Université du Québec
Institut national de la recherche scientifique
Centre Eau Terre Environnement

MÉTHODE BAYÉSIENNE DE MISE À L'ÉCHELLE (DOWNSCALING) SPATIALE

Par

Mohamed Ali Ben Alaya

Mémoire présenté pour l'obtention
du grade de Maître ès sciences (M.Sc.)
en sciences de l'eau

Jury d'évaluation

Examinateur interne	Fateh Chebana INRS-ETE
Directeur de recherche:	Taha B.M.J. Ouarda INRS-ETE
Examinateur externe	Mhamed Mesfioui UQTR

Mars 2012

REMERCIEMENTS

Je désire exprimer ma gratitude aux personnes qui d'une façon ou de l'autre m'ont aidé et encouragé durant mes travaux de recherche.

Particulièrement, à mon directeur de recherche, le professeur Taha Ouarda pour le support et la confiance qu'il a manifesté à mon égard.

À Dominique Fasbender, stagiaire postdoctoral dans l'équipe de mon directeur, pour sa collaboration.

À mes parents, mes frères, ma sœur et ma femme qui m'ont supporté durant ma maîtrise.

RÉSUMÉ

Les sorties des modèles couplés de circulation générale Atmosphère-Océan, AOGCM (Atmosphere-Ocean General Circulation Models) sont largement utilisées pour établir des projections du climat futur. Malheureusement, le maillage grossier de ces modèles ne permet pas de fournir les informations nécessaires pour les études d'impact en hydrologie ou pour les applications de modélisation ou d'adaptations. Dans ce contexte, diverses techniques de réduction d'échelle «Downscaling» (voir Benestad *et al.* 2008) ont été développées afin d'affiner les données AOGCM à des échelles plus pertinentes. Parmi ces techniques, les méthodes basées sur la régression sont couramment utilisées en raison de leur faible temps de calcul. Cependant, un inconvénient majeur de ces méthodes est que l'estimation est généralement reproduite seulement sur les sites mesurés. Dans ce mémoire, un modèle spatial bayésien, présenté dans un travail antérieur, est adapté et appliqué pour la réduction d'échelle des données AOGCM aux températures maximale et minimale quotidiennes dans la province de Québec. Le modèle est proposé en vue de contourner l'incapacité des méthodes de régression classiques de produire des estimations spatiales à des sites non jaugés. Dans cette méthode, la distribution de la moyenne à priori est estimée en utilisant les caractéristiques locales dans un modèle de régression géographique (GRM). L'approche employée repose sur un cadre bayésien afin de combiner un modèle spatial mensuel commun pour les températures minimales et maximales avec les fluctuations quotidiennes induites par les prédicteurs atmosphériques. Les produits des réanalyses sont utilisés dans ce travail afin d'évaluer le potentiel de la méthode proposée. Les résultats de l'application de cette procédure bayésienne ont été satisfaisants par rapport à une méthode classique fondée sur la régression.

TABLE DES MATIÈRES

LISTE DES TABLEAUX

LISTE DES FIGURES

1. INTRODUCTION ET REVUE DE LITTÉRATURE

Les modèles couplés de circulation générale Atmosphère-Océan, AOGCM (Atmosphere-Ocean General Circulation Models) constituent la méthode la plus efficace pour simuler les caractéristiques à grande échelle du climat de la terre. Les informations fournies par ces modèles sont largement utilisées pour établir des projections du climat futur. Malheureusement, les données AOGCM sont généralement produites sur des grilles régulières avec une faible résolution horizontale autour de 2.5° de longitude et de latitude (d'environ 250 à 300 km). Ce maillage grossier n'est pas approprié pour tenir compte des effets des forçages locaux et régionaux, et ne peut donc fournir les informations nécessaires pour mesurer les impacts en hydrologie et les applications de modélisation ou d'adaptation (Grotch et MacCracken 1991; Huth et Kysely 2000). Face à ce problème, diverses techniques de réduction d'échelle «Downscaling» (voir Benestad *et al.* 2008) ont été développées afin d'affiner les données AOGCM et fournir des informations à des échelles plus pertinentes. Ces techniques peuvent être réparties en deux grandes classes (Benestad *et al.* 2008; Herrera *et al.* 2006): les méthodes dynamiques et les méthodes statistiques.

Les méthodes dynamiques utilisent des modèles climatiques régionaux (RCM) qui ont les mêmes principes de base que le AOGCM avec une résolution élevée entre 25 et 50 km, mais il ne couvre qu'une portion limitée du globe. De plus ces méthodes nécessitent des ressources informatiques et humaines relativement importantes (voir Herrera *et al.* 2006).

Les méthodes statistiques considèrent une relation statistique entre les variables à grande échelle (le prédicteur) et les variables à petite échelle (le predictant). Un des avantages premiers de ces méthodes est qu'elles sont plus économiques en coût de calcul, offrant donc une alternative appropriée en cas d'insuffisance des ressources pour appliquer les méthodes dynamiques (Herrera *et al.* 2006). Il existe un grand nombre de modèles et de techniques de réduction d'échelle statistique qui peuvent être regroupés en trois grandes familles : les générateurs stochastiques de séquences météorologiques (Wilks et Wilby, 1999), les types de temps(Conway *et al.* 1996) et les méthodes de régression (Wilby *et al.* 2002). Chacune de ces familles a ses avantages et ses inconvénients pour reproduire la variabilité du prédictant à l'échelle locale. Nous mettons l'accent sur les méthodes de régression qui intéressent notre étude.

Les méthodes basées sur la régression permettent de trouver une relation directe entre les prédicteurs à grande échelle et les prédictants à l'échelle locale à travers des fonctions aléatoires ou déterministes, par exemple : la Régression Linéaire Multiple (Wilby, 1998; Von Storch *et al.*2000;Hellstrom *et al.*2001; Wilby *et al.*200 3), l'analyse des fonctions orthogonales empiriques (Buishand et Brandsma, 2001; Biau *et al.* 1999; Linderson *et al.* 2004), l'analyse canonique des corrélations (Huth et Pokorna, 2004), les réseaux de neurones artificiels (Wilby, 1998; Stehlik et Bárdossy, 2002; Hewitson et Crane, 1996; Von Storch *et al.* 2000), et la décomposition en valeurs singulières (Huth et Pokorna, 2004). En réalité les relations entre les prédicteurs et les prédictants sont très complexe. Les modèles de régression s'adaptent avec ces relations, et tiennent compte des effets locaux au niveau des stations, pour donner finalement des résultats consistants avec les valeurs observées de la variable d'intérêt (Fasbender et Ouarda, 2010).

Même si les modèles de régressions donnent de bons résultats pour réduire l'échelle, leur inconvénient majeur est que l'estimation est généralement limitée aux sites mesurés, et ils sont incapables d'étendre les résultats à des endroits non jaugés. Un autre inconvénient des modèles de régression c'est qu'ils reproduisent généralement la moyenne ou la partie centrale des prédictants conditionnellement à des prédicteurs sélectionnés. De ce fait, la variabilité de la régression est toujours inférieure à la variabilité observée, à cet égard les approches probabilistes ont apporté des contributions pertinentes dans les applications de réduction d'échelle. En effet, elles permettent de modéliser toute la distribution du prédictant et rendre compte de l'incertitude des paramètres estimés (Bates et al 1998; Bellone *et al.* 2000; Tebaldi *et al.* 2004; Vrac et Naveau 2007; Vrac *et al.* 2007).

Pour obtenir l'estimation aux endroits non jaugés, Michelangeli *et al.* (2009) ont interpolé les prédicteurs AOGCM sur l'emplacement des stations de mesure avant la régression. Une autre solution c'est d'interpoler les résultats à échelle réduite (Benestad 2002, 2007; Hundecha et Bárdossy 2005). Une troisième solution est d'interpoler les prédictants dans les stations de mesure sur la grille cible régulière puis faire la régression (Lim *et al.* 2007; Baigorria *et al.* 2008). Cette méthode donne des bons résultats et reproduit adéquatement les caractéristiques climatiques observées. Cependant, son inconvénient majeur est sa grande exigence en calcul qui dépend directement de la taille de la grille cible régulière.

Dans une étude plus récente, Fasbender et Ouarda (2010) ont proposé une solution alternative pragmatique. L'idée est de prendre directement en compte la dépendance spatiale des prédictants au cours du processus de réduction d'échelle. Dans cette étude Fasbender et Ouarda (2010) ont employé un modèle spatial bayésien pour la réduction d'échelle spatiale des données AOGCM aux températures maximale et minimale dans la partie sud du Québec. Ce modèle repose sur un cadre bayésien afin de relier les prédicteurs et les prédictants par des prédictants à grande échelle qui ont la même résolution et le même emplacement que les données AOGCM. L'information *a priori* a été décrite au moyen d'un modèle de régression géographique (GRM) qui tient compte des caractéristiques locales. Ce model a l'avantage de fournir à la fois une estimation à des sites non jaugés et la distribution des prédictants. De plus, les résultats ont été satisfaisants avec les valeurs observées au niveau des stations. Cette étude ayant été réalisée sur une zone relativement petite (environ 1000x1000 km²), les paramètres estimés du modèle ne peuvent pas être utilisés pour les grandes surfaces. De plus, le faible nombre des stations utilisées pour la calibration du modèle, n'a pas permis de tenir compte d'autres caractéristiques locales que la latitude et l'altitude pour exprimer l'*a priori* du GRM.

Dans ce travail, la méthode bayésienne proposée par Fasbender et Ouarda (2010) est adaptée sur une grande surface qui couvre la majorité de la province de Québec au Canada. La méthode est employée pour réduire l'échelle des prédicteurs AOGCM aux températures maximales et minimales quotidiennes. Dans ce travail, le jeu de donnée utilisé est plus grand, par conséquent, le GRM proposé par Fasbender et Ouarda (2010) est amélioré par l'ajout de deux autres variables : la longitude et la distance à la mer. Après une brève description de la région d'étude et de la base de données utilisée, le modèle spatial bayésien et la méthode d'évaluation de la qualité du modèle sont présentés. Les résultats sont validés et comparés à ceux obtenus avec une Régression Linéaire Multiple Multivariée (MMLR) en utilisant le même ensemble des prédicteurs.

2. ZONE D'ÉTUDE ET BASE DE DONNÉES

La zone d'étude se situe dans la province du Québec au Canada, entre les latitudes 45°N et 60°N, et les longitudes 60°W et 80°W (voir figure 1). On dispose comme prédictants, de 22 séries des températures maximales et minimales observées quotidiennement (voir figure1). Ces séries proviennent des 22 stations météorologiques d'Environnement Canada et couvrent la période entre le 1er janvier 1961 et le 31 décembre 2000. Elles ont été homogénéisées suivant la méthode préconisées par Vincent *et al.* (2002).

Les produits de réanalyse du NCEP/NCAR (Kalnay *et al.* 1996; Kistler *et al.* 2001) sont utilisés afin d'évaluer le potentiel de la méthode de réduction d'échelle. Toutes les données du NCEP/NCAR sont moyennées sur une base quotidienne à partir de données aux 6 heures sur la grille régulière originale de 2,5° lat. x 2,5° long. Les prédicteurs obtenus sont interpolés linéairement sur la grille gaussienne 3,75 ° x 3.75 ° du MCCG3 (DAI Prédicteurs MCCG3 2008) correspondant à la troisième version du modèle couplé du climat du globe (Scinocca *et al.* 2008). Ces prédicteurs (sauf la direction du vent) sont ensuite normalisés par rapport à la période de référence 1961-1990. 24 points de grille AOGCM qui couvrent la zone d'étude sont sélectionnés (voir figure 1), et pour chaque point de grille, 25 prédicteurs NCEP sont fournis (voir tableau 2). Pour chaque jour, 600 variables sont ainsi disponibles pour le processus de réduction d'échelle. L'ensemble des données totales est divisé en deux sous-ensembles indépendants. Les données entre 1961 et 1990 sont utilisées pour la calibration du modèle, tandis que les données entre 1991 et 2000 sont utilisées pour la validation.

