j \leq p}} \sum_{i: X_i^j < d} (Y_i - \frac{1}{\#(\{i : X_i^j < d\})} \sum_{i: X_i^j < d} Y_i)^2 \\ + \sum_{i: X_i^j > d} (Y_i - \frac{1}{\#(\{i : X_i^j > d\})} \sum_{i: X_i^j > d} Y_i)^2 \quad (2.115)$$

Ensuite, l'algorithme s'applique à chacun des nœuds-fils afin de continuer la construction de l'arbre, comme schématisé Fig. 2.2. Une fois l'arbre construit, on peut voir chaque feuille soit comme une valeur de Y (s'il ne reste qu'une année d'observation dans cette feuille), soit comme une réalisation de la distribution de probabilité de Y au sein de cette feuille (Breiman et al. 1984). Notons que si l'on pousse un arbre jusqu'à sa profondeur maximale, c'est-à-dire que chaque feuille ne contient qu'une seule année d'observation, ceci revient à appliquer une méthode de k plus proche voisin avec $k = 1$. Afin qu'un arbre se distingue de cette méthode, on pose au préalable un critère d'arrêt. Dans le cas d'une régression, on s'arrête dès lors que l'un des nœuds contient un nombre s (ou moins) d'années d'observations (Breiman et al. 1984).

Le problème des arbres de décision, c'est qu'ils sont extrêmement sensibles aux données sur lesquelles on les entraîne. Ainsi, en partant d'un échantillon d'apprentissage donné, le fait d'ajouter ou de retirer quelques variables (ici enregistrements paléoclimatiques) ou quelques années d'observations peut très largement modifier la reconstruction donnée par un arbre (Breiman 2001). En effet, si l'on supprime la variable permettant de couper la racine de l'arbre, alors ce dernier risque d'être construit de manière totalement différente, et donc de donner des reconstructions très différentes sur un même jeu d'années d'observation. Autrement dit, les arbres sur-apprennent les données sur lesquels ils ont été construits, c'est-à-dire qu'ils sont trop dépendants de l'échantillon d'apprentissage et donc peu efficaces sur l'échantillon à reconstruire. C'est de ce constat qu'est née la méthode de RF. L'idée proposée par Breiman (2001) est de générer un très grand nombre d'arbres (d'où le terme de « forêt »), pour lesquels on va tirer aléatoirement (d'où le terme « aléatoire ») $m < p$ enregistrements paléoclimatiques pour construire chacun d'entre eux. Le but de cette procédure est d'obtenir un algorithme de reconstruction qui n'est pas trop dépendant de un ou de quelques enregistrements.

Reconstruction

Une fois que B arbres ont été construits, il suffit, à partir des mesures d'enregistrements paléoclimatiques de la matrice X' , de parcourir chacun des arbres pour chaque pas de temps de T' . Pour un arbre donné dont les valeurs de X' terminent dans la feuille $\mathcal{F}$ pour l'année $k \in T'$, la reconstruction est donnée par $\frac{1}{\#(\mathcal{F})} \sum_{\substack{i: Y_i \in \mathcal{F} \\ i \in T}} Y_i$, c'est à dire la moyenne de l'indice climatique pour les années d'observation ayant été placées dans cette feuille au cours de la phase de modélisation. La reconstruction de RF finale est alors donnée par la moyenne des prédictions des arbres (Fig. 2.2). Ainsi, si l'on considère un ensemble de B arbres, et que l'on note $\hat{Y}_i^{(1)}, \dots, \hat{Y}_i^{(B)}$ la reconstruction de chacun de ces arbres pour l'année i , la reconstruction donnée par

la RF pour la même année est :

$$\hat{Y}_i = \frac{1}{B} \sum_{k=1}^B \hat{Y}_i^{(k)} \quad (2.116)$$

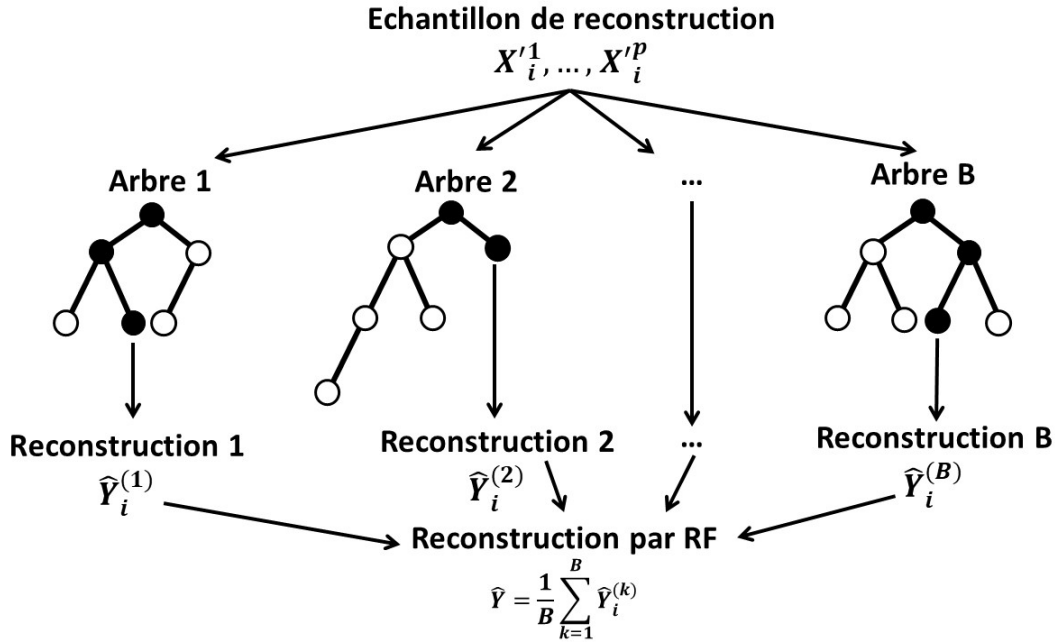


FIGURE 2.2 | Schéma de la phase de reconstruction d'un modèle de forêt aléatoire. On se place dans le cas d'une forêt à B arbres pour la reconstruction d'une année donnée $i \in T'$. Les feuilles et nœuds noirs représentent la partie parcourue par les données paléoclimatiques de l'année X'_i , selon les arbres construits sur la matrice X . Les feuilles et nœuds blancs sont ceux non parcourus par les données de X'_i .

Le fait même de générer un grand nombre d'arbres rend par nature la méthode de RF relativement coûteuse en matière de ressources et de temps de calcul, d'autant plus si ses paramètres doivent être optimisés. Le très grand avantage qu'a cette méthode en matière de complexité algorithmique est que bien qu'elle possède *a priori* 3 paramètres (s , m et B), un seul requiert d'être optimisé. En effet, Breiman (2001) stipulait dès la publication de sa méthode que pour différents jeux de données testés, un critère d'arrêt $s = 5$ donnait toujours des résultats très satisfaisants, et que le gain de son optimisation en terme de robustesse était négligeable au vu du temps de calcul qu'elle implique. Ceci a également été confirmé sur une grande variété de jeux de données par Goldstein et al. (2011). De la même manière, Oshiro et al. (2012) ont montré qu'au delà de $B = 128$ arbres générés, la reconstruction donnée par la RF converge (pour une valeur de m donnée). Ainsi, seul le paramètre m requiert une optimisation par validation croisée, ce qui rend très satisfaisant le temps et les ressources de calculs requis pour la paramétrisation de cette méthode.

2.4.6 Optimisation de paramètres de contrôle par validation croisée

Si l'on reprend l'exemple de la PCR (section 2.4.2) parmi les études précédentes l'ayant utilisé, il n'y a pas de consensus sur la méthode employée pour le choix du paramètre q . Tandis que certaines se fixent un seuil en terme de valeur absolue (Cook et al. 2002) ou encore de fraction (Wang et al. 2017) de la variance expliquée par les q PC sélectionnées. Dans le cas de Ortega et al. (2015), ces derniers

appliquent la règle N de Preisendörfer (1988) pour sélectionner q . Il s'agit d'une méthode de Monte-Carlo (section 2.3.5) ayant pour but de ne garder que les PC dont la fraction de variance expliquée dépasse dans 95% des cas le poids de PC provenant d'un grand nombre de matrices de faux enregistrements, dont les colonnes sont des bruits blancs simulés. Bien que ces méthodes soient adaptées pour la réduction de dimensions par ACP, elles ne permettent pas de déterminer quelles PC sont adéquates pour régresser un indice donné. C'est, entre autres, dans ce but de déterminer quelle est la configuration de paramétrisation optimale d'un modèle statistique, que les méthodes de validations croisées existent (Stone 1974, Geisser 1975). Les paramètres, inhérents à chaque méthode d'apprentissage statistique, se nomment les paramètres de contrôle, ou hyperparamètres. Pour les méthodes que nous allons employer dans cette thèse, les paramètres de contrôles sont explicités dans les sections précédentes.

La méthode de validation particulière qui sera utilisée se nomme la validation croisée *K-folds*. Elle a pour but de séparer les observations en K groupes, qui permettent alors de créer K découpages de deux ensembles : un ensemble d'apprentissage composé de $K - 1$ groupes, et un échantillon test composé d'un groupe. On va alors tester toutes les combinaisons possibles de paramètres sur les K ensembles d'apprentissage, et regarder quelle paramétrisation est la plus performante sur les échantillons tests. Dans le cas de paramètres définis dans des espaces continus (cf. *Enet*), on va discrétiser ces espaces dans le but d'avoir un nombre fini de paramètres à tester. Plus cette discrétisation est fine, plus la paramétrisation sera optimale, mais plus les temps de calculs en seront impactés. La validation croisée permet ainsi, pour une méthode donnée, de déterminer les paramètres appropriés de manière objective et transparente, ce qui manque cruellement dans les études précédentes de reconstruction d'indices climatiques. Les notations mathématiques de la validation croisée seront présentées de manière détaillée dans le chapitre 3.

2.4.7 Évaluation d'un modèle par *hold-out*

Puisqu'il n'y a, *a priori*, aucune raison de penser qu'une méthode de régression est meilleure qu'une autre (e.g. théorème du *No free lunch*), la méthode de *hold-out* permet, pour un certain jeu de données, d'estimer laquelle parmi différentes configurations de reconstruction est la meilleure. En effet, on a vu que la validation croisée permettait de déterminer une paramétrisation optimale pour une méthode de régression donnée. En revanche, ceci ne permet pas de déterminer, parmi deux reconstructions produites avec deux méthodes optimisées mais différentes, laquelle est la meilleure. Le *hold-out* est en réalité relativement proche de la validation croisée dans sa philosophie. Plutôt que de découper les années d'observation en un nombre donné de groupes, on va, un nombre donné de fois, découper les années d'observations en deux groupes : l'un sur lequel on va construire un modèle optimisé par validation croisée, et l'autre sur lequel on va tester ce modèle optimisé au travers d'une métrique mathématique, qui permet d'obtenir ce que l'on appelle un score (d'où le terme de *hold-out* qui signifie "laisser de côté"). On parle également de découpage apprentissage/test. Puisque les résultats obtenus peuvent être très différents pour deux découpages d'un même jeu de données, le but va être de réaliser cette opération un grand nombre de fois, ce qui permet d'obtenir un grand nombre de scores pour un modèle statistique optimisé donné. Ceci va alors permettre de

s'affranchir du biais impliqué par chacun des découpages. Dans un contexte de reconstruction paléoclimatiques, cette approche est d'autant plus intéressante qu'elle permet également de comparer les scores obtenus pour différents jeux d'enregistrements paléoclimatiques ou différentes fenêtres d'apprentissage. Les notations mathématiques de la méthode de *hold-out*, ainsi que des métriques associées pour calculer les scores, seront également présentées de manière détaillée dans le chapitre 3.

Chapitre 3

Développement d'un outil de reconstruction robuste et automatique

Ce chapitre a pour but de présenter un outil de reconstruction de signaux climatiques utilisable par tous. Cet outil permet d'utiliser des méthodes de régression avancées et peu répandues en climatologie, mais aussi d'évaluer la robustesse d'une reconstruction qu'il produit, notamment par des méthodes de validations croisées et de *hold-out*. La section 3.1 présente un état de l'art des reconstructions de signaux climatiques du dernier millénaire, et soulève le problème du manque d'une méthodologie robuste et vérifiée pour ce genre de question. La section 3.2 est une étude parue dans *Geoscientific Model Development* (Michel et al. 2020) qui présente cet outil, et l'applique à la reconstruction d'un indice de NAO. Enfin, la section 3.3 présente une discussion sur le potentiel et les limites de cet outil.

3.1 Reconstruction par apprentissage statistique

Dans la littérature, on trouve de nombreuses reconstructions d'indices climatiques (Luterbacher et al. 2001, Cook et al. 2002, Gray et al. 2004, Trouet et al. 2009, Li et al. 2011, Ortega et al. 2015, Wang et al. 2017, Cook et al. 2019). Cependant, il n'existe pas de méthodologie bien définie pour mettre en place ces reconstructions. Ainsi, chaque étude produit à sa façon une reconstruction. Certaines études se restreignent notamment à un nombre très limité de données paléo-climatiques (<3) (Trouet et al. 2009). L'incertitude de la reconstruction obtenue est donc très dépendante des incertitudes associées aux données utilisées, ce qui peut engendrer des interprétations erronées (Lehner et al. 2012). D'autres études ont implémenté des méthodes de régression plus élaborées, comme la PCR (section 2.4.2). Bien que cette méthode de régression soit abondamment répandue dans la reconstruction de signaux climatiques, il en existe de nombreuses autres, n'ayant jamais été utilisées pour ce type de reconstruction, comme par exemple celles présentées sections 2.4.3 (PCR), 2.4.4 (PLS) et 2.4.5. (Enet)

Au-delà de la méthode statistique, une reconstruction donnée dépend également des enregistrements utilisés comme prédicteurs pour la produire. La technique de sélection quasiment toujours employée est géographique : les enregistrements les plus proches des régions associés à l'indice climatique sont utilisés. Néanmoins, ces études souffrent souvent du fait qu'elle ne vérifient pas toujours le lien statistique entre l'indice climatique et l'enregistrement paléoclimatique en question. Si l'on prend pour exemple l'étude de Wang et al. (2017), qui proposent une reconstruction de la NASST par PCR (section 2.4.2), on voit que certains des enregistrements utilisés comme prédicteurs ont un coefficient de corrélation avec l'indice qui

n'est pas significatif, ce qui est susceptible d'apporter du bruit supplémentaire dans la reconstruction.

Par ailleurs, alors même que l'on se place dans un contexte de reconstruction, des études faisant appel dans leur procédure à une combinaison linéaire des enregistrements (dont la PCR, la PLS ou l'Enet font parties, sections 2.4.2 à 2.4.4), produisent en général des métriques non adaptées au contexte de régression. La statistique fréquemment donnée est le R^2 (section 2.2.2), qui ne renseigne que sur la qualité du modèle sur la période instrumentale. En revanche, cette mesure de qualité ne renseigne en aucun cas sur la propension du modèle linéaire à reconstruire des points qui n'ont pas été utilisés pour l'entraîner. C'est à ce moment là que l'échantillonnage apprentissage/test comme présenté section 2.4.7. devient crucial pour déterminer la qualité d'un modèle statistique de reconstruction.

Malgré ces différentes critiques faites aux études précédentes, l'étude de Ortega et al. (2015) propose une approche élaborée de manière objectivement robuste pour reconstruire un indice NAO. Ils utilisent notamment une sélection d'enregistrements paléoclimatiques par test de corrélation, mais aussi une évaluation par échantillonnage apprentissage/test comme présentée section 2.4.7. Néanmoins, bien que les découpages apprentissages/tests permettent de confirmer que leur modèle de PCR produit une reconstruction fiable, ils ne sont pas utilisés pour confronter la méthode qu'ils emploient à d'autres méthodes de régression. De plus, d'autres paramètres, choisis de manière arbitraire, pourrait être optimisés par des méthodes d'évaluation de modèles statistiques. Par exemple, le seuil de confiance de sélection des enregistrements paléoclimatiques est de 90% et n'est confronté à aucun autre critère de sélection.

Le manque d'homogénéité dans la méthodologie d'approche de reconstruction d'indice donne l'explication même de l'existence des méthodes de *machine learning* : Il existe tellement de configurations possibles pour mettre en place la reconstruction d'un même indice qu'il serait maladroit d'en sélectionner une sans l'avoir confrontée à d'autres. La méthode de régression employée ne représente qu'une composante de ces différentes possibilités de configuration. La stratégie la plus sage ici serait de considérer tout un ensemble de configurations (méthode de régression, critère de sélection d'enregistrements, période sur laquelle sont construits les modèles). Le but est alors de les évaluer et de sélectionner la meilleure, selon une métrique donnée, pour produire une reconstruction.

Dans la section qui suit, nous présenterons l'article Michel et al. (2020) publié dans le journal *Geoscientific Model Development*, qui propose l'outil de reconstruction ClimIndRec (pour *Climate Index Reconstruction*), implémenté dans les langages bash et R. Celui-ci remédie notamment aux différentes difficultés soulevées dans les paragraphes précédents. Tout d'abord, il permet à toute personne de produire la reconstruction d'un indice climatique qu'il souhaite, avec le jeu d'enregistrements qui lui convient. L'idée est de simplement remplir des fichiers d'entrées en renseignant l'indice, les enregistrements, la période de reconstruction, la méthode statistique, et plus d'une dizaine d'autres paramètres, notamment associés aux configurations de validation croisée et de découpage apprentissage/test (sections 2.4.6 et 2.4.7). L'outil produit alors une reconstruction évaluée par des ensembles de scores, calculés à

l'aide de deux métriques appliquées sur les périodes tests (non utilisées dans les modèles d'apprentissage, section 2.4.7) : la corrélation et le coefficient d'efficacité (*Coefficient of Efficiency*, CE). Le CE est la différence entre 1 et le rapport entre la somme des carrés des écarts des valeurs de l'échantillon de test avec leur prédiction, et la somme des carrés des écarts des valeurs de l'échantillon de test avec la moyenne des valeurs de l'échantillon d'apprentissage (Nash and Sutcliffe 1970). L'interprétation de cette métrique est relativement intuitive et utile : si elle est négative, utiliser la moyenne de l'échantillon d'apprentissage (une droite) est plus efficace que le modèle statistique, et inversement quand celle-ci est positive. Autrement dit, un CE négatif implique que même sans enregistrements paléoclimatiques, on peut faire mieux que le modèle statistique en donnant simplement l'état moyen de l'indice comme reconstruction pour tout pas de temps. À l'inverse, un CE positif induit que le modèle statistique a pu tirer plus d'informations des enregistrements paléoclimatiques que ce que l'on peut observer "à l'oeil nu" sur l'indice climatique. Le CE se veut être une approche plus efficace pour évaluer les modèles statistiques que la corrélation. En effet, en plus de vérifier la cohérence de phase entre l'indice reconstruit et les vraies valeurs de l'indice comme le fait la corrélation, celui-ci renseigne également sur les écart absolus entre ces derniers. Après calculs, l'outil produit toutes les informations nécessaires à l'utilisateur pour étudier le suivi de la reconstruction ainsi que les résultats obtenus. Il disposera ainsi de : la reconstruction ; les échantillons apprentissage/test tirés ; les scores calculés sur les échantillons test ; le poids des enregistrements utilisés pour la reconstruction ; les test de normalité des résidus sur les échantillons test...

3.2 Article publié dans *Geoscientific Model Development*

Title : Reconstructing climatic modes of variability from proxy records using ClimIndRec version 1.0

Simon Michel¹, Didier Swingedouw¹, Marie Chavent¹, Pablo Ortega¹, Juliette Mignot⁴, Myriam Khodri⁴

1 : Environnements et Paléoenvironnements Oceaniques et Continentaux (EPOC), UMR CNRS 5805 EPOC-OASU-Université de Bordeaux, Allée Geoffroy Saint-Hilaire, Pessac 33615, France.

2 : Institut National de la Recherche en Informatique et Automatique (INRIA), CQFD, 33400 Talence, France.

3 : BSC, Barcelona, Spain.

4 : Sorbonne Universités (UPMC, Univ. Paris 06)-CNRS-IRD-MNHN, LOCEAN Laboratory, 4 place Jussieu, F-75005 Paris, France.

Abstract :

3.2.1 Introduction

The interdependent components of the climate system, such as the atmosphere and the ocean, vary at different timescales. The interactions between those components (Mitchell et al. 1966) lead the climate to vary from the hourly to the multidecadal timescales. Preindustrial control simulations of global coupled climate models

have evidenced that such a variability is still present without any modulation of the external forcings, which is frequently referred to as internal variability (Hawkins and Sutton 2009). External factors such as volcanic aerosols (Mignot et al. 2011, Swingedouw et al. 2015, Khodri et al. 2017), anthropogenic aerosols (Evan et al. 2009, Evan et al. 2011, Booth et al. 2012), solar irradiance (Swingedouw et al. 2011, Seidenglanz et al. 2012), and greenhouse gas concentrations (Stocker et al. 2013), also influence the variations and dynamics of the climate system by altering the Earth's radiation balance. By only considering the impact of external forcings which are not due to the human activity, we can characterise the so-called natural climate variability.

An unequivocal synchronous rise in both the greenhouse gas concentration in the atmosphere and the global mean temperature has been observed in instrumental measurements (Stocker et al. 2013). However for temperatures, fluctuations around this trend from a decade to another (Kosaka and Xie 2013, Santer et al. 2014, Swingedouw et al. 2017) highlight the modulating role of natural variability at decadal to multidecadal scales. Improving our knowledge about past natural climate variability and its sources is therefore essential to better understand the potential coming changes in climate.

Physics driving the climate system induce large-scale variations, organised around recurring climate patterns with specific regional impacts and temporal properties. These variations are known as climate modes of variability. Their evolution is usually quantified by an index that can be calculated from a specific observed climate variable. These indices provide an evaluation of the corresponding climate variations and their regional impacts (Hurrell 1995, Neelin et al. 1998, Trenberth and Shea 2006). As an example, the North Atlantic Oscillation (NAO), is the leading mode of atmospheric variability in the North Atlantic basin (Hurrell et al. 2003). Generally defined as the sea level pressure (SLP) gradient between the Azores high and the Icelandic low, the NAO describes large-scale changes of winter atmospheric circulation in the northern hemisphere and controls the strength and direction of westerly winds and storm tracks across the Atlantic (Hurrell 1995). A stronger than normal SLP gradient between the two centers of action induces a northward shift of the eddy-driven jet-stream. Such large scale changes in atmospheric circulation lead to precipitation and temperature variations in various regions (North Africa, Eurasia, North America and Greenland (Casado et al. 2013). Moreover, these meteorological impacts have major influences on many ecological processes, including marine biology (Drinkwater et al. 2003) as well as terrestrial ecosystems (Myrsterud et al. 2001) This mode also affects the oceanic convection in the Labrador Sea and the Greenland-Iceland-Norwegian Seas through changes in atmospheric heat, freshwater and momentum fluxes (Dickson et al. 1996, Visbeck et al. 2003). These changes may lead in turn to modifications in the Atlantic Meridional Overturning Circulation (AMOC) which then affect the poleward heat transport and the related Sea Surface Temperatures (SST) pattern over the Atlantic (Trenberth and Fasullo 2017).

The dynamics of these modes are still not fully understood due to the relatively short duration of the instrumental records, which prevents robust statistical evaluation of their properties (e.g. spectrum, stability of teleconnections, underlying mechanisms). To partly overcome this limitation, reconstructions of climate beyond the period of direct measurements have been performed in numerous studies that

combine appropriate statistical methods and information from proxy records. Proxy records provide indirect estimates of past local or regional climate, derived from natural archives coming for instance from sediment cores, speleothems, ice cores or tree rings. According to its nature, each proxy record has a specific temporal resolution, from years to millennia, and can cover a specific period : from hundreds to millions of years. New proxy records are continuously gathered extending the available datasets and allowing paleoclimatologists to build increasingly consistent reconstructions (PAGES 2k Consortium 2013, PAGES 2k Consortium 2017).

Based on the assumption that climate modes such as the NAO affect climate conditions in different locations, some studies have used regression-based methods on temperature and drought-sensitive proxy records to reconstruct the variability of these modes over the last thousand years. Luterbacher et al. (2001) firstly proposed a partly monthly and seasonal reconstruction of the NAO that extends back to 1500 using the Principal Component Regressions (Hotelling 1957) (PCR) method. Another study have reconstructed the SLP fields in Europe further back to 1500 using a PCR approach as well (Luterbacher et al. 2002), and have found consistencies with the Luterbacher et al. (2001) NAO reconstruction. Cook et al. (2002) have also proposed a complete methodology of nested PCR using annually resolved proxy records bounding the North Atlantic to reconstruct the NAO variability further back to 1400. More recently, Ortega et al. (2015) performed a NAO reconstruction from 1073 to 1969, also based on the PCR, using 48 proxy records that were significantly correlated with the historical NAO index on their common time window. Instead of nesting reconstructions of different sizes, which can lead to inhomogeneities between time windows using different proxy selections, this study used several random calibration/validation samplings of the overlap period of the NAO index and the proxy records to perform individual reconstructions on the same time frame. Regression-based methods have also been used for reconstructing other climate modes indices than NAO, such as for instance El-Niño Southern Oscillation index (Li et al. 2013) and the Atlantic Multidecadal Variability index (Gray et al. 2004, Wang et al. 2017).

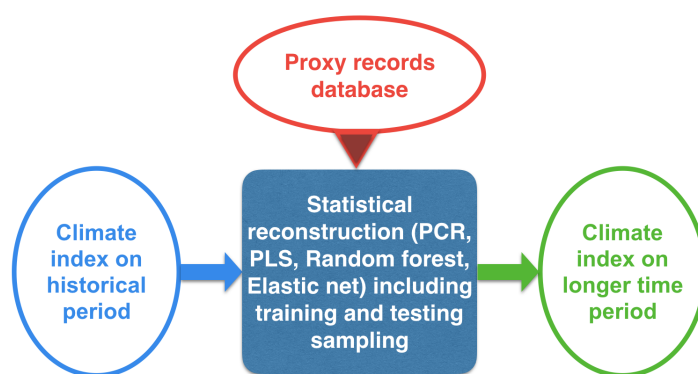


FIGURE 3.1 | Scheme summarizing the main features of ClimIndRec.

More recent algorithms than PCR provide alternative regression methods that can also be used to reconstruct climate modes, and may possibly further improve the quality and the robustness of these reconstructions. In this paper, we present the computer tool ClimIndRec (Climate Index Reconstruction) version 1.0, which includes multiple statistical approaches, for reconstructing climate modes indices. It is based on four regression methods :

the PCR (Hotelling 1957), the Partial Least Squares regression (PLS) (Wold et al. 1984), the Elastic-net regression (Enet) (Zou and Hastie 2005) and the Random Forest (RF) (Breiman 2001). It communicates with a large proxy database, that contains various types of proxy records distributed worldwide and which are sensitive to

different climate variables. ClimIndRec is thus designed to reconstruct the past variability of different climate modes (Fig. 3.1). It should be stressed that ClimIndRec will only be useful with climate indices for which there are enough proxy records representing their regional climate imprints, and that have the appropriate time resolution to capture its preferred timescale of variability. Besides the climate modes, ClimIndRec can also be used to reconstruct other kinds of climate timeseries such as temperature or precipitation in a given location.

Section 3.2.2 develops the functioning and the value-added of ClimIndRec for climate timeseries reconstruction. Section 3.2.3 compares the four regression methods by reconstructing the NAO index over the last millenium and investigates the reconstruction sensitivity to methodological choices such as the method used, the learning period or the proxy records selection for regression. Finally section 3.2.4 presents a discussion including some outlooks for next version of ClimIndRec and the conclusions of this study.

3.2.2 Data, notations and methodologies

3.2.2.1 General methodology of ClimIndRec

We here compare four models that all consist in regression methods among which the PCR has been used in many paleoclimate studies (Luterbacher et al. 2001, Luterbacher et al. 2002, Cook et al. 2002, Gray et al. 2004, Ortega et al. 2015, Wang et al. 2017). The methods we added (PLS, Enet and RF) aim at exploring alternative approaches than PCR and comparing different reconstructions using relevant metrics. PLS is a similar approach to PCR where the difference is that the matrix of Empirical Orthogonal Functions (EOFs) is calculated by maximising the variance of the projected proxies on the EOFs and the targeted climate index instead of the variance of the projected proxies (Wold et al. 1984). Enet belongs to the regularized regression methods family that are not usually used in paleoclimate reconstructions (Zou and Hastie 2005). It is here investigated in order to find if this kind of regression approach is relevant for climate index reconstruction. Finally, the RF method is an aggregation of multiple predictors called "regression trees", which are non-linear regression approaches (Breiman 2001). The mathematical details for each method are developed in supplementary material 1. Given a climate index and proxies, ClimIndRec optimizes a given regression method with cross-validation-based techniques and can thus be extrapolated to other regression-based approaches. Hence, updates of ClimIndRec will be dedicated to propose other regression methods such as adaptive lasso regression (Zou 2006).

In the case of the reconstruction of climate indices, regression methods seek to establish for each common time step the relationships between the proxies and the climate index to be reconstructed over the period of instrumental measurements. This set of relationships constitutes a statistical model of the considered climate index. The paleo-variations of proxy records are then translated into a climate index in the past using the relationships previously established by the statistical model. Since they all use unknown parameters, they must be optimized to make the reconstruction as robust as possible. In the case of PCR, for example, the number of principal components of the proxies used to regress the climate index directly affects the reconstruction since it modifies the set of predictors. The term "control parameter" is

used to design this ensemble of parameters inherent to each method. They are identified for each method in supplementary material 1. Their tuning (or optimization) using cross-validation technics (Stone 1974, Geisser 1975) are developed later in this section.

Reconstruction of a same climate index obtained from different regression methods may significantly differ. Thus, if the same index is reconstructed using different regression methods that each suggest different interpretations of the past, it may be difficult to compare them directly. A common approach is to separate the observation years (called learning period) in two to evaluate a statistical model. The first period, called the training (or calibration) period, is used to build the model using control parameter tuning, and thus to establish relationships between the climate index and proxies. The proxies of the second period, called the testing (or validation) sample, are then translated into a climate index over the years of observations of this period. The actual values of the climate index can then be compared with the reconstructed climate index over the testing period using a given metric which will be defined in section 3.2.2.3.2. It gives a score estimating the model ability to reconstruct the climate index using first-seen data of proxies. This procedure is called "hold-out approach" (McCornack 1959).

The scores obtained for different regression methods for a given training/testing sample might be impacted by the specific sampling. This is overcome by repeating the hold-out approach several times where years of observations between the training and the testing samples are shuffled. An ensemble of scores is obtained, yielding an evaluation of the methods' ability to reconstruct the climate index. The most robust regression model is the one that has the highest scores as it means this is the most accurate to reconstruct the climate index using first-seen data of proxies. This most robust regression method is then applied to the whole learning period to build a final model and infer the paleo-variations of the climate index from proxy records. In our study, and by default in ClimIndRec, the determination of the testing samples is performed using a block-style approach over time. This means that the first testing period of a given size encompasses the first time steps of the learning period. This testing period is then shifted by one time step which gives the second testing period of same size. And so on until each time step of the learning period has been used at least once for testing. The reason is that for climate timeseries, autocorrelation is often large, so that one get skills from persistence alone. Thus sampling is usually used with a block-style approach for climate timeseries.

The reconstruction might also largely differ for a same reconstruction method according to both the proxy records used and the years of observations used. Here, the sources of uncertainties associated with the proxy selection as well as the learning period used can be reduced using the same hold-out approach with evaluation and comparison of the most optimal sets using scores.

The number of proxy records and the reconstruction period are thus fixed for the different training/testing period sections and the final model, in contrast with some previous studies which used nested approaches (Cook et al. 2002, Wang et al. 2017). We make this choice because the aim of this study is mainly focused on optimizing the methodological approach for the reconstruction and not the reconstruction itself. Nevertheless, ClimIndRec can be used to perform reconstructions on different time

windows which can then be aggregated to perform a nested reconstruction, with associated scores for each portion of time.

It should be stressed that the approach of ClimIndRec implicitly assumes that the climate index to reconstruct is a linear combination of the proxy records. It means assuming that the climate reacts to proxies while the correct etiological relationship is the other way around (Tingley et al. 2012). Hence, it has to be specified that since climate variations affect proxies variations, we can attempt to estimate past climate fluctuations using statistical methods. Another caveat to highlight is that the proxy records used have their own uncertainties that can come from various sources such as the measurement methodologies, the dating uncertainties or the transfer function used to infer bio/geochemical data into paleoclimate variations. This inevitably leads to an underestimation of the true link between the climate index and the climate variable associated to the proxy record and therefore leads to a biased reconstruction with loss of variance (Isobe et al. 1990). To overcome this issue, previous climate index reconstruction studies (Ortega et al. 2015, Wang et al. 2017) rescaled the variance of the reconstruction according to the observed climate index variance. However it implies that the variance of the climate index is stationary which might not be true. In this study we thus present the raw reconstructions and the loss of variance will be quantified and specified.

ClimIndRec is developed using both bash and R scripts. It uses different R packages (presented supplementary table 5) that could be used independently to blindly perform reconstructions of any climate index. The added-value of ClimIndRec is to integrate synchronous hold-out approach and cross-validation according to the user inputs (proxy records, regression method, reconstruction period targeted, proxy records pre-selection). It therefore allows several inputs to be tested and provides relevant metrics that can be used to determine the optimal regression model.

3.2.2.2 Step-by-step procedure for reconstruction and model evaluation

The general reconstruction and model evaluation procedure follows 12 steps (Fig. 3.1), applied sequentially as follows :

1. An observational time series representing modulations of the targeted mode of variability is chosen to be used as the predictand
2. A target time period $\mathcal{T}$ for the reconstruction is selected
3. The statistical reconstruction method to be applied is selected
4. The proxy records that overlap with the selected reconstruction period are extracted to be used as predictors
5. The common period T between the observed climate index and the selected proxy records is identified and extracted for evaluating the reconstruction method
6. This common period is split in two, one for training the model (training period), and one for testing it (testing period). This is repeated R times following a block-style approach to perform splits, R depending on the size of the learning period and the size of testing periods determined by the user.
7. The proxy records that have a significant correlation at a given threshold with the climate index over the training period are selected to train the statistical model

8. Each of the R sets of periods and proxies is calibrated over the training window for all the different sets of control parameters of the given method selected in 3, and the best performing one is identified
9. The corresponding optimal setup is then applied to extend the reconstruction over the testing period for each member
10. Validation scores are computed by comparing each of the observation-based testing series and each training sample-based individual reconstruction over the corresponding testing period
11. The corresponding control parameters are tuned over the whole learning period T and the final model is built
12. The final reconstruction is obtained by applying the final model to the proxies over the reconstruction period $\mathcal{T}$

Thus ClimIndRec provides the final reconstruction with associated uncertainties (supplementary material 3) and a vector with of R validation scores following different metrics as final outputs.

3.2.2.3 Model evaluation and optimization

This section aims to clarify the technical details of the methodology presented in sections 2.1 and 2.2. It will thus call on the elements mentioned above.

3.2.2.3.1 Data notation

To simplify the mathematical notation, we make the assumption that the proxy record selection and truncation to their common time window with the climate index have already been made (see section 3.2.2.2, steps 4 and 5). In this study, it is important that all proxy records are truncated on the same time window to make them mergeable in the same matrix. Each record has to cover at least the chosen reconstruction time window $\mathcal{T}$ and it is excluded otherwise (section 3.2.2.2, step 2). Hence, the proxy records matrix does not contain missing values.

Fig. 3.2 illustrates how the proxy data are organised in the input matrix X . We denote $X^1 = (X_t^1)_{t \in \mathcal{T}}, \dots, X^p = (X_t^p)_{t \in \mathcal{T}}$, where t stands for the time (with N annual time steps), and p is the number of proxy records on the same period $\mathcal{T}$. X is thus a $N \times p$ matrix grouping the individual records : $X = [X^1, \dots, X^p]$. $Y = (Y_t)_{t \in T}$ is the target climate index, defined on the historical time window T called the learning period, that

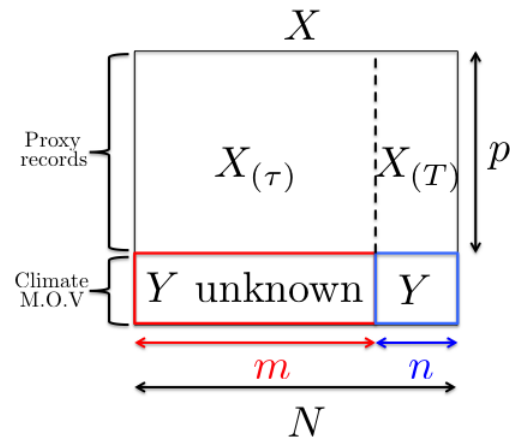


FIGURE 3.2 | Scheme of the initial data. X and Y are respectively the proxy records matrix and the index of the considered mode of variability. N is the size of the common period of all proxy records. n is the size of the common period of all proxy records and the index of the mode of variability. m is the size of the common period of all proxy records, where the mode of variability is not known. p is the number of proxy records. $X_{(T)}$ is the sub-matrix of X where the mode of variability is known. $X_{(\tau)}$ is the sub-matrix of X where the mode of variability is not known.

contains n annual time steps. The period where Y is not known is denoted τ , containing m annual time steps (Fig. 3.2). Thus $\mathcal{T} = T \cup \tau$ is the entire reconstruction period, which contains $N = n + m$ annual time steps. With these notations, the dimensions of the different matrices and vectors are : $X \in \mathbb{R}^{N \times p}$, $X_{(T)} \in \mathbb{R}^{n \times p}$, $X_{(\tau)} \in \mathbb{R}^{m \times p}$, $Y \in \mathbb{R}^n$. The learning set is denoted $\{X_{(T)}, Y\}$, and the reconstruction set is denoted $\{X_{(\tau)}\}$.

3.2.2.3.2 Terms and validation metrics

We denote the chosen reconstruction method by $\mathcal{M}$. Each method is defined by a specific number of control parameters q , contained in the vector denoted θ . We can denote the function $\mathcal{M}$ as a function of : (i) a set on which the model is built ($\{X, Y\}$), (ii) observations of the predictors on the reconstruction period ($X_{(rec)}$), and (iii) an control parameters vector (θ) :

$$\mathcal{M} : (\{X, Y\}, X_{(rec)}, \theta) \rightarrow \hat{Y}_\theta \quad (3.1)$$

$$(\{\mathbb{R}^{n \times p}, \mathbb{R}^n\}, \mathbb{R}^{m \times p}, \mathbb{R}^s) \rightarrow \mathbb{R}^m \quad n, p, m, s \in \mathbb{N} \text{ (not fixed)} \quad (3.2)$$

Hence, the $\mathcal{M}$ function gives an entire reconstruction of size $m \in \mathbb{N}$, depending on θ .

We introduce S as the score function, or validation metric. This function is an indicator that estimates the quality of a reconstruction $\hat{Y}$ with respect to the observed values $Y_{(obs)}$:

$$S : (Y_{(obs)}, \hat{Y}) \rightarrow s \quad (3.3)$$

$$(\mathbb{R}^m, \mathbb{R}^m) \rightarrow \mathbb{R} \quad (3.4)$$

In this paper, three kind of validation metrics are used for different tasks. The first is a correlation function, the second is a root mean squared error (RMSE) function and the third is a Nash-Sutcliffe coefficient of efficiency (Nash and Sutcliffe 1970) :

$$S_{cor}(Y_{(obs)}, \hat{Y}) = Cor(Y_{(obs)}, \hat{Y}) \quad (3.5)$$

$$S_{RMSE}(Y_{(obs)}, \hat{Y}) = \|Y_{(obs)} - \hat{Y}\| = \sqrt{\sum_{i=1}^m (Y_{i(obs)} - \hat{Y}_i)^2} \quad (3.6)$$

$$S_{CE}(Y_{(obs)}, \hat{Y}) = 1 - \frac{\sum_{i=1}^m (Y_{i(obs)} - \hat{Y}_i)^2}{\sum_{i=1}^m (Y_{i(obs)} - \bar{Y}_{(obs)})^2}, \text{ with } \bar{Y}_{(obs)} = \frac{1}{m} \sum_{i=1}^m Y_{i(obs)} \quad (3.7)$$

S_{CE} is used to validate the reconstruction methods over the testing period, and S_{RMSE} allows to determine the optimal control parameters (θ) for the reconstruction. We use S_{cor} because it is used in the last NAO reconstruction of Ortega et al (2015), with which we will compare our results. S_{CE} is a metric defined between $-\infty$ and 1, values lower than 0 mean that using the mean over the training period is better than the proposed statistical model (Nash and Sutcliffe 1970), additional informations about this metric are presented in the supplementary material 2. Here, we will consider that a final reconstruction is robust and reliable when its R CE scores are

significantly positive at the 99% confidence level using a Student test. As the possible values of the CE score is not symmetric around 0, the best reconstruction is identified as the one that has the higher median of CE scores.

3.2.2.3.3 Control parameter tuning by cross-validation and final reconstruction

As mentioned above, the initial learning sample is split into R partitions of two subsets : $\{T_{(train)}^{(r)}, T_{(test)}^{(r)}\}$, $\forall 1 \leq r \leq R$ (section 3.2.2.2, step 6). For a given method $\mathcal{M}$, R reconstructions are build on the R training samples. $\forall 1 \leq r \leq R$, we denote $\{X_{(train)}^{(r)}, Y_{(train)}^{(r)}\}$ the training set, and $\{X_{(test)}^{(r)}, Y_{(test)}^{(r)}\}$ the test set. At each step, the columns of X , $X_{(train)}$ and X_{test} are normalized to the mean and the standard deviation of the respective columns of $X_{(train)}$.

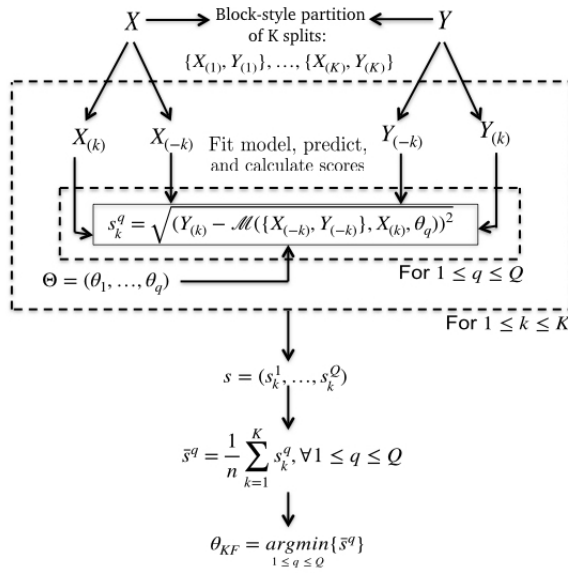


FIGURE 3.3 | Scheme of a K-fold cross-validation procedure to select the optimal control parameter of a specific learning method $\mathcal{M}$. X is the input set of predictors and Y the corresponding variability mode index. $\forall 1 \leq k \leq K$, $(X_{(k)}, Y_{(k)})$ is the k^{th} block-style-based group of observation and $(X_{(-k)}, Y_{(-k)})$ contains all observations except the k^{th} group. $\Theta = (\theta_1, \dots, \theta_s)$ is the ensemble of possible values of the s control parameters $\theta \in \mathbb{R}^s$.

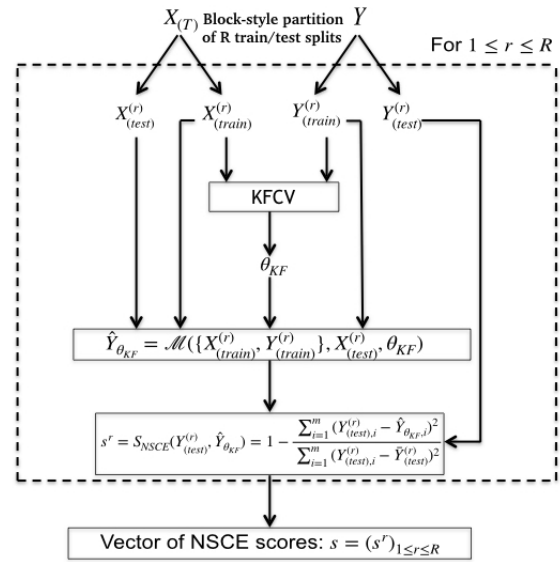


FIGURE 3.4 | Scheme of the whole procedure for scores calculation for a given method $\mathcal{M}$. Y is the index of the chosen mode of variability. $X_{(T)}$ is the proxy dataset restricted to the period where Y is known. $\{X_{(train)}^{(r)}, Y_{(train)}^{(r)}\}$ is the r^{th} training sample and $\{X_{(test)}^{(r)}, Y_{(test)}^{(r)}\}$ is the r^{th} testing sample. θ_{KF} is the empirically optimal set of parameters obtained by applying the KFCV (Fig 3.3, section 3.2.3.3.3).

To estimate the optimal set of control parameters θ_{opt} on a given training set $\{X_{train}, Y_{train}\}$, we use the K-fold cross validation approach (KFCV, section 3.2.2.2, step 8 and 9) (Stone 1974, Geisser 1975). Cross Validation (CV) methods, are in general, widely used as parametrization and model validation techniques (Kohavi 1995, Browne 2000, Homrighausen and McDonald 2014, Zhang and Yang 2015). Here, it is used as an optimization method to empirically determine an optimal set of control parameters for θ . As presented in Fig. 3.3, the KFCV splits the observations into a partition of n groups of same sizes (or approximately same sizes if the length of the training set is not divisible by K). $\forall 1 \leq k \leq K$, we denote $\{X_{(k)}, Y_{(k)}\}$, containing

only information for the k^{th} drawn sample. Then, $\{X_{(-k)}, Y_{(-k)}\}$ is the set containing all the $K-1$ other sets. For all possible values of θ contained in Θ , we scan the K models based on the sets $s \{X_{(-k)}, Y_{(-k)}\}_{1 \leq k \leq K}$. The empirical optimal set of control parameters is obtained by minimizing the averaged S_{RMSE} functions on the K splits by considering all possible combinations of θ (Stone 1974). Mathematically, the optimal KFCV set of control parameters θ_{KF} is determined by :

$$\theta_{KF} = \arg \min_{\theta \in \Theta} \frac{1}{K} \sum_{k=1}^K S_{RMSE}(Y_{(k)}, \mathcal{M}(\{X_{(-k)}, Y_{(-k)}\}, X_{(k)}, \theta)) \quad (3.8)$$

It should be noted that if $\dim(\theta) > 1$, then the different control parameters need to be optimized simultaneously, with nested KFCVs.

Using this approach, we retain the control parameters vector $\hat{\theta}_{opt} = \theta_{KF}$ for the given method $\mathcal{M}$ and a given learning set $\{X, Y\}$. KFCV is applied to build a unique optimized reconstruction for every training sets and any given method. Then, for all the corresponding and independent testing periods, the associated testing series $Y_{(test)}^{(r)}$ are compared to the individual reconstructions using the S_{CE} function. This way, R CE scores are obtained for $\mathcal{M}$. In section 3.2.3, the distributions of the CE scores will be used as a metric to compare different reconstructions. Fig. 3.4 shows the the calculations that gives the CE scores for a given method $\mathcal{M}$.

It should be stressed that the K -fold cross validation sampling is also implemented following a block-style approach in ClimIndRec for the same reasons than for the hold-out approach. This means that the K groups of observations are constructed along time instead of being randomly split. Also, the choice of K can have implications for the estimation of optimal control parameters. Large K leads to more diverse training samples, thereby bringing more variable estimates of RMSE. On the other hand, small K leads to a low number of samples used, thereby increasing the bias due to the particular way splits have been made. Additional works have shown that this choice is poorly influencing the final reconstruction obtained (not shown) so that we decided to set it to $K=5$ for this study. It is set at $K=5$ by default in ClimIndRec but it can certainly be changed in order to produce alternative reconstructions.

Once the model has been evaluated, it is launched over the whole learning set $\{X_{(T)}, Y\}$ with a K -fold cross-validation to optimize the control parameters such as done previously for training samples.

3.2.2.3.4. Data

The assessment of the proposed reconstruction techniques is investigated for the NAO index, as it is probably the mode of variability that has been observed for the longest time period. This index is indeed relatively simple to calculate from SLP timeseries as it only requires two locations with instrumental records : one within the center of action of the Azores anticyclone (typically Gibraltar) and one within the Icelandic low (typically Reykjavik). The reference NAO index is then calculated as the normalized SLP difference between these two locations. We here use the Jones et al. (1997) index spanning the whole historical period since 1856.

In terms of proxies, we use the state-of-the-art PAGES 2k database (PAGES 2k Consortium 2017) in its latest 2017 version (hereafter P2k2017). Proxy records with resolutions lower than annual were removed. Even if these proxy records could be interpolated to a finer temporal scale and used for the reconstruction, their use is not recommended as the interpolated timeseries will present high auto-correlation coefficients, which could inflate the correlations with the NAO and thus their weight in the final reconstruction, potentially leading to spurious results (Hanhijarvi et al. 2013). We added 44 annually-resolved proxy records used in Ortega et al. (2015) and not present in P2k2017 (see supplementary table 1). We end up with a database of 554 well-verified and worldwide distributed annually-resolved proxy records.

3.2.3 Results

3.2.3.1 Methodological sources of uncertainty in the reconstruction

We apply ClimIndRec with the four methods presented above to the reconstruction of the NAO. In the following, each reconstruction is obtained by averaging R individual reconstructions performed for R training/testing splits. R depends on the size of the testing samples relative to the size of the learning period as we perform block-style splits of the data to produce training and testing samples (section 3.2.2.1 and 3.2.2.2). Here, we set the relative length of the training splits as 80% of the learning period. CE scores are thus produced and stored in a vector of R elements. This vector will thus be used as a quality metric to characterize the methodological uncertainty in the reconstruction. The following actions were undertaken to minimize the reconstruction uncertainty identified in section 3.2.2.1, and estimate its sensitivity :

1. Pre-selecting the most relevant proxy records
2. Selecting the best learning period

These two steps are described below, before assessing the reconstruction itself.

3.2.3.1.1 Proxy pre-selection

Among the previous climate reconstruction studies, Ortega et al. (2015) have performed a proxy selection over the training periods at the 90% confidence level using the correlation test from McCarthy et al. (2015) while Cook et al. (2002) and Wang et al. (2017) have selected their proxies by focusing on the regions affected by the modes they respectively reconstructed. Here we run 4 reconstructions of $R = 50$ individual members for each method. These reconstructions are respectively performed with different significance levels for the proxy selection by correlation over the training periods. These levels are 0% (which means that all the records are used at each training/testing split), 80%, 90% and 95%. The reconstructions are performed for the reconstruction period $\mathcal{T} = \llbracket 1000, 1970 \rrbracket$ and the learning period $\mathcal{T} = \llbracket 1856, 1970 \rrbracket$ encompassing 110 available proxy records with $n = 115$.

Fig. 3.5 shows that RF method, particularly useful for larger datasets, is more efficient using the proxy records correlated at the 80% confidence level with $med(S_{CE}) = 0.15$ (med is the median function), even if using proxy records uncorrelated with the NAO or not located in regions affected by NAO variations. On the other hand, the 3 other regression methods are more adapted when the finest proxy selection (95%) is applied, as highlighted by Ortega et al. (2015) for the PCR. Fig. 3.5 is also evidencing that the widely used PCR method and PLS have to be employed cautiously with a statistically-based proxy selection over the training periods in further studies. Indeed the reconstructions performed with the widely used PCR method is never significantly robust at the 99% confidence level, whatever the pre-selection of proxies. On the opposite, for RF and Enet methods, the proxy selection is not affecting the statistical robustness of the reconstruction, with reconstructions significantly robust at the 99% confidence level for every choice of proxy selection.

Overall, RF gives the best CE scores. Nevertheless, we stress that these results have been obtained for a particular learning period (1856-1970). The sensitivity to this is assessed in the next section.

3.2.3.1.2 Sensitivity to the learning period

In this section, we keep for each method the optimal selection of proxy records over the training periods (see section 3.1.1). We explore the impact of the reconstruction period. This affects the final reconstruction in two different ways, both related

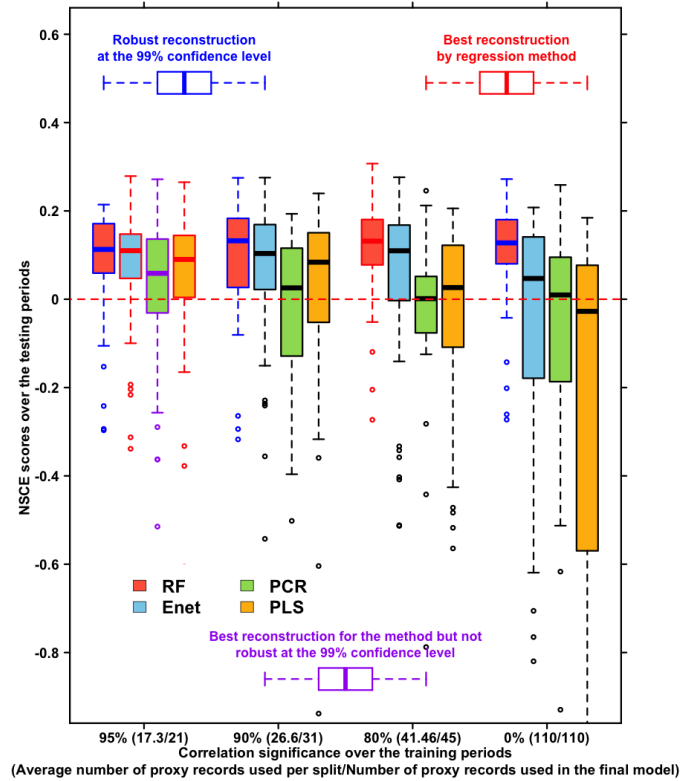


FIGURE 3.5 | Boxplot of CE scores obtained for the four methods and different groups of proxy records by reconstructing the NAO index of the period 1000–1970 with $R = 50$ training/testing samples. Green boxplots are the CE scores obtained for the PCR method. Yellow boxplots are the CE scores obtained for the PLS method. Red boxplots are the CE scores obtained for the RF method. Blue boxplots are the CE scores obtained for the Enet method. The first cluster of boxplots is the CE scores obtained by using all the available proxy records over the period (110 proxy records). The second cluster of boxplots is the CE scores obtained by using only proxy records significantly correlated with the NAO index at the 80% confidence level over the training periods. The third cluster of boxplots is the CE scores obtained by using only proxy records significantly correlated with the NAO index at the 90% confidence level over the training periods. The fourth cluster of boxplots is the CE scores obtained by using only proxy records significantly correlated with the NAO index at the 95% confidence level over the training periods. Boxplots with blue edges are the scores significantly positive at the 99% confidence level. Boxplots with red edges correspond to the scores associated with the best reconstruction for each method.

to the final proxy selection, as explained in section 2.1.

