

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

LES PROBLÈMES DISTINCTS ET POTENTIELS DE BIAIS MULTIVARIÉS ET
D'INCOHÉRENCE PHYSIQUE IMPLIQUÉS DANS L'AJUSTEMENT DE BIAIS
DES SIMULATIONS CLIMATIQUES

MÉMOIRE

PRÉSENTÉ

COMME EXIGENCE PARTIELLE

DE LA MAÎTRISE EN SCIENCES DE L'ATMOSPHÈRE

PAR

MÉGANE ALAVOINE

MAI 2022

UNIVERSITÉ DU QUÉBEC À MONTRÉAL
Service des bibliothèques

Avertissement

La diffusion de ce mémoire se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.04-2020). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

REMERCIEMENTS

De prime abord, je tiens à remercier mon directeur principal de mémoire, Patrick Grenier, pour la confiance qu'il m'a accordée dès le début et tout au long de cette recherche. Je le remercie chaleureusement d'avoir su croire en mes capacités et en notre collaboration. De plus, j'aimerais le remercier de sa disponibilité, de sa bienveillance et de la rigueur dont il a fait preuve tout au long de cette recherche, ce qui a grandement contribué à l'accomplissement du mémoire que vous avez sous les yeux.

J'aimerais également remercier mon co-directeur de mémoire, René Laprise, pour son support depuis le début de ma formation en sciences de l'atmosphère.

De plus, je remercie Ouranos et le programme Mitacs Accélération pour la bourse de maîtrise qui a rendu cette étude possible.

Finalement, je remercie infiniment Thibault pour son soutien indéfectible tout au long de cette aventure.

TABLE DES MATIÈRES

LISTE DES FIGURES.....	v
LISTE DES TABLEAUX.....	x
LISTE DES ABRÉVIATIONS, DES SIGLES ET DES ACRONYMES	xi
LISTE DES SYMBOLES ET DES UNITÉS	xiii
RÉSUMÉ	xv
CHAPITRE I INTRODUCTION.....	16
1.1 Problématique générale de l’ajustement de biais	17
1.1.1 Aspects techniques	17
1.1.1.1 Enjeu de biais	17
1.1.1.2 Enjeu d’incohérence physique.....	23
1.1.2 Aspects épistémologiques.....	25
1.1.2.1 Les débats climatologiques entourant l’ajustement de biais	25
1.1.2.2 Choix terminologiques	26
1.2 Description succincte de l’étude et de ses objectifs	28
1.3 Structure du mémoire	31
CHAPITRE II NOTIONS STATISTIQUES PRÉALABLES	33
2.1 Procédures d’ajustement de biais	33
2.1.1 Transformation quantile-quantile	34
2.1.2 Procédure de brassage des éléments	39
2.1.3 Transformation géométrique par vecteurs propres	41
2.2 Métriques d’évaluation.....	44
2.2.1 Corrélation inter-variables	45
2.2.2 Variabilité interannuelle des moyennes mensuelles	46
2.2.3 Autocorrélation mensuelle de décalage 1	47

CHAPITRE III THE DISTINCT PROBLEMS OF PHYSICAL INCONSISTENCY AND OF MULTIVARIATE BIAS POTENTIALLY INVOLVED IN THE STATISTICAL ADJUSTMENT OF CLIMATE SIMULATIONS	53
3.1 Introduction	56
3.2 Data and methods	60
3.2.1 Data sets.....	60
3.2.2 Nomenclature.....	62
3.2.3 Bias adjustment.....	64
3.2.4 Overarching methodology	67
3.2.4.1 Physical inconsistency.....	67
3.2.4.2 Statistical properties	68
3.3 Results and analysis.....	69
3.3.1 Physical inconsistency	69
3.3.2 Intra-annual cycle	70
3.3.3 Inter-annual variability	71
3.3.4 Inter-variable correlation	72
3.3.5 Lag-1 auto-correlation	73
3.3.6 Climate change signals	74
3.4 Summary with concluding remarks	75
CHAPITRE IV CONCLUSION	99
4.1 Contexte.....	99
4.2 Résumé de la méthodologie.....	101
4.3 Vérifications systématiques.....	103
4.4 Conclusions et recommandations	108
RÉFÉRENCES.....	116

LISTE DES FIGURES

Figure	Page
1.1 Carte des 12 villes à l'étude (Figure 1 de Grenier, 2018).....	32
2.1 a) les centiles de la distribution de la variable T pour la simulation 9 (SIM-09) en rouge et pour le produit de référence CFSR en noir, à San Francisco, pour le 11 août avec une fenêtre mobile de 31 jours (soit du 27 juillet au 26 août), avec l'option de régression 'none' et l'option de technique d'ajustement de biais univariée 'QM-only'. Les symboles carrés montrent les centiles de 1 à 99 par incrément de 2 alors que les extrema de distribution sont associés aux centiles 0 (minimum) et 100 (maximum) ; et b) fonction de transfert (F_{QM}) additive avec interpolation linéaire entre les nœuds (symboles ronds) et extrapolation constante au-delà des nœuds limites (0 et 100).....	50
2.2 Exemple simple de <i>shuffling</i> pour les variables de température (T) et pression (P). Valeurs et rangs associés de T et P a) pour le produit de référence ; b) après ajustement univarié des biais (QM) ; c) après application du <i>shuffling</i> avec T (en gras) comme variable maître ; et d) après application du <i>shuffling</i> avec P (en gras) comme variable maître. Les éléments soulignés dans a) et b) sont les éléments qui servent à construire les tableaux c) et d) (rangs pour le tableau a et valeurs pour le tableau b)..	51
2.3 Exemple à 2 dimensions pour T (en °C) et P (en hPa) de la transformation par vecteurs propres à Mexico City. Les 930 points noirs (fenêtre mobile de 31 jours x 30 années) représentent les valeurs du produit de référence (CFSR) alors que les 930 points représentent les valeurs de la simulation numéro 10 (SIM-10), pour le 1 ^{er} octobre. Les carrés représentent respectivement la moyenne des données pour CFSR (noir) et celle pour SIM-10 (rouge). Les lignes rouges (noires) sont les vecteurs propres de la simulation (CFSR), qui représentent le système de coordonnées $\mathcal{M}(\mathcal{O})$. Les gros points rouge (point $m = m^*$) et bleu (point $t = t^*$) représentent respectivement l'élément 17 (sur 930) de la simulation initiale brute (SIM-10) et du scénario climatique final (point ajusté). La station de Mexico City	

- se situe à 2 571 m d'altitude, ce qui explique des pressions atmosphériques proches de 750 hPa. Il est à noter qu'il y a un écart entre l'altitude assignée dans CFSR (tuile de grille d'environ 31 km de côté incluant la station de Mexico City) et la station elle-même (résolution ponctuelle). Le formalisme mathématique ici utilisé est celui de Hnilica et al. (2017). 52
- 3.1 Diagram of the connections between the raw simulated (with subscripts 'sim') and the bias-adjusted (with either 'uni' or 'multi' subscripts, referred to by the more general 'ba' subscript elsewhere) variables. The role of the corresponding CFSR variables is implicit in this diagram. The variable q_{so90} is post-calculated to be consistent with T_{ba} , P_{ba} and RH_{ba} 85
- 3.2 Diagram of the bias adjustment algorithm used in this study, showing the place of each of the optional decisions. 86
- 3.3 General diagram of the variables' categorization and transformations necessary to prevent bias adjustment from generating physical inconsistency of the out-of-bound type. Vertical blue bands represent the potential extension of the empirical distribution. Here variables are referred to by their conventional GCM or RCM labels, and variables used in this study are highlighted in red ('tas' for T , 'ps' for P , 'huss' for q , and 'hurs' for RH). 87
- 3.4 Two-dimensional counts for q_{ba} and q_{so90} , for one study site (Montreal) and for 9 of the 18 alternative adjustment techniques (TP option is activated). Counts include the ten simulations over the full application period (1981-2100), for a total of 438,000 time steps. Bin size is 0.25 g/kg x 0.25 g/kg. Black lines indicate the 1:1 ratio. On each panel, the root-mean-square error (RMSE) between q_{ba} and q_{so90} is provided as a numerical indication of physical inconsistency (departure from the 1:1 ratio). 88
- 3.5 Frequency (per mil) and amplitude (%) of supersaturation (SS) occurrences in RH_{so90} , for one study site (Miami) and for each of the 18 bias adjustment techniques. Each distribution is built from the SS cases among 438,000 considered time steps (10 simulations over the application period 1981-2100), with the solid line positioned at the SS frequency (x axis) and extending over the full range of obtained SS amplitudes (along the y axis), and the empty (or filled) circle representing the 50th (or 90th) percentile. Lowercase (or uppercase) letters are used to represent the deactivated (or activated) TP option, and the matching regarding the

- dimensionality and grouping options is the following: a/A for QM-only and annual; b/B for QM-shuf and annual; c/C for eig-QM and annual; d/D for QM-only and monthly; e/E for QM-shuf and monthly; f/F for eig-QM and monthly; g/G for QM-only and window; h/H for QM-shuf and window; i/I for eig-QM and monthly. 89
- 3.6 RH average annual cycle over the calibration period (1981-2010), for one study site (Iqaluit) and for 9 of the 18 alternative adjustment techniques (TP option is activated). The black line (corresponding to the reference product CFSR) and the red lines (each corresponding to one of the 10 raw simulations) are the same for all panels. Each blue line is the adjusted version of a raw simulation. Tick marks of the x axis indicate the first day of each month. 90
- 3.7 RH monthly average over the calibration period (1981-2010), for one study site (Iqaluit), for raw simulations (red), adjustment with QM-only (blue), adjustment with QM-shuf (green) and adjustment with eig-QM (purple). Each circle indicates one of the 10 (raw or adjusted) simulations (with superimposition in adjustment cases). Monthly grouping and TP options are activated. Black lines represent RH monthly averages for the reference product CFSR. 91
- 3.8 Standard deviation (σ) of monthly averages over the calibration period (1981-2010), for one study site (Denver), for variables: a) T ; b) P ; c) RH; and d) q . Each circle indicates one of the 10 (raw or adjusted) simulations, with same color code as in Figure 7. Window grouping and TP options are activated. Black lines represent CFSR σ values, and grey boxes delimit the 1st and 99th percentiles from 10,000 bootstrapping re-samplings of the 30 CFSR monthly averages. 92
- 3.9 Spearman's (rank) correlation coefficient (r_{rank}) between RH and T , for all days in each month-of-the-year over the calibration period (1981-2010), for one study site (Mexico City). Each circle indicates one of the 10 (raw or adjusted) simulations, with same color code as in Figure 7. Monthly grouping is activated and TP is deactivated. Black lines represent CFSR r_{rank} values. 93
- 3.10 Average of the 30 intra-month r_{rank} values over the calibration period (1981-2010), for one study site (Mexico City), for pairs of variables: a) RH and T ; b) RH and P ; and c) RH and q . Each circle indicates one of the 10 (raw or adjusted) simulations, with same color code as in Figure 7.

<p>Window grouping and TP options are activated. Black lines represent CFSR average intra-month r_{rank} values, and grey boxes delimit the 1st and 99th percentiles from 10,000 bootstrapping re-samplings of the 30 CFSR intra-month r_{rank} values.</p> <p>3.11 Average of the 30 intra-month lag-1 auto-correlation (AC) values over the calibration period (1981-2010), for one study site (El Paso), for: a) T; b) P; c) RH; and d) q. Each circle indicates one of the 10 (raw or adjusted) simulations, with same color code as in Figure 7. Window grouping and TP options are activated. Black lines represent CFSR average intra-month lag-1 AC values, and grey boxes delimit the 1st and 99th percentiles from 10,000 bootstrapping re-samplings of the 30 CFSR intra-month lag-1 AC values.</p> <p>3.12 RH time series (at 12 UTC each day) for August 1981 at one site (El Paso), for raw SIM-01 and after different adjustments (window grouping and TP options activated in each case).</p> <p>3.13 Relative long-term changes for September average specific humidity after adjustment, $\Delta_{\text{rel}}(q_{\text{ba}})$, in function of corresponding simulated changes, $\Delta_{\text{rel}}(q_{\text{sim}})$. Each panel contains 120 red symbols for adjustment with activated TP, and 120 blue symbols for adjustment without this procedure (12 sites x 10 simulations). Relative changes follow the general form [$\Delta_{\text{rel}}(q) \equiv 100(q(2071-2100) - q(1981-2010)) / q(1981-2010)$], in units of [% (90 yr)⁻¹]. Black lines indicate the 1:1 proportion, and printed root-mean-square differences (RMSD) are calculated over pairs of corresponding $\Delta_{\text{rel}}(q_{\text{ba}})$ and $\Delta_{\text{rel}}(q_{\text{sim}})$ values.</p> <p>3.14 Root-mean-square differences (RMSD) over 120 pairs of corresponding bias-adjusted and raw simulated 1981-2010-to-2071-2100 change (Δ or Δ_{rel}) values (12 sites x 10 simulations), for variables a) T_{ba}; b) P_{ba}; c) RH_{ba}; d) q_{ba}; e) RH_{so90}; and f) q_{so90}; and for all months (panel columns) and techniques (panel rows). Blue-shade upper-left (or red-shade lower-right) triangles correspond to technique without (or with) the trend preservation (TP) procedure. Labels ‘a’, ‘m’ and ‘w’ refer to annual, monthly and window grouping options, respectively. For each panel, RMSD values are presented as the fraction (white for 0 to 0.1, darker shade for 0.9 to 1) of the maximum of the 240 involved values. An empty (or filled) black dot indicates the associated with-TP RMSD value is larger than half its corresponding no-TP value (or larger than the corresponding no-TP value).</p>	<p>94</p> <p>95</p> <p>96</p> <p>97</p> <p>98</p>
---	---

- 4.1 Cycle annuel moyen de la température (T en °C) à Iqaluit. La ligne noire (T_{cfsr}) représente le cycle annuel moyen du produit de référence (CFSR) sur la période de calibration historique (1981-2010). Les 10 lignes bleues (T_{ba} 1981-2010) correspondent à celui des scénarios climatiques passés (sur 1981-2010) alors que les 10 lignes vertes (T_{ba} 2071-2100) représentent le cycle annuel moyen des scénarios climatiques futurs, sur la période 2071-2100. Les scénarios sont produits par l'alternative correspondant aux options 'loess', 'fenêtre mobile' et 'QM-only'. 113
- 4.2 Série temporelle de la température (T en °C) à Iqaluit sur la période d'application (1981-2100). La ligne noire (T_{cfsr}) représente la série temporelle du produit de référence (CFSR) sur la période de calibration historique (1981-2010) seulement (il n'existe évidemment pas de produit de référence futur). Les 10 lignes rouges correspondent aux séries temporelles des 10 simulations climatiques (T_{sim}) alors que les 10 lignes bleues représentent celles des scénarios climatiques (T_{ba}). Le diamant noir représente la moyenne sur 1981-2010 de la série temporelle de CFSR. La barre rouge (bleue) représente l'intervalle, le long de l'axe des y, des moyennes sur 1981-2010 pour les 10 simulations climatiques (scénarios climatiques). Les scénarios sont produits par l'alternative correspondant aux options 'loess', 'fenêtre mobile' et 'QM-only' 114
- 4.3 Statistique de Kolmogorov-Smirnov (D_{KS}) entre d'une part q_{cfsr} (l'humidité spécifique q du produit de référence) et d'autre part q_{sim} (q de la simulation climatique, en rouge), q_{ba} (q directement ajusté, en bleu) ou q_{so90} (q déduit des 3 autres variables, en cyan), pour chacune des 10 simulations. Ce cas est celui de la ville d'Iqaluit pour les scénarios produits par l'alternative univariée 'loess', 'fenêtre mobile' et 'QM-only'. Pour une meilleure appréciation visuelle, les lignes connectent les résultats des différents mois pour une même simulation. 115

LISTE DES TABLEAUX

Tableau	Page
3.1 Specifications for simulations used in this study. Expansions for names of modelling institutes and models can be found at http://www.ametsoc.org/PubsAcronymList	82
3.2 Specifications for study sites.	83
3.3 Synthesis of bias adjustment effects on simulated time series properties for each alternative technique, which is a combination of a dimensionality option (QM-only, QM-shuf or eig-QM), a temporal grouping option (annual, monthly or window), and an optional trend preservation procedure that is activated (TP) or not (-). For the PI problem: dark green indicates respect of the thermodynamic constraint (no PI); and dark red indicates a quasi-guaranteed breaking. For biases: grey indicates that simulated results are roughly kept; dark green stands for realistic results also expected <i>a priori</i> ; light green for prevalent improvement noticed <i>a posteriori</i> ; light red for prevalent deterioration noticed <i>a posteriori</i> ; and dark red for clear deterioration also expected <i>a priori</i> . For long-term changes: grey indicates that simulated changes are roughly kept; and light red stands for frequent strong alteration noticed <i>a posteriori</i> . Note that for QM-shuf combinations, some red cases concern only the shuffled variables (here P_{ba} , RH_{ba} and q_{ba}), whereas results for the master variable (here T_{ba}) are in line with the corresponding QM-only results.	8

LISTE DES ABRÉVIATIONS, DES SIGLES ET DES ACRONYMES

Liste française :

ACP	Analyse en composantes principales
GES	Gaz à effet de serre
HR	Humidité relative
ID	Identifiant
IP	Incohérence physique
MGC	Modèle global du climat
MRC	Modèle régional du climat
PAVICS	Pôle d'analyse et de visualisation de l'information climatique et scientifique
SIM- x	Simulation numéro x
TAB	Technique d'ajustement de biais

Liste anglaise :

AC	Auto-correlation
CFSR	Climate forecast system reanalysis
CMIP5	Coupled model intercomparison project - phase 5
CRCM5	Canadian regional climate model - version 5
dOTC	Dynamical optimal transport correction (algorithm)
GCM	Global climate model
MBCn	Multivariate bias correction with N -dimensional transform (algorithm)
PI	Physical inconsistency
QM	Quantile mapping
RCM	Regional climate model
RCP	Representative concentration pathways
RH	Relative humidity
RMSD	Root-mean-square difference
RMSE	Root-mean-square error
R^2D^2	Rank resampling for distributions and dependences (algorithm)
so90	Sonntag-1990 (fit empirique pour calculer la pression partielle de vapeur d'eau)
SS	Supersaturation
TP	Trend preservation

UTC
WMO

Universal time coordinated
World meteorological organization

LISTE DES SYMBOLES ET DES UNITÉS

D_{KS}	Statistique de Kolmogorov-Smirnov
e_{sat}	Pression partielle saturante de vapeur d'eau
ETJ	Écart thermique journalier
F_{QM}	Fonction de transfert
P	Pression
P_{ba}	Pression après ajustement de biais
P_{cfsr}	Pression du produit de référence CFSR
P_{sim}	Pression de la simulation
q	Humidité spécifique
q_{ba}	Humidité spécifique après ajustement de biais
q_{cfsr}	Humidité spécifique du produit de référence CFSR
q_{sat}	Humidité spécifique saturante
q_{sim}	Humidité spécifique de la simulation
q_{so90}	Humidité spécifique calculée avec le fit empirique so90
RH	Humidité relative
RH_{ba}	Humidité relative après ajustement de biais
RH_{cfsr}	Humidité relative du produit de référence CFSR
RH_{sim}	Humidité relative de la simulation
RH_{so90}	Humidité relative calculée avec le fit empirique so90
r_{rank}	Coefficient de corrélation de Spearman
T	Température
T_{ba}	Température après ajustement de biais
T_{cfsr}	Température du produit de référence CFSR
T_{max}	Température journalière maximale
T_{min}	Température journalière minimale
T_{sim}	Température de la simulation
w	Rapport de mélange
w_{sat}	Rapport de mélange saturant
σ	Déviation standard
Δ	Changement à long-terme (delta)
Δ_{rel}	Changement relatif à long-terme

g	Gramme
hPa	Hectopascal
kg	Kilogramme
m ²	Mètre carré
W	Watt
°C	Degré Celsius
%	Pourcentage
‰	Pour mille

RÉSUMÉ

L'ajustement de biais des simulations climatiques implique plusieurs arguments techniques et épistémologiques dans lesquels la notion d'incohérence physique est souvent évoquée, soit pour rejeter la légitimité de l'ajustement du biais en général, soit pour justifier la nécessité de techniques multivariées sophistiquées. Cependant, cette notion est généralement utilisée sans être définie. Dans ce contexte, l'objectif central de cette étude est de clarifier et d'illustrer la distinction entre incohérence physique et biais multivarié, en étudiant l'effet de l'ajustement de biais sur deux différents types de relations inter-variables, à savoir une contrainte physique censée se maintenir à chaque pas de temps d'une série temporelle et des propriétés statistiques qui émergent avec un biais potentiel sur une échelle de temps climatique. Pour ce faire, 18 alternatives d'ajustement de biais, qui sont obtenues en croisant trois options uni/multivariées avec deux options de gestion des tendances et avec trois options de regroupement temporel, sont appliquées sur 10 simulations climatiques et sur 12 sites à travers l'Amérique du Nord avec CFSR comme produit de référence. Les variables ajustées sont la température, la pression, l'humidité relative et l'humidité spécifique, liées entre elles par une contrainte thermodynamique. L'analyse des résultats suggère d'une part qu'un cas sans équivoque d'incohérence physique potentielle peut être évité avec une technique univariée ou multivariée, si et seulement si la stratégie d'ajustement de biais considère explicitement que la contrainte physique doit être préservée. D'autre part, les résultats suggèrent également que les techniques multivariées sophistiquées peuvent ne pas être des stratégies complètes en elles-mêmes, car elles ne peuvent pas remplacer le codage explicite d'une contrainte physique. Comme objectif supplémentaire, cette étude rapporte des procédures d'ajustement communes avec des effets probables sur diverses propriétés statistiques de base, dans le but de guider les utilisateurs d'informations climatiques dans la détermination de stratégies d'ajustement de biais adéquates pour leurs objectifs de recherche.

Mots clés : simulations climatiques, relations physiques, biais statistiques, techniques d'ajustement multivarié, humidité de l'air.

CHAPITRE I

INTRODUCTION

Le phénomène des changements climatiques s'impose de plus en plus dans notre quotidien. De par sa nature, l'espèce humaine est constamment en quête de rétrospection, pour faire le point sur le présent, mais surtout de projection, pour apprivoiser l'avenir. Par conséquent, il y a une demande grandissante pour des études d'impact, de vulnérabilité et d'adaptation aux changements climatiques (Mote et al., 2011 ; Lourenço et al., 2016). Pour ce faire, des scénarios climatiques, c'est-à-dire des trajectoires climatiques réalistes quant au passé et plausibles quant au futur, sont nécessaires. Dans ce contexte, on utilise souvent des simulations climatiques, qui sont issues de l'intégration d'un modèle numérique du climat basé le plus possible sur la physique du système climatique. Or, ces simulations peuvent comporter des biais dans certaines propriétés statistiques par rapport à un produit de référence, qui est un jeu de données fortement basé sur les observations et sélectionné pour servir de repère climatologique réaliste. Autrement dit, ces simulations peuvent ne pas être réalistes, à certains égards, par rapport à un produit de référence sélectionné. De ce fait, des techniques d'ajustement de biais (TAB) sont développées, qui, comme leur nom l'indique, reflètent l'objectif d'ajuster les biais. Néanmoins, ces techniques peuvent générer d'autres problèmes, notamment celui de l'incohérence physique (IP). La présente étude se penchera sur la distinction entre les problèmes de biais et d'IP souvent soulevés dans la littérature scientifique, ainsi que sur des tentatives de minimisation de ces enjeux.

1.1 Problématique générale de l'ajustement de biais

1.1.1 Aspects techniques

1.1.1.1 Enjeu de biais

Actuellement, les modèles climatiques sont les meilleurs outils sur lesquels se baser pour construire des climats futurs plausibles sur maintes échelles temporelles et spatiales. Il existe une multitude de modèles climatiques, qui diffèrent notamment selon le degré de simplification de la physique, le schéma numérique, la grille, les paramétrages et les champs géophysiques (topographie, types de sol, types de végétation, contours continentaux, bathymétrie, etc.). Entre autres, les modèles globaux du climat (MGC) servent à appréhender les processus en jeu dans la réponse du système climatique aux forçages anthropiques, sur l'entièreté du globe. En effet, les MGC couvrent toute la planète, avec une résolution horizontale allant typiquement de 300 à 100 km. Néanmoins, à cause de leur faible résolution, ils ne peuvent pas décrire avec précision les effets des forçages locaux (par exemple une topographie complexe) qui modulent le climat à des échelles plus fines. Cela peut impliquer, par exemple, un manque de représentation des phénomènes extrêmes comme les vagues de chaleur (2003 en Europe de l'Ouest) ou les tempêtes de verglas (janvier 1998 dans le nord-est de l'Amérique du Nord), ou encore les séquences météorologiques menant à des inondations (1996, 2017 et 2019 au Québec). De ce fait, des modèles régionaux du climat (MRC) couvrant une zone limitée du globe ont été développés depuis une trentaine d'années (Giorgi, 2019), avec des résolutions typiques de 50 à 10 km. D'autre part, l'évolution de la connaissance du système climatique et l'accroissement des capacités de calcul conduisent la communauté scientifique à développer de nouveaux modèles à très haute résolution, allant de 3 à 1 km (Giorgi, 2019). Ce nouveau type de

modèle permet la résolution de la convection (*convection permitting*) et donc des processus de très fine échelle comme ceux que l'on retrouve dans les nuages (Rockel, 2015).

Un scénario climatique, qui correspond à une simulation climatique aux biais ajustés, est le produit d'information le plus utilisé comme intrant dans les études d'impact, de vulnérabilité et d'adaptation aux changements climatiques. L'élaboration d'un tel produit nécessite 3 éléments indispensables : 1) un produit de référence (station météorologique, produit d'interpolation ou produit de réanalyse), qui représente de façon réaliste le climat pendant une longue période historique (typiquement 30 ans) ; 2) une simulation climatique, qui fournit surtout des valeurs de changements plausibles pour le futur ; et 3) une technique d'ajustement de biais (TAB), qui vise à ajuster les biais de la simulation climatique par rapport au produit de référence choisi. De ce fait, les TAB, en ajustant les biais de la simulation climatique pour la rendre plus réaliste quant au passé, élaborent les scénarios climatiques.

Les scénarios climatiques doivent avoir une certaine crédibilité physique afin de pouvoir être jugés plausibles quant au futur. Cela ne signifie pas qu'ils s'avèrent exacts après coup, mais plutôt que la chaîne d'opérations constituant son mode d'obtention soit sensée (ce qui s'évalue par jugement d'expert). Le concept de plausibilité s'inscrit ici dans une logique de projection (plutôt que de prédiction), ce qui englobe inévitablement des incertitudes, qui constituent l'un des concepts les plus importants dans le défi considérable des changements climatiques futurs. En effet, ces incertitudes émanent du fait que différentes simulations climatiques suggèrent diverses valeurs plausibles de changements futurs. Il existe plusieurs raisons pouvant expliquer cela : 1) l'incertitude sur les *Representative Concentration Pathways* (RCP), autrement dit la trajectoire future d'émissions et de concentrations de gaz à effet de serre (GES), utilisés comme forçage dans les modèles climatiques ; 2) la diversité des formulations de la

physique du système climatique dans les modèles numériques, ce qui leur confère des sensibilités diverses pour une trajectoire d'émission fixe ; 3) la variabilité interne des modèles, qui reflète la variabilité naturelle du système climatique réel, et qui fait en sorte qu'une trajectoire de GES et un modèle fixes vont mener à des changements climatiques différents, pour peu que l'on varie les conditions initiales d'une simulation à l'autre (théorie du chaos de Lorenz, 1963) ; et 4) l'existence de biais dans les simulations, qui implique leur ajustement par des TAB, et qui peut mener à une altération des changements futurs, surtout lors de l'application d'une technique sans procédure de préservation de tendance.

Les scénarios climatiques doivent aussi être réalistes quant au passé. À première vue, les simulations climatiques pourraient satisfaire à cette nécessité. Cependant, ces dernières comportent des biais dans certaines propriétés statistiques, par rapport à un produit de référence (Wilcke et al., 2013 ; Maraun, 2016). Par propriété statistique, on entend ici une quantité synthétique généralement exprimée en un seul nombre, qui reflète un comportement d'une série temporelle sur une longue période (typiquement au moins 30 ans), et qui peut évoluer dans le temps. Par exemple, la moyenne, la déviation standard, les corrélations inter-variables et spatiales, l'autocorrélation, etc. sont des propriétés statistiques. Les biais présents dans certaines de ces propriétés occasionnent le fait que les simulations climatiques ne sont pas considérées comme étant réalistes quant au passé, du moins à certains égards, par rapport au produit de référence. Le concept de réalisme s'inscrit dans une logique de précision, qui tend à minimiser les écarts avec le produit de référence. Il existe plusieurs sources pouvant expliquer ces écarts : 1) la différence de représentativité spatiale entre la simulation et le produit de référence, comme un décalage d'échelle, de grille et/ou d'altitude ; 2) les imperfections dans les structures des modèles climatiques, étant donné que la physique représentée ne correspond pas exactement à celle du monde réel (même si elle peut être qualitativement convenable), qui se reflètent inmanquablement dans les simulations

climatiques générées ; 3) les imperfections des produits de référence (station météorologique, produit d'interpolation et produit de réanalyse) (Diaconescu et al., 2017) ; et 4) la désynchronisation entre les variabilités naturelles réelles et simulées à basse fréquence, étant donné que les simulations climatiques issues de conditions initiales légèrement différentes sont désynchronisées les unes les autres ainsi que par rapport à la séquence observée, et que le système climatique a probablement des modes de variabilité de période supérieure à 30 ans (qui est la durée considérée pour établir les normales climatiques).