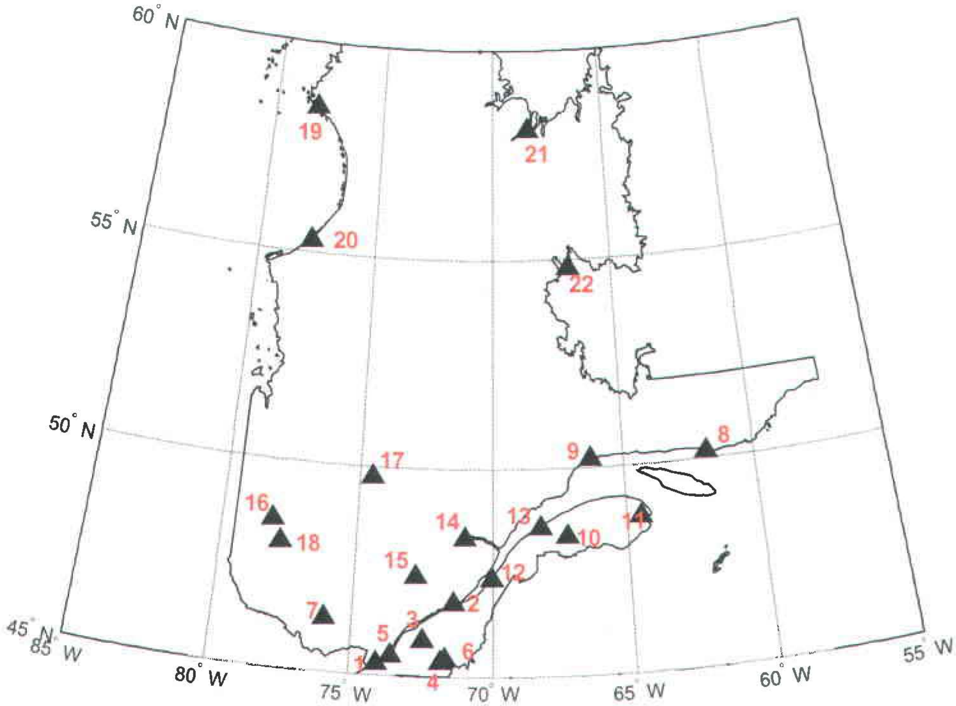

Figure 1. Les points de la grille MCGAO et les stations météorologiques choisies pour la zone d'étude : (1) Cesdres, (2) Quebec, (3) Drummondville, (4) Lennoxville, (5) McTavish (Montréal), (6) Sherbrooke A, (7) ManiwakiAirport, (8) Natashquan, (9) Sept-Iles A, (10) Causapscal, (11) Gaspe, (12) La Pocatière, (13) Mont-Joli A, (14) Bagotville A, (15) La Tuque, (16) Amos, (17) Chibougamau Chapais, (18) Val-D'Or A, (19) Inukjuak, (20) Kuujjuarapik, (21) Kuujjuaq, (22) Schefferville.

Tableau 1. **Les statistiques annuelles des 22 stations météorologiques.**

Station	Lat	Long	Élévation (m)	Tmax (°C)		Tmin (°C)	
				MA	SD	MA	SD
Inukjuak	58.47	-78.08	25	-2.9	13.45	-10.23	13.98
Kuujjuaq	58.1	-68.42	39	-1.08	14.13	-10.10	13.92
Kuujjuarapik	55.28	-77.77	10	0.06	13.74	-8.89	14.04
Schefferville	54.8	-66.82	522	-0.12	13.76	-9.61	14.34
Sept-Iles A	50.22	-66.27	55	5.53	10.98	-3.64	11.85
Natashquan	50.18	-61.82	11	5.63	10.10	-3.36	11.39
Chibougamau Chapais	49.77	-74.53	387	4.90	13.76	-6.74	14.40
Gaspe	48.78	-64.48	33	8.61	11.34	-2.76	10.70
Mont-Joli A	48.6	-68.22	52	7.56	11.79	-0.74	10.46
Amos	48.57	-78.13	310	6.48	13.74	-4.20	13.69
Causapscal	48.37	-67.23	168	8.13	12.52	-3.54	12.05
Bagotville A	48.33	-71	159	7.88	13.40	-2.39	12.69
Val-D'Or A	48.05	-77.78	337	7.20	13.52	-3.46	13.38
La Tuque	47.4	-72.78	152	9.54	13.09	-3.19	13.25
La Pocatiere	47.35	-70.03	31	8.81	12.15	-0.38	11.21
Québec	46.78	-71.38	74	9.18	12.56	-0.05	11.46
ManiwakiAirport	46.3	-76	200	9.90	12.64	-1.99	12.50
Drummondville	45.88	-72.48	82	10.49	12.61	0.89	12.11
McTavish (Montréal)	45.5	-73.58	73	11.14	12.29	3.49	11.52
Sherbrooke A	45.43	-71.68	241	10.34	12.13	-1.06	11.81
Lennoxville	45.37	-71.82	181	11.10	12.17	-0.20	11.93
Les Cedres	45.3	-74.05	47	10.77	12.25	2.37	11.72

Tableau 2. Prédicteurs NCEP sur la grille MCCG3.

No	Predicteurs	No	Predicteurs
1	Pression moyenne au niveau de la mer	14	Divergence à 500 hPa
2	Vitesse du vent 1000 hPa	15	Vitesse du vent 850 hPa
3	Composante U à 1000 hPa	16	Composante U à 850 hPa
4	Composante V à 1000 hPa	17	Composante V à 850 hPa
5	Vorticité à 1000 hPa	18	Vorticitéà 850 hPa
6	Direction du vent à 1000 hPa	19	Géopotentielle à 850 hPa
7	Divergence à 1000 hPa	20	Direction du vent à 850 hPa
8	Vitesse du vent à 500 hPa	21	Divergence à 1000 hPa
9	Composante U à 500 hPa	22	Humidité spécifique à 500 hPa
10	Composante V à 500 hPa	23	Humidité spécifique à 850 hPa
11	Vorticitéà 500 hPa	24	Humidité spécifique à 1000 hPa
12	Géopotentielle à 500 hPa	25	Température à 2m
13	Direction du vent à 500 hPa		

3. MÉTHODOLOGIE

Le but de ce travail est de fournir un modèle qui nous permet pour un jour donné, d'estimer des cartes des températures maximale et minimale pour la zone d'étude à partir des sorties AOGCM pour ce même jour. Avant tous, une analyse en composantes principales (ACP) est réalisée afin de réduire le nombre des prédicteurs NCEP à r prédicteurs décorelés qui conservent 95% de la variabilité des prédicteurs originaux. On note le vecteur de ces prédicteurs décorelés par

$$\mathbf{Y} = (Y_1, \quad \dots, \quad Y_r)', \tag{1}$$

où t désigne la transposé du vecteur. On considère maintenant la grille cible de réduction d'échelle, il s'agit d'une grille régulière plus fine que la grille AOGCM et qui couvre la zone d'étude. On désigne par $\mathbf{a}_i$ les coordonnées d'un point spécifique de cette grille, et par $\mathbf{a}_0$ les coordonnées du point cible pour la procédure de réduction d'échelle. Pour un jour donné, si on dispose du vecteur $\mathbf{Y}$, on cherche à estimer les températures maximale $T_{\max}(\mathbf{a}_0)$ et minimale $T_{\min}(\mathbf{a}_0)$ à l'emplacement $\mathbf{a}_0$. On désigne alors, par $\mathbf{X}_0$ le vecteur des prédictants à l'emplacement $\mathbf{a}_0$, et par $\mathbf{X}_i$ le vecteur des prédictants à l'emplacement $\mathbf{a}_i$,

$$\mathbf{X}_0 = (T_{\max}(\mathbf{a}_0) \quad T_{\min}(\mathbf{a}_0))' \ , \tag{2}$$

$$\mathbf{X}_i = (T_{\max}(\mathbf{a}_i) \quad T_{\min}(\mathbf{a}_i))' \ . \tag{3}$$

Dans un contexte probabiliste, on est donc à la recherche de la distribution conditionnelle $f(\mathbf{X}_0 \mid \mathbf{Y})$. La méthode bayésienne proposée par Fasbender et Ouarda (2010) est appliquée dans ce travail pour l'estimation de $f(\mathbf{X}_0 \mid \mathbf{Y})$. L'application de cette méthode pour tous les emplacements $\mathbf{a}_i$ permet d'estimer toutes les distributions conditionnelles $f(\mathbf{X}_i \mid \mathbf{Y})$. Ceci permet ainsi, d'estimer les deux cartes des températures maximale et minimale.

Dans les sections suivantes, la méthode bayésienne utilisée est présentée ainsi que la méthode d'évaluation de la qualité de réduction d'échelle.

3.1 Modèle spatial bayésien

Dans une approche bayésienne, tous les paramètres inconnus sont considérés comme des variables aléatoires. Pour tenir compte du caractère aléatoire d'un paramètre inconnu, l'information dont on dispose sur ce paramètre est modélisée par une densité de probabilité *a priori*. Puis en utilisant le théorème de Bayés cette distribution *a priori* est mise à jour par les observations pour avoir finalement une distribution *a posteriori*.

3.1.1 Modèle de régression géographique

L'information *a priori* est spécifiée en utilisant un modèle de régression géographique (GRM) qui tient compte de la longitude, la latitude, l'altitude et la racine carrée de la distance à la mer à l'endroit d'estimation $\mathbf{a}_0$ (voir Benestad 2002; Benestad *et al.* 2008 pour plus de détails). On désigne par α le vecteur des paramètres de la distribution *a priori* de $\mathbf{X}_0$ et par $m_0(\alpha)$ son vecteur moyen. Dans ce cas le modèle GRM s'écrit :

$$m_{0;d}(\alpha) = \alpha_{1;d} + \alpha_{2;d}(\lambda - m_\lambda) + \alpha_{3;d}(\phi - m_\phi) + \alpha_{4;d}\sqrt{D} + \alpha_{5;d}H + \varepsilon_{m;d}\ ,\ d = 1,\ldots,12 \quad (4)$$

où $d = 1,\ldots,12$ correspondant au mois actuel, (λ , ϕ , D , et H sont respectivement la longitude, la latitude, la distance à la mer et l'altitude de l'emplacement cible, m_ϕ et m_λ sont respectivement les moyennes des latitudes et des longitudes observées aux stations météorologiques, et (iv) les vecteurs $\alpha_{1:d},\ldots,\alpha_{5:d}$ sont les sous-vecteurs de α . Les paramètres $\alpha_{1,d},\ldots,\alpha_{4;d}$ sont ajustés à l'aide de la méthode des moindres carrés ordinaires (OLS). Tandis que les paramètres $\alpha_{5:d}$ sont ajustés conformément à l'atmosphère type international (voir ISO 2533:1975, ISO 2009), ils sont définis de telle sorte que la température diminue en fonction de l'altitude de 6,5 °C par km ce qui correspond au gradient thermique adiabatique dans la troposphère. Les estimateurs OLS sont asymptotiquement multivariés gaussiens. Notons par Σ_α la matrice de covariance du vecteur α . Afin de respecter les caractéristiques saisonnières observées dans les stations météorologiques, le GRM est corrigé en ajoutant les erreurs $\varepsilon_{m;d}$ par la méthode d'interpolation inverse distance, les paramètres de cette interpolation sont estimés en utilisant une procédure «leave-one-out» pour chaque mois. On suppose dans ce modèle que la variance des erreurs $\varepsilon_{m;d}$ est proportionnelle à un modèle de semi-variogramme gaussien qui utilise la distance à la station la plus proche. La matrice de covariance diagonale Σ_m est construite selon ce modèle.

3.1.2 Modèle spatial pour la température

Un modèle spatial commun pour les températures minimales et maximales quotidiennes est estimé en utilisant un modèle linéaire de corégionalisation (LMC) (voir Chilès et Delfiner 1999). Pour tenir compte du comportement spatial lisse de la température, ce modèle spatial combine un modèle gaussien et un modèle de pépite. Les distances spatiales sont calculées en utilisant les coordonnées UTM (Universal Transverse Mercator). Ceci permet d'éviter les déformations spatiales dues à des latitudes élevées.