We run the reconstruction for 31 periods $\mathcal{T}$: from 1000-1970 to 1000-2000, with an increment of one year. By doing so, the number of available proxy records is not the same for each of the periods (see Fig. 3.6). Fig. 3.6.a shows the CE scores obtained for the different reconstruction/learning periods. Using the CE metric, we find that the best reconstruction time window is 1000-1972 for PLS and RF methods and 1000-1971 for Enet and PCR methods. Following the optimal setup for each method from section 3.1.1, RF uses 47 records and the three others uses 21 records. (Fig. 3.6b). Among these four optimized reconstructions which are the final ones of this study, the RF gives the highest CE scores with $med(S_{CE}) \simeq 0.16$ and $S_{CE} \in [-0.4, 0.4]$ (Fig. 3.6a).

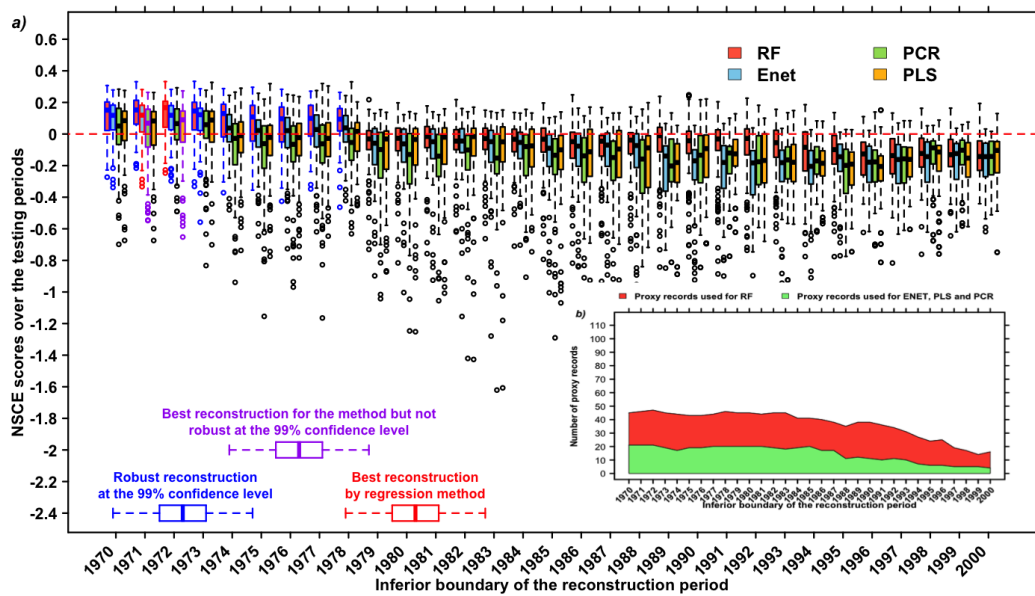


FIGURE 3.6 | Boxplot of CE scores for different reconstruction periods. a) Reconstructions are performed using 31 reconstruction period for the four methods : from 1000-1970 to 1000-2000 by moving the superior born by 1 with $R = 50$ training/testing samples. RF reconstructions are performed using the proxy records significantly correlated at the 80% confidence level with the NAO over the training periods (section 3.2.3.1.1). PCR reconstructions are performed by selecting the proxy records significantly correlated at the 95% confidence level with the NAO over the training periods (section 3.2.3.1.1). PLS and Enet reconstructions are performed using by selecting the proxy records significantly correlated at the 95% confidence level with the NAO over the training periods (section 3.2.3.1.1). Red boxplots are the CE scores obtained using RF method. Blue boxplots are the CE scores obtained using Enet method. Red green are the CE scores obtained using PCR method. Yellow boxplots are the CE scores obtained using PLS method. Boxplots with blue edges are the scores significantly positive at the 99% confidence level. Boxplots with red edges correspond to the scores associated with the best reconstruction for each method. b) Proxy records available/used by reconstruction period. Red area gives : Number of records used for. Green area : Number of records used for Enet, PCR and PLS for each reconstruction period.

Results show that the four methods are strongly affected by the choice of the reconstruction period. Thus, we recommend to determine this period carefully with different simulations on different time windows, following the approach we presented here, easily performable using ClimIndRec. Overall, this study shows that for each optimisation, PCR and PLS are less reliable to reconstruct the NAO than RF and Enet (section 3.2.3.1.1 and this section).

3.2.3.2 Scientific results

We compare and investigate the reconstruction with highest scores for each method following section 3.1. The four optimized reconstructions are obtained by using the full set of proxy records for RF and only using the proxy records significantly correlated at the 95% confidence level with the NAO index over the learning period for the other methods (see section 3.1.1). RF and Enet reconstructions are performed for the period 1000-1972 while PCR and PLS reconstructions are performed for the period 1000-1970 (section 3.1.2).

3.2.3.2.1 Comparison with previous work

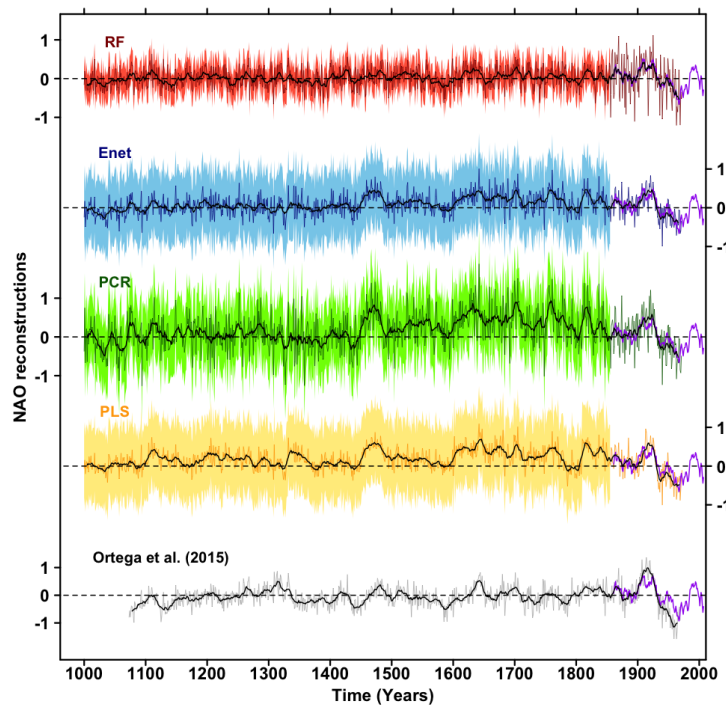


FIGURE 3.7 | Comparison of different NAO reconstructions from this study and Ortega et al. 2015. Red line : RF reconstruction on the period 1000-1972 (section 3.2.3.1.2), using proxy records significantly correlated at the 80% confidence level with the NAO over the training periods (section 3.2.3.1.1). Blue line : Enet reconstruction on the period 1000-1971 (section 3.2.3.1.2) by selecting the proxy records significantly correlated with the NAO index at the 95% confidence level over the training periods (section 3.2.3.1.1). Green line : PCR reconstruction on the period 1000-1971 (section 3.2.3.1.2) by selecting the proxy records significantly correlated with the NAO index at the 95% confidence level over the training periods (section 3.2.3.1.1). Orange line : PLS reconstruction on the period 1000-1972 (section 3.2.3.1.2) by selecting the proxy records significantly correlated with the NAO index at the 95% confidence level over the training periods (section 3.2.3.1.1). Black line (tiny) : Ortega et al. (2015) calibration constrained reconstruction (Ortega et al. 2015). Red area : Regression uncertainties (see supplementary material 3) for RF reconstruction. Blue area : Regression uncertainties for Enet reconstruction. Blue area : Regression uncertainties for PCR reconstruction. Orange area : Regression uncertainties for PLS reconstruction. Heavy black lines are the corresponding 11-year filtered reconstructions for each method. Purple lines : superposed 11-years filtered Jones et al. (1997) NAO index.

Fig. 3.7 shows the different reconstructions of the NAO, including the Ortega et al. (2015) calibration constrained reconstruction (only proxy-based), and Tab. 3.1 exhibits the paired correlations between the 5 reconstructions. The regression uncertainties (see supplementary material 3) are also shown for the four reconstructions

	RF	Enet	PCR	PLS	Ortega
RF	1.00	0.7	0.65	0.54	0.55
Enet	0.7	1.00	0.92	0.88	0.65
PCR	0.65	0.92	1.00	0.8	0.48
PLS	0.54	0.88	0.8	1.00	0.68
Ortega	0.55	0.65	0.48	0.68	1.00

TABLE 3.1 | Table of correlations between five NAO reconstructions. Includes Ortega et al. (2015) reconstruction; RF reconstruction on the period 1000-1972 using the proxy records significantly correlated with the NAO at the 80% confidence level; Enet reconstruction on the period 1000-1972 only using the proxy records significantly correlated with the 95% confidence level; PCR reconstruction on the period 1000-1970 only using the proxy records significantly correlated with the NAO at the 95% confidence level; PLS reconstruction on the period 1000-1970 only using the proxy records significantly correlated with the NAO at the 95% confidence level.

of this study on Fig. 3.7. The normality of the residuals for the four methods has been verified for both the models built over the training samples and the final model as demonstrated in Fig. 3.8. Tab. 3.1 and Fig. 3.7 shows that the NAO reconstruction based on RF is distinguishable from the four others including Ortega et al. (2015). Indeed its correlation with the other indices ranges between 0.49 and 0.67 (Tab. 3.1) while the paired correlations obtained between the others are greater than 0.88. Additionally Fig. 3.9 shows that the RF reconstruction has a higher correlation with the Jones et al. (1997) NAO index than the other indices : $r=0.98$ ($p<0.01$), while Ortega et al. (2015) reconstruction has a correlation of 0.45 ($p<0.01$).

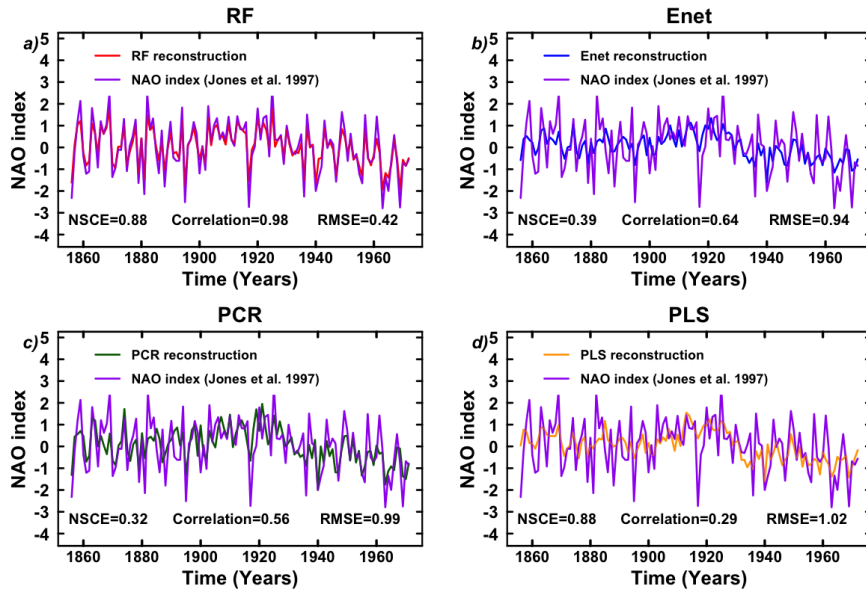


FIGURE 3.8 | Comparison of reconstructions from this study with the original Jones et al. (1997) NAO index (Purple line) over their common period. a) RF reconstruction. b) Enet reconstruction. c) PCR reconstruction. d) PLS reconstruction. CE, RMSE and correlation statistics are provided.

The RF reconstruction that uses 46 proxy records (22 common proxies with Ortega et al. 2015) presented in Fig. 3.10, has the best CE scores ($med(S_{CE}) = 0.16$, $S_{CE} \in [-0.24, 0.33]$) (section 3.2.3.1.1) and its correlation scores ($med(S_{cor}) \simeq 0.43$, $S_{cor} \in [0.06, 0.63]$) are significantly higher at the 99% confidence level than Ortega et al. (2015) calibration constrained reconstruction ($S_{cor} \in [-0.14, 0.58]$, $med(S_{cor}) \simeq 0.24$) and model constrained reconstruction ($S_{cor} \in [0.14, 0.64]$, $med(S_{cor}) \simeq 0.43$). We thus

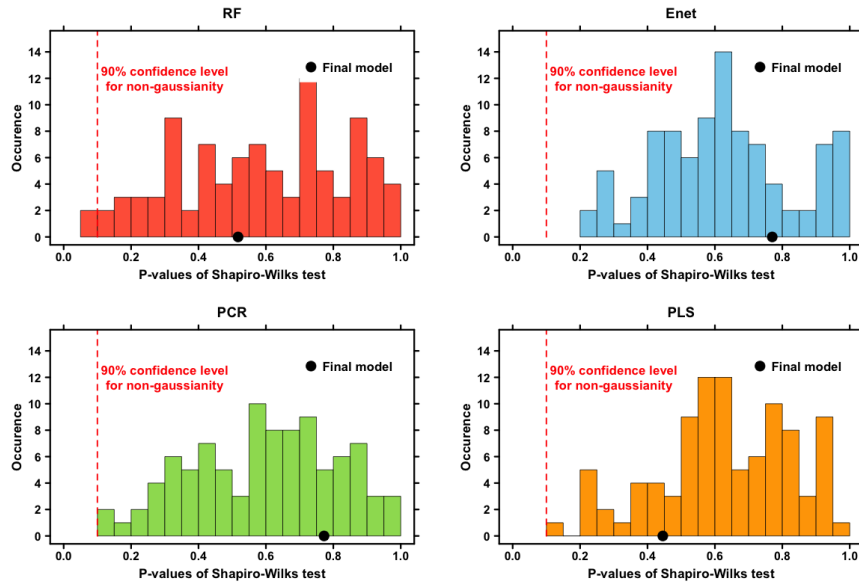


FIGURE 3.9 | P-values obtained from Shapiro-Wilk normality tests on the residuals from each reconstruction of Fig 3.7. For a), b) c) and d), the repartition of the 50 p-values obtained for each training/testing split are presented. Red dashed lines indicates the 90% confidence level for non-normality. For $0 \leq \alpha \leq 1$, if $p\text{-value} \leq \alpha$, it means that the residuals distributions is significantly not gaussian at the $1 - \alpha$ confidence level (see shapiro.test R documentation). Black dots indicates the p-values of the residuals obtained for the final models.

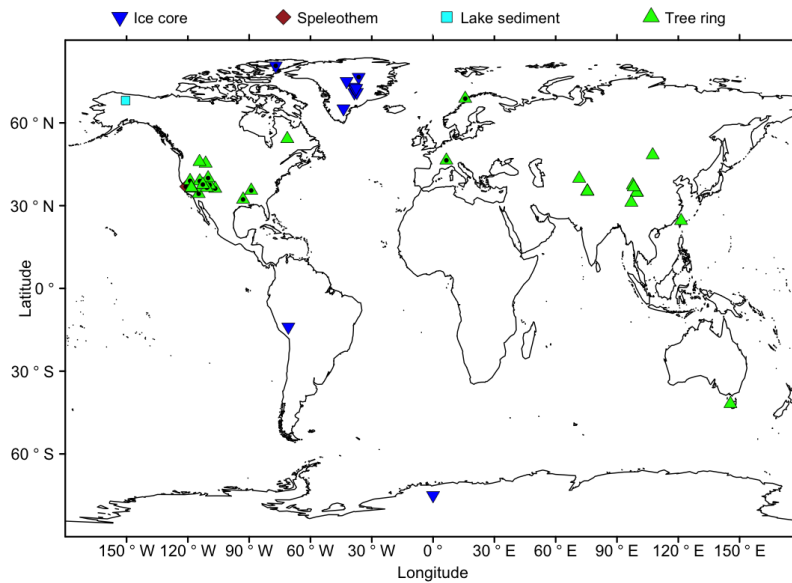


FIGURE 3.10 | Map of the 46 proxy records used for the reconstruction of the NAO index from Jones et al. (1997) on the time window 1000-1972 using the RF method. Points with a black dot are the proxy records also used in Ortega et al. (2015).

statistically verified that the best reconstruction from this study is more robust and reliable than those from Ortega et al. (2015). This improvement in performance may arise from the inclusion of new relevant proxy records into the reconstruction, but also from the use of a new statistical regression methods for climate index reconstruction : the RF. Finally, it has to be stressed that the 5 reconstructions presented in Fig. 3.7, including Ortega et al. (2015), do not show a predominant positive NAO phase during the Medieval Climate Anomaly, contrary to the hypothesis formulated

by Trouet et al. (2009).

3.2.3.2.2 Response to external forcing

No significant correlation is found between the NAO reconstruction based on RF method and the Total Solar Irradiance (TSI) reconstruction from Vieira et al. (2011) ($r \simeq -0.11$, $p > 0.18$). The same is true for the best reconstruction of the other methods (not shown) and Ortega et al. (2015). None of the reconstructions (including Ortega et al. (2015) shows clear negative phases during the Maunder and the Spörer minima as suggested by some model simulations (Shindell et al. 2004). In addition, no significant correlation on the pre-industrial era has been found with the CO₂ reconstruction based on a Law Dome (East Antarctica) ice core (Etheridge et al. 1996) (not shown), indicating that the NAO is not linearly associated with CO₂ variations over this time frame.

Ortega et al. (2015) suggested that a positive NAO phase is triggered two years after strong volcanic eruptions, a response that is not reproduced over the last millennium by model simulations (Swingedouw et al. 2017). We use the 10 large volcanic eruptions selected in Ortega et al. (2015) and a second selection (see supplementary table 2) of the 11 largest volcanic eruptions from the well-verified reconstruction of Sigl et al. (2015). By using a superposed epoch analysis and Rao et al. (2019) Monte-Carlo approach to calculate significance (see supplementary material 4), we find that using the same set of eruptions than Ortega et al. (2015) leads to the same result : a significant positive response of the NAO two years after the eruption. However, for RF this result is not significant with its p-value just under 0.1 (Fig. 3.11). On the opposite, by using the Sigl et al. (2015) 11 largest volcanic eruptions, we find a significant response at the 90% confidence level for PLS, but one year after the eruption with a p-value under 0.05 (Fig. 3.11). For RF, Enet and PCR, the positive NAO response is significant 1 to 3 years after the eruption (Fig. 3.11). Here again, the significance for the RF composite is smaller than for the other methods while this reconstruction is associated with the highest CE scores. Individual response analysis shows that for the RF reconstruction, this result is particularly significant for the 2 largest eruptions of the millennium (Samalas, 1257 and Kuwae, 1458) and not so clear for the 9 others (not shown). This result suggests that the positive NAO response might be mainly associated to volcanic eruptions with very large and rare intensities such as Samalas or Kuwae eruptions and concerns less eruptions with weaker intensities. However, further studies might be useful to verify the statistical robustness of this result, as this kind of event (eruption at least as strong as Kuwae, 1453) is very rare, thus only providing two events for this study.

3.2.4 Discussion and conclusion

3.2.4.1 Discussion, caveats and outlooks

The results presented above regarding the NAO have all been obtained using ClimIndRec. Indeed, they require advanced programming and statistical knowledge to ensure a good estimation of the reliability of the reconstruction performed. This is possible because ClimIndRec offers an integrated package through which parameters and methods can be efficiently tested and compared, together with reliable validation metrics such as the CE. Nevertheless, the methodology proposed in ClimIndRec could be further improved in different ways.

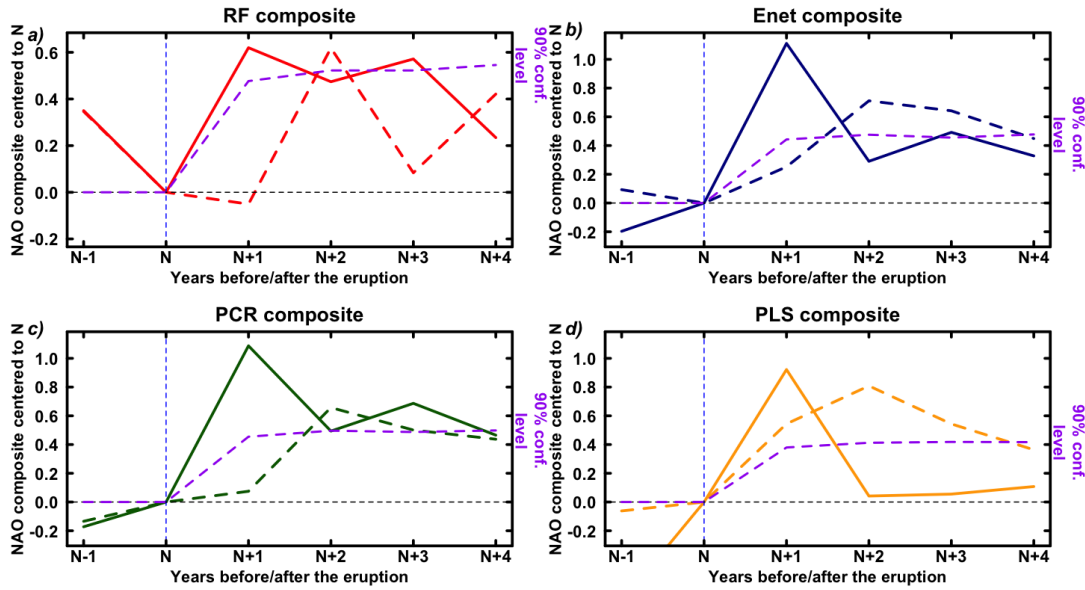


FIGURE 3.11 | Superposed epoch analysis of the NAO response from two years (N-1) before to five years after (N+4) to the largest volcanic eruptions used by Ortega et al. (2015) (10 eruptions) and the 11 largest from Sigl et al. (2015).. All of the composites are centred to their values at the year of the volcanic eruption occurrence. a) Red line : Composite for RF reconstruction response to Sigl et al. (2015) volcanic eruptions. Dashed red line : Composite for RF reconstruction response to Ortega et al. (2015) volcanic eruptions. Dashed purple line : Monte-Carlo 90% confidence level (Rao et al., 2019, supplementary material 4). b) Blue line : Composite for Enet reconstruction response Sigl et al. (2015) volcanic eruptions. Dashed blue line : Composite for Enet reconstruction response to Ortega et al. (2015) volcanic eruptions. Dashed purple line : Monte-Carlo 90% confidence level (Rao et al. 2019, supplementary material 4). c) Green line : Composite for PCR reconstruction response Sigl et al. (2015) volcanic eruptions. Dashed green line : Composite for PCR reconstruction response to Ortega et al. (2015) volcanic eruptions. Dashed purple line : Monte-Carlo 90% confidence level (Rao et al. 2019, supplementary material 4). d) Orange line : Composite for PLS reconstruction response Sigl et al. (2015) volcanic eruptions. Dashed orange line : Composite for PLS reconstruction response to Ortega et al. (2015) volcanic eruptions. Dashed purple line : Monte-Carlo 90% confidence level (Rao et al. 2019, supplementary material 4).

ClimIndRec does not deal with missing data in proxy records. This implies selecting exclusively the proxy records that entirely cover the reconstruction period, which thus excludes some existing proxy records. Also, proxy records with gaps are not used in the present version of ClimIndRec as their use in an interpolated version would artificially increase their weight in the reconstruction and thus possibly induce spectral artefacts in the reconstruction (Hanhijarvi et al. 2013). The most optimal way to develop a statistical model over the instrumental period is to use as many proxies as possible and as many years of observations as possible. This leads to a paradox since periods that are well covered by observation data are the most recent ones, which are generally less well covered by proxies. However, future versions of ClimIndRec will be dedicated to develop other probabilistic-based reconstruction approaches to deal with missing data such as Bayesian Hierarchical Models (Tingley and Huybers 2010a, Tingley and Huybers 2010b, Tingley 2012, Tingley and Huybers 2013, Cahill et al. 2016) or regularized Expectation-Maximization algorithms (Schneider 2001, Mann et al. 2008, Guillot et al. 2015). Another point that is limiting the capacities of ClimIndRec is that it is based on the assumption that teleconnections of the reconstructed mode are stationary in time, while they may depend on the state of the climate system. This is a classical limit for statistical climate reconstructions but it can be evaluated by use of pseudo-proxy methods (e.g.

Lehner et al. 2012, Ortega et al. 2015). On this aspect, more complex methods like data assimilation can clearly overcome this weakness by combining model and data. The use of such approaches for last millennium remains nevertheless very complex primarily because of their computational cost and the lack of data. They are however emerging (e.g. Hakim et al. 2016, Singh et al. 2018b). Data assimilation techniques can be very model dependent as highlighted for the ocean over the recent period (Karspeck et al. 2015) so that their reconstruction of a given regional climatic modes can suffer from interferences with reconstructions of other aspect of the climate. Thus, dedicated approaches like the ones developed here can be seen as very complementary approach and may increase our confidence in the reconstructions. Indeed, if different approaches provide very similar results, this can be interpreted as a source of robustness for a given result or reconstruction.

Another caveat concerns the fact that the present version of ClimIndRec does not account for dating uncertainties in proxy records. Future developments of ClimIndRec may allow to take into account these uncertainties and to provide their estimation along time. For doing so, deeper investigations for each proxy record are needed as these sources of uncertainty are not exhaustively provided in P2k2017. Also, we found that the reconstructions performed by ClimIndRec provide a clear loss of variance over the learning period and the reconstructed period (before 1856) (see supplementary table 4). The RF method is the only one that reproduces adequately the NAO amplitude only over the learning period but also provide a significant loss of variance over the reconstructed period. This indicates that the loss of variance over the reconstruction period could partly be due to the proxy records themselves and not only to the statistical approach.

A key aspect that has been found within this study is the sensitivity of the results to the validation metric used. Indeed, we also used correlation as the main score for the test period. It appears that this metric was mainly capturing the phasing of the modes in their reconstruction (not shown) (Wang et al. 2014). By using CE, we improved the strength of our reconstruction since other aspects than the synchronisation were accounted for. This latter metric, which is more classical in prediction evaluation further highlights that the RF method outperforms most of the others methods, and notably the PCR which is a classical method used in paleoclimatology (Cook et al. 2002, Gray et al. 2004, Ortega et al. 2015, Wang et al. 2017). Other metrics of prediction validation exist (e.g. Continuous Ranked Probability Score, Gneiting and Raftery 2007) so that a more extensive analysis of the sensitivity of the reconstruction to other metrics for the validation period might be very useful. Thus, the development of other validation metrics in next versions of ClimIndRec appears as an interesting avenue to explore.

3.2.4.2 Conclusions

We have proposed and described here four statistical methods for reconstructing modes of climate variability and have compared them for a particular example : the reconstruction of the NAO. By identifying and minimizing the sources of reconstruction uncertainty due to the method used (section 3.2.3.1.1, 3.2.3.1.2), the time frame considered (section 3.2.3.1.2) and the proxy selection (sections 3.2.3.1.1), we found the optimal NAO reconstructions. It was obtained for the RF method over the

time frame 1000-1972 using the 46 proxy records available on this time frame (section 3.2.3.2.1). This method has not been used yet to our knowledge for climate index reconstructions while it clearly outperforms the other methods (section 3.2.3.1) and seems thus promising. The reconstruction we obtained is distinguishable from the Ortega et al. (2015) reconstruction but remains significantly correlated with it ($r=0.49$, $p<0.01$ over the period 1073-1855).

We have shown that for Enet, PLS and particularly PCR which is frequently used in paleoclimatology, selecting proxy records with a strong correlation with the index to be reconstructed over the training periods is a good way to improve the CE scores, and hence it allows more reliable reconstructions (section 3.2.3.1.1). Contrarily, RF gives more reliable reconstructions using the proxy records significantly correlated at the 80% confidence level with the NAO (section 3.2.3.1.1). This may be due to the fact that it has been mainly developed for large datasets [Breiman, 2001]. For both cases, gathering new proxy records to the 554 available proxy records collected, may be a reliable source of reconstruction improvement. The inclusion of new NAO-sensitive proxy records in the future may thus lead to better reconstructions. ClimIndRec should allow to easily perform such new reconstructions.

3.2.5 Supplementary informations

Supplementary table 1

N °	Code	Location	Longitude (°E)	Latitude (°N)	First year	Last year	Archive	Proxy type	Related variable	Seasonality	Ref.
1	accr-crete-yr-annual	Crete	-38.50	71.00	1000	1973	Ice core	Snow accumulation	Precip.	Annual	Andersen et al. 2006
2	accr-GISP2-yr-annual	GISP2	-38.50	72.60	1000	1988	Ice core	Snow accumulation	Precip.	Annual	Cuffey et al. 1995
3	d18O-agassiz-79-annual	Agassiz	-77.00	80.70	1000	1972	Ice core	$\delta^{18}\text{O}$	SAT	Annual	Fisher et al. 1995
4	icecore-GISP2-ssNA-annual	GISP2	-39.00	73.00	1000	1986	Ice core	Sea salt Na	SLP	DJF	Meeker and Mayewski 2002
5	speleo-crystal-d18O-annual	Crystal cave	-121.00	36.90	1000	2007	Speleothem	$\delta^{18}\text{O}$	SAT	Annual	McCabe-Glynn et al. 2013
6	tree-alps-Tjas-annual	European Alps	9.00	46.00	1000	2004	Tree ring	Tree ring MXD	SAT	JJAS	Büntgen et al. 2006
7	tree-AR050-stab-annual	Black Swamp	-91.30	35.15	1019	1980	Tree ring	Tree ring width	SAT	Annual	Stahle 1996a
8	tree-AR052-stab-annual	Mayberry Slough	-89.00	35.50	1000	1990	Tree ring	Tree ring width	SAT	Annual	Stahle and Cleaveland 2005a
9	tree-CA051-tosh-annual	San Geronio	-116.82	33.40	1000	1970	Tree ring	Tree ring width	SAT	Annual	Tosh 1994
10	tree-ca605-king-annual	Mammoth Peak	-119.00	41.00	1000	1996	Tree ring	Tree ring width	Precip.	NDJFM	Bunn et al. 2005
11	tree-ca640-graum-annual	Hamilton	-118.92	39.00	1000	1988	Tree ring	Tree ring width	Precip.	NDJFM	Bunn et al. 2005
12	tree-co572-woodho-annual	Lily Lake	-105.60	40.30	1000	1998	Tree ring	Tree ring width	SAT	Annual	Woodhouse and Brown 2006
13	tree-FL001-stab-annual	Choctawhatchee River	-85.92	30.45	1000	1992	Tree ring	Tree ring width	SAT	Annual	Stahle and Cleaveland 2005b
14	tree-forfjo-cloud-annual	Forfjördalen	15.73	68.80	1000	2001	Tree ring	Tree ring $\delta^{13}\text{C}$	Cloud %	JJA	Young et al. 2012
15	tree-LA001-stab-annual	Big Cypress	-92.97	32.25	1000	1988	Tree ring	Tree ring width	SAT	Annual	Stahle 1996b
16	tree-mor-pdsi-annual	Morocco	-5.00	33.75	1049	2001	Tree ring	Tree ring width	PSI	FMAMJ	Esper et al., 2007
17	tree-mt113-wagon-annual	Yellow Mountain Ridge (Entire Bark Trees)	-109.80	45.60	1000	1998	Tree ring	Tree ring width	Precip.	NDJFM	Graumlich et al. 2003
18	tree-NM584-touch-annual	Mesa Alta	-106.60	36.20	1000	2007	Tree ring	Tree ring width	Precip.	ONDJFMAMJ	Touchan et al. 2011
19	tree-NV516-grayb-annual	Hill 10842	-114.20	38.90	1000	1984	Tree ring	Tree ring width	SAT	Annual	Graybill 1994a
20	tree-nv517-grayb-annual	Spring Mountains Lower	-114.70	34.30	1000	1984	Tree ring	Tree ring width	SAT	Annual	Graybill 1994b
21	tree-SCpla-precip-annual	S. Colorado Plateau I	-109.30	37.50	1000	1987	Tree ring	Tree ring width	Precip.	October-July	Salzer and Kipfmüller 2005
22	tree-SCpla-temp-annual	S. Colorado Plateau II	-110.70	36.50	1000	1996	Tree ring	Tree ring width	SAT	Annual	Salzer and Kipfmüller 2005
23	tree-siber-temp-annual	Taimyr-Putoran	103.00	71.29	1000	1996	Tree ring	Tree ring width	SAT	Annual	Naurzbaev et al. 2002
24	tree-swit177-schwein-annual	Lauenen + div. Stao	6.50	46.42	1000	1976	Tree ring	Tree ring width	SAT	JJA	Schweingruber, 1988
25	tree-UT508-grayb-annual	Wild Horse Ridge	-110.10	40.00	1000	1985	Tree ring	Tree ring width	SAT	Annual	Graybill 1994c
26	tree-UT509-grayb-annual	Mammoth Creek	-112.67	37.65	1000	1989	Tree ring	Tree ring width	SAT	Annual	Graybill 1994d
27	tree-silbermale-trw-annual	Albermale Sound	-76.00	36.00	934	2005	Tree ring	Tree ring width	PDSI	July	Stahle et al. 2013
28	tree-arjeplog-bi-annual	Arjeplog	17.90	66.50	1200	2010	Tree ring	Tree ring BI	SAT	JJA	Björklund et al. 2014
29	tree-jamtland-mxd-annual	Jamtland	15.00	63.10	800	2011	Tree ring	Tree ring MXD	SAT	AMJJAS	Zhang et al. 2016
30	tree-colzad-trw-annual	Col du Zad	-5.10	33.00	984	1984	Tree ring	Tree ring width	PSI	FMAMJ	Esper et al., 2007
31	tree-forf-mxd-annual	Forfjördalen-x	15.70	68.80	978	2005	Tree ring	Tree ring MXD	SAT	AMJJAS	McCarroll et al. 2013
32	tree-khibiny-bi-annual	Khibiny	33.50	67.50	821	2005	Tree ring	Tree ring BI	SAT	JJA	McCarroll et al. 2013
33	tree-lamila-mxd-annual	Laanila	27.30	68.50	800	2005	Tree ring	Tree ring MXD	SAT	JJA	McCarroll et al. 2013
34	tree-manitoba-trw-annual	S Manitoba	-97.10	49.50	1409	1998	Tree ring	Tree ring width	Precip.	Annual	George and Nielsen 2002
35	tree-mesoamerica-trw-annual	Mesoamerica	-100.00	20.00	800	2008	Tree ring	Tree ring width	PDSI	June	Stahle et al. 2011
36	accr-NGRIPs-yr-annual	NGRIP-s	-42.00	76.00	800	1995	Ice core	Snow accumulation	Precip.	Annual	Andersen et al. 2007
37	tree-potoriv-trw-annual	Potomac River	-77.50	39.30	950	2001	Tree ring	Tree ring width	Stream flow	MJJAS	Maxwell et al. 2011
38	tree-quebec-mxd-annual	Quebec-x	-77.50	39.30	1373	1988	Tree ring	Tree ring MXD	SAT	MJJAS	(Not found)
39	tree-SCengland-trw-annual	SC England	-1.40	51.50	950	2009	Tree ring	Tree ring width	Precip.	MJJ	Wilson et al. 2013
40	tree-sodankyla-thi-annual	Sodankyla	27.00	67.00	800	2007	Tree ring	Tree height increment	SAT	JJA	Lindholm and Jalkanen 2011
41	tree-southfin-mxd-annual	Southern Finland	28.50	61.50	800	2000	Tree ring	Tree ring MXD	SAT	MJJAS	Helama et al. 2014
42	tree-SWturkey-trw-annual	SW Turkey	31.00	37.00	1339	1998	Tree ring	Tree ring width	Precip.	MJ	Touchan et al. 2003
43	tree-tyrol-mxd-annual	Tyrol	12.50	48.00	1053	2003	Tree ring	Tree ring MXD	SAT	JAS	Esper et al., 2006
44	d18O-Nilshel-yr-annual	North Icelandic Shelf	66.53	-18.20	953	2000	Ice core	$\delta^{18}\text{O}$	SAT	Annual	Reynolds et al. 2016

TABLE 3.2 | Proxy records not in PAGES 2k.

Supplementary table 2

Volcano	Country	Date
Unknown		1229
Samalas	Indonesia	1257
Unknown		1285
Huaynaputina	Peru	1600
Parker	Phillippines	1640
Serua	Indonesia	1693
Unknown		1809
Tambora	Indonesia	1815
Cosiguina	Nicaragua	1834
Krakatau	Indonesia	1883

TABLE 3.3 | Ten largest volcanic eruptions from Ortega et al. 2015.

Supplementary table 3

Volcano	Country	Date
Unknown		1108
Unknown		1171
Unknown		1230
Samalas	Indonesia	1258
Kuwae	Vanuatu	1458
Huaynaputina	Peru	1601
Parker	Phillippines	1641
Serua	Indonesia	1695
Lakagigar	Iceland	1783
Unknown		1809
Tambora	Indonesia	1815

TABLE 3.4 | Ten largest volcanic eruptions from Sigl et al. 2015.

Supplementary table 4

Method	Mean test variance [range]	Mean train variance [range]	Variance 1000-1855	Variance 1856-1970 for Enet and PCR (1856-1972 for RF and PLS)
RF	0.24 [0.08,0.47]	0.74 [0.64,0.84]	0.16	0.72
Enet	0.26 [0.09,0.96]	0.4 [0.27,0.7]	0.18	0.33
PCR	0.35 [0.06,0.77]	0.56 [0.37,0.86]	0.56	0.62
PLS	0.33 [0.08,1.04]	0.57 [0.4,0.8]	0.17	0.43
NAO (1000-1970)	1.44 [1.23,1.61]	1.46 [0.72,2.27]	1.44	
NAO (1000-1973)	1.44 [1.27,1.57]	1.57 [0.73,2.33]	1.46	

TABLE 3.5 | Variance de la NAO reconstruite.

Supplementary table 5

Package	Reference
glmnet	Friedman et al. 2010
pls	Mevik et al. 2007
randomForest	Liaw and Wiener 2002
ncdf4	Pierce 2017
stringr	Wickham 2017

TABLE 3.6 | R packages used in ClimIndRec.

Supplementary material 1 : Regression methods

Voir section 2.4.

Supplementary material 2 : Correlation test

Voir section 2.3.3.

Supplementary material 3 : Uncertainties

The uncertainties we provide are calculated as in Ortega et al. (2015) but using the residuals calculated for the final model. These regression uncertainties are represented by the standard errors (s.e.) of the regression, calculated as the root of the sum of the squared residuals over the training periods divided by the degree of freedom :

$$s.e = \sqrt{\frac{\sum_{i=1}^n (Y - \hat{Y})^2}{n - 2}} \quad (3.9)$$

Where n_{train} is the length of the learning sample, Y the true values of the NAO index over the learning period, and $\hat{Y}$ the fitted NAO by the regression model over the training period. An uncertainty band $2 * s.e.$ is calculated which is our estimate of the total uncertainty range of the reconstruction.

Supplementary material 4 : Superposed Epoch Analysis

Voir section 2.3.5

3.3 Devenir de ClimIndRec : perspectives et améliorations possibles

En plus de proposer un outil de reconstruction, cette étude montre que son utilisation, en le soumettant à différentes configurations et en les comparant, permet d'obtenir une reconstruction plus robuste que celle de référence de Ortega et al. (2015), si l'on se base sur la même métrique que ces derniers. L'idée de développer un tel outil est également de s'assurer de son bon fonctionnement futur et de proposer des mises à jour. Notamment, dans une nouvelle version de l'outil (1.1), postérieure à celle présentée dans Michel et al. (2020), il est proposé deux nouvelles approches de régression qui sont toutes deux des variantes de l'Enet : la première est connue sous le nom de Lasso adaptatif, et la seconde est une régression Ridge après sélection de variables par Lasso (section 2.4.4). Cette nouvelle version de ClimIndRec possède une nouvelle métrique d'évaluation. Il s'agit de la réduction d'erreur (*Reduction of error*, RE), très fréquemment utilisée en dendroclimatologie (notamment Macias-Fauria et al. 2012). L'outil brut présenté dans la section précédente est disponible sur le site de partage de codes scientifique Zenodo : <https://zenodo.org/record/3464293.Xy1EBS1h1fi>. Sa version maintenue et mise à jour est disponible sur GitHub : <https://github.com/SimMiche/ClimIndRec>. Les prochaines versions peuvent être dédiées à ajouter des méthodes de régression, des métriques ou des options de reconstruction.

Si l'étude présentée ici se voulait méthodologique, nous avons également remarqué de nombreuses sources d'amélioration en matière de robustesse, indépendamment de la méthode d'évaluation et de sélection de modèle statistique présentée. Un point important souligné dans la discussion de Michel et al. (2020) est qu'il n'est pas possible d'introduire des enregistrements qui ne couvrent pas entièrement la période d'apprentissage et la période de reconstruction. C'est pourquoi dans la reconstruction finale, on n'a que 90 enregistrements disponibles, dont 48 sont sélectionnés par test de corrélation au seuil de significativité optimal de 80% (Michel et al. 2020). Puisqu'ici il s'agit de détailler une méthode de comparaison de modèles statistiques pour une période de reconstruction donnée, la version actuelle n'est pas vouée à directement répondre à ces problèmes de données manquantes. Une manière de pallier à ce problème serait d'appliquer la méthode de sélection de modèle statistique présentée ci-avant, mais sur plusieurs fenêtres imbriquées (Wang et al. 2017). Le principe est ensuite de mettre ces reconstructions les unes après les autres afin d'utiliser un nombre croissant d'enregistrements à mesure qu'une année reconstruite est proche dans le temps. Notons que puisque ceci n'a pas été appliqué dans l'étude Michel et al. (2020), il y'a pas ici la possibilité de produire une nouvelle reconstruction de NAO, en reprenant la configuration utilisée et en l'appliquant sur toutes les fenêtres possibles avant les observations de la NAO : de 1000-1972 à 1855-1972 dans notre cas. Une nouvelle reconstruction de NAO pourrait ainsi être proposée en imbriquant ces différentes reconstructions au niveau de chaque pas de temps. Maintenant que l'on dispose de la méthode présentée, les prochains chapitres feront des reconstructions imbriquées à partir de configurations de reconstruction optimales préalablement déterminées.

Enfin, une méthode émergente pour la validation de reconstruction est l'approche par *pseudo-proxy* (Ortega et al. 2015, Neukom et al. 2018), qui n'a pas été appliquée ici puisque cela sort du champ de notre étude. L'objectif d'une approche par

pseudo-proxy est de reproduire, au sein d'une ou plusieurs simulations de modèles, la méthode de reconstruction établie sur les données réelles. Le but est ainsi de vérifier la reproductibilité de la reconstruction au sein d'un modèle, notamment en terme de téléconnexion entre les enregistrements utilisés et l'indice reconstruit. Le prochain chapitre s'intéressera à produire une reconstruction de l'AMV et s'appliquera à étudier les propriétés de la reconstruction obtenue. Une approche de *pseudo-proxy* sera alors appliquée afin de gagner encore en confiance dans la reconstruction obtenue, en se basant sur un nouvel élément de validation issu d'un modèle de climat. Ceci garantit la cohérence de la méthode avec ce que l'on connaît de la dynamique du climat au sein des modèles.

Chapitre 4

Reconstruction des fluctuations de l'AMV : implications pour son évolution passée et future

Ce chapitre présente une reconstruction des variations internes de la SST de l'Atlantique Nord (AMV). L'objectif est de mieux comprendre son comportement et d'évaluer l'état de ses variations actuelles au vu de ses variations passées. La section 4.1 expliquera en quoi la procédure de reconstruction spécifique à cette étude est novatrice et peut potentiellement permettre de répondre à des questions scientifiques non élucidées. La section 4.2 est un article en correction pour le journal *Nature Geoscience*. Dans la section 4.3 nous esquisserons l'importance et l'implication des résultats présentés dans la section 4.2, en particulier au sujet d'un signal d'alerte de l'émergence d'un point de bascule dans le bassin de l'Atlantique Nord.

4.1 L'AMV comme mode de variabilité interne : une stratégie de reconstruction particulière

Comme évoqué dans la section 1.3.1, la nature des variations de l'AMV ne fait pas consensus. Bien qu'il semblerait que les facteurs internes comme externes jouent tout deux un rôle pour produire les fluctuations multi-décennales observées de la NASST (Wanatabe and Tatebe 2019), leurs contributions relatives restent encore incertaines.

Pour tenter de répondre à ce problème, Wang et al. (2017) ont d'abord reconstruit la NASST brute à base d'une méthode de PCR, comprenant les composantes de forçages et celles de variabilité interne. Ils ont ensuite retiré la composante forcée de leur reconstruction à l'aide d'un modèle linéaire inverse appliqué à une estimation du forçage du dernier millénaire, pour en déduire une reconstruction d'AMV. Comme soulevé dans les difficultés identifiées dans le chapitre 3, la reconstruction de Wang et al. (2017) souffre d'approximations méthodologiques, notamment dues à des choix de reconstruction arbitraires. L'exemple le plus éloquent est le choix du nombre de PC utilisées pour régresser la NASST sur l'historique. Ces derniers utilisent les PC de la matrice des enregistrements paléoclimatiques normalisées dont la valeur propre est supérieure à 1. Dans un cadre de PCA où l'unique but est de réduire les dimensions de la matrice d'enregistrements, ce choix a du sens. En effet, puisque ces derniers normalisent la matrice d'enregistrements paléoclimatiques, la somme des variances de ces derniers (qui valent toutes 1), est exactement égale à leur nombre. Ainsi, puisque la somme des valeurs propres est égale à la somme des

variances des données initiales, on aurait, sous l'hypothèse que tous les enregistrements sont orthogonaux (corrélations deux-à-deux de 0), toutes les valeurs propres auraient un poids de 1. De ce fait, comme les enregistrements ne sont dans la réalité pas orthogonaux, on peut considérer que toute PC dont la valeur propre est supérieure à 1 explique plus de variance du jeu de données initial que s'il n'y avait aucune variabilité commune entre les enregistrements. Cela aurait du sens pour faire de la sélection de PC par variance expliquée, mais ceci n'en a pas nécessairement dans un cadre de PCR. En effet, Wang et al. (2017) se basent ici sur l'hypothèse forte que pour un jeu d'enregistrement donné, toute PC qui explique une part significative de la variance initiale est nécessairement bonne à utiliser pour la régression de la NASST. De plus, Wang et al. (2017) proposent la reconstruction de la NASST pour un jeu d'enregistrements qui ne semble pas réellement adapté. En effet, si l'on regarde les données qu'ils mettent à disposition, on voit que près de la moitié des enregistrements utilisés n'atteignent pas un niveau de significativité de corrélation de 80% avec l'indice reconstruit. C'est une raison de plus pour suggérer que les PC construites, qui expliquent une part importante de la variabilité initiale des enregistrements, ne sont pas nécessairement appropriées pour reconstruire la NASST.

Ici, nous chercherons, tout comme Wang et al. (2017), à proposer une reconstruction de la composante de variabilité interne de la NASST, mais avec une approche sensiblement différente. Tout d'abord, les nombreux indices d'AMV proposés sur la période historique sont déjà construits de manière à isoler la composante de variabilité interne de la NASST, ce qui suggère donc que l'AMV n'est pas forcée. Puisqu'on ne veut pas faire d'*a priori* sur l'indice reconstruit, on testera la reconstruction pour plusieurs de ces indices (quatre d'entre eux) et on sélectionnera celui qui se reconstruit le mieux à l'aide des méthodes de comparaison proposées par Michel et al. (2020). Par ailleurs, puisque les variations des enregistrements couvrant la période historique sont également très influencées par les forçages, reconstruire une série temporelle associée à de la variabilité interne (AMV) à l'aide de telles séries (comprenant une composante forcée) aurait peu de sens. En effet, il nous semble qu'utiliser la NASST comme indice à reconstruire implique de sélectionner des prédicteurs qui sont sensibles aux forçages externes et pas forcément aux variations de l'AMV. Ainsi, une grande particularité des travaux présentés dans ce chapitre est que nous avons également retiré la composante de variabilité forcée de l'Atlantique Nord à chacun des enregistrements de la base de données utilisées (composée à plus de 90% des enregistrements de la base de données PAGES 2k).

En plus de tester la reconstruction pour plusieurs indices, à l'inverse de Wang et al. (2017), on soumettra nos reconstructions aux différentes approches de régression présentées section 2.4 et chapitre 3, mais aussi aux différentes sélections d'enregistrements ainsi que différentes fenêtres d'apprentissage. Afin d'utiliser un nombre satisfaisant d'enregistrements, on procédera, tout comme Wang et al. (2017), à une reconstruction imbriquée, impliquant ainsi l'utilisation d'un nombre croissant d'enregistrements à mesure que l'on se rapproche de la période historique. Enfin, pour rendre crédible les résultats scientifiques que nous exposerons dans la suite de ce chapitre, nous reproduirons la reconstruction d'AMV avec une approche par *pseudo-proxy* au sein de 12 membres de simulations millénaires du modèle de circulation globale CESM1. Ainsi, puisque ces simulations millénaires remontent à l'année 850, c'est jusqu'à cette date que remontera la reconstruction présentée, afin que celle-ci puisse être exactement reproduite par pseudo-proxy.

4.2 Article soumis au journal *Nature Geoscience*

Title : Internal abrupt change in the Atlantic Multidecadal Variability as the onset of the Little Ice Age.

Simon Michel¹, Didier Swingedouw¹, Juliette Mignot², Guillaume Gastineau², Pablo Ortega³, Myriam Khodri², Gerard McCarthy⁴

1 : Environnements et Paleoenvironnements Oceaniques et Continentaux (EPOC), UMR CNRS 5805 EPOC-OASU-Universite de Bordeaux, Allée Geoffroy Saint-Hilaire, Pessac 33615, France.

2 : Sorbonne Universités (UPMC, Univ. Paris 06)-CNRS-IRD-MNHN, LOCEAN Laboratory, 4 place Jussieu, F-75005 Paris, France.

3 : BSC, Barcelona, Spain.

4 : Irish Climate Analysis and Research UnitS (ICARUS), Department of Geography, Maynooth University, Maynooth, Ireland

Abstract :

The Atlantic Multidecadal Variability (AMV) has crucial impacts on human societies (Defrance et al. 2017) and ecosystems (Edwards et al. 2013). Its periodicity (Moore et al. 2017, Singh et al. 2018b) and drivers (Muir and Fedorov 2015, Yan et al. 2018, Clement et al. 2015, Wanatabe and Tatebe 2019) are controversial due to the short temporal extent of instrumental observations and competing impacts of external and internal forcing, which may limit our ability to predict it. Here, we use a well-verified set of proxy records (PAGES 2k Consortium 2017) and four regression methods (Michel et al. 2020) to perform and compare reconstructions since 850 C.E. of AMV indices reflecting the internal variability of the Atlantic (Enfield et al. 2001, Trenberth and Shea 2006, Ting et al. 2009, Enfield and Cid-Serrano 2010). An optimal reconstruction of internal AMV variations, which is consistently verified by pseudo-proxy experiments, is obtained using a non-linear method (random forest) newly applied to paleodata (Michel et al. 2020). It exhibits large multi-decadal variations in the range 20-80 years, a larger spectrum variability than the prominent 50 to 80-year timescale of the instrumental era (Kushnir 1994). Furthermore, this reconstruction suggests an early onset of the Little Ice Age in the North Atlantic, which is thus not related with the 13th century cluster of volcanic eruptions. This early onset highlights that the AMV can undergo rapid noise-induced shifts, almost irreversible over centuries. With the contemporary cold temperatures in the North Atlantic, our results question the possibility of near-term rapid and centuries-long shift in the Atlantic. The recent changes in spectral properties of the AMV can indeed be interpreted as an early warning for the proximity of a tipping point (Boulton et al. 2014).