Les centres de services climatiques aspirent à fournir aux utilisateurs des simulations climatiques ajustées, par l'entremise des TAB. Si la décision d'ajuster les biais présents dans les simulations climatiques est prise, alors cela signifie implicitement que le produit de référence choisi est jugé réaliste sur une durée de 30 ans et que cette durée est suffisante pour que les statistiques accumulées représentent le climat. Par conséquent, l'importance relative des sources de biais numérotées 3 et 4 au paragraphe précédent est implicitement considérée comme faible. Les biais observés entre la simulation climatique et le produit de référence sont alors associés à la différence de représentativité spatiale (source 1), aux imperfections du modèle (source 2) ou à une combinaison de ces deux éléments. Si c'est seulement la différence de représentativité spatiale qui engendre des biais, alors la solution serait une mise à l'échelle (*rescaling*) qui alignerait l'échelle de la simulation sur celle du produit de référence. En revanche, si les biais proviennent uniquement des imperfections du modèle ou s'ils persistent après la mise à l'échelle effectuée pour pallier les biais provenant de la source 1, alors ces biais seront ajustés par une TAB. En pratique, les deux étapes que sont la mise à l'échelle et l'ajustement de biais sont souvent amalgamées en une procédure conjointe que l'on nomme ajustement de biais par abus de langage (Wilcke et al., 2013). Dans cette étude, ces deux procédures vont effectivement être appliquées conjointement,

étant donné que les biais décrits ci-dessus peuvent aussi comporter une composante de disparité d'échelle.

Pour générer les scénarios climatiques, les TAB amalgament les contenus jugés informatifs du produit de référence et de la simulation climatique, au niveau des statistiques sur de longues périodes (Gennaretti et al., 2015). Une même technique peut être exposée sous différentes terminologies et peut aussi présenter différentes versions. Cela implique alors que de nombreuses TAB sont détaillées dans la littérature. Parmi elles, certaines sont communément utilisées : la méthode des analogues (Von Storch, 1999), la régression linéaire multiple (Von Storch, 1999 ; Hay et Clark, 2003) et la transformation quantile-quantile (l'acronyme QM sera utilisé, référant à l'expression anglaise *quantile mapping*) (Wood et al., 2002). Cependant, ces techniques comportent certaines lacunes, comme le traitement des extrêmes (surtout en ce qui concerne la précipitation) ou les fréquences d'occurrence de valeurs égales à 0 (notamment dans les zones de mousson). De plus, en ajustant les statistiques à long terme (biais), les TAB ont le potentiel de générer d'autres problèmes, comme celui de l'IP, qui est l'un des enjeux sur lequel l'accent est mis dans cette étude (voir Section 1.1.1.2).

La TAB utilisée dans une étude dépend habituellement de l'objectif final visé. En effet, plusieurs propriétés statistiques peuvent être ajustées comme la moyenne, la déviation standard, les corrélations inter-variables et spatiales, l'autocorrélation, etc. Si la moyenne et la déviation standard sont les seules propriétés qui sont importantes à ajuster pour un objectif donné, alors la technique utilisée pourra être en mode univarié puisqu'elle ajustera uniquement les biais des distributions (Trasher et al., 2012 ; Wilcke et al., 2013 ; Rocheta et al., 2014). Cela signifie alors que la simulation est ajustée indépendamment pour chaque variable et chaque emplacement donnés, pour correspondre aux distributions du produit de référence choisi (Cannon, 2016), comme le fait la transformation univariée quantile-quantile. Cette dernière est appliquée à un

grand nombre de variables, sous différentes variantes et noms (Themeßl et al., 2012 ; Trasher et al., 2012 ; Wilcke et al., 2013 ; Rocheta et al., 2014 ; Lehtonen et al., 2016), ce qui en fait sans doute la TAB la plus populaire. Le principe central de cette transformation, ainsi que ses nombreuses options et paramètres, seront expliquées dans le Chapitre 2.

D'autre part, les TAB multivariées ajusteront les biais corrélationnels (corrélations inter-variables et/ou spatiales) entre au moins deux variables (il peut en fait s'agir de différentes variables physiques en un même emplacement, de la même variable physique à différents emplacements ou de la combinaison de ces deux possibilités mathématiques) en plus des biais distributionnels. Ces dernières techniques connaissent un engouement dans la mesure où elles permettraient d'ajuster plus de propriétés statistiques, ce qui est cohérent avec le souhait d'avoir des scénarios climatiques toujours plus réalistes quant au passé. À cet effet, beaucoup de techniques multivariées sont développées et promues dans la littérature scientifique. Parmi elles, les techniques de type « successive conditionnelle » traitent une variable en mode univarié avant qu'une deuxième variable ne soit conditionnellement corrigée sur cette première variable (Piani et Haerter, 2012 ; Gennaretti et al., 2015). D'autre part, les techniques multivariées de type « marginale / dépendance » sont plus complexes car elles traitent chaque variable séparément en mode univarié pour traiter les marginales, puis appliquent une procédure d'une autre nature pour traiter les dépendances (corrélations inter-variables et/ou spatiales). Les procédures proposées dans la littérature pour gérer la dépendance impliquent par exemple le brassage de l'ordre des éléments dans les séries temporelles (*shuffling*) (Vrac et Friederichs, 2015 ; Vrac, 2018) ou le recours à une transformation par vecteurs propres (Hnilica et al., 2017) (plus de détails seront fournis dans la Section 2.1).

1.1.1.2 Enjeu d'incohérence physique

Les TAB sont principalement développées pour minimiser, voire éliminer, les biais des simulations climatiques (par rapport à un produit de référence choisi). Hélas, en voulant résoudre ce problème, elles peuvent en générer d'autres. Parmi eux, la destruction de la cohérence physique, autrement dit la génération d'incohérence physique (IP). En effet, ces techniques sont des manipulations statistiques non basées sur la physique, qui visent l'ajustement des propriétés des distributions ; elles sont donc insensibles face à la cohérence physique attendue à chaque pas de temps. Dans la littérature scientifique, la notion d'IP est souvent utilisée dans des argumentaires à l'encontre ou en défense de l'ajustement de biais, mais sans être définie. Grenier (2018) propose de clarifier le concept en l'associant aux relations attendues à chaque pas de temps, plutôt qu'aux propriétés statistiques émergeant sur le long terme. L'IP peut alors être classée en deux types. L'IP de type 1 se rapporte à l'attribution d'une valeur impossible à une variable, comme une précipitation inférieure à 0 ou une humidité relative supérieure à 100 %. Quant à elle, l'IP de type 2 se réfère au bris d'un lien unissant plusieurs variables, à l'instar d'un écart thermique journalier négatif ($ETJ = T_{\max} - T_{\min}$, avec T_{\max} la température journalière maximale et T_{\min} la température journalière minimale). Dans la présente étude, la notion d'IP, qui est examinée indépendamment à chaque pas de temps, se rapportera toujours à cette même définition. De ce fait, l'altération des propriétés statistiques émergeant à long terme (moyenne, déviation standard, corrélations inter-variables et/ou spatiales, autocorrélation, etc.) s'inscrit dans un enjeu de biais, et non d'IP. Par ailleurs, il est à noter que la génération d'IP demeure inacceptable en toutes circonstances, considérant qu'une IP peut revenir à contredire la définition même des variables (IP de type 1), comme lorsque $T_{\min} > T_{\max}$ pour une journée, ou encore à rejeter des éléments de théorie physique bien reconnus et utilisés pour la construction même des modèles climatiques (IP de type 2), comme le lien

thermodynamique existant entre la température (T), la pression (P), l'humidité relative (acronyme HR - symbole RH) et l'humidité spécifique (q).

La génération d'IP par les TAB univariées est relativement bien documentée dans la littérature, contrairement aux techniques multivariées. Par exemple, Agbazo et Grenier (2019) et Trasher et al. (2012) se penchent sur la production d'IP entre T_{\min} et T_{\max} alors que l'étude de Grenier (2018) explore la génération d'IP entre l'HR et ses variables connexes que sont T , P et q . En outre, Rocheta et al. (2014) étudient la perte de cohérence physique entre des variables dynamiques comme les vents méridional et zonal ou la pression et la température potentielle. Bien heureusement, il existe des suggestions pour éviter la génération d'IP en forçant la TAB à respecter minimalement la physique en : 1) calibrant les paramètres et options de la technique pour minimiser la fréquence d'occurrence des valeurs impossibles ; et/ou 2) en plafonnant des seuils réalistes quant à la variable (comme ramener une HR de 110 % à 100 % ou une précipitation négative à 0) ; et/ou 3) en déduisant après coup un certain nombre de variables de celles qui sont directement ajustées, dans le cas où un lien mathématique existe. Par exemple, Grenier (2018) suggère que la destruction de la cohérence physique peut être évitée en déduisant q des 3 autres variables (T , P et RH). Pour l'écart thermique journalier, il est conseillé de déduire T_{\min} de T_{\max} et de l'ETJ (Trasher et al., 2012 ; Agbazo et Grenier, 2019). Néanmoins, pour les variables déduites, un biais résiduel (perte d'équivalence statistique) avec le produit de référence est souvent le prix à payer pour éviter de générer des IP, même si celui-ci est réduit par rapport au biais initial (Trasher et al., 2012 ; Grenier, 2018). Dans le cas des TAB multivariées, il n'existe pas d'étude décrivant à quel point ces dernières peuvent dégrader la cohérence physique, et c'est notamment pourquoi ce mémoire se concentre sur l'impact de l'ajustement multivarié des biais sur la cohérence physique.

1.1.2 Aspects épistémologiques

1.1.2.1 Les débats climatologiques entourant l'ajustement de biais

En pratique, les TAB sont souvent appliquées de concert avec les techniques de mise à l'échelle (*rescaling*). De ce fait, ces deux types d'outils sont susceptibles d'être amalgamés en une procédure conjointe (Wilcke et al., 2013), ce qui peut alimenter une certaine confusion. Par conséquent, l'ajustement de biais est souvent perçu comme une approche alternative ou rivale à la modélisation régionale. Cependant, ces deux types de procédés ne visent pas les mêmes objectifs. Les techniques de mise à l'échelle sont appliquées pour aligner l'échelle de la simulation climatique sur celle du produit de référence et ainsi éliminer la composante des différences due aux écarts de représentativité spatiale. D'autre part, les TAB servent à ajuster la composante des différences ne provenant pas des écarts de représentativité spatiale, mais plutôt des imperfections des modèles. En somme, ces dernières techniques devraient être vues plutôt comme complémentaires étant donné que les MRC ne règlent pas nécessairement le problème de biais (Giorgi, 2019), et qu'une étape d'ajustement de biais peut suivre l'étape de mise à l'échelle dynamique.

Les TAB sont souvent contestées de par leur nature non physique. En effet, elles sont plutôt basées sur la statistique, ce qui implique qu'elles ne considèrent pas les lois de la physique. Selon Ehret et al. (2012), cette violation des équations physiques (conservation de la masse, de l'énergie, de l'eau, de la quantité de mouvement, etc.) altère alors les relations entre les variables, ce qui ne serait pas le cas dans les simulations générées par les modèles climatiques, étant donné que ces derniers respecteraient les lois de la physique par construction. De plus, Ehret et al. (2012) expose le fait que les TAB négligent largement les mécanismes de rétroaction

(*feedbacks*). Ces éléments contribuent à la pensée que ces dites techniques ne fournissent pas de justification physique satisfaisante, ce qui entraîne une dissimulation et non une réduction de la plage d'incertitude des simulations climatiques (Ehret et al., 2012).

1.1.2.2 Choix terminologiques

La littérature scientifique comporte de nombreux débats au sujet des biais et de l'IP. Il existe divers usages pour cohérence / incohérence physique, mais ces termes ne sont que rarement explicitement définis (quoiqu'on puisse généralement deviner une définition implicite). Plusieurs auteurs définissent implicitement l'IP comme l'attribution d'une valeur au-delà des limites physiques inférieures et supérieures (valeur non physique) à une variable (Quintana Seguí et al., 2010 ; Hempel et al., 2013 ; Wilcke et al., 2013 ; Ruane et al., 2015 ; Lange, 2019), comme une humidité relative (HR) supérieure à 100 % ou une précipitation négative. Aussi, l'IP peut être identifiée lors du bris d'une relation physique unissant plusieurs variables (Rocheta et al., 2014 ; Grenier, 2018 ; Agbazo et Grenier, 2019), comme $T_{\min} > T_{\max}$. Par ailleurs, certains auteurs établissent l'IP comme un manque de cohérence entre les distributions conjointes de différents types de données (Su et al., 2016 ; Frank et al., 2018). Autrement dit, ils définissent l'IP par les différences observées entre deux sets de données, ce que nous avons établi dans cette étude comme étant des biais. De toute évidence, les notions de cohérence et d'incohérence physique sont souvent associées au concept de biais multivariés, toutefois, ces deux enjeux sont bien distincts dans cette étude, comme expliqué dans la Section 1.1.1.

Les débats exposés dans la littérature sont nourris, entre autres, par de la confusion terminologique et conceptuelle. Ils pourraient bénéficier d'un travail de clarification

des notions fondamentales que sont les biais et l'IP. Les choix terminologiques proposés dans ce mémoire peuvent être mis en parallèle avec ceux des critères d'évaluation ou de comparaison de théories scientifiques utilisés en philosophie des sciences, notamment par Kuhn (1977). Selon cet auteur, il existe 5 critères standards pour évaluer la conformité d'une théorie scientifique : 1) la précision ; 2) la cohérence ; 3) la portée globale, qui doit s'étendre bien au-delà des observations, lois ou sous-théories particulières élaborées pour l'expliquer ; 4) la simplicité, qui permet de mettre de l'ordre dans des phénomènes qui, en son absence, seraient isolés individuellement et, en tant qu'ensemble, confus ; 5) les retombées scientifiques en termes de nouveaux résultats de recherche. Deux de ces critères, à savoir la précision et la cohérence, vont nous aider à appuyer et à clarifier les choix terminologiques effectués dans cette étude.

Pour Kuhn (1977), la précision d'une théorie réside dans le fait que les conséquences déductibles de cette même théorie doivent être en accord démontré avec le résultat d'expériences et d'observations existantes. Dans notre cas, le modèle climatique représente la théorie scientifique et les simulations représentent les conséquences déductibles. En ce sens, la notion de biais telle qu'exposée dans cette étude se rapporte à cette définition puisque les propriétés statistiques simulées doivent être en accord avec les observations (à un bémol près, soit l'ampleur de la variabilité naturelle même sur 30 ans, ce qui est souvent négligé dans les services climatiques opérationnels). Le deuxième critère de la conformité d'une théorie scientifique selon Kuhn (1977) est la cohérence. Selon lui, une théorie doit être cohérente non seulement en interne (absence de contradiction avec elle-même), mais aussi avec d'autres théories dominantes, et donc acceptées, de son époque. Dans notre cas, les modèles, qui représentent le fonctionnement du système climatique, doivent être cohérents avec la physique dominante de notre époque, c'est-à-dire avec la loi des gaz parfaits, la thermodynamique, les équations de Navier-Stokes, etc. De ce fait, la définition de

cohérence physique telle qu'utilisée dans ce mémoire est appuyée par le texte de Kuhn (1977).

Il semble important de présenter un exemple historique illustrant le rôle philosophique des critères exposés par Kuhn (1977). Depuis toujours, la majeure partie du monde occidental pensait que le système astronomique était géocentrique, c'est-à-dire que la Terre est au centre de l'Univers et que le Soleil tourne autour de celle-ci. Cette proposition faisait partie de la théorie astronomique de l'astronome Ptolémée. Si l'on s'appuie sur le critère de cohérence de Kuhn (1977), alors le modèle astronomique de Ptolémée était physiquement cohérent à son époque, car il était en accord avec la physique dominante de cette époque, celle d'Aristote, qui requérait le géocentrisme. Cependant, au XVI^{ème} siècle, un autre astronome nommé Copernic apporta une nouvelle théorie selon laquelle le système serait héliocentrique, autrement dit que le Soleil est au centre de l'Univers et que les autres planètes, dont la Terre, tournent autour de ce dernier. La révolution copernicienne qui en résulta fut longue et difficile, entre autres parce que le critère de cohérence favorisait le géocentrisme. C'est plutôt un autre critère de Kuhn (1977), soit celui de la simplicité, qui a initié l'acceptation progressive du système de Copernic. Cependant, au fur et à mesure que les éléments de physique de Galilée, Descartes et Newton remplaçaient ceux d'Aristote, le critère de cohérence défavorisait de moins en moins l'héliocentrisme, et le système de Copernic put s'imposer, avec des modifications ultérieures comme le remplacement des orbites circulaires par des elliptiques.

1.2 Description succincte de l'étude et de ses objectifs

L'étude effectuée dans le cadre de ce mémoire s'apparente à celle réalisée par Grenier (2018), dans la mesure où elle s'applique aux mêmes emplacements, mêmes variables,

mêmes simulations et même produit de référence. Néanmoins, une valeur ajoutée est apportée puisque les analyses sont étendues à des TAB multivariées, alors que son étude était concentrée sur une seule technique univariée. Comme le montre la Figure 1.1, les 12 villes investiguées sont concentrées sur l'Amérique du Nord (voir Tableau 3.2). Malgré tout, elles représentent un large éventail de climats regroupant un climat tropical (Miami), des climats secs (El Paso et Denver), tempérés (San Francisco, Mexico City, New Orleans, Vancouver et Saint Louis), continentaux (Montréal, Saint John's et Yellowknife) et un climat polaire (Iqaluit), selon la classification de Köppen-Geiger (Peel et al., 2007). Les 4 variables impliquées dans cette étude sont la température, la pression, l'humidité spécifique et l'humidité relative, qui sont reliées entre elles par une relation thermodynamique (voir équations 3.3 à 3.5). Un ensemble de 10 simulations climatiques (voir Tableau 3.2), issues de la phase 5 du projet d'inter-comparaison des modèles couplés (CMIP5), est assujéti à l'ajustement de biais. La plupart de ces simulations proviennent de MGC (SIM-01 à SIM-09), et l'une provient d'un MRC (SIM-10). Le produit de référence utilisé dans cette étude est le *Climate Forecast System Reanalysis* (CFSR).

Cette présente étude est de nature méthodologique et a pour vocation de tirer des conclusions générales non limitées à ce mémoire. À cet effet, beaucoup de situations différentes vont être analysées afin d'offrir une perspective sur le sort de la cohérence physique dans le contexte de l'ajustement de biais. Conséquemment, une grande difficulté est imposée, soit celle de croiser 3 options différentes pour notre algorithme d'ajustement de biais (plus de détails sont disponibles dans la Section 3.2.3) : 1) une option pour le groupement des jours de l'année ('annuel', 'mensuel' ou 'fenêtre mobile') pour réaliser l'ajustement de biais ; 2) une option de préservation ou non de la tendance à long terme ('loess' ou 'none') de la simulation climatique ; et 3) une option de TAB qui regroupe une technique univariée (ci-après nommée 'QM-only') et deux techniques multivariées (désignées par 'QM-shuf' et 'eig-QM'). Au total, 18

alternatives (3 x 2 x 3) d'ajustement de biais seront appliquées, qui sont des combinaisons de procédures reflétant des choix typiquement effectués par les praticiens de l'ajustement de biais dans les centres de services climatiques.

Les objectifs principaux de cette étude sont ici explicités. Le premier est de clarifier et de distinguer les enjeux de biais et de cohérence / incohérence physique, étant donné les débats présents dans la littérature scientifique (voir Section 1.1.2.2). Dans un deuxième temps, ce mémoire vise à développer et à recommander des stratégies d'implantation d'une TAB qui garantit la conservation de la cohérence physique inter-variables tout en minimisant les biais résiduels avec le produit de référence. Toujours dans une volonté de distinguer les notions de biais et d'IP, les objectifs spécifiques sont séparés en deux catégories, soit ceux liés aux relations physiques qui concernent l'IP, et ceux liés aux relations statistiques qui concernent les biais. Le premier objectif spécifique en termes d'enjeu physique est de vérifier l'impact des 18 alternatives d'ajustement de biais sur le lien physique unissant les 4 variables. En pratique, il s'agira notamment d'investiguer si cette relation est respectée (pas de génération d'IP inter-variables) ou non (génération d'IP inter-variables), après l'application de l'ajustement de biais. Il est essentiel de mentionner qu'initialement le produit de référence et les simulations climatiques sont biaisés mais en cohérence avec la relation thermodynamique unissant T , P , q et RH, car celle-ci fait partie des éléments physiques pertinents qui sont précisément ciblés pour être prescrits dans les modèles climatiques. Le deuxième objectif spécifique concernant l'enjeu physique est d'investiguer les impacts en termes d'IP de la déduction d'une variable des 3 autres (par exemple explorer si une HR déduite peut être supérieure à 100%). À propos de l'enjeu statistique, l'objectif spécifique est d'investiguer la plus-value potentielle de l'utilisation d'une TAB multivariée par rapport à une technique univariée en analysant des propriétés statistiques comme le cycle annuel, la corrélation inter-variables, l'autocorrélation de décalage 1 et la tendance à long terme.

1.3 Structure du mémoire

La suite de ce mémoire est organisée comme suit. Le Chapitre 2 est un chapitre préliminaire, en français, composé d'éléments techniques aidant à la compréhension et la distinction de notions clés utilisées dans le Chapitre 3. La Section 2.1 détaille les procédures principales imbriquées dans les techniques d'ajustement de biais alors que la Section 2.2 précise les métriques statistiques utilisées subséquentement pour décrire les résultats d'application de cesdites techniques. Le Chapitre 3 contient un article scientifique rédigé en anglais. La Section 3.1 présente le contexte scientifique ainsi que les motivations de cette étude, alors que les éléments méthodologiques, à savoir les données utilisées ainsi que les méthodes d'investigation, sont détaillés dans la Section 3.2. Les résultats sont exposés et analysés dans la Section 3.3, alors que la Section 3.4 résume l'étude, tout en apportant des remarques supplémentaires importantes. Par la suite, les principales conclusions de cette étude sont énumérées et résumées en français dans le Chapitre 4. Le mémoire se termine avec la liste des références employées pour la rédaction de ce mémoire.



Figure 1.1 : Carte des 12 villes à l'étude (Figure 1 de Grenier, 2018).

CHAPITRE II

NOTIONS STATISTIQUES PRÉALABLES

Le projet de recherche présenté dans ce mémoire a impliqué le recours à plusieurs notions statistiques relativement complexes, qui ne peuvent pas toutes être détaillées dans l'article scientifique formant le Chapitre 3. Le but de ce présent chapitre est de préparer la lecture de ce dernier, en introduisant et en distinguant certaines notions statistiques clés. Celles-ci sont regroupées en 2 catégories, à savoir d'une part les procédures principales imbriquées dans les techniques d'ajustement de biais (TAB), et d'autre part, les métriques statistiques utilisées subséquentement pour décrire les résultats d'application de ces dites techniques.

2.1 Procédures d'ajustement de biais

Les TAB sont appliquées à une simulation climatique dans le but de supprimer, ou au moins de minimiser, les biais initiaux présents dans celle-ci par rapport à un produit de référence sélectionné. La TAB utilisée dans une étude dépend généralement de l'objectif final visé, qui va déterminer les propriétés statistiques à ajuster. En effet, plusieurs propriétés statistiques peuvent être ajustées, comme la moyenne, la déviation standard, les corrélations inter-variables et/ou spatiales, l'autocorrélation, etc. Lorsque les propriétés statistiques à ajuster concernent seulement les distributions, alors il est

de mise d'utiliser une technique univariée, qui ajuste les propriétés distributionnelles indépendamment pour chaque variable et chaque emplacement donnés. En revanche, lorsque l'objectif de recherche requiert d'ajuster de surcroît les corrélations inter-variables, alors une technique multivariée doit être utilisée, liant deux variables ou plus (il peut s'agir de variables physiques différentes à un même emplacement, d'une même variable physique en différents emplacements, ou de la combinaison de ces deux possibilités mathématiques). Dans l'article scientifique formant le cœur de ce mémoire (voir Chapitre 3), l'un des objectifs consiste à comparer l'impact de différentes procédures mathématico-numériques couramment introduites dans les TAB, avec une grande importance accordée à la distinction univarié - multivarié. Ainsi, les trois procédures suivantes sont ici préalablement décrites : 1) la transformation quantile-quantile qui est impliquée dans toutes les méthodes afin d'ajuster les propriétés univariées ; 2) le brassage des éléments (*shuffling*) dans la série temporelle, qui est une procédure multivariée généralement utilisée en conjonction avec la transformation quantile-quantile ; et 3) le recours aux vecteurs propres (*eigenvectors*), une autre procédure de nature multivariée généralement utilisée en conjonction avec la transformation quantile-quantile.

2.1.1 Transformation quantile-quantile

La transformation quantile-quantile, ou *quantile mapping* (QM) en anglais, est souvent présentée comme la TAB la plus populaire, étant appliquée à un grand nombre de variables, sous différentes variantes et différents noms (Thiemeßl et al., 2012 ; Trasher et al., 2012 ; Wilcke et al., 2013 ; Rocheta et al., 2014 ; Lehtonen et al., 2016). Cette notoriété dans la communauté scientifique repose notamment sur le succès de cette transformation quand le but d'une étude est d'ajuster une seule variable (Fang et al., 2015) ainsi que de son accessibilité sur le plan informatique (Piani et al., 2010 ;

Themeßl et al., 2012). Le principe général de la transformation QM se décline en 2 grandes étapes conceptuellement simples. D’abord, une fonction de transfert (F_{QM}) liant les quantiles d’une distribution simulée aux quantiles correspondants d’un produit de référence sélectionné est construite pour une période historique dite de calibration (1981-2010 dans ce mémoire). Dans un deuxième temps, la F_{QM} est appliquée à la distribution simulée sur une période dite d’application et qui comprend généralement un segment futur (1981-2100 ici). Malgré la simplicité apparente du principe général de la transformation QM, celle-ci cache des difficultés liées à un certain nombre de choix de paramètres et à des options de procédures auxiliaires, comme nous allons le voir dans le reste de cette section.

Par exemple, la transformation QM peut s’appliquer de façon paramétrique, c’est-à-dire que les distributions impliquées et/ou les fonctions de transfert (F_{QM}) peuvent être résumées à de petits nombres de paramètres (Piani et al., 2010 ; Piani et Haerter, 2012). Elle peut aussi s’appliquer de manière non-paramétrique, comme dans les études de Grenier (2018) et Wood et al. (2002), mais le principe central demeure le même dans les deux cas. Il convient d’ajouter que l’approche dite non-paramétrique est en fait plutôt « peu paramétrique », car les F_{QM} sont établies à un nombre de nœuds (quantiles) généralement plus petit que le nombre d’éléments entrant dans les distributions, et ces nœuds peuvent être interprétés comme des paramètres. Par exemple, si l’application de QM porte sur les jours de janvier pour une période de 30 années, cela implique des distributions de 930 éléments (31 jours dans ce mois x 30 années), alors que la F_{QM} n’est définie qu’à 52 nœuds (les quantiles 0, 1, 3, 5, ..., 95, 97, 99 et 100). Le nombre de nœuds est lui-même un paramètre, avec par exemple un choix de 50 dans l’étude de Boé et al. (2007) et de 52 dans l’étude de Grenier (2018). Ces choix peuvent d’ailleurs être déterminants pour les extrêmes (notamment en-deçà du 5^{ème} et au-delà du 95^{ème} centiles), mais ces derniers reçoivent rarement un traitement adapté, ce qui constitue une limite de maintes TAB.

Un autre exemple de choix lié au QM est la façon de couvrir l'année. En effet, on peut ajuster les biais en utilisant une seule et même F_{QM} pour l'année complète (Reiter et al., 2016), 4 F_{QM} soit une pour chaque saison (Maraun, 2016), 12 F_{QM} c'est-à-dire une pour chaque mois (Haerter et al., 2011) ou 365 F_{QM} soit une pour chaque jour (Themebfl et al., 2012). Dans ce dernier cas, il est indispensable de sélectionner une largeur de fenêtre mobile qui va déterminer le nombre de jours avant et après le jour de l'année à ajuster pour pouvoir définir la fonction de transfert. Des valeurs typiques de largeur de la fenêtre mobile sont : 15 jours (Abatzoglou et Brown, 2012), 31 jours (Trasher et al., 2012 ; Grenier, 2018), 41 jours (Gennaretti et al., 2015) ou 91 jours (Rajczak et al., 2016). Par exemple, si l'on veut ajuster le jour du 11 août (223^{ème} jour de l'année) avec une fenêtre mobile de 31 jours, alors la sélection des jours se ferait du 27 juillet (208^{ème} jour de l'année) au 26 août (238^{ème} jour de l'année), soit 15 jours avant le 11 août et 15 jours après cette même date. Dans la littérature, il est commun de rapporter que la F_{QM} spécifique à chaque jour de l'année constitue un avantage par rapport à une F_{QM} spécifique à chaque saison ou chaque mois puisque cette option permet d'éviter les discontinuités climatologiques aux transitions de mois en mois ou de saison en saison (Gennaretti et al., 2015). Les différents choix possibles concernant le nombre de F_{QM} sont caractérisés par l'option de groupement des jours de cette étude, soit 'annuel' pour une seule F_{QM} , 'mensuel' pour 12 F_{QM} ou 'fenêtre mobile' de 31 jours pour 365 F_{QM} .

Une autre procédure auxiliaire au QM peut découler du choix de forcer ou non la préservation des tendances à long terme, qui proviennent de la simulation climatique. En effet, on peut appliquer le QM directement sur les valeurs simulées « complètes », laissant ainsi la tendance à long terme simulée libre d'évoluer à la guise des F_{QM} , ce qui pourrait entraîner leur modification. En réponse à cela, plusieurs auteurs isolent la tendance et ne travaillent que sur les résidus autour de cette même tendance, avant de réintégrer cette dernière aux résultats obtenus après l'ajustement de biais (Hempel et al., 2013, Agbazo et Grenier, 2019). Dans ce cas, plusieurs types de régression peuvent

être choisis, comme une régression linéaire ou quadratique, mais dans cette étude-ci, le choix s'est porté sur la régression non-paramétrique 'loess', expliquée par Cleveland (1979). Dans ce mémoire, la préservation ou non des tendances à long terme va être explorée à l'aide des deux options de régression suivante : 1) option 'none', ce qui signifie que la tendance à long terme n'est pas préservée et que le QM est appliqué sur les valeurs simulées « complètes » ; et 2) option 'loess' qui implique un ajustement de biais directement sur les résidus des valeurs simulées (la tendance est rajoutée après l'ajustement de biais). Encore une fois, il existe beaucoup de débats dans la littérature concernant la procédure de préservation (ou non) de la tendance, notamment car, selon certains auteurs, l'altération de la tendance n'est qu'une conséquence normale du caractère non-constant de la F_{QM} à travers les centiles (Gobiet et al., 2015).