On suppose maintenant que chaque point $\mathbf{A}_i$ de la grille du modèle MCGAO correspond à une certaine zone spatiale. On définit le vecteur aléatoire des prédictants à grande échelle par

$$\boldsymbol{\chi} = \left(T_{\max}(\mathbf{A}_1), \quad \ldots, \quad T_{\max}(\mathbf{A}_n), \quad T_{\min}(\mathbf{A}_1), \quad \ldots, \quad T_{\min}(\mathbf{A}_n) \right)' \quad , \tag{5}$$

où $T_{\max}(\mathbf{A}_i)$ et $T_{\min}(\mathbf{A}_i)$ sont définis par les deux formules suivantes :

$$T_{\max}(\mathbf{A}_i) = \frac{1}{|\mathbf{A}_i|} \int_{\mathbf{A}_i} T_{\max}(\mathbf{a})\mathrm{d}\mathbf{a} \ , \tag{6}$$

$$T_{\min}(\mathbf{A}_i) = \frac{1}{|\mathbf{A}_i|} \int_{\mathbf{A}_i} T_{\min}(\mathbf{a})\mathrm{d}\mathbf{a} \ . \tag{7}$$

Avec $|\mathbf{A}_i|$ represente la surface de la zone $\mathbf{A}_i$. Les éléments $T_{\max}(\mathbf{A}_i)$ et $T_{\min}(\mathbf{A}_i)$ du vecteur $\boldsymbol{\chi}$ sont les différentes expressions des prédictants $T_{\max}$ et $T_{\min}$ à grande échelle sur la zone $\mathbf{A}_i$. Le vecteur des prédictants à grande échelle $\boldsymbol{\chi}$ est lié au vecteur $\mathbf{X}_0$ par leurs positions spatiales, et il est employé ici pour pouvoir rendre compte de la dépendance spatiale de la température au cours du processus de réduction d'échelle. La relation entre $\mathbf{X}_0$ et $\boldsymbol{\chi}$ est décrite par une fonction de covariance régularisée (Goovaerts 2008) qui est calculée en utilisant le modèle LMC. Ainsi, on peut estimer une série de $\boldsymbol{\chi}$ couvrant la période de calibration. Le GRM et les deux équations (6) et (7) sont utilisés pour garantir une estimation plus cohérente de la moyenne mensuelle $m_{\boldsymbol{\chi}}(\alpha)$ du vecteur $\boldsymbol{\chi}$.

3.1.3 Modèle de Régression Linéaire Multiple Multivarié

Une fois les séries de $\mathcal{X}$ ont été estimées pour la période de calibration à partir des données historiques, on identifie leur relation avec les prédicteurs décorelés $\mathbf{Y}$ en utilisant une analyse de Régression Linéaire Multiple Multivariée (MMLR) donnée par

$$\mathcal{X} = \beta_{0,d} + \beta_{1,d}\mathbf{Y} + \varepsilon \qquad d = 1,\ldots 12\,, \tag{8}$$

où $\beta = (\beta_{0,d}, \beta_{1,d})$ est le vecteur des paramètres du modèle, $\beta_{0,d}$ est le sous-vecteur de β correspondant à l'interception du $d^{ème}$ mois, $\beta_{1,d}$ est une matrice $2 \times r$ avec des colonnes correspondantes aux différents prédicteurs de $\mathbf{Y}$ et ε est un vecteur des résidus de moyenne nulle et de matrice de covariance Σ_{ε}.

Le vecteur des prédictants à grande échelle $\mathcal{X}$ est lié au vecteur $\mathbf{X}_0$ par leurs positions spatiales, il est aussi lié au vecteur des prédicteurs $\mathbf{Y}$ par le modèle MMLR.

Ainsi le modèle bayésien présenté combine trois modèles : (i) le modèle GRM qui est employé pour l'estimation des moyennes *a priori* des températures maximales et minimales, (ii) le modèle spatial commun pour les températures maximale et minimale qui détermine le lien entre les prédictants à l'échelle locale et les pérdictants à grande échelle en utilisant la fonction de covariance régularisée et (iii) le modèle MMLR qui détermine le lien entre les prédictants à grande échelle et les prédicteurs décorrélés.

La section suivante montre comment les informations issues de ces trois modèles sont combinées afin d'estimer la distribution *a posteriori* $f(\mathbf{X}_0 \mid \mathbf{Y})$.

3.1.4 Estimation a posteriori

Le vecteur des prédictants à grande échelle $\boldsymbol{\chi}$ et les prédicteurs originaux (les prédicteurs NCEP) sont situés au même endroit sur la grille MCGAO, et représentent aussi une même résolution spatiale, ce qui permet de négliger l'information fournie par les prédictants $\mathbf{X}_0$ sur les prédicteurs $\mathbf{Y}$ par rapport à l'information fournie par les prédictants à grande échelle $\boldsymbol{\chi}$. Le vecteur $\boldsymbol{\chi}$ peut donc être considéré comme un vecteur de variables cachées permettant le transfert d'informations à partir des prédicteurs $\mathbf{Y}$ aux prédictants $\mathbf{X}_0$.

On suppose une indépendance entre les prédicteurs $\mathbf{Y}$ et les prédictants $\mathbf{X}_0$ conditionnellement au vecteur $\boldsymbol{\chi}$:

$$\mathbf{X}_0 \perp \mathbf{Y} \mid \boldsymbol{\chi}. \tag{9}$$

On suppose aussi que β n'influence que le lien entre $\mathbf{Y}$ et $\boldsymbol{\chi}$, et α n'influence que les distributions *a priori*. En plus, on suppose que α et β sont indépendants. Ces hypothèses permettent de montrer, en utilisant le théorème de Bayes, que la distribution *a posteriori* $f(\mathbf{X}_0 \mid \mathbf{Y})$ s'écrit :

$$f(\mathbf{X}_0 \mid \mathbf{Y}) \propto \iiint \frac{f(\boldsymbol{\chi} \mid \mathbf{Y};\beta)}{f(\boldsymbol{\chi} \mid \alpha)} f(\mathbf{X}_0, \boldsymbol{\chi} \mid \alpha) f(\alpha) f(\beta) \, \mathrm{d}\boldsymbol{\chi} \, \mathrm{d}\alpha \, \mathrm{d}\beta. \tag{10}$$

On peut trouver les détails de ce calcul dans Fasbender et Ouarda (2010).

Pour tenir compte des effets saisonniers, toutes les distributions de cette équation sont modélisées par mois. Généralement les données de la température mensuelle suivent approximativement une distribution gaussienne. Par conséquent, les distributions $f(\mathbf{X}_0, \boldsymbol{\chi} \mid \beta)$, $f(\boldsymbol{\chi} \mid \mathbf{Y}; \beta)$et $f(\boldsymbol{\chi} \mid \alpha)$sont supposées êtres gaussiennes. On considère aussi que toutes les distributions de l'équation (10) sont stationnaires. Toutefois, les fluctuations non stationnaires sont induites par les évolutions non stationnaires des prédicteurs.

Ces hypothèses impliquent que le couple ($\mathbf{X}_0, \boldsymbol{\chi}$) suit une distribution gaussienne multivariée (voir Fasbender et Ouarda (2010) pour plus de détail) de vecteur moyen $\mathbf{M}(\alpha, \beta)$et de matrice de covariance $\mathbf{S}$ donnés par

$$\begin{cases} \mathbf{S}^{-1} & = \boldsymbol{\Sigma}^{-1} + \begin{pmatrix} 0 & 0 \\ 0 & \boldsymbol{\Sigma}_{\varepsilon}^{-1} - \boldsymbol{\Sigma}_{\boldsymbol{\chi}}^{-1} \end{pmatrix} \\ \mathbf{M}(\alpha, \beta) & = \mathbf{S}\left(\boldsymbol{\Sigma}^{-1} \begin{pmatrix} m_0(\alpha) \\ m_{\boldsymbol{\chi}}(\alpha) \end{pmatrix} + \begin{pmatrix} 0 \\ \boldsymbol{\Sigma}_{\varepsilon}^{-1}(\beta_{0,d} + \beta_{1,d})\mathbf{Y} - \boldsymbol{\Sigma}_{\boldsymbol{\chi}}^{-1} m_{\boldsymbol{\chi}}(\alpha) \end{pmatrix} \right) \end{cases}, \quad (11)$$

où, Σ est la matrice de covariance de $\mathbf{X}_0$et$\boldsymbol{\chi}$, $\Sigma_{\boldsymbol{\chi}}$ est la sous-matrice deΣ correspondante à$\boldsymbol{\chi}$ seulement, Σ_{ε}est la matrice de covariance des résidus. $m_0(\alpha)$ et $m_{\boldsymbol{\chi}}(\alpha)$sont respectivement les vecteurs moyens de$\mathbf{X}_0$et$\boldsymbol{\chi}$. Fasbender et Ouarda (2010) ont montré qu'il n'était pas nécessaire de trouver une expression analytique de la distribution *a posteriori* $f(\mathbf{X}_0 \mid \mathbf{Y})$ car il est possible d'estimer toutes les propriétés de cette distribution à partir d'un algorithme d'échantillonnage (par exemple écart-type, mode, médiane). Il s'agit de tirer des échantillons $\widetilde{\beta}_i$ dans $f(\beta)$(en utilisant le modèle MMLR de la section II.1.c), $\widetilde{\alpha}_i$ dans $f(\alpha)$ et $(\widetilde{\varepsilon}_{m,d})_i$ (en utilisant le GRM de la section II.1.a), ensuite, en utilisant les échantillons $\widetilde{\alpha}_i$ et$(\widetilde{\varepsilon}_{m,d})_i$ calculer les valeurs perturbées $(\widetilde{m}_0(\widetilde{\alpha}_i))_i$ (en utilisant les équations 6 et 7 de la section II.1.b) pour le GRM de l'équation (4). Puis, en utilisant les échantillons $\widetilde{\beta}_i$ et $(\widetilde{m}_0(\widetilde{\alpha}_i))_i$ prélever un échantillon $(\widetilde{\mathbf{X}}_0, \widetilde{\mathbf{\chi}})_i$ dans la distribution gaussienne avec les paramètres donnés dans l'équation (11). Enfin répéter ces différentes étapes jusqu'à atteindre le nombre des itérations désiré.

3.2 Évaluation de la qualité du modèle de réduction d'échelle

Pour l'évaluation de la qualité de réduction d'échelle, les données entre 1991 et 2000 sont utilisées. Deux approches sont considérées pour la validation du modèle spatial bayésien. La première approche repose sur une comparaison directe entre les valeurs estimées et les valeurs observées en utilisant des critères statistiques, tandis que la deuxième approche est basée sur le calcul d'indices climatiques qui sont comparés avec une méthode classique basée sur la régression.

Dans la première approche de validation, 22 séries des températures maximales et minimales sont estimées au niveau des stations. Trois critères statistiques sont employés pour la validation du modèle. Le premier critère est l'erreur moyenne (ME) qui est donnée par

$$ME = \frac{1}{N}\sum_{i=1}^{N}\left(x_{obs_i} - x_{sim_i}\right), \tag{12}$$

où N désigne le nombre des observations, x_{obs_i} désigne la valeur observée, x_{sim_i} désigne la valeur simulée et i désigne le jour.

Ce critère représente une mesure de l'exactitude et doit donc être minimisée. Le deuxième critère est la racine carrée de l'erreur quadratique moyenne (RMSE) qui est donnée par

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(x_{obs_i} - x_{sim_i}\right)^2} \;. \tag{13}$$

La RMSE représente une mesure inverse de la précision et doit donc être minimisé. Le troisième critère utilisé est le pourcentage des observations contenues dans l'intervalle de crédibilité à 95%. Ce pourcentage doit donc être proche de 95% pour garantir que la distribution estimée soit pertinente. Ces trois critères statistiques permettent une comparaison directe entre la distribution à échelle réduite $f(\mathbf{X}_0 \mid \mathbf{Y})$ et les valeurs observées.

