

Synthèse des techniques d'homogénéisation des séries climatiques et analyse d'applicabilité aux séries de précipitations

CLAUDIE BEAULIEU, TAHA B. M. J. OUARDA & OUSMANE SEIDOU

Chaire en hydrologie statistique Hydro-Québec/CRSNG, Chaire du Canada en Estimation des Variables Hydrologiques, Université du Québec, INRS-ETE, 490 Rue de la Couronne, Québec, Québec G1K 9A9, Canada
claudie_beaulieu@ete.inrs.ca

Résumé Au cours des dernières décennies, plusieurs méthodes d'homogénéisation ont été développées pour corriger les ruptures artificielles dans les séries climatiques. Elles ont été développées dans divers pays, pour différentes situations et avec des domaines d'application différents. Cet article présente une revue exhaustive des méthodes d'homogénéisation des séries climatiques publiées dans la littérature. Une analyse critique de ces méthodes ainsi qu'une réflexion sur leur applicabilité aux séries de précipitations totales annuelles sont également présentées.

Mots clés homogénéisation; série climatique; saut; tendance; test statistique; méthodes Bayésiennes

A review of homogenization techniques for climate data and their applicability to precipitation series

Abstract During the last decades, a considerable effort was spent on the development of homogenization techniques that can identify and correct anthropogenic bias in climatic series. These methods were developed in various countries, for different contexts and for different domains of applicability. The present paper is an exhaustive review of published homogenization techniques for climatic series. A critical analysis of the described methods and a discussion on their applicability to total annual precipitation are also presented.

Key words homogenization; climatic series; jump; trend; statistical test; Bayesian methods

INTRODUCTION

Les conditions de mesure d'une station peuvent varier au cours du temps. Les modifications des conditions de mesure (déplacement de stations, remplacement d'instruments de mesure, changement des heures d'observations ou encore modification de l'environnement immédiat de l'instrument de mesure) peuvent introduire des ruptures artificielles dans les données qui ne reflètent pas les variations réelles du climat. De ce fait, il arrive régulièrement que les séries de données climatiques contiennent des ruptures artificielles qui sont dues aux modifications des conditions de mesure et qui ne reflètent pas les variations réelles du climat.

Au cours des dernières décennies, plusieurs méthodes ont été développées pour corriger les ruptures artificielles des séries climatiques. En effet, le besoin de longues séries de données climatiques fiables se fait de plus en plus sentir dans divers domaines. Par exemple, les études des changements climatiques nécessitent la création de bases de données complètes avec lesquelles on pourra analyser de manière adéquate le signal climatique, suivre son évolution dans le temps et prévoir les changements futurs avec une incertitude minimale. Il est alors très important de développer des techniques robustes de détection de ces biais artificiels pour que les données utilisées se rapprochent le plus possible des observations qui auraient été faites sans perturbation des conditions de mesure. Le processus de détection et de correction des ruptures d'origine non climatique est appelé homogénéisation.

Ce travail présente une revue bibliographique des méthodes d'homogénéisation des données climatiques décrites dans la littérature. On en fera ensuite une analyse

critique, suivie d'une réflexion sur leur applicabilité aux séries de précipitations enregistrées au Québec. Cette analyse facilitera le choix de la méthode à utiliser étant donnée l'information disponible de la station à homogénéiser et de ses stations voisines. Cette revue de littérature constitue un travail préliminaire à une comparaison des méthodes d'homogénéisation des précipitations au Québec méridional.

Le reste de l'article est organisé comme suit: (a) explication de la problématique de l'homogénéisation des données; (b) présentation des méthodes d'homogénéisation utilisées pour les données climatiques; (c) présentation des études comparatives réalisées; (d) discussion des méthodes d'homogénéisation; (e) une analyse d'applicabilité pour l'homogénéisation des séries de précipitations totales annuelles et, finalement, (f) conclusions.

L'HOMOGENEISATION DES DONNEES

Le terme "inhomogénéités" représente les variations non naturelles qui sont causées par des modifications dans les réseaux d'observations. L'homogénéisation consiste à détecter puis à corriger ces variations. Les sauts de moyenne et les tendances sont les deux types d'inhomogénéités les plus communes. Par exemple, les mesures de précipitations sont perturbées par le vent et l'effet Venturi produit au dessus du cône du pluviomètre. Donc, toute modification ayant un effet sur le vent induira des sauts dans les données (e.g. un changement d'emplacement, de la forme du capteur ou de sa hauteur au dessus du sol). Une modification dans l'environnement immédiat d'une station, induite par exemple par l'urbanisation, la reforestation, ou l'industrialisation, suscite des changements plus graduels (e.g. Peterson *et al.*, 1998) et ce, en particulier sur les températures. Une tendance est plus délicate à quantifier car il faut identifier correctement le début et la fin de cette tendance ainsi que son amplitude.

En pratique, il est assez difficile de déterminer si une rupture dans une série représente un changement dans le climat régional ou une inhomogénéité. Deux types d'informations permettent de pencher pour l'une ou l'autre des explications: les enregistrements aux stations voisines (s'il y en a) et les métadonnées (s'il y en a également).