Enfin, la nature additive ou multiplicative des fonctions de transfert constitue un autre choix à faire avec l'utilisation du QM, choix propre à chaque variable. Dans la littérature, il n'y a pas d'homogénéité à cet égard, mais les variables sont souvent séparées en 2 catégories. D'une part, il y a les variables bornées par zéro, comme le cumul quotidien de précipitation, pour lesquelles une F_{QM} additive pourrait générer des valeurs négatives. En effet, une F_{QM} additive est construite en soustrayant les centiles de la simulation climatique à ceux du produit de référence sur la période de calibration. Pour éviter ces situations vides de sens physique, des F_{QM} multiplicatives, construites en divisant les centiles de la simulation à ceux du produit de référence (toujours sur la période de calibration), sont alors utilisées. D'autre part, il y a les autres variables, qui peuvent souvent s'accommoder d'une F_{QM} additive ou multiplicative, même si le cas additif est souvent choisi par défaut. Dans ce mémoire, un formalisme plus systématique a été adopté, en classifiant les variables en 4 catégories selon leurs types de bornes (voir Figure 3.3) et en procédant à des transformations mathématiques permettant d'utiliser seulement des F_{QM} additives. Les catégories sont les suivantes : 1) la distribution est soit non-bornée, soit déconnectée de sa ou ses borne(s) ; 2) la

distribution est connectée à une borne fixe ; 3) la distribution est connectée à deux bornes fixes ; ou 4) la distribution est connectée à une borne fixe et une borne mobile. Les variables de la catégorie 1, comme la température, la pression, le vent zonal et méridional, sont traitées avec une F_{QM} additive sans ajout de procédures connexes. En revanche, les variables des 3 autres catégories sont habituellement traitées avec des F_{QM} multiplicatives. Toutefois, il existe certains procédés pour contourner le problème du mode multiplicatif et rendre la distribution déconnectée des deux bornes, c'est-à-dire traiter ces variables comme des variables de catégorie 1 (et donc pouvoir opérer en mode additif). Entre autres, l'utilisation de la fonction logarithme pour les variables de catégorie 2, comme la précipitation ou l'écart thermique journalier, et l'utilisation de la fonction logit pour les variables de catégorie 3, comme l'humidité relative, l'humidité spécifique ou la couverture nuageuse :

$$\text{logit}(x) = \ln\left(\frac{x}{1-x}\right) \text{ où } x \in]0;1[\quad (2.1)$$

Les 4 variables impliquées dans ce mémoire, soit la température, la pression, l'humidité spécifique et l'humidité relative, sont alors toutes traitées en mode additif, spontanément pour T et P qui sont des variables de catégorie 1, et à l'aide de la fonction logit pour q et RH qui sont de catégorie 3. Cette catégorisation des variables selon leurs bornes physiques nous permet une meilleure harmonisation du traitement, puisque tout est traité en mode additif, ainsi que d'éviter certaines formes d'incohérence physique (IP de type 1) par construction, puisque l'HR ne peut alors pas dépasser son seuil de 100% (plus de détails sont fournis dans la Section 3.2.4.1).

La Figure 2.1 illustre un exemple de construction de la F_{QM} en mode additif pour les températures quotidiennes (variable T) à 12 UTC, à San Francisco, pour un jour donné (le 11 août) avec l'option 'fenêtre mobile' de 31 jours. Pour ce cas spécifique, on voit que l'atteinte de l'équivalence statistique (distributionnelle) entre la simulation et le

produit de référence impliquerait que les plus faibles valeurs simulées soient ajustées par un facteur de l'ordre de -7°C à -8°C et que les plus fortes valeurs simulées soient ajustées par un facteur de l'ordre de -11°C à -12°C , ce qui démontre l'intérêt de décomposer l'ajustement par quantiles.

2.1.2 Procédure de brassage des éléments

Le brassage temporel des éléments, ou *shuffling*, est une procédure multivariée d'ajustement des biais dans les corrélations inter-variables, utilisée suite à l'ajustement des biais distributionnels (en mode univarié). Dans la présente étude, le recours au *shuffling* suit toujours le QM pour les variables individuelles. Ce type d'ajustement multivarié est classé par Vrac (2018) dans la catégorie « marginale / dépendance », le mot « marginale » référant en statistiques aux distributions considérées isolément et le mot « dépendance » faisant référence aux propriétés statistiques liant différentes distributions. Le principe central du *shuffling* est de, comme son nom l'indique, brasser l'ordre des éléments dans les séries temporelles dans le but d'ajuster les corrélations inter-variables de la simulation sur celles du produit de référence choisi. Cette procédure multivariée est utilisée et décrite dans plusieurs études comme celles de Vrac et Friederichs (2015) et Vrac (2018). Pour utiliser cette procédure, il faut choisir l'une des variables comme étant « maître », c'est-à-dire que celle-ci demeure intouchée après l'application du QM, et l'on brasse chacune des autres variables de manière à ce que leur corrélation avec la variable maître adopte la valeur de corrélation entre les variables correspondantes du produit de référence. Numériquement, le rang de la variable maître dans le produit de référence est donc le point de départ pour mélanger les autres variables (des détails sont donnés dans la suite de cette section).

Le recours au *shuffling* implique deux grandes contraintes. D'abord, il faut que la période d'application de la TAB soit de la même durée que celle de la période de calibration (30 ans dans la présente étude), puisque les rangs des éléments (pas de temps) des variables du produit de référence sont directement pris pour dicter le brassage des éléments des séries temporelles des variables non-maîtres. De ce fait, il a fallu séparer la période d'application de 120 ans (de 1981 à 2100) en 4 périodes de 30 ans, soit 1981-2010, 2011-2040, 2041-2070 et 2071-2100 pour pouvoir appliquer la procédure de brassage des éléments (ce n'était pas nécessaire pour l'application des autres TAB). La seconde contrainte vient du fait que chaque période d'application (disons de 30 ans) voit la simulation adopter les mêmes valeurs de corrélation inter-variables (du produit de référence), ce qui force une stagnation à long terme de ces valeurs, alors que l'évolution de processus dynamiques pourrait *a priori* impliquer des évolutions dans ces propriétés statistiques.

La procédure de *shuffling* vise l'ajustement des coefficients de corrélation de Spearman, puisque ces coefficients sont évalués à partir des rangs des différentes valeurs de la série temporelle et non entre les valeurs de ces dernières, comme c'est le cas pour les coefficients de corrélation (linéaire) de Pearson. Si le coefficient de corrélation de Spearman est fortement positif (proche de 1) entre deux variables (disons A et B), alors cela signifie que les grandes (ou petites) valeurs de la variable A ont tendance à s'aligner dans le temps avec les grandes (ou petites) valeurs de la variable B. Au contraire, si le coefficient de corrélation de Spearman entre les variables A et B est fortement négatif (proche de -1), alors cela signifie que les grandes (ou petites) valeurs de la variable A ont tendance à s'aligner dans le temps avec les petites (ou grandes) valeurs de la variable B ; on parle d'ailleurs d'anti-corrélation. Dans un dernier cas de figure, si le coefficient de corrélation est près de zéro, alors on ne peut pas déduire de lien évident entre le comportement des deux variables. Ce que cela implique pour le *shuffling*, c'est que le principe général est de forcer la corrélation inter-

variables (de Spearman) simulée entre les variables A et B à adopter la valeur correspondante de corrélation (de Spearman) entre les variables correspondantes du produit de référence (disons A* et B*). Par exemple, si l'on choisit la variable A comme étant la variable maître et si la plus petite valeur de la variable A* (rang 1) est accompagnée dans le temps de la 245^{ème} plus petite valeur de B* (rang 245), alors on déplacera dans le temps la 245^{ème} plus petite valeur de B (rang 245) pour l'aligner sur la plus petite valeur de A (rang 1). Si la deuxième plus petite valeur de A* (rang 2) est accompagnée dans le temps de la 127^{ème} plus petite valeur de B* (rang 127), alors de même, on déplacera dans sa série temporelle la 127^{ème} plus petite valeur de B (rang 127) pour l'aligner sur la 2^{ème} plus petite valeur de A (rang 2). Et ainsi de suite, ce qui brasse la variable non-maître B.

La Figure 2.2 fournit un exemple simple de *shuffling*, avec seulement 4 rangs et 2 variables (*T* et *P*). Ce diagramme montre que le choix de la variable maître n'affecte pas les couples de valeurs (le pairage), mais affecte l'ordre des couples dans le temps. L'on peut aussi suspecter que l'autocorrélation de la variable non-maître est affectée d'une certaine manière, alors que pour la variable maître, la procédure QM a tendance à conserver l'autocorrélation de la simulation climatique initiale (Wilcke et al., 2013).

2.1.3 Transformation géométrique par vecteurs propres

L'une des procédures d'ajustement de biais utilisée à des fins multivariées est la transformation par vecteurs propres (*eigenvectors*) présentée par Hnilica et al. (2017). Dans celle-ci, de nouveaux systèmes de coordonnées sont générés à l'aide des vecteurs et des valeurs propres des données respectives (nuages de points) du produit de référence et de la simulation climatique. Les vecteurs propres sont des axes orthogonaux entre eux qui représentent la forme et l'extension d'un nuage de points

donné, c'est-à-dire la variance des données (apparente dans la matrice de covariance). L'utilisation de ces vecteurs comme nouveau système de coordonnées permet au nuage de point simulé d'être « poussé » vers celui du produit de référence, ce qui contribue indirectement à ajuster certains biais puisque certaines propriétés géométriques du nuage de points du produit de référence vont être adoptées par le nuage de points simulés. Pour ce faire, ce dernier va subir des rotations, translations et homothéties (agrandissement ou réduction), le rapport d'homothétie étant donné par les valeurs propres associées aux vecteurs propres. En pratique, pour arriver au résultat escompté, il faut placer l'origine des vecteurs propres à la moyenne de l'ensemble des données et les redimensionner pour qu'ils aient une longueur égale à la racine carrée de la valeur propre correspondante, c'est-à-dire à l'écart type des données. Il est indispensable de noter que, dans tous les cas, la structure générale des données simulées va rester la même.

Puisque cette étude inclut 4 variables (T , P , q et RH), cela entraîne un problème quadridimensionnel qu'il est difficile de représenter visuellement. Ainsi, la Figure 2.3 présente un exemple simple à 2 dimensions pour les variables T et P . Le système de coordonnées d'origine pour les données du produit de référence et de la simulation brute est la base canonique notée \mathcal{C} (axe x (P) et axe y (T) de la Figure 2.3). Le système de coordonnées basé sur les vecteurs propres pour les données du produit de référence est \mathcal{O} (vecteurs propres en noir) et celui basé sur les données de la simulation brute est \mathcal{M} (vecteurs propres en rouge). Soit $m = (m_1, m_2)$ un vecteur de données simulées dans le système \mathcal{C} (point rouge sur la Figure 2.3). Pour ajuster les biais, il faut trouver les coordonnées du point m dans le système \mathcal{M} , puis remplacer ce point par le point t , qui a les mêmes coordonnées relatives que le point m , mais cette fois par rapport à l'origine du système du produit de référence \mathcal{O} . Autrement dit, l'ajustement des biais par les vecteurs propres consiste à convertir le système de coordonnées \mathcal{M} en \mathcal{O} . Pour ce faire, il y a 3 étapes : 1) la transition du point $m = (m_1, m_2)$ du système \mathcal{C} au point $m^* =$

(m_1^*, m_2^*) du système \mathcal{M} ; 2) le remplacement du point m^* du système \mathcal{M} par le point $t^* = (t_1^*, t_2^*)$ qui a les mêmes coordonnées que m^* (i.e. $t_1^* = m_1^*$ et $t_2^* = m_2^*$) mais dans le système de coordonnées du produit de référence \mathcal{O} ; 3) la transition du point $t^* = (t_1^*, t_2^*)$ du système de coordonnées \mathcal{O} au point $t = (t_1, t_2)$ du système \mathcal{C} , pour revenir à la base canonique d'origine. Le point final $t = (t_1, t_2)$ du système \mathcal{C} (point bleu sur la Figure 2.3) représente alors la donnée ajustée, qui avec les autres éléments transformés (non-représentés sur cette figure), et suite à d'autres procédures éventuelles (*quantile mapping*, préservation de tendance, etc.), vont former le scénario climatique.

L'utilisation des vecteurs propres pour ajuster les biais comporte une problématique pratique non triviale. En effet, les algorithmes génériques de calcul de vecteurs et de valeurs propres doivent déterminer un sens pour les axes du nouveau système de coordonnées, et dans les algorithmes disponibles ceci est décidé sur des bases mathématiques sans lien avec le problème physique auquel on l'applique (la fonction $\text{eig}(\cdot)$ de Matlab® a été utilisée). De ce fait, pour ajuster les données, il faut déterminer un sens directionnel arbitraire pour chaque vecteur propre du produit de référence. Dans notre cas, il y a 16 (2^4 car 2 sens possibles pour chacune des 4 variables) possibilités (dans le cas de la Figure 2.3 servant à illustrer la transformation, il y a $2^2 = 4$ possibilités, car 2 sens possibles et 2 variables), qui mènent chacune à des résultats différents concernant les données ajustées. Comment choisir la bonne direction des axes des vecteurs propres pour s'assurer que les transformations ne soient pas inversées et que les données soient adéquatement ajustées ? Notre stratégie ici est d'essayer les 16 possibilités. La possibilité retenue sera celle qui maximise les corrélations entre les vecteurs d'entrée de la simulation climatique initiale et les vecteurs de sortie du scénario climatique, supposant que la « bonne » transformation est celle qui modifie le moins le nuage de points. Dans de rares cas, la solution retenue peut comporter des coefficients négatifs, ce qui implique une inspection visuelle systématique des résultats.

En outre, il semblait important de mentionner quelques recommandations concernant le recours à la transformation par vecteurs propres. La première est qu'il ne faut pas confondre cette transformation avec ce qui est fait dans une analyse en composantes principales (ACP) car les objectifs de ces deux procédés sont bien distincts, même s'ils ont, tous les deux, recours aux valeurs et vecteurs propres. En effet, l'ACP utilise les vecteurs et les valeurs propres pour résumer l'information en ne retenant qu'un nombre réduit de nouvelles dimensions, qui sont des combinaisons linéaires des dimensions initiales, ce qui en fait une procédure de réduction de dimensionnalité. Au contraire, la transformation par vecteurs propres utilisée dans ce mémoire ne réduit pas le nombre de dimensions. La deuxième recommandation est de toujours appliquer une procédure assurant un ajustement adéquat et systématique des propriétés univariées, comme le QM, après le recours à la transformation par vecteurs propres. En effet, après inspections visuelles des résultats, la procédure multivariée ayant recours aux vecteurs propres peut mener, dans de rares cas, à des résultats univariés insatisfaisants (si aucune procédure univariée supplémentaire n'est ajoutée).

2.2 Métriques d'évaluation

Une fois les résultats (c'est-à-dire les scénarios climatiques correspondant aux simulations ajustées) générés, on peut les évaluer de deux manières, soit par vérification et par validation. D'une part, on peut vérifier que les choix de procédures spécifiques sont bien rendus par l'algorithme correspondant à la technique dans son ensemble (c'est-à-dire à l'assemblage des diverses procédures sélectionnées). D'autre part, on peut procéder à une certaine validation du caractère approprié des choix de *design* de la technique, en évaluant par exemple ce qui se passe au niveau de propriétés statistiques importantes mais non directement ciblées par l'une des procédures. La vérification (de si l'algorithme rend bien les choix scientifiques) et la validation (de si

les choix scientifiques sont appropriés) sont deux volets d'évaluation distincts, mais dans les faits ils ne sont pas totalement séparables vu que les choix scientifiques initiaux et les possibilités mathématico-numériques s'influencent l'une l'autre.

Les procédures d'ajustement de biais détaillées dans la Section 2.1 ont été choisies pour ajuster des propriétés statistiques bien précises. Par exemple, si le *shuffling* est activé dans le but d'ajuster les corrélations entre q et T , on peut vérifier *a posteriori* un résultat attendu *a priori* (par *design*). Des métriques d'évaluation appropriés, comme la corrélation inter-variables dans ce cas (voir Section 2.2.1), vont servir à la vérification des résultats *a posteriori* car si l'on n'a pas le résultat espéré (ajustement), alors cela signifierait un bogue dans notre algorithme. Cependant, il existe d'autres propriétés statistiques non directement ciblées par les procédures d'ajustement de biais, mais qui pourraient être d'intérêt dans des études subséquentes. De ce fait, il devient important de valider *a posteriori* la pertinence des TAB en surveillant ce qu'il advient de ces propriétés choisies subjectivement (de par leur caractère important et du fait qu'il n'y a pas d'attente *a priori* ; s'il y a de l'attente *a priori*, alors on est dans de la vérification). Les Sections 2.2.2 et 2.2.3 présentent des exemples de validation au moyen de la variabilité interannuelle des moyennes mensuelles et de l'autocorrélation mensuelle de décalage 1, respectivement.

2.2.1 Corrélation inter-variables

La corrélation inter-variables est une propriété statistique directement ou indirectement ajustée par les TAB multivariées intégrant le *shuffling* ('QM-shuf') ou la transformation par vecteurs propres ('eig-QM'). Il est donc attendu, *a priori*, que les biais dans cette propriété soient ajustés pour les techniques en mode multivarié (bien que les procédures auxiliaires puissent rendre le résultats imparfait). En revanche, pour

la technique univariée utilisée dans ce mémoire ('QM-only'), aucun ajustement n'est assuré pour cette propriété. De ce fait, *a priori*, on s'attend davantage à ce que les biais corrélacionnels restent inchangés par rapport à la simulation climatique initiale, même si, de manière fortuite, ils pourraient être ajustés pour certains cas. Comme mentionné précédemment (voir Section 2.1.2), c'est la corrélation de Spearman qui est investiguée dans cette étude. Après l'analyse des résultats (voir Section 3.3.4), les attentes *a priori* ont été vérifiées, à savoir que les TAB multivariées ajustent bel et bien les corrélations inter-variables. Dans le cas de la technique univariée, la propriété de la simulation climatique initiale est essentiellement conservée, que la simulation soit biaisée ou non à cet égard. Quant à savoir si cela invalide la technique, la question est très subjective, et dépend de l'objectif final visé. Par exemple, si l'intention de recherche requiert une absence de biais à cet égard, alors la technique univariée pourrait être inappropriée pour satisfaire à cet objectif.

2.2.2 Variabilité interannuelle des moyennes mensuelles

La variabilité interannuelle des moyennes mensuelles n'est pas une propriété ajustée par les TAB impliquées dans cette étude. En effet, elles s'appliquent sur les distributions de valeurs journalières sur la période de calibration de 30 ans (1981-2010), ce qui implique un mélange de toutes les années. De ce fait, on ne s'attend pas *a priori* à ce que la variabilité interannuelle des moyennes mensuelles soit ajustée. En revanche, de manière fortuite, ou de manière insoupçonnée par le cerveau humain, elle pourrait l'être, mais il est difficile d'en avoir une garantie *a priori*. Néanmoins, il est possible de modifier les TAB pour qu'elles traitent la forme du cycle intra-annuel et, par ricochet, qu'elles améliorent (plus ou moins légèrement) la variabilité interannuelle des moyennes mensuelles, même si cette dernière n'est pas directement visée (voir Section 3.3.2). Par exemple, les options de groupement des jours 'mensuel' et 'fenêtre

mobile' utilisent respectivement 12 et 365 fonctions de transfert (F_{QM}) pour ajuster les biais, ce qui permet un travail sur le cycle intra-annuel, contrairement à l'option 'annuel', qui n'utilise qu'une seule F_{QM} , et qui peut par exemple mélanger des jours de printemps et d'automne qui se ressemblent par rapport à certaines variables.

Il serait possible d'incorporer une procédure supplémentaire à nos TAB qui s'occuperait explicitement de l'échelle temporelle interannuelle. C'est ce qui est proposé dans l'étude d'Haerter et al. (2011), dans laquelle il est souligné que l'échelle de temps à laquelle l'ajustement de biais est fait influence les résultats finaux. Par exemple, si l'ajustement de biais est fait sur des distributions quotidiennes (comme c'est le cas ici), alors la variabilité interannuelle après ajustement ne sera pas nécessairement « bonne » puisque c'est une propriété statistique non directement ciblée par les techniques. À cet effet, les auteurs de cette étude proposent d'appliquer des TAB en cascade, qui fonctionnent chacune sur une échelle de temps différente, pour ainsi ajuster les biais aux échelles concernées. L'implémentation de telles procédures n'a pas été privilégiée dans cette étude-ci. Toutefois, la variabilité interannuelle des moyennes mensuelles demeure pertinente à étudier, notamment pour mettre en lumière le caractère approprié ou non des choix initiaux.

2.2.3 Autocorrélation mensuelle de décalage 1

L'autocorrélation est la propriété statistique qui représente la corrélation de la variable avec elle-même. Elle dépend du décalage k de la série temporelle auquel on examine la corrélation. Par exemple, si le décalage k est de 0 pas de temps, alors l'autocorrélation est de 1 puisque l'on compare 2 séries temporelles ayant précisément les mêmes valeurs à chaque pas de temps (ce sont exactement les mêmes séries temporelles). En revanche, si le décalage k est supérieur à 0, alors l'autocorrélation de

décalage k correspond à la corrélation entre les valeurs de la série temporelle initiale séparées par k pas de temps. En langage technique (programmation Matlab), cela correspondrait à évaluer la corrélation entre la série temporelle comportant les pas de temps allant des positions 1 à $end-k$ de la série temporelle initiale et celle comportant les pas de temps allant des positions $1+k$ à end de la même série initiale (autrement dit, la série temporelle initiale est décalée de k pas de temps par rapport à elle-même).

Dans cette étude, on analyse uniquement l'autocorrélation de décalage 1 (c'est-à-dire d'un seul pas de temps). Cette propriété statistique n'a aucune garantie d'être ajustée par les TAB impliquées ici, que ce soit en mode univarié ou multivarié (bien qu'en univarié, une bonne compréhension de QM permet de s'attendre à une bonne préservation de cette propriété statistique simulée). En revanche, pour la technique multivariée ayant recours au *shuffling* ('QM-shuf'), on s'attend, *a priori*, à avoir une destruction de l'autocorrélation de décalage 1 (uniquement pour les variables mélangées). En effet, comme expliqué dans la Section 2.1.2., la procédure de *shuffling* mélange les rangs dans la série temporelle, ce qui peut détériorer l'autocorrélation pour les variables mélangées (non-mâtres) que sont la pression, l'humidité spécifique et l'humidité relative. En revanche, l'autocorrélation de décalage 1 de la simulation climatique (qu'elle soit biaisée ou non à cet égard) est conservée pour la température, qui est la variable maître (autant que QM la conserve). Les résultats attendus *a priori* ont été confirmés *a posteriori*, à savoir que la technique 'QM-shuf' détruit l'autocorrélation de décalage 1 (sauf pour la variable maître T) et que la technique 'QM-only' conserve relativement bien cette propriété. Pour 'eig-QM', il est plus difficile d'avoir des attentes *a priori*, vu le caractère plus abstrait de la procédure basée sur des transformations géométriques, mais *a posteriori* on obtient que l'autocorrélation de décalage 1 de la simulation climatique (qu'elle soit biaisée ou non à cet égard) est conservée (voir Section 3.3.5). Si l'objectif de recherche requiert une autocorrélation réaliste, alors la TAB intégrant le *shuffling* ('QM-shuf') ne devrait pas

être privilégiée. D'autres techniques d'un autre genre, comme celle univariée ('QM-only') ou celle intégrant la transformation par vecteurs propres ('eig-QM', multivariée), pourraient être utilisées lorsque la simulation est déjà adéquate à cet égard, puisqu'il a été prouvé dans ce mémoire (voir le Chapitre 3) qu'elles n'amélioreraient (ni ne détérioreraient) pas les résultats initiaux de la simulation climatique à l'égard de cette propriété.

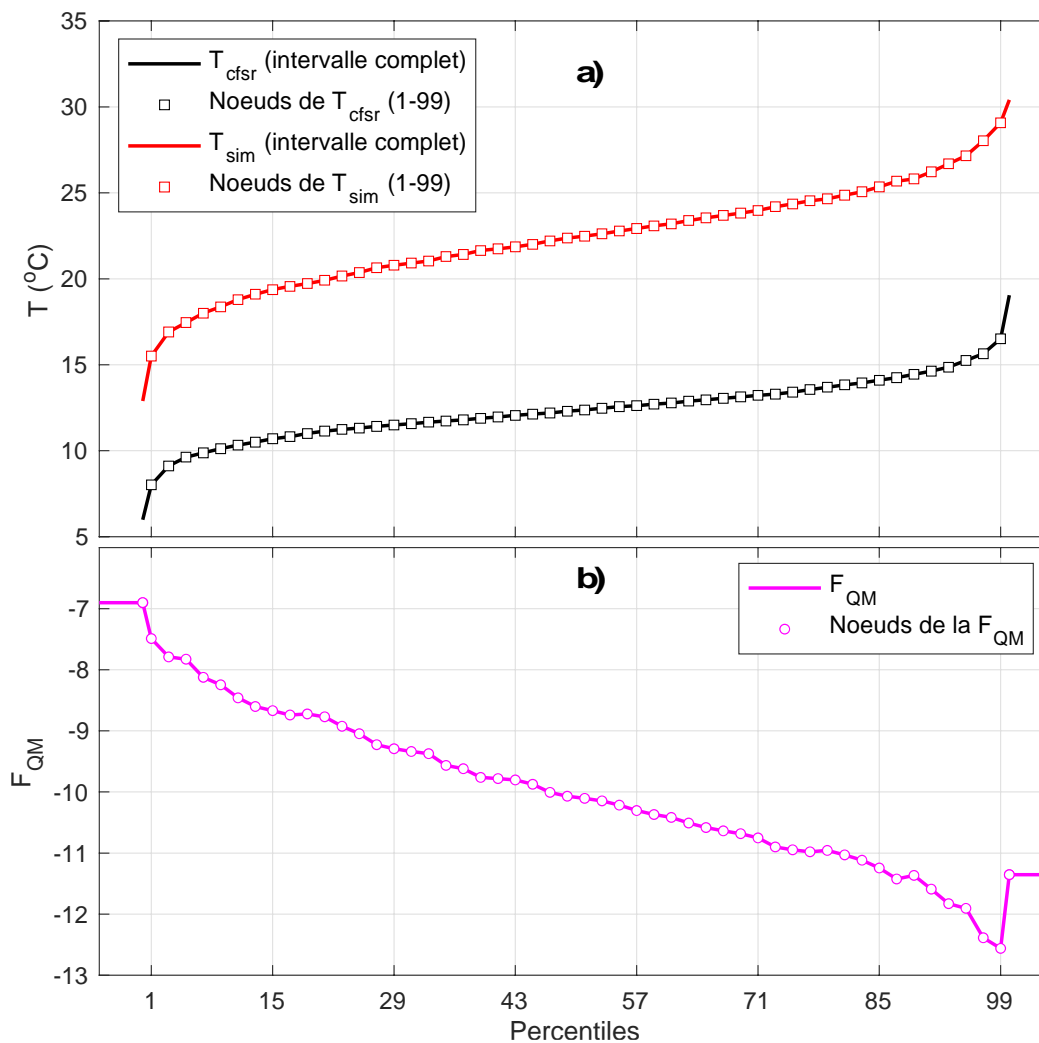


Figure 2.1 : a) les centiles de la distribution de la variable T pour la simulation 9 (SIM-09) en rouge et pour le produit de référence CFSR en noir, à San Francisco, pour le 11 août avec une fenêtre mobile de 31 jours (soit du 27 juillet au 26 août), avec l'option de régression 'none' et l'option de technique d'ajustement de biais univariée 'QM-only'. Les symboles carrés montrent les centiles de 1 à 99 par incrément de 2 alors que les extrema de distribution sont associés aux centiles 0 (minimum) et 100 (maximum) ; et b) fonction de transfert (F_{QM}) additive avec interpolation linéaire entre les nœuds (symboles ronds) et extrapolation constante au-delà des nœuds limites (0 et 100).

a) Produit de référence				b) Après QM (<i>quantile mapping</i>)			
T (°C)	Rang(T)	P (hPa)	Rang(P)	T (°C)	Rang(T)	P (hPa)	Rang(P)
8	<u>1</u>	1012	<u>4</u>	<u>11</u>	2	<u>1009</u>	4
16	<u>4</u>	999	<u>2</u>	<u>14</u>	3	<u>1004</u>	3
12	<u>2</u>	1005	<u>3</u>	<u>10</u>	1	<u>1000</u>	2
15	<u>3</u>	987	<u>1</u>	<u>17</u>	4	<u>994</u>	1

c) Après <i>shuffling</i> avec T maître				d) Après <i>shuffling</i> avec P maître			
T (°C)	Rang(T)	P (hPa)	Rang(P)	T (°C)	Rang(T)	P (hPa)	Rang(P)
11	2	1004	3	10	1	1009	4
14	3	994	1	11	2	1004	3
10	1	1009	4	17	4	1000	2
17	4	1000	2	14	3	994	1

Figure 2.2 : Exemple simple de *shuffling* pour les variables de température (T) et pression (P). Valeurs et rangs associés de T et P a) pour le produit de référence ; b) après ajustement univarié des biais (QM) ; c) après application du *shuffling* avec T (en gras) comme variable maître ; et d) après application du *shuffling* avec P (en gras) comme variable maître. Les éléments soulignés dans a) et b) sont les éléments qui servent à construire les tableaux c) et d) (rangs pour le tableau a) et valeurs pour le tableau b).