Dans la deuxième approche de validation, on considère un ensemble de plusieurs indices climatiques mensuels et saisonniers qui ont été proposés dans le cadre des climats nordiques. Les définitions de ces indices climatiques sont présentées au Tableau 3. Ces indices sont choisis pour évaluer la performance des modèles de réduction d'échelle et reflètent la fréquence, l'intensité, la durée et les températures extrêmes (voir Hessami *et al.* 2008; Gachon *et al.* 2005; Wilby *et al.* 1998, 2002.). L'approche consiste à comparer la concordance entre ces indices obtenus en utilisant le modèle spatial bayésien et ceux observés aux sites jaugés. Pour mesure cette concordance, les RMSE calculés à partir des différences observations-estimations sont utilisés. Afin de montrer que le modèle reproduit des résultats similaires à d'autres approches de désagrégation statistique, la même approche de validation est appliquée à un modèle MMLR utilisant les prédictants au niveau des stations et les mêmes prédicteurs décorrélés.

Tableau 3. Indices climatiques.

Indice	Définition	Unité	Échelle de temps
DTR	Moyenne de la plage de températures diurnes	°C	saison
FSL	Durée de la saison de gel: Nombre de jours entre 5 jours consécutifs $T_{moyenne}$<0 °C et 5 jours consécutifs $T_{moyenne}$> 0°C	jours	année
GSL	Durée de la saison de croissance : Nombre de jours entre 5 jours consécutifs $T_{moyenne}$> 5 °Cet 5 consécutifs $T_{moyenne}$<5 °C	jours	année
FR-Th	Nombre de jours avec cycle de gel et de dégel (T_{max}> 0°C,T_{min}<0°C)	jours	mois
Tmax90	90éme centile de la température maximale	°C	saison
Tmin90	90éme centile de la température minimale	°C	saison

$$T_{moyenne} = \frac{T_{max} + T_{min}}{2}$$

4. RÉSULTATS

4.1 Résultats de la calibration

Les caractéristiques locales (la longitude, la latitude, l'altitude et la distance à la mer) au niveau des 22 stations météorologiques ont été utilisées pour la calibration du modèle (GRM). Les données de l'altitude ont été fournies par un modèle numérique de terrain qui couvre la zone d'étude (voir Figure2).

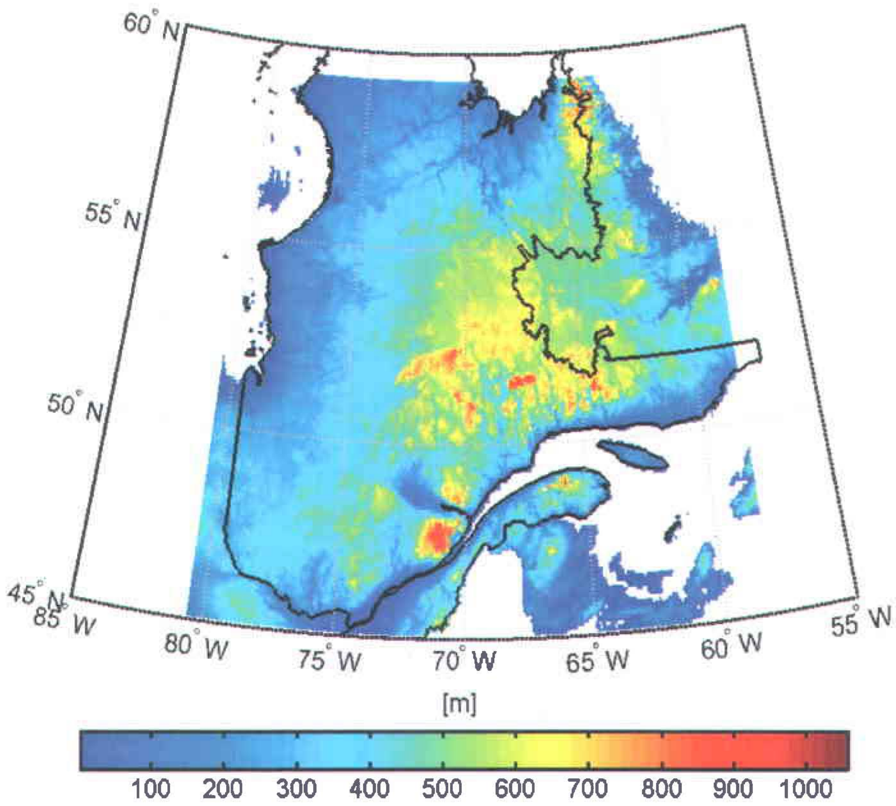

Figure 2. Modèle Numérique de Terrain de la zone d'étude.

Un deuxième modèle numérique de terrain qui couvre tout le Canada a été utilisé pour déterminer le contour des côtes canadiennes. Un second contour déterminant les côtes de la partie Est de la ville de New York a été ajouté. Le contour final a été utilisé pour calculer les distances à la mer (voir Figure3).

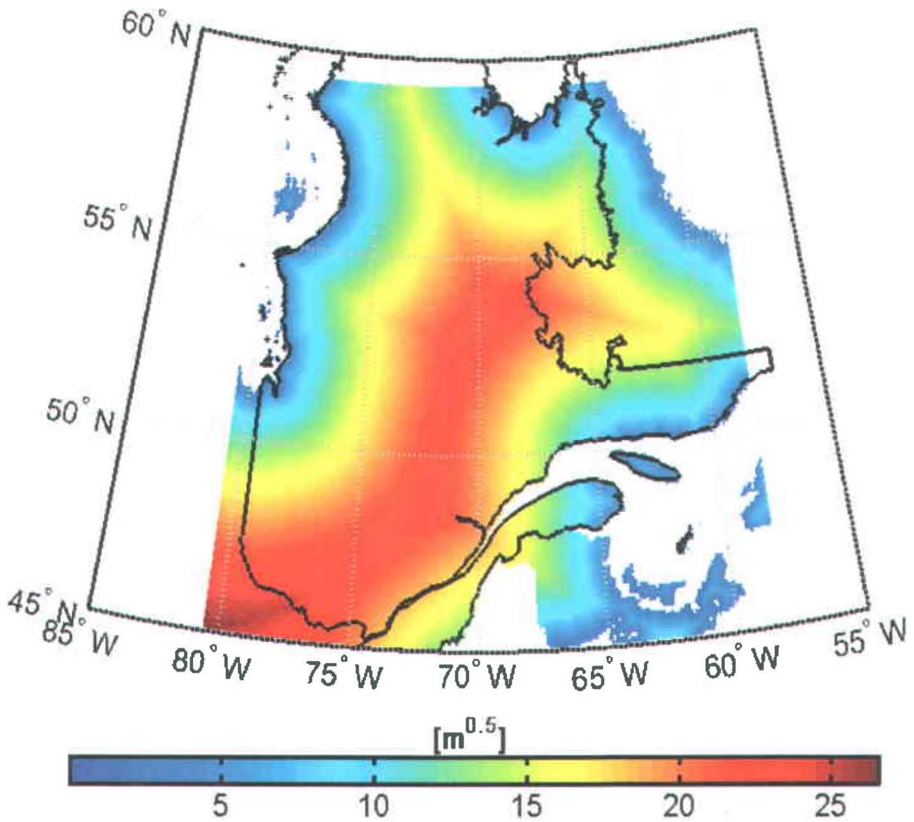

Figure 3. Racine carré de la distance à la mer de la zone d'étude.

Le Tableau 4 montre l'estimation mensuelle des paramètres du GRM pour les températures maximale et minimale. Les résidus de ce modèle de régression (voir Figure 4) ont été ensuite interpolés dans l'espace et ajoutés à l'aide d'une interpolation inverse distance. Les paramètres de puissance de l'interpolation inverse distance ont été estimés par mois, en utilisant une procédure leave one-out (voir Tableau 5). La variance de cette interpolation a été modélisée comme une fonction de variogramme gaussienne avec une variance égale à la variance observée des résidus du GRM et une portée égale à 100 km.

Tableau 4. Estimation mensuelle des paramètres du GRM pour les températures maximale et minimale.

Tmax												
	J	F	M	A	M	J	J	A	S	O	N	D
α_1	-8.41	-8.32	-2.99	4.06	11.37	16.59	19.88	19.37	15.78	10.35	3.39	-3.80
α_2	-1.24	-1.32	-1.24	-1.08	-1.07	-0.82	-0.70	-0.72	-0.81	-0.89	-0.81	-1.09
α_3	0.19	0.22	0.16	-0.01	-0.06	-0.01	0.01	0.01	-0.04	-0.07	0.02	0.11
α_4	-0.06	0.02	0.11	0.17	0.21	0.26	0.23	0.18	0.06	-0.01	-0.08	-0.12
Tmin												
	J	F	M	A	M	J	J	A	S	O	N	D
α_1	17.92	-18.56	-12.81	4.44	1.58	6.57	9.91	9.92	6.26	1.66	3.54	-12.10
α_2	-1.19	-1.31	-1.39	-1.17	-0.78	-0.74	-0.62	-0.57	-0.50	-0.53	-0.76	-1.11
α_3	0.12	0.21	0.21	0.07	-0.04	-0.02	0.01	-0.04	-0.11	-0.12	-0.02	0.08
α_4	-0.09	-0.02	0.05	0.07	0.12	0.17	0.16	0.10	0.05	0.01	-0.08	-0.14

Tableau 5. **Estimation mensuelle des paramètres de la méthode d'interpolation inverse distance pour les températures maximale et minimale.**

	J	F	M	A	M	J	J	A	S	O	N	D
Tmax	5.59	6.20	6.09	9.11	8.80	2.80	4.68	6.19	6.40	8.21	4.67	2.88
Tmin	4.31	4.59	6.01	9.08	3.58	3.95	9.12	7.11	3.06	7.39	9.05	7.52

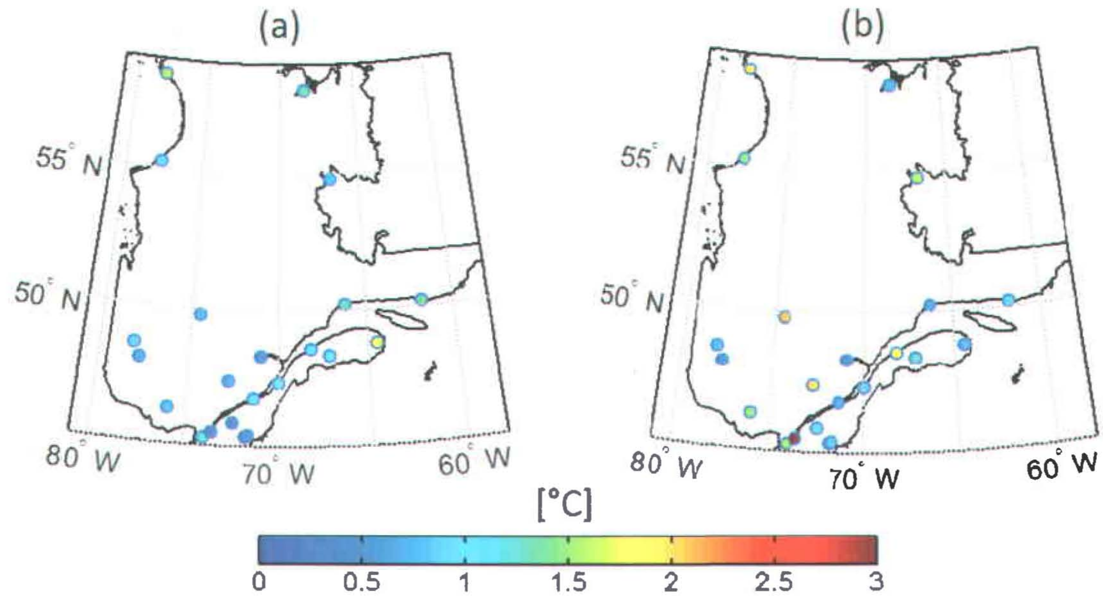

Figure 4. RMSE du GRM au niveau des stations pour les températures maximale (a) et minimale (b).