The North Atlantic region is characterized by successive decades of anomalously warm and cold phases in sea surface temperatures (Kushnir 1994) (SST) relative to the global average, which can thus contribute to amplification or damping of global warming effects in the Atlantic sector (Ting et al. 2009, Mann et al. 2014). The underlying variability mode, the Atlantic Multidecadal Variability (AMV), has diverse climatic effects in North Atlantic neighbouring regions (Sutton and Dong

2012, Azuz-Adeath et al. 2019). At the global scale, it also influences drought and rainfall in the Sahel (Defrance et al. 2017, Knight et al. 2006), Northeastern Brazil (Knight et al. 2006) and Central Asia (Knight et al. 2006); Atlantic hurricane frequency and intensity (Trenberth and Shea 2006, Enfield and Cid-Serrano 2010); sea ice thickness and extent over the Arctic (Miles et al. 2014); and is linked to Pacific climate variability (Ruprich-Robert et al. 2017). The processes driving the AMV and its resulting spectral properties remain a source of controversy. Disagreements come from the relatively short period over which the AMV is directly observed, a period which mainly encompasses climate changes significantly affected by human activity (Cook et al. 2016). Internal ocean variability is often cited as a driver of the AMV pattern through variations in the Atlantic Meridional Overturning Circulation (Muir and Fedorov 2015, Yan et al. 2018) (AMOC). These results found using coupled ocean-atmosphere general circulation models, have been further supported by a tide gauge-based reconstruction of the ocean circulation intensity in the North Atlantic intergyre region that exhibits strong correlations with the AMV index over the last 60 years (McCarthy et al. 2015). Several long paleoclimate records also indicate the existence of such multidecadal variability on longer time scales (Gray et al. 2004, Mjell et al. 2015). Nevertheless, it has also been shown that the AMV pattern and its red noise spectral properties can be reproduced in a slab-ocean model, without any influence of ocean dynamics (Clement et al. 2015), resulting only from the integration by the ocean of the stochastic noise embodied in the North Atlantic Oscillation related heat fluxes. It has also been highlighted that periods of negative AMV are dominated by both strong anthropogenic aerosol emissions in bordering regions of Europe and North America, and strong volcanic activity (Wanatabe and Tatebe 2019). The observed multidecadal behavior of North Atlantic SST (NASST) can be thus partly explained by variations in external forcing. Nevertheless, the potential impact from AMV changes driven by ocean circulation or external factors might be very different since the internal AMV (due to ocean dynamics and heat fluxes) strongly impact the meridional temperature gradient (Knight et al. 2005) and the atmospheric circulation, while external forcings have a more homogenous spatial impact. Thus, the two signals should be deconvoluted, especially when trying to reconstruct the AMV further back in time. In terms of spectral properties, there are discrepancies between the observed periodicity in AMV (50-80 years) and the periodicity found in a large part of the control runs of the models from the Climate Model Intercomparison Project Phase 5 (CMIP5) (10-30 years) (Lin et al. 2019). Grasping the relative contributions of the AMV driving factors and its resulting spectral properties could therefore be useful in assessing models and eventually developing emergent constraint approach based on these spectral properties to reduce uncertainty in decadal climate predictions (Hall et al. 2019). To reach this long-term goal, an improved knowledge of the internal variability of the NASST is crucial.

The AMV index definition is of paramount importance to tackle correctly those lingering questions. Since climate is currently in a period with a dominant anthropogenically induced warming signal, it is required to remove the global effect of radiative forcings to isolate the internal variability inherent to the North Atlantic basin. In this study, we chose four AMV indices where the external forcing from NASST has been removed using different techniques, in order to only account for its internal variations. We denote these indices AMV_{LD} (Enfield et al. 2001), AMV_{QD} (Enfield and Cid-Serrano 2010), AMV_{TS} (Trenberth and Shea 2006) and AMV_T (Ting et al. 2009) (see Methods and Fig. 4.1a). The observed widespread impacts of the AMV

(Fig. 4.1bc) suggest that its past variations may have played a decisive role in the structural evolution of the climate at the global scale. The last millennium is well-covered in space by high-resolution proxy records, making it very suitable for annually-resolved paleoclimate reconstructions. It is of particular interest since it encompasses two contrasting periods in temperature : the warm Medieval Climate Anomaly (MCA) and the cold Little Ice Age (LIA). The relative contribution of external forcing and internal climate variability to the transition between both sequences is unclear, as their spatio-temporal evolution (Schurer et al. 2014, Schleussner et al. 2015, Moreno-Chamorro et al. 2016).

Taking advantage of paleoclimate data covering the last millennium, there have been some attempts to reconstruct AMV indices. A recent study (Wang et al. 2017), reconstructed AMV over the last millennium using Principal Components Regression method and proxy records from continental borders of the North Atlantic. However, it is using NASST anomalies (Wang et al. 2017) for model fitting, while it includes a forced signal. The resulting reconstruction is thus overestimating the radiatively-induced response to strong volcanic eruptions, and mask part of the signal from internal variability, which may induce very specific imprints. In addition, this study is based on a single regression method (PCR, *cf.* Methods), and do not include an objective criterion for the selection of proxy records, meaning that some of them may have potentially no correlation with the target AMV index over the instrumental period. It also does not include any pseudo-proxy validation of the reconstruction to assess its reproducibility within a climate model.

4.2.1 Robust and validated AMV reconstruction

The recent gathering of annually-resolved proxy records of past temperature variations is providing a key opportunity to test and develop new objective reconstruction methodologies using advanced techniques. Here, we use the well-verified

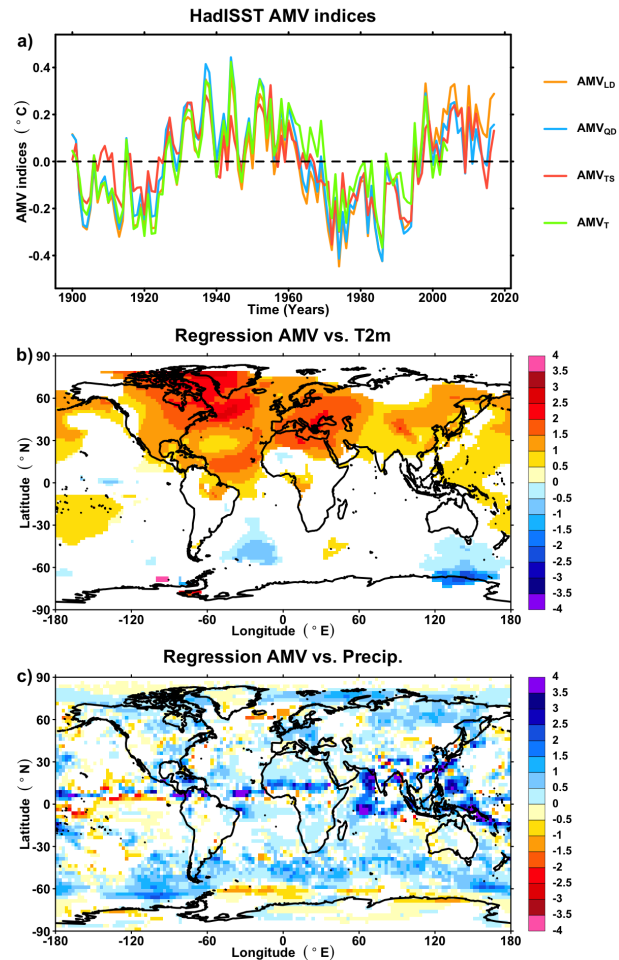


FIGURE 4.1 | Climatic impacts of the AMV over the historical era. a) Historical evolution of AMV indices investigated in this study for the period 1900-2017 calculated using the HadISST dataset (Rayner et al. 2003) (*cf.* Methods). b) Map of averaged regression coefficients between the four AMV indices from a) and GISTEMP v4 temperature dataset (Lensen et al. 2019) for the period 1900-2012 c) Map of averaged regression coefficients between the four 11-year running mean AMV indices from a) and 11-year running mean CMAP precipitation dataset (Xie and Arkin 1997) for the period 1979-2014. For b) and c), white grid points mean that at least 2 of the 4 indices do not have regression coefficients significantly different than 0 at the 90% confidence level, using a student test with corrected degrees of freedom (McCarthy et al. 2015) (Methods)

PAGES 2k database (PAGES 2k Consortium 2017) (P2k, 722 records), complemented with 41 other records from different sources (Tab. 3.2 - Extended Data Table 1), all labelled together as the PAGES 2k+ database (P2k+, 763 records). To avoid spectral artefacts in the AMV reconstruction, we isolate the annually-resolved proxy records (P2k-AR+, 550 records). As we aim at reconstructing the internally-driven variations in NASST, we regress out from each proxy record, the forced component of NASST as estimated from an ensemble of historical simulations from 37 models contributing to CMIP5 (Tab. 4.1 - Extended Data Table 2). This component is estimated using signal-to-noise maximizing empirical orthogonal function (Ting et al. 2009) (Methods, Fig. 4.5 - Extended Data Fig. 1). AMV indices over the historical era are estimated from the HadISST dataset (Rayner et al. 2003) (Methods) (Fig. 4.1a) and its widespread climate signature is estimated using the GISTEMP ERSST v5 (Lenssen et al. 2019) temperature and CMAP (Xie and Arkin 1997) precipitation datasets (Fig. 4.1bc).

Regression methods require that the predictors, here proxy data, have no gaps. This has a significant impact on the number of proxy data used when producing such long reconstructions. To remedy this, we proceed with a nested reconstruction approach (Wang et al. 2017) which consists in building a statistical model for each time step by sequentially using the available proxies. We set that the reconstruction covers periods back to 850 C.E. so that it encompasses the same period than general circulation models' last millennium simulations, thereby allowing to estimate its accuracy using pseudo-proxy experiment (Neukom et al. 2018) (PPE).

The Coefficient of efficiency (CE) metric (Nash and Sutcliffe 1970) is firstly used to compare reconstructions dating back to 850 C.E. (*i.e.* on the longer nested window) of each of the indices, for different proxy datasets (Methods) and four regression methods (Michel et al. 2020) : PCR, partial least squares, elastic-net and random forest (see Supplementary Informations). We use a novel climate index reconstruction software (Michel et al. 2020) to generate 496 reconstructions (124 for each AMV index, *cf.* Methods), with a score computed using the CE metric (Nash1970) (Methods) over 30 training and testing samples (Methods).

CE scores of the 496 reconstructions for each level of inputs (Methods, Fig. 4.6 - Extended Data Fig. 2) show that the random forest is on average the most robust regression technique among the four tested. It provides the highest CE scores for the reconstruction of the AMV_{QD} over the period 850-1989, using proxy records significantly correlated at the 95% confidence level with the latter. For this reconstruction, CE score is positive at the 99% confidence level, which validates its reliability ($CE \in [0.08, 0.64]$, $med(CE) = 0.47$, *cf.* Methods). We therefore perform reconstructions of AMV_{QD} with random forest regression for each time window from 850-1989 to 1850-1989 with an increment of one year. For each of them, only proxy records available and significantly correlated with the observed index at the 95% confidence level are used. The final nested reconstruction is thus obtained by placing one after the other the 1001 interlocking reconstructions (Fig. 4.2a). They provide increasing validation scores over time since the number of proxy records used is generally higher for the shortest (most recent) reconstructions (Fig. 4.2b). The nested reconstruction uses an ensemble of 102 proxy records. Their a posteriori weights, given by random forest importance (*cf.* Methods), and their temporal availability, are presented in Fig. 4.2cd. Proxy records having the highest weights belong to central Asia

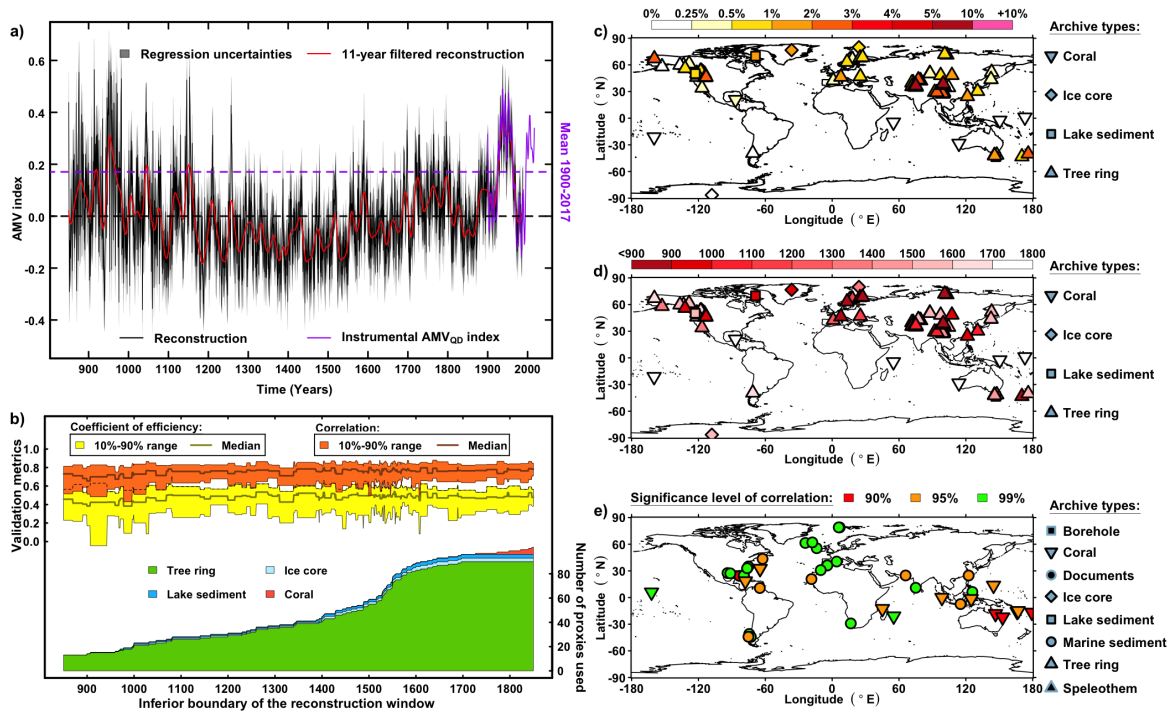


FIGURE 4.2 | Nested reconstruction of the AMV and related proxies. a) Black line : Annually-resolved nested reconstruction of AMV_{QD} using random forest (*cf.* Methods). Red line : Cubic spline smooth of the annually-resolved nested reconstruction (black line). The regression uncertainties of the annually-resolved nested reconstruction (black line) are defined for each time step of the nested reconstruction as ± 2 standard error of the regression. b) Validation metrics (CE and correlation) obtained over 30 training-testing splits, and proxy records types availability for the nested AMV_{QD} reconstruction. c) Proxies weights from the random forest method, relative to the proxy records temporal availability (see Methods) d) Temporal coverage of the availability of the proxy records. e) Significance levels of correlation between the nested AMV_{QD} reconstruction and the 43 ocean proxy records from P2k+ at least correlated with the latter at the 95% confidence level and not used to produce the reconstruction.

with two main clusters of records : one in Eastern Pakistan and one in southern central China (Fig. 4.2c), two regions where summer temperatures and precipitations are affected by variations of the AMV (Fig. 4.1bc). The remaining highly weighted proxies are mostly from Europe, Arctic and western North America, which are well known regions of AMV fingerprints (Sutton and Dong 2012, Azuz-Adeath et al. 2019). Proxies from other regions have lower weights ($<0.25\%$ of the total proxies' weight, Methods), except for the proxies of southern Oceania, a region where temperatures are significantly affected by AMV variations over the instrumental period (Fig. 4.1bc).

As for previous reconstructions (Gray et al. 2004, Wang et al. 2017), most of the proxies used are terrestrial proxies while the AMV is an oceanic mode. Our method indeed selected a poor amount of ocean proxies. To evaluate if our AMV reconstruction is in line with existing ocean proxy records, we consider those not used in the reconstruction. We find that 43 ocean records (22 from the North Atlantic) from P2k+ are significantly correlated at least at the 90% confidence level with our AMV reconstruction (Fig. 4.2e). This provides further confidence in the quality and consistency of this reconstruction of the AMV.

The consistency of the reconstruction has also been validated by PPE (Neukom

et al. 2018), using 12 members from the Last Millennium Large Ensemble of the Community Earth System Model (CESM1-LMLE, see Methods). For each reconstruction window, the available pseudo-proxies that are significantly correlated at the 95% confidence level with the simulated AMV are selected and an evaluation of the reconstruction is performed through 30 training/testing splits using the CE metric. For each member, we calculate CE scores, as well as the correlation between the AMV within the model and the AMV reconstructed from pseudo-proxies (Fig. 4.3a). The AMV is well reproduced among the different members, since each correlation between the reconstructed AMV and the model AMV is significant at least at the 95% confidence level, except for the 9th member for which the AMV is poorly reproduced ($r=0.12$, $p>0.2$). The other correlations range from 0.28 to 0.75. Way more homogeneity than in the real experiment is found for the weights of the selected pseudo-proxies within the different members (Fig. 4.3b) with similar main regions of pseudo-proxy selection than in the real experiment. However, although the main regions of the selected proxies are well represented in PPE, there are notable differences in their exact locations in the PPE and in the real experiment, which can be attributed to the biases of the model in how it represents AMV teleconnections at the regional scale. Thus, this PPE approach provides further confidence in our reconstruction and constitute the first model-based validation for an AMV reconstruction, which guarantees robustness.

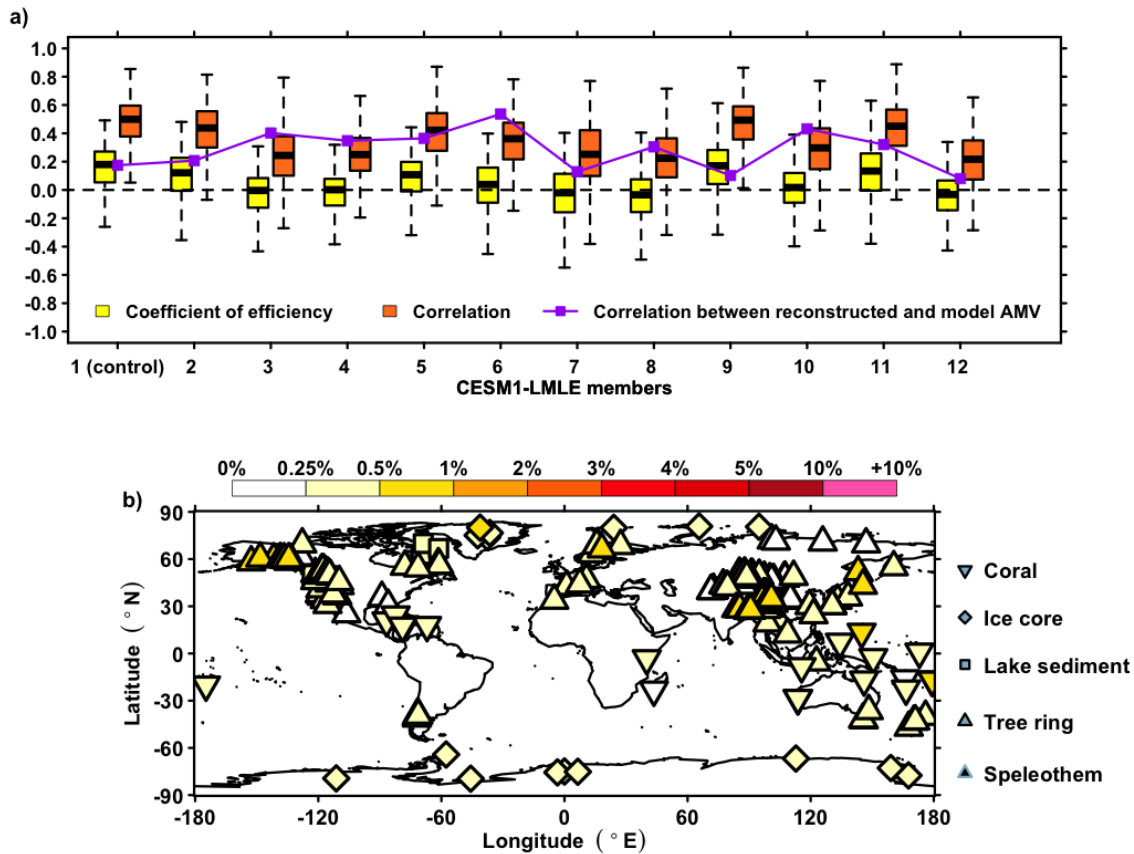


FIGURE 4.3 | Pseudo-Proxy validation. a) CE scores (yellow boxplots), correlation scores (orange boxplots) and correlation between the reconstructed AMV_{QD} and the model AMV_{QD} (purple line) for the 12 members of CESM1-LMLE. b) Pseudo-proxy weights given by the random forest method, relative to the pseudo-proxy temporal availability and CESM1-LMLE members (see Methods).

4.2.2 An early internally-driven onset of LIA

A striking feature of the reconstruction is the significant difference between the preindustrial (850-1850) and observed (1900-1989) mean states. This difference rises up to 0.17°C, which corresponds to 1 standard deviation over the whole period (850-1989, Fig. 4.4a). Hence, in the following, the mean state, and thus “positive” and “negative” states of AMV are defined with respect to its mean over the whole period (850-1989) instead of the instrumental one (1900-2018, Fig. 4.2a). It shows that from the late 9th century to the late 12th century, the index is mainly in a dominant positive phase (Fig. 4.4a). Around 1200, it exhibits a transition towards five hundred years of prevailing negative signal, with a mean state of approximately 1.5 standard deviation lower than the observed AMV. This extremely negative state has never been reached over the historical period and indicates that the North Atlantic was anomalously cold relative to the rest of the ocean, according to a classical definition of the AMV (Trenberth and Shea 2006). Yet, it is difficult to translate its variation in terms of cold or warm periods as this index is a relative balance of SST. To address this issue, we perform a reconstruction of NASST with no external forcing removal on proxies, as done in a recent study (Wang et al. 2017). For doing so, we use the same statistical model selection and the same nested approach than for the AMV to reconstruct the NASST back to 850 C.E. as well. The best reconstruction of NASST is obtained with the PCR approach. It exhibits significant correlation with our AMV reconstruction ($r=0.64$; $p<0.01$, Fig. 4.4a). This means that more than 40% of the variability of NASST over the last millennium can be explained by internal variability only, the rest being related with external forcing.

Using a recent reconstruction of volcanic activity (Sigl et al. 2015), we perform a superposed epoch analysis (Rao et al. 2019) (Methods) for both the NASST and the AMV reconstructions, in response to the 11 largest eruptions of the last millennium (Tab. 4.2 - Extended Data Table 2). While the reconstructed NASST has a similar response a decade after the eruption than a recent reconstruction also based on NASST (Wang et al. 2017) (Fig. 4.7 - Extended Data Fig. 3), no significant response is found for the AMV reconstruction (Fig. 4.7 - Extended Data Fig. 3). This highlights the fruitfulness of our approach to obtain an index which is not directly influenced by the external radiative forcing of the largest eruptions of the last millennium, and therefore seems to be mainly representing the internally generated component of NASST.

The late 12th century transition towards 500 years of negative anomalies is found in both the NASST and the AMV reconstruction. This transition may have thus been mainly caused by internal variability and has had a significant impact on the NASST, which also moved towards a prevailing negative phase. This view is supported by reconstructions of past Arctic cryosphere that report important glacier advances and sea ice expansion at the end of the 12th century (Kinnard et al. 2011, Jomelli et al. 2016, Halloran et al. 2020), in line with an abrupt transition in North Atlantic and Arctic climate during this period, which have happened before the cluster of strong volcanic eruptions of the 13th century (Sigl et al. 2015). The 15th and the 16th centuries show a substantial reinforcement in NASST with still an apparent drop in response to the 1458 very strong Kuwae (Vanuatu) eruption (Sigl et al. 2015), while the AMV remained negative (Fig. 4.4a). For the well-known cold period of 1600-1850, the AMV has a rather neutral behaviour instead of the prevailing negative one from

NASST reconstruction (Fig. 4.4a). This highlights that contrary to the earlier one, this second cold period, which matches with the classical definition of the LIA, may have been mainly a forced sequence of climate variations, forced by the large volcanic activity of the 17th century and the early 19th century, with little role from internal variability. Such an interpretation has also recently been suggested, based on other pieces of evidence (Brönnimann et al. 2019).

Our AMV reconstruction suggests that multidecadal variations still exist without any external forcings (Fig. 4.4c). AMV primarily varies in the 20-80 years band except for the 1600- 1800 period, which is more dominated by 20-40 years cycles (Fig. 4.4c). Thus, the 50-80 years peak of variability observed over the recent period may not be systematic. This result is key when evaluating model ability to reproduce the AMV spectral properties. Such an assessment can thus provide a source of emergent constraint, based on AMV representation (Hall et al. 2019). The recent increase of the AMV periodicity might could be indicative of a potential regime shift in ocean circulation (Boulton et al. 2014). Indeed, using a similar approach than a former model-based study evaluating early warning signal for an AMOC collapse (Boulton et al. 2014), we compute Kendall τ statistics, which indicate a significant increase in AMV spectral properties over the recent period (Fig 4.4b, Methods). According to the classical tipping points theory (Boulton et al. 2014), this constitutes the first observation-based estimate that the AMV may now be approaching a tipping point.

In conclusion, this new reconstruction of the AMV – using an advanced statistical approach applied to the newly released PAGES 2k database (PAGES 2k Consortium 2017) and robustly verified using pseudo-proxy approach – is showing preferential variability in the 20 to 80-year band. This is contradicting recent papers suggesting that the AMV might be purely red noise (Clement et al. 2015) or that the AMV might not even exist as an internal mode of variability (Mann et al. 2020). In that respect, this robust reconstruction is providing key insights on the nature of the AMV and of decadal variations over the North Atlantic, which might be likely related to AMOC variations given the existence of a prominent time scale of variability. To further validate our results, the same reconstruction approach has been applied to NASST, and comparative study of their response to volcanic eruptions proves that the AMV reconstruction we propose indeed constitutes an internally generated component of

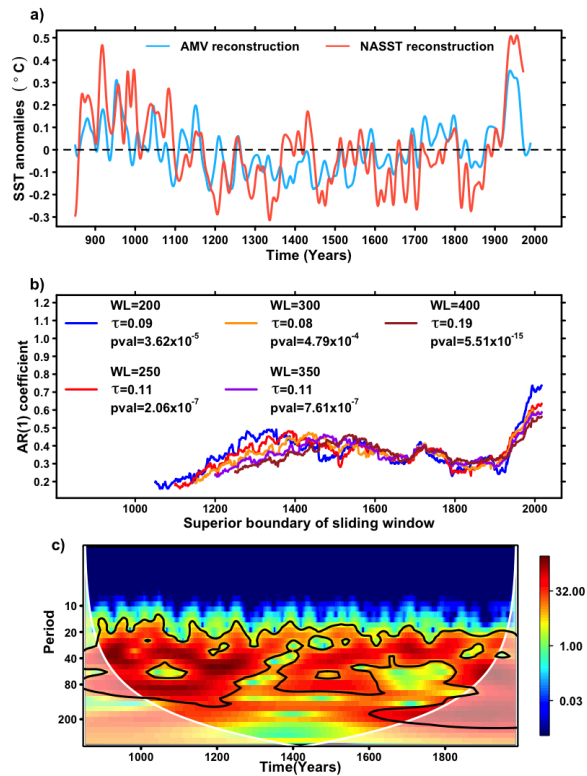


FIGURE 4.4 | Comparison with NASST and early warning signal test. a) Cubic spline smooth of the nested NASST and AMV_{QD} reconstructions from this study. b) Early warning signal test (Boulton et al. 2014) of the nested AMV_{QD} reconstruction (Methods) for different window length (WL). c) Discrete wavelet transform of the nested AMV reconstruction from this study. Contours provide 90% confidence level.

past NASST variability. In that sense, this is the first reconstruction of AMV over the last millennium that can robustly be considered as an internal mode of variability of the climate system. It is providing a new dynamical explanation to the onset of the LIA as a noise-induced shift of the AMV. The latter might have occurred during the late 12th century, and have led the whole NASST into negative anomalies for centuries. Our study thus highlights the potential strong influence of internal variability (*i.e.* not related with any external forcing, even natural) in shaping the variations of climate of the last millennium, especially in the 12th century when the climate shifted from the MCA to the LIA in the Northern Hemisphere, supporting that these sequences have not been a global event (Neukom et al. 2019). Also, the recent changes in spectral property of the AMV towards longer timescales can be related with the possible approach of a tipping point in the North Atlantic system (Boulton et al. 2014) (Fig. 4.4b), suggesting that the AMV, which recently moved towards a negative phase, might be able to switch into a long-term negative state soon.

4.2.3 Extended data figures and tables

Extended data figures

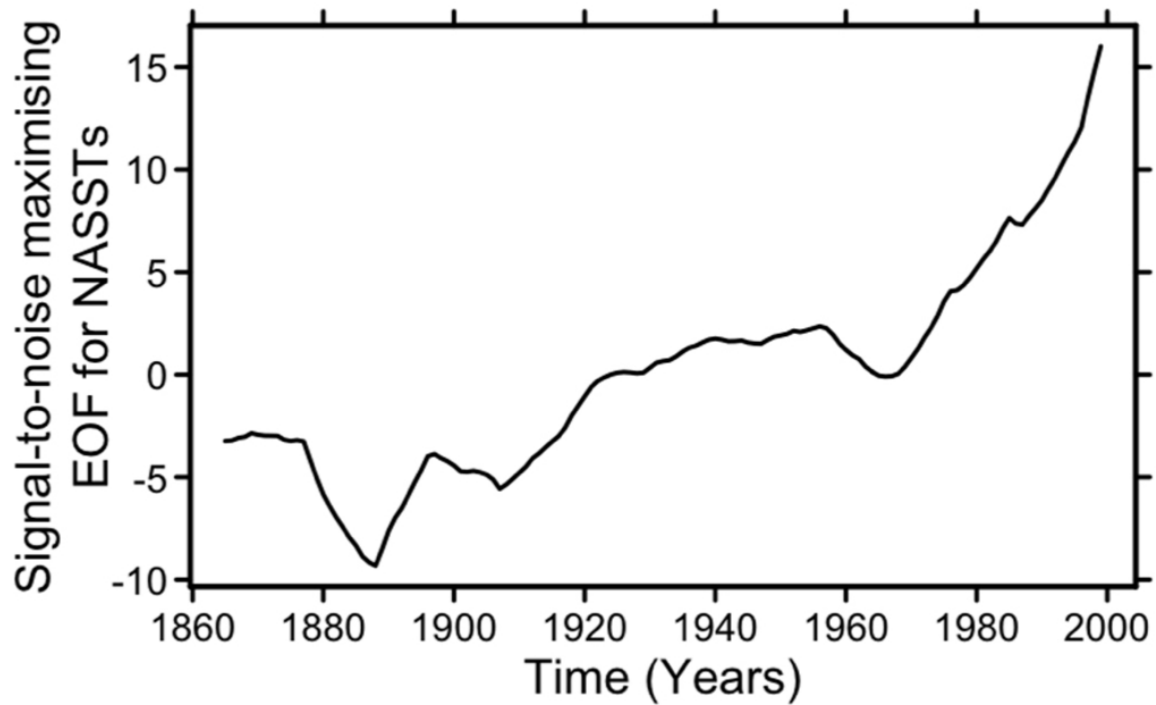


FIGURE 4.5 | Extended Data Figure 1 - Estimated forced component of historical NASST. Estimation is performed using signal-to-noise maximizing EOF (Ting et al. 2009) (*cf.* Methods), using historical simulations from 37 climate models (Tab. 4.1 - Extended Data Table 2). The timeseries is filtered with 11-years running mean.

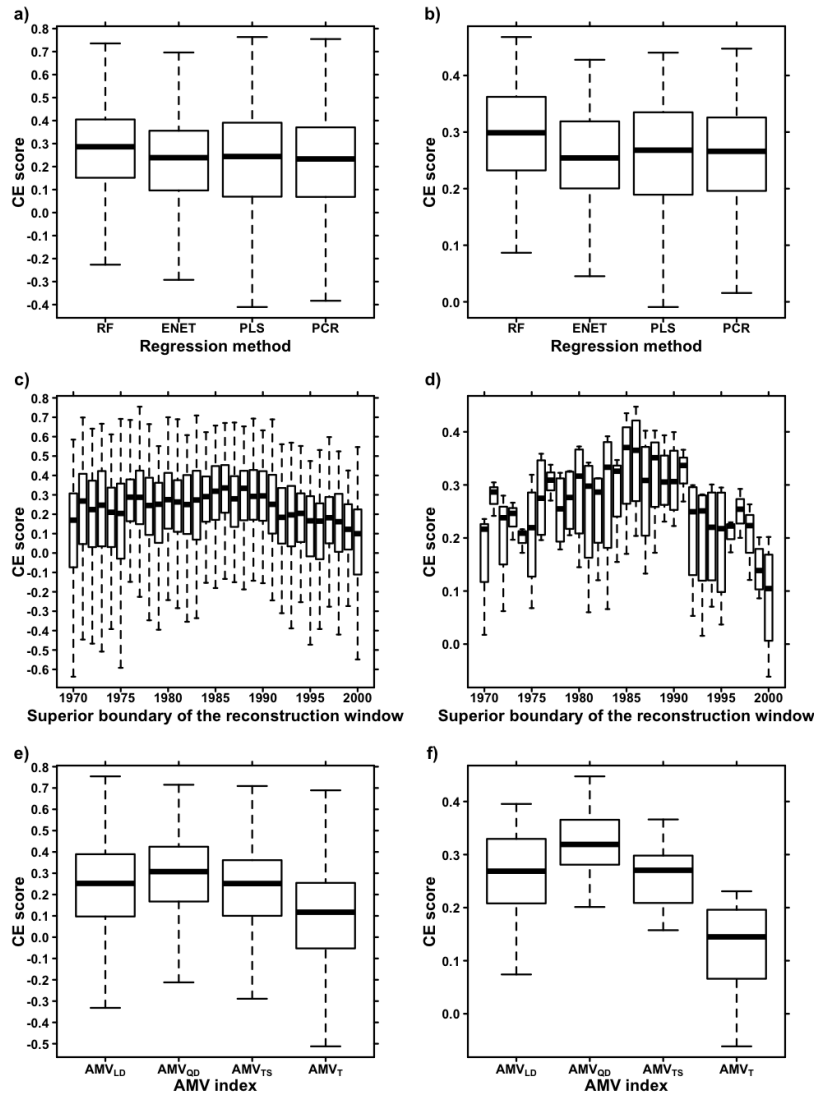


FIGURE 4.6 | Extended Data Figure 2 - Score by level of inputs for the 496 reconstructions. Coefficient of Efficiency scores (Nash and Sutcliffe 1970) for the 496 reconstructions compared in this study for the different sources of methodological choices (regression method, reconstruction frame and AMV index). a) and b) give the CE scores by regression methods (496/4=124 reconstructions by method). c) and d) give the CE scores by superior boundary of the reconstruction window (496/31=16 final reconstructions by window). e) and f) give the CE scores by AMV index (496/4=124 final reconstructions by index). a), c) and e) give the CE scores for all the training splits (496*30=14,800 scores) b), d) and f) give CE scores for each final reconstruction as the medians of the 30 corresponding individual reconstructions (496 scores). For all boxplots, median is shown as heavy dark line. Boxplots edges give first and third quartiles. Boxplot “whiskers” give the 10%-90% range. Outliers are not shown. A point from a boxplot is here considered as outlier when it is outside 1.5 times the interquartile range above the upper quartile and below the lower quartile.

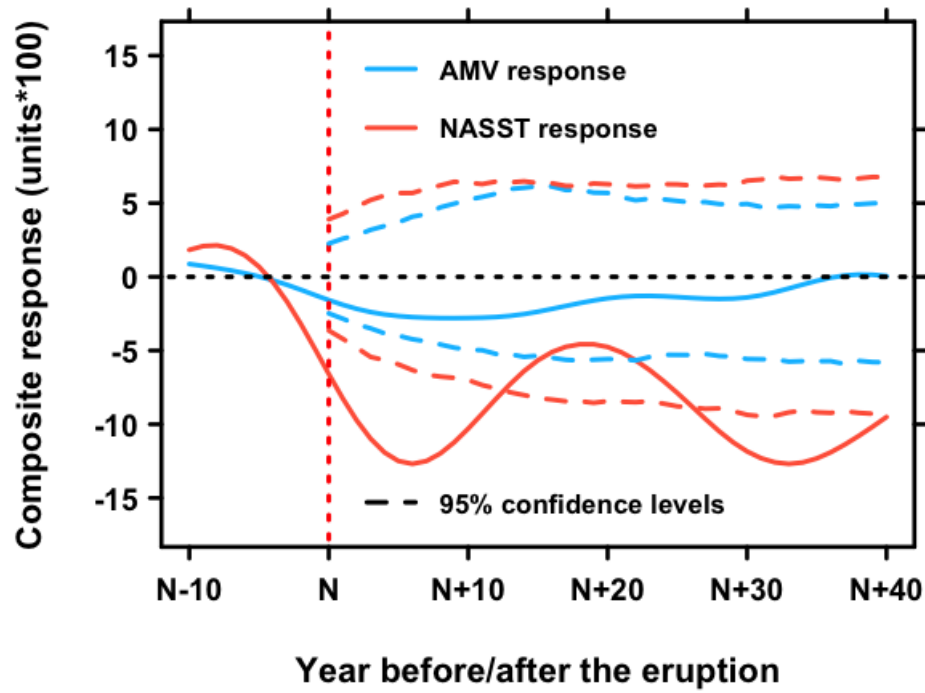


FIGURE 4.7 | Extended Data Figure 3 - NASST and AMV response to the 10 largest volcanic eruptions of the last millennium. Superposed epoch analysis of the AMV (blue line) and NASST (red line) responses to the eleven strongest volcanic eruptions of the last millennium (Sigl et al. 2015). 95% confidence levels with respective colors are calculated using Monte-Carlo approach (Rao et al. 2019).

Extended data tables

Voir Tab. 3.2 pour le tableau "Extended data table 1".

BCC-CSM1-1	GISS-E2-R-CC
BCC-CSM1-1-M	GISS-E2-R
BNU-ESM	HadCM3
CCSM4	HadGEM2-AO
CESM1-BGC	HadGEM2-CC
CESM1-CAM5-1-FV2	HadGEM2-ES
CESM1-CAM5	IPSL-CM5A-LR
CNRM-CM5	IPSL-CM5A-MR
CSIRO-MK3-6-0	IPSL-CM5B-LR
CanESM2	MIROC-ESM-CHEM
EC-EARTH	MIROC-ESM
FIO-ESM	MIROC5
GFDL-CM3	MPI-ESM-LR
GFDL-ESM2G	MPI-ESM-MR
GFDL-ESM2M	MRI-CGCM3
GISS-E2-H-CC	MRI-ESM1
GISS-E2-H	NorESM1-ME
	NorESM1-M

TABLE 4.1 | Extended Data Table 2 - List of the 37 models used for the estimation of the NASSTs forced component.

Volcano	Location	Year of the eruption
Unknown	Unknown	1108
Unknown	Unknown	1171
Unknown	Unknown	1230
Samalas	Indonesia	1258
Kuwae	Vanuatu	1458
Huaynaputina	Peru	1601
Parker	Phillippines	1641
Lakagigar	Iceland	1783
Unknown	Unknown	1809
Tambora	Indonesia	1815

TABLE 4.2 | Extended Data Table 3 - 10 last millennium largest volcanic eruptions (Sigl et al. 2015).

4.2.4 Methods

Instrumental AMV indices :

In this study, historical AMV indices have been calculated using annually-resolved values. The AMV reconstructions compared in this study are cubic spline smoothed reconstructions of these annually-resolved historical AMV indices. Additional results (not shown) highlight that using annually resolved indices provides better validation scores for the reconstruction than using the smoothed versions of the AMV indices investigated.

We have computed over the period 1900-2017 the four instrumental AMV indices (Enfield et al. 2001, Trenberth and Shea 2006, Ting et al. 2009, Enfield and Cid-Serrano 2010) of this study using the Hadley Center Global Sea Ice and Sea Surface Temperatures (HadISST) dataset (Rayner et al. 2003), where NASST is calculated as the spatially averaged SST over the grid points belonging to the North Atlantic along (0-60°N).

AMV_{LD} (Enfield et al. 2001) and AMV_{QD} (Enfield and Cid-Serrano 2010) indices are respectively calculated by removing a linear and a quadratic trend from the observed NASST. It was pointed out that external forcing can be more complex in its variations over time (Trenberth and Shea 2006), thus suggesting to use the global averaged SST anomalies as a proxy for the externally forced signal that is subtracted from NASST, resulting in the AMV_{TS} (Trenberth and Shea 2006) (Fig. 4.1a). This approach also has limitations since it does not account for regional variations in external forcings, such as the distribution of anthropogenic aerosols, which was denser in North America-Europe at the beginning of the last century and became more prominent in Asia in the recent decades (Smith et al. 2016). To account for this, the AMV_T (Ting et al. 2009) index is built using climate model historical simulations to isolate the forced component of NASST calculated by signal-to-noise maximizing empirical orthogonal function (EOF, Fig. 4.5 - Extended Data Fig. 1), which is then removed by a linear estimate of its effect (Ting et al. 2009).

Estimation of NASSTs forced component using signal-to-noise maximizing EOF :

The forced component of NASST is estimated from historical simulations of 37 climate models (Tab. 4.1 - Extended Data Table 2). For each, we extract the NASST anomalies (0-60°N over the Atlantic) that are merged as the columns of the same matrix. Using a Principal Component Analysis of the latter matrix, the first Principal Component associated to the first EOF is retained as the estimated forced component of NASST which is presented in Fig. 4.5 - Extended Data Fig. 1. For additional details, the reader is referred to the AMV_T related publication (Ting et al. 2009).

Generation of the 496 reconstructions compared in the study :

The 496 reconstructions compared in this study, are performed for 31 timeframes Γ : from 850-1970 to 850-2000 by incrementing the superior boundary by one year. We use this technique to generate different reconstruction, as for the shortest reconstruction periods (with the lowest superior boundary), more proxies are available but less observations of each are used to establish relationships between proxies

and AMV indices. On the opposite, using the longest period (with the highest superior boundary), leads to use less proxies, but provides a longer overlap period between proxies and AMV indices. These 31 frames are shuffled with the four AMV indices (Enfield et al. 2001, Trenberth and Shea 2006, Ting et al. 2009, Enfield and Cid-Serrano 2010) and the four regression methods (Hotelling 1957, Wold et al. 1984, Breiman 2001, Zou and Hastie 2005) presented above. All these setups are tested by only using proxies available and significantly correlated at the 95% confidence level with the respective AMV index. We thus end up with $31 \times 4 \times 4 = 496$ final reconstructions that are compared in this study using the CE metric. For the NASST reconstruction, 124 setups are compared, by shuffling the same 31 timeframes and the same 4 regression methods than the AMV indices.

Nested reconstruction

In this study, the best reconstruction found is the reconstruction of the AMV_{QD} index with the random forest method over the period 850-1989 by using the proxy records significantly correlated at the 95% confidence level with the AMV_{QD} index over the training periods (see Fig. 4.6 - Extended Data Fig. 2).

To overcome this limitation, using the same set of methodological choices, we perform 1001 reconstructions for the periods 850-1989 to 1850-1989 with an increment of one year for the inferior boundary, thereby allowing to use an increasing number of proxy records as soon as the reconstruction period is reduced. The nested reconstruction (*e.g.* the reconstruction presented in this study), is obtained by placing one after the other the 1001 interlocking reconstructions. This way, 102 proxy records are used for the reconstruction presented in this study

Random forest variable importance

The weights of the proxy records used for the nested reconstruction are presented in Fig. 4.2c. Those weights have been calculated using the random forest variable importance (Breiman 2001). Different importance metrics exists, in this study, we use the Gini importance or Mean Decrease in Impurity (MDI). MDI of a given proxy record is calculated as the sum of the number of splits where it is used across the trees (see Supplementary Information for details on the regression methods), proportionally to the numbers of split samples in all trees (*cf.* Supplementary Information). For Fig. 4.2c, each proxies importance over the 1001 reconstructions are summed with a weight of n/N , where n is the number of available proxies for a given time step, and N the total number of proxies used at the end for the reconstruction (*i.e.* $N=102$). Finally, Fig. 4.2c is computed by calculating each proxy importance as a fraction of the previously calculated importance and the total importance over the 1001 reconstructions. The same is done in Fig. 4.3b, but each importance is also averaged over the PPE performed on the 12 members of CESM1-LMLE used.

Pseudo-proxy experiment (PPE)

The nested reconstruction is tested for 12 members of the National Center of Atmospheric Research (NCAR) Community Earth System Model 1 (CESM1) Last Millennium Large Ensemble (LMLE) experiment. The fact that the calculation of

AMV_{QD} is time-dependent implies that for a given CESM1-LMLE member, the quadratic trend of the NASSTs calculated over the respective periods 850-1989 and 1900-1989 are different. This explains the inconsistency of the two estimates of AMV_{QD} over their common time frame, which means that the PPE can not be properly tested for the real experiment. To overcome this problem, the AMV_{TS} index is used in the PPE as its calculation is not time dependent. Using this index, we can properly compare the reconstructed AMV_{TS} index in the PPE from the period 1900-1989 to the period 850-1982, and the model AMV_{TS} over the period 850-1989. For the rest of the methodological setup, the same reconstruction method is performed than in the real experiment.

To mimic the proxy records in each member, the surface air temperatures at two meters (T2M) is used. The AMV_{TS} is extracted from the SST grid as the difference between the extracted NASST anomalies (0-60°N over the Atlantic) and the extracted global SST anomalies. As for the real reconstruction, we use the proxy records significantly correlated with the AMV at the 95% confidence level for the reconstruction in the PPE to avoid local scale model biases. As for the real experiment, the PPE is performed for the 1001 periods of the nested reconstruction that are then aggregated over their common window. The CE scores presented in Fig. 4.2b are the merged scores obtained for the 30 training-testing splits of the 1001 members of the nested reconstruction.

Early warning signal test

We base our approach on methods for the detection of incoming climate tipping point, recently validated for an AMOC slowdown within a general circulation model. The AMV reconstruction is firstly smoothed using a Kernel Gaussian filtering with a bandwidth of size 100 years. The latter is then regressed onto de former. AR(1) coefficients of the residuals from this regression are calculated for different sliding window length $WL=200,250,300,350,400$ years (Fig. 4.4b). Kendall τ is calculated for each of the AR(1) coefficient series. Contrary to a former study focusing on early warning signal applied to model-based investigation of an AMOC collapse (Lenton 2011, Boulton et al. 2014), we cannot use a Monte-Carlo estimate of the significance of Kendall τ . This is because they compare their statistic with other Kendall τ calculated from simulations where no AMOC collapse happens. Here, our reconstruction does not come from model simulation and thus such an approach cannot be applied. We thus use the classical Kendall τ significance test (statistics and p-values shown Fig 4.4b, see “Statistical information”).

Boxplots

For all boxplots of the study, median is shown as heavy darkline. Boxplots edges give first and third quartiles. Boxplot “whiskers” give the 10%-90% range. For better graphical representations, outliers are not shown. A point from a boxplot is here considered as outlier when it is outside 1.5 times the interquartile range above the upper quartile and below the lower quartile.

Statistical information

- Fig. 4.1bc For each grid point, a bilateral Student test is applied to the regression coefficients between the corresponding climate variable and the AMV indices. The degrees of freedom are corrected using timeseries autocorrelations as in McCarthy et al. (2015) and Michel et al. (2020).
- Fig. 4.4b The Kendall rank correlation coefficient, or Kendall τ coefficient, measures the ordinal association between two quantities, here AR(1) coefficients denoted $(x_i)_{1 \leq i \leq n}$ here, and time denoted $(y_i)_{1 \leq i \leq n}$. The statistic is given by :

$$\tau = \frac{n_c - n_d}{n_0} \quad (4.1)$$

Where considering $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ the ensemble of joint pairs, $\forall (i, j) \in \llbracket 1, n \rrbracket$:

$$n_c = \#_{i \neq j} \{ \{x_i > x_j \cap y_i > y_j\} \cup \{x_i < x_j \cap y_i < y_j\} \} \quad (4.2)$$

$$n_c = \#_{i \neq j} \{ \{x_i < x_j \cap y_i > y_j\} \cup \{x_i > x_j \cap y_i < y_j\} \} \quad (4.3)$$

$$n_0 = \frac{n(n-1)}{2} \quad (4.4)$$

For large sample ($n > 50$), as in this study, the distribution is approximated with a Gaussian distribution of mean 0 and variance $\frac{2(2n+5)}{9n(n-1)}$, under the null hypothesis $H_0 : " \tau = 0 "$ which is tested against the alternative hypothesis $H_1 : " r \neq 0 "$. The p-value (shown Fig. 4.4b) of the test is deducted from the quantile of this distribution.

- In this study, each reconstruction is performed using a selection of the proxies significantly correlated at the 95% confidence level with the corresponding AMV index. The same bilateral Student test for correlation than McCarthy et al. (2015) and Michel et al. (2020) is applied for the whole study, with corrected degrees of freedom using timeseries autocorrelation. The other p-values of correlations presented in this study are also based on this test, including Fig 6.3e.

4.3 Implications des résultats obtenus

Un résultat important de cette étude est que nous avons pu montrer qu'en retirant le forçage des enregistrements paléoclimatiques sur la période historique, on parvient à reconstruire un indice AMV insensible aux forçages solaires (non montré dans l'étude) et volcaniques (Fig. 4.7). À l'inverse, lorsqu'on applique la sélection de modèles statistiques (chapitre 3) à la reconstruction de la NASST à partir des enregistrements bruts, on obtient une réponse radiative significative de 10 ans en moyenne (Fig. 4.7), et ce pour les 11 éruptions les plus puissantes du dernier millénaire (Tab. 4.2). C'est le résultat auquel on devrait s'attendre. En effet, la NASST peut être directement affectée par un changement de forçage radiatif lié aux éruptions volcaniques, tandis que l'AMV étant, *a priori*, plutôt liée aux processus de variabilité interne du climat telle que la circulation océanique, on s'attend à une réponse moindre les quelques années suivant une éruption.

La reconstruction d'AMV obtenue par une imbrication de modèles de RF utilise un nombre satisfaisant d'enregistrements (102, Fig. 4.2cd). Un autre résultat important est qu'elle se corrèle significativement au seuil d'au moins 90% avec des enregistrements océaniques, non-utilisés pour la reconstruction. Au delà de la méthode employée qui se veut objectivement optimale (chapitre 3), la corrélation significative de la reconstruction avec 43 enregistrements océaniques indépendants, dont 22 proviennent de l'Atlantique Nord, apporte une confiance supplémentaire en la robustesse de la reconstruction (Fig. 4.2e). En terme de validation supplémentaire, l'approche par *pseudo-proxy* a également permis de montrer qu'au sein d'un modèle de circulation globale, la méthode de reconstruction, avec l'échantillonnage d'observations disponibles, reconstruit un indice AMV qui se corrèle significativement avec l'AMV calculée indépendamment dans le modèle. (Fig. 4.3). Toutes ces approches de validations complémentaires sont importantes pour donner de la valeur à l'ensemble de la méthodologie présentée dans le chapitre 3, ce qui permet de la valider, pour le cas de l'AMV, autant empiriquement que physiquement.

Au-delà des différentes validations de reconstruction que l'on propose, cette étude présente par ailleurs trois résultats importants.

1. Le premier est l'entrée précoce (dans les premiers siècles du millénaire) dans une phase froide de l'Atlantique Nord, suggérant que la séquence du MCA dans le bassin Nord-Atlantique se termine aux alentours de 1150. Ce changement climatique survient notamment avant le groupe d'éruptions du XIII^e siècle. De plus, les deux séries reconstruites, AMV et NASST, passent toutes deux en phase négative à ce moment là (Fig. 4.4a), ce qui soutient l'idée que cette transition n'est pas due aux forçages externes. En revanche, on observe une succession importante de phases froides dans la NASST à partir de 1450, pour laquelle on voit une variation négative du même ordre que celle de 1150. Dans le même temps, l'AMV reconstruite (insensible aux forçages externes) ne possède pas de réponse particulière à ce moment là, ce qui soutient que le déclin de l'indice NASST aux alentours de 1450 est très certainement une réponse à l'éruption du volcan Kuwae en 1458 (Vanuatu, caldeira située entre les îles *Epi* et *Tonga*), considérée comme la deuxième plus puissante du dernier millénaire (Sigl et al. 2015). Par la suite, la NASST reste dans un état significativement négatif jusqu'en 1900, à l'inverse de l'AMV qui se trouve dans un état neutre. Ces différents résultats suggèrent donc que le LIA peut très certainement être séparé en deux périodes dont les causes sont distinctes : une première d'environ 1150 à 1350, initialisée par un changement dans les régimes de circulation océanique, très probablement aggravée et prolongée par les éruptions du XIII^e siècle; et une seconde d'environ 1450 à 1900, essentiellement causée par une forte activité volcanique (Sigl et al. 2015, Brönnimann et al. 2019).
2. Le second résultat montre que le spectre de l'AMV est bien plus large que celui estimé par Kerr (2000) (40-80 ans), et celui simulé par la plupart des modèles CMIP5 (Lin et al. 2019) (10-30 ans). On trouve ici une périodicité de l'AMV qui fluctue de manière non stationnaire sur des échelles de 20 à 80 ans (4.4c). Ces différentes échelles de temps peuvent être expliquées par des processus de variabilité interne uniquement selon les simulations de Frankcombe et al. (2010), ce qui soutient la cohérence de notre résultat.