Utilisation de l'information aux stations voisines

En homogénéisation, on compare souvent les données d'une station avec celles des stations voisines pour éviter qu'un changement climatique soit interprété comme une inhomogénéité. La série dont on veut vérifier l'homogénéité est appelée *série de base*. Les *séries voisines* sont supposées être climatiquement similaires à la série de base (souvent des observations de stations géographiquement voisines). De plus, elles doivent être homogènes, sinon des inhomogénéités dans une de ces dernières pourraient être attribuées à la série de base. Or, il est difficile de garantir que les séries voisines sont exemptes d'inhomogénéités. Il existe des techniques qui permettent de minimiser les inhomogénéités potentielles et de former une *série de référence* homogène. Une *série de référence* est une fonction d'une ou de plusieurs séries voisines. On peut trouver dans Peterson *et al.* (1998) une revue des techniques utilisées pour minimiser les inhomogénéités potentielles de ces séries.

Pour les précipitations, il est souvent difficile de trouver des stations voisines qui appartiennent à la même région climatique que la série de base étant donnée la forte variabilité spatio-temporelle de cette variable. Lorsque les stations voisines utilisées sont trop éloignées ou d'altitudes très différentes, les inhomogénéités présentes dans la série de base risquent d'être masquées par une trop grande divergence entre ces caractéristiques. La qualité de l'homogénéisation des précipitations dépend alors de la densité du réseau d'observations et de la topographie de la zone où ont été prises les données. De plus, le degré d'importance de ces caractéristiques dépend fortement du pas de temps des données que l'on veut homogénéiser.

Les métadonnées

Les métadonnées sont des informations historiques sur les conditions dans lesquelles les données à une station ont été enregistrées. Elles sont constituées des enregistrements de la station, des annuaires météorologiques, des fiches d'inspection, des photographies de la station et de son environnement, d'une entrevue avec la personne responsable d'une station, etc. (Aguilar *et al.*, 2003). En homogénéisation, les métadonnées sont consultées pour expliquer la cause d'une inhomogénéité. Cependant, les informations sont souvent incomplètes, inexistantes, erronées ou remplies d'informations non pertinentes. Il se peut alors qu'une inhomogénéité détectée soit impossible à valider par les métadonnées. À l'inverse, un changement recensé dans les métadonnées peut ne pas avoir influencé sensiblement la mesure.

METHODES UTILISEES POUR HOMOGENEISER DES DONNEES CLIMATIQUES

Diverses techniques d'homogénéisation ont été développées pour accommoder différents types de facteurs tels que la variable à homogénéiser, la variabilité spatiale et temporelle des données selon l'endroit où les stations sont situées, la longueur des séries et le nombre de données manquantes, les métadonnées disponibles et la densité du réseau d'observations (Aguilar *et al.*, 2003). Les techniques d'homogénéisation varient également selon l'objectif pour lequel on les applique et la philosophie de chaque équipe de travail.

Malgré la grande diversité des méthodes d'homogénéisation, elles peuvent néanmoins être classifiées en deux catégories principales qui les rejoignent toutes: méthodes subjectives ou objectives. Lorsque l'emplacement d'une discontinuité est détecté à l'œil nu sur un graphique, la méthode appartient à la classe subjective même si des tests statistiques sont appliqués par la suite. Par contre, les méthodes objectives ne dépendent pas du jugement de l'utilisateur pour localiser les inhomogénéités. Les méthodes Bayésiennes, qui seront présentées plus loin, pourraient être considérées comme subjectives ou objectives selon qu'elles incorporent ou non l'opinion du modélisateur par le biais d'une distribution *a priori* informative. Comme aucune des méthodes Bayésiennes présentées dans cet article n'utilise de distributions *a priori* informatives, elles sont présentées parmi les approches objectives.

Cette section présente les approches subjectives et objectives pour l'homogénéisation des données climatiques rencontrées dans la littérature. Par la suite, les études

comparatives réalisées sur ces méthodes sont discutées dans le but de mieux les classer. Enfin, les méthodes utilisées pour homogénéiser les précipitations sont également présentées par pays.

Approches subjectives

La plupart des approches subjectives sont basées sur l'examen visuel d'observations cumulées, de différences ou de rapports, ou encore de déviations cumulées entre la série de base et des séries voisines pour que les inhomogénéités soient plus facilement détectables visuellement. Dans cette section, elles sont regroupées selon la variable que l'on inspecte sur le graphique.

Examen graphique des doubles accumulations Plusieurs techniques se basent sur une analyse visuelle de valeurs cumulées. Par exemple, l'analyse des doubles accumulations (Kohler, 1949) consiste à faire une régression linéaire simple sur les valeurs cumulées de la série de base (y) en fonction des valeurs accumulées d'une série de référence (x). Ensuite, on produit un graphique des couples de points (x, y) sur lesquels on superpose la droite de régression. Si les séries sont homogènes, les points sont disposés aléatoirement autour de la droite de régression. Par contre, un changement à l'une ou l'autre des deux stations se remarque par une cassure de la pente. La technique des ellipses de Bois (Bois, 1986) ainsi que la méthode du vecteur des indices annuels (Brunet-Moret, 1979) sont des extensions de l'analyse des doubles accumulations.