Ainsi l'incertitude du GRM a été considérée comme proportionnelle à la distance à la station météorologique la plus proche. Une validation croisée a été utilisée pour valider le GRM. La Figure 5 montre la comparaison, pour chaque mois, entre les écarts types estimés fournis par le GRM et les RMSE observées calculées à l'aide de la validation croisée. On peut également voir que les deux lignes sont très proches, de ce fait, les écarts types du GRM sont en accord avec la RMSE de la validation croisée.

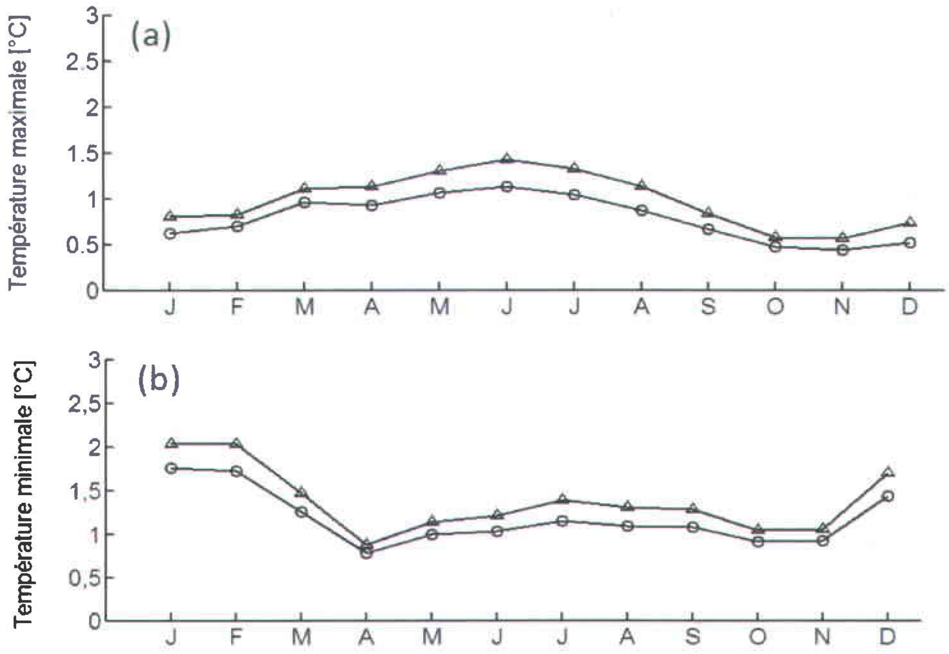

Figure 5. Comparaison, pour chaque mois, entre les RMSE du GRM (cercles) et les RMSE observées calculées à l'aide de la validation croisée (triangles) pour les températures maximales (a) et les températures minimales (b).

Le GRM a été appliqué sur la zone d'étude. La Figure 6 montre l'effet de la distance à la mer sur la moyenne *a priori* des températures maximale et minimale. On observe un effet saisonnier de la distance à la mer; durant l'été la température augmente en fonction de la distance à la mer; par contre en hivers on remarque un effet contraire bien que moins important. Les figures 6.e et 6.f montrent que durant l'été, l'effet de la distance à la mer sur la température maximale est plus marqué que celui sur la température minimale. Finalement, ces 8 figures montrent que l'effet de la distance à la mer ne peut pas être ignoré. De ce fait, la prise en compte de ce paramètre dans le modèle GRM permet d'améliorer les résultats obtenus par Fasbender et Ouarda (2010).

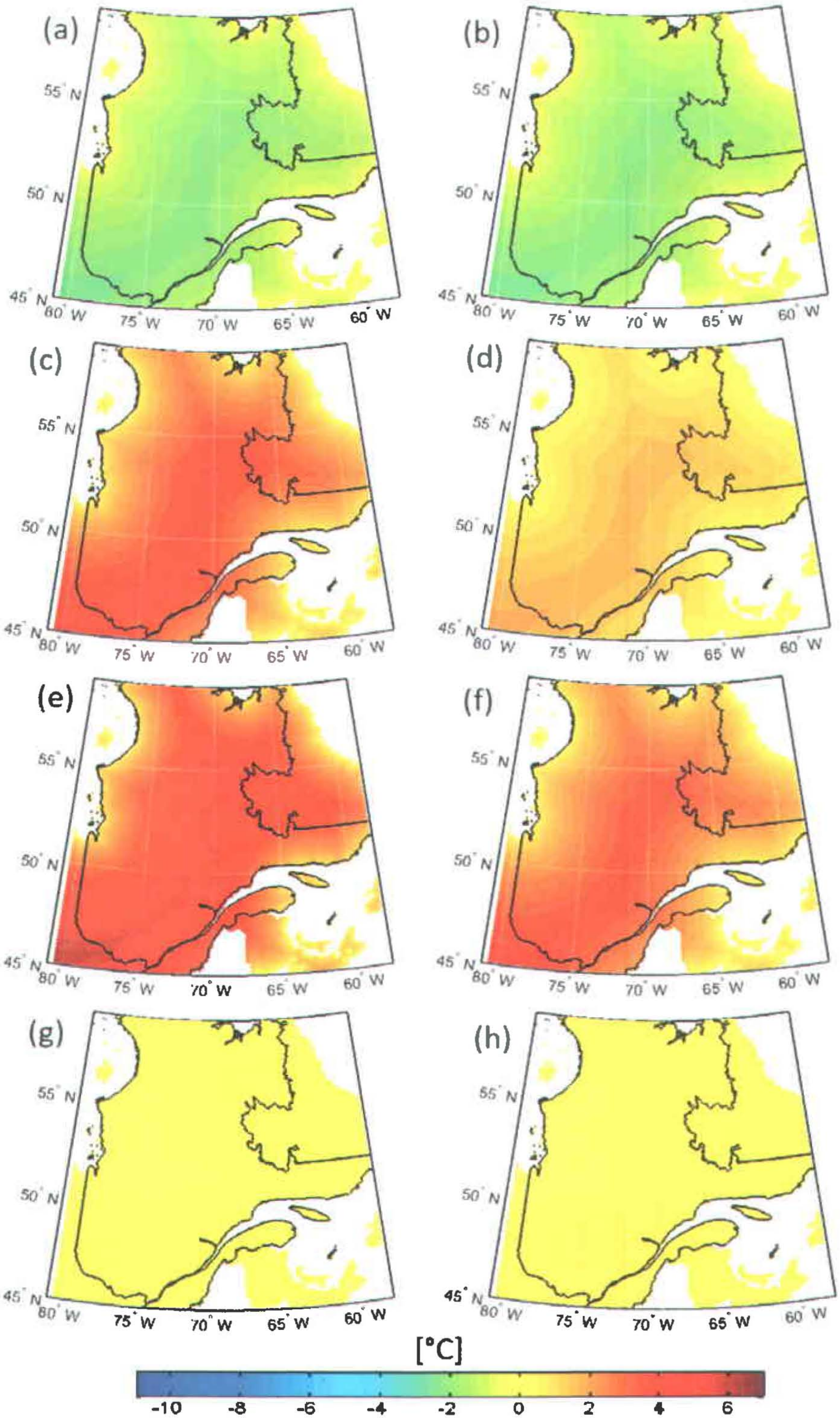

Figure 6. **Effet de la racine carré de la distance à la mer sur la moyenne *a priori* de la température maximale (colonne à gauche) el la température minimale (colonne à droite) pour les mois de Janvier (a et b), Avril (c et d), Juillet (e et f) et Octobre (g et h).**

La Figure 7 montre l'effet de la position géographique sur la moyenne *a priori* de la température maximale (colonne à gauche) et la température minimale (colonne à droite) pour les mois de Janvier (a et b), Avril (c et d), Juillet (e et f) et Octobre (g et h). La figure montre que la variation de la température en fonction de la position géographique est plus importante en hiver qu'en été. On peut voir aussi que l'effet de la longitude est plus marqué en hivers, et durant les autres saisons il est très faible et négligeable devant l'effet de la latitude.

Le modèle spatial commun pour les températures minimales et maximales quotidiennes a été estimé par mois en utilisant le modèle LMC (voir Figure 8). Les différents paramètres estimés pour ce modèle sont donnés à la Figure 9. On peut constater que la proportion de la composante gaussienne est plus faible au cours de l'été que pendant l'hiver, ceci est due au fait que la dépendance spatiale de la température est plus forte durant l'été, ce qui est compatible avec une grande variabilité en hiver (voir Gachon *et al.* 2005).

Ensuite les prédictants à grande échelle ont été estimées en utilisant le modèle spatial selon la théorie des fonctions de covariances régularisées (Goovaerts 2008), et leurs moyennes *a priori* ont été calculées par le GRM. Par la suite, le modèle MMLR de l'équation (8) a été ajusté en utilisant ces prédictants à grande échelle et les prédicteurs décorrélés.

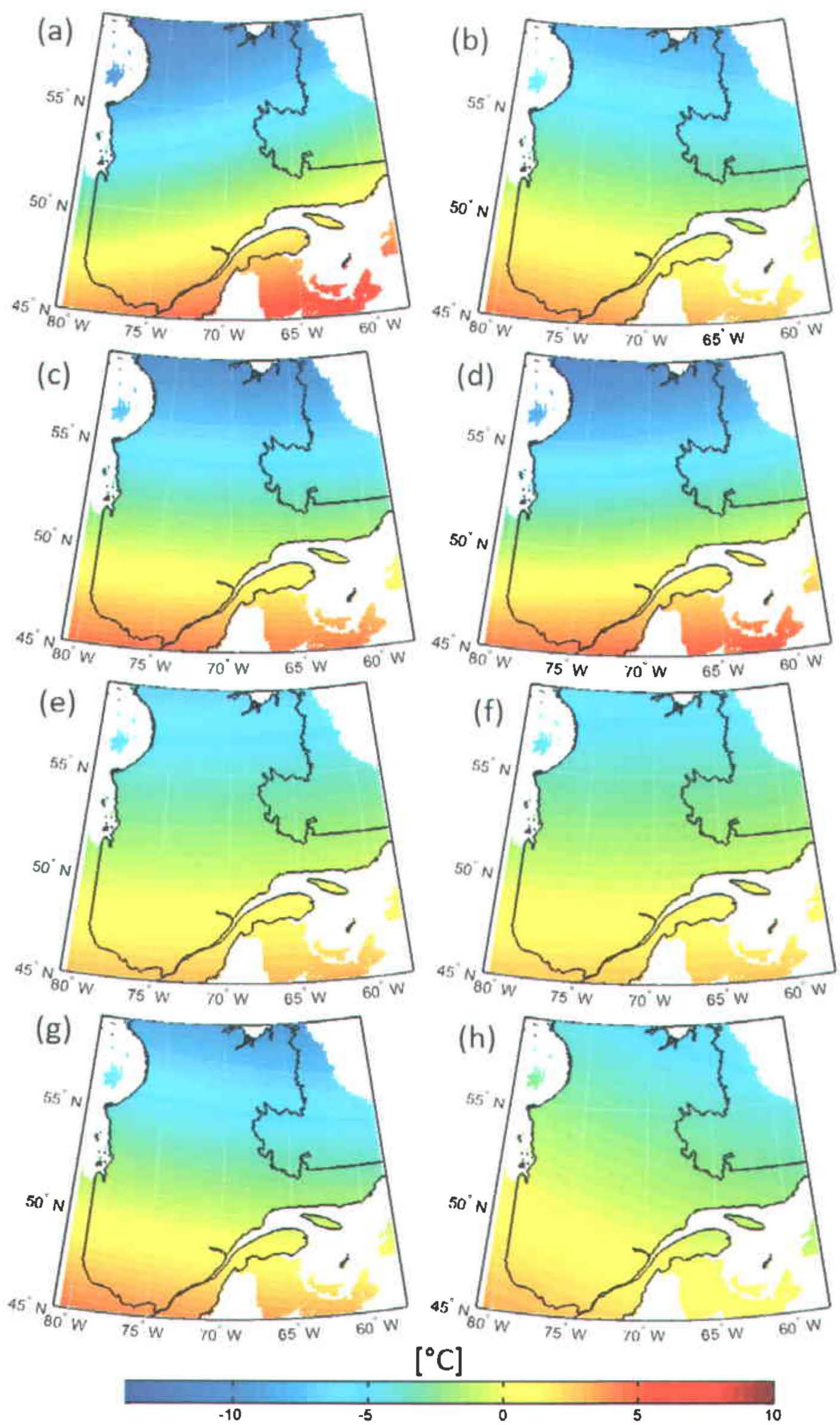

Figure 7. **Effet de la position géographique sur la moyenne *a priori* de la température maximale (colonne à gauche) el la température minimale (colonne à droite) pour les mois de Janvier (a et b), Avril (c et d), Juillet (e et f) et Octobre (g et h).**