3. Le troisième et dernier résultat important de notre étude concerne également les propriétés spectrales de l'AMV reconstruite. En se basant sur les travaux de Lenton (2011), qui propose un ensemble de tests pour la détection des points de bascule climatique, on trouve une statistique de test de "ralentissement critique" très significative (>99% quelle que soit la fenêtre de temps étudiée) une fois appliqué à la reconstruction de l'AMV que l'on propose (4.4c). Ce test semble robuste puisqu'il a notamment été appliqué avec succès pour la détection d'un ralentissement d'AMOC au sein d'un modèle climatique (Boulton et al. 2014). En matière de données réelles, ce test a de plus permis d'identifier un signal précoce au sein de carottes de glace, annonçant l'entrée dans la dernière déglaciation (Lenton et al. 2012). Sous l'hypothèse que notre AMV reconstruite possède une part importante de variabilité commune avec l'AMOC sur le dernier millénaire, ce résultat vient renforcer les hypothèses d'un refroidissement futur de l'Atlantique Nord suite à une diminution de la convection dans la région du SPG (Sgubin et al. 2017), potentiellement signe d'un changement d'état moyen de l'AMOC au long terme (Lenton 2011, Boulton et al. 2014). Ceci va plus loin que le ralentissement *très probable* suggéré par le GIEC (Collins et al. 2019), puisque le test proposé par Lenton (2011) est voué à déterminer l'approche d'un changement significatif de l'état moyen de l'AMOC au delà de son ralentissement. Notons que le fait que notre reconstruction permet de disposer d'une période suffisamment longue pour appliquer le test de Lenton (2011), et d'estimer un comportement futur du climat. Ceci souligne le grand intérêt que peuvent avoir des travaux de paléoclimatologie pour l'étude du changement climatique anthropique actuel.

Cette étude a en particulier permis de voir que l'approche la plus robuste pour reconstruire l'AMV s'avère être la méthode RF, tout comme pour la NAO dans le chapitre 3. Ceci est vrai pour la reconstruction ayant le plus haut score, mais est également vrai pour la moyenne des scores obtenus pour les 124 configurations de reconstructions testées pour chacune des méthodes (4.6). En revanche, on remarque que pour la reconstruction de la NASST, la reconstruction optimale s'avère être obtenue avec la méthode PCR, ce qui montre qu'il n'y a pas de méthode universellement meilleure que les autres sur toutes les configurations de reconstructions paléoclimatiques possibles (*cf.* théorème du *No free lunch*), justifiant également l'approche développée dans le chapitre 3 et utilisée dans ce chapitre. Néanmoins, les résultats prometteurs de la RF des chapitres 3 et 4 sont tels qu'on pourrait imaginer l'utiliser dans un cadre plus global en matière de reconstruction de signaux climatiques. Le prochain chapitre sera voué à étendre cette méthode de reconstruction sur l'ensemble de la grille de données historiques de SST (HadISST, Rayner et al. 2003). On cherchera alors à donner une interprétation encore plus exhaustive des variations climatiques du dernier millénaire.

Chapitre 5

Reconstruction d'une grille de SST basée sur une méthode de forêt aléatoire

Ce chapitre présente la reconstruction d'une grille de SST en étendant la méthode de RF avec une sélection d'enregistrements et une optimisation de paramètres pour chaque point de grille. La section 5.1 explique en quoi notre méthode se distingue de ce qui a été fait précédemment et présente l'intérêt de la démarche entreprise. La section 5.2 est une étude complète dont l'article sera incessamment soumis au journal *Climate Dynamics*. La section 5.3 résume les résultats obtenus et discute les assertions proposées.

5.1 Objectifs et motivations de ce chapitre

Nous présenterons l'extension d'une grille quasi-globale de SST de résolution $1^\circ \times 1^\circ$ jusqu'à l'année 850 C.E. et un ensemble d'analyses statistiques sur la dynamique océanique du dernier millénaire. Les travaux présentés dans ce chapitre constituent la partie la plus aboutie de cette thèse et de l'application de méthodes de *machine learning* telle que la RF.

Ici, on se distingue à nouveau des études précédentes par la méthode statistique employée. En effet, c'est la première fois que la méthode de RF, qui plus est optimisée par validation croisée, sera appliquée à l'ensemble d'une grille de température. Parmi les études précédentes, on peut citer celle de Mann et al. (2008), qui a longtemps fait référence. Ces derniers s'appuient sur un algorithme dit de maximisation de l'espérance appliqué dans un cadre de régression linéaire. Des reconstructions plus récentes de grilles de températures utilisent, elles, soit des régressions linéaires simples entre un point de grille et l'enregistrement le plus proche (Hakim et al. 2016), soit des moyennes des enregistrements proches du point de grille (Abram et al. 2016). Notre approche, couplant la RF avec des approches d'évaluation et d'optimisation, semble en général plus efficace que les méthodes usuelles employées pour la reconstruction paléoclimatique (chapitres 3 et 4).

Ce qui distingue particulièrement notre démarche des autres études est la méthode de RF, mais aussi l'évaluation de chaque reconstruction que l'on produit au travers de la méthodologie du chapitre 3, qui se veut la plus objective possible. Ainsi, tous les points pour lesquels les scores ne permettent pas de dire que la reconstruction est significativement robuste ne seront pas considérés dans l'analyse. Un fait intéressant est que cette reconstruction de grille a également été appliquée à

une grille historique de T2m continentale (GISTEMP, Lenssen et al. 2019), et ce, sans succès selon notre méthodologie d'évaluation (chapitre 3). En effet, nous trouvons que seulement 28,5% des points de grilles de ce jeu de données ont des scores significativement supérieurs à 0 (chapitre 3). À l'inverse, on obtient sur la reconstruction de SST près de 87,4% de points de grille dont les scores sont significativement supérieurs à 0. Autrement dit, les enregistrements à résolution uniquement annuelle que l'on utilise, et dont seulement quelques uns sont des coraux (chapitre 3 et 4), sont bien plus efficaces pour reconstruire des points de grille de températures océaniques de surface que pour les points de grille continentaux. Ce résultat est relativement étonnant et contre-intuitif puisqu'on aurait pu s'attendre à mieux reconstruire les points de grilles continentaux avec une majorité d'enregistrements paléoclimatiques terrestres. Une autre manière de le dire du point de vue de la méthode que l'on utilise est que le lien statistique que l'on établit entre les données historiques et paléoclimatiques est bien plus fort avec les données historiques océaniques que les données continentales. On pourrait émettre l'hypothèse que les variations des enregistrements terrestres, souvent représentatives d'une saison, sont mieux représentées dans les moyennes annuelles de SST que de T2m, de par la forte inertie de l'océan. Quoi qu'il en soit, les scores bien plus importants sur l'océan que sur les continents expliquent le choix de ne pas avoir présenté de reconstruction de températures continentales dans ce qui suit.

L'ensemble des analyses que nous présenterons à la suite de l'évaluation de la grille de reconstruction obtenue permettent également de nous distinguer des autres études puisque nous parcourons un ensemble de sujets distincts en matière de dynamiques et de réponses aux forçages, et ce, autant sur un plan spatial qu'au travers de modes océaniques connus. On s'intéressera alors à de nombreuses questions, en partie présentées dans la section 1.3, et pleinement explicitées dans l'introduction de la section suivante.

5.2 Article soumis au journal *Climate Dynamics*

Title : Analysis of last millennium ocean variability through a gridded random forest-based SST reconstruction.

Simon Michel¹, Didier Swingedouw¹, Juliette Mignot²

1 : Environnements et Paleoenvironnements Océaniques et Continentaux (EPOC), UMR CNRS 5805 EPOC-OASU-Université de Bordeaux, Allée Geoffroy Saint-Hilaire, Pessac 33615, France.

2 : Sorbonne Universités (UPMC, Univ. Paris 06)-CNRS-IRD-MNHN, LOCEAN Laboratory, 4 place Jussieu, F-75005 Paris, France.

Abstract :

Intrinsic variability of the ocean circulation and its response to external forcings is poorly known beyond the decadal scale due to the short time length of available instrumental observations. Sea surface temperatures (SST) data indicate pronounced multidecadal variability in regions such as the North Atlantic, which variations are linked to several observed large-scale climate fluctuations.

The Atlantic Meridional Overturning Circulation (AMOC) is believed to be a major element in shaping this decadal to multi-centennial SST variability. However, the 150 years of SST observations are far too short to correctly characterize this variability. Using a machine learning approach (random forest) applied to the state-of-the-art PAGES 2k paleoclimate database, the present study proposes a novel investigation of decadal to multidecadal SST variability through the production of a new gridded, and cautiously verified, SST reconstruction going back to 850 C.E. Results indicate a significant role of natural forcings on SST variability, with a particularly strong effect of volcanic eruptions. In terms of internal variability, results show that positive phases of the Atlantic Multidecadal Variability are significantly associated with a decadal negative response of the Interdecadal Pacific Variability, indicating, based on a long-enough observational dataset, that the North Atlantic might have played a key role in shaping the recently observed hiatus of global SST between 1998 and 2012. Finally, an extracted and well-verified AMOC index of the last millennium is showing a significant negative trend since the 1940s, which has reached an unprecedentedly low state these last few decades since at least 850 C.E.

5.2.1 Introduction

Understanding decadal climate variability is key in progressing on actual and future climate variations (Hawkins and Sutton 2009, Cassou et al. 2018), since it leads to large uncertainty in model projections notably at the continental to regional scales (Deser et al. 2012, Deser et al. 2020). In this context, a focus on the ocean is crucial as its impacts on climate variations at decadal to longer timescales have been evidenced in both general circulation models' simulations and instrumental observations (Schlesinger and Ramankutty 1994, Mann and Park 1994, Knight et al. 2006, Swingedouw et al. 2013). The study of the ocean in three spatial dimensions is essential for understanding global ocean circulation. Yet, it requires complex and costly measurements. As part of the large-scale oceanic circulation, the Atlantic Meridional Overturning Circulation (AMOC, Delworth and Mann 2000, Knight et al. 2006) is responsible for a large amount of meridional heat transport towards the north, bringing heat from the southern hemisphere and influencing the global climate (Buckley and Marsahll 2016). The main mechanism at the origin of the AMOC is oceanic convection. The latter is one of the few processes allowing exchanges between upper and deeper ocean layers, which are usually highly stratified due to density gradients, determined by temperature and salinity. It occurs in specific locations of the northern North Atlantic notably, when the vertical density balance between upper and deeper layers reverses, which directly impacts the AMOC strength. This strength thus regulates the advective transport of tropical Atlantic warm waters to the northern North Atlantic and Arctic regions. Studying the AMOC is crucial (Collins et al. 2019) as it is thought to be a tipping element of the climate system, able to undergo rapid and abrupt changes. The anomalous anthropogenically-induced melting of Greenland ice-cap and Arctic sea ice, gathering a significant input of freshwater in upper layers of North Atlantic convective regions, as well as increase in precipitations and runoff, is indeed expected to significantly weaken the AMOC (Swingedouw et al. 2007, Bakker et al. 2016).

It is only since the beginning of the RAPID project in 2004 that direct observations of the AMOC have become available (Cunningham et al. 2007). The few

observations of AMOC intensity at 26°N over the Atlantic show that its annual mean is significantly weaker from 2008 to 2017 than between 2004 and 2007 (Smeed et al. 2018). However, 14 years of AMOC observations are not enough to determine if this is a long-lasting abrupt decline. To study and understand AMOC variations on multi-decadal time scales from observations, most studies thus use paleo-reconstructions (Rahmstorf et al. 2015, Swingedouw et al. 2015, Thornalley et al. 2018, Ayache et al. 2018), models (Zhang 2008, Persechini et al. 2013, Escudier et al. 2013, Ortega et al. 2017) or instrumental fingerprints (Caesar et al. 2018). The first approach has notably suggested that the recent weakening of ocean circulation is possibly already unprecedented since up to two millennia (Rahmstorf et al. 2015, Thornalley et al. 2018). However, the sea surface temperature (SST) fingerprint-based AMOC index from Rahmstorf et al. 2015, has been extracted from a quite old and low-resolution surface temperature gridded reconstruction (Mann et al. 2008). It also does not account for the large amount of recently gathered proxy records (PAGES 2k Consortium 2017), and thus needs to be validated using up-to-date reconstructions. Based on the same SST fingerprint-based AMOC index, a recent study suggests that this weakening trend started in the 1960s (Caesar et al. 2018). Instrumental observations of the ocean surface indeed date from much further back (around 1870) than for deeper layers. SST is the longest instrumental oceanic indicator available to understand the behavior of ocean circulation from direct observations. These observations show a strong and coherent multidecadal SST variability in the North Atlantic ocean, named the Atlantic Multidecadal Variability (AMV), suspected to be driven by AMOC variability (Muir and Fedorov 2015, Yan et al. 2018, Zhang et al. 2019a). A recent reconstruction of an AMV index (Michel et al. rev) is not showing such a trend over the historical period. Nevertheless, an analysis of early warning signals (Lenton 2011, Boulton et al. 2014) in this AMV time series shows that it has recently experienced abnormally high amplitudes and low frequency as compared to the previous thousand years (Michel et al. rev), which is an indication for a potential bifurcation in the AMOC. On a lower spatial extent, a multi-model study found that 5 out of the 11 best models from CMIP5 in terms of observed Subpolar Gyre (SPG) stratification representativity simulate an abrupt convection collapse in the SPG (Sgubin et al. 2017). As SPG is only a component of the large scale AMOC phenomenon, a convection collapse in this region only is evaluated to lead to a significant AMOC weakening, but not a complete disruption since convection in the Nordic Seas might remain (Sgubin et al. 2017). Although there is still not enough evidence that the observed AMOC decline is unprecedented since two millennia, there is a large consensus for a 21st century AMOC weakening. It is evaluated to be *very likely* to occur even for very optimistic scenarios in terms of reduction in greenhouse gas emissions (Collins et al. 2019). On the other hand, a complete AMOC collapse is *very unlikely* to occur by 2100 (Collins et al. 2019).

Oceanic surface modes, such as the AMV and the Interdecadal Pacific Variability (IPV) have been characterized using Principal Component Analysis of observed SST (Nigam et al. 2011, Henley et al. 2015). Both of these modes are suspected to be key drivers of the recent 1998-2012 hiatus period observed in both global surface air temperature and SST (England et al. 2014). This period is indeed marked by no discernible warming at the global scale with a trend reaching the lower end of the range of simulations from CMIP5 models (Wanatabe et al. 2013). However, since atmospheric levels of greenhouse gas concentration are experiencing continuous growth, internal and natural climate variability have been mainly investigated as responsible for

this slowdown in the observed global temperature increase. Indeed, the 1998-2012 hiatus period correspond to strong negative anomalies over eastern equatorial Pacific, typically in Niño 3 and 4 regions, but also in the whole eastern tropical Pacific, and positive anomalies in northern and southern high latitudes Pacific Ocean. This pattern typically characterizes a specific tripole in Pacific SST, witnessing a negative phase of the IPV (Henley et al. 2015). Such IPV phases have thus been identified as key for the occurrence of hiatus periods (Kosaka and Xie 2013, England et al. 2014).

Potential drivers leading to such multidecadal-long negative IPV patterns have been extensively investigated. In terms of natural forcings, a single model-based study has suggested that volcanic eruptions has a higher probability to lead negative anomalies in the Niño 3.4 region (5°S-5°N, 170°W-120°W) the year following the eruption (McGregor and Timmermann 2011), which might enhance the probability for a negative IPV-like state in the next years. This result is not consensual as other studies using observations and models have shown that eastern equatorial Pacific SST rather has positive anomalies the year following the eruption due to weakening Walker circulation and diminished equatorial upwelling (Ohba et al. 2013, Khodri et al. 2017). These discrepancies may partly be due to the lack of a sufficient number of large stratospheric eruptions over the historical period to have a robust statistical view of the Pacific response to this kind of forcing. On the other hand, Kosaka and Xie (2013) have shown that internal variability is able to explain such a long-term negative IPV state. They proposed a positive feedback mechanism implicating a zonal SST gradient between western and central to eastern equatorial Pacific, and a strengthened Walker circulation, responsible for equatorial cold waters upwelling (Kosaka and Xie 2013). Their suggestion is indeed consistent with observed anomalously strong easterly winds over this region during the hiatus period (England et al. 2014). Complementary studies suggest that the strengthened Walker circulation and negative equatorial anomalies in the Pacific, might have been caused by the anomalously warm North Atlantic (*i.e.* positive AMV) since the early 1990s (McGregor et al. 2014, Chafik et al. 2016). Such a lagged negative IPV response to a positive AMV is indeed significant over the whole period of instrumental observations (Marini and Frankignoul 2013), which however remain a short timeframe to evaluate such multi-decadal scales.

In spite of all these advances, studying decadal to multi-decadal climate variability only using instrumental data is indeed subject to serious issues of robustness. For instance, the internal climate mode AMV seems to have a 60-80 years cycle over the instrumental observations (Kerr 2000). However, there is no way to calculate any significance for such a peak in the spectrum since not even two cycles of the mode have been observed, which is not enough to build robust statistical conclusions. Indeed, extending AMV indices since 850 C.E., Michel et al. (rev) found a much wider spectrum of preferential variability (20-80 years), which can be entirely explained by internal climate processes, as demonstrated by a prior modeling study (*e.g.* Frankcombe et al. 2010). This issue in terms of robustness can be extrapolated to the wide variety of observation-based statistical estimations of multidecadal oceanic modes, variability and relationships. Reconstructions of indices of oceanic modes using large amounts of proxy records covering the last millennium provide far more robust estimates of the spectral properties, response to natural forcings, and implications in producing large scales climate patterns such as the Medieval Climate Anomaly (MCA) and the Little Ice Age (LIA) (Li et al. 2011, Wang et al. 2017, Michel

et al. rev). Gridded reconstructions of climate variables such as SST or surface air temperature are able to bring deeper understandings of decadal climate variability by adding the essential information concerning the spatial dimensions and stability of the modes. Previous gridded reconstructions of SST data have highlighted key results such as the unprecedented growth in historical temperatures as compared to the past two millennia (Mann et al. 2008, Abram et al. 2016), and the evidence that the MCA and LIA sequences were not as globally coherent as the recent anthropogenic warming (Neukom et al. 2019). These former studies, while sometimes testing different reconstruction techniques (Neukom et al. 2019), are often restricted to linear combinations of proxies (or modified versions of proxies). Non-linear techniques such as Random Forest (RF) have been recently applied to index-based reconstructions. It appears to generally overclass most of the widely used regression techniques in index reconstruction over the validation period (Michel et al. 2020) but also in pseudo-proxy approach (Michel et al. rev), and could thus be promising for gridded ones.

The present study focuses on providing a reconstruction of the SST field since 850 C.E. based on an RF approach. Section 5.2.2 describes the data and techniques used for producing the gridded SST reconstruction studied in the following sections. Section 5.2.3 analyses the SST reconstruction. In particular, the reconstruction is first validated using metrics and regression uncertainties. Secondly, the effects of natural forcings, including the strongest volcanic eruptions and long-term solar minima are assessed. The most contrasting climate sequences are also identified, and the well-known Pacific and North Atlantic modes are explored. Finally, the reliability of SST based AMOC indices is assessed and the relationship between AMV and IPV is estimated. Sections 5.2.4 provides discussions and conclusions of this study.

5.2.2 Method : Random forest-based gridded SST reconstruction applied to PAGES2K

5.2.2.1 Data used

This study uses instrumental SST data from the HadISST dataset (Rayner et al. 2003) and an annually-resolved proxy records database (Michel et al. rev). The latter contains the annually-resolved records from the PAGES 2k database (PAGES 2k Consortium 2017) (509 records) complemented by 41 annually-resolved proxy records from the Arctic, North America, and Europe (Michel et al. rev) (550 records in total). Because sea ice is associated to a temperature of -1.8°C , we noted (not shown) that the use of proxies located in areas possibly covered by sea ice along the year is biasing the model fitting for annual SST averages : this leads to reconstructions with inconsistent annual SST under -1.8°C . For this reason, we do not consider southern hemisphere grid points of historical SST dataset for which at least one monthly observation of sea ice fraction is higher than 0% over the historical period. For the northern hemisphere, we do not consider grid points above 60°N with at least one monthly observation of sea ice higher than 0% over the historical period. The rest of the monthly resolved SST dataset is annually-averaged.

5.2.2.2 The random forest method applied to paleoclimate data

Over the instrumental period, both proxies and instrumental data overlap. For a given grid point, instrumental time series and proxy records can be seen for each

time step as an ensemble of measures of climate made using direct measurements or derived from natural archives, respectively. The RF (Breiman 2001) method consists in generating and aggregating a large number of regression trees based on random subsets of proxy records with equal length. A regression tree (Breiman et al. 1984) is built by selecting the best proxy records allowing to split along time the corresponding SST data into two data clusters in which the sum of variances is minimized. The latter clusters constitute two child nodes on which the same algorithm is applied recursively until a node contains 5 values or less. The latter value is widely used in applications of RF and is known to generally produce robust results on a wide variety of datasets, while its tuning strongly increases memory use and computation time for a very weak gain of robustness (Breiman 2001, Goldstein et al. 2011). We thus made this default choice since we will generate more than five million RF models in this study. For a single RF model, we compute 300 trees, which is large enough to converge in terms of mean squared error, even for larger datasets than ours (Oshiro et al. 2012). The number of randomly picked candidate proxy records for splitting each node of each tree is the only control parameter we tune using 5-fold cross-validation (Michel et al. 2020).

Once 300 trees have been built by scanning each time step of the training data, a new vector of proxy records for a given year in the past is used to browse trees using the optimal splits previously set up from training data. A given tree returns an estimation of SST, which is the average of historical SST data that previously finished up in the same node during model fitting. At the end, we obtain one reconstruction for the corresponding year for each of the 300 trees, which are averaged to give the final estimation. Uncertainties are calculated as the standard error between the observations and the reconstructed data over the period 1870-1980 (*cf.* Ortega et al. 2015, Michel et al. 2020).

Tuned RF models are beforehand evaluated by splitting 20 times SST data and proxies into 70% of randomly drawn training data and 30% remaining testing data (Michel et al. 2020). For a given reconstruction time frame, a grid point has a significant reconstructability if its Coefficient of Efficiency (Nash and Sutcliffe 1970) score is significantly positive among the 20 training/testing splits (Michel et al. 2020).

Such a tuned and evaluated RF model is computed for each grid point over 8 time frames : 850-1980 and from 1000-1980 to 1600-1980 with a 100 years step. For each of these time frames and each point we use for the respective RF models, the proxy records that entirely cover the target reconstruction period and that are significantly correlated at the 90% confidence level with the corresponding historical SST data (see Methods). The final reconstruction of SST for the period 850-1980 is obtained by nesting the 8 reconstructions (Wang et al. 2017, Michel et al. rev) made over the overlapping time frames. The nesting is done by putting one after the other the 8 overlapping reconstructions. We here fit RF models on absolute °C values. In the following, we focus on studying the anomalies (w.r.t. 850-1980) from this reconstruction.

5.2.3 Results

5.2.3.1 Reconstruction validation

Fig. 5.1 gives the spatial distribution of the CE score for the 8 timeframes introduced above (*i.e.* 850-1980 and 1000-1980 to 1600-1980 with a 100 years step). It shows a higher fraction of significantly reconstructable points for shortest reconstructions, which is due to the fact that the number of available proxy records increases too (Fig. 5.2a). It means that each grid point only covers the time frames over which it has significantly positive CE scores (Fig. 5.1a-h). For the longest time frame (850-1980) for which we only use 50 proxy records, 36% of the SST grid points have non-significant positive CE scores (Fig. 5.1). These missing points cover particular areas : more than 2/3 of Indian ocean ; Niño 1 and 2 regions ; the southern part of the Niño 3 and 4 regions ; the Pacific warm pool ; central northern and southern Pacific ; half of North Atlantic between 55°N and 70°N (Fig. 5.1a-i, 5.2a). By contrast, the rest of the North Atlantic and South Atlantic oceans have more than 95% of reconstructible grid points.

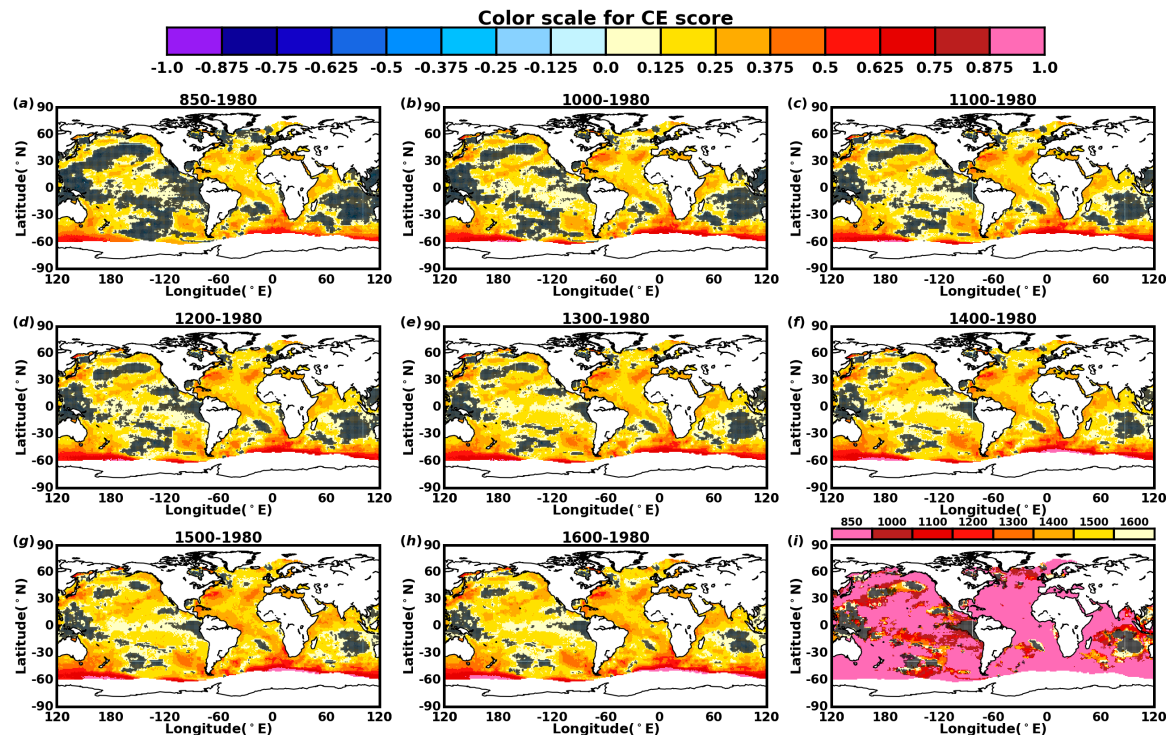


FIGURE 5.1 | CE score by time-frame of reconstruction. (a)-(h) Median validation CE scores (unitless) for the 20 training/testing splits. A black dot indicates that the grid point has not significantly positive CE scores at the 90% confidence level and is considered as missing values over the corresponding time frame. (i) Summary of the temporal availability of each grid point (starting calendar year when they are used) according to their reconstructability (cf. (a)-(h)).

For the shortest time frame (1600-1980), we selected 392 proxy records from the original set of 550, using the approach from section 5.2.2, and only 12% of grid points cannot be reconstructed (Fig. 5.1a-i, 5.2a). Except for small clusters in the Northern and Southern Pacific, we find three main clusters of grid points that have non-significant positive CE scores in the Pacific warm pool, the Niño 1 and 2 regions, and the eastern Indian Ocean. This poor reconstructability can be explained

by the lack of paleoclimate data in the southern hemisphere (PAGES 2k Consortium 2017). Indeed, the warm pool and the eastern Indian Ocean variabilities are known to be significantly affected by variations of southern hemisphere extratropical modes such as the southern oscillation (Picault et al. 1996) and the southern annular mode (Yeo and Kim 2015).

The 392 worldwide distributed proxy records that have been selected by the RF algorithm for the period 1600-1980 include 90.3% tree rings, 6.4% ice cores, 2% lake sediments, 0.8% corals and 0.5% historical documents (Fig. 5.2ab). The poor amount of ocean data is explained by the very short temporal availability of coral data and the very low temporal resolution of sediment data, the latter being removed here due to this time resolution limitation (*cf.* section 5.2.2.2).

The North and South Atlantic basins have the highest fractions of significant reconstructability for all time frames, and have on average higher validation scores than the Pacific and Indian Oceans (Fig. 5.3a). The ratio between regression standard error and respective SST standard deviation (Fig. 5.3c) never exceeds 0.5 (Fig. 5.3b) all over oceans, which shows that the reconstruction is providing relatively weak errors as compared to SST variations.

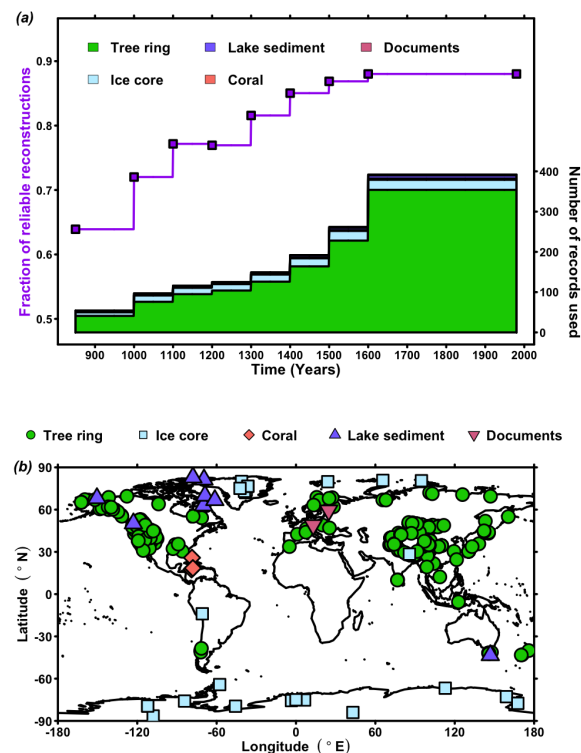


FIGURE 5.2 | Proxy records' types and temporal availability. (a) Purple line : the fraction of points with validation CE scores (unitless) significantly positive at the 90% confidence level. Areas : amount, types and temporal availability of the number proxy records used for each time frame of reconstruction (y-axis on the right inside). (b) Geolocation and types of the 392 proxy records used along the 8 timeframes of reconstruction.

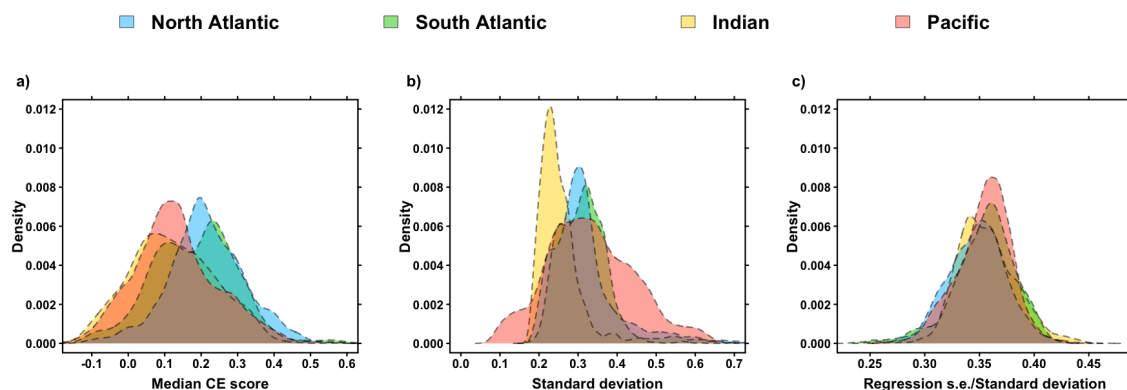


FIGURE 5.3 | Score, uncertainties and standard deviations for 4 oceanic basins Densities by grid points of 3 different metrics for North Atlantic, South Atlantic, Indian and Pacific basins : (a) Median CE score; (b) Standard deviation; (c) Ratio between regression standard error and standard deviation...

5.2.3.2 External forcings and last millennium climate sequences

The extracted globally averaged SST anomalies (w.r.t 850-1980, Fig. 5.4) are significantly positive for the period 850-1200 ($p < 0.01$) and significantly negative between 1200 and 1850 ($p < 0.01$). In this section we address the spatial and temporal SST response to natural radiative forcings, that are the modulation of the solar insolation and the volcanic eruptions, as noted in Fig. 5.4, and we identify the spatial differences between the warmest and coldest periods at the global scale.

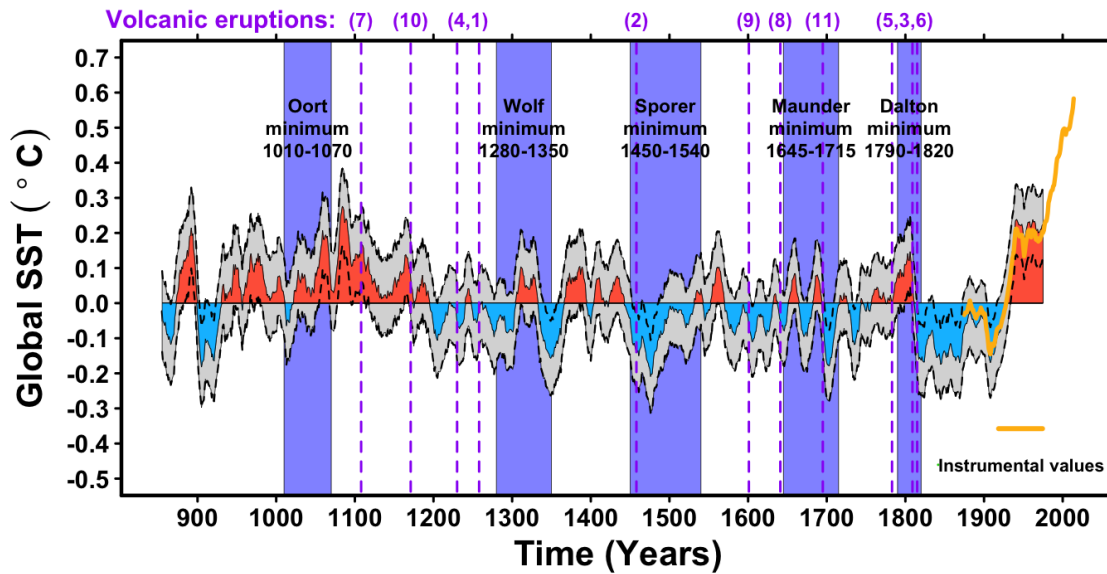


FIGURE 5.4 | Global SST anomalies, solar minima and large volcanic eruptions. 11-year running mean of the global SST anomalies. Heavy orange line : 11-year running mean of instrumental global SST anomalies. Grey area : Spatially-averaged regression uncertainties. Dashed purple lines : 11 largest volcanic eruptions of the past 1000 years (Sigl et al. 2015, details in Tab. 5.1). Purple numbers above dashed purple lines give the rank of the eruptions according to the loss of W.m^{-2} , (1) standing for the strongest, and (11) for the weakest (Sigl et al. 2015, details in Tab. 5.1). Blue areas : Solar minima (Lean 2018).

5.2.3.2.1 Response to volcanic eruptions

Using a well-verified reconstruction of last millennium volcanic activity (Sigl et al. 2015), we perform a composite response of reconstructed SST to the 11 largest tropical volcanic eruptions of the last millennium (Tab. 5.1). Here, we mention N as the year of the eruption, and $N-i/N+i$ respectively stand for the i -th year before/after the eruption. For the North Atlantic, we find a very strong negative response that gradually decreases for 2 decades (Fig. 5.5), which is in line with findings from a coupled model-based study (Mignot et al. 2011), and other basin-scale averages of North Atlantic SST reconstructions (Wang et al. 2017, Michel et al. rev). For the second decade after the eruption, though, we only find a few negative signals in the tropical Atlantic (Fig. 5.5). On the other hand, the tropical South Atlantic does not have a significant response from over these two decades. This absence of significant response while northern tropical Atlantic cools, might be indicative of a changing balance in tropical Atlantic SST after strong eruptions, leading to a southward shift of the Atlantic Inter-Tropical Convergence Zone possibly in response to an AMOC

weakening (Knight et al. 2006) as highlighted in Mignot et al. 2011. The lack of signal in the northern subpolar gyre might be related to the large variability in this region, which hampers to have significant response due to the large internal noise there.

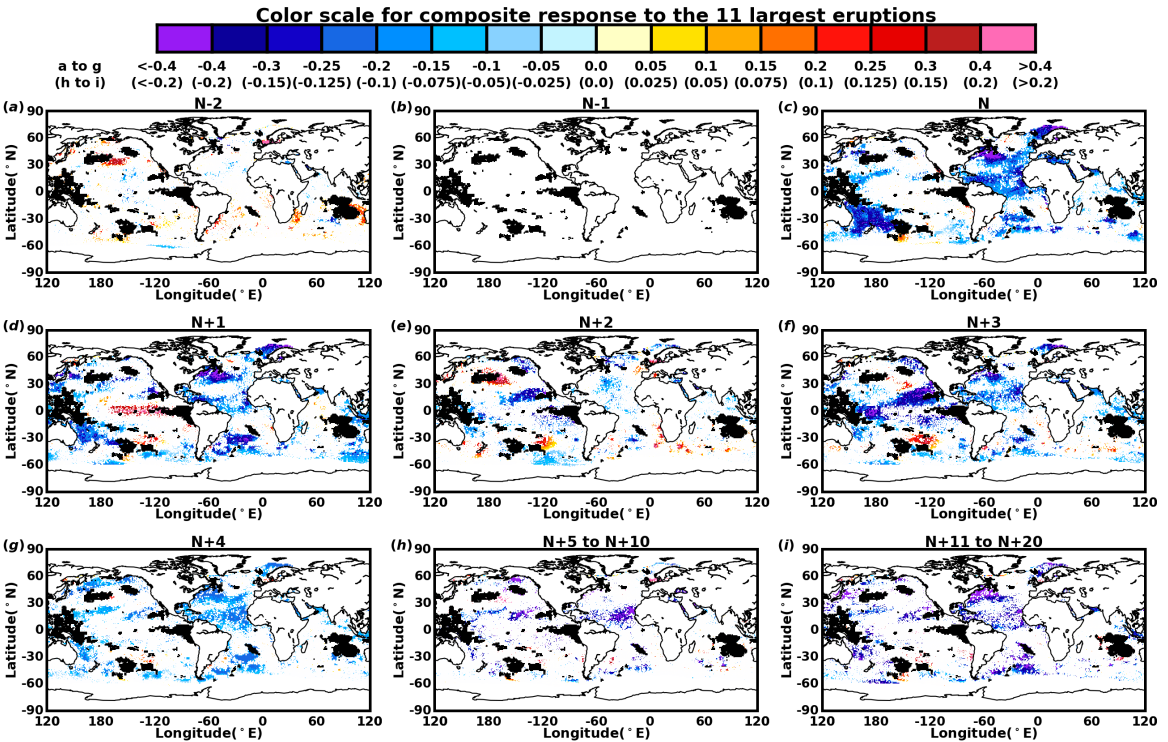


FIGURE 5.5 | Spatial response of SST to large volcanic eruptions. Composite response of SST to the 11 largest volcanic eruptions of the last millennium (Sigl et al. 2015, see Tab. 5.1), respectively for year N-2 to N+4 (a-g), N+5 to N+10 (h) and N+11 to N+20 (i). (a-i) Black dots are grid points not considered because of their poor reconstructability (Fig 1). Grid points colored in white indicates no significance at the 90% confidence level (See Rao et al. 2019 and Michel et al. 2020 for significance test).

For the Pacific Ocean, an interesting feature is the significant positive anomalies found in the eastern equatorial Pacific in response to volcanic eruption in year N+1 (Fig. 5.5), which is in line with modeling results from Ohba et al. (2013) and Khodri et al. (2017). The latter studies indeed suggest that tropical volcanic eruptions might help to trigger an El Niño event within the year following a large volcanic eruption. We find a few positive but decreasing anomalies from N+2 and N+4 in central northern and central southern Pacific (Fig. 5.5). From N+3 to N+4, we find a relatively large number of grid points with negative anomalies covering the Pacific basin elsewhere but with no discernible well-known pattern. From N+5 to N+10 and N+11 to N+20, negative anomalies are still present but they largely decrease over time.

Volcano	Country	Date
Unknown		1108
Unknown		1171
Unknown		1230
Samalas	Indonesia	1258
Kuwae	Vanuatu	1458
Huaynaputina	Peru	1601
Parker	Phillippines	1641
Serua	Indonesia	1695
Lakagigar	Iceland	1783
Unknown		1809
Tambora	Indonesia	1815

TABLE 5.1 | Eleven largest volcanic eruptions of the last millennium and location (Sigl et al. 2015).

The decadal to bidecadal responses of the Pacific Ocean do not suggest any lagged negative IPV-like response. Thus, this does not validate the hypothesis that the recent negative IPV-induced hiatus can be solely explained as a lagged response to the Pinatubo eruption (1991) (Smith et al. 2016), which was a largely weaker eruption than those investigated here. From this paleo-reconstruction perspective, the signal might be far too small to explain this recent feature.

The Indian Ocean has a very weak response to volcanic eruptions as compared to the other oceans with just a few grid points with negative anomalies mainly located in the Arabian Sea and west of Indonesia. This result is not validating findings from Izumo et al. (2018) who proposed that large tropical volcanic eruptions enhance the west-east SST gradient resulting in a negative Indian Ocean Dipole index. However, the latter was only based on the response to the 5 largest eruptions of the last 150 years that do not even include one of the 11 we investigate here. Also, as we have shown, the Indian Ocean has relatively low scores of reconstructability in our reconstruction (Fig. 5.1, 5.3), which might be a source of misestimation of its response to external forcings.

5.2.3.2.2 Response to solar variability

We now assess the climate impacts of long-term minima in Total Solar Irradiance (TSI) of the last millennium (Lean 2018, Tab. 5.2) using composite time-averaged anomalies for solar minima and for the remaining periods, out of solar minima (Fig. 5.6a-c). We find a relatively weak absolute difference between averages of both groups of anomalies (0.024°C) but with strong significance ($p < 0.01$, Student t-test) (Fig. 5.6c).

Name	Start	End
Oort	1010	1070
Wolf	1280	1350
Spörer	1450	1540
Maunder	1645	1715
Dalton	1790	1820

TABLE 5.2 | Largest solar minima of the last millennium (Lean 2018).

The same result is found for the difference between variances of time-averaged anomalies, with significantly stronger variance during solar minima ($p < 0.01$, Fisher-Snedecor test). In terms of spatial distribution, we find positive anomalies in the North Atlantic subpolar gyre region during solar minima, which might be indicative of enhanced AMOC. Indeed, the latter has been found to be anti-correlated with solar variations in a climate model (Swingedouw et al. 2011). We also detect positive anomalies in the equatorial Pacific, mainly in the Niño 3 region. Northern and Southern mid-latitudes in Pacific are notably characterized by negative anomalies, shaping a positive IPV-like pattern in the Pacific basin, which suggest more occurrence of the latter during solar minima than out of them as suggested in a model simulation in Swingedouw et al. (2011). For the rest of the oceans, we find negative anomalies in the extratropical northern and southern Atlantic and in the northern and eastern Indian Ocean. Thus, this new reconstruction validates two model results concerning the response of the North Atlantic and Tropical Pacific to variation in solar forcing in a long enough time frame. Furthermore, it is providing a quantification of the response, which might be of use to evaluate such a strength of the response in climate models.

5.2.3.2.3 Characterization of MCA and LIA

In the previous sections, we have shown that the cooling of temperatures during solar minima periods is less than 0.025°C in terms of globally averaged anomalies.

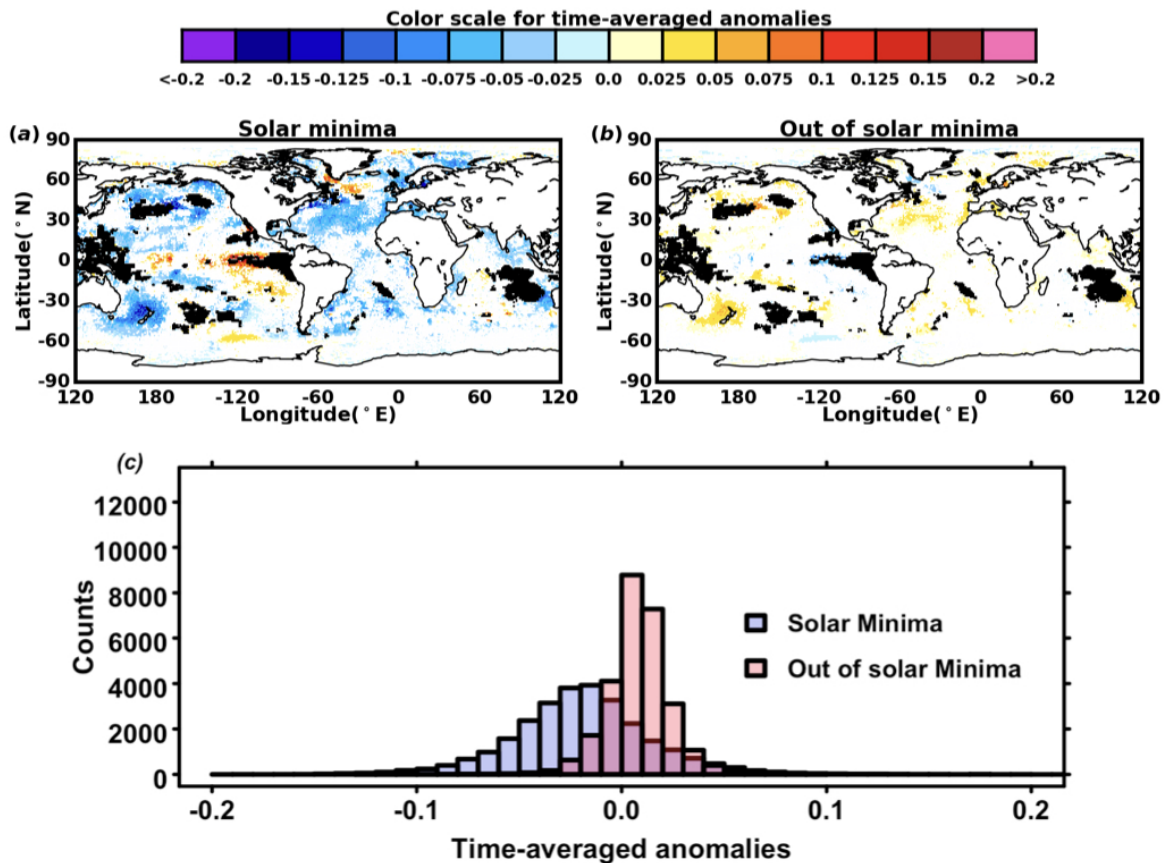


FIGURE 5.6 | Spatial response of SST to grand solar minima. (a) Composite values of SST (in °C) for the 5 longest solar minima of the last millennium (see Tab. 5.2) (b) Composite values of SST out of the largest solar minima of the last millennium. (a,b) Black dots are grid points not considered because of their poor reconstructability (Fig 1). Grid points colored in white indicate no significance at the 90% confidence level according to a Student t-test. (c) Histograms of time-averaged SST anomalies respectively during (blue) and out of (red) the 5 longest solar minima of the last millennium (see Tab. 5.2)

This very low response might be also due to the fact that the equatorial Pacific shows a positive response to solar minima. We have also shown that the temperature response to large volcanic eruptions is particularly strong the first ten years following the eruption and significantly decreases for the following ten years. These results can thus partly but not solely explain multi-centennial-long climate sequences such as MCA and LIA. Defining these time periods as 750-1350 and 1400-1850, respectively, Neukom et al. (2019) have shown a non-spatially synchronous warming (resp. cooling) during the MCA (resp. LIA). On the other hand, studies have shown that Arctic glaciers and sea ice have anomalously extended during the middle of the 12th century (Kinnard et al. 2011, Jomelli et al. 2016), supposedly part of the MCA. Comparing forced and unforced reconstructions of North Atlantic SST, Michel et al. (rev) have suggested that this cold period might be the result of internally-driven changes by emphasizing that it occurred before the 13th century cluster of strong volcanic eruptions.

Here, given the uncertainty in LIA and MCA definition, we propose to determine objectively, in terms of global SST averages, the most contrasting and distinguishable periods of 200, 300 and 400 years of the last millennium (Fig. 5.7). We

first find that for all possible time windows, the warmest period lies around the beginning of the last millennium and the coldest one much later, consistent with the general picture of MCA and LIA. We also find that the warmest periods of 200 and 300 years (respectively 971-1170 and 872-1171) (Fig. 5.7) do not include the second part of the 12th century and the following decades. This result matches with prevailing negative anomalies in global SST (Fig. 5.4) and supports results from Kinnard et al. (2011), Jomelli et al. (2016) and Michel et al. (rev) highlighting that the end of MCA might start as early as the end of the 12th century. The warmest 400-years long period (935-1334) does include the late 12th and the 13th centuries (Fig. 5.7), but as highlighted before, the shift might occur earlier. Indeed, the differences between periods for such a long 400 years scale are far lower than for 200 years period (0.14°C for the former vs. 0.08°C for the latter).

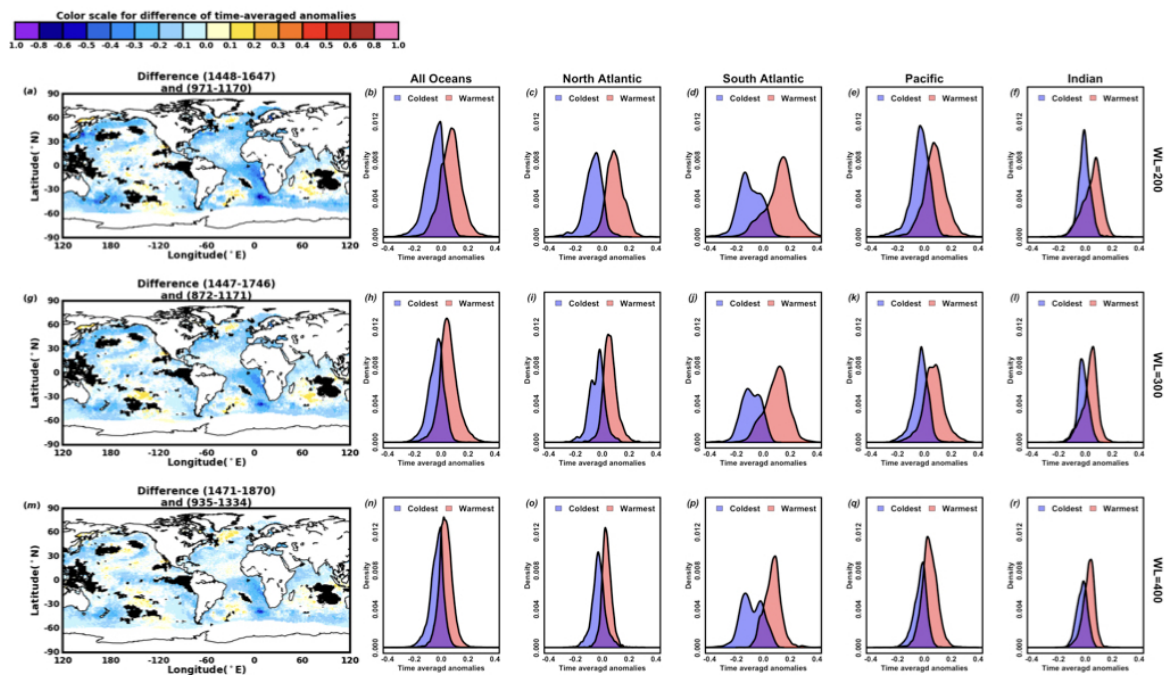


FIGURE 5.7 | Optimal global-scale contrasting periods for SST over the last millennium. Optimal global-scale contrasting periods are calculated for : (a-f) 200 years time frames ; (g-l) 300 years time frames ; (m-r) 400 years time frames and are expressed in $^{\circ}\text{C}$. For these periods, time-averaged anomalies are provided spatially (a,g,m) and as densities for : All Oceans (b,h,n) ; North Atlantic Ocean (c,i,o), South Atlantic Ocean (d,j,p), Pacific Ocean (e,k,q) and Indian Ocean (f,l,r).

For the three window length, we investigate four basins (North and South Atlantic, Pacific, and Indian) which all have significantly lower anomalies during LIA than during MCA at the 95% confidence level (Fig. 5.7). Our results, therefore, support a relative spatial homogeneity for MCA and LIA definitions, which differs with Neukom et al. (2019) conclusion, although they focus on minimum and maximum temperature averages over a shorter time frame length (50 years), which largely differ from the present analysis. Here we focus on the ocean, which is a low pass filter of climatic variability, and might therefore highlight better coherency spatially. Performing the same analysis than Neukom et al. 2019, we also found heterogeneity of 50-years long maxima and minima, but with a significantly different spatial distribution (not shown). Furthermore, we emphasize that while 950-1150 is strongly marked by positive anomalies, the remaining preindustrial period (1200-1850) is rather fluctuating on multi-decadal to multi-centennial timescales in terms of anomalies

w.r.t. 850-1980 (Fig. 5.4). It is marked by three particularly cold periods that suggest a gradual decrease in long-term averaged temperatures over the millennium : 1170-1350 ($m=-0.024^{\circ}\text{C}$, $p<0.01$), 1450-1700 ($m=-0.043^{\circ}\text{C}$, $p<0.01$) and 1800-1880 ($m=-0.098^{\circ}\text{C}$, $p<0.01$), with m standing for the average of global SST anomalies over the period, p measures the significance of a negative average of global SST anomalies (w.r.t. 850-1980) over the period using a Student t-test.