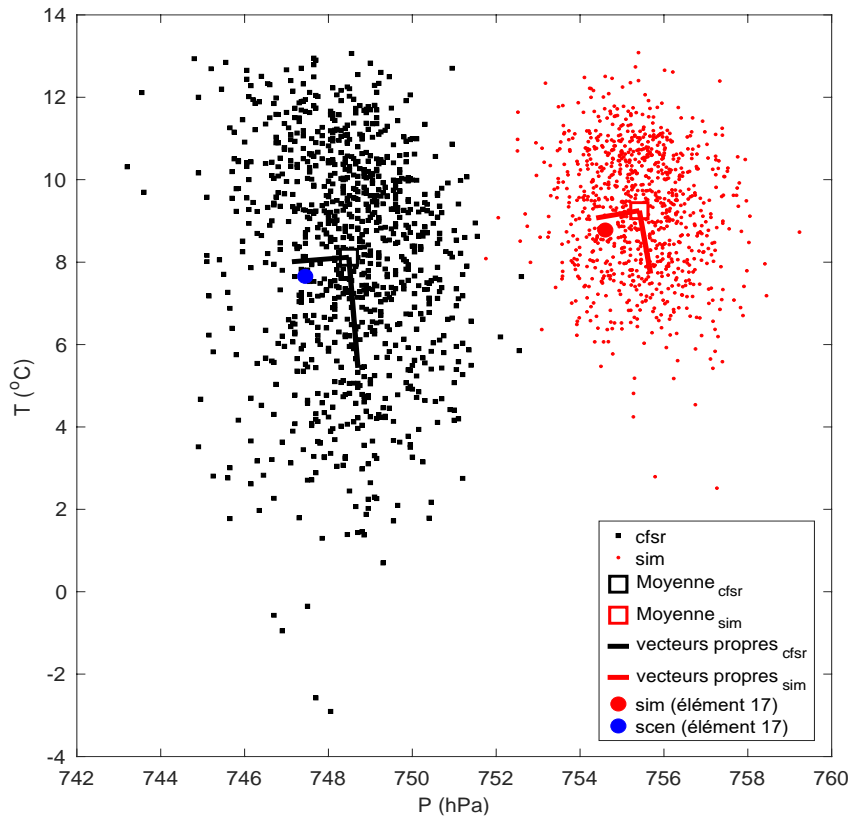


Figure 2.3 : Exemple à 2 dimensions pour T (en $^{\circ}\text{C}$) et P (en hPa) de la transformation par vecteurs propres à Mexico City. Les 930 points noirs (fenêtre mobile de 31 jours \times 30 années) représentent les valeurs du produit de référence (CFSR) alors que les 930 points représentent les valeurs de la simulation numéro 10 (SIM-10), pour le 1^{er} octobre. Les carrés représentent respectivement la moyenne des données pour CFSR (noir) et celle pour SIM-10 (rouge). Les lignes rouges (noires) sont les vecteurs propres de la simulation (CFSR), qui représentent le système de coordonnées $\mathcal{M}(\mathcal{O})$. Les gros points rouge (point $m = m^*$) et bleu (point $t = t^*$) représentent respectivement l'élément 17 (sur 930) de la simulation initiale brute (SIM-10) et du scénario climatique final (point ajusté). La station de Mexico City se situe à 2 571 m d'altitude, ce qui explique des pressions atmosphériques proches de 750 hPa. Il est à noter qu'il y a un écart entre l'altitude assignée dans CFSR (tuile de grille d'environ 31 km de côté incluant la station de Mexico City) et la station elle-même (résolution ponctuelle). Le formalisme mathématique ici utilisé est celui de Hnilica et al. (2017).

CHAPITRE III

THE DISTINCT PROBLEMS OF PHYSICAL INCONSISTENCY AND OF
MULTIVARIATE BIAS POTENTIALLY INVOLVED IN THE STATISTICAL
ADJUSTMENT OF CLIMATE SIMULATIONS

Ce chapitre est présenté sous forme d'un article scientifique rédigé en anglais. Il est basé sur mes analyses et mes graphiques, et a été rédigé sous la supervision de mes co-directeurs de recherche. Les paragraphes 4, 5, 7 et 8 de la Section 3.4 (*Summary with concluding remarks*), qui concernent des notions avancées en matière d'interprétation des méthodes d'ajustement de biais, ont été directement rédigés par Patrick Grenier, co-auteur et directeur principal du mémoire. L'article est actuellement en révision, après avoir été soumis le 19 septembre 2021 à l'*International Journal of Climatology*.

THE DISTINCT PROBLEMS OF PHYSICAL INCONSISTENCY AND OF
MULTIVARIATE BIAS POTENTIALLY INVOLVED IN THE STATISTICAL
ADJUSTMENT OF CLIMATE SIMULATIONS

Mégane Alavoine¹, and Patrick Grenier¹

¹*Centre pour l'Étude et la Simulation du Climat à l'Échelle Régionale (ESCER),
Université du Québec à Montréal, Montréal, Québec, Canada*

Corresponding author address and e-mail:

Patrick Grenier
Université du Québec à Montréal – Centre ESCER
Pavillon Président-Kennedy – local PK-6151
201, avenue du Président-Kennedy
Montréal (QC)
H2X 3Y7, Canada
grenier.patrick.2@courrier.uqam.ca

ABSTRACT

Bias adjustment of numerical climate model simulations involves several technical and epistemological arguments wherein the notion of physical inconsistency is often referred to, either for rejecting the legitimacy of bias adjustment in general or for justifying the necessity of sophisticated multivariate techniques. However, this notion is often mishandled, in part because the literature generally proceeds without defining it. In this context, the central objective of this study is to clarify and illustrate the distinction between physical inconsistency and multivariate bias, by investigating the effect of bias adjustment on two different kinds of inter-variable relationships, namely a physical constraint expected to hold at every step of a time series and statistical properties that emerge with potential bias over a climatic time scale. The study involves the application of 18 alternative bias adjustment techniques on 10 climate simulations and over 12 sites across North America. Adjusted variables are temperature, pressure, relative humidity and specific humidity, linked by a thermodynamic constraint. The analysis suggests on the one hand that a clear instance of potential physical inconsistency can be avoided with either a univariate or a multivariate technique, if and only if the bias adjustment strategy explicitly considers the physical constraint to be preserved. On the other hand, it also suggests that sophisticated multivariate techniques alone aren't complete adjustment strategies in presence of a physical constraint, as they cannot replace its explicit consideration. As a supplementary objective, this study relates common optional adjustment procedures with likely effects on diverse basic statistical properties, as an effort to guide climate information users in the determination of adequate bias adjustment strategies for their research purposes.

Keywords: climate simulations, physical relationships, statistical biases, multivariate adjustment techniques, air humidity.

3.1 Introduction

Supply and demand for climate information have been continuous and arguably growing for decades (Hecht 1984; Lourenço et al., 2016; Lugen 2020; Findlater et al., 2021). In this sector of activity, simulations from physics-based global and regional climate models (GCMs and RCMs) play an important role, models being generally viewed as the best source for plausible values of future change in key climate variables (Flato et al., 2013). Yet, simulated statistical properties such as averages, extreme values and inter-variable correlations often show differences relative to observation-based reference products over the recent past. Therefore, it is often necessary to adjust simulations before using them, for example in impact models or for estimating plausible evolutions in threshold-dependent climate indices (Maraun et al., 2017; Lanzante et al. 2018; Zscheischler et al., 2019; Martins et al., 2021).

There exist at least four potential sources for reference-simulation statistical differences, which are not easy to disentangle and whose relative roles are context-dependent: 1) mismatch in spatial representativeness (scale, grid location and/or altitude); 2) imperfections in climate models' physics formulation; 3) imperfections in reference products; and 4) non-synchronicity between real and simulated natural fluctuations over long time scales (e.g., Addor and Fischer, 2015; Chen et al., 2016; Diaconescu et al., 2017; Kotlarski et al., 2019). In operational climate services, the third and fourth sources are often found or assumed to play lesser roles, and differences are often managed by means of a single joint statistical technique addressing simultaneously scale mismatch as well as biases stemming from GCMs' and RCMs' imperfect physics (e.g., Wilcke et al., 2013; Gennaretti et al., 2015; Lehtonen et al., 2016). In such cases (and in this study), expressions like 'bias adjustment' and 'bias correction' are thus convenient misnomers, as the involved technique is in fact dealing with more than the sole bias problem. It must also be emphasized that the bias problem

is *a priori* potential, as there exist contexts (i.e., specific simulation, location, variable and statistical property) where reference and raw simulated time series match relatively well (for a mixture of right and wrong reasons). In such contexts, a suitable adjustment technique is expected to leave raw data essentially untouched (adding ~ 0 or multiplying by ~ 1).

A distinct problem involved in the use of climate simulations is the risk of generating physical inconsistency (PI) through the adjustment technique. Here the concept is used in the sense of internal or external inconsistency, occurring when a theory, a model or a description is self-contradictory or contradicts theories in neighbor fields, in line with how it is used in traditional philosophy of science (e.g., Kuhn 1977; Laudan, 1984; Steel 2010). In the context of climate simulation bias adjustment, one illustrative example would be the breaking of the thermodynamic relationship between temperature (T), pressure (P), specific humidity (q) and relative humidity (RH). This specific relationship has already been used to illustrate how PI may be generated by a typical univariate bias adjustment algorithm (Grenier, 2018), and the present study extends the methodology notably to multivariate algorithms. The nomenclature of Grenier (2018) refers to PI of type 1 when an out-of-range value is attributed to an individual variable (hence contradicting the very meaning of the variable), and to PI of type 2 when several variables have values collectively contradicting their meanings or contradicting a well-credentialed element of physics (while each individual value may respect its range).

Normally, raw simulations are devoid of many potential instances of such inconsistencies, as relevant physics elements are precisely targeted as building blocks for climate models (Jacobson, 2005; Laprise, 2008; Winsberg 2018; Hewitt et al., 2021), but exceptions linked with numerical artefacts may occur (Laprise, 2008; Ruosteenoja et al., 2017). Investigating the potential PI problem having in mind the

model-as-a-whole (rather than specific constitutive elements) would be a dead-end, as there would be no other science to be consistent or inconsistent with. Indeed, physics does not generate numerical climate models, but rather fundamental principles and relationships that climate modelers select and combine, along with less principled considerations (discretization schemes, parameterizations, domain management). Moreover, in the context of bias adjustment it is worth recalling that climate models are known *a priori* not to be in isomorphism with the real world (see: Petersen 2000; Parker, 2009; Giere, 2010; Lehnard and Winsberg, 2010), which rules out the argument that so-called model internal consistency does confer full physical consistency to simulated long-term statistics (this argument is often used against bias adjustment legitimacy). Current models and simulations are *based on* physics, but not *holistically consistent with* physics.

Framed this way, the physical inconsistency (PI) problem is distinguished from the bias problem in at least two important ways. First, PI is identified from individual time steps and grid tiles. For example, PI instantiated by daily minimal temperature exceeding the maximal daily temperature ($T_{\min} > T_{\max}$) would be diagnosed for each day separately (Thrasher et al., 2012; Agbazo and Grenier, 2019). In contrast, a bias concerns statistics (either univariate or multivariate) computed from a large number of time steps. Secondly, PI is tied to a particular definition or to a specific element of physics. For example, occurrences of adjusted RH (with respect to liquid water) substantially exceeding 100 % could be viewed as physically inconsistent for Earth's atmosphere, because the aerosol load is known as sufficient to host condensation whenever supersaturation occurs (Pruppacher and Klett, 1997). Current models generally do not resolve detailed aerosol-water processes, but typical parameterizations of cloud properties attribute meaningful roles to the 100 % threshold (e.g., Del Genio et al., 1996). In contrast, a 30-year bias in a RH time series could hardly be linked to any well-identified physics element, as it emerges from the model-as-a-whole. These

distinctions are excluded from several concurrent (and generally implicit) definitions of the expression ‘physical inconsistency’ in the context of bias adjustment. For example, some authors implicitly tag this expression (or synonyms) onto any alteration of simulated multivariate statistical dependencies (e.g., Chen et al., 2016; Sippel et al., 2016; Gómez-Navarro et al., 2018), whereas, in a rather contrary perspective, others implicitly tag this expression onto simulated gaps relative to reference multivariate properties (e.g., Vrac and Friederichs, 2015).

There exists a quasi-continuum of bias adjustment techniques, including complex ones that combine several more basic mathematical-numerical procedures. Techniques differ notably by the statistical properties they adjust, by options like attempting (or not) to preserve the simulated long-term trend, and by how days (or other temporal units) are grouped within the adjustment technique. One popular procedure is quantile mapping (QM), which adjusts the shape of simulated distributions in a univariate way (e.g., Lehtonen et al., 2016; Martins et al., 2021). The central idea of this procedure is to build a transfer function mapping the simulated onto the reference distribution over a calibration period, and then to apply it on any wanted simulation segment. Many multivariate extensions have been developed during the last decade, for purposes that also require inter-variable and/or inter-site correlations to be adjusted (e.g., Su et al., 2020). These can consist of strategic successive conditional applications of univariate QM (e.g., Piani and Haerter, 2012), or in complementing univariate QM with another procedure to adjust inter-variable dependences as well, resorting for example to shuffling (e.g., Vrac, 2018), to eigenvector-based geometric transformations (e.g., Hnilica et al., 2017), or to a complex mix of akin procedures (e.g., Cannon, 2018). Each technique is adequate when the final purpose requires its specific constitutive procedures to be activated.

In this context, the central objective of this study is to clarify and illustrate the distinction between the bias problem and the physical inconsistency problem, notably by showing that multivariate bias adjustment procedures do have the potential to break a fixed thermodynamic relationship while adjusting inter-variable correlations fairly well. To this end, 18 alternative bias adjustment techniques are applied on 10 daily climate simulations at 12 sites over North America. This offers a wide range of bias adjustment contexts, with adjustment alternatives obtained by crossing three univariate or multivariate options with two trend management options and with three temporal grouping options. The thermodynamic relationship used to illustrate the distinction is the same as in Grenier (2018), namely the instantaneous constraint linking temperature (T), pressure (P), specific humidity (q) and relative humidity (RH) for a homogeneous air parcel. Physical inconsistency is monitored following the two types (1 and 2) already mentioned, while monitoring of the statistical effects covers intra-annual cycles, inter-annual variability, inter-variable correlations, lag-1 auto-correlations and long-term climate change values. Monitoring these key statistical properties also serves a supplementary objective, namely to help bias adjustment practitioners judging whether techniques involving certain promoted procedures are adequate for specific purposes.

3.2 Data and methods

3.2.1 Data sets

Because this study essentially extends that of Grenier (2018) (from 1 to 18 bias adjustment techniques, and now including two multivariate procedures), the same data sets for near-surface temperature (T), pressure (P), relative humidity (RH) and specific humidity (q) have been used, namely the 1-hourly Climate Forecast System Reanalysis

(CFSR; Saha et al., 2010) as the reference product, and an ensemble of ten 3-hourly simulations from phase 5 of the Coupled Model Intercomparison project (CMIP5; Taylor et al., 2012) as the time series to be adjusted. One of the simulations is dynamically downscaled with the Canadian Regional Climate Model version 5 (CRCM5; Šeparović et al, 2013). Table 3.1 contains identification information for all simulations as well as internal codes for convenient referencing within this study (e.g., SIM-08). Note that same-model simulations share the same data up to year 2005; therefore certain results concerning the calibration period are based on the four RCP8.5 simulations only. Time series from grid tiles including twelve different cities over North America are used, with coordinates corresponding to that of the World Meteorological Organization (WMO)'s respective stations (although station data themselves are not used; see Table 3.2 for more detailed site-related information). This sites selection includes tropical (Miami), arid (e.g., El Paso), temperate (e.g., Vancouver) and cold (e.g., Yellowknife) environments (Peel et al., 2007). The bias adjustment calibration period is 1981-2010, and the application period is 1981-2100. Bias adjustment temporal frequency is daily, with each day represented by its variables' values at 12 UTC.

It is worth mentioning that the selection of CFSR and of the ten simulations does not necessarily meet operational climate services concerns, such as identifying the best gridded reference product for a given purpose or adequately covering uncertainty in future climate change. This is justified by the methodological nature of this study, which focuses notably on showing how promoted multivariate bias adjustment procedures may act on simulations that are initially consistent regarding one specific thermodynamic aspect. For such a purpose, it is not necessary that the level of realism of the reference product be optimal. Moreover, the potential for generalization of the results is sufficiently addressed by covering a fair variety of climatological situations,

in terms of types of climates as well as of simulated bias structures and future change values.

3.2.2 Nomenclature

For any site, reference time series are referred to as T_{cfsr} , P_{cfsr} , RH_{cfsr} and q_{cfsr} , whereas time series from any of the ten simulations are referred to as T_{sim} , P_{sim} , RH_{sim} and q_{sim} . At any time step, each of these time series respects a same thermodynamic constraint between the four variables, which can be symbolized by the pair of reverse functions $f(\cdot)$ and $g(\cdot)$ (subscript absence denotes a variable in general):

$$\text{RH} = f(T, P, q) \quad (3.1)$$

$$q = g(T, P, \text{RH}) \quad (3.2)$$

The function $f(\cdot)$ can be obtained through the following basic thermodynamic equations (Wallace and Hobbs, 2006):

$$\text{RH} \equiv 100 w / w_{\text{sat}} \quad (3.3)$$

$$w = q / (1 - q) \quad (3.4)$$

$$w_{\text{sat}} = 0.622 e_{\text{sat}} / (P - e_{\text{sat}}) \quad (3.5)$$

where w is the mixing ratio, w_{sat} the saturation mixing ratio, and e_{sat} the saturation partial pressure of water vapor. The empirical fit of Sonntag (1990), hereafter referred to by the subscript ‘so90’, has been chosen to calculate e_{sat} from temperature:

$$e_{\text{sat}}(T) = 100 \exp(a_1 + a_2 + a_3 + a_4 + a_5) \quad (3.6a)$$

$$a_1 = -6096.9385 T^{-1} \quad (3.6b)$$

$$a_2 = 16.635794 \quad (3.6c)$$

$$a_3 = -2.711193 \times 10^{-2} T \quad (3.6d)$$

$$a_4 = 1.673952 \times 10^{-5} T^2 \quad (3.6e)$$

$$a_5 = 2.433502 \ln(T) \quad (3.6f)$$

where T is in kelvins and e_{sat} in pascals. The constraint $g(\cdot)$ is straightforwardly obtained by isolating q in the system of equations (3.3) to (3.5) and using the same empirical fit for $e_{\text{sat}}(T)$. Because reanalysis or model outputs (for T , P , RH and q) may possibly not respect this $f(\cdot)$ relationship, due either to the use of another empirical fit or to a univariate extrapolation to approximate the near-surface variables from the lowest model level (Ruosteenoja et al., 2017; Grenier, 2018), RH has been recalculated with $f(\cdot)$ for all situations, and new values exceeding 100% have been capped at this threshold, with corresponding q values recalculated with $g(\cdot)$. Hence, time series that serve as input to the adjustment techniques are consistent regarding $f(\cdot)$ or $g(\cdot)$ at each time step, assuming that the involved variables represent instantaneous and spatially homogeneous quantities. Also, the study assumes RH with respect to liquid water, even for temperatures below the freezing point.

Time series directly obtained from the application of any of the 18 bias adjustment techniques (described in Section 3.2.3) are referred to as T_{ba} , P_{ba} , RH_{ba} and q_{ba} . Connections between simulated and adjusted quantities are schematized in Figure 3.1. Among the tested techniques, 6 are univariate and 12 are multivariate, and the diagram

distinguishes these two classes of alternatives by subdividing the ‘ba’ subscript into ‘uni’ (univariate procedure) and ‘multi’ (joint procedure); the latter subscripts are only used in Figure 3.1. After direct bias adjustment, the quantities q_{so90} and RH_{so90} are calculated through the relations:

$$RH_{so90} \equiv f(T_{ba}, P_{ba}, q_{ba}) \quad (3.7)$$

$$q_{so90} \equiv g(T_{ba}, P_{ba}, RH_{ba}) \quad (3.8)$$

to meet the methodological objectives presented in Section 3.2.4.

3.2.3 Bias adjustment

The 18 alternative adjustment techniques are obtained by crossing three options related to dimensionality (either univariate, shuffling-based bivariate or eigenvector-based quadrivariate) with two options related to long-term trend management (either free to evolve or tentatively preserved through the isolation of a regression) and with three options related to temporal grouping of the daily values (either annual, monthly or resorting to a moving window). Quantile mapping (QM) is a procedure common to all alternatives, as this corresponds to a univariate step also embedded into the multivariate techniques. The QM step is performed with additive (by contrast with multiplicative) transfer functions in all cases, as variables RH and q (bounded by 0 and 1) are transformed into their logit function [of general form: $y = \ln(x / (1-x))$] prior to bias adjustment, and inversely transformed through the sigmoid function [of general form: $x = 1/(1+\exp(-y))$] at the end.

The bias adjustment algorithm is schematized in Figure 3.2. The first computing step is variable categorization, whereby a variable is either kept as is if its typical distribution is disconnected from any physical bounds (like T and P), transformed into its logarithm if lower-bounded at zero while there is no upper bound (this would be the case for precipitation), transformed into its logit function if it is bounded within 0 and 1 (like RH and q), or normalized prior to logit transformation if the upper bound is not fixed (this would be the case for incoming solar radiation). This pragmatic categorization, schematized in Figure 3.3, expands upon an already proposed recourse to the logit transformation (Cannon, 2018), and its rationale is to prevent occurrence of type 1 (out-of-bound) physical inconsistency for directly adjusted variables. Next comes the decision for temporal grouping of the days. The ‘annual’ option uses a single transfer function (F_{QM}) for all days of the time series, the ‘monthly’ option uses for example an F_{QM} based on 900 days for April (30 days in this month, times 30 calibration years over 1981-2010), and the ‘(moving) window’ option starts for example with a first F_{QM} based on 930 days for January 1st (31 days from December 17th to January 16th, times 30 calibration years). The next decision in the algorithm is whether F_{QM} is applied on complete values, hence letting the simulated long-term trend free to evolve as a consequence of the quantile-dependent character of the bias to adjust (Gobiet et al., 2015), or if there is rather a trend preservation (TP) attempt, by which a regression is isolated for letting the core of the algorithm (QM plus optional multivariate procedure) operate on the residuals (Hempel et al., 2013; Agbazo and Grenier, 2019). The TP procedure resorts to a local regression (Cleveland, 1979). Finally comes the core of the algorithm, followed by reverse steps such as adding back the regression (if the trend preservation option was activated) and transforming back the logit function (for RH and q).

The core of the algorithm corresponds to the dimensionality option, namely to a choice between the univariate and one of the two multivariate possibilities. The univariate

option simply consists of applying QM on the values selected by prior options (temporal grouping, and either the complete values or the residuals around the isolated regression). The F_{QM} is defined at 52 nodes (percentiles 0, 1, 3, ..., 97, 99 and 100). QM is widely applied in climate studies, under different variants and names, and further details as well as implications have been described notably by Themeßl et al. (2012), Wilcke et al. (2013) and Gennaretti et al. (2015).

The first multivariate option resorts to the shuffling procedure, used and explained notably by Vrac (2018). Here univariate QM is first applied on each variable separately, and T is chosen as the master (unshuffled) variable on which to align each of the other variables (separately, in a bivariate fashion) in order to obtain the reference inter-variable Spearman correlation coefficients. Shuffling is a time-exchange of individual values within the time series, thus durations of the application and the calibration periods must be the same. Shuffling-based techniques hence require the application period to be covered through four successive and subsequently concatenated blocks (1981-2010, 2011-2040, 2041-2070 and 2071-2100). For the window option, implementing the shuffling is conceptually more challenging than what was done by Vrac (2018), because a choice has to be made between shuffling all values involved in the construction of the current F_{QM} (the width of the window is 31 days), or shuffling only the currently processed day-of-the-year; the latter option was retained, to ensure each simulated day is “reflected” somewhere in the final adjusted time series.

The second multivariate option resorts to eigenvectors for geometrically translating, rotating and resizing altogether the simulated distributions to match the reference ones (in a quadrivariate fashion). This procedure is described by Hnilica et al. (2017) in the context of a precipitation multi-site application (here each study site is processed separately). The Matlab® *eig*(·) function is used, and the problem of arbitrariness in the eigenvector orientations described by Hnilica et al. (2017) is addressed by

computing all sixteen (2^4) orientation possibilities and retaining the one for which the sum of the four (univariate) correlation coefficients between the simulated and the adjusted series was larger. In rare cases, the retained solution incorporated negative coefficients, which prompted systematic visual inspection of the results. The procedure is followed by a QM procedure, as visual inspection showed that such guardrail can rectify univariate results that are otherwise unstable, again in rare cases.

Hereinafter, the labels ‘QM-only’, ‘QM-shuf’ and ‘eig-QM’ will refer respectively to the six univariate, the six shuffling-based and the six eigenvector-based techniques, either in general or for a specific alternative if the auxiliary options are also specified. Each technique corresponds to a row in Table 3.3, which is discussed in Section 3.4 as a synthesis of the results.

3.2.4 Overarching methodology

3.2.4.1 Physical inconsistency

To investigate whether a bias adjustment technique does cause inter-variable physical inconsistency (PI), q_{so90} is compared with q_{ba} (see Figure 3.1). Because q_{so90} is a quantity defined to be consistent with T_{ba} , P_{ba} and RH_{ba} , any discrepancy between q_{so90} and q_{ba} would indicate a breaking of the clear instance of inter-variable thermodynamic consistency expressed by $f(\cdot)$ or $g(\cdot)$. As PI is assessed at individual time steps, there is, for each adjustment alternative, each simulation and each study site, a total of 43,800 test occasions (120 years for the application period x 365 days per year). Related results will tell whether this specific instance of thermodynamic consistency survives bias adjustment, when the adjustment strategy does not explicitly prescribe it.

It will also be investigated whether RH_{so90} can exceed 100 %. The directly adjusted quantity RH_{ba} cannot exceed this threshold because of the logit transformation, but there is no such guarantee for RH_{so90} . Related results could give insight, in case bias adjustment is shown not to preserve thermodynamic consistency, about whether it is RH or q that should be explicitly calculated from the three other directly adjusted variables, in order to respect $f(\cdot)$ and $g(\cdot)$.

3.2.4.2 Statistical properties

The bias problem relates to long-term statistics, either univariate or multivariate, and the point of rendering an algorithm multivariate is to adjust inter-variable statistical dependencies. When the various mathematical-numerical procedures promoted in the scientific literature to achieve this specific objective (e.g., Piani and Haerter, 2012; Hnilica et al., 2017; Vrac, 2018, Cannon, 2018) are used in conjunction with auxiliary procedures (e.g., trend preservation, moving window, logit transformation, multiscale adjustment), a slight undoing of the work of first-acting procedures can occur, hence rendering the final effect on several relevant statistical properties not always obvious *a priori*. Therefore, the following properties (with related test metrics) are systematically monitored: annual cycle (through visual inspection); inter-annual variability (with 30-yr standard deviations of intra-month averages); inter-variable correlation (with 30-yr averages of intra-month Spearman rank coefficients); lag-1 auto-correlation (with 30-yr averages of intra-month Spearman rank coefficients); and climate change values (with so-called Δ , from 1981-2010 to 2071-2100).

3.3 Results and analysis

3.3.1 Physical inconsistency

The effect of bias adjustment on inter-variable physical consistency is investigated by comparing q_{ba} (the q value directly obtained with adjustment) and q_{so90} (the q value it takes to be consistent with T_{ba} , P_{ba} and RH_{ba}). Results for one study site (Montreal) are presented in Figure 3.4, for all dimensionality options ('QM-only', 'QM-shuf' or 'eig-QM') crossed with all grouping options ('annual', 'monthly' or 'window'); trend preservation (TP) is activated for all panels (deactivating it led to similar results). Each individual time step from each simulation provides a count, for a sum of 438,000 counts per panel. Results show noticeable discrepancies from the 1:1 ratio line for all techniques, hence indicating a breaking of the inter-variable thermodynamic consistency expressed by $f(\cdot)$ or $g(\cdot)$. Some techniques show lower RMSE (between q_{ba} and q_{so90}) than the others, but it is clear that procedures developed for adjusting 30-year inter-variable dependencies are not made to preserve thermodynamic consistency at specific time steps. The relationship $f(\cdot)$ or $g(\cdot)$ must hence be instructed explicitly within the adjustment strategy. Selecting q_{so90} instead of q_{ba} could however lead to increased residual biases for the q variable. Among 10,368 cases (18 alternatives x 12 sites x 12 months x 4 RCP8.5 simulations), q_{so90} distributions are more biased (with respect to q_{cfsr}) than their corresponding q_{ba} distribution with a frequency of 69.3 %, and less biased than their corresponding initial q_{sim} distribution with a frequency of 89.5 % (using the Kolmogorov-Smirnov distance metric to compare).

By construction, q_{so90} values cannot exceed their corresponding q_{sat} values, because the logit transformation maintains RH_{ba} within its mathematical and physical limits [0, 100%]. The converse is however not true, as the logit transformation used to maintain

q_{ba} within its mathematical limits $[0, 1]$ does not necessarily maintain it within its physical limits $[0, q_{sat}(T,P)]$. This is illustrated in Figure 3.5, which shows that supersaturation ($RH_{so90} > 100\%$) is generated for all investigated adjustment techniques (results for Miami are representative of all study sites). Thus, after Figure 3.4 showed that preserving the relation $f(\cdot)$ or $g(\cdot)$ requires adjusting three of the involved variables and post-calculating the fourth one, Figure 3.5 shows that keeping all variables within meaningful bounds requires including RH within the three directly adjusted variables.