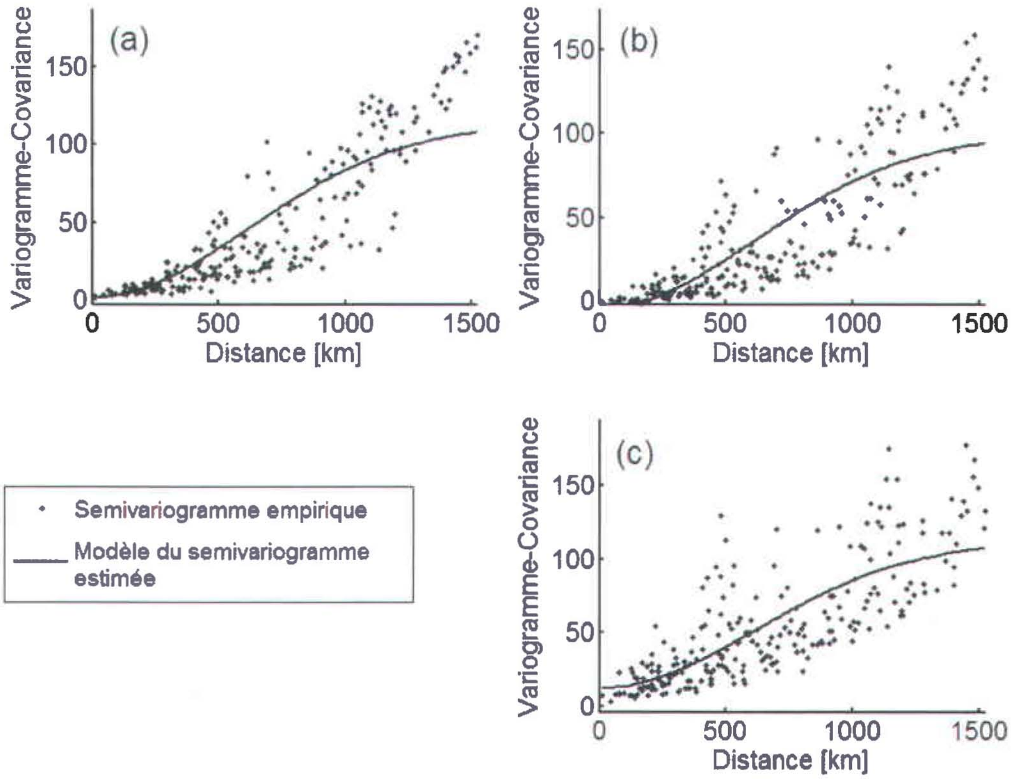

Figure 8. **Modèle estimé du semivariogramme multivarié, et semivariogramme empirique (les points), pour la température maximale (a), les températures maximale et minimale en couple (b) et la température minimale (c).**

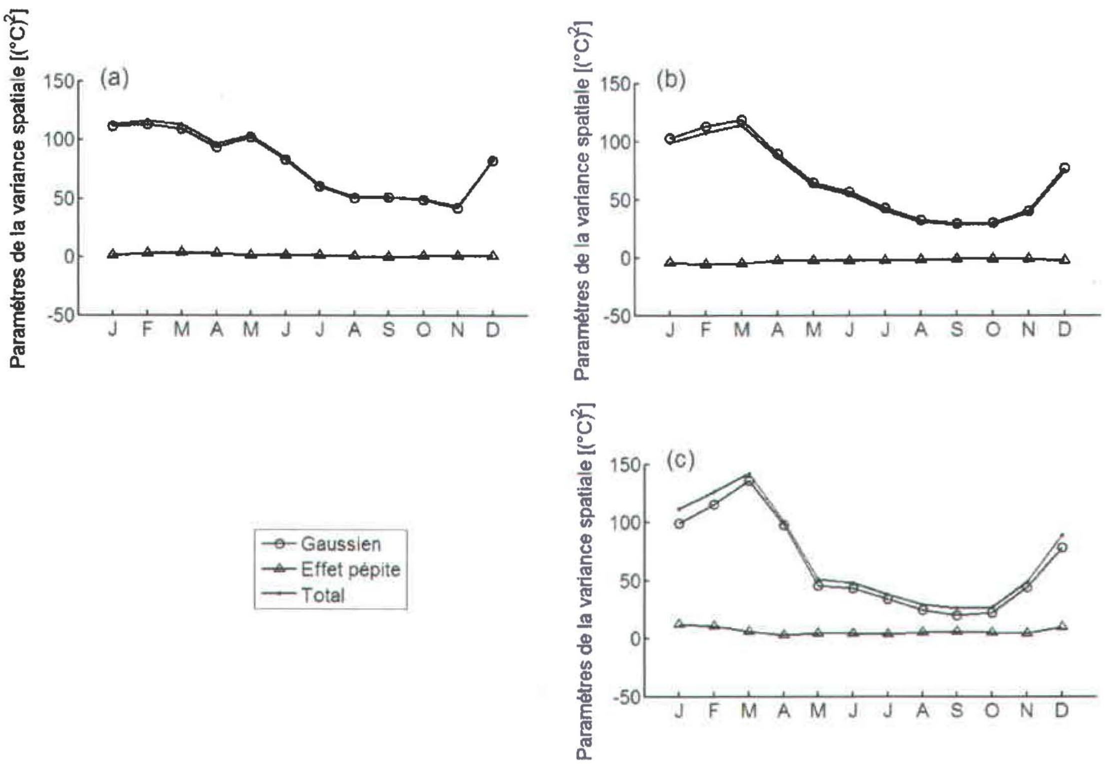

Figure 9. **Évolution des paramètres estimés des différentes composantes du modèle spatial pour chaque mois; (a) les variances mensuelles de la température maximale, (b) les covariances mensuelles entre les températures maximale et minimale et (c) les variances de la température minimale.**

4.2 Résultats de la validation

Le modèle spatial bayésien a été appliqué sur la zone d'étude. À titre d'exemple, quatre dates correspondantes à certaines anomalies observées ont été choisies. Les cartes de températures obtenues pour ces dates sont présentées à la Figure 10. Les figures 10.a et 10.b montrent les cartes estimées pour le 3 Janvier 1981. Cette journée est caractérisée par des températures très basses dans la partie sud de la province de Québec, les températures enregistrées dans les basses latitudes ont été inférieures à celles enregistrées dans des latitudes plus élevées. On peut voir que le modèle a été capable de reproduire l'anomalie observée pour cette journée. Les figures 10.c et 10.d montrent les cartes estimées pour le 17 Janvier 1996, qui est une journée caractérisée par des températures assez importantes dans les basses latitudes, on peut voir que grâce à l'information issue des prédicteurs, le modèle bayésien a été capable de reproduire les températures élevées pour cette journée. De même, pour le 24 Août 1994, l'approche bayésienne a été capable d'estimer correctement les températures maximales très basses. Les deux dernières figures 10.g et 10.h montrent les deux cartes estimées pour le 24 août 1983, pour laquelle le modèle bayésien a bien estimé les températures enregistrées.

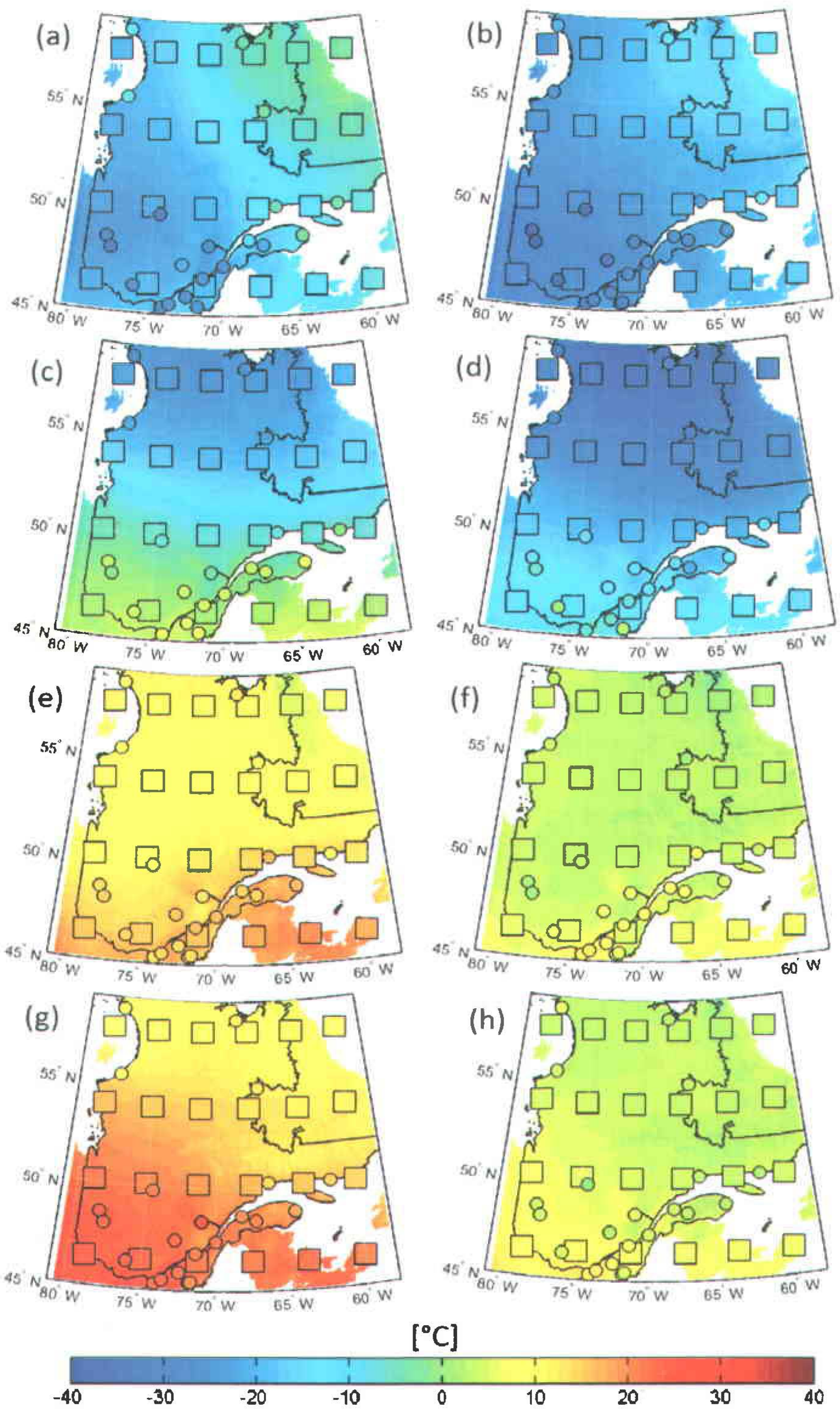

Figure 10. **Exemple de cartes d'estimation des températures maximales (colonne à gauche) et minimales (colonne à droite) pour le 3 Janvier 1981 (a et b), le 17 Janvier 1996 (c et d), le 24 Août 1994 (e et f) et le 24 Août 1983. Les cercles représentent les valeurs observées et les carrés représentent les valeurs à grande échelle sur les points de la grille MCGAO.**

Les résultats ont été validés ensuite par comparaison spatiale des températures observées et estimées sur les 22 stations. Des séries de températures maximales et minimales ont été estimées sur les 22 stations par le modèle spatial bayésien au cours de la période de validation (1991 - 2000).