5.2.3.3 North Atlantic and Pacific modes

We now focus on the multi-decadal to centennial variability modes that might be extracted from our gridded reconstruction of SST. For this purpose, we first standardize and apply a 10-year Nadaraya-Watson kernel smoothing (Franke et al. 2002) of Pacific and North Atlantic reconstructed data. Using a principal component analysis, we explore for each basin the first two principal components (PCs) that we compare with classical SST-based indices for the historical data (Fig. 5.8) :

- NASST : Spatially-averaged North Atlantic SST anomalies.
- PSST : Spatially-averaged Pacific SST anomalies.
- AMV_{TS} : Difference between NASST and global SST anomalies following Trenberth and Shea (2006).
- AMOC_{C} : Difference between SPG SST anomalies and global SST anomalies following Caesar et al. (2018).
- IPV_{H} : Tripole IPV index (see Fig. 5.8) following Henley et al. (2015).
- $\text{N}_{3.4}$: Spatially averaged SST anomalies in the Niño 3.4 region (see Fig. 5.8).

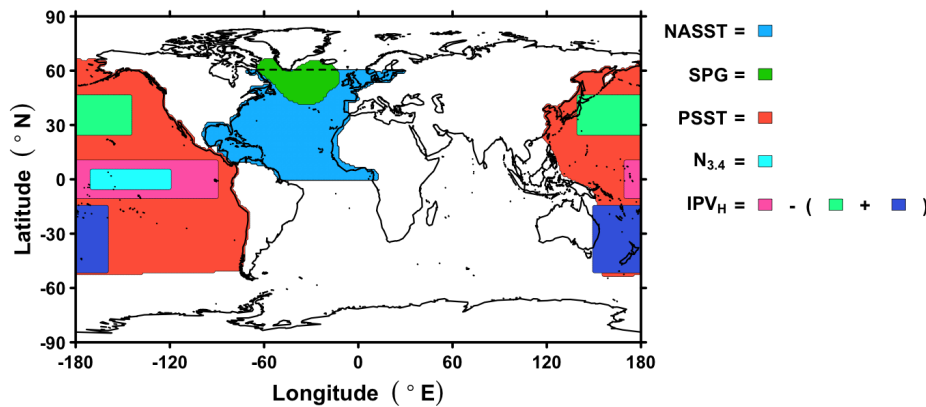


FIGURE 5.8 | Definition of the areas over which climate indices are calculated.

PCs and their associated empirical orthogonal functions (EOFs) of North Atlantic (resp. Pacific) are denoted PC1_{NA} , PC2_{NA} , EOF1_{NA} and EOF2_{NA} (resp. PC1_{Pac} , PC2_{Pac} , EOF1_{Pac} and EOF2_{Pac}). For each PC/EOF, we project the matrix of observed historical SST from the corresponding basin on the EOF basis to extend time series to 2019 (see Methods).

5.2.3.3.1 Pacific

PC1_{Pac} and PC2_{Pac} have relatively similar weights : 23.3% and 21.6%, respectively (Fig. 5.9). PC1_{Pac} is strongly correlated with the PSST ($r=0.98$, $p<0.01$, Fig. 5.9) *i.e.* the changes at the global scale in the Pacific. Its corresponding EOF is indeed characterized by positive values all over the ocean, except for two negative clusters

in central southern and central northern Pacific (Fig. 5.9). EOF2_{Pac} is a tripole of Pacific SST, with strong positive coefficient from central to eastern equatorial Pacific (Fig. 5.9), and negative regression coefficients in the rest of the basin. This EOF is very similar to the large-scale IPV pattern, and PC2_{Pac} is indeed strongly correlated with the IPV_H index ($r=0.97$, $p<0.01$, Fig. 5.9), and to a lesser extent with the N3.4 index ($r=0.81$, $p<0.01$). Projecting the historical data on both EOF1_{Pac} and EOF2_{Pac} also validates the consistency in terms of correlation with observed PSST ($r=0.99$, $p<0.01$) and IPV_H ($r=0.97$, $p<0.01$) in instrumental data (Fig. 5.9). Thus, this analysis validates for the Pacific the validity of simple index definitions that correlate very well with modes defined using a more advanced approach, here the PCA.

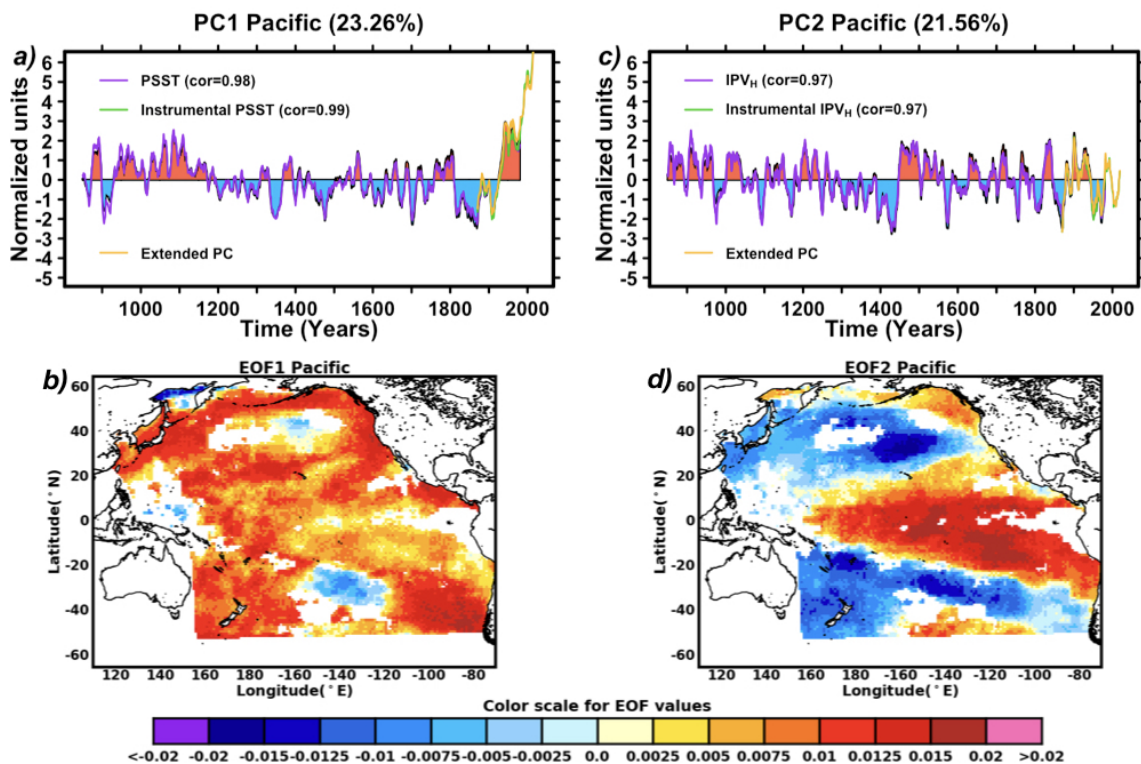


FIGURE 5.9 | **Analysis of the Pacific.** Left : First PC (a) and EOF (b) of 11-year running mean Pacific SSTs. Right : Second PC (c) and EOF (d) of 11-year running mean Pacific SSTs. In top panels (a and c), the purple curves represent the evolution of similar index but based on classical definition based on boxes (cf. Fig. 5.8); the green curves correspond to the same indices but computed using instrumental data; and the yellow curves stand for the projections of instrumental observations on the associated EOF.

5.2.3.3.2 North Atlantic

PC1_{NA} (45% explained variance, Fig. 5.9) has 0.99 ($p<0.01$) correlation with NASST and 0.90 ($p<0.01$) with global SST. EOF1_{NA} is a positive pattern all over the North Atlantic except in the northern SPG region with a few slightly negative values (Fig. 5.10). As for PC1_{Pac}, PC1_{NA} is associated with the North Atlantic average (NASST) which encompasses both responses to internal variability (Frankcombe et al. 2010, Clement et al. 2015, Muir and Fedorov 2015, Yan et al. 2018) and external forcings (Fig. 5.5, 5.6). Its correlation with AMV_{TS} (0.65, $p<0.01$) is similar to the one found between independently reconstructed NASST and AMV indices (Michel et al. in rev., where $r=0.62$, $p<0.01$). Concerning these previous index-based reconstructions,

while only based on indices temporal variability, they are consistent with our NASST and AMV_{TS} extracted from the gridded reconstruction ($r=0.7$, $p<0.01$ for both).

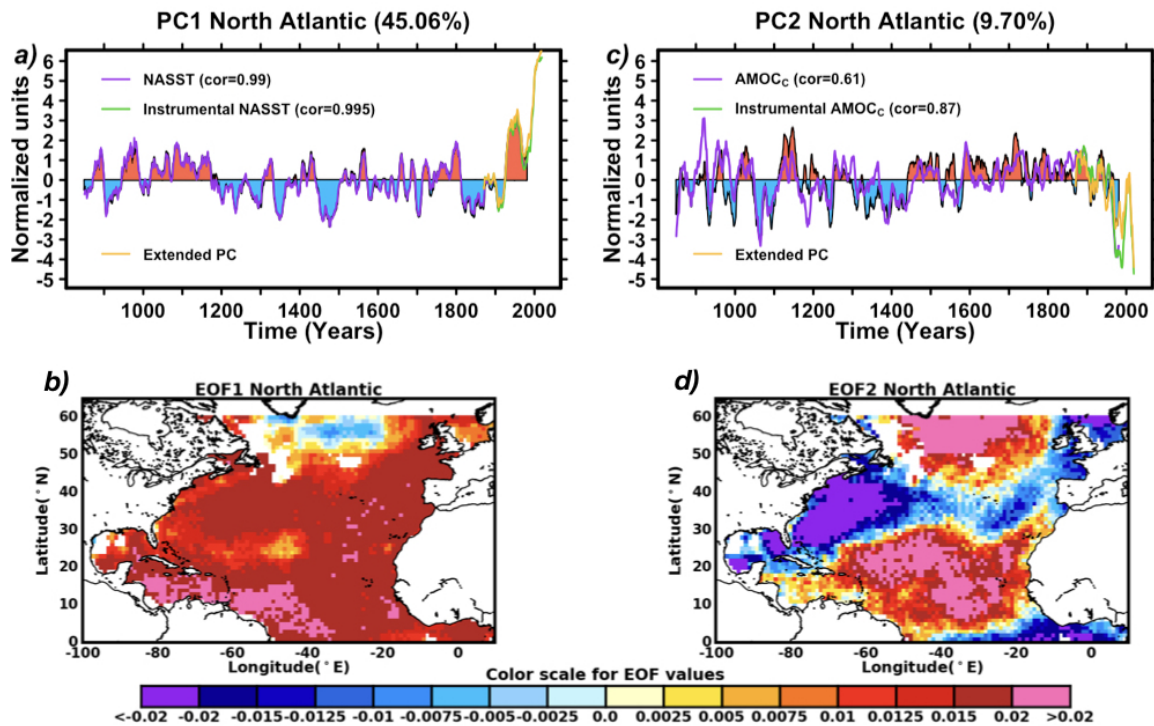


FIGURE 5.10 | PCA of the North Atlantic. Left : First PC (a) and EOF (b) of 11-year running mean NASST. Right : Second PC (c) and EOF (d) of 11-year running mean North Atlantic SSTs. In top panels (a and c), the purple curves represent the evolution of similar index but based on classical definition based on boxes (cf. Fig. 5.8); the green curves correspond to the same indices but computed using instrumental data; and the yellow curves stand for the projections of instrumental observations on the associated EOF.

EOF2_{NA} (10% explained variance, Fig. 5.10) represents a tripole over the North Atlantic. It is negative along the coastal Gulf Stream (GS) region and positive in the SPG and tropical North Atlantic. This pattern is consistent with AMOC SST fingerprints described in other studies (Knight et al. 2006, Zhang 2008, Persechino et al. 2013). PC2_{NA} is indeed significantly correlated with the extracted AMOC_C index over the period 850-1980 (0.61, $p<0.01$). In terms of time evolution, we find that PC2_{NA} reaches a minimum in 1065 after experiencing a strong negative trend and rapidly recovers towards a neutral state about 40 years later. The maximum is reached in 1147 and suddenly experiences a second drop. After this second drop, PC2_{NA} falls into 300 years with only 27% ($p<0.01$) of positive values (between 1150 and 1450). After 1450, PC2_{NA} has 80% of positive anomalies until 1870 (Fig. 5.10). These estimations of AMOC variations support those from a recent study which suggests a similar drop in both NASST and internal AMV during the late 12th century, arguing that it is not related to external forcings (Michel et al. rev). It is also in line with no role for oceanic circulation in shaping negative NASST after 1600 (Michel et al. rev), which might be rather related to volcanic forcing as suggested by Brönnimann et al. (2019).

5.2.3.4 Comparison of two potential last millennium AMOC estimates

To determine which of the PC2_{NA} or AMOC_C indices is the best candidate for describing the last millennium AMOC variability, we use regression coefficients to

estimate their SST fingerprints within the reconstruction for the period 850-1980 (Fig. 5.11). $AMOC_C$ has positive relationships in North Atlantic SPG regions and negative relationships with the rest of the Atlantic Ocean (Fig. 5.11). For the other oceans, except a cluster of positive regression coefficients in North-Eastern Pacific, a few sparse ones in the tropical Pacific, $AMOC_C$ only has negative relationships. Contrary to $AMOC_C$, $PC2_{NA}$ draws a very clear inter-hemispheric gradient of SST over the Atlantic, such as evidenced for the AMOC SST fingerprint by many previous studies using climate models (Knight et al. 2006, Persechino et al. 2013, Swingedouw et al. 2015).

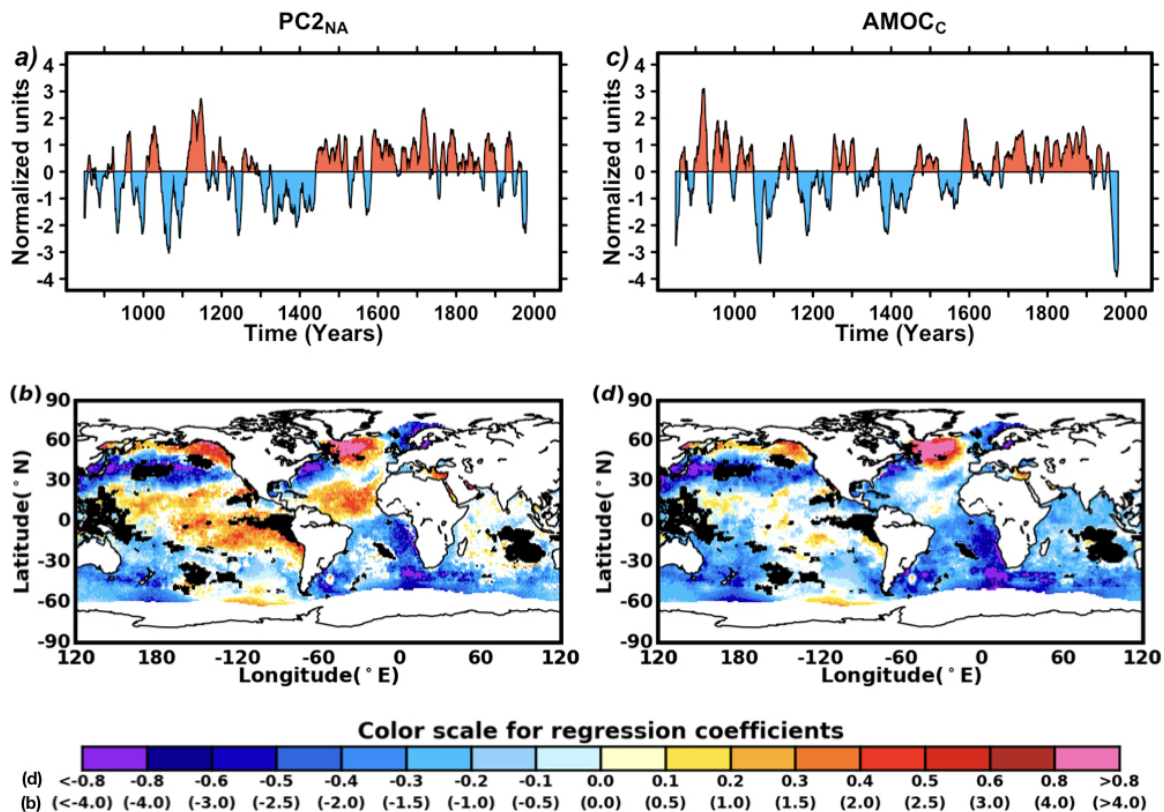


FIGURE 5.11 | Time series and regression maps (in °C per unitless PC value) for two candidate indices for AMOC. a-b) $PC2_{NA}$ c-d) $AMOC_C$ index.

In addition, $PC2_{NA}$ has positive relationships with the SPG SST and negative relationships with GS SST, which is also consistent with the observed and modeled AMOC fingerprint in these regions (Zhang 2008, Caesar et al. 2018). While significantly lacking reconstructed data (Fig. 5.1, 5.3), the Indian Ocean has mainly positive relationships with $PC2_{NA}$ except along the equator (Fig. 5.11). For eastern and western Indian Ocean, we respectively find positive and negative regression coefficients, which indicate a positive relationship of $PC2_{NA}$ with the Indian Ocean Dipole Index (Izumo et al. 2018) (Fig. 5.11). However, as the Indian Ocean has weaker validation scores than the other oceans (Fig. 5.1, 5.3), this statement needs more focused and deeper investigations. Regarding the Pacific we find much stronger positive relationships with $PC2_{NA}$ than $AMOC_C$ all over the tropic (from 20°S to 20°N) (Fig. 5.11). These positive regression coefficients also spread along with the extratropical coastal North and South America (from 40°S to 60°N) (Fig. 5.11). For the rest of the Pacific Ocean, we find negative relationships between $PC2_{NA}$ and SST (Fig.

5.11). We note that Pacific coefficients related to $PC2_{NA}$ (Fig. 5.11) have strong similarities with the IPV pattern given by $EOF2_{Pac}$ (Fig. 5.9). We indeed find a significant in-phase relationship between $PC2_{NA}$ and $PC2_{Pac}$ ($r=-0.49$, $p<0.01$), which thus emphasize another link between North Atlantic and Pacific. The $PC2_{NA}$ fingerprint in the Pacific is more in line with previous findings than the one of $AMOC_C$ (Knight et al. 2006, Zhang 2008, Persechino et al. 2013) which mostly has negative regression coefficients all over the basin.

The extended $PC2_{NA}$ (Fig. 5.10) is obtained by projecting instrumental HadISST data for the period 1870-2019 on $EOF2_{NA}$. The latter exhibits a strong correlation with $AMOC_C$ over the instrumental data ($r=0.87$, $p<0.01$). As suggested by many previous studies either based on proxies, models, or historical $AMOC_C$ (Rahmstorf et al. 2015, Thornalley et al. 2018, Caesar et al. 2018), $PC2_{NA}$ shows a very strong negative trend reaching an unprecedented state since at least 850 C.E. (Fig. 5.10). Yet, contrary to studies suggesting that the AMOC decline started during the 19th century (Rahmstorf et al. 2015, Thornalley et al. 2018), we here estimate that the large decline approximately started in the 1940s. Indeed, with respect to the average (m) and standard deviation (σ) of $PC2_{NA}$ over the period 850-1980, we find that the extended $PC2_{NA}$ is significantly positive between 1870 and 1945 ($CI_{95}(m) \subset [0.24\sigma, 0.54\sigma]$), where CI_{95} stands for confidence interval at the 95% confidence level). For 1946-2019, we find more than 90% of negative values of $PC2_{NA}$ which is strongly significant ($CI_{95}(m) \subset [-1.58\sigma, -1.1\sigma]$). We have found here some evidence that $PC2_{NA}$, and its extended version (Fig. 5.10), is a better estimate of the AMOC over the last millennium and the historical period than the $AMOC_C$ index. $PC2_{NA}$ and $AMOC_C$ are nevertheless very consistent in terms of correlation (Fig. 5.10), and our historical AMOC estimate is consistent with the recent large trend found by Rahmstorf et al. (2015) and Caesar et al. et al. (2018). It therefore confirms the possibility of a substantial recent AMOC weakening, estimated to start from the 1940s.

5.2.3.5 Inter-relationship between North Atlantic and Pacific basins

For our preferential AMOC estimate ($PC2_{NA}$), we only find in-phase positive relationships with the AMV ($r(PC2_{NA}, AMV_{TS})=0.4$, $0.01<p<0.05$). Since models usually find a lag between the two (e.g. Knight et al. 2005), such lack of time lag highlights that this index might be slightly not in phase with real AMOC variations. This seems logical since the index is based on SST fingerprints, which takes time to settle after heat transport variations (time of integration of the anomalies), usually of the order of 3-10 years (e.g. Zhang and Wang 2013, Ba et al. 2014).

Concerning the inter-basin linkages, we find in-phase positive relationships between the indirect AMOC estimate $PC2_{NA}$ and the IPV ($r(PC2_{NA}, PC2_{Pac})=0.49$, $p<0.01$; $r(PC2_{NA}, IPV_H)=0.47$, $p<0.01$) and a relatively weak but significant lagged correlations when AMV_{TS} leads both IPV_H ($r=-0.28$, lag=14, $0.05<p<0.1$) and $PC2_{Pac}$ ($r=-0.24$, lag=14, $0.05<p<0.1$) (Fig. 5.12). This is in line with Marini and Frankignoul (2013), McGregor et al. (2014) and Chafik et al. (2016) results. Also, as suggested by Marini and Frankignoul (2013), we do not find significance when IPV leads the AMV (Fig. 5.12). In the following, we only discuss the relationship between AMV_{TS} and IPV_H that is the strongest we find for the two IPV estimates (Fig. 5.12).

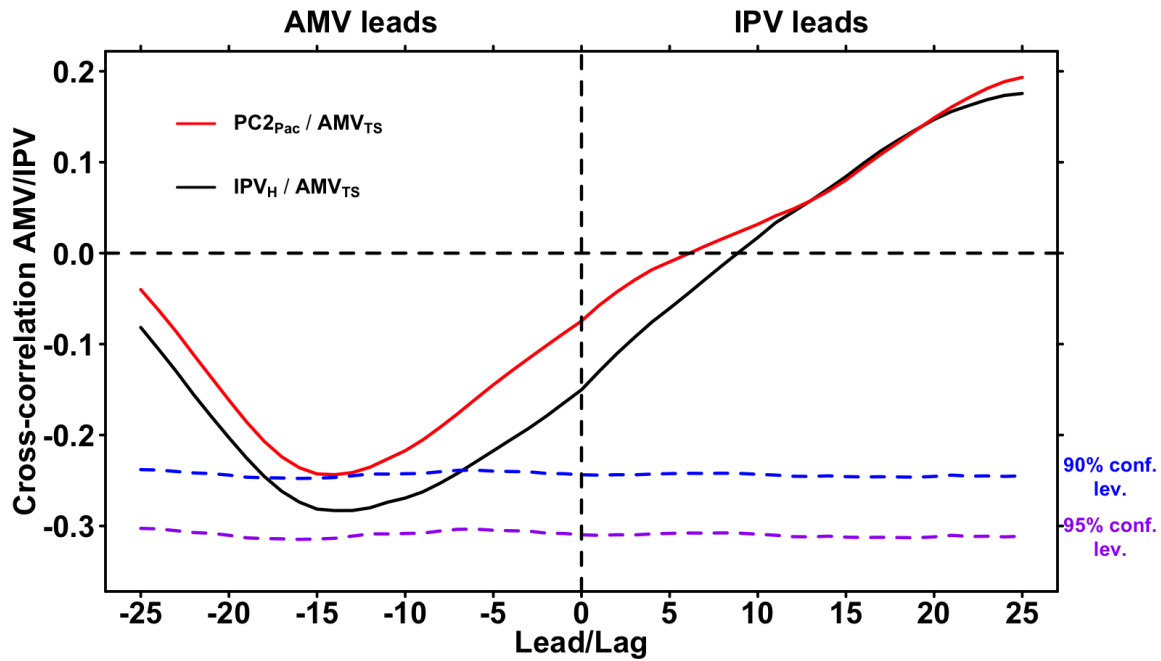


FIGURE 5.12 | Cross-correlations between AMV_{TS} and IPV_H (black) and $PC2_{NA}$. The confidence level is calculated for each time lag using the correlation test from McCarthy et al. (2015), Ortega et al. (2015), and Michel et al. (2020), with timeseries autocorrelations correction.

A deeper investigation shows that the lagged relationship between AMV and IPV appears to depend on the statistical properties of both modes in terms of index mean state and variance (Fig. 5.13). To gain insight into this aspect, we examine the minimum lagged correlation when AMV_{TS} leads IPV_H for 936 time-frames of 100 consecutive years taken between 850 and 1980. The results are plotted as a function of the standard deviation in Fig. 5.13a,b, and they suggest that strong amplitudes of either IPV and AMV are weakly but significantly associated with a strengthening of the link between the two modes. Modulations of the AMV amplitude have indeed been proven to significantly increase the strength of its teleconnections in different models (Qasmi et al. 2017), which seems logical since a larger signal might produce clearer linkages. Nevertheless, the relatively low R^2 statistics show that the relation between the AMV/IPV link and their respective standard deviations is not constant over time. It also supposes that many other processes, whose amplitude might also vary, also play a role in AMV and IPV variabilities (Fig. 5.13a,b). We also find that strong IPV amplitudes explain the strongest relationships with short lags (1 to 5 years), and strong AMV amplitudes rather explain the strongest decadal-scale lags (10 to 20 years) (Fig. 5.13a,b). This might indicate that the strength and timing of the linkage processes might be sensitive to the state of both AMV and IPV and. Although providing a detailed explanation of this is out of the scope of this paper, we can highlight that these differences might be related to the state of the Walker circulation at the equator, which is strongly dependent on the state of the AMV and IPV. The Walker circulation pattern is a subtle interplay between the different deep convection occurrences in the tropical belt (Li et al. 2016), which might be at the origin of the different time scale found depending on the state of the AMV and IPV.

We additionally find that the signs of both AMV and IPV over the 100 years

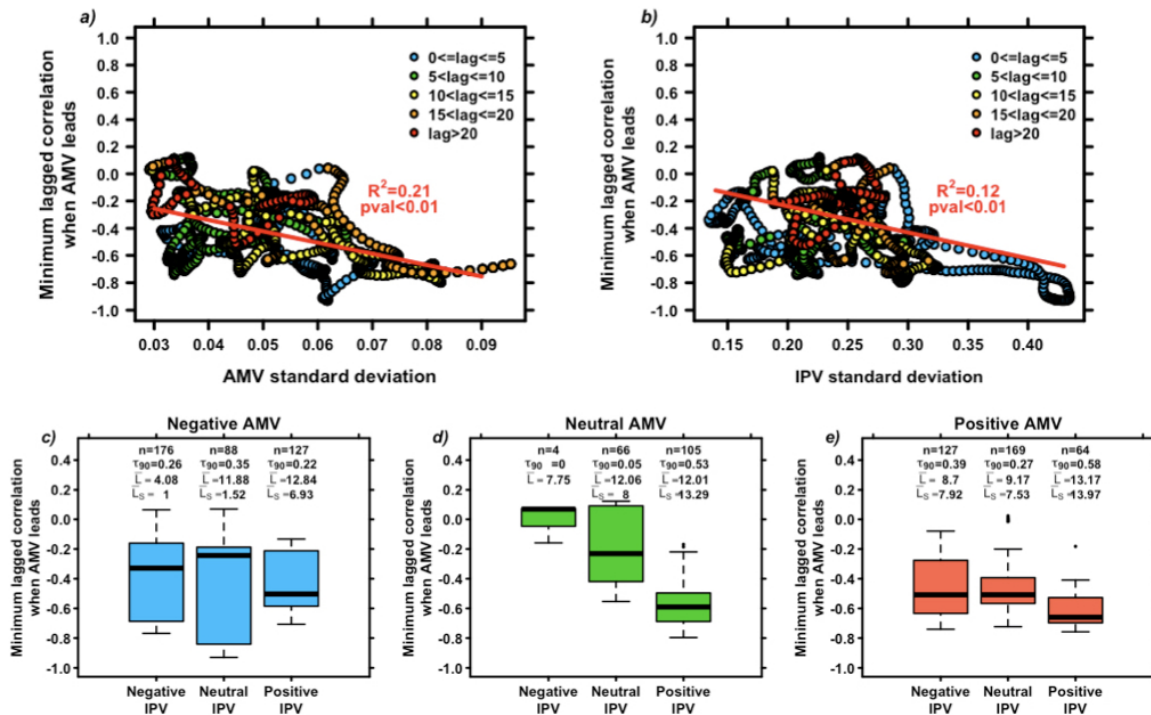


FIGURE 5.13 | Study of the link between AMV and IPV. a-b) Minimum lagged correlation (since the correlation is negative, which means that we search here for the maximum in absolute correlation for anti-phase linkages) when AMV_{TS} leads vs. AMV_{TS}/IPV_H standard deviation, respectively. c-e) Boxplots of minimum lagged correlations when AMV_{TS} leads for negative, neutral and positive AMV_{TS} , respectively. For each, boxplots from left to right are for negative, neutral and positive IPV_H , respectively. AMV_{TS}/IPV_H are considered positive/negative if their 10-years kernel smoothed time series are significantly positive/-negative at the 95% confidence level over the 100 years time frame. If the index is neither significantly positive nor negative over the 100 years time frame, it is considered neutral. For each boxplot, n is the number of time frames for the particular states of AMV and IPV. τ_{90} is the fraction of significant lagged correlations at the 90% confidence level among the n time frames. L is the average lag of the n time frames. LS is the average lag for time frames for which the lagged correlation is significant at the 90% confidence level. The confidence level is calculated for each time lag using the correlation test from McCarthy et al. (2015), Ortega et al. (2015) and Michel et al. (2020), with timeseries autocorrelations correction.

time frames do impact the strength of the lagged correlation (Fig. 5.13c-e). These results show that while there is only evidence for a leading effect of the AMV on IPV, phases of both play a role in strengthening or weakening these relationships. In particular, positive AMV is particularly subject to induce a shift in IPV as suggested by McGregor et al. (2014) and Chafik et al. (2016), through a strengthened inter-basin Walker circulation and trade winds leading to an eastern equatorial Pacific cooling through Ekman drift. This dependence in the response of the Pacific to the sign of the AMV can be related to the triggering of deep atmospheric convection in the Atlantic, which is a threshold mechanism dependent on the SST, and is favored when SST is warmer (e.g. Swingedouw et al. 2017). This is a non-linear response which might explain physically the asymmetry we found. However, the lagged decadal correlation is significant at the 90% confidence level for 37% of the 360 timeframe for which AMV is positive, which emphasizes that positive AMV might not systematically lead to an IPV shift as questioned by McGregor et al. (2014). Such an estimate might be very crucial to properly assess the prediction from climate system based on former cases on a long enough timescale

Another interesting feature of our analysis is that the sign of the modes also influences the optimal lag of the correlation. Indeed, by only accounting for significant lagged correlations, we find very short lags for negative AMV and negative IPV, if taken independently. For negative AMV, the average time lag of significant correlation is 1 and 1.5 years for negative and neutral IPV, respectively. On the contrary, while considering a negative AMV state, significant lagged correlations are reached for an average lag of 6.9 years, which is significantly different than for the two other IPV states at the 99% confidence level. For the negative AMV state, we rather find a 7.9 averaged time lag for significant lagged correlations and 13 and 14 years for neutral and positive AMV states, respectively. The potential physical reasons that might explain these differences are unclear at the moment but might be used as a potential roadmap for interpreting sensitivity experiments like the one of McGregor et al. (2017). These particularly strong AMV/IPV lagged anti-correlations when AMV is positive is a crucial result since McGregor et al. (2014) and Chafik et al. (2016) have suggested the the positive phase of AMV in the 90's have been the cause of the long-term negative IPV and related hiatus of global temperatures between 1998 and 2012. This is another piece of evidence that the AMV might have been the initial cause of the onset of the 1998-2012 hiatus.

Our results highlight a clear dependence between the statistical behavior of modes and their inter-relationship. Indeed, performing cross-correlations on long time frames (multi-centennial scales, *e.g.* Fig. 5.12) can hide very important features of the exact communication between modes. In our case, we find that mean states of modes and amplitudes are related to both the strength of the relationship and the time lag (Fig. 5.13c-e). We also emphasize the possibility that other climate processes (such as NAO or AMOC) and external factors (such as natural or anthropogenic forcings) can play a role in modifying the statistical properties of AMV and IPV that can thus weaken or strengthen their relationship. All the knowledge gained from this analysis will have important applications for understanding past and future changes in these modes and their decadal prediction, based on expert knowledge using observational constraint rather than only on numerical models predictions that still suffer from strong deficiencies (Yaeger and Robson 2017). This might also be a benchmark to evaluate the realism of the modeling of this interbasin teleconnection.

5.2.4 Discussion

This study has proposed to extend a regression-based approach of climate index reconstruction to a 1° resolution grid of SST data (HadISST, Rayner et al. 2003), with a collection of 392 well-verified proxies (PAGES 2k Consortium 2017). These records are mostly tree rings, which, although can be limited for transcribing ocean variability, can benefit from teleconnections with the ocean on land to reconstruct variations in the ocean. Because of very costly calculations, the robust multi-member pseudo-proxy validation approach (*e.g.* Ortega et al. 2015) has not been carried out. However, we applied here a machine learning algorithm (RF) coupled with cross-validation techniques (section 5.2.2, Michel et al. 2020), able to determine whether a grid point can be reconstructed with a satisfying level of robustness. Relevant validation metrics and selection of proxies have been applied to also ensure the reduction of uncertainties as much as possible for each grid point.

As for any reconstruction study, a lot of sources of uncertainties exist and need to be taken into account to evaluate the limits and caveats of our results. For instance, the combination of quantified and unquantified uncertainties in each proxy record can result in inaccurate reconstructions. This is why we used up-to-date statistical modeling techniques to reduce reconstruction uncertainties, which also include a selection of relevant records for reconstructing each grid point (section 5.2.2.2). We have also seen that for long time frames of reconstructions (>1000 years), a significant fraction of grid points can not be efficiently reconstructed after RF model validation (Fig. 5.2). This is mainly caused by the lack of proxy data and their unequal distribution in space for long time frames (Fig. 5.2). The resulting non-consideration of grid points that have poor reconstructability also should be taken into account, especially for further periods of reconstruction. Indeed, if not considering some points when calculating spatial averages and climate indices, their estimation might thus not be entirely accurate. The latter caveat could be for instance overtook using a partial data assimilation of SST, only using grid points with significantly positive scores of reconstruction.

Beyond these caveats, we highlight that we find consistencies between our reconstruction and numerous recent results from modeling studies, for a wide variety of topics : response to volcanic and solar forcings (section 5.2.3.2.1, 5.2.3.2.2), long-term warm and cold periods (section 5.2.3.2.3) and the lagged relationships between IPV and AMV (section 5.2.3.4). It is encouraging that many of the results we find in our reconstruction are physically reproducible within models. This in turn validates the dynamics of models using our reconstruction based on real-world data, and also our reconstruction in itself, since the sources of error between models and paleoclimate reconstructions are very different.

Reconstructions of paleoclimate are and will remain, increasingly relevant as long as the increasing number of both historical and paleoclimate data leads to increasing robustness in statistical model fitting. For instance, we find (not shown) in this study major discrepancies with the gridded temperature reconstruction of Mann et al. (2008) on regional scales but also in terms of globally averaged anomalies. Even without considering the differences in the statistical approaches of this study and the latter one, the use of a larger amount of new data collected since the study of Mann et al. (2008) provides here a largely updated gridded reconstruction. Nevertheless, we also find a very consistent result with Mann et al. (2008) regarding the growth in historical temperatures, which we validate here to be unprecedented since at least 850 C.E.

In terms of radiative forcings, we found that both long-term solar minima and strong volcanic eruptions appear to significantly impact the climate at the global scale (section 5.2.3.2, Fig. 5.5, 5.6). However we quantify the large solar minima to only account for a loss of $\sim 0.024^{\circ}\text{C}$ on average in the global oceanic surface temperature, while very strong volcanic eruptions, such as the Kuwae one in 1458, can trigger a global temperature loss of 0.25°C in 30 years (Fig. 5.4).

From the gridded reconstruction of Mann et al. (2008), Rahmstorf et al. (2015) extracted the AMOC_C index and found that the AMOC has been significantly declining over the last two centuries, reaching unprecedented low value over the historical period. In our reconstruction, we find the same drop in instrumental data when

extracting the same $AMOC_C$ index. However, we found that using $AMOC_C$ is only representing variations of SPG and does not satisfyingly reproduce the AMOC SST fingerprint over the whole Atlantic Ocean (section 5.2.3.3.2, Fig. 5.11). Variations of SST in the SPG can indeed be related to other processes than the AMOC, like heat fluxes variations linked to atmospheric variability, which is not necessarily related to the AMOC changes. Furthermore, modeling studies indicate that AMOC modulations impact a wider region of the North Atlantic than only the SPG region. We propose in this study an alternative index of the last millennium AMOC variations defined as the second PC of smoothed North Atlantic data ($PC2_{NA}$). Regression coefficients of $PC2_{NA}$ robustly reproduce the AMOC fingerprints over the Atlantic basin (interhemispheric seesaw, changes in the Gulf Stream region) and thus indicate that it might be a better estimation of last millennium AMOC variations than Rahmstorf et al. (2015) (section 3.3.2, Fig. 5.11). By projecting instrumental data in the North Atlantic EOF basis, we also find unprecedented low values for extended $PC2_{NA}$ (section 3.3.2, Fig. 5.11) over the recent period, which confirms the main result from studies based on $AMOC_C$ (Rahmstorf et al. 2015) or independent proxy of deep ocean circulation (Thornalley et al. 2018).

Furthermore, we examined the lagged relationship between AMV and IPV to gain insights on pan-tropical linkages between the different ocean basins, in order to verify the relationship for decadal variability that has been recently suggested, based on a very short time frame (*e.g.* Li et al. 2016). We have found a non-stationary lagged relationship when the AMV leads the IPV (section 3.4, Fig. 5.12, 5.13) by 14 years on average. We found that this non-regularity is partly related to centennial standard deviations and mean states of both modes, which both increase the strength of the lagged relationship in terms of correlation (Fig. 5.13). We additionally found that mean states of both modes also play a role in increasing the time lag of the relationship (Fig. 5.13c-e). Evaluating the ability of dynamical climate models to reproduce such a relationship between North Atlantic and Pacific basins might be an interesting feature for future studies, in order to better understand the underlying physical mechanisms through dedicated sensitivity experiments.

To conclude, we present here a reconstruction that provides an up-to-date SST field since 850 C.E. and aims to be used as a framework for understanding multidecadal variability and evaluating climate models. In addition to the different analysis we have made, many other uses of this reconstruction can be intended such as early warning analysis, the study of other indices, or last millennium data assimilation. Also, it might be interesting to check the reproducibility of this reconstruction through a pseudo-proxy experiment. For this purpose, it has to be applied to an ensemble of climate models' last millennium simulations to avoid a model bias issue. However, the present reconstruction requires relatively heavy computation resources and a long computation time (~ 10 days) and we estimated that such a validation needs large capacities of computation to be performed in an acceptable period of time and therefore left for a future study.

5.3 Intérêts apportés par les résultats d'une telle étude

La reconstruction, suite à une sélection d'enregistrements par test de corrélation à chaque point de grille, utilise, *in fine*, un total de 392 enregistrements paléoclimatiques (Fig. 5.2). Nous nous sommes servis de toute une batterie d'analyses statistiques pour explorer différentes hypothèses émises par des études antérieures, basées sur des modèles numériques et des observations instrumentales (donc trop courtes pour analyser correctement des variations multi-décennales). Ceci a notamment l'avantage de donner du crédit à notre reconstruction mais aussi de valider la capacité de certains modèles à reproduire ce que l'on observe dans notre reconstruction. On peut citer par exemple la réponse positive dans la région *El Niño* pour les années succédant à de puissantes éruptions, très significative dans notre reconstruction (Fig. 5.5). Ceci avait déjà été proposé à partir de modèles et d'observations (Ohba et al. 2013, Khodri et al. 2017), si bien qu'il est difficile d'imaginer que l'on trouve par hasard un signal aussi significatif dans notre reconstruction. On observe également, en réponse aux grosses éruptions volcaniques, du signal négatif dans l'Atlantique Nord (Fig. 5.5) comme suggéré par Mignot et al. (2011) à partir de simulations millénaires de modèles, ainsi que suggéré par les reconstructions de NASST de Wang et al. (2017) et Michel et al. (2020).

Nous avons également pu estimer que le LIA et le MCA furent des périodes relativement synchrones sur la SST aux échelles de temps pluri-centennales. Cela ne remet néanmoins pas le résultat de Neukom et al. (2019) en cause, puisque ces derniers s'intéressent à des échelles de temps beaucoup plus courtes (*e.g.* 50 ans). Encore une fois il semble se dessiner que la période que l'on appelle le LIA est en fait une succession de périodes froides, dont notamment celle de 1150-1350, identifiée dans le chapitre 4. Ceci se voit particulièrement bien sur la moyenne globale des SST reconstruites (Fig. 5.4), où la baisse de température semble graduelle, témoignant d'une certaine continuité dans la tendance climatique du dernier millénaire. On peut en effet voir que le LIA était à son paroxysme au début du XIX^e siècle, notamment en raison de l'éruption de Tambora, comme l'ont indiqué Brönnimann et al. (2019). Les minima ont eu un effet significatif mais relativement moindre en comparaison à l'effet des éruptions volcaniques et des variations climatiques internes (Fig. 5.5, 5.6), comme suggéré par Schurer et al. 2014.

Un autre résultat important est le lien que nous avons trouvé entre l'IPV et l'AMV reconstruites. Ce que l'on apporte notamment de nouveau c'est que l'on montre que cette relation n'est uniquement significative que pour quelques fenêtres de temps. Ceci témoigne d'une relation inter-bassin non-stationnaire (Fig. 5.11, 5.12, 5.13). Ceci peut s'expliquer par le fait que l'on n'a pas affaire à un système à seulement deux composantes, la variabilité de chacun des modes étant également impactée par d'autres phénomènes. En revanche, nous avons pu mettre en évidence que la relation entre les bassins semble s'intensifier à mesure que la variance et l'état moyen de chacun augmente. En particulier, on trouve que les variances jouent un rôle relativement faible bien que significatif (Fig. 5.13ab), tandis que l'état moyen semble déterminant pour expliquer le lien entre les deux modes (Fig. 5.13c-e). Autrement dit, le lien que l'on observe entre ces derniers sur la période instrumentale (Marini and Frankignoul 2013) est en fait très fortement influencé par le fait que l'un des deux modes soit en phase majoritairement positive, et en particulier l'AMV (Fig. 5.13c-e). Le fait qu'une AMV positive ait particulièrement une forte propension

à stimuler une phase négative d'IPV dix ans après a déjà été mis en avant par des études récentes basées sur des modèles et des observations (McGregor et al. 2014, Chafik et al. 2016, Li et al. 2016), qui sont en un sens incertaines étant donné les limites des deux approches (section 1.1.4). Ces études stipulent qu'une AMV positive est associée à une intensification du gradient thermique inter-bassin, et donc une intensification de la cellule de Walker dans le Pacifique, synonyme d'IPV négative (Kosaka and Xie 2013, England et al. 2014). Ce résultat important vient apporter un élément de preuve supplémentaire pour dire que la phase positive d'AMV ayant émergée dans les années 1990, à probablement été un catalyseur, voire le déclencheur, de la période de hiatus observée entre 1998 et 2012 (McGregor et al. 2014, Chafik et al. 2016, Li et al. 2016). Ceci rend d'autant plus attractif le fait de s'intéresser à l'attribution exacte des processus à l'origine des variations de l'AMV. Une étude intéressante et qui pourrait renforcer ou mesurer notre résultat serait de reproduire l'analyse statistique que nous avons présentée au sein de plusieurs simulations de modèles de circulation globale. Ceci permettrait notamment de voir dans quelles mesures et dans quelles proportions ces derniers reproduisent le résultat présenté ici.

Le dernier résultat que nous allons discuter offre des perspectives considérables. La PCA appliquée à la SST reconstruite nous permet en effet d'extraire une série en tant que seconde PC, dont on a toutes les raisons de penser que les variations sont étroitement liées au comportement de l'AMOC au cours du dernier millénaire. Ce qui nous permet notamment de suggérer que ceci est la signature de l'EOF associée à cette PC, qui dessine un tripôle de SST très clair dans l'Atlantique Nord (Fig. 5.10d). Cette signature est en effet similaire à celle de l'AMOC sur la SST évoquée par Knight et al. (2006), Zhang (2008) et Persechini et al. (2013). Si l'on étend la signature de cette PC sur l'ensemble de la SST, on voit qu'elle est négativement corrélée avec les points de grille de l'Atlantique Sud (Fig. 5.11b). Puisque la partie nord tropicale de l'Atlantique est quand elle corrélée positivement avec cette PC, ceci rejoint également les hypothèses de Knight et al. (2006) et Persechini et al. (2013), qui proposent que les variations de l'AMOC sont responsables de modifications dans le gradient inter-hémisphérique de SST de l'Atlantique, impliquant une variabilité multi-décennale dans le positionnement latitudinal moyen de l'ITCZ. On obtient un résultat également intéressant lorsque l'on projette les données historiques dans la base de la PCA (ce qui correspond à appliquer le produit matriciel entre les données de SST historiques et la matrice des EOF construite sur le millénaire). On trouve alors une seconde PC projetée avec une très forte tendance négative, quantifiée comme sans précédent depuis au moins l'année 850. Si l'on considère que la PC associée à la seconde EOF de notre analyse traduit bien des variations de l'AMOC, ce résultat apporte des éléments de preuve supplémentaires aux hypothèses d'un ralentissement sans précédent et d'origine anthropique de l'AMOC, proposées par Rahmstorf et al. (2015), Thornalley et al. (2018) et Caesar et al. (2018). Ceci étant dit, il reste néanmoins à montrer que cette série est une bonne estimation des variations de l'AMOC du dernier millénaire. Des travaux en cours de production, et qui devrait prochainement faire l'objet d'études scientifiques intéressantes, sont dédiés à comparer la série d'AMOC présentée ici, et des enregistrements paléoclimatiques voués à traduire l'intensité des courants océaniques. Ces types d'enregistrements sont par exemple obtenus à l'aide de mesures de la taille des grains prélevés dans des limons sédimentaires. Puisque nous avons fait le choix de ne pas inclure des

enregistrements sédimentaires pour entraîner nos modèles statistiques, la comparaison de notre indice AMOC avec ces enregistrements de circulation indépendants de nos modèles a d'autant plus d'intérêt. Les études qui produisent ce type d'enregistrements paléoclimatiques s'intéressent, entre autres, à reconstruire l'intensité d'un courant nommé l'ISOW (*Iceland-Scotland Overflow Waters*) (Mjell et al. 2016, Moffa-Sánchez and Hall 2017). Ce courant d'eaux profondes est situé dans une région idéale pour faire des prélèvements sédimentaires. Il est issu de la convection importante dans les mers Nordiques et constitue un des bras de la circulation profonde de l'AMOC. Des premiers résultats, qui feront l'objet d'études postérieures à cette thèse, montrent des variations significatives (corrélation > 0.4 , $p\text{-val} < 0.05$, sur la période 850-1980) entre notre indice AMOC et des reconstructions de l'intensité de l'ISOW à partir de l'enregistrement paléoclimatique évoqué précédemment.

Chapitre 6

Rappel vers la SST reconstruite sur 1000 ans dans le modèle IPSL-CM5A2

Ce chapitre reprend la reconstruction de SST de Michel et al. (sub) présentée chapitre 5, de laquelle sera déviée la trajectoire du modèle IPSL-CM5A2 à l'aide d'une méthode de rappel en SST (*nudging*). On notera par la suite *Rap.* la réanalyse obtenue. La section 6.1 est une brève introduction sur les méthodes d'assimilation de données et soulève l'intérêt qu'elles peuvent avoir dans le cadre d'une étude paléoclimatique. La section 6.2 décrit le modèle climatique utilisé et la méthode de rappel employée. La section 6.3 se centrera sur la comparaison entre la SST simulée et celle de la reconstruction sur laquelle est appliqué le rappel. On y comparera également la réanalyse obtenue avec celle issue d'une autre étude (Hakim et al. 2016), obtenue par méthode de KF. Dans la section 6.4, on comparera l'AMOC au sein de la réanalyse obtenue avec l'indice estimé à partir de la grille de reconstruction de SST présentée dans le chapitre 5. La section 6.5 présente des perspectives issues des résultats de ce chapitre.

6.1 Méthode d'assimilation de données : application simple dans un contexte paléoclimatique

De manière générale, les méthodes de rappel sont beaucoup moins coûteuses en temps de calcul que des méthodes d'assimilation de données plus complexes comme les filtres de Kalman d'ensemble (EKF) ou les filtres de Kalman (KF) (section 1.3.5). Il serait donc intéressant de savoir si le bénéfice en matière de robustesse et de précision des méthodes plus complexes par rapport à la méthode de rappel est suffisamment important. Si ce gain n'est pas significatif, il serait judicieux d'envisager des méthodes d'assimilation par rappel dans un cadre paléoclimatique, puisque les périodes de simulations sont très longues. Si l'efficacité des méthodes KF ou EKF s'avérait être relativement proche de celle des méthodes de rappel, il serait plus facile de mettre en place des études d'assimilation paléoclimatique multi-modèles, et ainsi d'estimer l'incertitude associée aux différents modèles pour ce type d'approche. Il faut tout de même souligner qu'une étude a montré que dans un contexte paléoclimatique et au sein du modèle de complexité intermédiaire (simplifié sur une très basse résolution spatiale) iLOVECLIM (Roche 2013, Caley and Roche 2013, Roche and Caley 2013), les PF et EKF sont largement plus efficaces que les méthodes de rappel (Dubinkina and Goosse 2013). Cependant, la méthode de rappel n'a pas encore été appliquée dans un contexte de reconstruction millénaire avec un modèle de type GCM. En revanche, Hakim et al. (2016) ont proposé une assimilation de données par KF sur le dernier millénaire. Ces derniers assimilent la température de

surface à 2 mètres d'altitude (*Temperature at 2 meters, T2m*) au sein du modèle de circulation globale CCSM4. Ce sont les résultats de cette réanalyse (notée *KFH16*) que nous comparerons dans la section 6.3.

6.2 Données et méthode

6.2.1 Le modèle IPSL-CM5A2-VLR comme base de simulation

Le modèle de circulation global utilisé pour cette étude est le modèle CMIP5 de l'Institut Pierre Simon Laplace (IPSL), dans sa version CM5A (CM signifie *Climate Model*, et 5A fait référence à la version du modèle d'atmosphère utilisé, (cf. LMDZ5A). Comme pour la plupart des modèles CMIP5, les projections pluri-décennales à l'échelle globale sont très dépendantes du scénario d'émissions considéré (Dufresne et al. 2013), comme l'avaient déjà souligné Hawkins and Sutton (2009) pour les modèles CMIP3. En tant que modèle de circulation globale du climat, ce modèle est le résultat du couplage entre différents sous modèles, ce qui représente un effort scientifique considérable (Fig. 6.1).

On distingue deux types de modèles couplés pour produire celui de l'IPSL-CM5A : des modèles physiques, implémentés selon des équations physiques connues, et des modèles biochimiques qui permettent notamment de faire communiquer les modèles physiques au travers d'échanges de CO_2 , de l'ozone et des aérosols (Fig. 6.1).

Pour l'atmosphère, le modèle de circulation générale atmosphérique LMDZ développé par le laboratoire de météorologie dynamique est utilisé. Ce dernier contient 39 couches d'atmosphère selon l'altitude, que l'on juge suffisantes pour résoudre la propagation d'ondes dans la stratosphère (Dufresne et al. 2013). Ici, on utilisera la version LMDZ5A (Hourdin et al. 2013). Le modèle biochimique atmosphérique REPROBUS (*Reactive Processes Ruling the Ozone Budget in the Stratosphere* (Jourdain et al. 2008) permet de calculer la distribution des différents gaz, des aérosols et des nuages au sein de la stratosphère du modèle LMDZ. Le modèle INCA, quant à lui, permet de simuler les concentrations de gaz et d'aérosols au sein de la troposphère du modèle LMDZ. Ces deux modèles (INCA et REPROBUS) ne seront pas activés ici. Ils sont souvent utilisés en *off line* pour calculer les concentrations aérosols. Le modèle de surface continental est le modèle ORCHIDEE (*ORganic Carbon and Hydrology In Dynamic EcosystEms*). Ce dernier est implémenté de manière à simuler le cycle énergétique et le cycle de l'eau au sein des sols et de la végétation.