3.3.2 Intra-annual cycle

The effect of the adjustment techniques on the annual cycle is investigated by visual inspection of its general shape and of potential month-to-month jumps, a caveat already anticipated by Hewitson et al. (2014) and illustrated in Gennaretti et al. (2015). Figure 3.6 shows related results for the RH average annual cycle at one study site (Iqaluit) over 1981-2010, for the three temporal grouping options crossed with the three dimensionality options (with activated TP option; to which results are practically insensitive). Using annual grouping (Figs. 3.6a, 3.6b and 3.6c) leads to unsatisfactory results regarding the general shape, as the transfer function (F_{QM}) may link simulated and reference daily values from different moments of the year. However, Fig. 3.6b reveals that the shuffling process can take advantage of the respective intra-annual shapes of the master (here T) and the shuffled (here RH) variables to improve the general shape of the latter. With monthly grouping (Figs. 3.6d, 3.6e and 3.6f), the most striking problem is month-to-month jumps, particularly severe in Fig. 3.6d. Jumps stem from the fact that QM cannot adjust the ascending or descending character of intra-month sequences when the simulation and the reference product differ thereupon (see for June). Using a moving window (Figs. 3.6g, 3.6h and 3.6i) leaves results devoid of

these two problems (bad cycle shape, and jumps). Qualitative conclusions for RH at Iqaluit are generally representative of those for other investigated variables and sites.

It is worth emphasizing that the month-to-month jumps stemming from the monthly temporal grouping of the days may easily go unnoticed when verifications are performed with monthly averages only. This is shown in Figure 3.7, where RH values (from Figs. 3.6d, 3.6e and 3.6f) are further averaged over the days of a same month; adjusted annual cycles look good from this perspective, which hides the intra-month ascending or descending prevalent character of the sequences.

3.3.3 Inter-annual variability

The effect of bias adjustment on inter-annual variability is investigated by calculating the standard deviation (σ) of the 30 intra-month averages over 1981-2010. Figure 3.8 shows this metric for each variable, for one study site (Denver), with window grouping and TP (results have rather weak sensitivity to these options). To estimate the natural variability envelop, the Matlab® *bootstrp*(\cdot) function has been used to re-sample the 30 CFSR intra-month averages 10,000 times, and the 1st and 99th percentiles of resulting σ values were retained (delimiting the grey boxes in Figure 3.8). Results illustrate that these techniques, operating on daily indices, do not necessarily adjust year-to-year variability, an effect already pointed out by Haerter et al. (2011) for a univariate technique. Biases in inter-annual variability are often reduced with QM-only and eig-QM, but there are cases of deterioration. For shuffled variables (P , RH and q) in QM-shuf, inter-annual variability is generally flattened (pushed towards zero), because time-exchange of daily elements involves above-average as well as below-average years. Among 1728 cases (12 sites x 12 months x 3 shuffled variables x 4 RCP8.5 simulations), σ is found {above; within; below} the bootstrap envelop with

respective frequencies of {22 %; 63 %; 15 %} with the raw simulation, {16 %; 77 %; 7 %} with QM-only, {4 %; 52 %; 43 %} with QM-shuf, and {16 %; 77 %; 7 %} with eig-QM (apparent differences from sums of 100 % are due to rounding). For the master variable (T) in QM-shuf, the small perceptible differences with QM-only (Fig. 3.8a) stem from the same-duration constraint for calibration/application periods with QM-shuf, which affects the regressions calculation.

3.3.4 Inter-variable correlation

The effect on inter-variable correlations has first been investigated by calculating the Spearman's (rank) correlation coefficients (r_{rank}) for all days in a same month-of-the-year over 1981-2010. Figure 3.9 shows the results for the pair of variables RH and T , for one study site (Mexico City), with monthly grouping and without TP (weak sensitivity to these options). Because in this case the metric is perfectly aligned with how the shuffling procedure works, Figure 3.9 is merely a verification that QM-shuf leads to a perfect match with CFSR. For eig-QM however, r_{rank} values are not *a priori* expected to match perfectly with those of CFSR, as the procedure acts in a quadrivariate fashion (there is translation, rotation and resizing of the simulated point cloud, but no reshaping); results illustrate *a posteriori* that eig-QM generally results in a fair but imprecise adjustment for pairwise correlations.

Next, a comparison with an estimate of natural variability has been performed, which required modifying the verification metric. For each time series involved, the ensemble of the 30 intra-month r_{rank} values was generated (e.g., r_{rank} for the 31 days of May 1981, r_{rank} for the 31 days of May 1982, and so on up to May 2010), and then the average of these 30 values calculated; analogously to the Section 3.3.3 situation, the bootstrap envelop is obtained by re-sampling 10,000 times among the 30 intra-month r_{rank} values,

and retaining the 1st and 99th percentiles of the resulting averages). Results are shown in Figure 3.10 for correlations between RH and each of the other variables, for one study site (Mexico City), with window grouping and TP activated (weak sensitivity to these options). For this example and this verification metric, large simulated biases are generally improved by multivariate procedures, but there are cases of deterioration for eig-QM (e.g., October on Fig. 3.10c). Among 3456 cases (12 sites x 12 months x 6 possible pairs of variables x 4 RCP8.5 simulations), the metric is found {above; within; below} the bootstrap envelop with respective frequencies of {29 %; 39 %; 33 %} with the raw simulation, {27 %; 40 %; 33 %} with QM-only, {13 %; 69 %; 18 %} with QM-shuf, and {11 %; 74 %; 15 %} with eig-QM (apparent differences from sums of 100 % are due to rounding).

3.3.5 Lag-1 auto-correlation

The effect on auto-correlation (AC) is investigated with the Spearman's (rank) correlation coefficient of lag 1. Calculations are performed analogously to those presented in Figure 3.10 for another property, with intra-month lag-1 AC calculated for each of the 30 instances of a given month-of-the-year over 1981-2010, and with the average subsequently taken. Figure 3.11 shows results for one site (El Paso), with window grouping and TP procedure (weak sensitivity to these options). Lag-1 AC values in raw simulations are roughly preserved by QM-only and eig-QM, whereas for shuffled variables (P , RH and q) in QM-shuf the values are pushed towards zero. Among 1728 cases (12 sites x 12 months x 3 shuffled variables x 4 RCP8.5 simulations), the metric is found {above; within; below} the bootstrap envelop with respective frequencies of {36 %; 50 %; 14 %} with the raw simulation, {34 %; 52 %; 14 %} with QM-only, {3 %; 10 %; 87 %} with QM-shuf, and {32 %; 55 %; 14 %} with eig-QM (apparent differences from sums of 100 % are due to rounding).

To further illustrate the effect of the shuffling on auto-correlation (AC), Figure 3.12 shows the RH time series for August 1981 at one site (El Paso), for raw and for adjusted SIM-01, with window grouping and TP activated. In this fairly representative example, the simulated 31-day segment presents a lag-1 AC value of 0.66, which changes to 0.64 with QM-only, to 0.01 with QM-shuf, and to 0.48 with eig-QM.

3.3.6 Climate change signals

The level of preservation of the simulated long-term changes is investigated with the so-called Δ (deltas) from 1981-2010 to 2071-2100. Figure 3.13 shows results for September $\Delta_{\text{rel}}(q_{\text{ba}})$ in function of corresponding $\Delta_{\text{rel}}(q_{\text{sim}})$ values, with the subscript ‘rel’ referring to relative changes. The justification for showing Δ_{rel} stems from recourse to the $\text{logit}(\cdot)$ transformation for q , which implicitly transfers the trend preservation (TP) effort from the absolute to the relative change; the situation is the same for RH, whereas for T and P it is the absolute changes that are tentatively preserved by the TP procedure. Figure 3.13 shows that $\Delta_{\text{rel}}(q_{\text{sim}})$ is often altered when TP is deactivated, a well-known general potential effect from quantile mapping. Activating TP generally leads to fair $\Delta_{\text{rel}}(q_{\text{sim}})$ preservation, with limitations for annual grouping and multivariate options (Figs. 3.13b and 3.13c) as the RMSD between $\Delta_{\text{rel}}(q_{\text{ba}})$ and $\Delta_{\text{rel}}(q_{\text{sim}})$ then does not decrease much (higher RMSD indicates higher discrepancy from the 1:1 ratio line).

A broader view of the results may be found in Figure 3.14, which presents the RMSD values for the study’s main variables (T_{ba} , P_{ba} , RH_{ba} , q_{ba} , RH_{so90} , q_{so90}). Each panel has its own normalization, with all 216 RMSD values (12 months x 18 techniques) being divided by the maximum among these values. Results show that the most problematic no-TP situations (darker blue shades) are generally much improved by activating the

TP procedure, except the striking case of P_{ba} when shuffling allows inter-month exchanges (annual grouping). For RH, the TP procedure does not look effective, but this mostly reflects perturbations of RMSD values already small even without TP. For q , the panel essentially generalizes the September situation presented in Figure 3.13. No simple explanation has been found for why the problematic QM-shuf / annual grouping combination is pronounced for only one of the three shuffled variables; reasons possibly involve particularities in annual cycles of recent-past states and Δ 's, as well as in Δ -to- σ (signal-to-noise) ratios. One interesting feature in Figure 3.14 is that the TP effect partly carries over to post-calculated $RH_{so90} \equiv f(T_{ba}, P_{ba}, q_{ba})$ and $q_{so90} \equiv g(T_{ba}, P_{ba}, RH_{ba})$ values, possibly due to a stabilizing effect from T_{ba} .

3.4 Summary with concluding remarks

The central objective of this study was to illustrate the distinction between two problems potentially involved in the statistical adjustment of climate simulations, namely multivariate biases and physical inconsistency. A multivariate bias problem occurs when a GCM or a RCM simulation presents inter-variable dependencies that are markedly different from those of a trusted reference (observation-based) product, if the research purpose requires these dependencies to be realistic. This potential problem concerns statistical properties that emerge over a very large number of time steps and from the model-as-a-whole (whose formulation is not in isomorphism with real-world physics), and here it has been monitored with pairwise correlations among temperature (T), pressure (P), specific humidity (q) and relative humidity (RH). A physical inconsistency problem occurs when a contradiction with a variable's definition or with a well-identified element of physics is generated. This potential problem concerns physics-based expectations at individual time steps, and here it has

been monitored with the instantaneous thermodynamic relationship that links T , P , q and RH.

A supplementary objective was to help bias adjustment practitioners in identifying techniques that are adequate for specific research purposes; this was addressed by working with several typical procedures combined into 18 alternative techniques, and by investigating several statistical properties aside that involved in the central objective (inter-variable correlation). Main qualitative conclusions, based on 10 daily simulations at 12 sites across North America, are summarized in Table 3.3, where each row corresponds to the situation for either the raw simulation or one of the 18 bias adjustment techniques.

Regarding the thermodynamic relationship, results show that all investigated techniques break this instance of physical consistency when applied on simulations that are consistent thereupon (Figure 3.4). In particular, two procedures promoted to fix the multivariate bias problem, namely shuffling and eigenvector-based geometric transformations, have generated physical inconsistency, no matter the auxiliary decisions on trend preservation and on temporal grouping of the days. Of course there is an easy solution to this PI problem, which is to adjust only three of the variables and to post-calculate the fourth one, which amounts to explicitly considering the thermodynamic relationship in the bias adjustment strategy (note that RH shouldn't be the post-calculated variable; see Figure 3.5). In other words, a sophisticated bias adjustment technique may not be in itself a complete strategy, when the involved variables are constrained by a specific instantaneous physical link.

Note that this account opposes arguments from at least two widely held positions. The first position casts doubt upon legitimacy of bias adjustment in general, advancing as a main argument that it presumably destroys simulated physical consistency (e.g., Ehret

et al., 2012; Chen et al. 2016; Sippel et al., 2016). Yet, as already stated, no model is *fully* or *holistically* consistent with physics, hence there is no such thing as an automatic destruction of physics as soon as a simulated emerging property is altered. Instead, a clear instance of contradiction with physics must be identified before raising the specter of physics destruction. And if a specific relationship can be pointed out, then it should be easy to incorporate it into the bias adjustment strategy. The second position promotes multivariate procedures, advancing inter alia the argument that mere univariate adjustment could lead to outputs with inappropriate physical laws or to inconsistency in underlying model physics (e.g., Cannon et al., 2015; François et al., 2020). Yet, this is based on a rather vague association between multivariate adjustment possibilities and model physics (or physics *tout court*). In fact, it is clear that any alteration to a simulation conceptually disconnects it from its generator (the model). But it is difficult to grasp why, in the particular case of multivariate statistical procedures, such disconnection would amount to a form of reconnection with the model physics formulation or with any specific physical law. Multivariate techniques add no physics to simulations, at least not in a reductionist way. And just like a good strategy involving a univariate technique may preserve a clear instance of physical consistency, a bad strategy involving a multivariate technique may destroy it.

One potential advantage of resorting to multivariate techniques is that, when a variable of interest is post-calculated to keep consistency with its directly-adjusted parent variables, its marginal residual biases could be lower if parent variables are adjusted in a multivariate way. Such advantage has been illustrated by Zscheischler et al. (2019) for the wet bulb globe temperature, a function of temperature and relative humidity. It must however be emphasized that in this case, it is not the involved multivariate technique (MBCn; Cannon, 2018) that brings instantaneous consistency between the variables; it is rather another piece of the adjustment strategy that plays this role, namely the explicit instruction to respect the function.

Regarding biases, Table 3.3 reveals an interesting portrait. Performing adjustment by grouping all days of the year may lead to unrealistic intra-annual cycles, and grouping them on a monthly basis may lead to important month-to-month jumps (Figure 3.6) that go unnoticed through monthly averages (Figure 3.7); such caveats are avoided when resorting to a moving window (Thiemeßl et al., 2012). Inter-annual variability is not directly targeted for adjustment by the techniques tested in this study, but *a posteriori* a prevalent improvement was found for QM-only and eig-QM alternatives. As for QM-shuf, the prevalent flattening effect expected *a priori* has been verified (Figure 3.8). Daily techniques could eventually be subsumed into the cascade approach proposed by Haerter et al. (2011), which also favors adjustment of inter-annual intra-month averages. Inter-variable correlation is the hallmark of bivariate QM-shuf (Figure 3.9), whereas QM-only essentially keeps the simulated values and quadrivariate eig-QM is shown *a posteriori* to have a prevalent improvement effect (Figure 3.10; for these results it must be recalled that the verification metric is de-aligned from the exact way the shuffling operates). Finally, simulated auto-correlation was essentially kept by QM-only and eig-QM alternatives (Figure 3.11). In contrast, QM-shuf has a prevalent suppressing effect on auto-correlation (except for the master variable), which is well illustrated by day-to-day sequences (Figure 3.12). The TP optional procedure generally had a weak effect on the calibration period's biases situation, and its general ability to preserve simulated monthly changes (Δ) can fairly be granted except when combining QM-shuf or eig-QM with annual grouping (Figures 3.13 and 3.14). Note that there exists a defensible position emphasizing that signal alteration is just a normal consequence of the quantile-dependent character of the bias (e.g., Gobiet et al., 2015).

This study was not intended to bring out any better bias adjustment technique. Instead, Table 3.3 should be viewed as a helpful chart matching several potential choices with likely effects on the input simulations. Assuming a preceding careful simulation selection step, this is in line with adequacy-for-purpose thinking (Parker, 2009; Parker,

2020), because procedures are not viewed as valid or not, but rather as adequate or not in a given context of use. This however does not hinder the identification of particular difficulties for certain procedures. For example, shuffling does not look good in the context of bias adjustment. Indeed, for the benefit of adjusting inter-variable correlations, it flattens inter-annual variability and suppresses auto-correlation, two properties for which climate models often have an interesting level of success (e.g., Figures 3.8 and 3.11). Moreover, resorting to shuffling limits the application period to the same duration as that of the calibration period, unless one works with multiple blocks (with conceivable period-to-period jumps, an issue however not investigated in this study). But the most particular difficulty is arguably that shuffling does not allow inter-variable correlations to evolve over time, at least not as used in the R^2D^2 technique (Vrac, 2018) whereas the situation thereupon is unclear for the MBCn technique (François et al., 2020). This is paradoxical for future climate change studies, as inter-variable correlations are thus considered important enough to be adjusted, but not important enough to retain an eventual simulated evolution in this property. Nevertheless, *a priori* it cannot be ruled out that shuffling be an adequate procedure for specific end-user purposes.

Finally, it must be mentioned that the covering of multivariate procedures in this study is not exhaustive. For example, the dOTC technique recently proposed by Robin et al. (2019) and which extends the QM concept to more than one dimension by minimizing a cost function, was not included. Idem for the MBCn technique of Cannon (2018), which combines random geometric rotations and back rotations, a QM variant and a final shuffling step. Including them would not have changed the way to illustrate the distinction between the notions of multivariate bias and physical inconsistency, but it would be interesting to investigate whether these techniques can preserve the T - P - RH - q thermodynamic relationship without being explicitly instructed to. Also, this study does not have the conceptual scope covered by Maraun et al. (2017), who advocate for

a process-based orientation of bias adjustment. Indeed, dynamic meteorological processes are more complex than instantaneous relationships, lying somewhere between basic prescribed physics elements and long-term emerging properties.

ACKNOWLEDGEMENTS

We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and thank the climate modelling institutes (listed in Table 3.1 of this paper) for producing and making their model outputs available. The U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and leads the development of software infrastructure for CMIP in partnership with the Global Organization for Earth System Science Portals. The CRCM5 simulation has been generated and supplied by Ouranos. We also thank the Climate Forecast System Reanalysis (CFSR) team for making its reanalysis product publicly available. A Master's fellowship from Ouranos and the Mitacs Accelerate program made this study possible. Finally, we would like to thank Prof. René Laprise (UQAM) for valuable comments on the manuscript.

Table 3.1: Specifications for simulations used in this study. Expansions for names of modelling institutes and models can be found at <http://www.ametsoc.org/PubsAcronymList>.

Code for present study	Modelling institute	Model	Emission scenario	Member	Spatial resolution (latitude x longitude)	Reference
SIM-01	MIROC	MIROC5	RCP 2.6	r1i1p1	1.4008 ° x 1.4062 °	Watanabe et al. (2010)
SIM-02	MIROC	MIROC5	RCP 4.5	r1i1p1	1.4008 ° x 1.4062 °	
SIM-03	MIROC	MIROC5	RCP 6.0	r1i1p1	1.4008 ° x 1.4062 °	
SIM-04	MIROC	MIROC5	RCP 8.5	r1i1p1	1.4008 ° x 1.4062 °	
SIM-05	MRI	MRI-CGCM3	RCP 2.6	r1i1p1	1.1215 ° x 1.1250 °	Yukimoto et al. (2012)
SIM-06	MRI	MRI-CGCM3	RCP 4.5	r1i1p1	1.1215 ° x 1.1250 °	
SIM-07	MRI	MRI-CGCM3	RCP 6.0	r1i1p1	1.1215 ° x 1.1250 °	
SIM-08	MRI	MRI-CGCM3	RCP 8.5	r1i1p1	1.1215 ° x 1.1250 °	
SIM-09	NOAA-GFDL	GFDL-ESM2G	RCP 8.5	r1i1p1	2.0225 ° x 2.5000 °	Dunne et al. (2012)
SIM-10*	Ouranos	CanESM2 / CRCM5	RCP 8.5	r2i1p1	0.22°**	Šeparović et al. (2013) Arora et al. (2011)

* This is the only regional simulation of the ensemble; Ouranos internal identification code is bbi/bbh.

** The CRCM5 grid is rotated relative to meridians and parallels.

Table 3.2: Specifications for study sites.

Site	WMO station ID	Latitude of station (°N)	Longitude of station (°W)	Altitude of station (m)	Altitude of CFSR nearest grid tile (m)
Miami	72202	25.8	80.4	5	1
El Paso	72270	31.8	106.4	1197	1314
Denver	72565	39.9	104.7	1640	1512
San Francisco	72494	37.6	122.4	3	62
Mexico City	76679	19.4	99.1	2235	2571
New Orleans	72231	30.0	90.3	5	-3
Vancouver	71892	49.2	123.2	4	56
St-Louis	72434	38.8	90.4	171	158
Montreal	71627	45.5	73.7	36	10
St-John's	71801	47.6	52.7	141	30
Yellowknife	71936	62.5	114.4	206	225
Iqaluit	71909	63.8	68.6	34	258

Table 3.3: Synthesis of bias adjustment effects on simulated time series properties for each alternative technique, which is a combination of a dimensionality option (QM-only, QM-shuf or eig-QM), a temporal grouping option (annual, monthly or window), and an optional trend preservation procedure that is activated (TP) or not (-). For the PI problem: dark green indicates respect of the thermodynamic constraint (no PI); and dark red indicates a quasi-guaranteed breaking. For biases: grey indicates that simulated results are roughly kept; dark green stands for realistic results also expected *a priori*; light green for prevalent improvement noticed *a posteriori*; light red for prevalent deterioration noticed *a posteriori*; and dark red for clear deterioration also expected *a priori*. For long-term changes: grey indicates that simulated changes are roughly kept; and light red stands for frequent strong alteration noticed *a posteriori*. Note that for QM-shuf combinations, some red cases concern only the shuffled variables (here P_{ba} , RH_{ba} and q_{ba}), whereas results for the master variable (here T_{ba}) are in line with the corresponding QM-only results.

bias adjustment technique			instantaneous thermodynamic consistency ($q_{ba} = q_{so90}$)	biases					monthly long-term changes
				intra-annual cycle		monthly verification metrics			
				shape	jumps	inter-annual variability	inter-variable correlation	lag-1 auto-correlation	
raw simulation									
QM-only	annual	TP	The problem can be solved by explicitly considering the thermodynamic relationship within the bias adjustment strategy.						
		-							
	monthly	TP							
		-							
	window	TP							
		-							
QM-shuf	annual	TP				shuffled		shuffled	shuffled
		-				shuffled		shuffled	
	monthly	TP				shuffled		shuffled	
		-				shuffled		shuffled	
	window	TP				shuffled		shuffled	
		-				shuffled		shuffled	
eig-QM	annual	TP							
		-							
	monthly	TP							
		-							
	window	TP							
		-							

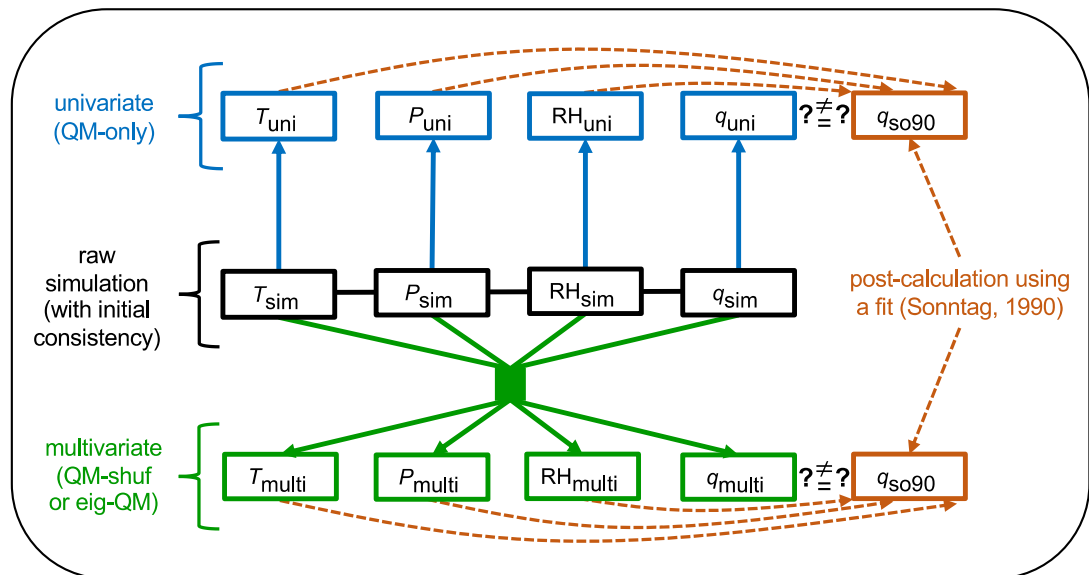


Figure 3.1: Diagram of the connections between the raw simulated (with subscripts ‘sim’) and the bias-adjusted (with either ‘uni’ or ‘multi’ subscripts, referred to by the more general ‘ba’ subscript elsewhere) variables. The role of the corresponding CFSR variables is implicit in this diagram. The variable q_{so90} is post-calculated to be consistent with T_{ba} , P_{ba} and RH_{ba} .

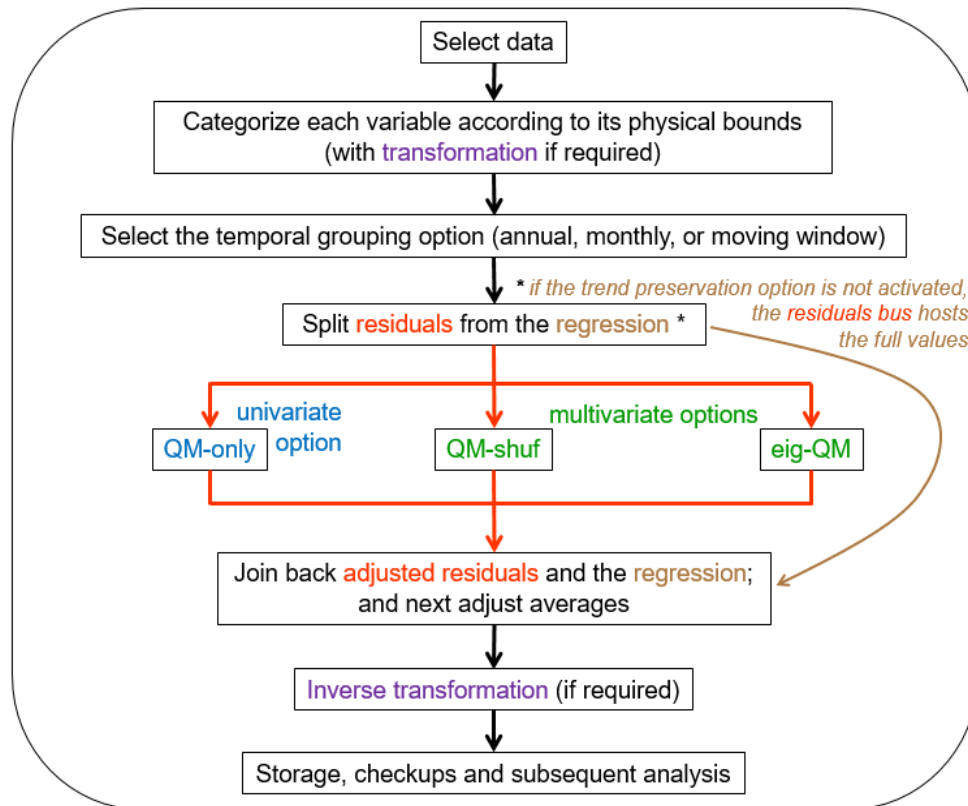


Figure 3.2: Diagram of the bias adjustment algorithm used in this study, showing the place of each of the optional decisions.

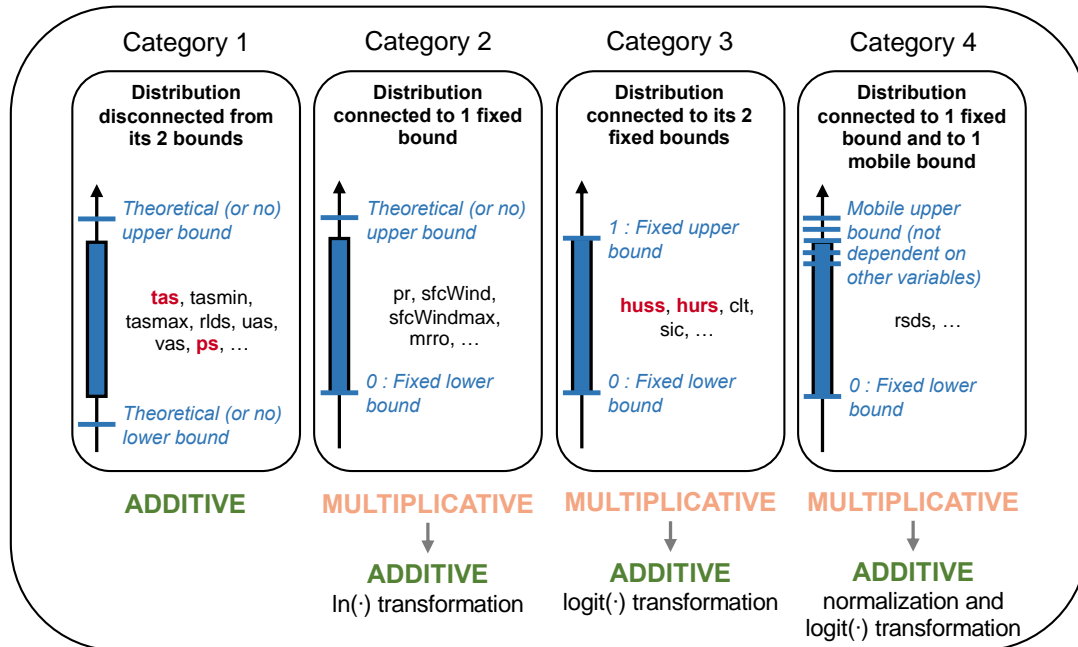


Figure 3.3: General diagram of the variables' categorization and transformations necessary to prevent bias adjustment from generating physical inconsistency of the out-of-bound type. Vertical blue bands represent the potential extension of the empirical distribution. Here variables are referred to by their conventional GCM or RCM labels, and variables used in this study are highlighted in red ('tas' for T , 'ps' for P , 'huss' for q , and 'hurs' for RH).