Le Tableau 6 montre les évaluations de la qualité de réduction d'échelle au moyen des trois critères statistiques employés : la ME, la RMSE, et le pourcentage des observations dans l'intervalle de crédibilité à 95% calculé à base des simulations.

Ces résultats indiquent que les séries estimées sont très proches des valeurs observées. La ME varie entre -0.03°C et 0.64°C pour la température maximale, et entre -0.19°C et 1.02°C pour la température minimale. Cependant, la RMSE varie entre 3.07°C et 3.97°C pour la température maximale et entre 3.16°C et 4.86°C pour la température minimale. Finalement, environ 95% des observations se situent dans les intervalles de crédibilité de 95%, ce qui assure que le modèle est capable de reproduire les fluctuations quotidiennes des températures maximales et minimales au Québec.

Tableau 6. Évaluation de la qualité des séries estimées pour la période de validation (1991-2000). Les critères sont l'erreur moyenne (ME), l'erreur quadratique moyenne (RMSE) et le pourcentage d'observations dans l'intervalle de crédibilité à 95% (% dans CI).

numéro de la station	Température maximale(°C)			Température Minimale(°C)		
	ME	RMSE	% dans CI	ME	RMSE	% dans CI
(1)	0.48	3.34	0.94	0.19	3.65	0.97
(2)	0.60	3.46	0.91	0.78	3.49	0.97
(3)	0.15	3.38	0.92	0.46	3.94	0.95
(4)	0.57	3.59	0.93	0.49	4.27	0.95
(5)	0.24	3.52	0.92	-0.19	3.26	0.98
(6)	0.19	3.58	0.92	0.28	4.33	0.95
(7)	0.30	3.47	0.92	0.60	4.02	0.95
(8)	0.20	3.07	0.93	0.28	3.54	0.96
(9)	0.58	3.23	0.90	0.49	3.26	0.97
(10)	0.25	3.33	0.91	0.35	4.39	0.90
(11)	0.26	3.30	0.91	0.06	3.52	0.96
(12)	0.57	3.48	0.90	-0.02	3.43	0.97
(13)	0.03	3.25	0.91	0.16	3.16	0.97
(14)	-0.03	3.62	0.89	0.24	3.57	0.96
(15)	0.38	3.62	0.89	0.73	4.86	0.89
(16)	0.47	3.74	0.93	0.39	4.20	0.95
(17)	0.39	3.72	0.91	0.49	4.24	0.94
(18)	0.19	3.65	0.93	0.02	4.15	0.95
(19)	0.64	3.49	0.92	0.79	3.64	0.95
(20)	0.26	3.97	0.93	1.02	3.85	0.95
(21)	0.59	3.73	0.93	0.79	3.69	0.95
(22)	0.004	3.46	0.92	0.03	3.85	0.94

Ensuite, pour chaque station, 1000 séries quotidiennes des températures maximales et minimales ont été simulées par la méthode bayésienne. Pour chaque série simulée les indices climatiques présentés au Tableau 4 ont été calculés. Ces mêmes indices ont été calculés pour les 22 séries observées pour avoir des indices climatiques observés. Ceci nous a permis d'avoir 1000 valeurs des RMSE de ces indices pour chaque station (voir Tableau 7). Similairement le Tableau 8 montre les valeurs moyennes des RMSE de ces mêmes indices calculés en utilisant le modèle MMLR. Rappelant que ce dernier a été ajusté pour directement modéliser les prédictants des mêmes 22 stations météorologiques à partir des mêmes prédicteurs décorrélés contrairement au modèle MMLR de la méthode bayésienne qui a été utilisé pour modéliser le lien entre les prédictants à grande échelle et les prédicteurs décorrélés. On peut voir que les résultats de la méthode classique basée (Tableau 8) sur la régression sont légèrement meilleurs à ceux du modèle spatial bayésien (Tableau 7). C'est tout à fait logique puisque le modèle MMLR détermine une relation directe entre les prédicteurs et les prédictants des 22 stations météorologiques, par contre le modèle spatiale bayésien tient compte des informations issues des prédicteurs par leurs liens avec les prédictants à grande échelle. Généralement le modèle MMLR donne de très bons résultats au niveau des stations, c'est pour cette raison qu'il est couramment utilisé en réduction d'échelle. En revanche, ce modèle est très spécifique au prédictants observées et se trouve incapable de fournir des estimations à des endroits non jaugés. Cependant les différences observées entre les résultats des deux méthodes ne sont pas significatives. De plus, la légère perte de précision aux endroits jaugés du modèle bayésien est compensée par l'avantage de fournir des estimations aux endroits non jaugés.

Tableau 7. **Moyenne des RMSE entre les séries observées et simulées pour le modèle spatial bayésien sur la base d'indices climatiques pour chacune des vingt-deux stations météorologiques au cours de la période de validation (1991-2000). MM désigne la moyenne mensuelle et MSD désigne l'écart type mensuel.**

Indices climatiques	Moyennes des RMSE										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
MMTmax (°C)	1.26	1.18	1.02	1.23	1.16	1.06	1.17	1.09	1.18	0.98	0.98
MSDTmax (°C)	0.89	0.84	0.78	0.82	0.83	0.86	0.80	0.76	0.82	0.77	0.83
MMTmin (°C)	1.38	1.40	1.32	1.44	1.19	1.35	1.60	1.31	1.14	1.27	1.10
MSDTmin (°C)	1.33	1.22	1.12	1.08	1.55	1.04	1.11	0.94	1.00	1.10	1.02
FR-TH (jour)	2.75	2.77	2.89	2.86	2.38	2.64	3.17	3.00	2.79	2.69	2.68
Tmax90 (°C)	1.35	1.19	1.12	1.30	1.26	1.25	1.31	1.17	1.36	1.26	1.47
Tmin90 (°C)	1.56	1.29	1.31	1.43	2.12	1.44	1.52	1.48	1.44	1.94	1.20
DTR (°C)	1.01	0.72	0.88	0.98	1.08	0.86	0.97	0.77	0.72	0.78	0.73
GSL (jour)	22.77	14.95	14.94	13.74	16.10	13.65	19.35	16.48	13.25	14.22	15.33
FSL (jour)	3.86	2.04	3.41	4.18	13.08	8.09	2.59	1.21	1.08	6.24	0.98

Indices climatiques	Moyennes des RMSE										
	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)
MM Tmax (°C)	1.21	0.95	1.08	1.20	1.20	1.25	1.13	1.69	1.27	1.40	0.97
MSD Tmax (°C)	0.85	0.78	0.88	0.85	0.88	0.88	0.82	0.94	0.98	0.90	0.75
MM Tmin (°C)	1.17	1.06	1.18	1.54	1.37	1.39	1.40	1.82	1.87	1.66	1.32
MSD Tmin (°C)	1.20	1.48	0.99	1.23	0.94	1.03	1.03	1.30	1.12	1.08	0.99
FR-TH (jour)	2.72	2.68	2.27	2.48	2.26	2.43	2.54	2.82	2.81	2.99	2.39
Tmax90 (°C)	1.30	1.33	1.21	1.62	1.61	1.28	1.25	2.13	1.92	1.78	1.34
Tmin90 (°C)	1.53	1.48	1.34	2.22	1.78	2.32	1.43	1.78	2.28	1.53	1.47
DTR (°C)	1.09	0.64	0.72	1.01	1.00	0.90	0.89	0.88	1.06	0.82	0.87
GSL (jour)	13.94	14.56	17.88	10.74	16.27	44.32	17.67	19.89	16.81	12.50	18.73
FSL (jour)	2.08	1.82	1.26	1.31	1.00	0.00	1.06	0.01	0.88	0.05	0.01

Tableau 8. Moyenne des RMSE entre les séries observées et simulées pour le MMLR sur la base d'indices climatiques pour chacune des vingt-deux stations météorologiques au cours de la période de validation (1991-2000).

Indices climatiques	Moyenne des RMSE										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
MMTmax (°C)	1.26	1.18	1.02	1.22	1.16	1.06	1.17	1.09	1.18	0.98	0.98
MSD Tmax (°C)	0.89	0.84	0.78	0.82	0.83	0.86	0.80	0.76	0.82	0.77	0.84
MM Tmin (°C)	1.38	1.40	1.32	1.44	1.19	1.34	1.60	1.31	1.13	1.28	1.10
MSD Tmin (°C)	1.33	1.22	1.12	1.08	1.55	1.04	1.10	0.94	0.99	1.10	1.02
FR-TH (jour)	2.76	2.77	2.90	2.85	2.38	2.65	3.16	2.99	2.79	2.70	2.68
Tmax90 (°C)	1.34	1.19	1.13	1.30	1.27	1.26	1.30	1.17	1.36	1.26	1.47
Tmin90 (°C)	1.56	1.29	1.31	1.43	2.12	1.44	1.52	1.48	1.44	1.94	1.20
DTR (°C)	1.01	0.72	0.87	0.99	1.07	0.85	0.97	0.77	0.72	0.78	0.73
GSL (jour)	22.56	14.88	14.95	13.69	16.16	13.74	19.37	16.53	13.20	14.08	15.35
FSL (jour)	3.82	2.01	3.34	4.11	13.06	8.27	2.59	1.19	1.11	6.01	1.00

Indices climatiques	Moyenne des RMSE										
	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)
MMTmax (°C)	1.21	0.95	1.07	1.20	1.20	1.25	1.13	1.69	1.26	1.40	0.97
MSD Tmax (°C)	0.85	0.78	0.87	0.85	0.87	0.88	0.82	0.94	0.97	0.91	0.74
MM Tmin (°C)	1.16	1.05	1.19	1.54	1.37	1.39	1.40	1.82	1.86	1.66	1.33
MSD Tmin (°C)	1.20	1.48	0.98	1.23	0.94	1.03	1.04	1.30	1.12	1.08	0.99
FR-TH (jour)	2.70	2.69	2.29	2.48	2.26	2.42	2.53	2.83	2.81	3.01	2.39
Tmax90 (°C)	1.29	1.33	1.21	1.62	1.61	1.27	1.26	2.13	1.93	1.79	1.34
Tmin90 (°C)	1.53	1.48	1.34	2.23	1.76	2.33	1.42	1.79	2.29	1.52	1.48
DTR (°C)	1.08	0.64	0.72	1.01	0.99	0.90	0.89	0.87	1.06	0.82	0.87
GSL (jour)	13.68	14.67	18.05	10.82	15.10	44.39	17.73	19.91	16.75	12.58	18.28
FSL (jour)	2.11	1.84	1.21	1.32	0.95	0.00	1.05	0.01	0.89	0.06	0.01

5. DISCUSSIONS

Les résultats ont montrés que le modèle spatial bayésien donne des valeurs consistantes avec les valeurs observés. En plus, les résultats on été satisfaisants par la comparaison avec une méthode classique de réduction d'échelle basée sur la régression.

L'avantage premier de ce modèle bayésien c'est qu'il permet de fournir des estimations à des endroits non jaugé, ce qui n'est pas le cas pour les méthodes classiques basées sur la régression. De plus, il est basé sur une approche complètement probabiliste qui lui permet de reproduire toute la distribution *a posteriori* des prédictants. En particulier, il est capable de reproduire la moyenne, ainsi que la variabilité des valeurs observées. De ce fait, il présente un autre avantage par rapport aux méthodes de régression. En effet les méthodes de régression sous-estiment la variance des prédictants, parce qu'elles reproduisent la partie centrale des valeurs observées conditionnellement à des prédicteurs sélectionnés. Ainsi, il n'est pas nécessaire de recourir à des techniques d'inflation ou de randomisation souvent appliquées aux méthodes de régression pour reproduire la variance observée (Fasbender et Ouarda 2010).