Examen graphique des déviations cumulées L'analyse graphique des déviations cumulées a été développée pour homogénéiser des séries de précipitations totales annuelles par Craddock (1979). On calcule les déviations de la série de base par rapport au ratio \bar{y}/\bar{x} , où y représente la série de base et x une série de référence homogène. Un graphique des déviations cumulées en fonction du temps est produit. Un saut de moyenne dans la série de base se remarque facilement par un changement de pente sur le graphique des déviations cumulées en fonction du temps. Craddock (1979) répète le processus plusieurs fois avec différentes séries de référence et compare les graphiques obtenus. La méthode des sommes cumulatives parallèles se base sur le même principe. De plus, une technique permettant d'estimer l'amplitude du saut ainsi que son intervalle de confiance est présentée par Rhoades & Salinger (1993). On y présente également diverses façons de déterminer le seuil significatif du saut dans le cas de stations isolées.

Examen graphique des séries de différences ou de rapports Jones *et al.* (1986) proposent une méthode qui consiste à examiner graphiquement les séries de différences entre la série de base et ses voisines. Les points de discontinuités sont identifiés par un examen graphique des différences mensuelles entre la série de base et chaque série de référence. Lorsque le même saut se répète sur plusieurs graphiques, alors il provient de la série de base. Ensuite, les métadonnées sont consultées et si les positions d'inhomogénéités détectées sont validées, la série est ajustée. Par ailleurs, Tayanc *et al.* (1998) proposent d'appuyer cette méthode par deux tests non paramétriques. Lorsqu'un saut est apparent par un examen visuel, les tests de Kruskal-Wallis et de Wald-Wolfowitz

décrits par Tayanc *et al.* (1998) sont appliqués sur les séries différenciées. Les résultats des deux tests donnent le niveau de signification des inhomogénéités détectées visuellement. Pour les précipitations, on peut analyser graphiquement les ratios de précipitations totales annuelles entre la série de base et une série de référence (Craddock, 1979). Pour une série homogène, les rapports devraient être distribués aléatoirement autour d'une certaine constante. Par contre, une discontinuité entraîne une déviation des valeurs par rapport à cette constante. Le but de cette méthode n'est pas de fournir un outil de correction, mais plutôt de montrer qu'il y a une discontinuité à expliquer dans la série.

Examen graphique des données filtrées Le filtre de Kolmogorov-Zurbenko est applicable à des données contenant des tendances et des cycles saisonniers (Zurbenko *et al.*, 1996). Ce filtre a été testé directement sur des données prises en altitude (température et humidité). Le filtre est appliqué sur les données brutes. Lorsque la fenêtre mobile s'approche d'une région dans laquelle il y a une discontinuité, elle peut être identifiée par une croissance de la variabilité ou de l'amplitude de la pente de la série filtrée. Donc, les discontinuités de la série deviennent évidentes en représentant la série filtrée sur un graphique. Leur amplitude est estimée en calculant la différence des moyennes des valeurs avant et après le saut des données filtrées. Le seuil significatif des sauts est obtenu par un test sur la variance de la série filtrée (Zurbenko *et al.*, 1996).

Approches objectives

Les approches objectives peuvent se classer selon le modèle qui représente la rupture dans la série de base. On distingue: (a) les modèles de régression linéaire avec des changements dans un ou plusieurs coefficients de régression; (b) les modèles avec une forme paramétrique différente de la régression linéaire; et (c) les modèles sans forme paramétrique explicitement spécifiée. À l'intérieur de chaque classe, des sous-classes sont définies en fonction du nombre de ruptures (une ou plusieurs) et de la nature de la rupture (changement abrupt de la moyenne ou de pente). Dans les sections suivantes, les méthodes inventoriées sont regroupées selon la classification décrite ci-haut.

(a) Modèles de régression linéaire avec des changements dans un ou plusieurs paramètres

(i) Un changement de moyenne

Plusieurs méthodes sont basées sur la recherche d'un saut dans une série de différences ou de rapports formée de la série de base avec une série de référence. Cela permet d'inclure l'information d'une série de référence et d'avoir un modèle très simple:

$$q_i = \begin{cases} \tau + e_i & i = 1, \dots, p \\ \tau + \delta + e_i & i = p + 1, \dots, n \end{cases} \quad (1)$$

où q_i est la valeur de la série de différences/ratios au temps i , τ est la moyenne, p est la position d'un saut potentiel et δ son amplitude. Les résidus du modèle sont représentés par e_i . Pour détecter la position du saut et/ou estimer son amplitude, plusieurs techniques peuvent être employées.

Dans un premier temps, lorsque la position p est connue, par l'information obtenue des métadonnées, on peut appliquer un test de différences de moyennes local pour

vérifier la significativité de l'amplitude (Karl & Williams, 1987). On forme des séries de différences (température) ou de différences entre les logarithmes des rapports (précipitation) de la série de base avec ses voisines. Ensuite, un intervalle de confiance basé sur le test de Student est calculé pour la température et sur le test de Wilcoxon pour les précipitations. Par la suite, une procédure est appliquée pour corriger la série de base des sauts significatifs.