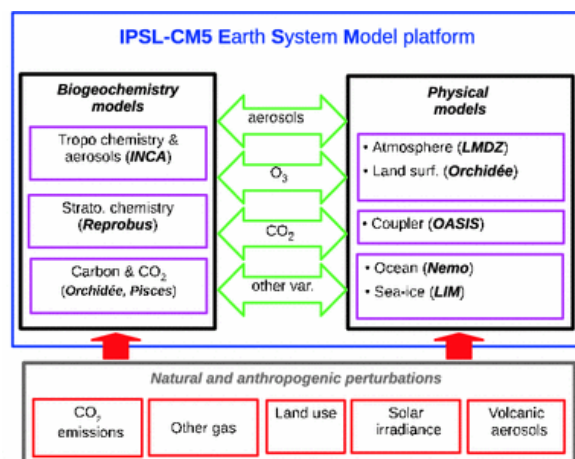


FIGURE 6.1 | Schéma de la structure du modèle IPSL-CM5 dans sa "configuration pleine", c'est-à-dire lorsque tous les modèles qui le composent sont actifs (Dufresne et al. 2013). Les sous-modèles sont représentés dans les rectangles violets, les échanges chimiques calculés par le modèle sont représentés dans les rectangles verts. Enfin, les variables prescrites, que l'on peut voir comme les conditions aux limites que l'on va imposer aux modèles, sont représentées dans les rectangles rouges. Source : figure de Dufresne et al. (2013).

Le modèle biochimique atmosphérique REPROBUS (*Reactive Processes Ruling the Ozone Budget in the Stratosphere* (Jourdain et al. 2008) permet de calculer la distribution des différents gaz, des aérosols et des nuages au sein de la stratosphère du modèle LMDZ. Le modèle INCA, quant à lui, permet de simuler les concentrations de gaz et d'aérosols au sein de la troposphère du modèle LMDZ. Ces deux modèles (INCA et REPROBUS) ne seront pas activés ici. Ils sont souvent utilisés en *off line* pour calculer les concentrations aérosols. Le modèle de surface continental est le modèle ORCHIDEE (*ORganic Carbon and Hydrology In Dynamic EcosystEms*). Ce dernier est implémenté de manière à simuler le cycle énergétique et le cycle de l'eau au sein des sols et de la végétation.

Pour simplifier, on utilisera ici ce modèle dans sa version non dynamique, c'est-à-dire que les types de sols restent fixes et que seule la quantité de feuilles change. Il simule également les cycles terrestres du carbone, ainsi que la composition et la répartition spatiale de la végétation, répartie en 12 types fonctionnels (Krinner et al. 2005). Lors d'une simulation du modèle IPSL-CM5A2, la grille de résolution spatiale d'ORCHIDEE est toujours la même que celle de LMDZ, de manière à faciliter le couplage des deux modèles via les flux d'eau et de chaleur.

Pour l'océan, IPSL-CM5A se compose du modèle de circulation globale océanique NEMO v3.2 (*Nucleus for European Modelling of the Ocean*, Madec 2008). Ce dernier est lui-même composé de différents modèles, dont le modèle de dynamique océanique OPA, le modèle de biogéochimie océanique PISCES ainsi que le modèle LIM, permettant de modéliser la dynamique et thermodynamique de la glace de mer.

6.2.2 Application de la méthode de rappel

On applique ici une méthode de rappel à γ constant (Servonnat et al. 2015), avec un rappel en surface de SST. Les données assimilées sont celles de la reconstruction de SST produite dans le chapitre 5 à l'aide de méthodes de RF. Les données de la reconstruction du chapitre 5 sont à résolution annuelle. Cependant, puisque la variabilité saisonnière, notamment aux hautes latitudes, implique des différences de température de plusieurs dizaines de degrés, ceci n'aurait physiquement aucun sens d'assimiler des moyennes annuelles. On calcule donc préalablement la SST mensuelle moyenne sur la grille ORCA2 au sein d'une simulation contrôle de 400 ans. Chaque donnée annuelle de reconstruction est ensuite recentrée sur les 12 moyennes mensuelles calculées précédemment, débouchant ainsi sur une grille d'anomalies de reconstruction mensuelle. Le rappel est ensuite appliqué annuellement à chaque point de grille entre les années 851 et 1970. Le rappel s'applique ici par un terme de restauration de Haney, qui agit tel un flux de chaleur supplémentaire dans l'équation de conservation de la chaleur du modèle :

$$\frac{\partial SST}{\partial t} = \dots + \frac{\gamma}{\rho C_p h} (SST_{model} - SST_{rec})$$

Où $\gamma = 40 \text{ W.m}^{-2}\text{K}^{-1}$ est le coefficient de rappel, dont la valeur, ici fixe, est basée sur la valeur physique correspondant à une échelle de temps de relaxation de 60 jours pour une couche de mélange de 50 mètres (Frankignoul et al. 2002), ce qui est considéré comme un choix physiquement cohérent (Servonnat et al. 2015). $\frac{\partial SST}{\partial t}$ est la dérivée partielle de la SST par rapport au temps. ρ est la densité moyenne de l'océan au point de grille (en kg.m^{-3}). C_p est la chaleur de l'océan au point de grille (en $\text{J.K}^{-1}.\text{kg}^{-1}$). h est l'épaisseur de la couche de surface au point de grille (en m). SST_{model} et SST_{rec} correspondent respectivement à la SST simulée par le modèle et la SST de la grille de reconstruction du chapitre 5 à laquelle est ajoutée la climatologie du modèle et le cycle saisonnier (paragraphe précédent), interpolées sur la grille océanique du modèle IPSL-CM52-LR (ORCA2). Finalement, les points de suspension représentent le reste de l'équation, incluant l'advection, la diffusion, les flux atmosphériques... (voir Madec 2008).

6.3 Validation et comparaison avec une étude précédente

Dans cette section, on compare les résultats obtenus pour les réanalyses *Rap.* et *KFH16*. *KFH16* est basée sur une assimilation de données paléoclimatiques par filtre de Kalman. Contrairement aux données de reconstruction sur lesquelles on base *Rap.* (SST), *KFH16* se base sur une reconstruction de T2m. La méthode de reconstruction sous-jacente utilisée est basée sur une régression linéaire simple (à une variable), entre les données historiques de T2m d'un point de grille et la donnée paléoclimatique qui en est la plus proche dans l'espace (Hakim et al. 2016).

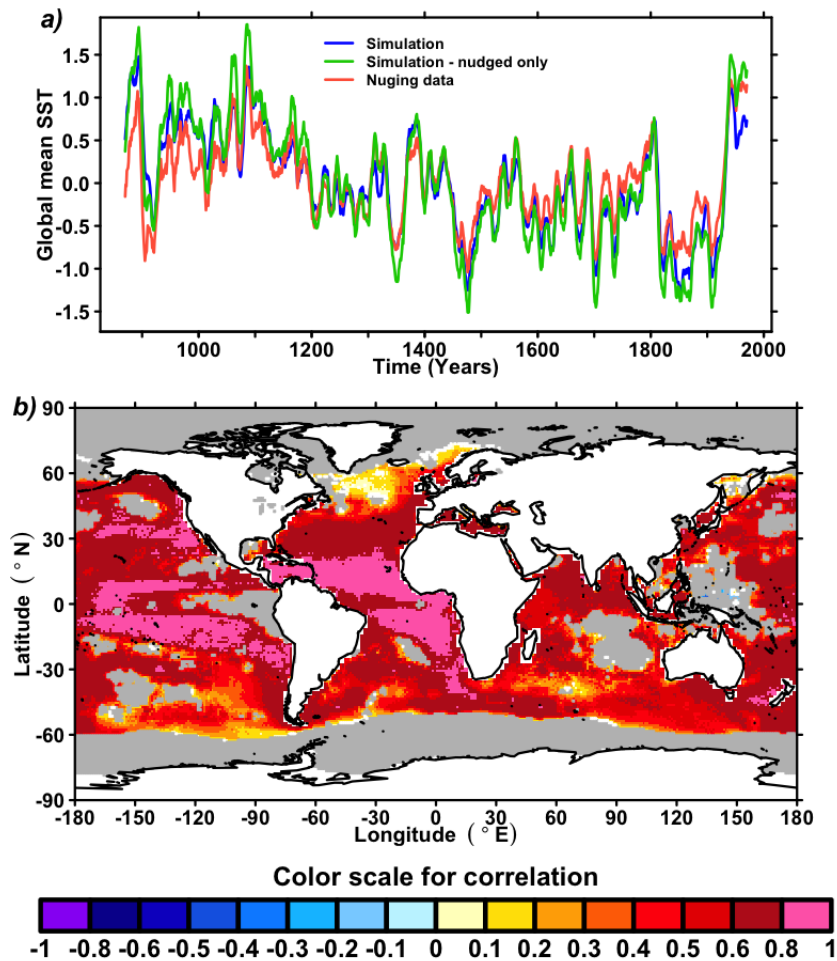


FIGURE 6.2 | Validation de la réanalyse. a) Comparaison entre la SST globale simulée (bleu) et la SST globale des données assimilées (rouge). La courbe verte représente la SST globale simulée en ne considérant que les points de grille sur lequel le rappel est appliqué (0,4% d'entre eux, cf. chapitre 5). b) Carte de corrélation entre la SST de la réanalyse et la SST de la reconstruction du chapitre 5 pour la période 870-1970. Les points de grille gris sont ceux pour lesquels le rappel n'a pas été appliqué.

Dans le but de vérifier le bon fonctionnement de la procédure d'assimilation, la Fig. 6.2 présente une comparaison entre la SST de *Rap.*, et les données que l'on a assimilées (cf. reconstruction du chapitre 5). En terme de moyenne globale, les sorties de réanalyse et les données assimilées sont très fortement corrélées ($r=0.87$, $p<0.01$, Fig. 6.2a). Puisque la réanalyse est basée sur un rappel partiel (i.e. certains points de grilles ne sont pas assimilés, cf. chapitre 5), la moyenne globale simulée est

d'autant plus fortement corrélée avec celle des points assimilés uniquement ($r=0.95$, $p<0.01$, Fig. 6.2a). Dans la lignée des réanalyses de Servonnat et al. (2015), basées sur le même modèle de climat, on trouve la même hétérogénéité spatiale dans la propension du modèle de climat à reproduire les données assimilées. En particulier, on trouve des corrélations bien plus élevées dans les régions tropicales, qui ont tendance à décroître avec la latitude (Fig. 6.2b).

On constate des cartes de corrélation similaires sur les océans entre les SST et les T2m issue des réanalyses *Rap.* et *KFH16* pour la période 870-1970. On remarque en particulier la cohérence entre les deux réanalyses dans l'océan Atlantique Nord, où l'on trouve les plus fortes corrélations (Fig. 6.3a). Sur les continents, on distingue trois régions où les deux réanalyses de T2m ne sont pas significativement corrélées à 90% : le sud de l'Amérique du sud, l'Australie, ainsi que l'ensemble du continent asiatique (Fig. 6.3b).

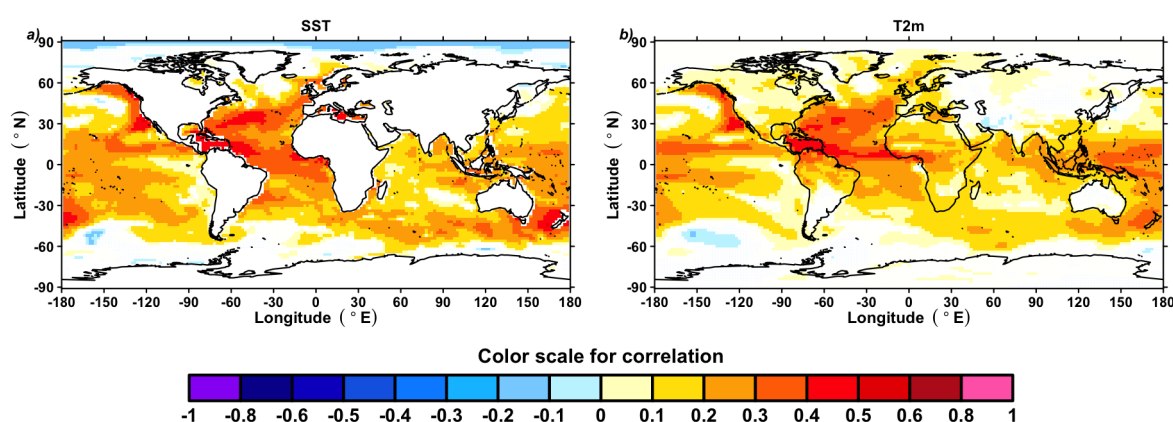


FIGURE 6.3 | Étude de la cohérence entre deux réanalyses distinctes. Cartes de corrélation entre la SST (a) et la T2m (b) de la réanalyse *Rap.* et de la réanalyse *KFH16*. Les points de grille blancs indiquent que la corrélation n'est pas significative au seuil de 90%.

Alors que Hakim et al. (2016) assimilent des données de T2m et que nous assimilons des données de SST, on remarque que chacune des approches tend à mieux reproduire que l'autre, sur sa période commune avec l'historique (1870-2000 pour *KFH16*, et 1870-1970 pour *Rap.*), la variable qu'elle a assimilée. En effet, la Fig. 6.4 montre que malgré les lacunes que l'on rencontre pour simuler la T2m observée, on simule avec une très grande précision la SST observée sur l'historique. On remarque toutefois que pour les deux réanalyses, on a très peu de points significativement corrélés en T2m dans l'océan Austral. Ceci peut venir du fait qu'on dispose de peu d'observations dans ces régions.

On remarque par ailleurs que la SST simulée par *Rap.* dans le Pacifique et l'Atlantique Nord se corrèle particulièrement bien avec les observations instrumentales sur la période historique, avec majoritairement des corrélations supérieures à 0.8, en particulier dans les régions tropicales (Fig. 6.4c). Cependant la région du SPG de l'Atlantique Nord est entièrement non corrélée avec les données historiques, que ce soit pour *KFH16* ou *Rap.* (Fig. 6.4cd). Ceci peut notamment provenir du fait que la région possède une variabilité rendue très complexe avec sa communication avec le *Gulf Stream*, la mer du Labrador et la mer d'Irminger, et son rôle dans la circulation océanique. À ce sujet, l'étude de Sgubin et al. (2017) montre en effet que peu de

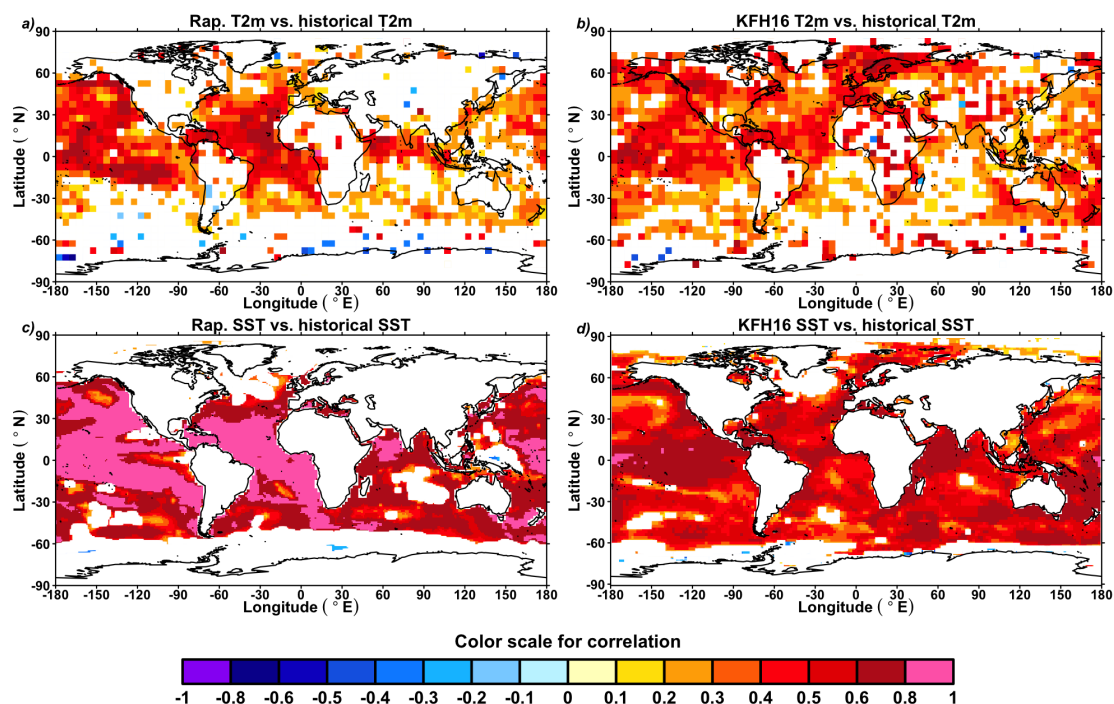


FIGURE 6.4 | Cartes de corrélation entre les observations instrumentales et les réanalyses *Rap.* et *KFH16*. a) Comparaison de *Rap.* avec les observations de T2m (HadCRUT4, Morice et al. 2012). b) Comparaison de *KFH16* avec les observations de T2m (HadCRUT4, Morice et al. 2012). c) Comparaison de *Rap.* avec les observations de SST (HadISST, Rayner et al. 2003). d) Comparaison de *KFH16* avec les observations de SST (HadISST, Rayner et al. 2003). Les points de grilles blancs indiquent que la corrélation n'est pas significative au seuil de significativité de 90%.

modèles CMIP5 représentent bien la stratification observée des couches océaniques dans la région. On remarque dans l'ensemble que *Rap.* peine à reproduire la variabilité des hautes latitudes comme souligné par Servonnat et al. (2015) en "modèle parfait" (Fig. 6.4). Pour *KFH16*, ce phénomène est moins marqué (Fig. 6.4). Il faut tout de même noter que pour des raisons de cohérence, la reconstruction sur laquelle on applique la méthode de rappel (*cf.* reconstruction du chapitre 5) ne comprend pas les points dont la latitude est supérieure à 60°N car les points de grille comprennent beaucoup de glace de mer, ce qui impose une borne inférieure aux séries (*i.e.* -1.8°C), chose que les modèles statistiques ne permettent pas de traiter. Ceci peut donc en partie expliquer que les points de hautes latitudes sont mal reproduits dans notre réanalyse, puisque non assimilés.

Cependant, même si *Rap.* se base sur de la SST, on garde une certaine cohérence dans les variations de la T2m avec celles de *KFH16* (Fig. 6.3b). Ceci est également vrai en terme de moyenne globale, où la T2m que l'on obtient avec *Rap.* a une corrélation de 0.4 avec celle de *KFH16* ($p < 0.01$) sur les séries à résolution annuelle. Une fois les séries filtrées à 10 ans avec un filtre à noyau de Nadaraya-Watson (Franke et al. 2002), la corrélation s'élève à 0.56 et reste significative au seuil de 95% ($0.01 < p < 0.05$) (Fig. 6.5).

Cette cohérence entre les deux réanalyses est un résultat fort et prometteur pour les reconstructions paléoclimatiques. En effet, malgré le fait que les méthodes pour les obtenir sont largement différentes (méthode de la reconstruction sous-jacente, méthode d'assimilation, variable assimilée, modèle de climat...), on simule, sur le

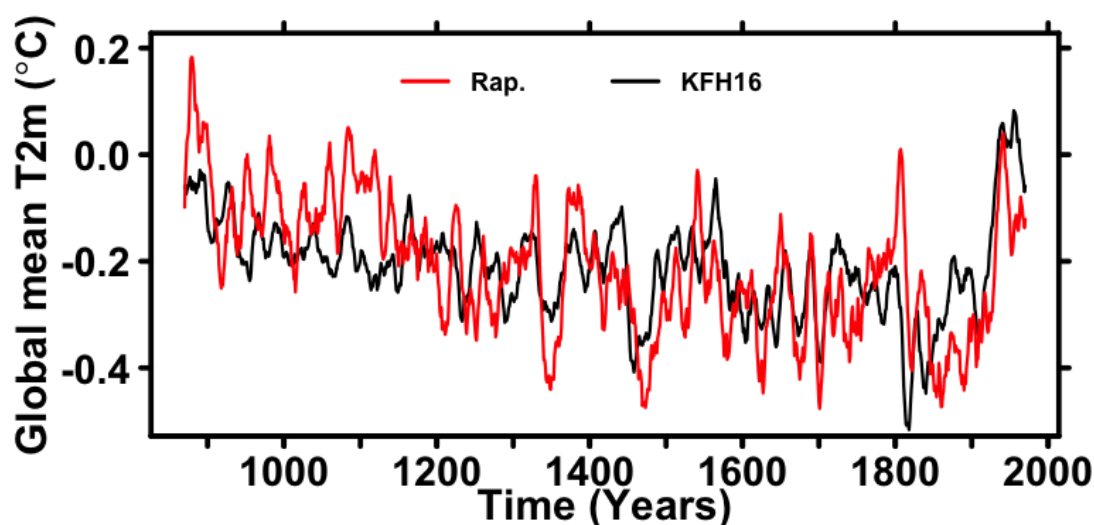


FIGURE 6.5 | Comparaison des moyennes globales issues des deux réanalyses. Moyenne globale de la T2m de *Rap.* (rouge), et de *KFH16* (noir).

plan global, des températures atmosphériques similaires pour la période 870-1970. Ceci montre qu'à mesure que les recherches avancent, et que de plus en plus de données paléoclimatiques sont disponibles et vérifiées (PAGES 2k Consortium 2017), les reconstructions qui les agrègent semblent converger. On remarque cependant que les séries filtrées ne sont pas significativement corrélées avant le début du XIV^e siècle ($r=0.16$, $p>0.2$). Ceci peut très certainement s'expliquer par le fait qu'il y a peu de données paléoclimatiques disponibles sur cette période (moins de 150) par rapport à la seconde partie du millénaire (plus de 400), comme le montrent très bien les Fig. 1.22, 4.2 et 5.2. On peut très fortement supposer que le nombre d'enregistrements disponibles sur une période de reconstruction donnée est déterminant pour que des approches de reconstructions distinctes rendent des résultats significativement similaires. Ceci encourage d'autant plus à poursuivre les études de terrain permettant de disposer de données paléoclimatiques, sans lesquelles les travaux de cette thèse n'auraient pu voir le jour. On remarque également que *KFH16*, qui assimile une reconstruction de T2m, suggère une différence de température globale beaucoup moins marquée entre le MCA et le LIA que pour notre reconstruction, bien que celle-ci reste significative au seuil de confiance de 99%. En effet, si l'on considère les définitions usuelles du MCA (900-1300) et du LIA (1400-1850), on trouve une différence des moyennes de l'ordre de 0.067°C pour la *KFH16*, et de l'ordre de 0.127°C pour *Rap.* Si l'on considère les définitions de MCA (870-1150) et du LIA (1200-1850) que suggèrent ces travaux de thèse, on trouve une différence de l'ordre de 0.081°C pour *KFH16*, et de 0.173°C pour *Rap.* On note que la différence moyenne entre les deux séquences est plus élevée dans la *KFH16* pour nos définitions de MCA et du LIA que pour les définitions usuelles ($p<0.01$), ce qui procure un gain de confiance dans les résultats présentés dans cette thèse.

Afin de vérifier si les deux approches ont tendance à reconstruire les variations paléoclimatiques enregistrées dans les archives naturelles, on considère la base de données d'enregistrements paléoclimatiques sur lesquelles celles-ci s'appuient en partie (P2k+, Sections 3 et 4, Michel et al. 2020, PAGES 2k Consortium 2017). P2k+ est majoritairement composée d'enregistrements continentaux et donc associés à de

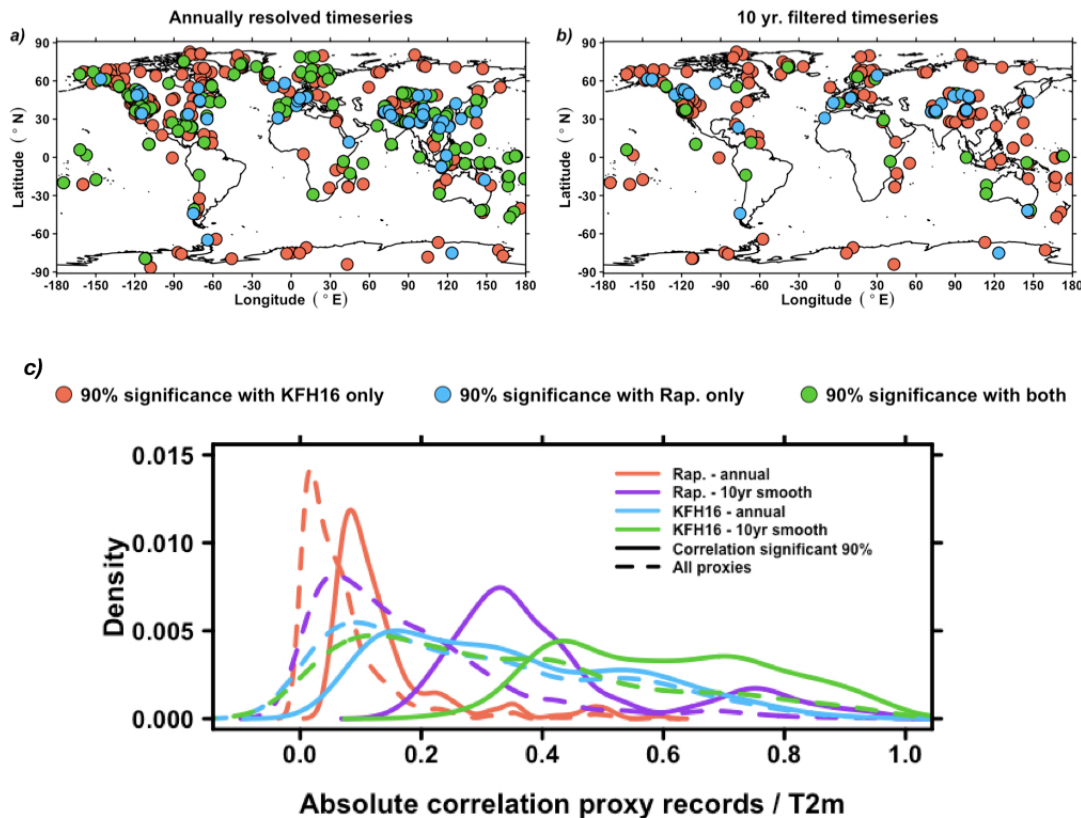


FIGURE 6.6 | Corrélation entre les enregistrements paléoclimatiques existants et les réanalyses étudiées. (a) Localisation des enregistrements paléoclimatiques significativement corrélés au seuil de 90% avec le point de grille de T2m annuel le plus proche dans *Rap.* uniquement (bleu), dans *KFH16* uniquement (rouge), ou les deux (verts). (b) Pareil que (a), excepté que les séries sont filtrées à 10 ans avec un filtre à noyau de Nadaraya-Watson (Franke et al. 2002). (c) Densités des corrélations entre les enregistrements paléoclimatiques, et le point de grille de T2m annuel le plus proche dans *Rap.* et *KFH16*.

la T2m. Ainsi, on trouve bien plus d'enregistrements corrélés au seuil de significativité de 90% pour *KFH16* (78% des 726 enregistrements pour les séries à résolution annuelle) que pour celle obtenue avec *Rap.* (30% des 726 enregistrements pour les séries à résolution annuelle) (Fig. 6.6).

6.4 Comparaisons de différentes estimations millénaires de l'AMOC

Dans cette section, nous allons explorer comment se comporte l'AMOC simulée. Puisque nous proposons un indice AMOC tiré de la grille présentée dans le chapitre 5, nous vérifierons si *Rap.* reproduit convenablement les variations de l'AMOC sur le dernier millénaire. Pour ce faire, on utilise à titre de comparaison la simulation millénaire réalisée par Myriam Khodri (Swingedouw et al. 2015), obtenue à partir du modèle IPSL-CM5A-LR. En terme de forçages externes, cette simulation utilise la reconstruction de la TSI de Vieira et al. (2011) et la reconstruction de l'activité volcanique de Gao et al. (2008), cette dernière possédant certaines erreurs de datation (Sigl et al. 2015).

On trouve ici que l'AMOC simulée au sein de *Rap.* n'est pas significativement corrélée avec l'indice que l'on propose dans le chapitre 5, à partir d'une analyse PCA/EOF ($\text{cor} = -0.15, p > 0.2$, pour les séries filtrées à 10 ans, Fig. 6.7). De même,

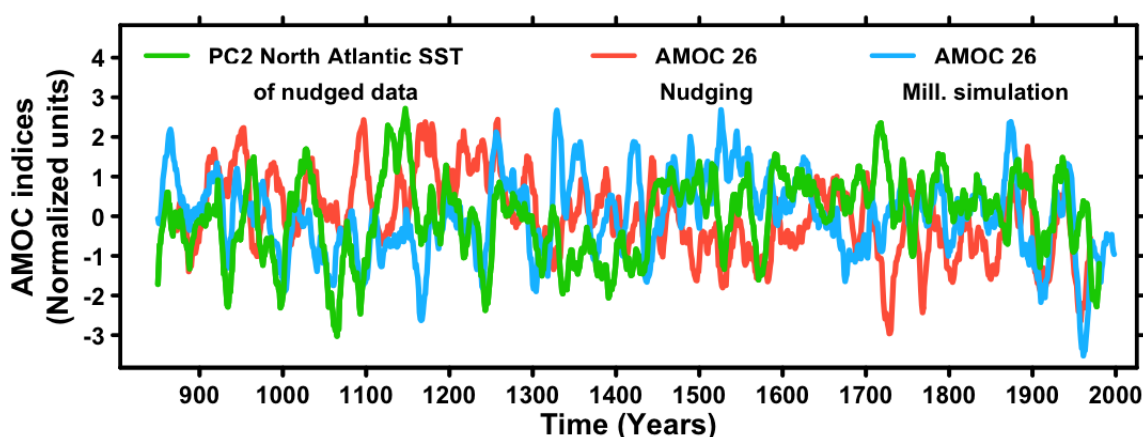


FIGURE 6.7 | Comparaison d'estimations des variations de l'AMOC sur le dernier millénaire. Intensité de l'AMOC extraites à 26°N pour *Rap.* (rouge). Intensité de l'AMOC extraites à 26°N dans une simulation millénaire du modèle IPSL-CM5A2-LR (Swingedouw et al. 2015) (bleu). Seconde PC de l'Atlantique Nord dans la grille de reconstruction de SST du chapitre 5 (vert).

l'AMOC obtenue dans la simulation millénaire n'est également pas significativement corrélée avec la seconde PC de la SST proposée dans le chapitre 5 ($\text{cor}=0.2$, $p>0.2$, pour les séries filtrées à 10 ans, Fig. 6.7). Sur la période historique (après 1870), on remarque tout de même une significativité de corrélation entre l'indice AMOC du chapitre 5 et l'AMOC de *Rap.* ($\text{cor}=0.69$, $p<0.1$). En revanche, sur cette même période historique, notre estimation statistique de l'AMOC n'est pas significativement corrélée avec l'AMOC de la simulation millénaire, qui n'assimile pas de données ($\text{cor}=0.4$, $p>0.2$). Du fait que l'AMOC de *Rap.* et celle estimée à partir de méthodes statistiques (chapitre 5) ne sont pas cohérentes, on pourrait supposer que le modèle ne parvient pas à simuler convenablement la convection océanique et donc l'AMOC. Ceci peut notamment être dû au fait que la densité des masses d'eaux est également très sensible aux variations de salinité (Swingedouw et al. 2007, Bakker et al. 2016, Sgubin et al. 2017, Collins et al. 2019). Pour vérifier ce qui vient d'être dit, il serait intéressant d'estimer les variations de la SSS sur le dernier millénaire pour qu'elles puissent être assimilées avec la SST. Cependant, bien qu'il existe certains enregistrements marins qui estiment les variations passées de cette variable climatique (Thornalley et al. 2009, Moffa-Sánchez et al. 2014, Moffa-Sánchez et al. 2019), elles existent dans une quantité bien moins importante que les enregistrements de paléo-températures (e.g. PAGES 2k Consortium 2017), si bien qu'elles constituent un échantillon trop petit pour envisager entraîner des modèles statistiques robustes.

6.5 Perspectives de ce chapitre

Au vu de ce qui a été dit dans la section 6.4, assimiler uniquement la SST ne semble pas suffisant pour reproduire les variations de l'AMOC convenablement. On pourrait imaginer proposer la procédure présentée ci-dessus dans un cadre où l'on assimile une ou plusieurs paires de grilles de reconstruction de SST et de SSS, notamment pour obtenir une série de variations de l'AMOC la plus robuste possible. En effet, nous avons vu que la SSS semble être un élément clef pour simuler convenablement les variations de l'AMOC. Pour valider cette hypothèse, il serait intéressant

de proposer une étude comparative entre l'AMOC de *Rap.*, et celle d'une réanalyse millénaire obtenue à partir d'une méthode de rappel appliquée à une grille de SST et une grille de SSS. Puisqu'il semble compliqué, notamment en terme de données disponibles (section 6.4), de reconstruire sur une grille une variable telle que la SSS, on pourrait extrapoler la forte corrélation que la SSS possède avec la SST dans les observations de la période historique, à partir des variations spatiales de SST reconstruites et présentées dans la chapitre 5. Dans ce cas, il n'y aurait ici qu'une variable explicative (SST) et qu'une variable à expliquer (SSS), si bien que seule une régression linéaire simple semble applicable pour estimer les variations du dernier millénaire. Une manière de vérifier que les méthodes de reconstruction d'une grille de SSS proposées ont du sens serait de tester leur reproductibilité au sein d'un modèle de circulation globale :

- En *pseudo-proxy* pour la reconstruction à partir de méthodes statistiques appliquées à des données paléoclimatiques
- En "modèle parfait" (Servonnat et al. 2015, Estella-Perez et al. 2020) au sein d'une simulation millénaire pour la reconstruction par extrapolation de la corrélation SSS/SST

En terme de performance des méthodes d'assimilation de données, on observe que mis à part pour la variable assimilée par rappel (SST), la reconstruction par filtre de Kalman employée par Hakim et al. (2016) est plus efficace pour reproduire les variations observées de températures atmosphériques continentales instrumentales ainsi que les variations d'enregistrements paléoclimatiques sur les périodes communes respectives qu'elle a avec eux. Néanmoins, puisque les deux réanalyses sont produites de manières très différentes (section 6.3), il serait intéressant de comparer les deux méthodes sur des configurations similaires, à savoir sur le même modèle de circulation globale, et la même grille de reconstruction climatique. Si l'on étend un peu plus loin cette idée, on pourrait imaginer une étude intéressante qui pourrait être menée à l'avenir. Dans l'idée de la comparaison de méthodes présentée dans le chapitre 3, on pourrait comparer les assimilations de données de différents modèles de climat, sur différentes grilles de reconstruction de mêmes variables climatiques, avec différentes méthodes d'assimilation. Une métrique intéressante pour comparer les différentes configurations d'assimilation serait de regarder leur propension à reconstruire des variations de températures similaires aux données paléoclimatiques de la base de données PAGES 2k, par exemple avec des scores de CE et de RE. Ceci représente bien sûr une importante quantité de temps et de travail et requière l'implication de plusieurs centres de recherche. Cependant, si l'on disposait d'un nombre satisfaisant de ces réanalyses aux configurations différentes, ceci permettrait de tirer de nombreuses conclusions. En effet, on pourrait déduire quels modèles, indépendamment de la grille de reconstruction et de la méthode d'assimilation, reproduisent le mieux les variations des enregistrements paléoclimatiques de la base de données PAGES 2k en moyenne. On pourrait de la même manière voir quelles méthodes d'assimilation reproduisent le mieux les variations des enregistrements paléoclimatiques de PAGES 2k en moyenne, indépendamment des modèles utilisés et des grilles assimilées. De la même manière, la grille reproduisant le mieux les observations paléoclimatiques indépendamment de la méthode et du modèle utilisés, pourrait être considérée comme la grille la plus cohérente avec ce que l'on connaît des lois de la physique. Au final, toujours en se basant sur une métrique globale de comparaison avec les enregistrements de PAGES 2k, on pourrait déterminer laquelle des configurations testées semble la plus robuste et la plus représentative de

ce que suggèrent les données paléoclimatiques existantes. Ceci constitue une autre perspective d'étude, beaucoup plus coûteuse que celle proposée au paragraphe précédent.

Chapitre 7

Conclusions, critiques et perspectives

Les données instrumentales permettent de voir qu’au delà de l’effet des forçages anthropiques, des processus et phénomènes naturels jouent un rôle important dans la variabilité climatique décennale, comme l’illustre le hiatus de la température globale observé entre 1998 et 2012 (Kaufmann et al. 2011, McGregor and Timmermann 2011, Kosaka and Xie 2013, England et al. 2014, McGregor et al. 2014, Chafik et al. 2016, Medhaug et al. 2017). Cependant la relativement courte période sur laquelle les observations sont disponibles (< 150 ans) ne permet pas de bien comprendre les processus internes et le rôle des forçages externes. L’objectif principal de cette thèse est de parfaire la compréhension de la variabilité climatique multi-décennale, et notamment des rôles relatifs de la variabilité climatique interne et des forçages radiatifs naturels. Pour répondre à cet objectif, on propose de s’intéresser à la période du dernier millénaire, afin de disposer de suffisamment de données pour étudier les processus qui participent à cette variabilité de manière statistiquement robuste. Pour ce faire, on applique de nouvelles approches de reconstruction climatique à l’aide de techniques de *machine learning* et d’assimilation par méthode de rappel appliquées à un nombre important de données paléoclimatiques. Dans le but de vérifier leur cohérence et de répondre à certaines questions, les reconstructions obtenues sont confrontées à des résultats provenant d’études basées sur différentes approches : simulations de modèles, enregistrements paléoclimatiques de terrain, reconstructions paléoclimatiques, études d’observations instrumentales... La section 7.1 présente un résumé de l’ensemble des résultats présentés dans cette thèse. La section 7.2 esquisse une critique objective des approches proposées en montrant leurs limites et les hypothèses sous-jacentes sur lesquelles elles se basent. La section 7.3 expose pour finir différentes perspectives en ouvrant la porte à de futures études.

7.1 Résumé des résultats présentés

La procédure de reconstruction présentée dans le chapitre 3, basée sur des algorithmes de *machine learning*, a pour ambition d’être aussi objective que possible en se basant sur des critères d’optimisation et de sélection de modèles fiables qui n’ont jamais été utilisés dans les études de reconstructions paléoclimatiques précédentes. Ces critères de sélection ont notamment permis de révéler la forêt aléatoire (RF) comme un outil puissant de reconstruction paléoclimatique, qui, selon les métriques d’évaluation que l’on propose, surclasse la PCR, abondamment utilisée (sans optimisation) dans les études précédentes. Ceci montre exactement en quoi la méthodologie proposée dans le chapitre 3 est essentielle pour garantir une grande robustesse de reconstruction en comparant les résultats donnés par différentes approches sur différents jeux de données paléoclimatiques. Il faut noter que seules

quatre méthodes sont explorées dans cette thèse et que la plus performante (RF) est la seule qui soit non linéaire. Il y aurait donc tout intérêt à poursuivre la démarche en confrontant de nouvelles méthodes non linéaires à la méthode de RF présentée ici : les réseaux de neurones, les arbres extrêmement randomisés ou encore les méthodes du gradient boosté. Elles pourraient éventuellement donner des résultats encore meilleurs, ce qui requiert l'implémentation de leur optimisation. Ceci représente une perspective intéressante pour l'outil développé dans le chapitre 3.

À l'aide de cet outil statistique appliqué à la base de données PAGES 2k, nous avons proposé les reconstructions par RF des indices NAO (chapitre 3) et AMV (chapitre 4). La reconstruction la plus robuste de l'AMV a été validée à l'aide d'une approche en *pseudo-proxy*. On a en effet montré qu'au sein de 12 membres de simulation de CESM1-LMLE, on parvient avec notre approche à reconstruire une AMV significativement corrélée avec celle simulée pour 11 d'entre eux (Fig. 4.3). On a de plus montré qu'elle est significativement corrélée au seuil de 90% avec 22 enregistrements coralliens et sédimentaires de l'Atlantique Nord qui n'ont pas été utilisés dans la reconstruction. Afin d'obtenir une vision plus globale, nous avons étendu la méthode de RF à l'ensemble des points d'une grille de données de SST (chapitre 5). À l'aide de la métrique du CE, nous avons montré que les modèles statistiques produisent des reconstructions significativement robustes pour plus de 87% des points de grilles, nous permettant ainsi de disposer d'une reconstruction de SST annuelle de 1130 ans (850-1980) et de résolution $1^\circ \times 1^\circ$. De nombreux résultats obtenus provenant de ces reconstructions sont validés par des simulations de modèles physiques provenant d'études antérieures. Ces résultats peuvent ainsi prétendre éclaircir certaines questions faisant actuellement débat concernant la variabilité climatique interne, ses réponses aux forçages naturels, ainsi que le devenir de l'évolution du climat dans les années à venir.

En matière de forçages radiatifs naturels, nous avons pu montrer que les variations de la TSI possèdent un effet significatif sur la SST (chapitre 5, Fig. 5.6). Cependant, la diminution de température, à l'échelle globale comme locale, reste relativement faible. Ceci suggère que les amplitudes de variation de la TSI au cours du dernier millénaire étaient également relativement faibles, ce qui est cohérent avec ce que proposent Schurer et al. (2014). Après avoir fait le constat que certaines des reconstructions de TSI existantes diffèrent de près d'un ordre de magnitude (Schurer et al. 2014), ces derniers constatent que les amplitudes de variation d'enregistrements paléoclimatiques ne sont pas comparables à celles de sorties de modèles auxquels on impose de fortes amplitudes de variations de la TSI.

En matière d'éruptions volcaniques, nous avons pu montrer de nombreuses réponses significatives aux plus importantes d'entre elles. On montre par exemple dans le chapitre 3 une réponse de la NAO un an après celles-ci, comme cela a pu être suggéré par Shindell et al. (2004). Cette étude soulignait par ailleurs le besoin de reconstructions paléoclimatiques pour vérifier ses hypothèses, basées sur des significativités statistiques assez faibles au vu de la relativement modeste activité volcanique de la période de mesures instrumentales. Ceci montre parfaitement l'intérêt de notre reconstruction de NAO, qui, en moyenne, répond positivement aux puissantes éruptions. Notons que ce résultat est nouveau car c'est la première fois qu'est présentée une telle réponse de la NAO à partir d'une reconstruction millénaire. Ortega et al. (2015) avaient montré à partir de leur reconstruction de NAO que la réponse significative se trouve à 2 ans. Puisque les hivers se situent entre décembre

et février, ce que l'on considère ici comme réponse de la NAO un an après les très grosses éruptions est en réalité une réponse positive l'hiver boréal qui leur succède. Le mécanisme le plus souvent proposé pour une telle réponse est que l'absorption par les aérosols volcaniques d'une partie des radiations solaires ré-émises au sein la stratosphère sous forme de rayons infra-rouges va avoir tendance à réchauffer cette couche de l'atmosphère. On observe alors un accroissement du gradient de température méridien dans la stratosphère, accompagné d'une réduction du gradient de température méridien au sein de la troposphère dû à l'effet radiatif des aérosols volcaniques (Robock and Mao 1995). Ceci conduirait alors à une intensification du vortex polaire et un décalage du courant jet vers le nord lors de l'hiver boréal (Robock and Mao 1995), synonyme d'une phase de NAO positive. Cependant, Stenchikov et al. (2002) indiquent que l'effet des aérosols volcaniques sur les températures de la stratosphère est en général restreint aux basses latitudes, si bien que leur impact direct sur la circulation du vortex polaire est faible. Ils proposent des mécanismes plus indirects pouvant intensifier le vortex polaire, telle qu'une baisse des flux d'ondes planétaires due à la réduction du gradient méridien de la troposphère. Le chapitre 3 souligne que pour la reconstruction de NAO obtenue par RF, la réponse moyenne est significative pour les 11 plus grosses éruptions mais surtout qu'elle est particulièrement forte pour les deux plus importantes (Samalas en 1258 et Kuwae en 1458). Pour les 9 autres éruptions, la réponse positive de la NAO n'est pas systématique et peut également dépendre de l'état du système au moment de l'éruption ou de l'ampleur de cette éruption, comme cela a pu être suggéré par Bittner et al. (2016). Ces derniers auteurs montrent en effet qu'au sein d'un ensemble de membres de simulation de GCM, l'effet d'une éruption type Pinatubo (1991) sur la NAO n'est pas significatif, à l'inverse d'une éruption de type Tambora (1815). En particulier, Bittner et al. (2016) soulignent que l'intensification du vortex polaire n'implique pas forcément une NAO positive, notamment lorsque la perturbation stratosphérique induite par l'éruption est faible en comparaison de la forte variabilité interne de la NAO. Ceci indique donc que la réponse de la NAO aux éruptions volcaniques n'est pas systématique, et dépend notamment de la puissance de l'éruption en question (chapitre 3).

En terme de variabilité interne, de nombreux éléments tendent à montrer qu'il existe un lien négatif entre les variations d'ENSO et de la NAO (Cassou and Terray 2001, Brönnimann 2007, Graf and Zanchettin 2012, Zhang et al. 2019c). Cependant, une récente étude montre que de nombreux paramètres, tels que les phases de l'AMV et les forçages externes, jouent également sur les variations de chacun des modes, rendant ainsi leur relation non-stationnaire (Zhang et al. 2019b). Les événements *El Niño* sont par ailleurs proposés comme source de désintensification du vortex polaire (Garfinkel and Hartmann 2008). Ainsi, la réponse non-systématique d'une NAO positive pourrait éventuellement s'expliquer par le fait que l'on trouve également une réponse positive d'*El Niño* l'année ou l'hiver boréal suivant les grosses éruptions (chapitre 5, Fig. 5.5). Le fait que l'on trouve en moyenne des réponses significativement positives de la NAO et de la SST dans la région *El Niño* à la suite des plus puissantes éruptions peut sembler contre-intuitif au vu de l'anti-corrélation déduite des observations et des simulations de modèles (Cassou and Terray 2001, Brönnimann 2007, Graf and Zanchettin 2012, Zhang et al. 2019c). Cependant, cette réponse significative d'événements *El Niño* l'année suivant les puissantes éruptions avait déjà été mise en évidence dans des simulations de modèles antérieures, et s'explique par une réduction de gradient zonal de SST dans le Pacifique, et donc une

désintensification de la cellule de Walker et un ralentissement des remontées d'eaux froides au large de l'Amérique du Sud (Ohba et al. 2013, Khodri et al. 2017). La non-stationnarité du lien négatif qui existe entre la NAO et les événements *El Niño* a récemment été attribué aux fluctuations de l'AMV (Zhang et al. 2019b). Cependant, nous avons montré qu'en moyenne, les éruptions volcaniques stratosphériques ont un effet significatif sur l'occurrence d'événements *El Niño* mais aussi sur une NAO positive. Une étude des réponses individuelles aux différentes éruptions pourrait permettre de déceler si la réponse non-systématique de NAO positive est due à un événement *El Niño*, ou bien si la non-stationnarité de l'anti-corrélation qui existe entre les modes (Zhang et al. 2019b) peut en partie s'expliquer par le forçage volcanique.

Pour la SST de l'Atlantique Nord, on a montré dans le chapitre 4 que l'on pouvait discerner la variabilité interne de la variabilité forcée en deux séries distinctes. La composante de variabilité interne ne présente pas de réponse particulière aux éruptions puisqu'elle a été construite à l'aide d'un modèle statistique entraîné sur des données où l'on a préalablement retiré les effets des forçages. En matière de variabilité forcée, ces résultats sont cohérents avec les simulations de Mignot et al. (2011), qui suggèrent un refroidissement de 10 à 20 ans dans l'Atlantique selon les éruptions (chapitre 4 Fig. 4.7, chapitre 5, Fig. 5.5). Ces différents résultats significatifs de réponses climatiques aux éruptions, cohérents avec la physique des modèles, permettent ainsi d'identifier quelles périodes du dernier millénaire ont pu être largement forcées par les éruptions volcaniques. On trouve en effet que les quatre siècles les plus froids du millénaire en matière de SST globale sont les XIII^e, XV^e, XVII^e et XIX^e siècles, ce qui correspond aux siècles avec les activités volcaniques les plus intenses. Néanmoins, il reste encore beaucoup à étudier sur la réponse climatique à ces éruptions au cours du dernier millénaire. Un fait marquant qui montre le besoin d'autres travaux pour éclaircir les résultats obtenus est que la réponse à la plus puissante éruption (Samalas en 1258) ne semble pas avoir eu les effets *a priori* escomptés au vu de la moyenne d'ensemble des réponses de la SST globale. En effet, bien que celle-ci soit en moyenne significativement négative au XIII^e siècle, les signaux négatifs observés aux XV^e et XIX^e siècles, respectivement en réponse aux volcans Kuwae en 1458 et Tambora en 1815, semblent bien plus importants (chapitre 5, Fig. 5.4 et 5.5). Une piste intéressante pour résoudre cette énigme serait de s'intéresser aux saisonnalités et aux localisations des éruptions, ce qui n'a pas été fait ici. En effet, comme souligné par Stevenson et al. (2017), ces paramètres pourraient très largement influencer la réponse du système et sa temporalité. Cependant ces informations ne sont disponibles que pour un nombre restreint des éruptions identifiées et étudiées dans cette thèse (Sigl et al. 2015), ce qui encourage d'autant plus les recherches en glaciologie, volcanologie et en modélisation du climat pour disposer de plus d'informations à ce sujet.

Les résultats obtenus impliquent également qu'au-delà de l'effet des éruptions volcaniques et au vu du peu de forçages anthropiques et solaires, la variabilité climatique interne a joué un rôle crucial dans l'évolution climatique du dernier millénaire. En particulier, la reconstruction d'AMV que l'on propose dans le chapitre 4, insensible au forçage, a fait l'objet d'une importante transition au cours de la fin du XII^e siècle, l'entraînant vers une phase négative de plus de 400 ans, synonyme d'une entrée précoce dans le LIA dans les bassins Nord-Atlantique et Arctique. Cette transition est un résultat important puisqu'elle survient plus de 50 ans

avant la succession d'importantes éruptions volcaniques du XIII^e siècle (chapitre 4). Bien que cette transition semble prématurée, celle-ci coïncide cependant avec des reconstructions de la cryosphère Arctique, qui documentent d'anormales avancées de glaciers et de glace de mer au cours du XIII^e siècle (Kinnard et al. 2011, Jomelli et al. 2016). On distingue également cette transition sur la NASST comprenant les forçages, ce qui suggère bien que l'Atlantique est devenu anormalement froid au cours de cette période. Si l'on se réfère aux impacts observés de l'AMV, on peut donc émettre l'hypothèse qu'à ces époques, les températures estivales européennes et nord-américaines étaient anormalement basses (Sutton and Hodson 2005, Sutton and Dong 2012, O'Reilly et al. 2017).

Dans le chapitre 5, nous avons également pu apporter des éléments de compréhension quant au lien qui existe entre les bassins de l'Atlantique Nord et du Pacifique, ainsi que ses implications. On y montre en effet qu'à l'échelle du dernier millénaire, l'AMV a un effet négatif à 10 ans légèrement significatif ($0.05 < p < 0.1$) sur les variations de l'IPV (Fig. 5.12). Une étude plus détaillée de ce lien hors-phase entre les deux modes sur de plus petites échelles de temps (100 ans) a montré qu'il est particulièrement fort lorsque l'AMV est significativement dans une phase positive (Fig. 5.13). Ce résultat vient donc soutenir les hypothèses de McGregor et al. (2014) et de Chafik et al. (2016), qui proposent que la phase positive de l'AMV dans les années 1990 a pu jouer un rôle non-négligeable dans la mise en place du hiatus de la température globale observé entre 1998 et 2012, puisqu'il coïncide exactement avec une phase d'IPV négative (Kosaka and Xie 2013, England et al. 2014).

L'AMV reconstruite sur le dernier millénaire (chapitre 4) possède également des propriétés spectrales cohérentes (20-80 ans) avec ce qui a été proposé par Frankcombe et al. (2010), en impliquant uniquement des processus de variabilité indépendants des forçages. Ce spectre montre par ailleurs que l'AMV possède de larges amplitudes sur les observations récentes (60-90 ans, chapitre 4, Fig. 5.4), en comparaison de ces variations sur le dernier millénaire. Selon les méthodes de détection de point de bascule proposées par Lenton (2011), Lenton et al. (2012) et Boulton et al. (2014), ceci est un signal significatif de l'approche d'un tel événement de bascule de l'AMV. Puisque de nombreux éléments montrent que l'AMV est en partie l'expression en surface des variations de l'AMOC (Knight et al. 2006, Frankcombe et al. 2008, Frankcombe and Dijkstra 2009, Muir and Fedorov 2015, Sévellec and Huck 2015, Yan et al. 2018, Zhang et al. 2019a), ceci pourrait être, de manière plus large, le signal qu'un processus de ralentissement et de changement d'état de l'AMOC pourrait advenir à l'avenir. Ceci pourrait également être le signal d'un futur refroidissement anormal de l'Atlantique Nord suite à un ralentissement important de la circulation dans le SPG, comme proposé par Sgubin et al. (2017). Ce résultat semble se vérifier à l'aide de la deuxième PC de l'Atlantique Nord, obtenue depuis la grille de SST reconstruite et présentée dans le chapitre 5. Alors que cette série extraite de la grille connaît une empreinte spatiale sur la SST très similaire à celle de l'AMOC au sein des modèles (Knight et al. 2005, Zhang 2008, Persechini et al. 2013, chapitre 5, Fig. 5.10 et 5.11), on y distingue une forte tendance négative depuis environ les années 1940 (chapitre 5, Fig. 5.10 et 5.11). On voit même que le niveau d'intensité de l'AMOC a atteint un niveau sans précédent ces dernières décennies, depuis au moins l'année 850 (chapitre 5, Fig. 5.10). Ce résultat entre dans la lignée de nombreux travaux qui soutiennent un ralentissement actuel important et anormal de l'AMOC (Rahmstorf et al. 2015, Caesar et al. 2018, Smeed et al. 2018, Thornalley

et al. 2018). Ces études sont par ailleurs basées sur des méthodes aux approches relativement différentes, ce qui renforce la crédibilité de leur résultat commun. On peut citer : Smeed et al. (2018) à partir des observations directes, mais très courtes (14 ans), de l'AMOC (*cf.* RAPID); Rahmstorf et al. (2015) à partir d'une ancienne grille de reconstruction de températures qui s'étend jusqu'à l'an 850 (Mann et al. 2008); Caesar et al. (2018) à partir de l'empreinte de l'AMOC sur la SST observée depuis 1870; Thornalley et al. (2018) à partir d'un enregistrement paléoclimatique de circulation océanique qui remonte à l'an 400. À ce sujet, les simulations de modèles de circulation globale suggèrent que l'AMOC va *très probablement* ralentir de manière significative avant 2100 (Collins et al. 2019). Au vu de nos résultats et de ceux des études antérieures, ce ralentissement sans précédent aurait déjà largement commencé et ne pourrait en être qu'à ses débuts. Ainsi, le caractère inédit de ce phénomène accrédite l'idée selon laquelle il est, au moins en partie, d'origine anthropique. À l'avenir, cette tendance pourrait s'aggraver notamment parce que la fonte rapide de la banquise Arctique et de la calotte groenlandaise devrait entraîner un important apport d'eau douce. (Swingedouw et al. 2007, Bakker et al. 2016). Ce résultat a des implications importantes en matière de changements climatiques qui adviendront dans les prochaines décennies, tant les impacts des variations de l'AMOC sont nombreux (Buckley and Marsahll 2016). Si les hypothèses présentées ici venaient à se confirmer, ceci rendrait d'autant plus urgent le besoin de mettre en place des politiques de réduction des émissions de gaz à effet de serre et d'adaptation rationnelles et efficaces, et ce à l'échelle mondiale.