L'application de l'ACP sur l'ensemble des prédicteurs NCEP originaux permet d'avoir des prédicteurs décorelés. Cette étape est très importante, parce que les prédicteurs vont être introduits dans un modèle de régression linéaire (le modèle MMLR). En effet, dans un modèle de régression linéaire, si les prédicteurs sont corrélés ils visent à expliquer la même chose. Dans ce cas la manière de la discrimination de la variance entre les prédicteurs peut ne pas être significative.

En outre l'ACP permet aussi de réduire le nombre des prédicteurs, et donc réduire le nombre des paramètres dans le modèle MMLR.

Ce travail peut être amélioré par une étude préliminaire qui vise à sélectionner les prédicteurs les plus pertinents. En fait, il est préférable d'utiliser des prédicteurs primaires directement fournis par le modèle AOGCM, au lieu d'utiliser des prédicteurs paramétrés qui sont dépendants d'autres variables. Aussi, il est plus judicieux de choisir des prédicteurs qui corrèlent le mieux les prédictants pour pouvoir reproduire la plupart de la variabilité locale. Cependant, il se peut qu'un prédicteur ne semble pas être important pour simuler le climat local actuel, mais le changement de ce même prédicteur dans des conditions climatiques futures porte le signale d'un changement climatique local.

Les données NCEP sont utilisées pour la calibration et la validation de ce modèle spatial bayésien. Ces données renferment des incertitudes qui peuvent donc influencer négativement les résultats lors de l'utilisation de ce modèle bayésien. Par conséquent les études de sensibilité de ce modèle bayésien au prédicteurs NCEP sont très bénéfiques pour l'élaboration des scénarios.

6. CONCLUSIONS

Un modèle spatial bayésien a été présenté dans ce travail dans le but d'affiner les données AOGCM aux températures maximale et minimale dans la province de Québec.

La moyenne *a priori* des températures est décrite pour chaque mois, au moyen du GRM qui tient compte des effets locaux. Par conséquent, l'information issue du GRM permet de reproduire les évolutions saisonnières à petite échelle spatiale.

Des prédictants à grande échelle qui ont la même résolution spatiale et se situent aux mêmes endroits que les données AOGCM, sont utilisés par le modèle bayésien pour transférer les informations issues des prédicteurs vers les prédictants originaux. Ceci permet au prédicteurs d'induire les fluctuations quotidiennes à grande échelle à travers leur relation avec les prédictants à grande échelle. Cette relation est décrite par le modèle MMLR. Du fait que le modèle bayésien combine les deux informations issues des deux modèles GRM et MMLR, la distribution *a postriori* résultante reflète à la fois les évolutions saisonnières à petite échelle et les fluctuations quotidiennes à grande échelle.

D'autre part, les prédictants à grande échelle sont reliés aux prédictants originaux par leurs positions spatiales. Ainsi la relation est décrite à l'aide d'une fonction de covariance régularisée qui a été calculée en utilisant le modèle LMC. Ceci permet au modèle spatial bayésien de fournir des estimations à n'importe qu'elle endroit de la zone d'étude.

Le modèle a été appliqué pour estimer les températures maximale et minimale dans une zone qui couvre presque la totalité de la province de Québec, Canada. Les résultats obtenus ont été pertinents. En effet, le modèle est capable de bien estimer les cartes de température et ce même pour des dates correspondantes à certaines anomalies observées. En plus les résultats on été satisfaisants par la comparaison avec une méthode classique de réduction d'échelle statistique. Aussi, les séries simulées par le modèle ont été comparées avec les séries observées au moyen de trois critères statistiques. Les résultats ont montrés que le modèle donne des valeurs consistantes avec les valeurs observés.

Enfin, le modèle spatial bayésien présenté dans ce travail est basé sur une approche purement statistique probabiliste, et donc, malgré ces avantages, il ne peut pas échapper à l'inconvénient majeur de toutes les méthodes de réduction d'échelle statistique, c'est de supposer avant de les utiliser que dans des conditions du climat futur, la relation prédictants / prédicteurs demeure valide.

BIBLIOGRAPHIE

Baigorria, G. A., J.W.Hanse, N. Ward, J.W. Jones, and J. J.O'Brien, 2008. Assessing predictability of cotton yields in the Southeastern United States based on regional atmospheric circulation and surface temperatures. *Journal of Applied Meteorology and Climatology*, **47**, 76-91.

Bates, B., S. Charles, and J. Hughes, 1998. Stochastic downscaling of numerical climate model simulations. *Environmental Modelling and Software*, **13 (3-4)**, 325-331.

Bellone, E., J. Hughes, and P. Guttorp, 2000. A hidden Markov model for downscaling synoptic atmospheric patterns to precipitation amounts. *Climate Research*, **15**, 1-12.

Benestad, R. E., D. Chen, and I. Hanssen-Bauer, 2008. Empirical-Statistical Downscaling. *World Scientific Publishing Company.*

Benestad, R. E., 2002. Empirically downscaled temperature scenarios for Northern Europe based on a multi-model ensemble. *Climate Research*, **21**, 105-125.

Benestad, R. E., 2007. Novel methods for inferring future changes in extreme rainfall over Northern Europe. *Climate Research*, **34**, 195-210.

Biau, G., E. Zorita, H. Von Storch and H. Wackernagel, 1999. Estimation of precipitation by kriging in EOF space. *Journal of Climate,* **12**, 10701085.

Buishand, T.A. and T. Brandsma, 2001. Multi-site simulation of daily precipitation and temperature in the Rhine basin by nearest-neighbour resampling. *Water Resources Research,* **37**, 2761-2776.

Chilès, J. P., and P. Delfiner, 1999. Geostatistics: Modeling Spatial Uncertainty. *Wiley-Interscience*, 720 pp.

Conway, D., R.L. Wilby, and P.D. Jones, 1996. Precipitation and airflow indices over the British Isles. *Climate Research,* **7**, 169-183.

DAI CGCM3 Predictors, 2008. Sets of Predictor Variables Derived From CGCM3 T47 and NCEP/NCAR Reanalysis, version 1.1, April 2008, Montreal, QC, Canada, 15 p.

Fasbender, D., T.B.M.J. Ouarda, 2010. Spatial Bayesian Model for Statistical Downscaling of AOGCM to Minimum and Maximum Daily Temperatures. *Journal of Climate,* **23**, 5222–5242.

Gachon, P., A. St-Hilaire, T. B. M. J. Ouarda, VTV Nguyen, C. Lin, J. Milton, D. Chaumont, J. Goldstein, M. Hessami, T.D. Nguyen, F. Selva, M. Nadeau, P. Roy, D. Parishkura, N. Major, M. Choux & A. Bourque, 2005. A first evaluation of the strength and weaknesses of statistical downscaling methods for simulating extremes over various regions of eastern Canada.*Sub-component, Climate Change Action Fund (CCAF), Environment Canada, Final report, Montréal, Québec, Canada*, 209 p.

Goovaerts, P., 2008. Kriging and semivariogram deconvolution in the presence of irregular geographical units. *Mathimatical Geosciences*, **40**,101-128.

Grotch, S. L. and M. C. MacCracken, 1991. The use of General Circulation Models to predict regional climatic change. *Journal of Climate*, **4 (3)**, 286-303.

Hellstrom, C., D. Chen, C. Achberger and J. Raisanen, 2001. A comparison of climate change scenarios for Sweden based on statistical and dynamical downscaling of monthly precipitation. *ClimateResearch,* **19**, 45-55.

Herrera, E., T.B.J.M. Ouarda and B. Bobée, 2006. Méthodes de désagrégation appliquées aux Modèles du Climat Global Atmosphère-Océan (MCGAO). *Journal of Water Science*, **19 (4)**, 297-312.

Hessami, M., P. Gachon, T. B. M. J. Ouarda, and A. St.-Hilaire, 2008. Automated regression-based statistical downscaling tool. *Environmental Modelling and Software*, **23**, 813-834.

Hewitson, B.C. and R.G. Crane, 1996. Climate downscaling: Techniques and applications, *Climate Research,* **7**, 85-95.

Hundecha, Y., and A. Bárdossy, 2005. Trends in daily precipitation and temperature extremes across western Germany in the second half of the 20th century. *International Journal of Climatology*, 25, 1189-1202.

Huth, R. and J. Kysely, 2000. Constructing site-specific climate change scenarios on a monthly scale. *Theorical and Applied Climatology.*, **66 (1-2)**, 13-27

Huth, R. and Pokoma, 2004. Parametric versus non-parametric estimates of climatic trends. *Theoretical and Applied Climatology,* **77**, 107-112.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bulletin of the American Meteorological Society*, **77**, 437-471.

Kistler, R., and Coauthors, 2001. The NCEP–NCAR 50-Year Reanalysis: Monthly means CD-ROM and documentation. *Bulletin of the American Meteorological Society*, **82**, 247-267.

Lim, Y.-K., D. W. Shin, S. Cocke, T. E. LaRow, J. T. Schoof, J. J. O'Brien, and E. P. Chassignet, 2007. Dynamically and statistically downscaled seasonal simulations of maximum surface air temperature over the southeastern United States. *Journal of Geophysical Research*, **112**, D24102, doi:10.1029/2007 JD008764.

Linderson M.L., C. Achberger and D. Chen, (2004). Statistical downscaling and scenario construction of precipitation in Scartia, southern Sweden. *Nordic Hydrology,* **35(3)**, 261-278.

Michelangeli, P. A., M. Vrac, and H. Loukos, 2009. Probabilistic downscaling approaches: Application to wind cumulative distribution functions. *Geophysical Research Letters*, **36**, L11708, doi:10.1029/2009GL038401.

Scinocca, J. F., N. A. McFarlane, M. Lazare, J. Li, and D. Plummer, 2008. The CCCma third generation AGCM and its extension into the middle atmosphere. *Atmospheric Chemistry and Physics*, **8 (23)**, 7055-7074.

Stehlik, J., and A. Bárdossy, 2002. Multivariate stochastic downscaling model for generating daily precipitation series based on atmospheric circulation, *Journal of Hydrology,* **256**, 120-141.

Tebaldi, C., L. Mearns, D. Nychka, and R. Smith, 2004. Regional probabilities of precipitation change: A Bayesian analysis of multimodel simulations. *Geophysical Research Letters*, **31**, L24213, doi:10.1029/2004GL021276.

Vincent, L., X. Zhang, B. Bonsal, and W. Hogg, 2002. Homogenization of daily temperatures over Canada. *Journal of Climate*, **15**,1322-1334.

Von Storch, H., H. Langeberg, and F. A. Feser, 2000. A Spectral nudging technique for dynamical downscaling purposes. *Monthly Weather Review,* **128(10)**, 3664-3673.

Von Storch., H., H. Langeberg, and F. A. Feser, 2000. Spectral nudging technique for dynamical downscaling purposes. *Monthly Weather Review,* **128(10)**, 3664-3673.

Vrac, M., and P. Naveau, 2007. Stochastic downscaling of precipitations: From dry events to heavy rain falls. *Water Resources Research*, **43**, W07402, doi:10.1029/2006WR005308.

Vrac, M., M. Stein, and K. Hayhoe, 2007. Statistical downscaling of precipitations through non homogeneous stochastic weather typing. *Climate Research*, **34**, 169-184.

Wilby, R.L., C.W. Dawson, and E.M. Barrow, 2002. SDSM -a decision support tool for the assessment of regional climate change impacts. *Envirornmental Modelling & Software,* **17,** 147-159.

Wilby, R.L., 1998. Statistical downscaling of daily precipitation using daily airflow and seasonal teleconnection indices. *Climate Research,* **10**, 163-178.

Wilby, R.L., O.J. Tomlinson and C.W. Dawson, 2003. Multi-site simulation of precipitation by conditional resampling. *Climate Research,* **23**, 1199-1202.

Wilks, O.S. and R.L. Wilby, 1999. The weather generation game: a review of stochastic weather models. *Progress in Physical Geography,* **23**, 329-357.