Lorsque la position du saut est inconnue, plusieurs méthodes peuvent résoudre ce problème. La méthode Bayésienne univariée pour un saut proposée par Lee & Heghinian (1977) et expliquée par Ouarda *et al.* (1999) peut être employée. La distribution *a priori* est déterminée par la série de base. La procédure calcule, pour chaque année, la probabilité que le saut se produise à ce moment étant données les observations recueillies. De la même manière, l'amplitude du saut possède sa fonction de densité *a posteriori*. Le seuil significatif de l'amplitude du saut est déterminé en calculant un intervalle de crédibilité Bayésien. Une extension de cette méthode pour détecter un saut dans une série de précipitations ou de débits a été développée (Perreault *et al.*, 1999). Selon cette méthode, on ne doit pas faire l'hypothèse qu'il y a eu un changement de moyenne dans la série. Dans cette version révisée de la méthode, les distributions *a priori* sont plus flexibles.

Des tests statistiques classiques peuvent également résoudre ce problème. Le test le plus utilisé est celui développé par Alexandersson (1986) pour homogénéiser des séries de précipitations totales annuelles ou de températures moyennes annuelles (Alexandersson & Moberg, 1997) en se basant sur un test du rapport des vraisemblances. On calcule une statistique pour toutes les positions de sauts possibles. La statistique maximale correspond à l'emplacement le plus probable pour un saut et son seuil significatif se détermine à partir d'un tableau de valeurs critiques (Alexandersson, 1986; Khaliq & Ouarda, 2006). En principe, les mathématiques du test sont applicables à une discontinuité par série. En pratique, on applique successivement le test jusqu'à ce que tous les segments de la série soient considérés homogènes. La procédure consistant à chercher le maximum de la statistique en fonction de la position s'applique à d'autres statistiques. La méthode de Jaruskova (1996) et les tests sur les déviations cumulées (Buishand, 1982, 1984) se basent sur le même principe. Enfin, le test de Student séquentiel consiste à appliquer un test de Student pour comparer cinq observations avant et après chaque position de saut potentielle (Gullett *et al.*, 1990). Les endroits où la statistique calculée dépasse la valeur critique de la distribution de Student avec huit degrés de liberté correspondent aux positions de sauts significatifs.

Une autre façon de procéder est de détecter un saut dans une série de base qui ne se produit pas dans la série de référence. Un modèle représentant une telle série peut s'exprimer ainsi:

$$y_i = \begin{cases} \tau + \beta x_i + e_i & i = 1, \dots, p \\ \tau + \delta + \beta x_i + e_i & i = p+1, \dots, n \end{cases} \quad (2)$$

où y_i est la valeur de la série de base au temps i , x_i est la valeur de la série de référence au temps i . Pour détecter la position du saut et/ou estimer son amplitude, on peut employer l'approche bivariée et la méthode Bayésienne bivariée pour un saut.

L'approche bivariée est une technique statistique classique basée sur le rapport des vraisemblances (Maronna & Yohai, 1978; Potter, 1981; Young, 1993). La série de

base et la série de référence appartiennent à une distribution normale bivariée et on teste l'hypothèse que les moyennes sont différentes. Pour détecter la position du changement, on utilise la procédure consistant à chercher le maximum de la statistique en fonction de la position.

La méthode Bayésienne bivariée pour un saut constitue une extension de la méthode Bayésienne univariée pour un saut. Asselin *et al.* (1999) ont développé ce modèle qui diffère du précédent en incluant une série de référence. Les observations de la série de base et de la série de référence sont distribuées selon la loi normale bivariée. Les distributions *a priori* des paramètres de ce modèle sont les mêmes que celles du modèle univarié présenté dans la section précédente. Les détails concernant les densités marginales et la densité conjointe *a posteriori* de la date du changement et de son amplitude sont présentés par Asselin *et al.* (1999). Une autre approche Bayésienne pour un saut basée sur la distribution normale multivariée a été proposée pour détecter un changement de moyenne commun dans plusieurs séries hydrologiques corrélées (Perreault *et al.*, 2000). Dans ce modèle, on teste l'hypothèse qu'il n'y a pas de changement dans les séries contre l'hypothèse qu'il y a un saut simultané à chacun des sites. Par la suite, Seidou *et al.* (2007) ont développé un modèle de régression Bayésienne. Un changement de moyenne dans un modèle de régression multivariée est considéré. Cette méthode est basée sur des hypothèses *a priori* non informatives et les distributions *a posteriori* sont obtenues par l'échantillonnage de Gibbs. Encore une fois, la méthode vérifie la présence d'une rupture concomitante à plusieurs séries.