Après les données instrumentales et paléoclimatiques, les modèles et les reconstructions paléoclimatiques, l'assimilation de données de surface pourrait constituer une preuve supplémentaire pour vérifier ce ralentissement sans précédent de l'AMOC. Cependant, la réanalyse obtenue par une méthode de rappel appliquée à la grille de SST présentée dans le chapitre 5 au sein du modèle IPSL-CM5A2-LR n'a pas permis de vérifier ce résultat. Malgré le fait que les températures et paléotempératures disponibles semblent relativement bien reproduites (chapitre 6, Fig. 6.5 et 6.6), l'AMOC que l'on extrait de cette réanalyse est bien différente de celle extraite à partir de la grille de SST. Ceci amène à penser que l'AMOC n'est pas convenablement simulée lorsque la SST seule est assimilée, tel que proposé par Estella-Perez et al. (2020). Ceci a de fortes implications pour l'assimilation de données sur le dernier millénaire. Notamment, il semble déterminant de considérer l'assimilation d'une reconstruction de grille de SSS en plus de la SST, afin de convenablement forcer la convection aux hautes latitudes et les variations de l'AMOC sur la période.

7.2 Regard critique et réflexions sur les travaux présentés

Les travaux présentés sont des approches très statisticiennes de la paléoclimatologie et de l'étude de la variabilité climatique décennale. Cette section vient discuter de l'approche et de la méthodologie mobilisées pendant cette thèse, ce qui vient relativiser les résultats présentés et expliquer dans quelles mesures ils peuvent avoir de l'intérêt.

Les statistiques, précieux outils dans de nombreux domaines de la recherche, connaissent évidemment des limites. C'est notamment le cas lorsque les données recèlent une part relativement importante d'incertitudes, qu'elles soient quantifiées ou

non quantifiées. En particulier, entraîner un modèle de *machine learning* sur des données incertaines est susceptible d'établir des associations, et donc de produire des reconstructions qui peuvent s'éloigner significativement de la réalité. C'est pourquoi il est important, lorsque l'on produit des travaux tels que ceux présentés dans cette thèse, d'évoquer les différentes sources d'incertitudes et d'approximations qui leur sont associées, afin de faire preuve d'une nécessaire humilité scientifique. Pour les données paléoclimatiques, on peut par exemple évoquer les approximations statistiques dont elles sont issues, en particulier liées aux modèles d'âges ou aux fonctions de transferts permettant de passer de mesures géochimiques ou biochimiques à la reconstruction locale ou régionale d'une variable climatique. Il existe également des incertitudes associées aux outils et protocoles de mesures, qui, en tant qu'estimations, ne sont, par définition, pas la mesure exacte de la réalité mais une approximation de celle-ci. Max Planck disait dans une critique philosophique faite à la science empirique : "[u]ne expérience est une question que la science pose à la nature, et une mesure est l'enregistrement de la réponse de la nature" (Planck 1949). Bien que formulée il y a plus de 80 ans, cette phrase prend d'autant plus de sens de nos jours, à l'heure où l'on mesure une quantité d'informations considérable sur le monde qui nous entoure. Ceci souligne que toute science qui se base sur des mesures de la réalité, surtout quand elles sont indirectes comme en paléoclimatologie, se doit d'avoir un certain recul sur les résultats qu'elle obtient et sur les paradigmes qu'elle propose. La nature, complexe et continue dans l'espace-temps, ne peut être mesurée de manière exhaustive. En effet, il faudrait une capacité de mémoire infinie pour stocker l'ensemble des informations détenues dans n'importe quel système évoluant dans un espace continu tel que l'espace-temps. Ceci est toujours vrai, que ce soit pour une fraction de seconde ou plusieurs milliards d'années. Les données paléoclimatiques, très largement discrétisées dans l'espace-temps, sont calculées à partir de nombreuses approximations afin de rendre compte au mieux de la réalité. Évidemment, ces critiques sur les données paléoclimatiques n'enlèvent rien aux mérites des travaux réalisés par les scientifiques de terrain qui les ont produites. Ils ont permis d'accumuler de nombreuses connaissances sur l'histoire de l'évolution du climat et des environnements.

Il faut également être conscient du fait que les données instrumentales qui sont, dans cette thèse, implicitement considérées comme la réalité, possèdent également des incertitudes de mesures, bien que celles-ci soient moins importantes que pour les données paléoclimatiques et qu'elles ont tendance à diminuer à mesure que les techniques pour les recueillir s'améliorent. Notamment, il est légitime d'avoir un doute raisonnable sur les données qui remontent relativement loin dans le temps (fin du XIX^e et début du XX^e). Ce doute n'enlève rien à l'engagement admirable des personnes qui à l'époque ont fait ces précieuses mesures. Ce doute vient de la faible couverture spatio-temporelle de ces données ainsi que de l'imprécision relative des instruments de mesure de l'époque qui ont permis de les produire. Notons que pour disposer d'un nombre satisfaisant de données instrumentales, certains jeux de données tel que HadISST (Rayner et al. 2003), très largement utilisés dans cette thèse, sont, pour les périodes les plus anciennes, des interpolations linéaires. C'est-à-dire que les données de ces époques sont en partie des estimations statistiques appliquées aux données réellement mesurées à l'aide d'outils appropriés. Ceci rajoute une certaine couche d'incertitude pour les reconstructions présentées ici, puisque les données que l'on considère ici comme mesures du réel ne sont pas *stricto sensu* des mesures instrumentales dans leur intégralité.

Dans un cadre d'application de méthodes de *machine learning*, plus il y a d'incertitudes et d'approximations dans la collecte des données utilisées, plus l'on a de risques que le modèle statistique donne une interprétation biaisée de la réalité. En particulier, il faut noter que pour les reconstructions présentées ici, on a besoin que les données paléoclimatiques et les données instrumentales couvrent intégralement la même période. Cependant, les données paléoclimatiques s'accumulent depuis des décennies et ne peuvent en toute logique remonter au plus tard qu'à la date où elles ont été prélevées. Ainsi, plus l'on souhaite utiliser de données paléoclimatiques, moins l'on va pouvoir considérer de données instrumentales, et inversement. Ceci suggère donc qu'on doit en partie s'affranchir d'un nombre considérable de données instrumentales, et en particulier des plus récentes, qui sont pourtant celles possédant le moins d'incertitudes, qui de ce fait permettraient d'obtenir des modèles statistiques de reconstruction plus robustes.

Au-delà des incertitudes et biais évoqués précédemment, les modèles statistiques ne sont pas non plus parfaits et font l'objet d'approximations. En effet, même si l'on disposait de mesures qui ne possèdent aucune incertitude, les modèles statistiques peuvent quand même se tromper dans leur interprétation de la réalité. Ceci peut être dû à de nombreuses raisons, comme par exemple le fait qu'on ne peut pas disposer d'un ensemble exhaustif d'informations de la réalité. En d'autres termes, il est difficile d'imaginer qu'un modèle statistique décrive quelque chose qu'il n'a jamais vu sur les données sur lesquelles il est construit. Si l'on place ceci dans le contexte des présents travaux, ce constat a des implications très fortes. En effet, les variations climatiques sont bien plus influencées par l'activité humaine sur la période instrumentale que sur la période reconstruite. On peut alors très facilement imaginer que ces périodes sont significativement différentes et qu'on cherche à reconstruire des variations climatiques à l'aide de modèles statistiques qui n'en ont jamais vu de telles. Néanmoins, ceci peut en partie se compenser avec le problème de la période commune des enregistrements et des observations, nécessaire pour entraîner les modèles statistiques, qui a été évoquée au paragraphe précédent. En effet, puisque l'on a tendance à plutôt utiliser les données les plus anciennes, celles-ci sont moins influencées par le forçage anthropique, et donc sont plus enclines à posséder des variations similaires avec ce que l'on cherche à reconstruire. De plus, il faut noter que ces travaux reposent sur l'hypothèse forte, difficilement vérifiable, et physiquement peu cohérente, que les liens établis par les modèles statistiques entre les enregistrements paléoclimatiques et les observations sont stationnaires. Autrement dit, ceci repose sur l'hypothèse que ces liens n'ont pas changé entre aujourd'hui et il y a 300 ans, 500 ans ou 1000 ans par exemple.

Malgré ces précautions critiques, les assertions proposées au cours de cette thèse prennent des positions relativement tranchées au sein de controverses scientifiques actuelles. On peut par exemple citer les résultats concernant l'actuel ralentissement sans précédent de la circulation océanique qui devrait perdurer (chapitres 4 et 5), ce qui ne fait pas nécessairement l'unanimité. Par exemple, Lambert et al. (2016) montrent qu'à partir d'un "modèle en boîte" de circulation thermohaline de Stommel (Stommel 1961), que l'effet de l'apport d'eau douce seul ne peut influencer que dans une moindre mesure sur la stabilité de l'AMOC. Par ailleurs, un des résultats majeurs de cette thèse est que le LIA aurait commencé bien plus tôt (aux XII^e – XIII^e siècles) que ce qui est généralement proposé (aux XV^e – XVI^e siècles, Bradley and Jones 1993, Goosse et al. 2005, Mann et al. 2009). Ce résultat pourrait s'expliquer par

le fait qu'aujourd'hui on dispose de beaucoup plus de données paléoclimatiques, permettant de construire des modèles statistiques plus robustes qu'à l'époque où l'on suggérait que le début du LIA était aux alentours de 1400. En d'autres termes, à mesure que les connaissances scientifiques s'accumulent et que plus de données sont disponibles, les paradigmes évoluent. De la même manière, ce résultat pourrait très bien se voir réfuter dans quelques années ou décennies, au profit de nouvelles définitions du MCA et du LIA, lorsque de nouvelles données paléoclimatiques seront à disposition. Mais l'objet de la science n'est pas de produire des vérités définitives, mais plutôt de faire avancer les connaissances à partir des faits et des données disponibles à un moment donné, pour se rapprocher de la vérité. On peut prendre l'exemple de l'étude de Trouet et al. (2009), dans laquelle ils suggèrent que la NAO fut positive pendant tout le début du millénaire jusqu'en 1400, et qui en concluent que cette variation est la cause de la transition entre le MCA et le LIA. De nombreuses études ont ensuite proposé des arguments solides pour faire tomber cette hypothèse (Lehner et al. 2012, Ortega et al. 2015, Cook et al. 2019, Michel et al. 2020). Ainsi, bien que les preuves s'accumulent pour contredire les hypothèses de Trouet et al. (2009), il n'en demeure pas moins que les questions soulevées par cette étude ont, *in fine*, permis de parfaire notre compréhension du climat du dernier millénaire. Pour autant, il ne s'agit pas ici de dire que les hypothèses présentées dans cette thèse ont vocation à nécessairement être réfutées. Celles-ci se basent sur plus de données et sur des modèles statistiques qui semblent plus robustes que précédemment. Ces hypothèses découlent en effet de significativités statistiques consciencieusement calculées et peuvent donc être considérées, au vu de l'information et des connaissances disponibles à l'heure actuelle, comme fiables. De plus, le résultat d'une entrée précoce dans le LIA ne vient pas entièrement bousculer les résultats des précédentes études. En particulier, on trouve que les températures les plus froides à l'échelle globale sont atteintes entre les XVII^e et XIX^e siècles (Fig. 5.4), ce qui a déjà été identifié précédemment, en réponse notamment à l'important forçage radiatif externe sur cette période (Goosse et al. 2005, Brönnimann et al. 2019).

Il faut bien souligner que les critiques soulevées dans cette section sont également valables pour les études antérieures et postérieures à cette thèse, qui se basent ou se baseront sur le même type d'approche. On peut également relativiser ces critiques par le fait que la science appliquée au monde réel ne peut jamais être parfaite, ce qui suppose que tout scientifique, statisticien ou non, doit toujours prendre du recul sur les résultats qu'il trouve.

En conclusion de cette section, la paléoclimatologie est un domaine scientifique jeune et très vaste que l'on étudie selon de nombreuses approches aux techniques très différentes, qui possèdent chacune leurs qualités et leurs défauts. C'est pourquoi chacun des contributeurs se doit d'identifier et de renseigner les limites de son approche, tout en posant des hypothèses claires, réfutables ou vérifiables, afin de faire avancer les connaissances dans ce domaine, et pour espérer à terme approcher au mieux la vérité. Finalement, pour illustrer ce qui vient d'être dit, on peut se référer à Victor Hugo, selon lequel : "[l]a science est l'asymptote de la vérité, elle approche sans cesse et ne touche jamais" (Hugo 1864).

7.3 Perspectives

Cette thèse ouvre différentes perspectives, qui pourraient faire l'objet de plusieurs études. En particulier, on peut esquisser des pistes intéressantes pour pallier à certains problèmes évoqués dans la section précédente.

En ce qui concerne les données paléoclimatiques, une manière de remédier aux incertitudes temporelles dans le cadre d'application de méthodes de *machine learning* serait de préalablement établir le pas de temps pour lequel la corrélation croisée entre un enregistrement et la variable climatique que l'on veut reconstruire est maximale (en valeur absolue). En effet, pour les reconstructions produites dans les chapitres 3 à 5, les modèles statistiques les plus robustes sont en général ceux pour lesquels on pré-sélectionne les enregistrements par test de corrélation. Ainsi, il serait intéressant d'établir une sélection d'enregistrements en maximisant (en valeur absolue) la corrélation croisée avec l'indice reconstruit plutôt que simplement s'intéresser à la corrélation en phase. Ceci permettrait de s'affranchir en partie de potentielles incertitudes temporelles d'enregistrements sur la période historique, mais aussi de détecter de réelles téléconnexions qui ne sont pas en phase.

En terme d'assimilation de données, certaines méthodes, telle que la méthode de KF, permettent de prendre en compte les incertitudes temporelles des enregistrements pour produire les réanalyses (Hakim et al. 2016). À ce sujet, les paléotempératures (PAGES 2k Consortium 2017) sont moins bien reproduites dans la réanalyse obtenue par méthode de rappel (chapitre 6) que dans celle obtenue par la méthode de KF (Hakim et al. 2016, chapitre 6). Néanmoins, le fait que les réanalyses diffèrent au niveau du modèle utilisé, de la variable climatique assimilée, et des enregistrements paléoclimatiques utilisés pour produire la grille de reconstruction assimilée, ne permet pas de rigoureusement les comparer (chapitre 6). Il y a donc ici un double intérêt à produire à l'avenir une réanalyse du dernier millénaire au sein du modèle IPSL-CM5A2-LR à l'aide d'une méthode de KF : s'affranchir des incertitudes temporelles de la grille de reconstruction sous-jacente, et comparer rigoureusement cette méthode à une méthode dont la complexité computationnelle est moins importante (*e.g.* la méthode de rappel).

Dans le chapitre 3, il est souligné que la méthodologie de reconstruction d'indices climatiques présentée n'a été employée que dans le cas d'une période de reconstruction fixe, amenant à proposer une reconstruction de NAO basée sur un nombre relativement faible d'enregistrements paléoclimatiques (*i.e.*, ceux dont la couverture spatiale est la plus importante). Cela permet d'envisager qu'appliquer la méthodologie de reconstruction présentée dans le chapitre 3, dans un cadre de reconstruction sur fenêtres imbriquées tel que dans les chapitres 4 et 5, pourrait produire une reconstruction de NAO encore plus robuste, puisque basée sur un nombre beaucoup plus important d'enregistrements paléoclimatiques adéquats (chapitres 3 à 5).

Comme soulevé dans la section 7.1, il serait intéressant de voir les réponses simultanées de la reconstruction de NAO (chapitre 3) et de la grille de reconstruction de SST dans la région *El Niño* (chapitre 5), pour chacune des éruptions volcaniques étudiée dans les chapitres 3 à 5 plutôt que sur une moyenne d'ensemble. En effet, sur les moyennes d'ensemble, on observe une réponse positive de chacun de ces phénomènes (chapitres 3 Fig. 3.10, chapitre 5 Fig. 5.5), alors qu'ils sont connus

pour être anti-corrélés (Cassou and Terray 2001, Brönnimann 2007, Graf and Zanchettin 2012, Zhang et al. 2019c). Une étude plus approfondie de leurs réponses aux éruptions étudiées dans cette thèse pourrait permettre de savoir dans quelles mesures celles-ci peuvent jouer sur la non-stationnarité du lien qui existe entre ces deux modes (*cf.* Zhang et al. 2019b).

La grille de reconstruction de SST présentée dans le chapitre 5 a permis de trouver des résultats intéressants en terme de dynamiques et de réponses aux forçages, qui sont notamment cohérents avec des simulations de modèles (chapitre 5, chapitre 7.1). Cependant, celle-ci pourrait être améliorée, en employant une approche de reconstruction et de validation plus rigoureuse :

1. La procédure de reconstruction par RF n’a pas été vérifiée en *pseudo-proxy* (Neukom et al. 2018), tel que ceci a pu être fait dans le chapitre 4 pour la reconstruction de l’AMV.
2. Il a été décidé de manière arbitraire de directement utiliser la méthode RF, aux dépens des méthodes lui ayant été confrontées dans les chapitres 3 et 4 pour les reconstructions des variations respectives de la NAO et de l’AMV (*e.g.* PCR, PLS et Enet).

Comme indiqué dans le chapitre 5, plus de 5 millions de modèles de RF sont générés pour optimiser la reconstruction de chaque point de grille du jeu de données HadISST (Rayner et al. 2003). En conséquence, le temps d’exécution du code qui produit cette reconstruction est d’une dizaine de jours lorsque celui-ci est parallélisé sur 48 coeurs. Pour répondre au problème 1 soulevé ci-dessus, il faudrait reproduire la reconstruction en *pseudo-proxy* sur un nombre satisfaisant de membres de simulation (au moins une dizaine). Pour répondre au problème 2, on pourrait envisager d’appliquer une comparaison des reconstructions de grille obtenues pour différentes configurations, tel que ceci a été appliqué pour des indices de modes dans les chapitres 3 et 4, dont les reconstructions sont très largement moins coûteuses en terme de ressources de calcul. De ce fait, si l’on voulait répondre aux deux problèmes exposés ci-dessus, ce qui permettrait de faire gagner en robustesse les résultats obtenus, il faudrait envisager l’utilisation de 5 à 10 fois plus de coeurs pour produire les calculs requis dans un temps acceptable.

L’indice AMOC que nous proposons dans le chapitre 5 possède une empreinte spatiale sur la SST qui semble cohérente avec ce que suggèrent des simulations de modèles, notamment en terme de gradient inter-hémisphérique de SST induit par les variations de l’AMOC (Knight et al. 2005, Zhang 2008, Persechini et al. 2013, Fig. 5.10d et 5.11b). Cependant, cette reconstruction n’a pas été comparée à des enregistrements paléoclimatiques de circulation océanique profonde (section 1.1.4) indépendants (non utilisés dans la reconstruction), ce qui constitue un élément de clé pour valider les assertions qui en ont été déduites (*e.g.* un ralentissement sans précédent sur l’historique). Il existe plusieurs enregistrements de circulation en Atlantique Nord, dont une collection intéressante est répertoriée dans l’article de synthèse de Moffa-Sánchez et al. (2019). Des comparaisons préliminaires entre certains de ces enregistrements avec l’indice AMOC proposé dans le chapitre 5 donnent de premiers résultats prometteurs. Une étude plus exhaustive de ces enregistrements pour valider la cohérence de l’indice AMOC proposé dans le chapitre 5 devrait faire l’objet d’une nouvelle étude à la suite de cette thèse.

Finalement, il est clair qu'un approfondissement des travaux présentés dans le chapitre 6 pourrait produire des réanalyses millénaires plus intéressantes que celle qui y est actuellement proposée. Comme expliqué au début de la section, il y a tout d'abord de nombreux intérêts à envisager l'application de méthodes d'assimilation plus robustes, telle que la méthode de KF. Au-delà de la méthode, nous avons également montré dans le chapitre 6, que l'AMOC de la réanalyse obtenue par méthode de rappel a une corrélation non significative avec l'indice AMOC que nous avons estimé avec des méthodes statistiques (chapitre 5). Sous l'hypothèse que l'indice proposé dans le chapitre 5 soit une bonne estimation des variations de l'AMOC (cf. paragraphe précédent), ceci suggérerait qu'assimiler la SST seule au sein d'un modèle de circulation ne semble pas suffisant pour convenablement forcer la convection aux hautes latitudes, et donc les variations de l'AMOC (chapitre 6, sections 7.1). Une piste intéressante serait donc d'envisager une assimilation simultanée de la SSS et la SST sur le dernier millénaire. Pour ce faire, il faudrait donc produire une estimation des variations dans l'espace et dans le temps de la SSS sur le dernier millénaire. À cet égard, Estella-Perez et al. (2020) ont récemment montré que pour la méthode de rappel appliquée au modèle IPSL-CM5A2-LR, l'assimilation de la SSS dans un cadre de réanalyse en modèle parfait (*i.e.* en assimilant les données provenant d'une simulation antérieure) reproduit bien mieux l'AMOC de la simulation sous-jacente par rapport au cas où seule la SST est assimilée. Les résultats de Estella-Perez et al. (2020) sont d'autant plus intéressants qu'ils sont obtenus en assimilant uniquement la SSS sur l'Atlantique Nord, et ce, sur une grille dont la résolution spatiale est relativement basse. Ainsi, envisager la reconstruction d'une grille de SSS telle que celle qu'ils proposent semble un objectif atteignable. On pourrait en première approche songer à appliquer le même type de reconstruction que la SST, par exemple en appliquant des méthodes de *machine learning* à des enregistrements paléoclimatiques appropriés. Cependant, bien qu'il existe quelques enregistrements paléoclimatiques de SSS en Atlantique Nord qui couvrent une partie du dernier millénaire (Moffa-Sánchez et al. 2019), leur nombre est relativement faible en comparaison du nombre d'enregistrements de paléo-températures existants (PAGES 2k Consortium 2017), si bien que l'application de méthodes de *machine learning* semble peu envisageable. L'application de ces méthodes sur des données *in situ* uniquement pourrait de surcroît s'avérer particulièrement compliqué car les observations de salinité ont une couverture temporelle bien moins importante que la SST, les plus anciennes remontant aux début des années 1990. Une piste intéressante serait toutefois d'établir au sein de la grille utilisée par Estella-Perez et al. (2020) des liens de covariations entre la SSS et la SST à partir des observations, et d'extrapoler ces liens à partir la grille de reconstruction présentée dans le chapitre 5 pour produire une estimation des variations millénaires de la SSS. Une telle étude pourrait s'avérer d'un grand intérêt pour obtenir une réanalyse du dernier millénaire dans laquelle on peut supposer que l'AMOC pourrait être mieux reconstruite.

Bibliographie

- Abram, N. J., McGregor, H. V., Tierney, J. E., Evans, M. N., McKay, N. P., and Kaufman, D. S. (2016). Early onset of industrial era-warming across the oceans and continents. *Nature*, 536 :411–418.
- Alexandru, A., de Elia, R., Laprise, R., Separovic, L., and Biner, S. (2009). Sensitivity Study of Regional Climate Model Simulations to Large-Scale Nudging Parameters. *Monthly Weath. Rev.*, 137 :1666–1686.
- Andersen, K., Ditlevsen, P., Rasmussen, S., Clausen, H., Vinther, B., Johnsen, S., and Steffensen, J. (2006). Retrieving a common accumulation record from Greenland ice cores for the past 1800 years. *Journal of geophysical research*, 111 :D15106.
- Andersen, K. K., Bigler, M., Buchardt, S. L., Clausen, H. B., Dahl-Jensen, D., Davies, S. M., Fischer, H., Goto-Azuma, K., Hansson, M. E., Heinemeier, J., Johnsen, S. J., Larsen, L. B., Mischeler, R., Olsen, G. J., Rasmussen, S. O., Röthlisberger, R., Ruth, U., Seierstad, I. K., Siggaard-Andersen, M.-L., Steffensen, J. P., Svensson, A. M., and Vinther, B. M. (2007). Greenland Ice Core Chronology 2005 (GICC05) and 20 year means of oxygen isotope data from ice core NGRIP.
- Apaéstegui, J., Cruz, F. W., Sifeddine, A., Vuille, M., Espinoza, J. C., Guyot, J. L., Khodri, M., Strikis, N., Santos, R. V., Cheng, H., Edwards, L., Carvalho, E., and Santini, W. (2014). Hydroclimate variability of the northwestern Amazon Basin near the Andean foothills of Peru related to the South American Monsoon System during the last 1600 years. *Clim. Past*, 10 :1967–1981.
- Ayache, M., Swingedouw, D., Mary, Y., Eynaud, F., and Colin, C. (2018). Multi-centennial variability of the AMOC over the Holocene : A new reconstruction based on multiple proxy-derived SST records. *Glob. Plan. Change*, 170 :172–189.
- Azuz-Adeath, I., González-Campos, C., and Cuevas-Corona, A. (2019). Predicting the Temporal Structure of the Atlantic Multidecadal Oscillation (AMO) for Agriculture Management in Mexico's Coastal Zone. *Journ. Coast. Res.*, 35 :210–226.
- Ba, J., Keenlyside, S., N., Latif, M., Park, W., Ding, H., Lohmann, K., Mignot, J., Menar, M., Ottera, O. H., Wouters, B., Salas y Melia, D., Oka, A., Belucci, A., and Volodin, E. (2014). A multi-model comparison of Atlantic multidecadal variability. *Clim. Dyn.*, 43 :2333–2348.
- Bakker, P., Schmittner, A., Lenaerts, J. T. M., Abe-Ouchi, A., Bi, D., van der Broeke, M. R., Chan, W.-L., Hu, A., Beadling, R. L., Marsland, S. J., Mernild, S. H., Saenko, O. A., Swingedouw, D., Sullivan, A., and Yin, J. (2016). Fate of the Atlantic Meridional Overturning Circulation : Strong decline under continued warming and Greenland melting. *Geophys. Res. Lett.*, 43 :12252–12260.
- Biraben, J.-N. (1979). Essai sur l'évolution du nombre des hommes. *Population*, 34 :13–25.

- Bittner, M., Timmreck, C., Schmidt, H., Toohey, M., and Krüger, K. (2016). The impact of wave-mean flow interaction on the northern hemisphere polar vortex after tropical volcanic eruptions : NH response to volcanic eruptions. *Journ. Geophys. Res. Atm.*, 121 :5281–5297.
- Bjerknes, J. (1969). Atmospheric teleconnections from the equatorial Pacific. *Monthly Weather Review*, 97 :163–172.
- Björklund, J. A., Gunnarson, B. E., Seftigen, K., Esper, J., and Linderholm, H. W. (2014). Blue intensity and density from northern Fennoscandian tree rings, exploring the potential to improve summer temperature reconstructions with early-wood information. *Clim. Past.*, 10 :877–885.
- Boé, J. and Habets, F. (2014). Multi-decadal river flow variations in France. *Hydrol. Earth Syst. Sci.*, 18 :691–708.
- Bojariu, R. (2003). Predictability and numerical modelling of the North Atlantic Oscillation. *Earth. Sci. Rev.*, 63 :145–168.
- Booth, B. B. B., Dunstone, N. J., Halloran, P. R., Andrews, T., and Bellouin, N. (2012). Aerosols implicated as a prime driver of twentieth-century North Atlantic climate variability. *Nature*, 484 :228–232.
- Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P., Kerminen, V.-M., Kondo, Y., Liao, H., Lohmann, U., Rasch, P., Satheesh, S. K., Sherwood, S., Stevens, B., and Zhang, X. Y. (2013). Clouds and Aerosols. In Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., editors, *Climate Change 2013 - The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, page 571. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, Intergovernmental Panel on Climate Change edition.
- Boulton, C. A., Allison, L. C., and Lenton, T. M. (2014). Early warning signals of Atlantic Meridional Overturning Circulation collapse in a fully coupled climate model. *Nat. Com.*, 5 :5752.
- Bradley, R. S. and Jones, P. D. (1993). 'Little Ice Age' summer temperature variations : their nature and relevance to recent global warming trends. *The Holocene*, 3 :367–376.
- Breiman, L. (2001). Random forests. *Mach. Learn.*, 45 :5–32.
- Breiman, L., Friedman, J. H., Olshen, R., and Stone, C. J. (1984). Classification and Regression Trees. *Monterey, CA :Wadsworth and Brooks*.
- Bretherton, C. S., Widmann, M., Dymnikov, V. P., Wallace, J. M., and Bladé, I. (1999). The effective number of spatial degrees of freedom on time-varying field. *J. Clim.*, 12 :1990–2009.
- Broecker, W. S. (1999). What If the Conveyor Were to Shut Down? Reflections on a Possible Outcome of the Great Global Experiment. *GSA Today*, 9 :1–7.
- Brönnimann, S. (2007). Impact of El Niño-Southern Oscillation on European climate : ENSO impact on europe. *Rev. Geophys.*, 45.

- Brönnimann, S. and Krämer, D. (2016). *Tambora and the "year without a summer" of 1816 : a perspective on earth and human systems science*. Geographica Bernensia, Geographisches Institut der Universität Bern, Bern. OCLC : 968739143.
- Browne, M. W. (2000). Cross-Validation Methods. *Astronomy, Astrophysics*, 44 :108–132.
- Brönnimann, S., Franke, J., Nussbaumer, S. U., Zumbühl, H., Steiner, D., Trachsel, M., Hegerl, G. C., Shurer, A., Worni, M., Malik, A., Flückiger, J., and Raible, C. C. (2019). Last phase of the Little Ice Age forced by volcanic eruptions. *Nat. Geosci.*, 12 :650–656.
- Bucci, A. and Raurich, X. (2017). Population and Economic Growth Under Different Growth Engines. *Germ. Eco. Rev.*, 18 :182–211.
- Buckley, M. W. and Marsahll, J. (2016). Observations, inferences, and mechanisms of the Atlantic Meridional Overturning Circulation : A review. *Rev. Geophys.*, 54 :5–63.
- Buffon, V. (1756). *Histoire naturelle*. volume 6, page 60. Paris, l'impression royale.
- Bunn, A. G., Graumlich, L. J., and Urban, D. L. (2005). Trends in twentieth-century tree growth at high elevations in the Sierra Nevada and White Mountains, USA. *The Holocene*, 15 :481–488.
- Büntgen, U., Franck, D. C., Nievergelt, D., and Esper, J. (2006). Summer Temperature Variations in the European Alps, a.d. 755-2004. *J. Clim.*, 19 :5606–5623.
- Caesar, L., Rahmstorf, S., Robinson, A., Feulner, G., and Saba, V. (2018). Observed fingerprint of a weakening Atlantic Ocean overturning circulation. *Nature*, 556 :191–195.
- Cahill, N., Kemp, A. C., Horton, B. P., and Parnell, A. C. (2016). A Bayesian hierarchical model for reconstructing relative sea level : from raw data to rates of change. *Climate of the Past*, 12 :525–542.
- Caian, M., Koenigk, T., Döscher, R., and Devasthale, A. (2018). An interannual link between Arctic sea-ice cover and the North Atlantic Oscillation. *Clim. Dyn.*, 50 :423–441.
- Caley, T. and Roche, D. M. (2013). $\delta^{18}\text{O}$ water isotope in the iLOVECLIM model (version 1.0) – Part 3 : A palaeo-perspective based on present-day data-model comparison for oxygen stable isotopes in carbonates. *Geosci. Model Dev.*, 6 :1505–1516.
- Carré, M., Azzoug, M., Zaharias, P., Camara, A., Cheddadi, R., Chevalier, M., Fiorillo, D., Gaye, A. T., Janicot, S., Khodri, M., Lazar, A., Lazareth, C. E., Mignot, J., Mitma García, N., Patris, N., Perrot, O., and Wade, M. (2019). Modern drought conditions in western sahel unprecedented in the past 1600 years. *Clim. Dyn.*, 52 :1949 – 1964.
- Casado, M., Ortega, P., Masson-delmotte, V., Risi, C., Swingedouw, D., Daux, V., Genty, D., Maignan, F., Solomina, O., Vinther, B., Viovy, N., and Yiou, P. (2013). Impact of precipitation intermittency on NAO-temperature signals in proxy records. *Climate of the Past*, 9 :871–886.

- Cassou, C., Kushnir, Y., Hawkins, E., Pirani, A., Kucharski, F., I.-S., K., and Caltabiano, N. (2018). Decadal climate variability and predictability : challenges and opportunities. *BAMS*, 99 :479–490.
- Cassou, C. and Terray, L. (2001). Dual influence of Atlantic and Pacific SST anomalies on the North Atlantic/Europe Winter Climate. *Geophys. Res. Lett.*, 28 :3195–3198.
- Cassou, C., Terray, L., Hurrell, J. W., and Deser, C. (2004). North Atlantic Winter Climate Regimes : Spatial Asymmetry, Stationarity with Time, and Oceanic Forcing. *J. Clim.*, 17 :14.
- Cattiaux, J., Vautard, R., Cassou, C., Yiou, P., Masson-Delmotte, V., and Codron, F. (2010). Winter 2010 in Europe : A cold extreme in a warming climate : Cold winter 2010 in europe. *Geophys. Res. Lett.*, 37 :044613.
- Chafik, K., Häkkinen, S., England, M. H., Carton, J. A., Nigam, S., Ruiz-Barradas, A., Hannachi, A., and Miller, L. (2016). Global linkages originating from decadal ocean variability in the subpolar North Atlantic. *Geophys. Res. Lett.*, 43 :10909–10919.
- Chen, X. and Ishwaran, H. (2012). Random forests for genomic data analysis. *Genomics*, 99 :323–329.
- Christiansen, J. H., Krishna Kumar, K., Aldrian, E., An, S. I., Cavalcanti, I. F. A., de Castro, M., Dong, W., Goswami, P., Hall, A., Kanyanga, J. K., Kitoh, H., Kosin, J., Lau, N.-C., Renwick, J., Stephenson, D. B., Xie, S.-P., and Zhou, T. (2013). Climate Phenomena and their Relevance for Future Regional Climate Change. In Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., editors, *Climate Change 2013 - The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, pages 1217–1308. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, Intergovernmental Panel on Climate Change edition.
- Clement, A., Bellomo, K., Murphy, L. N., Cane, M. A., Mauritsen, T., Rädcl, G., and Stevens, B. (2015). The Atlantic Multidecadal Oscillation without a role for ocean circulation. *Science*, 350 :320–324.
- Collins, M., Sutherland, M., Bouwer, L., Cheong, S.-M., Frölicher, T., Jacot Des Combes, H., Koll Roxy, M., Losada, I., McInnes, K., Ratter, B., Rivera-Arriaga, E., Susanto, R. D., Swingedouw, D., Tibig, L. [Pörtner, H.-O., Roberts, D. C., Masson-Delmotte, V., Zhai, P., Tignor, M., Poloczanska, E., Mintenbeck, K., Alegria, A., Nicolai, M., Okem, A., Petzold, J., Rama, B., and Weyer, N. M. e. (2019). Extremes, Abrupt Changes and Managing Risk. In : IPCC Special Report on the Ocean and Cryosphere in a Changing Climate. *in press*.
- Cook, E. R., D'Arrigo, R. D., and Mann, M. E. (2002). A Well-Verified, Multiproxy Reconstruction of the Winter North Atlantic Oscillation Index since A.D. 1400. *J. Clim.*, 15 :1754–1764.
- Cook, E. R., Kushnir, Y., Smerdon, J. E., Williams, A. P., Anchukaitis, K. J., and Wahl, E. R. (2019). A Euro-Mediterranean tree-ring reconstruction of the winter NAO index since 910 C.E. *Clim. Dyn.*, 53 :1567–1580.

- Cook, J., Oreskes, N., Doran, P. T., Anderegg, W. R. L., Verheggen, B., Maibach, E. W., Carlton, J. S., Lewandowsky, S., Skuce, A. G., Green, S. A., Nuccitelli, D., Jacobs, P., Richardson, M., Winkler, B., Painting, R., and Rice, K. (2016). Consensus on consensus : a synthesis of consensus estimates on human-caused global warming. *Environ. Res. Lett.*, 11 :048002.
- Crowley, T. J., Zielinski, G., Vinther, B., Udisti, R., Kreutz, K., Cole-Dai, J., and Castellano, E. (2008). Volcanism and the Little Ice Age. *PAGES news*, 16 :22–23.
- Cuffey, K. M., Clow, G. D., Alley, R. B., Stuiver, M., Waddington, E. D., and Saltus, R. W. (1995). Large Arctic temperature change at the Wisconsin-Holocene glacial transition. *Science*, 270 :455–458.
- Cunningham, S. A., Kanzow, T., Rayner, D., Baringer, M. O., Johns, W. E., Maotzke, J., Longworth, H. R., Grant, E. M., Hirshi, J.-M., Beal, L. M., Meinen, C. S., and Bryden, H. L. (2007). Temporal variability of the Atlantic Meridional Overturning Circulation at 26.5°N. *Science*, 317 :935–938.
- Czaja, A. and Frankignoul, C. (2002). Observed Impact of Atlantic SST Anomalies on the North Atlantic Oscillation. *J. Climate*, 15 :18.
- Defrance, D., Ramstein, G., Charbit, S., Vrac, M., Famien, A. M., Sultan, B., Swingedouw, D., Dumas, C., Gemenne, F., Alvarez-Solas, J., and Vanderlinden, J.-P. (2017). Consequences of rapid ice sheet melting on the Sahelian population vulnerability. *PNAS*, 114 :6533–6538.
- Delworth, T. L. and Mann, M. E. (2000). Observed and simulated multidecadal variability in the Northern Hemisphere. *Clim. Dyn.*, 16 :661–676.
- Delworth, T. L., Zeng, F., Vecchi, G. A., Yang, X., Zhang, L., and Zhang, R. (2016). The North Atlantic Oscillation as a driver of rapid climate change in the Northern Hemisphere. *Nat. Geosci.*, 9 :509–512.
- Deser, C., Lehner, F., Rodgers, K. B., Ault, T. L., D. T., DiNezio, P. N., Fyfe, J. C., E., H. D., Kay, J. E., Knutti, R., Lovenduski, N. S., Marotzke, J., McKinnon, K. A., Minobe, S., Randerson, J., Screen, J. A., Simpson, I. R., and Ting, M. (2020). Insights from Earth system model initial condition large ensembles and future prospects. *Nat. Clim. Change*, 10 :277–286.
- Deser, C., Phillips, A., Bourdette, V., and Teng, H. (2012). Uncertainty in climate change projection : the role of internal variability. *Clim. Dyn.*, 38 :527–546.
- Dickson, R., Lazier, J., Meincke, J., Rhines, P., and Swift, J. (1996). Long-term coordinated changes in the convective activity of the North Atlantic. *Progress in Oceanography*, 38 :241–295.
- Drinkwater, K. F., Belgrano, A., Borja, A., Conversi, A., Edwards, M., Greene, C. H., Ottersen, A., Pershing, J., and Walker, H. A. (2003). The North Atlantic Oscillation : Climate significance and meteorological impacts. In Hurrell, J. W., Kushnir, Y., Ottersen, G., and Visbeck, M., editors, *The response of marine ecosystems to climate variability with the North Atlantic Oscillation*, chapter 10. American Geophysical Union.
- Dubinkina, S. and Goosse, H. (2013). An assessment of particle filtering methods and nudging for climate state reconstructions. *Clim. Past*, 9 :1141–1152.

- Dufresne, J.-L., Foujols, M.-A., Denvil, S., Caubel, A., Marti, O., Aumont, O., Balkanski, Y., Bekki, S., Bellenger, H., Benshila, R., Bony, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Cheruy, F., Codron, F., Cozic, A., Cugnet, D., de Noblet, N., Duvel, J.-P., Ethé, C., Fairhead, L., Fichefet, T., Flavoni, S., Friedlingstein, P., Grandpeix, J.-Y., Guez, L., Guilyardi, E., Hauglustaine, D., Hourdin, F., Idelkadi, A., Ghattas, J., Joussaume, S., Kageyama, M., Krinner, G., Labetoulle, S., Lahellec, A., Lefebvre, M.-P., Lefevre, F., Levy, C., Li, Z. X., Lloyd, J., Lott, F., Madec, G., Mancip, M., Marchand, M., Masson, S., Meurdesoif, Y., Mignot, J., Musat, I., Parouty, S., Polcher, J., Rio, C., Schulz, M., Swingedouw, D., Szopa, S., Talandier, C., Terray, P., Viovy, N., and Vuichard, N. (2013). Climate change projections using the IPSL-CM5 Earth System Model : from CMIP3 to CMIP5. *Clim. Dyn.*, 40 :2123–2165.
- Dunstone, N., Smith, D., Scaife, A., Hermanson, L., Eade, R., Robinson, N., Andrews, M., and Knight, J. (2016). Skilful predictions of the winter north atlantic oscillation one year ahead. *Nat. Geosci.*, 9 :809–814.
- Eden, C. and Jung, T. (2001). North Atlantic Interdecadal Variability : Oceanic Response to the North Atlantic Oscillation (1865–1997). *J. Climate*, 14 :16.
- Edwards, M., Beaugrand, G., Helaouët, P., Alheit, J., and Coombs, S. (2013). Marine ecosystems response to the Atlantic Multidecadal Oscillation. *PLOS ONE*, 8 :e57212.
- Eliassen, A. (1951). Slow thermally or frictionally controlled meridional circulation in a circular vortex. *Astr. Norv.*, 5 :19–59.
- Ely, C. R., Brayshaw, D. J., Methven, J., Cox, J., and Pearce, O. (2013). Implications of the North Atlantic Oscillation for a UK–Norway Renewable power system. *Energy Policy*, 62 :1420–1427.
- Emanuel, K. (2005). Increasing destructiveness of tropical cyclones over the past 30 years. *Nature*, 436 :686–688.
- Enfield, D. and Cid-Serrano, L. (2010). Secular and multidecadal warmings in the North Atlantic and their relationships with major hurricane activity. *Int. Journ. Clim.*, 30 :174–184.
- Enfield, D., Mestas-Nuñez, A., and Trimble, P. (2001). The Atlantic multidecadal oscillation and its relation to rainfall and river flows in the continent U.S. *Geophys. Res. Lett.*, 28 :2077–2080.
- England, M. F., McGregor, S., Spence, P., Meehl, G. A., Timmermann, A., Cai, W., Gupta, A. S., McPhaden, M. J., Purish, A., and Santoso, A. (2014). Recent intensification of wind-driven circulation in the Pacific and the ongoing warming hiatus. *Nat. Clim. Change*, 4 :222–227.
- Escudier, R., Mignot, J., and Swingedouw, D. (2013). A 20-year coupled ocean-sea ice-atmosphere variability mode in the North Atlantic in an AOGCM. *Clim. Dyn.*, 40 :619–636.
- Esper, J., Büntgen, U., Frank, D., Verstege, A., Nievergelt, D., and Liebhold, A. (2006). 1200 years of regular outbreaks in alpine insects. *Proc. Biol. Sci.*, 274 (1610) :671–679.

- Esper, J., Frank, D., Büntgen, U., Verstege, A., Luterbacher, J., and Xoplaki, E. (2007). Long-term drought severity variations in Morocco. *Geophys. Res. Lett.*, 34 :L17702.
- Estella-Perez, V., Mignot, J., Guilyardi, E., Swingedouw, D., and Reverdin, G. (2020). Advances in reconstructing the AMOC using sea surface observations of salinity. *Clim. Dyn.*, 55 :975–992.
- Etheridge, D. M., Steele, L. P., Langenfelds, R. L., and Francey, R. J. (1996). Natural and anthropogenic changes in atmospheric CO₂ over the last 1000 years from air in Antarctic ice and firn. *Jour. Geophys. Res.*, 101 :4115–4128.
- Evan, A. T., Foltz, G., Zhang, D., and Vimont, D. J. (2011). Influence of African dust on ocean-atmosphere variability in the tropical Atlantic. *Nat. Geosci.*, 4 :762–765.
- Evan, A. T., Vimont, D. J., Heidinger, A. K., Kossin, J. P., and Bennartz, R. (2009). The Role of Aerosols in the Evolution of Tropical North Atlantic Ocean Temperature Anomalies. *Science*, 324 :778–781.
- Fisher, D. A., Koerner, R. M., and Reeh, N. (1995). Holocene climatic records from Agassiz Ice Cap, Ellesmere Island, NWT, Canada. *The Holocene*, 5 :19–24.
- Frajka-Williams, E., Beaulieu, C., and Ducheze, A. (2017). Emerging negative Atlantic Multidecadal Oscillation index in spite of warm subtropics. *Sci. Rep.*, 7 :11224.
- Frankcombe, L. M. and Dijkstra, H. A. (2009). Coherent multidecadal variability in North Atlantic sea level : Multidecadal variability in sea level. *Geophys. Res. Lett.*, 36 :n/a–n/a.
- Frankcombe, L. M., Dijkstra, H. A., and von der Heydt, A. (2008). Sub-surface signatures of the Atlantic Multidecadal Oscillation. *Geophys. Res. Lett.*, 35 :L19602.
- Frankcombe, L. M., von der Heydt, A., and Dijkstra, H. (2010). North Atlantic multidecadal climate variability : An investigation of dominant time scales and processes. *J. Climate*, 23 :3626–3638.
- Franke, J., Kreiss, J.-P., and Mammen, E. (2002). Bootstrap of kernel smoothing in non-linear time series. *Bernouilli*, 8 :1–37.
- Frankignoul, C., Gastineau, G., and Kwon, Y.-O. (2017). Estimation of the SST Response to Anthropogenic and External Forcing and Its Impact on the Atlantic Multidecadal Oscillation and the Pacific Decadal Oscillation. *J. Climate*, 30 :9871–9895.
- Frankignoul, C. and Hasselmann, K. (1977). Stochastic climate models, Part II Application to sea-surface temperature anomalies and thermocline variability. *Tellus*, 29 :289–305.
- Frankignoul, C., Kestenare, É., and Mignot, J. (2002). The surface heat flux feedback. Part II : direct and indirect estimates in the ECHAM4/OPA8 coupled GCM. *Clim. Dyn.*, 19 :649–655.
- Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, 33 :1–22.
- Friedman, J. H. (1999). Stochastic gradient boosting. *Computational Statistics & Data Analysis*, 38 :367–378.

- Gadgil, S. (2018). The monsoon system : Land–sea breeze or the ITCZ? *J. Earth Syst. Sci.*, 127 :1.
- Ganachaud, A. and Wunsch, C. (2000). Improved estimates of global ocean circulation, heat transport and mixing from hydrographic data. *Nature*, 408 :453–457.
- Gao, C., Robock, A., and Ammann, C. (2008). . *Volcanic forcing of climate over the past 1500 years : an improved ice core-based index for climate models*, 113 :D23111.
- Garfinkel, C. I. and Hartmann, D. L. (2008). Different ENSO teleconnections and their effects on the stratospheric polar vortex. *Journ. Geophys. Res. Atm.*, 113.
- Geisser, S. (1975). The predictive sample reuse method with applications. *Journal of the Royal Statistical Society*, 70 :320–328.
- George, S. S. and Nielsen, E. (2002). Hydroclimatic Change in Southern Manitoba Since A.D. 1409 Inferred from Tree Rings. *Quaternary Research*, 58 :103–111.
- Geurts, P., Ernst, D., and Wehenkel, L. (2006). Extremely randomized trees. *Mach. Learn.*, 63 :3–42.
- GIEC (2014). Annex II - Glossary. In Agard, J., Shipper, E. L. F., Birkmann, J., Campos, M., Dubeux, C., Nojiri, Y., Olsson, L., Osman-Elasha, B., Pelling, M., Prather, M. J., Rivera-Ferre, M. G., Ruppel, O. C., Sallenger, A., Smith, K. R., and St. Clair, A. L., editors, *Climate Change 2013 - The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, pages 1757–1776. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, intergovernmental panel on climate change edition.
- Gneiting, T. and Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102 :359–378.
- Goldstein, B. A., Polley, E. C., and Briggs, F. B. S. (2011). Random forests for genetic association studies. *Stat. Appl. Genet. Mol. Bio.*, 10 :32.
- Goosse, H., Crespin, E., Dubinkina, S., Loutre, M.-F., Mann, M. E., Renssen, H., Sallaz-Damaz, Y., and Shindell, D. (2012). The role of forcing and internal dynamics in explaining the "Medieval Climate Anomaly". *Climate Dynamics*, 39 :2847–2866.
- Goosse, H., Renssen, H., Timmermann, A., and Bradley, R. S. (2005). Internal and forced climate variability during the last millennium : a model-data comparison using ensemble simulations. *Quaternary Science Reviews*, 24 :1345–1360.
- Graf, H.-F. and Zanchettin, D. (2012). Central Pacific El Niño, the “subtropical bridge,” and Eurasian climate : TWO EL NIÑO TYPES AND EURASIAN CLIMATE. *J. Geophys. Res.*, 117 :n/a–n/a.
- Graumlich, L. J., Pisaric, M. F. J., Waggoner, L. A., Littell, J. S., and King, J. C. (2003). Upper Yellowstone river flow and teleconnections with Pacific basin climate variability during the past three centuries. *Climatic change*, 59 :245–262.
- Gray, S. T., Graumlich, L. J., Betancourt, J. L., and Pederson, G. T. (2004). A tree-ring based reconstruction of the Atlantic Multidecadal Oscillation since 1567 A.D. *Geophys. Res. Lett.*, 31 :1–4.

- Graybill, D. A. (1994a). International Tree-ring Data Bank NV516.
- Graybill, D. A. (1994b). International Tree-ring Data Bank NV517.
- Graybill, D. A. (1994c). International Tree-ring Data Bank UT508.
- Graybill, D. A. (1994d). International Tree-ring Data Bank UT509.
- Griffiths, M. L., Kimbrough, A. K., Gagan, M. K., Drysdale, R. N., Cole, J. E., Johnson, K. R., Zhao, J.-X., Cook, B. I., Hellstrom, J. C., and Hantoro, W. S. (2016). Western Pacific hydroclimate linked to global climate variability over the past two millennia. *Nat. Commun.*, 7 :11719.
- Grinsted, A., Moore, J. C., and Jevrejeva, S. (2004). Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlin. Processes Geophys.*, 11 :561–566.
- Guillot, D., Rajaratnam, B., and Emile-Geay, J. (2015). Statistical paleoclimate reconstructions via Markov random fields. *Ann. of App. Stat.*, 1 :324–352.
- Guiot, J. (2012). A robust spatial reconstruction of April to September temperature in Europe : Comparisons between the medieval period and the recent warming with a focus on extreme values. *GPC*, 84-85 :14–22.
- Hakim, G. J., Emile-Geay, J., Steig, E. J., Tardif, R., Steiger, N., and Perkins, W. A. (2016). The last millennium climate reanalysis project : Framework and first results. *Jour. Geophys. Res. Atmos.*, 121 :6745–6764.
- Hall, A., Cox, P. M., Hutingford, C., and Klein, S. (2019). Progressing emergent constraints on future climate change. *Nat. Clim. Chang.*, 9 :269–278.
- Halloran, P. R., Hall, I. R., Menary, M., Reynolds, D. J., Scourse, J. D., Screen, J. A., Bozzo, A., Dunstone, N., Phipps, S., Shurer, A. P., Sueyoshi, T., Zhou, T., and Garry, F. (2020). Natural drivers of multidecadal Arctic sea ice variability over the last millennium. *Nat. Comm.*, 10 :688.
- Hanhijarvi, M., Tingley, M. P., and Korhola, A. (2013). Pairwise Comparisons to Reconstruct Mean Temperature in the Arctic Atlantic Region Over the Last 2000 Years. *Climate dynamics*, 41 :2039–2060.
- Hasselmann, K. (1976). Stochastic climate models Part I. Theory. *Tellus*, 28 :473–485.
- Hawkins, E. and Sutton, R. (2009). The potential to narrow uncertainty in regional climate predictions. *American Meteorological Society*, August :1095–1107.
- Helama, S., Holopainen, J., Timonen, M., and Mielikäinen, K. (2014). An 854-Year Tree-ring chronology of Scots Pine for South-West Finland. *Studia Quaternaria*, 31 :61–68.
- Henley, B. J., Gergis, J., Karoly, D. J., Power, J., Kennedy, J., and Folland, C. K. (2015). A tripole index for the Interdecadal Pacific Oscillation. *Clim. Dyn.*, 45 :3077–3090.
- Hermanson, L., Bilbao, R., Dunstone, N., Ménégos, M., Ortega, P., Pohlmann, H., Robson, J. I., Smith, D. M., Strand, G., Timmreck, C., Yeager, S., and Danabasoglu, G. (2020). Robust Multiyear Climate Impacts of Volcanic Eruptions in Decadal Prediction Systems. *J. Geophys. Res. Atmos.*, 125.