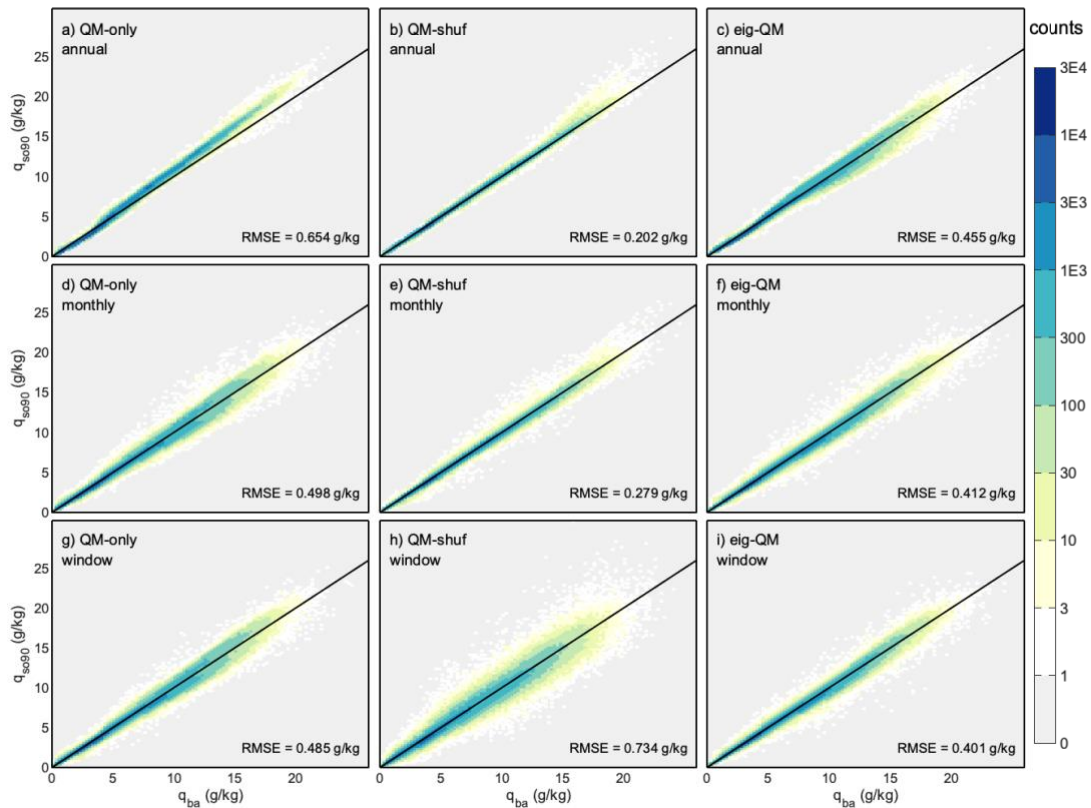


Figure 3.4: Two-dimensional counts for q_{ba} and q_{so90} , for one study site (Montreal) and for 9 of the 18 alternative adjustment techniques (TP option is activated). Counts include the ten simulations over the full application period (1981-2100), for a total of 438,000 time steps. Bin size is 0.25 g/kg x 0.25 g/kg. Black lines indicate the 1:1 ratio. On each panel, the root-mean-square error (RMSE) between q_{ba} and q_{so90} is provided as a numerical indication of physical inconsistency (departure from the 1:1 ratio).

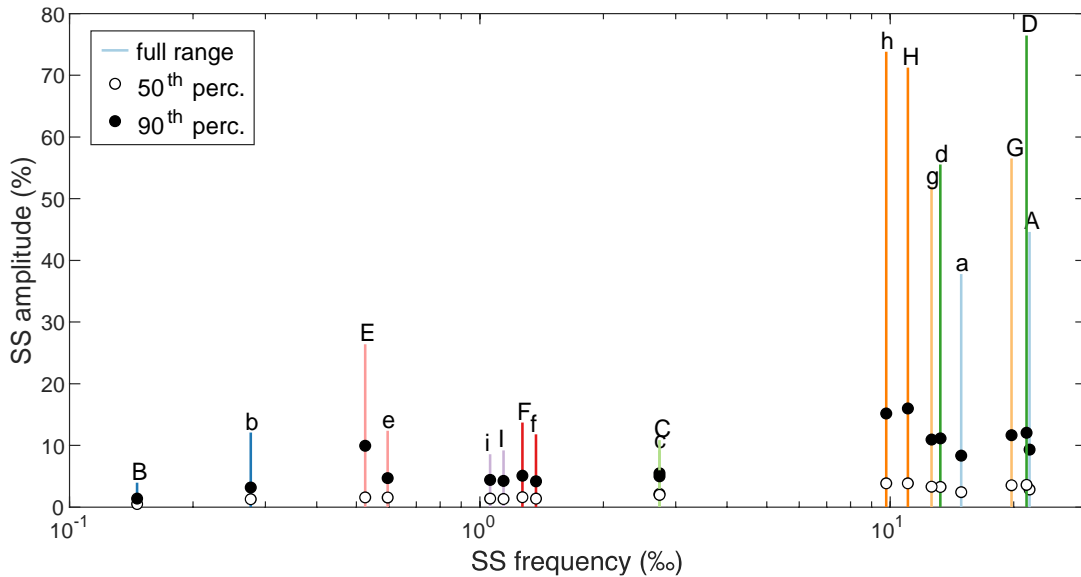


Figure 3.5: Frequency (per mil) and amplitude (%) of supersaturation (SS) occurrences in RH_{so90} , for one study site (Miami) and for each of the 18 bias adjustment techniques. Each distribution is built from the SS cases among 438,000 considered time steps (10 simulations over the application period 1981-2100), with the solid line positioned at the SS frequency (x axis) and extending over the full range of obtained SS amplitudes (along the y axis), and the empty (or filled) circle representing the 50th (or 90th) percentile. Lowercase (or uppercase) letters are used to represent the deactivated (or activated) TP option, and the matching regarding the dimensionality and grouping options is the following: a/A for QM-only and annual; b/B for QM-shuf and annual; c/C for eig-QM and annual; d/D for QM-only and monthly; e/E for QM-shuf and monthly; f/F for eig-QM and monthly; g/G for QM-only and window; h/H for QM-shuf and window; i/I for eig-QM and monthly.

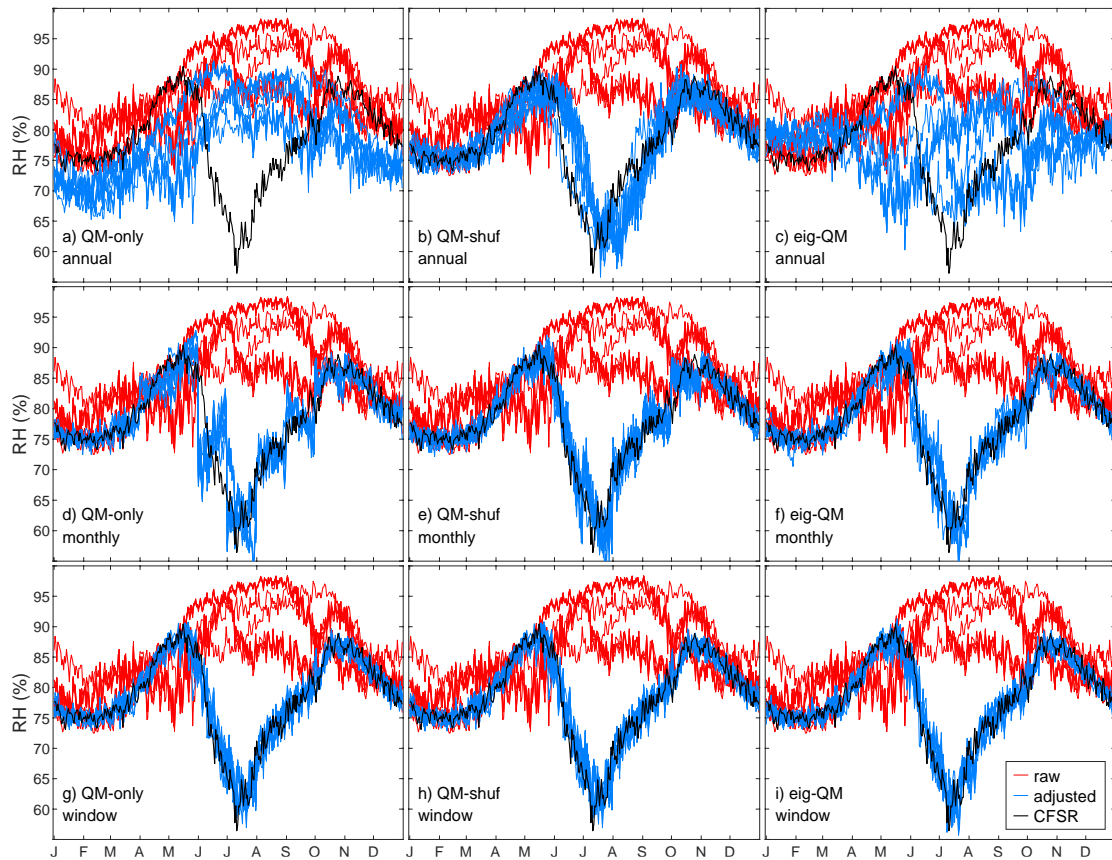


Figure 3.6: RH average annual cycle over the calibration period (1981-2010), for one study site (Iqaluit) and for 9 of the 18 alternative adjustment techniques (TP option is activated). The black line (corresponding to the reference product CFSR) and the red lines (each corresponding to one of the 10 raw simulations) are the same for all panels. Each blue line is the adjusted version of a raw simulation. Tick marks of the x axis indicate the first day of each month.

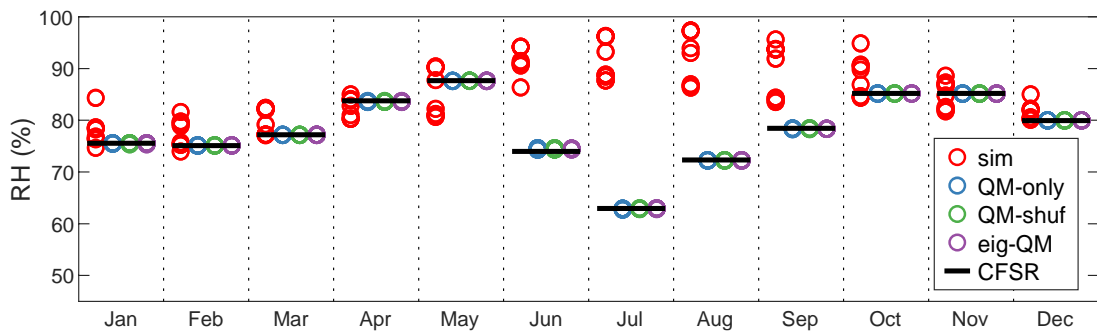


Figure 3.7: RH monthly average over the calibration period (1981-2010), for one study site (Iqaluit), for raw simulations (red), adjustment with QM-only (blue), adjustment with QM-shuf (green) and adjustment with eig-QM (purple). Each circle indicates one of the 10 (raw or adjusted) simulations (with superimposition in adjustment cases). Monthly grouping and TP options are activated. Black lines represent RH monthly averages for the reference product CFSR.

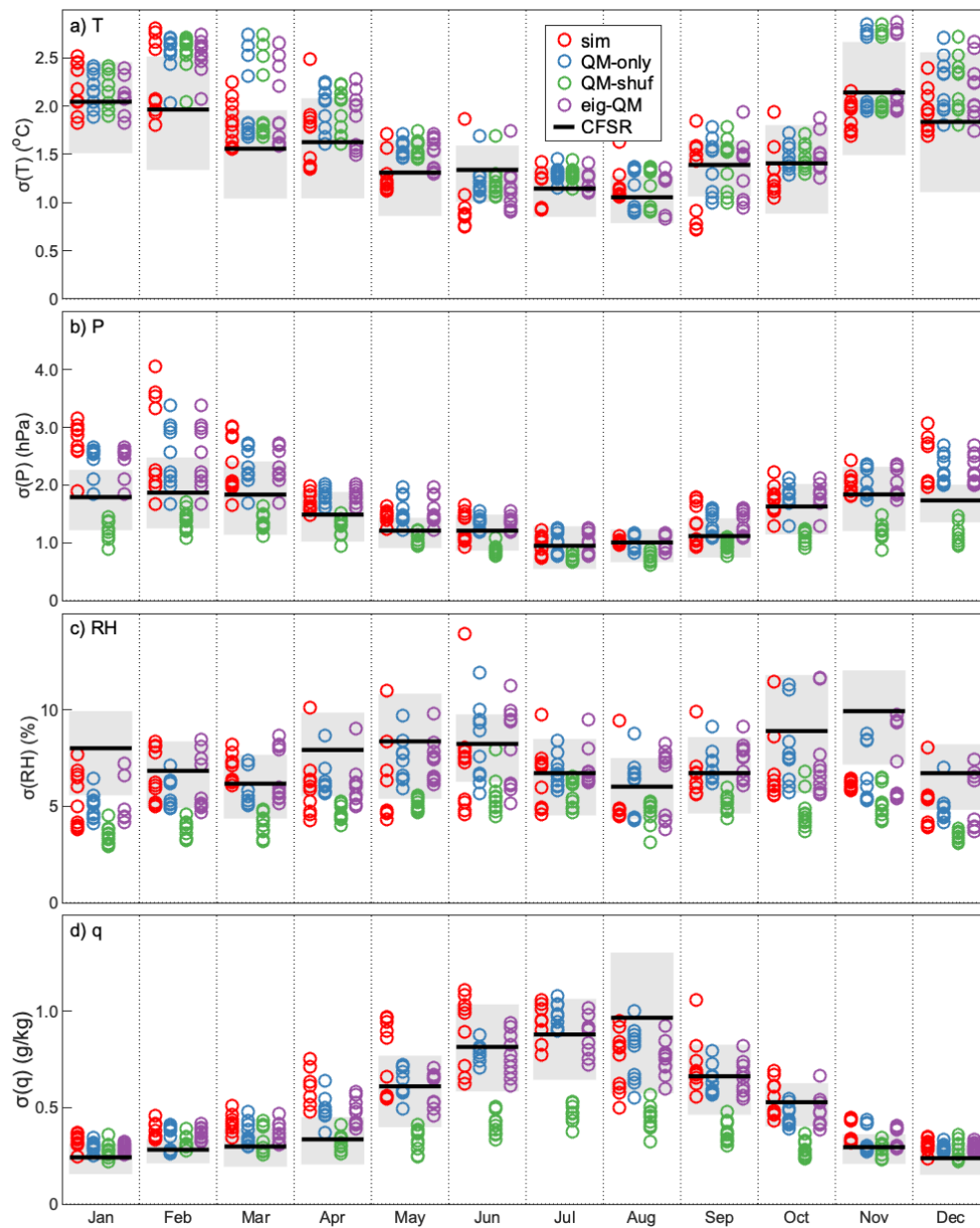


Figure 3.8: Standard deviation (σ) of monthly averages over the calibration period (1981-2010), for one study site (Denver), for variables: a) T ; b) P ; c) RH; and d) q . Each circle indicates one of the 10 (raw or adjusted) simulations, with same color code as in Figure 7. Window grouping and TP options are activated. Black lines represent CFSR σ values, and grey boxes delimit the 1st and 99th percentiles from 10,000 bootstrapping re-samplings of the 30 CFSR monthly averages.

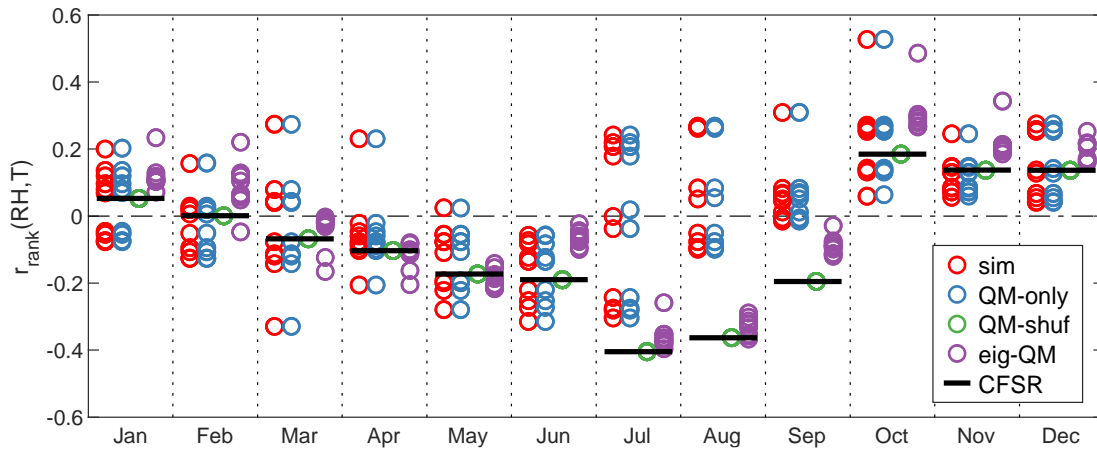


Figure 3.9: Spearman's (rank) correlation coefficient (r_{rank}) between RH and T , for all days in each month-of-the-year over the calibration period (1981-2010), for one study site (Mexico City). Each circle indicates one of the 10 (raw or adjusted) simulations, with same color code as in Figure 7. Monthly grouping is activated and TP is deactivated. Black lines represent CFSR r_{rank} values.

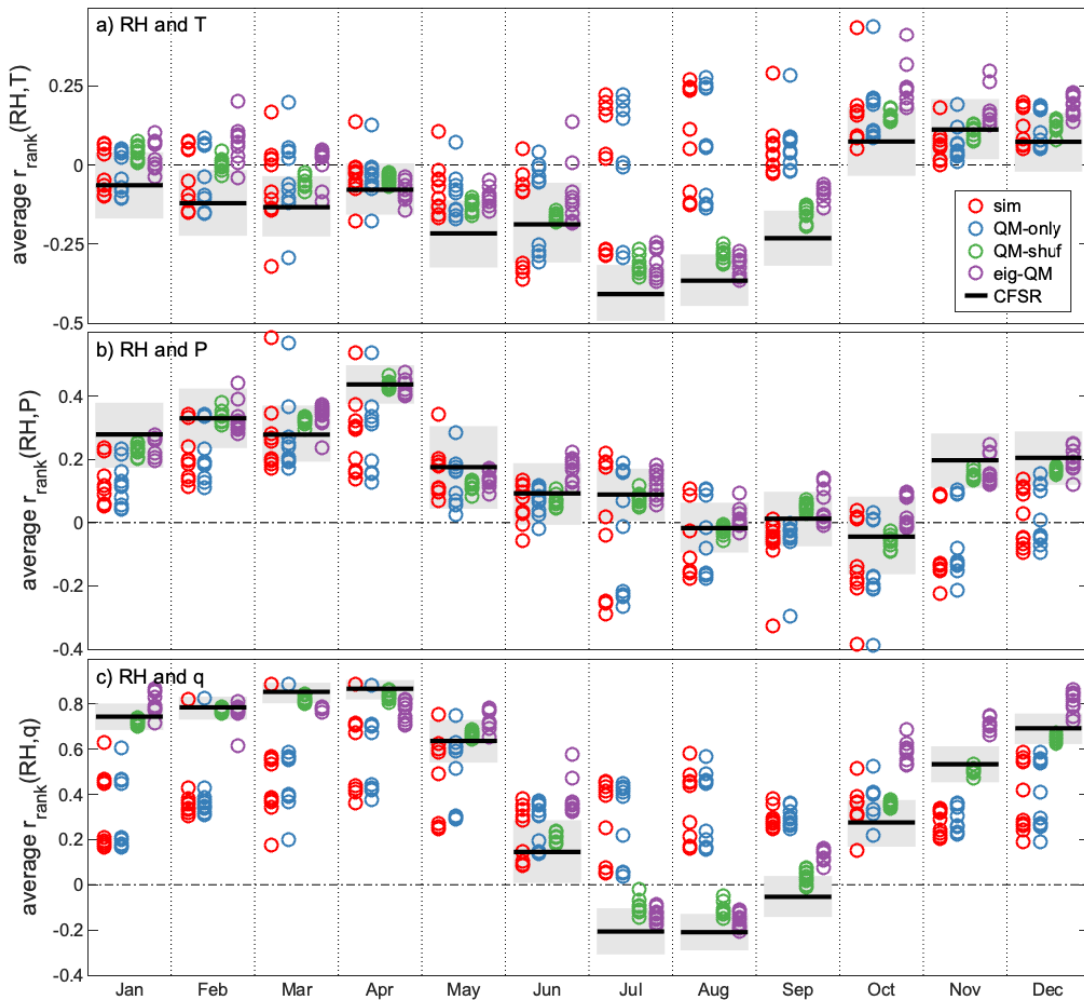


Figure 3.10: Average of the 30 intra-month r_{rank} values over the calibration period (1981-2010), for one study site (Mexico City), for pairs of variables: a) RH and T ; b) RH and P ; and c) RH and q . Each circle indicates one of the 10 (raw or adjusted) simulations, with same color code as in Figure 7. Window grouping and TP options are activated. Black lines represent CFSR average intra-month r_{rank} values, and grey boxes delimit the 1st and 99th percentiles from 10,000 bootstrapping re-samplings of the 30 CFSR intra-month r_{rank} values.

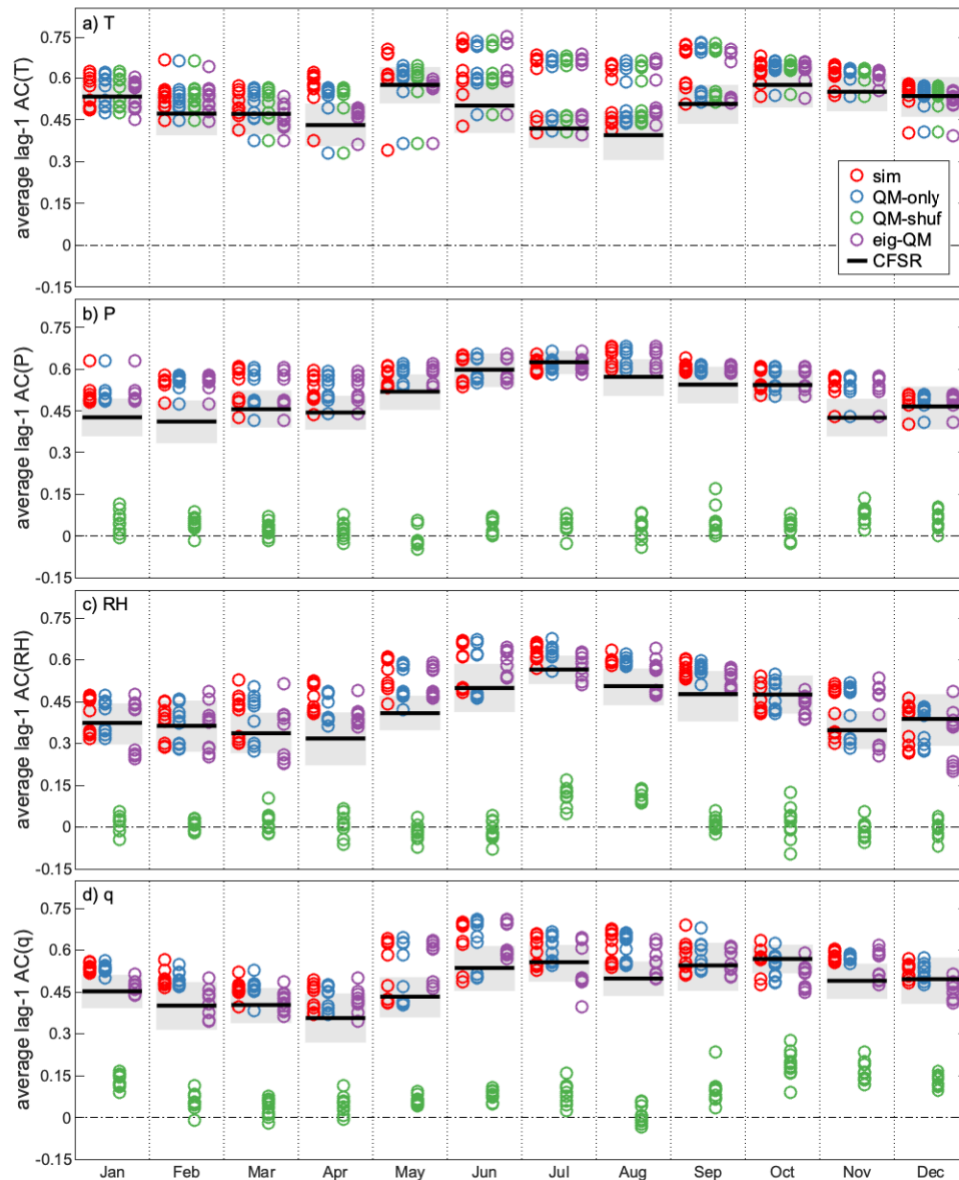


Figure 3.11: Average of the 30 intra-month lag-1 auto-correlation (AC) values over the calibration period (1981-2010), for one study site (El Paso), for: a) T ; b) P ; c) RH; and d) q . Each circle indicates one of the 10 (raw or adjusted) simulations, with same color code as in Figure 7. Window grouping and TP options are activated. Black lines represent CFSR average intra-month lag-1 AC values, and grey boxes delimit the 1st and 99th percentiles from 10,000 bootstrapping re-samplings of the 30 CFSR intra-month lag-1 AC values.

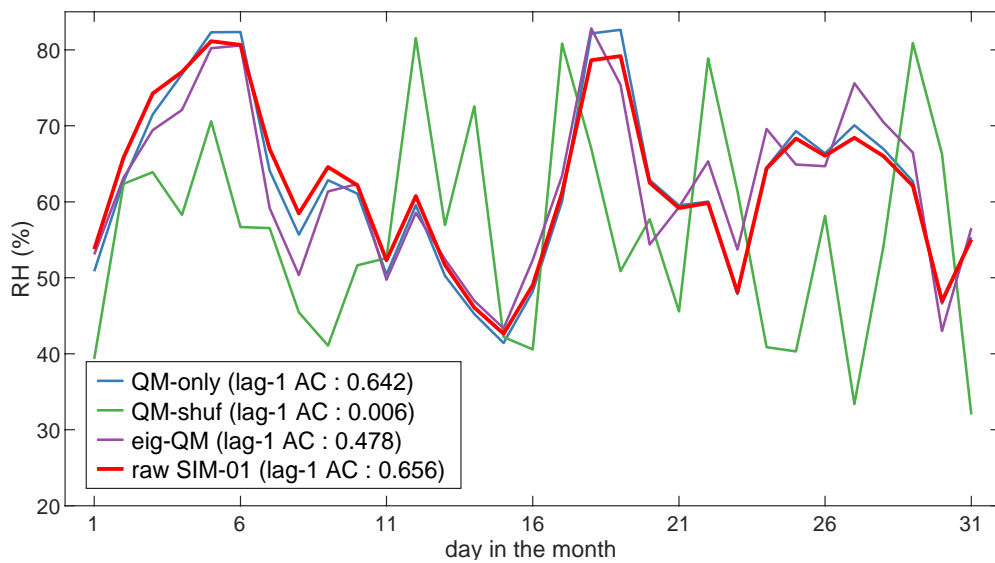


Figure 3.12: RH time series (at 12 UTC each day) for August 1981 at one site (El Paso), for raw SIM-01 and after different adjustments (window grouping and TP options activated in each case).

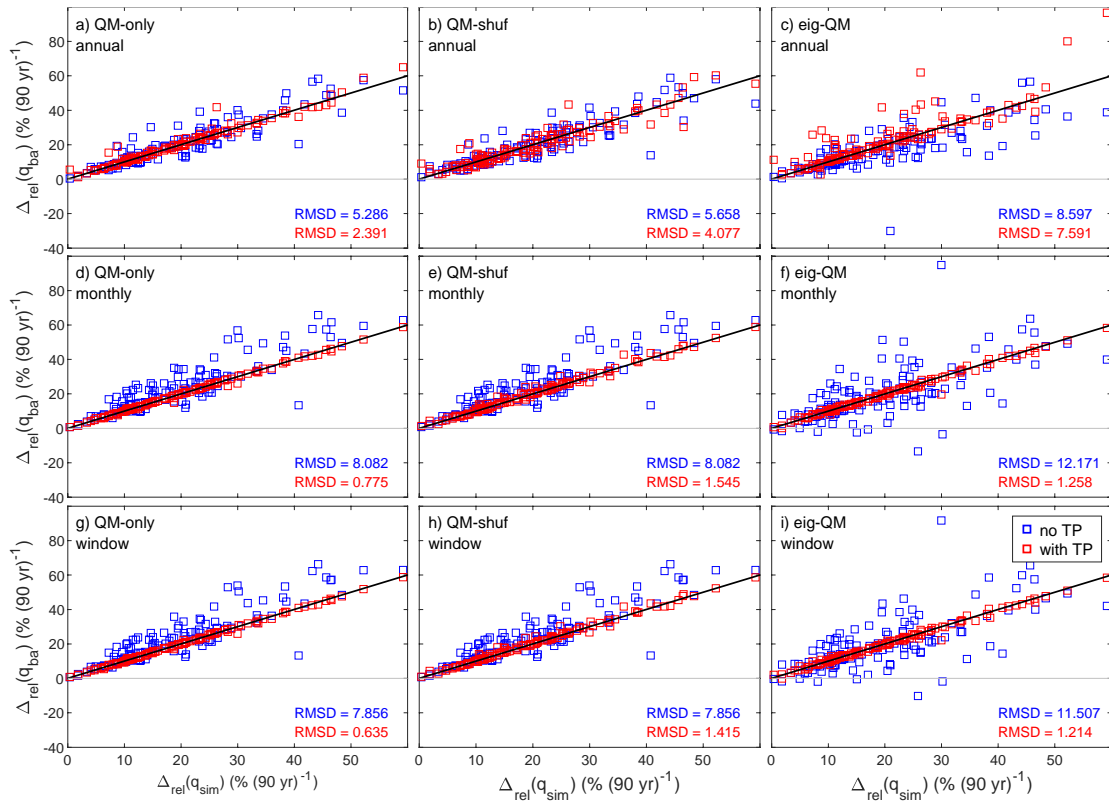


Figure 3.13: Relative long-term changes for September average specific humidity after adjustment, $\Delta_{\text{rel}}(q_{\text{ba}})$, in function of corresponding simulated changes, $\Delta_{\text{rel}}(q_{\text{sim}})$. Each panel contains 120 red symbols for adjustment with activated TP, and 120 blue symbols for adjustment without this procedure (12 sites x 10 simulations). Relative changes follow the general form $[\Delta_{\text{rel}}(q) \equiv 100(q(2071-2100) - q(1981-2010)) / q(1981-2010)]$, in units of [% (90 yr) $^{-1}$]. Black lines indicate the 1:1 proportion, and printed root-mean-square differences (RMSD) are calculated over pairs of corresponding $\Delta_{\text{rel}}(q_{\text{ba}})$ and $\Delta_{\text{rel}}(q_{\text{sim}})$ values.