(ii) *Un changement de moyenne et/ou de tendance*

Il existe des méthodes d'homogénéisation basées sur des modèles de régression linéaire qui recherchent à la fois des sauts et des tendances. Ce sont des modèles dans lesquels on suppose les propriétés de normalité, d'indépendance et d'hétéroscédasticité des données, et qui sont résolus par des techniques moindres carrés standard. Ceci s'applique à toutes les techniques basées sur des modèles de régression présentées dans cette section. Par exemple, la régression multiple se base sur l'application de plusieurs modèles de régression pour homogénéiser des séries de températures (Vincent, 1998). Le premier représente une série homogène. Lorsque les résidus sont indépendants, le modèle appliqué ajuste bien les données. Dans le cas contraire, le modèle n'ajuste pas bien les données et il pourrait y avoir une inhomogénéité dans la série de base. Le test de Durbin-Watson ou un intervalle de confiance sur l'autocorrélation des résidus permettent de vérifier cette hypothèse. Lorsque la série de base est considérée inhomogène, on ajuste un autre modèle pour déterminer le type de discontinuité dans la série de base. Pour identifier un saut, le modèle suivant est appliqué:

$$y_i = \begin{cases} \tau + \beta_1 x_{i1} + \dots + \beta_k x_{ik} + e_i & i = 1, \dots, p \\ \tau + \delta + \beta_1 x_{i1} + \dots + \beta_k x_{ik} + e_i & i = p+1, \dots, n \end{cases} \quad (3)$$

où y_i est la valeur de la série de base au temps i , x_{ik} est la valeur de la série de référence k au temps i . Il y a n observations et k séries de référence. L'emplacement du saut, p , est déterminé en ajustant le modèle pour toutes les positions possibles et en sélectionnant celui avec la plus petite somme des carrés résiduelle. Le choix de la position du saut se valide avec le test de Fisher. L'estimation de l'amplitude du saut est donnée par δ et son seuil significatif se calcule avec la statistique de Student. Des

modèles représentant une série avec une tendance et une série qui contient une tendance et un saut sont ajustés par la suite (Vincent, 1998).

Un modèle de régression à deux phases permet également de détecter un changement de moyenne et/ou de tendance dans une série (Solow, 1987). D'abord, un modèle représentant une série homogène est ajusté. Le deuxième modèle ajusté représente une série dans laquelle il y a une discontinuité à un certain point p :

$$y_i = \begin{cases} \tau_1 + \lambda_1 i + e_i & i = 1, \dots, p \\ \tau_2 + \lambda_2 i + e_i & i = p+1, \dots, n \end{cases} \quad (4)$$

où y_i est la valeur de la série de base au temps i , τ_1 et τ_2 sont les moyennes avant et après le changement, λ_1 et λ_2 sont les tendances avant et après le changement et p est la position du changement. Les résidus du modèle sont représentés par e_i . L'emplacement du saut est déterminé selon les moindres carrés. Plusieurs modifications ont été apportées:

- Easterling & Peterson (1995) appliquent le modèle sur une série de différences créée avec des stations voisines, ne forcent pas les deux droites de régression à se joindre, appliquent la technique itérativement pour détecter plusieurs sauts et évaluent les seuils significatifs par une procédure de permutation multiple (Mielke *et al.*, 1981).
- Lund & Reeves (2002) proposent une statistique de Fisher révisée.
- Wang (2003) propose un modèle dans lequel les pentes sont égales avant et après la rupture.

Par ailleurs, la régression multiphase a été développée afin de détecter des sauts et des tendances dans des données de température mensuelle (Gullett *et al.*, 1991). Cette méthode consiste à ajuster sept modèles de régression différents dans lesquels la série de base constitue la variable dépendante et les variables explicatives sont le temps et des séries voisines. Le premier représente une série homogène tandis que les six autres évoquent une série de base qui contient de une à six discontinuités. Celui qui comprend une discontinuité se représente par:

$$y_i = \begin{cases} \tau_1 + \lambda_1 i + \beta_1 x_{i1} + \dots + \beta_k x_{ik} + e_i & i = 1, \dots, p \\ \tau_2 + \lambda_2 i + \beta_1 x_{i1} + \dots + \beta_k x_{ik} + e_i & i = p+1, \dots, n \end{cases} \quad (5)$$

où y_i est la valeur de la série de base au temps i , x_{ik} est la valeur de la série de référence k au temps i . À chaque étape, on détermine les positions qui ajustent le mieux les données au sens des moindres carrés. Le nombre de discontinuités dans la série est déterminé en choisissant le modèle le plus simple qui améliore significativement le dernier modèle ajusté. Par la suite, le seuil significatif de chaque inhomogénéité potentielle est déterminé par la statistique de Student.

(iii) Plusieurs sauts

Dans le cas où on recherche plusieurs inhomogénéités dans la série de base, on ne fait plus face à un problème de test d'hypothèse, mais plutôt au choix de modèle ainsi qu'à la combinatoire. Dans cette section, les techniques de détection de sauts multiples sont présentées en séparant les problèmes de choix de modèle et de combinatoire.