- Hoerl, A. E. and Kennard, R. W. (1970). Ridge regression : Biased estimation of nonorthogonal problems. *Technometrics*, 12 :55–67.
- Homrighausen, D. and McDonald, D. J. (2014). Leave-one-out cross-validation is risk consistent for lasso. *Machine Learning*, 97 :65–78.
- Höskuldsson, A. (1988). PLS regression methods. *J. Chemometrics*, 2 :211–228.
- Hotelling, H. (1933). Analysis of a complex of statistical variables into Principal Components. *Journal of Education Psychology*, 24 :498–520.
- Hotelling, H. (1957). The relations of the newer multivariate statistical methods to factor analysis. *British Journal of Statistical Psychology*, 10 :69–76.
- Hourdin, F., Foujols, M.-A., Codron, F., Guemas, V., Dufresne, J.-L., Bony, S., Denvil, S., Guez, L., Lott, F., Ghattas, J., Braconnot, P., Marti, O., Meurdesoif, Y., and Bopp, L. (2013). Impact of the LMDZ atmospheric grid configuration on the climate and sensitivity of the IPSL-CM5A coupled model. *Clim. Dyn.*, 40 :2167–2192.
- Houtekamer, P. L. and Zhang, F. (2016). Review of the Ensemble Kalman Filter for Atmospheric Data Assimilation. *Monthly Weath. Rev.*, 144 :4489–4532.
- Hugo, V. (1864). L'Art et la Science - Partie 1 - Chapitre 3. In *William Shakespeare*. Library of Alexandria.
- Huntingford, C., Jeffers, E. S., Bonsall, M. B., Christensen, H. M., Lees, T., and Yang, H. (2019). Machine learning and artificial intelligence to aid climate change research and preparedness. *Environ. Res. Lett.*, 14 :124007.
- Hurrell, J. W. (1995). Decadal Trends in the North Atlantic Oscillation : Regional Temperatures and Precipitation. *Science*, 269 :676–679.
- Hurrell, J. W., Kushnir, Y., Ottersen, G., and Visbeck, M. (2003). An overview of the North Atlantic Oscillation. *Geophysical Monograph*, 134 :1–35.
- Isobe, T., Feigelson, E. D., Akritas, M. G., and Babu, G. J. (1990). Linear regression in astronomy. *Astrophysical Journal, Part 1*, 364 :104–113.
- Izumo, T., Khodri, M., Lengaigne, M., and Suresh, I. (2018). A subsurface Indian Ocean Dipole response to tropical volcanic eruptions. *Geophys. Res. Lett.*, 45 :9150–9159.
- Jacobel, A. W., McManus, J. F., Anderson, R. F., and Winckler, G. (2016). Large deglacial shifts of the Pacific Intertropical Convergence Zone. *Nat. Commun.*, 7 :10449.
- Jomelli, V., Lane, T., Favier, V., Masson-Delmotte, V., Swingedouw, D., Rinterknecht, V., Schimmelpfennig, I., Brunstein, D., Verfaillie, D., Adamson, K., Leanni, L., Mokadem, F., and Team, A. (2016). Paradoxical cold conditions during the medieval climate anomaly in the Western Arctic. *Sci. Rep.*, 6 :32984.
- Jones, P. D., Jonsson, T., and Wheeler, D. (1997). Extension to the North Atlantic Oscillation using early instrumental pressure observations from Gibraltar and south-west Iceland. *Int. J. Climatol.*, 17 :1433–1450.

- Jourdain, L., Bekki, S., Lott, F., and Lefèvre, F. (2008). The coupled chemistry-climate model LMDz-REPROBUS : description and evaluation of a transient simulation of the period 1980–1999. *Ann. Geophys.*, 26 :1391–1413.
- Jouzel, J., Vimeux, F., Caillon, N., Delaygue, G., Hoffmann, G., Masson-Delmotte, V., and Parrein, F. (2003). Magnitude of isotope/temperature scaling for interpretation of central Antarctic ice cores. *J. Geophys. Res.*, 108 :4361.
- Kaniewski, D., Van Campo, E., Paulissen, E., Weiss, H., Bakker, J., Rossignol, I., and Van Lerberghe, K. (2011). The medieval climate anomaly and the little Ice Age in coastal Syria inferred from pollen-derived palaeoclimatic patterns. *Global and Planetary Change*, 78 :178–187.
- Karspeck, A. R., Stammer, D., Köhl, A., Danabasoglu, G., Balmaseda, M., Smith, D. M., Fujii, Y., Zhang, S., Giese, B., Tsujino, H., and Rosati, A. (2015). Comparison of the Atlantic meridional overturning circulation between 1960 and 2007 in six ocean reanalysis products. *J. Clim.*, 26 :7392–7413.
- Kaufmann, R. K., Kauppi, H., Mann, M. L., and Stock, J. H. (2011). Reconciling anthropogenic climate change with observed temperature 1998-2008. *PNAS*, 108 :11790–11793.
- Kelly, M. H. and Gore, J. A. (2008). Florida river flow patterns and the Atlantic multidecadal oscillation. *River Reaseach and Applications*, 24 :598–616.
- Kenigson, J. S., Han, W., Rajagopalan, B., Yanto, and Jasinski, M. (2018). Decadal Shift of NAO-Linked Interannual Sea Level Variability along the U.S. Northeast Coast. *J. Climate*, 31 :4981–4989.
- Kerr, R. A. (2000). A North Atlantic Climate Pacemaker for the Centuries. *Science*, 288 :1984–5.
- Khodri, M., Izumo, T., Vialard, J., Janicot, S., Cassou, C., Lengaigne, M., Mignot, J., Gastineau, G., Guilyardi, E., Lebas, N., Robock, A., and McPhaden, M. J. (2017). Tropical explosive volcanic eruptions can trigger El Niño by cooling tropical Africa. *Nature Communications*, 8 :No. 778.
- Kinnard, C., Zdanowicz, C. M., Fisher, D. A., Isaksson, E., de Vernal A., and Thompson, L. (2011). Reconstructed changes in Arctic sea ice over the past 1,450 years. *Nature*, 479 :509–512.
- Knight, J., Folland, C., and Scaife, A. (2006). Climate impacts of the Atlantic Multidecadal Oscillation. *Geophys. Res. Lett.*, 33 :L11706.
- Knight, J. R., Allan, R. J., Folland, C. K., Vellinga, M., and Mann, M. E. (2005). A signature of persistent natural thermohaline circulation cycles in observed climate. *Geophys. Res. Lett.*, 32 :L20708.
- Koenigk, T., Caian, M., Nikulin, G., and Schimanke, S. (2016). Regional Arctic sea ice variations as predictor for winter climate conditions. *Clim. Dyn.*, 46 :317–337.
- Kohavi, R. (1995). A study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection.

- Kosaka, Y. and Xie, S.-p. (2013). Recent global-warming hiatus tied to equatorial Pacific surface cooling. *Nature*, 501 :403–407.
- Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S., and Prentice, I. C. (2005). A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system : DVGM for coupled climate studies. *Global Biogeochem. Cycles*, 19.
- Kushnir, Y. (1994). Interdecadal variations in North Atlantic Sea Surface Temperatures and associated atmospheric conditions. *Journ. Clim.*, 7 :1–20.
- LaMarche, V. C. J. (1974). Paleoclimatic Inferences from Long Tree-Ring Records. *Science*, 183 :1043–1048.
- Lamb, H. H. (1965). The early medieval warm epoch and its sequel. *Paleogeography, Paleoclimatology, Paleoecology*, 1 :13–37.
- Lambert, E., Eldevik, T., and Haugan, P. M. (2016). How northern freshwater input can stabilise thermohaline circulation. *Tellus A : Dynamic Meteorology and Oceanography*, 68 :31051.
- Lavigne, F., Degeai, J.-P., Komorowski, J.-C., Guillet, S., Robert, V., Lahitte, P., Oppenheimer, C., Stoffel, M., Vidal, C. M., Surono, Pratomo, I., Wassmer, P., Hajdas, I., Hadmoko, D. S., and de Belizal, E. (2013). Source of the great A.D. 1257 mystery eruption unveiled, Samalas volcano, Rinjani Volcanic Complex, Indonesia. *PNAS*, 110 :16742–16747.
- Lean, J. L. (2018). Estimating solar irradiance since 850 CE. *Earth Spac. Sci.*, 5 :133–149.
- Lehner, F., Raible, C. C., and Stocker, T. F. (2012). Testing the robustness of a precipitation proxy-based North Atlantic Oscillation reconstruction. *Quaternary Science Reviews*, 45 :85–94.
- Lenssen, N., Schmidt, G., Hansen, J., Menne, M., Persin, A., Ruedy, R., and Zyss, D. (2019). Improvement in the GISTEMP uncertainty model. *J. Geophys. Res. Atmos.*, 124 :6307–6326.
- Lenton, T. M. (2011). Early warning of climate tipping point. *Nat. Clim. Change*, 1 :201–208.
- Lenton, T. M., Livina, V. N., Dakos, V., and Scheffer, M. (2012). Climate bifurcation during the last deglaciation? *Clim. Past*, 8 :1127–1139.
- Li, C., Xie, S.-P., Gille, S. T., and Yoo, C. (2016). Atlantic-induced pan-tropical change over the past three decades. *Nat. Clim. Change*, 6 :275–279.
- Li, J., Xie, S., Cook, E. R., Morales, M. S., Christie, N. C. J., Chen, F., D'Arrigo, R., Fowler, A. M., and Gou, X. (2013). El Niño modulations over the past seven centuries. *Nature climate change*, 3 :822–826.
- Li, J., Xie, S.-P., Cook, E. R., Huang, G., D'Arrigo, R., Liu, F., Ma, J., and Zheng, X.-T. (2011). Interdecadal modulation of El Niño amplitude during the past millennium. *Nat. Clim. Change*, 1 :114–118.

- Liaw, A. and Wiener, M. (2002). Classification and regression by randomforest. *R News*, 2 :18–22.
- Lin, P., Yu, Z., Lü, J., Ding, M., Hu, A., and Liu, H. (2019). Two regimes of Atlantic Multidecadal Oscillation : cross-basin dependant or Atlantic-intrinsic. *Sci. Bull.*, 64 :198–204.
- Lindholm, M. and Jalkanen, R. (2011). Subcentury scale variability in height-increment and tree-ring width chronologies of Scots pine since AD 745 in northern Finland. *The Holocene*, 22 :571–577.
- Liu, P., Tsimpidi, A. P., Hu, Y., Stone, B., Russell, A. G., and Nenes, A. (2012). Differences between downscaling with spectral and grid nudging using WRF. *Atmos. Chem. Phys.*, 12 :3601–3610.
- Luo, W., Nguyen, T., Nichols, M., Tran, T., Rana, S., Gupta, S., Phung, D., Venkatesh, S., and Allender, S. (2015). Is Demography Destiny ? Application of Machine Learning Techniques to Accurately Predict Population Health Outcomes from a Minimal Demographic Dataset. *PLoS ONE*, 10 :e0125602.
- Luterbacher, J., Neukom, R., González-Rouco, F., Fernandez-Donado, L., Raible, C., and Zorita, E. (2011). Reconstructed and simulated Medieval Climate Anomaly in southern South America. *PAGES news*, 19 :20–21.
- Luterbacher, J., Xoplaki, E., Dietrich, D., Jones, P. D., Davies, T. D., Portis, D., Gonzalez-Rouco, J. F., von Storch, H., Gyalistras, D., Casty, C., and Wanner, H. (2001). Extending North Atlantic Oscillation Reconstructions Back to 1500. *Atmos. Sci. Lett.*, 2 :114–124.
- Luterbacher, J., Xoplaki, E., Dietrich, D., Rickli, R., Jacobeit, J., Beck, G., Gyalistras, D., Schmutz, C., and Wanner, H. (2002). Reconstruction of Sea Level Pressure fields over the Eastern North Atlantic and Europe back to 1500. *Clim. Dyn.*, 18 :545–561.
- Macias-Fauria, M., Grinsted, A., Helama, S., and Holopainen, J. (2012). Persistence matters : Estimation of the statistical significance of paleoclimatic reconstruction statistics from autocorrelated time series. *Dendrochronologia*, 30 :179–187.
- Madec, G. (2008). NEMO Ocean Engine. Note du Pole de modélisation, Institut Pierre-Simon Laplace.
- Mann, M. E. and Park, J. (1994). Global-scale modes of surface temperature variability on interannual to century timescales. *Journ. Geophys. Res.*, 99 :25819–25833.
- Mann, M. E., Steinman, B. A., and Miller, S. K. (2014). On forced temperature changes, internal variability, and the AMO. *Geophys. Res. Lett.*, 41 :3211–3219.
- Mann, M. E., Steinman, B. A., and Miller, S. K. (2020). Absence of internal multi-decadal and interdecadal oscillations in climate model simulations. *Nat. Comm.*, 11 :49.
- Mann, M. E., Zhang, Z., Hughes, M. K., Bradley, R. S., Miller, S. K., Rutherford, S., and Ni, F. (2008). Proxy-based reconstructions of hemispheric and global surface temperature variations over the past two millennia. *PNAS*, 35 :13252–13257.

- Mann, M. E., Zhang, Z., Rutherford, S., Bradley, R. S., Hughes, M. K., Shindell, D., Amman, C., Faluvegi, G., and Ni, F. (2009). Global signature and dynamical origins of the Little Ice Age and Medieval Climate Anomaly. *Science*, 326 :1256–1260.
- Marini, C. and Frankignoul, C. (2013). An attempt to deconstruct the Atlantic Multidecadal Oscillation. *Clim. Dyn.*, 43 :607–625.
- Massé, G., Rowland, S. J., Sicre, M. A., Jacob, J., Jansen, E., and Belt, S. T. (2008). Abrupt climate changes for Iceland during the last millennium : evidence from high resolution sea ice reconstructions. *Earth and Planetary Science Letters*, 269 :565–569.
- Maxwell, R. S., Hessler, A. E., Cook, E. R., and Pederson, N. (2011). A multispecies tree ring reconstruction of Potomac River streamflow (950-2001). *Water resources research*, 47 :W05512.
- McCabe-Glynn, S., Johnson, K. R., Strong, C., Berkelhammer, M., Sinhan, A., Cheng, H., and Edwards, R. L. (2013). Variable North Pacific influence on drought in southwestern North America. *Nat. Geosci.*, 6 :617–621.
- McCarroll, D., Loader, N. J., Jalkanen, R., Gagen, M. H., Hakan Grudd, H., and Gunnarson, B. E. (2013). Fennoscandia 1200 Year Tree Growth Data and Summer Temperature Reconstruction. *The Holocene*, 23 :471–484.
- McCarthy, G. D., Haigh, I. D., Hirshi, J. J.-M., Grist, J. P., and Smeed, D. A. (2015). Ocean impact on decadal Atlantic climate variability revealed by sea-level observations. *Nature*, 521 :508–512.
- McCormack, R. L. (1959). An evaluation of two methods of cross-validation. *Psychological Reports*, 5 :127–130.
- McGowan, H., Callow, J. N., Soderholm, J., McGrath, G., Campbell, M., and Zhao, J.-x. (2018). Global warming in the context of 2000 years of Australian alpine temperature and snow cover. *Sci. Rep.*, 8 :4394.
- McGregor, S., Gupta, A. S., Dommengat, D., Lee, T., McPhaden, M. J., and Kessler, W. S. (2017). Factors influencing the skill of synthesized satellite wind products in the tropical Pacific. *Journ. Geophys. Res. Oc.*, 122 :1072–1089.
- McGregor, S. and Timmermann, A. (2011). The effect of explosive tropical volcanism on ENSO. *J. Climate*, 24 :2178–2191.
- McGregor, S., Timmermann, A., Stecker, M., England, M. H., Merrifield, M., Jin, F.-F., and Chikamoto, Y. (2014). Recent Walker circulation strengthening and Pacific cooling amplified by Atlantic circulation. *Nat. Clim. Change*, 4 :888–892.
- Medhaug, I. M., Slope, B., Fischer, E. M., and Knutti, R. (2017). Reconciling controversies about the “global warming hiatus”. *Nature*, 545 :41–47.
- Meeker, L. D. and Mayewski, P. A. (2002). A 1400-year high-resolution record of atmospheric circulation over the North Atlantic and Asia. *The Holocene*, 12 :257–266.
- Mevik, B., Wehrens, R., and Liland, K. H. (2007). The pls package : Principal component and partial least squares regression in r. *Journal of Statistical Software*, 18 :1–23.

- Michel, S., Swingedouw, D., Chavent, M., Ortega, P., Mignot, J., and Khodri, M. (2020). Reconstructing climatic modes of variability using ClimIndRec version 1.0. *Geosci. Model Dev.*, 13 :841–858.
- Michel, S., Swingedouw, D., and Mignot (sub.). Analysis of last millennium ocean variability through a gridded random forest-based SST reconstruction.
- Michel, S., Swingedouw, D., Mignot, J., Gastineau, G., Ortega, P., Khodri, M., and McCarthy, G. D. (in rev.). Internal abrupt change in the Atlantic Multidecadal Variability as the onset of the Little Ice Age.
- Mignot, J., Garcia Serrano, J., Swingedouw, D., Germe, A., Nguyen, S., Ortega, P., Guilyardi, E., and Ray, S. (2016). Decadal prediction skill in the ocean with surface nudging in the IPSL-CM5A-LR climate model. *Clim. Dyn.*, 47 :1225–1246.
- Mignot, J., Khodri, M., Frankignoul, C., and Servonnat, J. (2011). Volcanic impact on the Atlantic ocean over the last millenium. *Clim. Past. Discuss.*, 7 :2511–2554.
- Miles, M., Divine, D., Furevik, T., Jansen, E., Moros, M., and Ogilvie, A. E. J. (2014). A signal of persistent Atlantic Multidecadal variability in Arctic sea ice. *Geophys. Res. Lett.*, 41 :463–469.
- Miller, G. H., Geirsdóttir, A., Zhong, Y., Larsen, D. J., Otto-Bliesner, B. L., Holland, M. M., Bailey, D. A., Refsnider, K. A., Lehman, S. J., Southon, J. R., Anderson, C., Björnsson, H., and Thordarson, T. (2012). Abrupt onset of the Little Ice Age triggered by volcanism and sustained by sea-ice/ocean feedbacks. *Geophys. Res. Lett.*, 39 :1–5.
- Mitchell, J. M. J., Dzerdzeevskii, B., Flohn, H., Hofmeyr, W. L., Lamb, H. H., Rao, K. N., and Wallén, C. C. (1966). *Climatic change : Technical note No. 79, report of a working group for the commission of climatology*. World Meteorological Organization, Geneva, Switzerland.
- Mjell, T. L., Ninnemann, U. S., Kleiven, H. F., and Hall, I. R. (2016). Multidecadal changes in Iceland Scotland Overflow Water vigor over the last 600 years and its relationship to climate : 600 years of isow vigor variability. *Geophys. Res. Lett.*, 43 :2111–2117.
- Mjell, T. R., Ninnemann, U. S., Eldevik, T., and Kleiven, H. K. F. (2015). Holocene multidecadal-to millennial-scale variations in Iceland-Scotland overflow and their relationship to climate. *Paleoceanography and Paleoclimatology*, 30 :558–569.
- Moffa-Sánchez, P., Born, A., Hall, I. R., Thornalley, D. J. R., and Barker, S. (2014). Solar forcing of North Atlantic surface temperature and salinity over the past millennium. *Nat. Geosci.*, 7 :275–278.
- Moffa-Sánchez, P. and Hall, I. R. (2017). North Atlantic variability and its links to European climate over the last 3000 years. *Nat. Commun.*, 8 :1726.
- Moffa-Sánchez, P., Moreno-Chamarro, E., Reynolds, D. J., Ortega, P., Cunningham, L., Swingedouw, D., Amrhein, D. E., Halfar, J., Jonkers, L., Jungclaus, J. H., Perner, K., Wanamaker, A., and Yeager, S. (2019). Variability in the Northern North Atlantic and Arctic Oceans Across the Last Two Millennia : A Review. *Paleoceanography and Paleoclimatology*, 34 :1399–1436.

- Moore, G. W. K., Halfar, J., Majeed, H., Adey, W., and Kronz, A. (2017). Amplification of the Atlantic Multidecadal Oscillation associated with the onset of the industrial-era warming. *Sci. Rep.*, 7 :40861.
- Moreau, R. (2013). *L'air et l'eau*. Edp sciences, collection grenoble sciences edition.
- Moreno-Chamorro, E., Zanchettin, D., Lohmann, K., and Jungclaus, J. H. (2016). An abrupt weakening of the subpolar gyre as trigger of LIA-type episodes. *Clim. Dyn.*, 48 :727–744.
- Morice, C. P., Kennedy, J. J., Rayner, N. A., and Jones, P. D. (2012). Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates : the HadCRUT4 dataset. *Journ. Geophys. Res.*, 117 :D08101.
- Muir, L. C. and Fedorov, A. V. (2015). How the AMOC affects ocean temperatures on decadal to centennial timescales : the North Atlantic versus an interhemispheric seesaw. *Clim. Dyn.*, 45 :151–160.
- Mysterud, A., Stenseth, N. C., Yoccoz, N. G., Langvatn, R., and Steinheim, G. (2001). Nonlinear effects of large-scale climatic variability on wild and domestic herbivores. *Nature*, 410 :1096–1099.
- Nash, J. E. and Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I : A discussion of principles. *J. Climatol.*, 10 :282–290.
- Naurzbaev, M. M., Vaganov, E. A., Sidorova, O. V., and Schweingruber, F. H. (2002). Summer temperatures in eastern Taimyr inferred from a 2427-year late-Holocene tree-ring chronology and earlier floating series. *The Holocene*, 12 :727–736.
- Neelin, J. D., Anthony, S. B., Hirst, A. C., Jin, F.-f., Wakata, Y., Yamagata, T., and Zebiak, S. E. (1998). ENSO theory. *Journal of Geophysical Research*, 103 :14261–14290.
- Neukom, R., Shurer, A. P., Steiger, N. J., and Hegerl, G. C. (2018). Possible causes of data model discrepancy in the temperature history of the last Millennium. *Sic. Rep.*, 8 :7572.
- Neukom, R., Steiger, N., Gómez-Navarro, J. J., Wang, J., and Werner, J. P. (2019). No evidence for globally coherent warm and cold periods over the preindustrial Common Era. *Nature*, 571 :550–554.
- Nigam, S., Guan, B., and Ruiz-Barradas, A. (2011). Key role of the Atlantic Multidecadal Oscillation in 20th century drought and wet periods over the Great Plains. *Geophys. Res. Lett.*, 38 :L16713.
- Ntampaka, M., Avestruz, C., Boada, S., Caldeira, J., Cisewski-Kehe, J., Di Stefano, R., Dvorkin, C., Evrard, A. E., Farahi, A., Finkbeiner, D., Genel, S., Goodman, A., Goulding, A., Ho, S., Kosowsky, A., La Plante, P., Lanusse, F., Lochner, M., Mandelbaum, R., Nagai, D., Newman, J. A., Nord, B., Peek, J. E. G., Peel, A., Poczos, B., Rau, M. M., Siemiginowska, A., Sutherland, D. J., Trac, H., and Wandelt, B. (2019). The Role of Machine Learning in the Next Decade of Cosmology. *Astrophysics*, Preprint. arXiv : 1902.10159.

- Ohba, M., Shiogama, H., Yokohata, T., and Wanatabe, M. (2013). Impact of strong tropical volcanic eruptions on ENSO simulated in a coupled GCM. *J. Climate*, 26 :5169–5182.
- O'Reilly, C. H., Woollings, T., and Zanna, L. (2017). The Dynamical Influence of the Atlantic Multidecadal Oscillation on Continental Climate. *J. Clim.*, 30 :7213–7230.
- Ortega, P., Lehner, F., Swingedouw, D., Masson-Delmotte, V., Raible, C. C., Casado, M., and Yiou, P. (2015). A model-tested North Atlantic Oscillation reconstruction for the past millennium. *Nature*, 523 :71–74.
- Ortega, P., Robson, J., Sutton, R. T., and Andrews, M. B. (2017). Mechanisms of decadal variability in the Labrador sea and the wider North Atlantic in a high-resolution climate model. *Clim. Dyn.*, 49 :2625–2647.
- Oshiro, T. M., Perez, P. S., and Baranauskas, J. A. (2012). How many trees in a random forest? *Lect. Note Comp. Sci.*, 7376 :154–168.
- PAGES 2k Consortium (2013). Continental-scale temperature variability during the past two millennia. *Nat. Geosci.*, 6 :339–346.
- PAGES 2k Consortium (2017). A global multiproxy database for temperature reconstructions of the Common Era. *Scientific Data*, 4.
- Paillard, D. (2008). From atmosphere, to climate, to Earth system science. *Interdisc. Sci. Rev.*, 33 :25–35.
- Parker, D., Folland, C., Scaife, A., Knight, J., Colman, A., Baines, P., and Dong, B. (2007). Decadal to multidecadal variability and the climate change background. *J. Geophys. Res.*, 112 :D18115.
- Peings, Y. and Magnusdottir, G. (2014). Forcing of the wintertime atmospheric circulation by the multidecadal fluctuations of the North Atlantic ocean. *Environ. Res. Lett.*, 9 :034018.
- Persechino, A., Mignot, J., Swingedouw, D., Labetoulle, S., and Guilyardi, E. (2013). Decadal predictability of the Atlantic meridional overturning circulation and climate in the IPSL-CM5A-LR model. *Clim. Dyn.*, 40 :2359–2380.
- Picault, J., Ioualalen, M., Menkes, C., Delcroix, T., and McPhaden, M. J. (1996). Mechanism of the zonal displacements of the Pacific warm pool : implications for ENSO. *Science*, 274 :1486–1489.
- Pierce, D. (2017). *ncdf4 : Interface to Unidata netCDF (Version 4 or Earlier) Format Data Files*. R package version 1.16.
- Planck, M. (1949). The Meaning and Limits of Exact Science. *Science*, 110 :319–327.
- Poole, M. A. and O'Farrell, P. (1971). The assumption of the linear regression model. In *Transactions of the Institute of the British Geographers*, page 145.
- Preisendorfer, R. W. (1988). Rule N. In *Principal Components Analysis in Meteorology and Oceanography*, chapter 3.d., pages 199–204.

- Qasmi, S., Cassou, C., and Boé, J. (2017). Teleconnection between Atlantic multidecadal variability and european temperature : diversity and evaluation of the Coupled Model Intercomparison Project Phase 5 models. *Geophys. Res. Lett.*, 44 :11140–11149.
- Qu, K., Guo, F., Liu, X., Lin, Y., and Zou, Q. (2019). Application of Machine Learning in Microbiology. *Front. Microbiol.*, 10 :827.
- Rahmstorf, S. (2002). Ocean circulation and climate during the past 120,000 years. *Nature*, 419 :207–214.
- Rahmstorf, S. (2006). Thermohaline Ocean Circulation. In *Encyclopedia of Quaternary Sciences*. Elsevier edition.
- Rahmstorf, S., Box, J. E., Feulner, G., Mann, M. E., Robinson, A., Rutherford, S., and Schaffernicht, E. J. (2015). Exceptional twentieth-century slowdown in Atlantic ocean overturning circulation. *Nat. Clim. Change*, 5 :475–480.
- Rao, M. P., Cook, E. R., Cook, B I an Anchukaitis, K. J., D'Arrigo, R. D., Krusic, P. J., and LeGrande, A. N. (2019). A double bootstrap approach to Superposed Epoch Analysis to evaluate response uncertainty. *Dendrochronologia*, 55 :119–124.
- Ravestein, P., van der Schrier, G., Haarsma, R., Scheele, R., and van den Broek, M. (2018). Vulnerability of European intermittent renewable energy supply to climate change and climate variability. *Renewable and Sustainable Energy Reviews*, 97 :497–508.
- Rayner, N. A., Parker, D. E., Horton, E. B., Folland, C. K., Alexander, L. V., Rowell, D. P., Kent, E. C., and Kaplan, E. (2003). Globale analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth. *Journ. Geophys. Res.*, 108 :D14.
- Reinsel, D., Gantz, J., and Rydning, J. (2018). The digitization of the world from edge to core. *IDC White Paper*.
- Reynolds, D. J., Scourse, J. D., Halloran, P. R., Nederbragt, A. J., Wanamaker, A. D., Butler, P. G., Richardson, C. A., Heinemeier, J., Eiriksson, J., Knudsen, K. L., and Hall, I. R. (2016). Annually resolved North Atlantic marine climate over the last millennium. *Nature Communications*, 7.
- Robock, A. (2000). Volcanic eruptions and climate. *Rev. Geophys.*, 38 :191–219.
- Robock, A. and Mao, J. (1995). The Volcanic Signal in Surface Temperature Observations. *J. Climate*, 8 :1086–1103.
- Robson, J., Sutton, R. T., Archibald, A., Cooper, F., Christensen, M., Gray, L. J., Holliday, N. P., Macintosh, C., McMillan, M., Moat, B., Russo, M., Tilling, R., Carslaw, K., Desbruyères, D., Embury, O., Feltham, D. L., Grosvenor, D. P., Josey, S., King, B., Lewis, A., McCarthy, G. D., Merchant, C., New, A. L., O'Reilly, C. H., Osprey, S. M., Read, K., Scaife, A., Shepherd, A., Sinha, B., Smeed, D., Smith, D., Ridout, A., Woollings, T., and Yang, M. (2018). Recent multivariate changes in the North Atlantic climate system, with a focus on 2005-2016. *Int. J. Climatol.*, 38 :5050–5076.

- Robson, J. I., Sutton, R. T., and Smith, D. M. (2012). Initialized decadal predictions of the rapid warming of the North Atlantic Ocean in the mid 1990s : PREDICTIONS OF 1990s ATLANTIC WARMING. *Geophys. Res. Lett.*, 39.
- Roche, D. M. (2013). $\delta^{18}\text{O}$ water isotope in the iLOVECLIM model (version 1.0)– Part 1 : Implementation and verification. *Geosci. Model Dev.*, 6 :1481–1491.
- Roche, D. M. and Caley, T. (2013). $\delta^{18}\text{O}$ water isotope in the iLOVECLIM model (version 1.0) – Part 2 : Evaluation of model results against observed $\delta^{18}\text{O}$ in water samples. *Geosci. Model Dev.*, 6 :1493–1504.
- Rubino, M., Etheridge, D. M., Thornton, D. P., Howden, R., Allison, C. E., Francey, R. J., Langenfelds, R. L., Steele, L. P., Trudinger, C. M., Spencer, D. A., Curran, M. A. J., van Ommen, T. D., and Smith, A. M. (2019). Revised records of atmospheric trace gases CO_2 , CH_4 , N_2O , and $\delta^{13}\text{C}-\text{CO}_2$ over the last 2000 years from Law Dome, Antarctica. *Earth Syst. Sci. Data*, 11 :473–492.
- Ruprich-Robert, Y., Msadek, R., Castruccio, F., Yaeger, S., Delworth, T., and Danabasoglu, G. (2017). Combined influences of seasonal East Atlantic Pattern and North Atlantic Oscillation to excite Atlantic multidecadal variability in a climate model. *J. Clim.*, 3 :2785–2810.
- Salzer, M. W. and Kipfmüller, K. F. (2005). Reconstructed Temperature and Precipitation on a Millennial Timescale from Tree-Rings in the Southern Colorado Plateau, U.S.A. *Climatic change*, 70 :465–487.
- Santer, B. D., Bonfils, C., Painter, J. F., Zelinka, M. D., Mears, C., Solomon, S., Schmidt, G. A., Fyfe, J. C., Cole, J. N. S., Nazarenko, L., Taylor, K. E., and Wentz, F. J. (2014). Volcanic contribution to decadal changes in tropospheric temperatures. *Nat. Geosci.*, 7 :185–189.
- Schlesinger, M. E. and Ramankutty, N. (1994). An oscillation in the global climate system of period 65–70 years. *Nature*, 367 :723–726.
- Schleussner, C., Divine, D., Donges, J., Miettinen, A., and Donner, R. V. (2015). Indications for a North Atlantic Ocean regime shift at the onset of the Little Ice Age. *Clim. Dyn.*, 45 :3623–3633.
- Schneider, T. (2001). Analysis of Incomplete Climate Data : Estimation of Mean Values and Covariance Matrices and Imputation of Missing Values. *J. Clim.*, 14 :853–871.
- Schurer, A. P., Tett, S. F. B., and Hegerl, G. C. (2014). Small influence of solar variability on climate over the past millennium. *Nat. Geosci.*, 7 :104–108.
- Schweingruber, F. H. (1998). International Tree-ring Data Bank SWIT177.
- Seidenglanz, A., Prange, M., Varma, V., and Schulz, M. (2012). Ocean temperature response to idealized Gleissberg and de Vries solar cycles in a comprehensive climate model. *Geophys. Res. Lett.*, 39 :1–6.
- Servonnat, J., Mignot, J., Guilyardi, E., Swingedouw, D., Séférian, R., and Labetoulle, S. (2015). Reconstructing the subsurface ocean decadal variability using surface nudging in a perfect model framework. *Clim. Dyn.*, 44 :315–338.

- Sévellec, F. and Huck, T. (2015). Theoretical Investigation of the Atlantic Multidecadal Oscillation. *Journ. Phys. Oc.*, 45 :2189–2208.
- Sgubin, G., Swingedouw, D., Drijfhout, S., Mary, Y., and Bennabi, A. (2017). Abrupt cooling over the North Atlantic in modern climate models. *Nat. Commun.*, 8 :14375.
- Shanahan, T. M., Overpeck, J. T., Anchukaitis, K. J., Beck, J. W., Cole, J. E., Dettman, D. L., Peck, J. A., Scholz, C. A., and King, J. W. (2009). Atlantic Forcing of Persistent Drought in West Africa. *Science*, 324 :377–380.
- Shindell, D. T., Schmidt, G. A., Mann, M. E., and Faluvegi, G. (2004). Dynamic winter climate response to large tropical volcanic eruptions since 1600. *Journal of Geophysical Research*, 109 :D05104.
- Sigl, M., Winstrup, M., McConnell, J. R., Welten, K. C., Plunkett, G., Ludlow, F., Büntgen, U., Caffee, M., Chellman, N., Dahl-Jensen, D., Fisher, H., Kipfstuhl, S., Kostick, C., Maselli, J., Mekhaldi, F., Mulvaney, R., Muscheler, R., Pasteris, D. R., Pilcher, J. R., Salzer, M., Schüpbach, S., Steffensen, J. P., Vinther, B. M., and Woodruff, T. E. (2015). Timing and climate forcing of volcanic eruptions for the past 2,500 years. *Nature*, 523 :543–549.
- Singh, D., Seager, R., Cook, B. I., Cane, M., Ting, M., Cook, E., and Davis, M. (2018a). Climate and the Global Famine of 1876–78. *J. Climate*, 31 :9445–9467.
- Singh, H. K. A., Hakim, G. J., Tardif, R., Emile-Geay, J., and Noone, D. C. (2018b). Insights into Atlantic multidecadal variability using the Last Millennium Reanalysis framework. *Jour. Geophys. Res. Atm.*, 14 :157–174.
- Sinha, A., Cannariato, K. G., Stott, L. D., Cheng, H., Edwards, R. L., Yadava, M. G., Ramesh, R., and Singh, I. B. (2007). A 900-year (600 to 1500 A.D.) record of the Indian summer monsoon precipitation from the core monsoon zone of India : INDIAN SUMMER MONSOON PRECIPITATION. *Geophys. Res. Lett.*, 34.
- Smeed, D. A., Josey, S. A., Beaulieu, C., Johns, W. E., Moat, B. I., Frajka-Williams, E., Rayner, D., Meinsen, C. S., Baringer, M. O., Bryden, H. L., and McCarthy, G. D. (2018). The North Atlantic Ocean is in a state of reduced overturning. *Geophys. Res. Lett.*, 45 :1527–1533.
- Smeed, D. A., McCarthy, G. D., Cunningham, S. A., Frajka-Williams, E., Rayner, D., Johns, W. E., Meinen, C. S., Baringer, M. O., Moat, B. I., Duche, A., and Bryden, H. L. (2014). Observed decline of the Atlantic meridional overturning circulation 2004–2012. *Ocean Sci.*, 10 :29–38.
- Smith, D. M., Booth, B. B. B., Dunstone, N. J., Eade, R., Hermanson, L., Jones, G. S., Scaife, A. A., Sheen, K. L., and Thompson, V. (2016). Role of volcanic and anthropogenic aerosols in the recent global surface warming slowdown. *Nat. Clim. Change*, 6 :936–940.
- Smith, D. M., Scaife, A. A., Eade, R., Athanasiadis, P., Bellucci, A., Bethke, I., Bilbao, R., Borchert, L. F., Caron, L.-P., Counillon, F., Danabasoglu, G., Delworth, T., Doblas-Reyes, F. J., Dunstone, N. J., Estella-Perez, V., Flavoni, S., Hermanson, L., Keenlyside, N., Kharin, V., Kimoto, M., Merryfield, W. J., Mignot, J., Mochizuki, T., Modali, K., Monerie, P.-A., Müller, W. A., Nicolí, D., Ortega, P., Pankatz, K., Pohlmann, H., Robson, J., Ruggieri, P., Sospedra-Alfonso, R., Swingedouw, D., Wang,

- Y., Wild, S., Yeager, S., Yang, X., and Zhang, L. (2020). North Atlantic climate far more predictable than models imply. *Nature*, 583 :796–800.
- Smith, T. M., Reynolds, R. W., Peterson, T. C., and Lawrimore, J. (2008). Improvements to NOAA's Historical Merged Land–Ocean Surface Temperature Analysis (1880–2006). *J. Climate*, 21 :2283–2296.
- Stahle, D. K., Burnette, D. J., and Stahle, D. W. (2013). A Moisture Balance Reconstruction for the Drainage Basin of Albemarle Sound, North Carolina. *Estuaries and Coasts*, 36 :1340–1353.
- Stahle, D. W. (1996a). International Tree-ring Data Bank AR050.
- Stahle, D. W. (1996b). International Tree-ring Data Bank LA001.
- Stahle, D. W. and Cleaveland, M. K. (2005a). International Tree-ring Data Bank AR052.
- Stahle, D. W. and Cleaveland, M. K. (2005b). International Tree-ring Data Bank FL001.
- Stahle, D. W., Villanueva Diaz, J., Brunette, D. J., Cerano Paredes, J., Heim Jr., R. R., Fye, F. K., Acuna Soto, R., Therell, M. D., Cleaveland, M. K., and Stahle, D. K. (2011). Major Mesoamerican droughts of the past millennium. *Geophys. Res. Lett.*, 38 :L05703.
- Stauffer, D. R. and Seaman, N. L. (1990). Use of Four-Dimensional Data Assimilation in a Limited-Area Mesoscale Model. Part I : Experiments with Synoptic-Scale Data. *Monthly Weath. Rev.*, 118 :1250–1277.
- Steinman, B. A., Mann, M. E., and Miller, S. K. (2015). Atlantic and Pacific multidecadal oscillations and Northern Hemisphere temperatures. *Science*, 347 :988–991.
- Stenchikov, G., Robock, A., Ramaswamy, V., Schwarzkopf, M. D., Hamilton, K., and Ramachandran, S. (2002). Arctic Oscillation response to the 1991 Mount Pinatubo eruption : Effects of volcanic aerosols and ozone depletion. *J. Geophys. Res.*, 107 :4803.
- Stevenson, S., Fasullo, J. T., Otto-Bliesner, B. T., and Thomas, R. A. (2017). Role of eruption season in reconciling model and proxy responses to tropical volcanism. *PNAS*, 114 :1822–1826.
- Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M. M. B., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M. (2013). Climate Change 2013, The Physical Science Basis. Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.
- Stommel, H. (1961). Thermohaline Convection with Two Stable Regimes of Flow. *Tellus*, 13 :224–230.
- Stone, M. (1974). Cross-Validatory choice and assesment of statistical predictions. *Journal of the Royal Statistical Society*, 36 :111–147.
- Sutton, R. T. and Dong, B. (2012). Atlantic Ocean influence on a shift in European climate in the 1990s. *Nat. Geosci.*, 5 :788–792.

- Sutton, R. T. and Hodson, D. L. R. (2005). Atlantic Ocean Forcing of North American and European Summer Climate. *Science*, 309 :115–118.
- Swingedouw, D., Bopp, L., Matras, A., and Braconnot, P. (2007). Effect of land-ice melting and associated changes in the AMOC result in little overall impact on oceanic CO₂ uptake. *Geophys. Res. Lett.*, 34 :L23706.
- Swingedouw, D., Mignot, J., Labetoule, S., Guilyardi, E., and Madec, G. (2013). Initialisation and predictability of the AMOC over the last 50 years in a climate model. *Clim. Dyn.*, 40 :2381–2399.
- Swingedouw, D., Mignot, J., Ortega, P., Khodri, M., Menegoz, M., Cassou, C., and Hanquiez, V. (2017). Impact of explosive volcanic eruptions on the main climate variability modes. *Global and Planetary Change*, 150 :24–45.
- Swingedouw, D., Ortega, P., Mignot, J., Guilyardi, E., Masson-delmotte, V., Butler, P. G., Khodri, M., and Séférian, R. (2015). Bidecadal North Atlantic ocean circulation variability controlled by timing of volcanic eruptions. *Nature Communications*, 6 :No. 6545.
- Swingedouw, D., Terray, L., Cassou, C., Voldoire, A., Salas-Mélia, D., and Servonnat, J. (2011). Natural forcing of climate during the last millennium : fingerprint of solar variability. *Climate Dynamics*, 36 :1349–1364.
- Tanaka, H. L., Ishizaki, N., and Kitoh, A. (2004). Trend and interannual variability of Walker, monsoon and Hadley circulations defined by velocity potential in the upper troposphere. *Tellus*, 56 :250–269.
- Tandon, N. F. and Kushner, P. J. (2015). Does External Forcing Interfere with the AMOC's Influence on North Atlantic Sea Surface Temperature? *J. Climate*, 28 :6309–6323.
- Tans, P. and Keeling, R. (2020). Mauna Loa CO₂ data - NOAA.
- Thornalley, D. J. R., Elderfield, H., and McCave, I. N. (2009). Holocene oscillations in temperature and salinity of the surface subpolar North Atlantic. *Nature*, 457 :711–714.
- Thornalley, D. J. R., Oppo, D. W., Ortega, P., Robson, J. I. and Brierly, C. M., Davis, R., Moffa-Sanchez, P., Rose, N. L., Spooner, P. T., Yashayaev, I., and Keigwin, L. D. (2018). Anomalously weak Labrador Sea convection and Atlantic overturning during the past 150 years. *Nature*, 556 :227–230.
- Tibshirani, R. (1996). Regression shrinkage and selection via Lasso. *Journal of the Royal Statistical Society*, 58 :267–288.
- Ting, M., Kushnir, Y., Saeger, R., and Cuihua, L. (2009). Forced and internal twentieth-century SST trends in the North Atlantic. *Journ. Clim.*, 22 :1469–1481.
- Tingley, M. P. (2012). A Bayesian ANOVA Scheme for Calculating Climate Anomalies, with Applications to the Instrumental Temperature Record. *J. Clim.*, 25 :777–791.
- Tingley, M. P., Craigmile, P. F., Haran, M., Li, B., Mannshardt, E., and Rajaratnam, B. (2012). Piecing together the past : statistical insights into paleoclimatic reconstructions. *Quaternary Science Reviews*, 35 :1–22.

- Tingley, M. P. and Huybers, P. (2010a). A Bayesian Algorithm for Reconstructing Climate Anomalies in Space and Time. Part I : Development and Applications to Paleoclimate Reconstruction Problems. *J. Clim.*, 23 :2759–2781.
- Tingley, M. P. and Huybers, P. (2010b). A Bayesian Algorithm for Reconstructing Climate Anomalies in Space and Time. Part II : Comparison with the Regularized Expectation-Maximization Algorithm. *J. Clim.*, 23 :2782–2800.
- Tingley, M. P. and Huybers, P. (2013). Recent temperature extremes at high northern latitudes unprecedented in the past 600 years. *Nature*, 496 :201–5.
- Tosh, R. (1994). International Tree-ring Data Bank CA051.
- Touchan, R., Garfin, G. M., Meko, D. M., Funkhouser, G., Erkan, N., Hughes, M. K., and Wallin, B. S. (2003). Preliminary reconstructions of spring precipitation in southwestern Turkey from tree-ring width. *Int. J. Climatol.*, 23 :157–171.
- Touchan, R., Woodhouse, C. A., Meko, D. M., and Allen, C. (2011). Millennial precipitation reconstruction for the Jemez Mountains, New Mexico, reveals changing drought signal. *Int. J. Climatol.*, 31 :896–906.
- Trenberth, K. E. and Fasullo, J. T. (2017). Atlantic meridional heat transports computed from balancing Earth's energy locally. *Geophys. Res. Lett.*, 44 :1919–1927.
- Trenberth, K. E., Fasullo, J. T., and Kiehl, J. (2009). Earth's Global Energy Budget. *Bull. Amer. Meteor. Soc.*, 90 :311–324.
- Trenberth, K. E. and Shea, D. J. (2006). Atlantic hurricanes and natural variability in 2005. *Geophys. Res. Lett.*, 33 :1–4.
- Trigo, R. M., Vaquero, J. M., Alcoforado, M.-J., Barriendos, M., Taborda, J., García-Herrera, R., and Luterbacher, J. (2009). Iberia in 1816, the year without a summer. *Int. J. Climatol.*, 29 :99–115.
- Trouet, V., Esper, J., Graham, N. E., Baker, A., Scourse, J. D., and Frank, D. C. (2009). Persistent positive North Atlantic oscillation mode dominated the Medieval Climate Anomaly. *Science*, 324 :78–80.
- Undorf, S., Polson, D., Bollasina, M. A., Ming, Y., Schurer, A., and Hegerl, G. C. (2018). Detectable impact of local and remote anthropogenic aerosols on the 20th century changes of west african and south asian monsoon precipitation. *J. Geophys. Res. Atmos.*, 123 :4871–4889.
- Vapnik, V. N. (2000). *The Nature of Statistical Learning Theory*. Statistics for Engineering and Information Science.
- Vieira, L. E. A., Solanki, S. K., Krivova, N. A., and Usoskin (2011). Evolution of the solar irradiance during the Holocene. *Astronomy, Astrophysics*, 531 :A6.
- Visbeck, M., Chassignet, E. P., Curry, R. G., Delworth, T. L., Dickson, R. R., and Krahnemann, G. (2003). The Ocean's Response to North Atlantic Oscillation Variability. In Hurrell, J. W., Kushnir, Y., Ottersen, G., and Visbeck, M., editors, *The North Atlantic : Climatic Significance and Environmental impacts*.
- Wald, L. (2007). Solar Radiation Energy (Fundamentals). *Sol. Energy Conserv. and Photoenerg. Syst.*, 1.

- Wanatabe, M., Kamae, Y., Yoshimori, M., Oka, A., Sato, M., Ishii, M., Mochizuki, T., and Kimoto, M. (2013). Strengthening of ocean heat uptake efficiency associated with the recent climate hiatus. *Nat. Geosci.*, 40 :3175–3179.
- Wanatabe, M. and Tatebe, H. (2019). Reconciling roles of sulphate aerosol forcing and internal variability in Atlantic multidecadal climate changes. *Clim. Dyn.*, 53 :1–15.
- Wang, J., Emile-Geay, J., Guillot, D., Smerdson, J. E., and Rajaratnam, B. (2014). Evaluating climate field reconstruction techniques using improved emulations of real-world conditions. *Clim. Past.*, 10 :1–19.
- Wang, J., Yang, B., Ljungqvist, F. C., Luterbacher, J., Osborn, T. J., Briffa, K. R., and Zorita, E. (2017). Internal and external forcing of multidecadal Atlantic climate variability over the past 1,200 years. *Nat. Geosci.*
- Wen, N., Liu, Z., Liu, Q., and Frankignoul, C. (2005). Observations of SST, heat flux and North Atlantic Ocean-atmosphere interaction. *Geophys. Res. Lett.*, 32 :L24619.
- Wickham, H. (2017). *stringr : Simple, Consistent Wrappers for Common String Operations*. R package version 1.2.0.
- Wilson, R., Miles, D., Loader, N. J., Cooper, R., and Briffa, K. (2013). A millennial long March-July precipitation reconstruction for southern-central England. *Climate Dynamics*.
- Wold, S., Ruhe, A., Wold, H., and Dunn III, W. J. (1984). The collinearity problem in linear regression. The Partial Least Squares (PLS) approach to generalized inverses. *J. Sci. Stat. Comput.*, 5 :735–743.
- Wolpert, D. H. (1996). The Lack of A Priori Distinctions Between Learning Algorithms. *Neural Computation*, 8 :1341–1390.
- Woodhouse, C. A. and Brown, P. M. (2006). International Tree-ring Data Bank CO572.
- Wunsch, C. (2000). Moon, tides and climate. *Nature*, 405 :743–744.
- Wunsch, C. (2002). What is the thermohaline circulation ? *Science*, 298 :1179–1181.
- Xie, P. and Arkin, P. A. (1997). Global precipitation : A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical outputs. *Bull. Amer. Meteor. Soc.*, 78 :2539–2558.
- Yaeger, S. G. and Robson, J. I. (2017). Recent progress in understanding and predicting Atlantic decadal variability. *Curr. Clim. Change Rep.*, 3 :112–127.
- Yan, X., Zhang, R., and Knutson, T. (2018). Underestimated AMOC variability and implications for AMV and predictability in CMIP models. *Geophys. Res. Lett.*, 45 :4319–4328.
- Yeo, S.-R. and Kim, K.-Y. (2015). Decadal changes in the southern hemisphere sea surface temperature in association with El-Niño-Southern Oscillation and Southern Annular Mode. *Clim. Dyn.*, 45 :3227–3242.

- Young, G. H. F., McCarroll, D., Loader, N. J., Gagen, M., Kirchhefer, A. J., and Demmler, J. C. (2012). Changes in atmospheric circulation and the Arctic Oscillation preserved within a millennial length reconstruction of summer cloud cover from northern Fennoscandia. *Climate Dynamics*, 39 :495–507.
- Zerefos, C. S., Gerogiannis, V. T., Balis, D., Zerefos, S. C., and Kazantzidis, A. (2007). Atmospheric effects of volcanic eruptions as seen by famous artists and depicted in their paintings. *Atmos. Chem. Phys.*, page 16.
- Zhang, K., Wan, H., Liu, X., Ghan, S. J., and Kooperman, G. J. (2014). Technical Note : On the use of nudging for aerosol-climate model intercomparison studies. page 14.
- Zhang, L. and Wang, C. (2013). Multidecadal North Atlantic sea surface temperature and Atlantic meridional overturning circulation variability in CMIP5 historical simulations. *Journ. Geophys. Res. Oc.*, 118 :5772–5791.
- Zhang, P., Linderholm, H. W., Gunnarson, B. E., Björklund, J. A., and Chen, D. (2016). 1200 years of warm-season temperature variability in central Scandinavia inferred from tree-ring density. *Clim. Past.*, 12 :1297–1312.
- Zhang, R. (2008). Coherent surface-subsurface fingerprint of the Atlantic meridional overturning circulation. *Geophys. Res. Lett.*, 35 :L20705.
- Zhang, R. and Delworth, T. L. (2006). Impact of Atlantic multidecadal oscillations on India/Sahel rainfall and Atlantic hurricanes. *Geophys. Res. Lett.*, 33 :L17712.
- Zhang, R., Sutton, R., Danabasoglu, G., Kwon, Y.-O., Marsh, R., Yeager, S. G., Amrhein, D. E., and Little, C. M. (2019a). A review of the role of Atlantic meridional overturning circulation in Atlantic multidecadal variability and associated climate impacts. *Rev. Geophys.*, 57 :316–375.
- Zhang, W., Mei, X., Geng, X., Turner, A. G., and Jin, F.-F. (2019b). A Nonstationary ENSO–NAO Relationship Due to AMO Modulation. *J. Climate*, 32 :33–43.
- Zhang, W., Wang, Z., Stuecker, M. F., Turner, A. G., Jin, F.-F., and Geng, X. (2019c). Impact of ENSO longitudinal position on teleconnections to the NAO. *Clim. Dyn.*, 52 :257–274.
- Zhang, Y. and Yang, Y. (2015). Cross-validation for selecting a model selection procedure. *Journal of Econometrics*, 187 :95–112.
- Zou, H. (2006). The Adaptive Lasso and its Oracle Properties. *Journal of the American Statistical Association*, 101 :1418–1429.
- Zou, H. and Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society*, 67 :301–320.