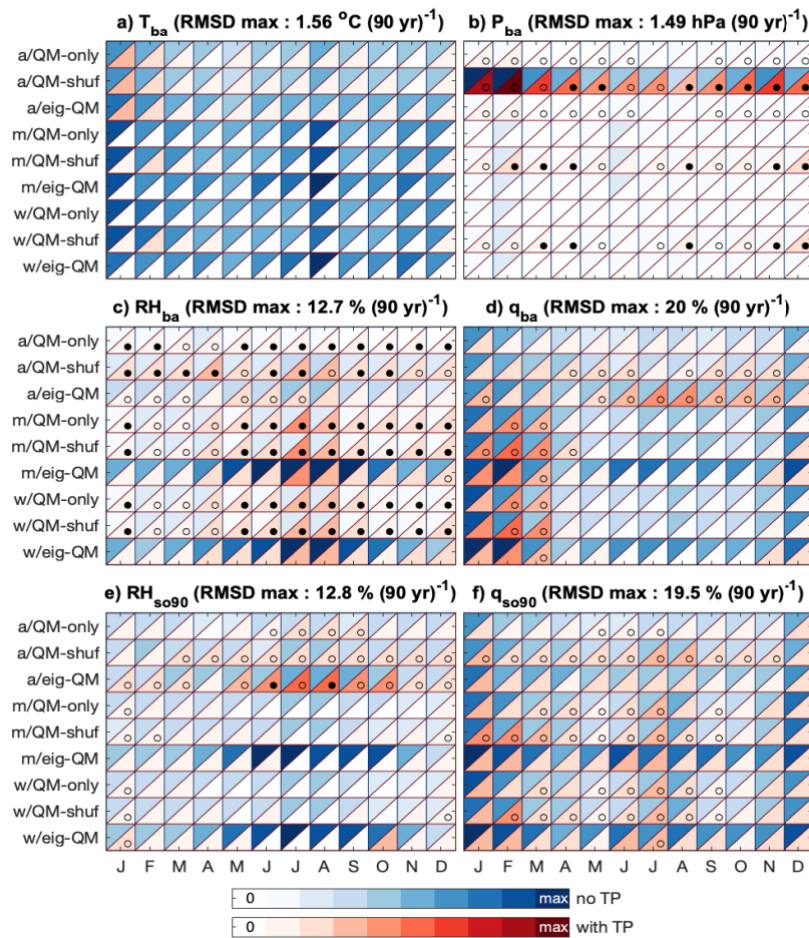


Figure 3.14: Root-mean-square differences (RMSD) over 120 pairs of corresponding bias-adjusted and raw simulated 1981-2010-to-2071-2100 change (Δ or Δ_{rel}) values (12 sites \times 10 simulations), for variables a) T_{ba} ; b) P_{ba} ; c) RH_{ba} ; d) q_{ba} ; e) RH_{so90} ; and f) q_{so90} ; and for all months (panel columns) and techniques (panel rows). Blue-shade upper-left (or red-shade lower-right) triangles correspond to technique without (or with) the trend preservation (TP) procedure. Labels ‘a’, ‘m’ and ‘w’ refer to annual, monthly and window grouping options, respectively. For each panel, RMSD values are presented as the fraction (white for 0 to 0.1, darker shade for 0.9 to 1) of the maximum of the 240 involved values. An empty (or filled) black dot indicates the associated with-TP RMSD value is larger than half its corresponding no-TP value (or larger than the corresponding no-TP value).

CHAPITRE IV

CONCLUSION

4.1 Contexte

Les centres de services climatiques ont pour mission principale de produire et de diffuser des informations climatiques de qualité et aisément interprétables, et ce de manière efficace. Cette information peut alors être utilisée comme intrant dans des études d'impact, de vulnérabilité et d'adaptation aux changements climatiques, pour ultimement aider à la prise de décision. C'est le cas des scénarios climatiques élaborés à partir de simulations climatiques, d'un produit de référence (station météorologique, produit d'interpolation ou produit de réanalyse) et d'une technique d'ajustement de biais (TAB). On vise alors à ce que ces scénarios soient réalistes quant au passé (sans biais majeurs) et plausibles quant au futur (aux tendances basées sur la physique). Le contexte actuel et futur des changements climatiques incite, depuis quelques années, des organisations internationales comme l'Organisation Météorologique Mondiale, à promouvoir de bonnes pratiques scientifiques dans les centres de services climatiques à des échelles locales, régionales et nationales partout dans le monde (Lugen, 2020).

Par ailleurs, la demande grandissante auprès des centres de services climatiques les conduit naturellement à vouloir mettre à disposition des utilisateurs des produits climatiques toujours plus crédibles et détaillés. Dans ces circonstances, on remarque

une sophistication croissante des techniques utilisées pour générer les produits climatiques, comme les TAB. En effet, jusqu'au début des années 2000, ces techniques étaient souvent très simples et ajustaient principalement les moyennes, voire aussi les déviations standards (Boé et al., 2007 ; Déqué, 2007). Progressivement, la communauté scientifique remarque l'importance d'utiliser un ajustement de biais mathématiquement sophistiqué, comme l'ajustement par quantile des simulations climatiques (ou *quantile mapping* en anglais - QM). C'est de manière plutôt simple que fut initialement utilisée la procédure QM (e.g., Mpelasoka et Chiew, 2009), sans la coupler par exemple à une procédure de fenêtre mobile ou de préservation de tendance. Depuis quelques années, des procédures numérique-mathématiques plus complexes voient le jour, comme le brassage des éléments (*shuffling*) (Vrac et Friederichs, 2015 ; Vrac, 2018) ou la transformation par vecteurs propres (*eigenvectors*) (Hnilica et al., 2017), avec pour objectif d'ajuster les biais multivariés. Les nouvelles procédures, en voulant traiter le maximum d'enjeux possibles, peuvent toutefois générer d'autres problèmes comme celui des incohérences physiques (IP).

Les effets et les impacts des TAB sur la construction de scénarios climatiques ne sont pas triviaux, à cause de l'imbrication des différentes procédures les unes dans les autres qui rend leurs actions difficilement prévisibles. Pour cette raison, il est important d'avoir un pas de recul sur les différentes TAB et de s'assurer de bien comprendre ce qui est fait dans l'algorithme, plutôt que d'imbriquer ensemble différentes procédures, car le résultat final pourrait ne pas être adéquat (même si les résultats de chaque procédure prise individuellement peuvent l'être). Aussi, il est important de distinguer divers problèmes, comme ceux de biais multivarié et d'IP, car de la confusion à cet égard peut mener à des argumentaires inadéquats, soit à l'encontre de l'ajustement de biais en général, soit en faveur ou en défaveur des procédures multivariées en particulier.

4.2 Résumé de la méthodologie

Dans cette étude, le besoin de clarifier et de distinguer les notions de biais et d'IP a été ciblé, étant donné que la littérature scientifique comporte de nombreux débats au sujet de ces deux enjeux. Le problème de biais est décelé lorsque certaines propriétés statistiques simulées présentent des différences par rapport à un produit de référence sur des décennies (généralement 30 ans). C'est d'ailleurs la raison pour laquelle un ajustement de biais est effectué sur les simulations climatiques avant de les utiliser dans d'autres études, comme des études d'impact, de vulnérabilité et d'adaptation aux changements climatiques. Un biais peut être univarié, comme des différences entre T simulée et observée (produit de référence), ou multivarié, comme des différences entre la corrélation simulée et observée de T avec RH . À noter que cet emploi du mot « biais » est conceptuellement élargi par rapport à son sens traditionnellement lié seulement aux moyennes (Galmarini et al., 2019).

Par ailleurs, le problème d'IP peut être causé par l'utilisation des TAB. Dans ce mémoire, le concept d'incohérence est le même que celui généralement utilisé dans la philosophie traditionnelle des sciences (Kuhn, 1977), à savoir une incohérence interne ou externe se produisant lorsqu'une théorie est auto-contradictoire ou contredit des éléments théoriques fondamentaux dans des domaines voisins. En pratique, cette étude reprend une classification simple proposée par Grenier (2018) où l'on a deux types d'IP : l'IP de type 1, qui se rapporte à l'attribution d'une valeur impossible à une variable (IP univariée) et l'IP de type 2, qui réfère au bris d'un lien fondamental unissant plusieurs variables (IP multivariée). Il est important de noter que la génération d'IP demeure inacceptable en toutes circonstances, considérant qu'une IP peut revenir à contredire la définition même des variables (IP de type 1) ou encore à rejeter des éléments de théorie physique bien reconnus et justement utilisés pour la construction même des modèles climatiques (IP de type 2).

Les problèmes de biais et d'IP se distinguent sur au moins deux aspects importants. Premièrement, une IP est identifiée indépendamment pour chaque ensemble de données à des pas de temps et des tuiles de grille spécifiques, alors qu'un biais concerne les différences statistiques calculées à partir d'un grand nombre de pas de temps (par exemple, sur 30 ans) et/ou de tuiles de grille. Deuxièmement, une IP est liée à une définition particulière ou à un élément spécifique de la physique tandis qu'un biais dans la série temporelle de la T sur 30 ans (par exemple) pourrait difficilement être lié à un élément physique bien identifié, puisqu'il émerge du modèle dans son ensemble (qui comporte aussi des aspects numériques, et auquel il manque certains éléments de la « physique du monde réel »).

L'étude présentée dans ce mémoire prolonge essentiellement celle de Grenier (2018), dans la mesure où elle s'applique aux mêmes 12 villes (Figure 1.1) et utilise les mêmes jeux de données pour les mêmes variables impliquées. Les 12 villes, concentrées sur l'Amérique du Nord, ont été choisies subjectivement pour représenter une grande diversité de situations climatologiques regroupant des climats tropicaux, secs, tempérés, continentaux et polaires, selon la classification de Köppen-Geiger (Peel et al., 2007). Les 4 variables, soit la température (T), la pression (P), l'humidité spécifique (q) et l'humidité relative (acronyme HR - symbole RH), ont été choisies car elles sont liées par une relation physique instantanée (voir équations 3.3 à 3.5 dans la Section 3.2.2), ce qui permet d'illustrer notre propos concernant la différence entre les biais (notion d'instantanéité) et les IP (notion de statistique). Le produit de référence choisi dans cette étude est le *Climate Forecast System Reanalysis* (CFSR) et un ensemble de 10 simulations climatiques (voir Tableau 3.1) issues de la phase 5 du projet d'inter-comparaison des modèles couplés (CMIP5) est assujetti à l'ajustement de biais. Ces simulations, qui proviennent soit d'un modèle global du climat (MGC) (SIM-01 à SIM-09) soit d'un modèle régional du climat (MRC) (SIM-10), sont généralement

dépourvues d'IP car les éléments physiques pertinents sont précisément ciblés pour être prescrits dans les modèles climatiques.

Cette étude-ci ajoute de la valeur par rapport à l'étude de Grenier (2018), puisque 18 TAB différentes, impliquant entre autres des procédures multivariées, ont été analysées afin d'offrir une perspective sur le sort de la cohérence physique dans le contexte de l'ajustement de biais (l'étude de Grenier (2018) était concentrée sur une seule technique, univariée de surcroît). De ce fait, une grande difficulté est imposée, soit celle de croiser 3 options différentes pour notre algorithme d'ajustement de biais : 1) une option pour le groupement des jours de l'année ('annuel', 'mensuel' ou 'fenêtre mobile') pour réaliser l'ajustement de biais ; 2) une option de préservation ou non de la tendance à long terme ('none' ou 'loess') de la simulation climatique ; et 3) une option de TAB qui regroupe une technique univariée utilisant le QM (nommée 'QM-only') et deux techniques multivariées, l'une utilisant la procédure de brassage des éléments (*shuffling*) expliquée dans Vrac (2018) et le QM (désignée par 'QM-shuf'), et l'autre utilisant la procédure de transformation géométrique par vecteurs propres (*eigenvectors*) présentée par Hnilica et al. (2017) et le QM (désignée par 'eig-QM'). Au total, 18 (3 x 2 x 3) techniques alternatives d'ajustement de biais, qui sont des combinaisons de procédures reflétant les choix généralement effectués par les praticiens de l'ajustement de biais dans les centres de services climatiques, sont appliquées aux 10 simulations climatiques et aux 12 emplacements.

4.3 Vérifications systématiques

L'algorithme d'ajustement de biais est très complexe, notamment à cause de l'imbrication des différentes procédures employées dans ce mémoire. Conséquemment, le risque de commettre des erreurs est fortement accru. Aussi, l'effet

de synergie des différentes procédures imbriquées peut mener à des artefacts difficilement prévisibles. De ce fait, une vigilance perpétuelle est de mise, ce qui nécessite, entre autres, de faire des vérifications systématiques. Par exemple, il faut s'assurer que l'algorithme ne détériore pas les statistiques « élémentaires », qui ne sont pas forcément visées pour être ajustées par ces techniques mais qui peuvent subir des effets collatéraux à l'ajustement de biais. L'étape des vérifications systématiques ne devrait pas être vu comme une perte de temps, mais plutôt comme une étape préliminaire à l'analyse des résultats visés par l'étude, pour gagner en efficacité si un problème était décelé préalablement dans le code (c'est-à-dire « perdre » du temps en amont mais pour en gagner après s'il advenait une difficulté dans le code). Dans cette section, des exemples de vérifications sur la variable de température vont être présentées. Le choix de la variable T a été fait car cette dernière a des tendances claires connues dans la littérature, il sera donc plus accessible de voir si les résultats des vérifications sont en accord avec les attentes usuelles.

La vérification du cycle annuel moyen devrait être faite systématiquement, car elle peut révéler beaucoup d'informations, dont des problèmes potentiels qui pourraient être difficiles à anticiper, ou des problèmes soupçonnés dont l'ampleur pourrait être plus grande qu'anticipé (voir les panels d, e et f de la Figure 3.6). La Figure 4.1 donne un exemple de cycle annuel moyen de T pour la ville d'Iqaluit. Le cycle annuel moyen correspond à la moyenne pour chacun des 365 jours de l'année sur 30 ans, soit 1981-2010 pour le produit de référence (CFSR, en noir) et pour les scénarios passés (en bleu), ou 2071-2100 pour les scénarios futurs (en vert). En observant la figure, on remarque que les cycles annuels moyens des 10 simulations climatiques ajustées sur 1981-2010 forment un ensemble de courbes statistiquement indiscernables du cycle annuel moyen de CFSR sur la même période. Ce résultat était attendu car, sur la période de calibration (période historique de 1981 à 2010), les propriétés statistiques du scénario s'ajustent sur celles du produit de référence (rappelons tout de même que c'est ce pourquoi on

fait de l'ajustement de biais). En revanche, si l'on regarde les 10 courbes représentant les cycles annuels moyens des 10 simulations climatiques ajustées sur la période future de 2071 à 2100, alors on ne remarque plus un seul et unique paquet. Conséquemment, on peut considérer que les valeurs de changement de chaque courbe bleue vers sa courbe verte correspondante sont relativement indépendantes entre elles, bien que pour la période de calibration, les simulations issues d'un même modèle partagent les mêmes données sur 1981-2005. Cette dissemblance significative s'explique principalement par les différences liées à la variabilité naturelle, aux différences inter-modèles dans la formulation de la physique ainsi qu'aux divers *Representative Concentration Pathways* (RCP) utilisés dans les 10 simulations climatiques de l'année 2006 à l'année 2100. Les RCP sont des trajectoires futures d'émissions et de concentrations de gaz à effet de serre (GES), utilisés comme forçage dans les modèles climatiques. Dans cette étude, les simulations 1 et 5 (SIM-01 et SIM-05) utilisent le RCP2.6, SIM-02 et SIM-06 le RCP4.5, SIM-03 et SIM-07 le RCP6.0 et les simulations 4, 8, 9 et 10 le RCP8.5, sachant que ce dernier est considéré comme le plus pessimiste (forçage anthropique global de 8.5 W/m^2 en 2100).

D'autre part, sur la Figure 4.1, on constate que le réchauffement à Iqaluit pour l'horizon 2071-2100 est très important, surtout pour les mois d'hiver. Ce résultat n'est pas surprenant, dans la mesure où Iqaluit est située dans le Nunavut (Canada), à une latitude d'environ 64°N . Dans cette région, et dans les régions situées dans les hautes latitudes en général, le réchauffement entraîne la fonte des glaciers, de la glace de mer et de la neige, ce qui entraîne une diminution de l'albédo. Plus l'albédo est faible, plus la radiation solaire est absorbée, ce qui implique encore plus de réchauffement. Il s'ensuit alors une boucle de rétroaction positive entre l'albédo et le réchauffement initial. Toutefois, cette boucle de rétroaction positive n'est potentiellement pas la boucle la plus importante pour l'hiver, étant donné le peu de lumière solaire disponible dans les régions de hautes latitudes en cette saison. Un autre mécanisme, celui de la rétroaction

de Planck, joue sans doute un rôle plus crucial dans l'amplification du réchauffement hivernal dans le nord par rapport aux plus basses latitudes. La loi de Stefan-Boltzmann établit que la puissance totale rayonnée par un corps est proportionnelle à la puissance 4 de la température absolue de ce corps. La rétroaction de Planck découle directement de cette loi. En effet, pour un même forçage par les GES, une surface froide doit élever davantage sa température qu'une surface chaude pour atteindre l'équilibre énergétique (énergie entrante = énergie sortante). Une description plus complète des boucles de rétroaction impliquées dans le phénomène d'amplification arctique peut être trouvée dans l'article de Pithan et Mauritsen (2014).

Toutefois, on remarque aussi (voir Figure 4.2) que certains scénarios futurs (en vert) peuvent être plus proches des scénarios passés (en bleu), ce qui correspond à un réchauffement moindre dans cette zone (par rapport aux autres scénarios climatiques futurs). Cela peut s'expliquer par la situation géographique d'Iqaluit. En effet, la ville d'Iqaluit est située non loin de la zone de plongée des eaux de l'Atlantique Nord (dans la mer du Labrador). À cet endroit, l'eau plonge dans les profondeurs à cause des différences de densité de l'eau et active la circulation océanique générale ou thermohaline. En contexte de réchauffement global, l'inlandsis du Groenland fond, ce qui ajoute de l'eau douce (non salée) et donc diminue la salinité des eaux de cette région, ce qui modifie la densité de l'eau et entraîne une stratification océanique plus importante. Tous ces enjeux mis bout à bout impliquent un ralentissement de la plongée des eaux, et donc de la circulation thermohaline. Cela influence aussi l'apport de chaleur du *Gulf Stream*, qui à son tour, est diminué. On observe alors une zone de refroidissement dans la mer du Labrador, entourée de zones où l'on remarque un réchauffement (Rahmstorf et al., 2015 ; Sgubin et al., 2017). Cet « œil » de refroidissement ou de stagnation des températures pourrait s'étendre parfois jusqu'à la hauteur d'Iqaluit, et c'est possiblement ce que l'on constate sur la Figure 4.2 pour le scénario futur le plus proche des scénarios passés.

La Figure 4.2 donne un exemple de série temporelle de T à Iqaluit sur la période d'application (de 1981 à 2100). La série temporelle du produit de référence (CFSR, en noir) n'est disponible que jusqu'à 2010 alors que les séries temporelles des simulations climatiques (en rouge) et des scénarios climatiques (en bleu) sont disponibles sur toute la période d'application. On remarque que, de 1981 à 2005, certaines simulations climatiques sont exactement les mêmes (il y a seulement 4 paquets de simulations climatiques alors que l'on devrait en avoir 10). Comme expliqué précédemment, les RCP sont appliqués aux simulations seulement à partir du 1er janvier 2006, c'est pourquoi certaines simulations sont semblables de 1981 à 2005 (SIM-01 à SIM-04 et SIM-05 à SIM-08) mais elles se séparent à partir de 2006. La moyenne sur 1981-2010 de la série temporelle de CFSR (diamant noir), est un point unique (car il y a seulement un produit de référence) qui correspond à -9°C . Les moyennes des 10 simulations climatiques sur 1981-2010 s'étendent sur un intervalle allant de -15°C à -9°C (barre rouge), ce qui traduit bien la présence de biais initiaux dans les simulations par rapport au produit de référence. En revanche, les moyennes des 10 scénarios climatiques sur 1981-2010 forment un point unique (« barre » bleue) qui correspond exactement à la moyenne de CFSR sur 1981-2010. Cela montre bien comment les scénarios climatiques sont construits, c'est-à-dire en amalgamant les contenus jugés informatifs du produit de référence sur la période de calibration historique (1981-2010) et ceux de la simulation climatique sur la période future 2011-2100. En effet, les scénarios climatiques s'alignent sur CFSR pour les statistiques distributionnelles sur 1981-2010, ce qui est traduit par la correspondance de la barre bleue avec le diamant noir, et elles suivent les valeurs de changements simulés de la simulation climatique (ici c'est l'option de préservation de tendance 'loess' qui est sélectionnée). Il est important de mentionner que si les séries temporelles de la simulation climatique et du scénario climatique se suivent, alors cela signifie que la simulation initiale était peu biaisée c'est-à-dire que sa moyenne sur 1981-2010 était proche de -9°C (moyenne de CFSR).

4.4 Conclusions et recommandations

Les résultats de cette étude sont résumés dans un tableau-synthèse (voir Tableau 3.3) qui caractérise les effets généraux de l'ajustement des biais à travers les 10 simulations, les 12 sites d'étude et les 4 variables. Chacune des 18 alternatives testées conduit à des résultats différents selon les propriétés choisies (propriétés statistiques de biais ou IP). Les couleurs foncées représentent les résultats attendus *a priori*, ce qui signifie que l'alternative d'ajustement de biais correspondante était censée garantir une amélioration (vert foncé) ou une détérioration (rouge foncé) de la propriété choisie. Au contraire, les couleurs claires indiquent les résultats obtenus *a posteriori*, ce qui signifie que l'ajustement de biais peut conduire à une potentielle correction (vert clair) ou détérioration (rouge clair) incontrôlée de la propriété choisie. La couleur grise, quant à elle, indique que les résultats simulés (ligne 'raw' dans le tableau) sont essentiellement conservés après l'ajustement de biais (la simulation dicte *grosso modo* la propriété statistique). Comme le suggère le tableau-synthèse, chaque alternative a ses forces (vert foncé et vert clair) et ses faiblesses (rouge foncé et rouge clair) en fonction des propriétés que la TAB cherche à ajuster.

Cette présente étude atteste que les centres de services climatiques doivent considérer plus que la simple sophistication statistique dans leurs stratégies d'ajustement de biais, qu'elles soient univariées ou multivariées, en raison du problème d'IP. En effet, les liens physiques unissant plusieurs variables doivent être explicitement programmés dans les TAB, car les sophistications statistiques ne comportent aucune physique. Les résultats présentés dans le tableau-synthèse décrit précédemment apportent la vérification que les procédures généralement insérées dans une TAB multivariée pour ajuster les dépendances inter-variables de nature statistique, comme le brassage des éléments (*shuffling*) et la transformation géométrique par vecteurs propres (*eigenvectors*), fonctionnent pour la tâche prévue, même en conjonction avec d'autres

procédures comme la préservation des tendances ou une fenêtre mobile. Cependant, d'autres propriétés statistiques ne sont pas bien ajustées, voire peuvent être fortement dégradées, comme la variabilité interannuelle des moyennes mensuelles et les moyennes mensuelles elles-mêmes. À propos, la procédure de *shuffling* devrait être évitée autant que possible dans la mesure où les séquences sont détruites (météorologie et systèmes synoptiques). Cependant, il faut souligner qu'il pourrait y avoir des fins pour lesquelles cette procédure est vraiment adéquate, comme lorsqu'une désagrégation spatiale est nécessaire pour les précipitations, en lien avec le problème de l'inflation, par lequel diverses sous-parties d'une tuile de modèle grossière peuvent connaître leurs extrêmes de précipitations en même temps, ce qui peut mener à des conséquences hydrologiques inadéquates (par exemple, Su et al., 2020). Par ailleurs, le *shuffling* et la transformation géométrique par vecteurs propres, tout comme la procédure d'ajustement de biais univariée utilisée ici (QM), ne garantissent pas la préservation de certains éléments physiques sur lesquels les modèles climatiques sont basés (cohérence physique), comme ici la relation thermodynamique entre T , P , q et RH.

À la lumière des résultats de cette étude, plusieurs recommandations finales peuvent être données. La première est d'incorporer des fonctions, comme la logit ou le logarithme dans le but de borner les distributions de chaque variable pour un traitement uniforme de la fonction de transfert en mode additif. De plus, cela conduit aussi à éviter la génération d'IP de type 1 pour les variables directement ajustées (par exemple, des valeurs d'HR supérieures à 100 % sont alors évitées par construction). Cette procédure est dans l'esprit de celle utilisée par Cannon (2016), et nous plaidons ici pour un cadre raisonné encore plus large, qui favorise l'uniformité dans le traitement des variables entre les différents centres de services climatiques.

Dans un deuxième temps, nous recommandons de choisir une alternative d'ajustement de biais dans l'esprit du concept de « *adequacy-for-purpose* ». Cela signifie que le

choix de l'alternative doit être fait en fonction des propriétés statistiques que l'on souhaite ajuster et de l'objectif final de l'étude. À cette fin, le tableau-synthèse (voir Tableau 3.3) pourrait être un outil pour éclairer ce genre de choix, par lequel une technique n'est pas considérée bonne ou mauvaise en elle-même, mais plutôt adéquate ou inadéquate *pour un objectif donné*. Notre quatrième recommandation, qui est détaillée dans la section précédente, est de faire des vérifications systématiques.

La dernière recommandation est, lorsqu'une relation inter-variable de nature physique est identifiée, d'ajuster directement des variables considérées comme indépendantes, puis de calculer subséquemment les autres, considérées comme dépendantes (autant en mode univarié que multivarié), dans le but d'empêcher la génération d'IP de type 2. Cela signifie que la relation physique entre les variables doit être explicitement prise en compte dans la stratégie d'ajustement de biais (la stratégie est plus large que la technique elle-même, dans le sens où l'on réfléchit à l'emploi de la technique d'une manière à ne pas créer d'IP). Dans notre cas (4 variables unies par 1 relation thermodynamique), plusieurs possibilités sont envisageables pour le choix de la variable déduite, mais pour des raisons pratiques, le choix de la variable dépendante s'est fait ici entre les deux variables d'humidité (q ou RH). Or, grâce à l'emploi de la fonction logit pour RH directement ajustée, la génération de cas où $RH > 100\%$ est empêchée par construction. De ce fait, il est plus judicieux d'ajuster directement T , P et RH et de calculer q ensuite, car dans ce cas, les deux types d'IP sont évités. Idéalement, cela serait fait pour q même si T , P ou RH n'est pas fourni, pour s'adapter plus facilement aux changements ultérieurs de la portée de l'étude et favoriser l'uniformité entre les projets dans les méthodes et les résultats. La prise en compte du lien physique unissant les variables dans l'algorithme implique de concéder un biais résiduel dans la variable calculée après coup (q) plutôt que d'avoir des formes claires d'IP. En effet, la question du choix d'avoir un problème d'IP ou d'avoir du biais résiduel avec le produit de référence sélectionné ne devrait pas se poser, car les IP sont

inacceptables en toutes circonstances en raison de leur définition physique. De plus, le biais résiduel est, en général, beaucoup plus petit que le biais initial (Trasher et al., 2012 ; Grenier, 2018), ce qui peut contribuer à l'acceptation de cette concession.

La Figure 4.3 compare le biais initial et les biais finaux observés sur la période de calibration (1981-2010) pour la variable d'humidité spécifique (q) à Iqaluit, à l'aide de la statistique de Kolmogorov-Smirnov (D_{KS}). Le biais initial, représenté par les 10 lignes rouges, correspond aux biais entre le produit de référence (CFSR) et les simulations climatiques. Le biais final en bleu correspond aux biais entre q_{cfsr} et q directement ajustée (q_{ba}) et celui en cyan correspond aux biais entre q_{cfsr} et q calculée après coup à l'aide des 3 autres variables directement ajustées (q_{so90}). La statistique de Kolmogorov-Smirnov (D_{KS}) détermine à quel point deux distributions sont similaires. Plus la D_{KS} est proche de 0, plus les deux distributions sont considérées comme similaires alors que plus la D_{KS} est proche de 1, plus elles sont considérées comme distinctes. De ce fait, à $D_{KS} = 0$, on peut dire que les deux distributions sont identiques tandis qu'à $D_{KS} = 1$, elles ne se chevauchent même pas. Comme anticipé, on remarque que c'est les D_{KS} en rouge qui sont le plus loin de 0 (relativement loin pour les 10 courbes) ce qui signifie que les biais initiaux entre le produit de référence et les simulations climatiques étaient importants. Si l'on se penche sur les D_{KS} en bleu, alors on remarque que le biais final est presque nul, ce qui signifie que les TAB ont bien effectué la tâche pour laquelle elles sont appliquées aux simulations climatiques, c'est-à-dire d'ajuster les biais avec le produit de référence (bien que l'ajustement ne soit pas parfait lorsqu'évalué avec cette métrique D_{KS} , dû aux procédures auxiliaires). En effet, pour la ville d'Iqaluit, les biais sont améliorés (par rapport aux biais initiaux) quand q est directement ajusté (q_{ba}) pour 116 cas sur 120 (12 mois x 10 simulations), ce qui représente environ 96,67 % des cas. Enfin, les D_{KS} en cyan, qui représentent le biais final dans la variable calculée après coup q_{so90} , sont plus importants que ceux dans le cas d'un ajustement direct de q (q_{ba} , en bleu), pour 81,67 % des cas (98/120). De ce

fait, si l'on calcule l'une des variables à partir des 3 autres directement ajustées, alors le biais résiduel est plus important que celui obtenu après ajustement direct de cette même variable. En revanche, le biais final dans la variable calculée après coup est, dans 94,17 % des cas (113/120), inférieur aux biais initiaux présents dans les simulations climatiques. La situation pour la ville d'Iqaluit représente bien celle des autres villes (vérifié mais non montré).