Choix de modèle Le choix des emplacements de sauts et l'estimation de leurs amplitudes peut se faire en utilisant une procédure de vraisemblance pénalisée par un

critère comme le critère de Schwarz (1978), d'Akaike (1973) ou de Caussinus & Lyazrhi (1997). Les critères de Schwarz (BIC) et d'Akaike (AIC) sont des classiques en choix de modèle et sont utilisés pour faire de la sélection de prédicteurs en régression multiple. Le critère de Caussinus & Lyazrhi est spécifiquement conçu pour les modèles à rupture. D'ailleurs, il performe beaucoup mieux que AIC et BIC sur le modèle en marche d'escalier (Caussinus & Mestre, 2004). La méthode développée par Caussinus & Mestre (2004) se base sur ce critère pour le choix de modèle combinée à un algorithme de programmation dynamique. Les séries sont homogénéisées en étant tour à tour série de base alors que toutes les autres deviennent des séries voisines. Des séries de différences ou de rapports sont formées et lorsqu'un changement revient dans toutes les séries, on considère qu'il est attribuable à la série de base.

Une méthode Bayésienne ayant pour paramètres le nombre de changements et la position de ces derniers peut également être utilisée pour détecter des sauts multiples. L'approche Bayésienne multi-sauts est une adaptation de la méthode présentée dans Fearnhead (2006) qui cherche à détecter un changement de paramètres dans une série de données distribuée selon une loi quelconque. Ce modèle a été adapté par Seidou & Ouarda (2007) au cas où les données sont normales et la moyenne de la série climatique considérée varie en fonction de plusieurs covariables. En sortie, le modèle fournit la densité de probabilité *a posteriori* du nombre de changements, et la position des sauts étant donné le nombre de changements le plus probable.

Pour détecter des sauts multiples, il est également possible d'appliquer des tests localement comme le test de Student ou le test Scheffé. Il est alors nécessaire de posséder des métadonnées complètes qui peuvent indiquer les positions à tester. Par ailleurs, on peut également appliquer une méthode développée pour détecter plusieurs sauts à la fois comme la méthode de Szentimrey (1997, 1999). Cette dernière se base sur le principe du rapport de vraisemblance, mais dans le cas de sauts multiples. La particularité de cette méthode est le traitement des séries voisines qui ne sont pas considérées homogènes. La méthode consiste à attribuer un poids aux séries de différences formées à partir des séries voisines. Le poids optimal à allouer est obtenu en minimisant la variance des résidus des séries de différences. La procédure de détection s'applique et retourne les emplacements des changements potentiels. Il est également possible d'obtenir des intervalles de confiance pour l'emplacement du saut et son amplitude.

Combinatoire Ensuite, on fait face au problème de combinatoire. Puisqu'il serait impraticable de faire une recherche exhaustive de toutes les combinaisons de ruptures possibles, la plupart des auteurs utilisent une technique pas-à-pas simple. Cela consiste à détecter une rupture, la conserver en mémoire, en détecter une seconde, et ainsi de suite. La plupart des techniques développées pour détecter un seul saut s'appliquent de cette façon en pratique. Par exemple, la procédure de segmentation de série hydrométéorologique (Hubert *et al.*, 1989) utilise cet algorithme. La technique cherche la segmentation optimale au sens des moindres carrés, celle pour laquelle l'écart entre la série et la segmentation considérée est minimale. La procédure consiste à calculer l'écart quadratique entre la série et la segmentation considérée. Une fois que la segmentation optimale est identifiée, l'égalité des moyennes des segments consécutifs est testée au moyen du test des contrastes de Scheffé. La segmentation obtenue est acceptée seulement si le test de Scheffé est significatif. Une généralisation de cette

méthode a été formulée par Kehagias *et al.* (2006). Dans cette dernière, on choisit le nombre de sauts en se basant sur le critère de Schwarz. Les techniques pas-à-pas deviennent problématiques lorsque les ruptures consécutives sont trop proches.

Lorsque les données sont gaussiennes, la technique de programmation dynamique fournit un algorithme équivalent à la recherche exhaustive de toutes les combinaisons de ruptures (Hawkins, 2001). Cette technique est plus chère que du pas-à-pas simple, mais optimale.

(b) Modèles avec une forme paramétrique différente de la régression linéaire

Dans cette section, d'autres techniques qui peuvent être utilisées pour l'homogénéisation des données sont présentées. Thompson (1984) a développé une méthode pour homogénéiser une série de précipitations totales mensuelles basée sur un processus aléatoire de Poisson. La fréquence des précipitations est caractérisée par un processus de Poisson et son amplitude suit indépendamment une loi exponentielle. La série est divisée en plusieurs segments selon les dates de déplacement de la station. Le segment avec le plus d'années d'observations constitue la série de référence et les autres séries seront comparées à celle-ci. Le test d'homogénéité basé sur le test du rapport de vraisemblance donné par Thompson (1984), vérifie si les estimateurs du maximum de vraisemblance varient entre les segments. Pour vérifier que la série corrigée est effectivement homogène, la statistique de Mann-Whitney est calculée afin de tester l'hypothèse que les deux séries font partie de la même population. Ce processus est répété pour tous les segments.

Une méthode basée sur l'analyse des ondelettes a été proposée pour ajuster les observations de températures anciennes qui sont souvent mal documentées (Datsenko *et al.*, 2002). Elle se définit par trois idées principales. En premier lieu, les séries climatiques sont décomposées en différentes échelles de temps et l'ajustement se fait sur ces différentes échelles. En deuxième lieu, on suppose que les inhomogénéités dans les séries sont nombreuses et qu'elles peuvent varier continuellement. En troisième lieu, on fait l'hypothèse que les données prises avant que l'on adopte des méthodes de calcul modernes sont de mauvaise qualité. La méthode proposée ajuste les séries même si l'amplitude des homogénéités détectées est inférieure à un certain seuil significatif en se basant sur l'idée que les ajustements apportés sur des données de bonne qualité seront pratiquement négligeables. L'analyse se réalise simultanément sur la plus longue période d'observations commune de la série de base ainsi que plusieurs séries prises à des stations voisines à l'aide de fonctions d'ondelettes. Les séries sont divisées en deux parties : la première contient des données qui renferment des biais et des erreurs tandis que les données de la partie récente sont supposées homogènes. L'information historique peut servir à déterminer la date de séparation des séries. La méthode consiste à estimer les statistiques de la partie récente des séries, et à déterminer les ajustements optimaux à faire sur la première partie des séries par une stratégie Bayésienne.

(c) Modèles sans forme paramétrique explicitement spécifiée

Les modèles basés sur l'application successive de tests non paramétriques entrent dans cette catégorie. Par exemple, la procédure du test de Mann-Whitney séquentiel corrige les inhomogénéités pas-à-pas (Lanzante, 1996). Cette section inclut également les

méthodes hybrides qui sont constituées de l'application simultanée de plusieurs tests paramétriques et/ou non paramétriques successifs qui ne vérifient pas les mêmes hypothèses. Ceci permet de réduire le nombre de fausses alarmes et d'éviter de confondre un saut et une tendance. Boroneant & Tomozeiu (1999) proposent d'appliquer successivement les tests de Pettitt, de Mann-Kendall et d'Alexandersson pour l'homogénéisation de séries de températures mensuelles. Le test de Pettitt est une approche non paramétrique pour identifier un changement de moyenne ou de variance dans une série (Pettitt, 1979). Le test de Mann-Kendall présenté dans Boroneant & Tomozeiu (1999) permet également de tester la stabilité de la moyenne et de la variance d'une série d'observations. Il s'agit d'une procédure non paramétrique dans laquelle on compare chaque observation de la série de base avec ses observations précédentes. Le test d'Alexandersson, tel que présenté précédemment, est à son tour appliqué. La méthode proposée par Wijngaard *et al.* (2003) consiste à appliquer quatre tests d'homogénéité différents (test d'Alexandersson, test sur l'étendue des déviations cumulées standardisées, test de Pettitt et rapport de Von Neumann) sur des données annuelles et classer les séries selon le nombre de tests significatifs. Les tests sont présentés en détails dans Wijngaard *et al.* (2003). La station est classifiée selon le nombre de tests significatifs au seuil de 1%. Elle appartient à la première catégorie lorsqu'au plus un test est significatif, à la deuxième catégorie lorsque deux tests rejettent l'hypothèse nulle et à la troisième catégorie lorsque trois ou quatre des tests sont significatifs. Les séries classées dans la catégorie 1 peuvent être utilisées ultérieurement car elles semblent homogènes. Les séries appartenant aux catégories 2 ou 3 doivent être vérifiées en consultant les métadonnées.

ETUDES COMPARATIVES REALISEES

Quelques études comparatives ont été effectuées sur des séries climatiques simulées (Ducré-Robitaille *et al.*, 2003; Slonosky *et al.*, 1999; Lubes-Niel *et al.*, 1998; Zurbenko *et al.*, 1996; Easterling & Peterson, 1992). Étant donné que les résultats de ces études dépendent fortement de l'effectif et des caractéristiques des séries synthétiques générées, les conclusions doivent être considérées avec prudence.

La plus récente (Ducré-Robitaille *et al.*, 2003) compare huit techniques différentes: la méthode d'Alexandersson pour les sauts et pour les tendances, la régression multiple, la régression à deux phases (version proposée par Easterling & Peterson; 1995), le test de Student séquentiel, le test de Mann-Whitney séquentiel et la méthode Bayésienne univariée pour un saut. Le test de Mann-Whitney séquentiel utilisé ici est similaire au test de Student séquentiel en remplaçant seulement le test de Student par le test de Mann-Whitney. Ces méthodes ont été comparées sur des séries synthétiques de température avec un nombre aléatoire de sauts. Les résultats se résument ainsi:

- (a) Pour identifier les séries homogènes, la méthode d'Alexandersson pour les sauts, la régression multiple et l'approche Bayésienne ont été performantes.
- (b) Les sauts de grandes amplitudes sont toujours bien détectés. Par contre, les petits sauts sont plus facilement détectables par les méthodes d'Alexandersson et Bayésienne en utilisant une série de référence.
- (c) Les méthodes possédant la capacité de déceler une tendance (Alexandersson pour tendances et régression multiple) sont portées à mal classer les petits sauts.