Dans un contexte de démocratisation de l'accès aux données climatiques, des systèmes ouverts comme les plateformes PAVICS (pôle d'analyse et de visualisation de l'information climatique et scientifique) d'Ouranos ou Copernicus de l'Union Européenne sont développés. Conséquemment, des non-experts peuvent soudainement avoir un accès facilité à l'information, ce qui est, selon moi, une bonne chose. Cependant, il n'en est pas moins préjudiciable quand l'utilisateur n'a pas toutes les clefs nécessaires pour comprendre, utiliser et interpréter l'information climatique qu'il a devant les yeux. En ce sens, il est à espérer que ce mémoire contribuera à clarifier, dans la littérature scientifique, la notion d'IP dans le contexte de l'ajustement de biais. Toutefois, il est important de rappeler que même si notre définition de l'IP nous semble pertinente puisqu'elle s'inscrit dans des notions fondamentales de la philosophie traditionnelle des sciences, elle n'en demeure pas moins subjective.

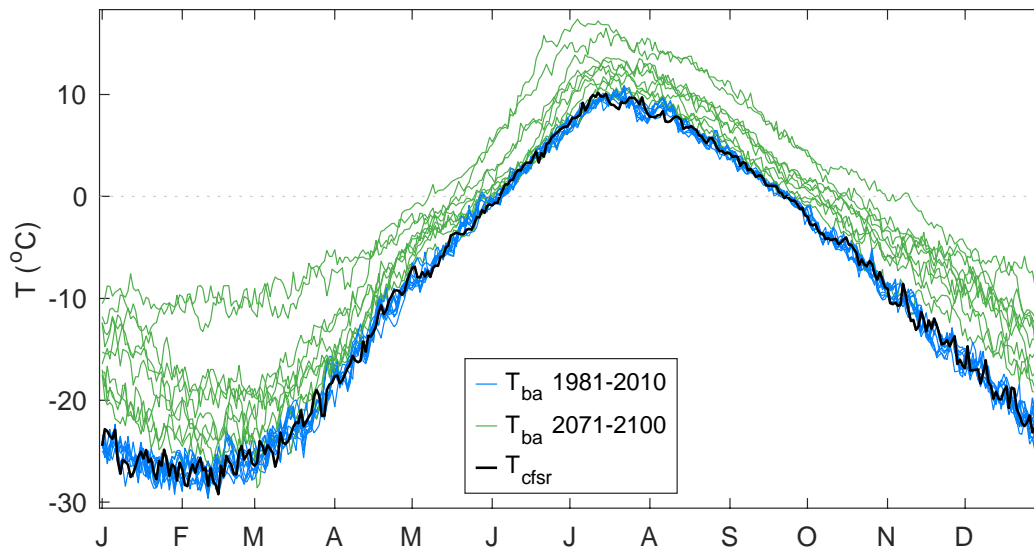


Figure 4.1 : Cycle annuel moyen de la température (T en $^{\circ}\text{C}$) à Iqaluit. La ligne noire (T_{cfsr}) représente le cycle annuel moyen du produit de référence (CFSR) sur la période de calibration historique (1981-2010). Les 10 lignes bleues (T_{ba} 1981-2010) correspondent à celui des scénarios climatiques passés (sur 1981-2010) alors que les 10 lignes vertes (T_{ba} 2071-2100) représentent le cycle annuel moyen des scénarios climatiques futurs, sur la période 2071-2100. Les scénarios sont produits par l'alternative correspondant aux options 'loess', 'fenêtre mobile' et 'QM-only'.

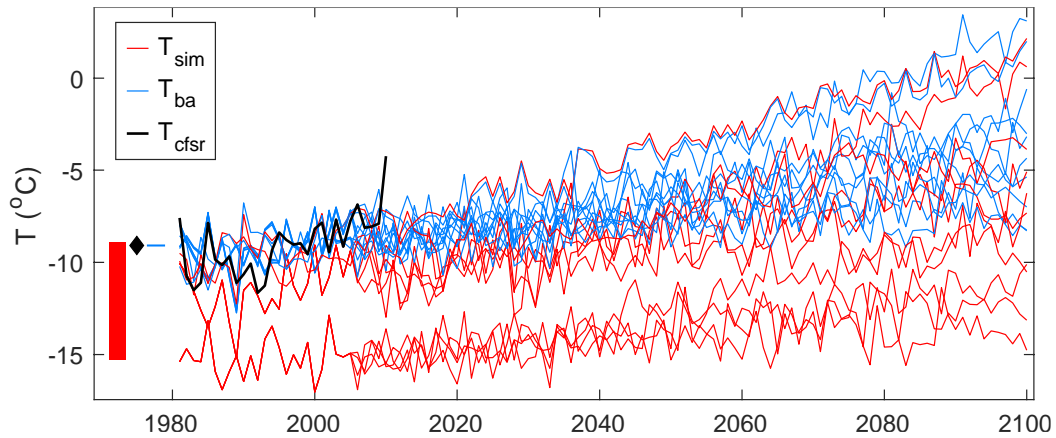


Figure 4.2 : Série temporelle de la température (T en °C) à Iqaluit sur la période d'application (1981-2100). La ligne noire (T_{cfsr}) représente la série temporelle du produit de référence (CFSR) sur la période de calibration historique (1981-2010) seulement (il n'existe évidemment pas de produit de référence futur). Les 10 lignes rouges correspondent aux séries temporelles des 10 simulations climatiques (T_{sim}) alors que les 10 lignes bleues représentent celles des scénarios climatiques (T_{ba}). Le diamant noir représente la moyenne sur 1981-2010 de la série temporelle de CFSR. La barre rouge (bleue) représente l'intervalle, le long de l'axe des y, des moyennes sur 1981-2010 pour les 10 simulations climatiques (scénarios climatiques). Les scénarios sont produits par l'alternative correspondant aux options 'loess', 'fenêtre mobile' et 'QM-only'.

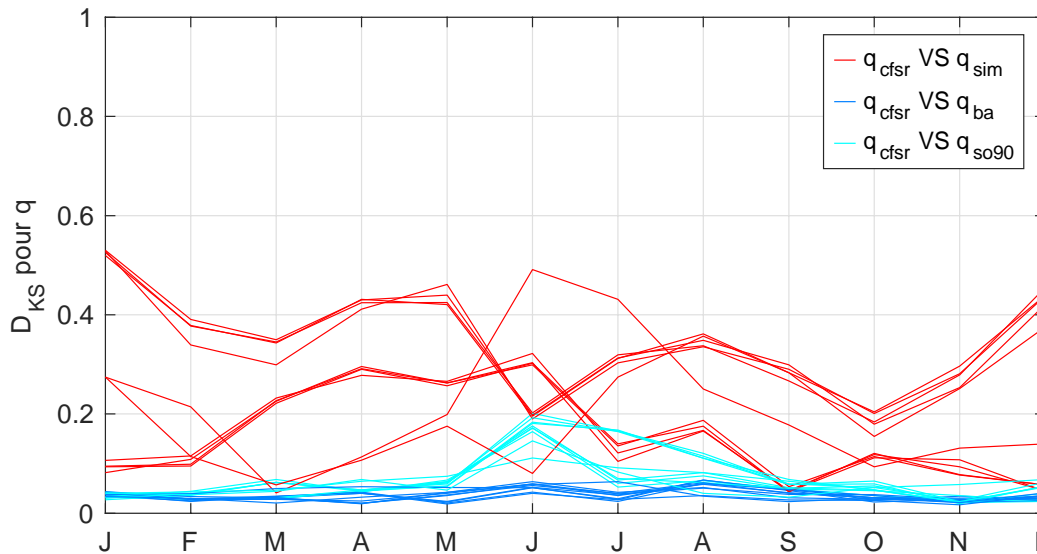


Figure 4.3 : Statistique de Kolmogorov-Smirnov (D_{KS}) entre d'une part q_{cfsr} (l'humidité spécifique q du produit de référence) et d'autre part q_{sim} (q de la simulation climatique, en rouge), q_{ba} (q directement ajusté, en bleu) ou q_{so90} (q déduit des 3 autres variables, en cyan), pour chacune des 10 simulations. Ce cas est celui de la ville d'Iqaluit pour les scénarios produits par l'alternative univariée 'loess', 'fenêtre mobile' et 'QM-only'. Pour une meilleure appréciation visuelle, les lignes connectent les résultats des différents mois pour une même simulation.

RÉFÉRENCES

Abatzoglou J. T. and Brown T. J. (2012). A comparison of statistical downscaling methods suited for wildfire applications. *International Journal of Climatology*, 32 (5), p. 772-780, doi: 10.1002/joc.2312.

Addor N. and Fischer E. M. (2015). The influence of natural variability and interpolation errors on bias characterization in RCM simulations. *Journal of Geophysical Research: Atmospheres*, 120 (19), p. 10180-10195, doi: 10.1002/2014JD022824.

Agbazo M. N. and Grenier P. (2019). Characterizing and avoiding physical inconsistency generated by the application of univariate quantile mapping on daily minimum and maximum temperatures over Hudson Bay. *International Journal of Climatology*, 40 (8), p. 3868-3884, doi: 10.1002/joc.6432.

Arora V. K. and Coauthors (2011). Carbon emission limits required to satisfy future representative concentration pathways of greenhouse gases. *Geophysical Research Letters*, 38 (5), doi: 10.1029/2010GL046270.

Boé J., Terray L., Habets F. and Martin E. (2007). Statistical and dynamical downscaling of the Seine basin climate for hydrometeorological studies. *International Journal of Climatology*, 27 (12), p. 1643-1655, doi: 10.1002/joc.1602.

Cannon A. J., Sobie S. R. and Murdock T. Q. (2015). Bias Correction of GCM Precipitation by Quantile Mapping: How Well Do Methods Preserve Changes in Quantiles and Extremes? *Journal of Climate*, 28 (17), p. 6938-6959, doi: 10.1175/JCLI-D-14-00754.1.

Cannon A. J. (2016). Multivariate bias correction of climate model output: Matching marginal distributions and intervariable dependence structure. *Journal of Climate*, 29 (19), p. 7045-7064, doi: 10.1175/JCLI-D-15-0679.1.

Cannon A. J. (2018). Multivariate quantile mapping bias correction: An N -dimensional probability density function transform for climate model simulations of multiple variables. *Climate dynamics*, 50 (1), p. 31-49, doi: 10.1007/s00382-017-3580-6.

Chen J., St-Denis B. G., Brissette F. and Lucas-Picher P. (2016). Using Natural Variability as a Baseline to Evaluate the Performance of Bias Correction Methods in Hydrological Climate Change Impact Studies. *Journal of Hydrometeorology*, 17 (8), p. 2155-2174, doi: 10.1175/JHM-D-15-0099.1.

Cleveland W. S. (1979). Robust Locally Weighted Regression and Smoothing Scatterplots. *Journal of the American Statistical Association*, 74 (368), p. 829-836.

Del Genio A.D., Yao M.-S., Kovari W. and Lo K. K.-W. (1996). A prognostic cloud water parameterization for global climate models. *Journal of Climate*, 9 (2), p. 270-304, doi: 10.1175/1520-0442(1996)009<0270:APCWPF>2.0.CO:2.

Déqué M. (2007). Frequency of precipitation and temperature extremes over France in an anthropogenic scenario: Model results and statistical correction according to observed values. *Global and Planetary Change*, 57 (1-2), p. 16-26, doi: 10.1016/j.gloplacha.2006.11.030.

Diaconescu E. P., Mailhot A., Brown R. and Chaumont D. (2017). Evaluation of CORDEX-Arctic daily precipitation and temperature-based climate indices over Canadian Arctic land areas. *Climate Dynamics*, 50 (5), p. 2061-2085, doi: 10.1007/s00382-017-3736-4.

Dunne J.P. and Coauthors (2012). GFDL's ESM2 global coupled climate-carbon earth system models. Part I: Physical formulation and baseline simulation characteristics. *Journal of Climate*, 25 (19), p. 6646-6665, doi: 10.1175/JCLI-D-11-00560.1.

Ehret U., Zehe E., Wulfmeyer V., Warrach-Sagi K. and Liebert J. (2012). Should we apply bias correction to global and regional climate model data? *Hydrology and Earth System Sciences*, 16 (9), p. 3391-3404, doi: 10.5194/hess-16-3391-2012.

Fang G. H., Yang J., Chen Y. N. and Zammit C. (2015). Comparing bias correction methods in downscaling meteorological variables for a hydrologic impact study in an arid area in China. *Hydrology and Earth System Sciences*, 19 (6), p. 2547-2559, doi: 10.5194/hess-19-2547-2015.

Findlater K., Webber S., Kandlikar M. and Donner S. (2021). Climate services promise better decisions but mainly focus on better data. *Nature Climate Change*, 11, p. 731-737, doi: 10.1038/s41558-021-01125-3.

Flato G. and Coauthors (2013). Evaluation of climate models. In *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, p. 741-882, doi: 10.1017/CBO9781107415324.020.

François B., Vrac M., Cannon A. J., Robin Y. and Allard D. (2020). Multivariate bias corrections of climate simulations: which benefits for which losses? *Earth System Dynamics*, 11 (2), p. 537-562, doi: 10.5194/esd-11-537-2020.

Frank C. W., Wahl S., Keller J. D., Pospichal B., Hense A. and Crewell S. (2018). Bias correction of a novel European reanalysis data set for solar energy applications. *Solar Energy*, 164, p. 12-24, doi: 10.1016/j.solener.2018.02.012.

Galmarini S. and Coauthors (2019). Adjusting climate model bias for agricultural impact assessment: How to cut the mustard. *Climate services*, 13, p. 65-69, doi: 10.1016/j.cliser.2019.01.004.

Gennaretti F., Sangelantoni L. and Grenier P. (2015). Toward daily climate scenarios for Canadian Arctic coastal zones with more realistic temperature-precipitation interdependence. *Journal of Geophysical Research: Atmospheres*, 120 (23), p. 11862-11877, doi: 10.1002/2015JD023890.

Giere R. N. (2010). An agent-based conception of models and scientific representation. *Synthese*, 172 (2), p. 269-281, doi: 10.1007/s11229-009-9506-z.

Giorgi F. (2019). Thirty years of regional climate modeling: where are we and where are we going next? *Journal of Geophysical Research: Atmospheres*, 124 (11), p. 5696-5723, doi: 10.1029/2018JD030094.

Gobiet A., Suklitsch M. and Heinrich G. (2015). The effect of empirical-statistical correction of intensity-dependent model errors on the temperature climate change signal. *Hydrology and Earth System Sciences*, 19 (10), p. 4055-4066, doi: 10.5194/hess-19-4055-2015.

Gómez-Navarro J. J., Raible C. C., Bozhinova D., Martius O., García Valero J. A. and Montávez J. P. (2018). A new region-aware bias-correction method for simulated precipitation in areas of complex orography. *Geoscientific Model Development*, 11 (6), p. 2231-2247, doi: 10.5194/gmd-11-2231-2018.

Grenier P. (2018). Two types of physical inconsistency to avoid with univariate quantile mapping: a case study over North America concerning relative humidity and its parent variables. *Journal of Applied Meteorology and Climatology*. 57 (2), p. 347-364, doi: 10.1175/JAMC-D-17-0177.1.

Haerter J. O., Hagemann S., Moseley C. and Piani C. (2011). Climate model bias correction and the role of timescales. *Hydrology and Earth System Sciences*, 15 (3), p. 1065-1079, doi: 10.5194/hess-15-1065-2011.

Hay L. E. and Clark M. P. (2003). Use of statistically and dynamically downscaled atmospheric model output for hydrologic simulations in three mountainous basins in the western United States. *Journal of Hydrology*, 282 (1-4), p. 56-75, doi: 10.1016/S0022-1694(03)00252-X.

Hecht A. D. (1984). Meeting the challenge of climate service in the 1980s. *Bulletin of the American Meteorological Society*, 65 (4), p. 365-366, doi: 10.1175/1520-0477(1984)065<0365:MTCOCS>2.0.CO:2.

Hempel S., Frieler K., Warszawski L., Schewe J. and Piontek F. (2013). A trend-preserving bias correction - the ISI-MIP approach. *Earth System Dynamics*, 4 (2), p. 219-236, doi: 10.5194/esd-4-219-2013.

Hewitson B. C., Daron J., Crane R. G., Zermoglio M. F. and Jack C. (2014). Interrogating empirical-statistical downscaling. *Climatic Change*, 122 (4), p. 539-554, doi: 10.1007/s10584-013-1021-z.

Hewitt C. D. and Coauthors (2021). Recommendations for Future Research Priorities for Climate Modeling and Climate Services. *Bulletin of the American Meteorological Society*, 102 (3), p. E578-E588, doi: 10.1175/BAMS-D-20-0103.1.

Hnilica J., Hanel M. and Puš V. (2017). Multisite bias correction of precipitation data from regional climate models. *International Journal of Climatology*, 37 (6), p. 2934-2946, doi: 10.1002/joc.4890.

Jacobson M. Z. (2005). *Fundamentals of atmospheric modelling (Second Edition)*. Cambridge University Press.

Kotlarski S. and Coauthors (2019). Observational uncertainty and regional climate model evaluation: A pan-European perspective. *International Journal of Climatology*, 39 (9), p. 3730-3749, doi: 10.1002/joc.5249.

Kuhn T. (1977). Objectivity, value judgment, and theory choice. In *The Essential Tension: Selected studies in scientific tradition and change*, University of Chicago press, p. 320-339.

Lange S. (2019). Trend-preserving bias adjustment and statistical downscaling with ISIMIP3BASD (v1.0). *Geoscientific Model Development*, 12 (7), p. 3055-3070, doi: 10.5194/gmd-12-3055-2019.

Lanzante J. R., Dixon K. W., Nath M. J., Whitlock J. E. and Adams-Smith D. (2018). Some pitfalls in statistical downscaling of future climate. *Bulletin of the American Meteorological Society*, 99 (4), p. 791-803, doi: 10.1175/BAMS-D-17-0046.1.

Laprise R. (2008). Regional climate modelling. *Journal of Computational Physics*, 227 (7), p. 3641-3666, doi: 10.1016/j.jcp.2006.10.024.

Laudan L. (1984). Kuhn's critique of methodology. In: *Science and values: The aims of science and their role in scientific debate*, University of California Press, p. 87-102.

Lehtonen I., Kämäräinen M., Gregow H., Venäläinen A. and Peltola H. (2016). Heavy snow loads in Finnish forests respond regionally asymmetrically to projected climate change. *Natural Hazards and Earth System Sciences*, 16 (10), p. 2259-2271, doi: 10.5194/nhess-16-2259-2016.

Lenhard J. and Winsberg E. (2010). Holism, Entrenchment, and the Future of Climate Model Pluralism. *Studies in History and Philosophy of Science*, 41 (3), p. 253-262, doi: 10.1016/j.shpsb.2010.07.001.

Lorenz, E. N. (1963). Deterministic nonperiodic flow. *Journal of Atmospheric Sciences*, 20 (2), p. 130-141, doi: 10.1175/1520-0469(1963)020<0130:DNF>2.0.CO:2.

Lourenço T., Swart R., Goosen H. and Street R. (2016). The rise of demand-driven climate services. *Nature Climate Change*, 6 (1), p. 13-14, doi: 10.1038/nclimate2836.

Lugen M. (2020). Framing Climate Services: Logics, Actors, and Implications for Policies and Projects. *Atmosphere*, 11 (10), p. 1047, doi: 10.3390/atmos11101047.

Maraun D. (2016). Bias correcting climate change simulations - A critical review. *Current Climate Change Reports*, 2 (4), p. 211-220, doi: 10.1007/s40641-016-0050-x.

Maraun D. and Coauthors (2017). Towards process-informed bias correction of climate change simulations. *Nature Climate Change*, 7 (11), p. 764-773, doi: 10.1038/nclimate3418.

Martins J., Fraga H., Fonseca A. and Santos J. A. (2021). Climate Projections for Precipitation and Temperature Indicators in the Douro Wine Region: The Importance of Bias Correction. *Agronomy*, 11 (5), p. 990, doi: 10.3390/agronomy11050990.

Mote P., Brekke L., Duffy P. B. and Maurer E. (2011). Guidelines for constructing climate scenarios. *Eos, Transactions American Geophysical Union*, 92 (31), p. 257-258, doi: 10.1029/2011EO310001.

Mpelasoka F. S. and Chiew F. H. (2009). Influence of rainfall scenario construction methods on runoff projections. *Journal of Hydrometeorology*, 10 (5), p. 1168-1183, doi: 10.1175/2009JHM1045.1.

- Parker W. S. (2009). Confirmation and Adequacy-for-Purpose in Climate Modeling. *Proceedings of the Aristotelian Society*, Supplementary Volume 83 (1), p. 233-249, doi: 10.1111/j.1467-8349.2009.00180.x.
- Parker W. S. (2020). Model Evaluation: An Adequacy-for-Purpose View. *Philosophy of Science*, 87 (3), p. 457-477, doi: 10.1086/708691.
- Peel M. C., Finlayson B. L. and McMahon T. A. (2007). Updated world map of the Köppen-Geiger climate classification. *Hydrology and Earth System Sciences*, 11 (5), p. 1633-1644, doi: 10.5194/hess-11-1633-2007.
- Petersen A. C. (2000). Philosophy of climate science. *Bulletin of the American Meteorological Society*, 81 (2), p. 265-272, doi: 10.1175/1520-0477(2000)081<0265:POCS>2.3.CO;2.
- Piani C., Haerter J. O. and Coppola E. (2010). Statistical bias correction for daily precipitation in regional climate models over Europe. *Theoretical and Applied Climatology*, 99 (1), p. 187-192, doi: 10.1007/s00704-009-0134-9.
- Piani C. and Haerter J. O. (2012). Two dimensional bias correction of temperature and precipitation copulas in climate models. *Geophysical Research Letters*, 39 (20), doi: 10.1029/2012GL053839.
- Pithan F. and Mauritsen T. (2014). Arctic amplification dominated by temperature feedbacks in contemporary climate models. *Nature geoscience*, 7 (3), p. 181-184.
- Pruppacher H. R. and Klett J. D. (1997). Microphysics of clouds and precipitation (Second Edition). *Kluwer academic Publishers*, p. 954.
- Quintana Seguí P., Ribes A., Martin E., Habets F. and Boé J. (2010). Comparison of three downscaling methods in simulating the impact of climate change on the hydrology of Mediterranean basins. *Journal of Hydrology*, 383 (1-2), p. 111-124, doi: 10.1016/j.jhydrol.2009.09.050.
- 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. *Nature climate change*, 5 (5), p. 475-480, doi: 10.1038/nclimate2554.
- Rajczak J., Kotlarski S., Salzmann N. and Schär C. (2016). Robust climate scenarios for sites with sparse observations: a two-step bias correction approach. *International Journal of Climatology*, 36 (3), p. 1226-1243, doi: 10.1002/joc.4417.

Reiter P., Gutjahr O., Schefczyk L., Heinemann G. and Casper M. (2016). Bias correction of ENSEMBLES precipitation data with focus on the effect of the length of the calibration period. *Meteorologische Zeitschrift*, p. 85-96, doi: 10.1127/metz/2015/0714.

Robin Y., Vrac M., Naveau P. and Yiou P. (2019). Multivariate stochastic bias corrections with optimal transport. *Hydrology and Earth System Sciences*, 23 (2), p. 773-786, doi: 10.5194/hess-23-773-2019.

Rocheta, E., Evans J. P. and Sharma A. (2014). Assessing atmospheric bias correction for dynamical consistency using potential vorticity. *Environmental Research Letters*, 9 (12), doi: 10.1088/1748-9326/9/12/124010.

Rockel B. (2015). The regional downscaling approach: a brief history and recent advances. *Current Climate Change Reports*, 1 (1), p. 22-29, doi: 10.1007/s40641-014-0001-3.

Ruane A. C., Goldberg R. and Chryssanthacopoulos J. (2015). Climate forcing datasets for agricultural modeling: Merged products for gap-filling and historical climate series estimation. *Agricultural and Forest Meteorology*, 200, p. 233-248, doi: 10.1016/j.agrformet.2014.09.016.

Ruosteenoja K., Jylhä K., Räisänen J. and Mäkelä A. (2017). Surface air relative humidities spuriously exceeding 100% in CMIP5 model output and their impact on future projections. *Journal of Geophysical Research: Atmospheres*, 122 (18), p. 9557-9568, doi:10.1002/2017JD026909.

Saha S. and Coauthors (2010). The NCEP Climate Forecast System Reanalysis. *Bulletin of the American Meteorological Society*, 91 (8), p. 1015-1058, doi: 10.1175/2010BAMS3001.1.

Šeparović L., Alexandru A., Laprise R., Martynov A., Sushama L., Winger K., Tete K. and Valin M. (2013). Present climate and climate change over North America as simulated by the fifth-generation Canadian regional climate model. *Climate Dynamics*, 41 (11), p. 3167-3201, doi: 10.1007/s00382-013-1737-5.

Sgubin G., Swingedouw D., Drijfhout S., Mary Y. and Bennabi A. (2017). Abrupt cooling over the North Atlantic in modern climate models. *Nature Communications*, 8 (1), p. 1-12, doi: 10.1038/ncomms14375.

Sippel S., Otto F. E. L., Forkel M., Allen M. R., Guillod B. P., Heimann M., Reichstein M., Seneviratne S. I., Thonicke K. and Mahecha D. (2016). A novel bias correction methodology for climate impact simulations. *Earth System Dynamics*, 7 (1), p. 71-88, doi: 10.5194/esd-7-71-2016.

Sonntag D. (1990). Important new values of the physical constants of 1986, vapor pressure formulations based on ITS-90 and psychrometer formulae. *Zeitschrift für Meteorologie*, 40 (5), p. 340-344.

Steel D. (2010). Epistemic values and the argument from inductive risk. *Philosophy of Science*, 77 (1), p. 14-34, doi: 10.1086/650206.

Su Y.-F., Cheng C.-T., Liou J.-J., Chen Y.-M. and Kitoh A. (2016). Bias correction of MRI-WRF dynamic downscaling datasets. *TAO: Terrestrial, Atmospheric and Oceanic Sciences*, 27 (5), p. 649-657, doi: 10.3319/TAO.2016.07.14.01.

Su T., Chen J., Cannon A. J., Xie P. and Guo Q. (2020). Multi-site bias correction of climate model outputs for hydro-meteorological impact studies: An application over a watershed in China. *Hydrological Processes*, 34 (11), p. 2575-2598, doi: 10.1002/hyp.13750.

Taylor K. E., Stouffer R. J. and Meehl G. A. (2012). An Overview of CMIP5 and the Experiment Design. *Bulletin of the American Meteorological Society*, 93 (4), p. 485-498, doi: 10.1175/BAMS-D-11-00094.1.

Themeßl M. J., Gobiet A. and Heinrich G. (2012). Empirical-statistical downscaling and error correction of regional climate models and its impact on the climate change signal. *Climatic Change*, 112 (2), p. 449-468, doi: 10.1007/s10584-011-0224-4.

Thrasher B., Maurer E. P., McKellar C. and Duffy P. B. (2012). Technical Note: Bias correcting climate model simulated daily temperature extremes with quantile mapping. *Hydrology and Earth System Sciences*, 16 (9), p. 3309-3314, doi: 10.5194/hess-16-3309-2012.

Von Storch H. (1999). On the use of "inflation" in statistical downscaling. *Journal of Climate*, 12 (12), p. 3505-3506, doi: 10.1175/1520-0442(1999)012<3505:OTUOII>2.0.CO:2.

Vrac M. and Friederichs P. (2015). Multivariate - Intervariable, Spatial, and Temporal - Bias Correction. *Journal of Climate*, 28 (1), p. 218-237, doi: 10.1175/JCLI-D-14-00059.1.

Vrac M. (2018). Multivariate bias adjustment of high-dimensional climate simulations: the Rank Resampling for Distributions and Dependences (R²D²) bias correction. *Hydrology and Earth System Sciences*, 22 (6), p. 3175-3196, doi: 10.5194/hess-22-3175-2018.

Wallace J. M. and Hobbes P. V. (2006). Atmospheric Science: An introductory Survey (Second Edition). *Elsevier Academic Press*, 92, p. 483.

Watanabe M. and Coauthors (2010). Improved Climate Simulation by MIROC5: Mean States, variability, and Climate Sensitivity. *Journal of Climate*, 23 (23), p. 6312-6335, doi: 10.1175/2010JCLI3679.1.

Wilcke R. A. I., Mendlik T. and Gobiet A. (2013). Multi-variable error correction of regional climate models. *Climatic Change*, 120 (4), p. 871-887, doi: 10.1007/s10584-013-0845-x.

Winsberg E. (2018). Philosophy and climate science. *Cambridge University Press*, doi: 10.1017/9781108164290.

Wood A. W., Maurer E. P., Kumar A. and Lettenrnaier D. P. (2002). Long-range experimental hydrologic forecasting for the eastern United States. *Journal of Geophysical Research: Atmospheres*, 107 (20), p. 1-15, doi: 10.1029/2001JD000659.

Yukimoto S. and Coauthors (2012). A New Global Climate Model of the Meteorological Research Institute: MRI-CGCM3 - Model Description and Basic Performance. *Journal of the Meteorological Society of Japan*, 90, p. 23-64, doi: 10.2151/jmsj.2012-A02.

Zscheischler J., Fischer E. M. and Lange S. (2019). The effect of univariate bias adjustment on multivariate hazard estimates. *Earth System Dynamics*, 10 (1), p. 31-43, doi: 10.5194/esd-10-31-2019.