- (d) La méthode d'Alexandersson pour les sauts, la régression multiple et la régression à deux phases ont offert la meilleure performance pour identifier des sauts multiples.
- (e) Le modèle de régression à deux phases présenté par Easterling & Peterson (1995) se montre très sensible, mais la version révisée de ce modèle (Lund & Reeves, 2002) pourrait améliorer la performance du modèle.
- (f) Le test de Student séquentiel est désavantagé parce que la fenêtre mobile n'est pas assez large.
- (g) Les méthodes incluant une ou des séries de référence performant mieux du fait qu'elles expliquent une part de la variabilité de la série de base.

Une étude comparative de trois méthodes d'homogénéisation a été réalisée sur des séries de pression moyenne en surface en Europe (Slonosky *et al.*, 1999). Les inhomogénéités ont été détectées au pas de temps annuel et l'ajustement des séries est fait mensuellement. Les techniques utilisées sont l'analyse graphique des séries de différences, la méthode de Caussinus & Mestre et la méthode d'Alexandersson. La documentation a été consultée pour valider les inhomogénéités identifiées. Les grandes conclusions de cette étude sont:

- La méthode de Caussinus & Mestre requiert un grand nombre de stations voisines pour fournir des facteurs de correction robustes. La méthode est problématique lorsque la station de base se trouve à un endroit isolé ou avec seulement quelques voisines (moins de six).
- La méthode d'Alexandersson est performante avec de bonnes séries de référence. L'efficacité de cette technique dépend alors de la disponibilité de séries de référence.
- L'examen graphique des séries de différences a donné les meilleurs résultats.

Lubes-Niel *et al.* (1998) ont comparé le test de corrélation sur les rangs de Kendall, le test de Pettitt, le test de Buishand, la procédure Bayésienne de Lee & Heghinian et la procédure de segmentation des séries hydrométéorologiques de Hubert *et al.* (1989) sur des séries synthétiques de précipitations annuelles. Ces méthodes ne fournissent pas les mêmes éléments de réponse quant à l'homogénéité d'une série. En effet, le test de Kendall sert à vérifier la présence d'une tendance dans les données, les tests de Pettitt et de Buishand permettent de positionner une rupture de moyenne et d'évaluer son seuil significatif, la méthode Bayésienne permet d'estimer la distribution de la position de la rupture et la procédure de segmentation est conçue pour la recherche de changements multiples de moyenne dans une série. Le but de ce travail n'était pas de choisir la meilleure technique, mais plutôt d'étudier la puissance et la robustesse de ces cinq méthodes dans des conditions contrôlées. La puissance des méthodes a été évaluée en introduisant des ruptures de moyenne (saut et tendance) dans des séries normales indépendantes. La robustesse a été évaluée en générant des séries normales indépendantes, lognormales indépendantes, gamma indépendantes, normales autocorrélées et des séries avec des ruptures d'écart-type. Voici les conclusions générales de cette étude:

- (a) Les tests de Kendall, de Pettitt et de Buishand donnent des résultats conformes à l'erreur de type 1 utilisée sur les séries stationnaires.
- (b) La méthode Bayésienne avec une loi *a priori* non informative détecte un fort pourcentage (40%) de ruptures sur les séries stationnaires. Elle devrait être utilisée

seulement lorsque les séries étudiées ont une forte probabilité de rupture *a priori*.

- (c) Avec la procédure de segmentation, le pourcentage de séries rejetées alors qu'elles respectent l'homogénéité est acceptable.
- (d) La puissance de toutes les méthodes est supérieure à 50% pour détecter des sauts de moyenne dépassant 75% de la valeur de l'écart-type.
- (e) Les méthodes reposant sur l'indépendance des observations (Pettitt, Buishand, méthode Bayésienne) ne sont pas robustes lorsque les données contiennent une tendance. Ces méthodes montrent également un manque de robustesse lorsque les données sont autocorrélées.
- (f) Les méthodes basées sur la normalité des données (Buishand, méthode Bayésienne, procédure de segmentation) offrent une bonne performance lorsque cette hypothèse n'est pas respectée.
- (g) Les méthodes exigeant l'invariance de l'écart-type (procédure Bayésienne, Pettitt, Buishand) sont relativement robustes lorsque les séries contiennent des ruptures d'écart-type et que la moyenne est stationnaire.

Easterling & Peterson (1992) ont choisi un ensemble de méthodes indépendantes des métadonnées disponibles, facilement automatisables et qui considèrent une série de référence formée de stations voisines, et les ont appliquées sur des séries de températures réelles et synthétiques. La méthode d'Alexandersson, le test bivarié, le test de Student séquentiel, l'analyse des doubles accumulations, les sommes cumulatives parallèles et la régression à deux phases ont été utilisées. La performance de chacune des techniques à identifier correctement l'année d'une discontinuité et à estimer l'amplitude du facteur de correction à appliquer a été évaluée. Les méthodes qui ont bien identifié une discontinuité ont ensuite été testées pour détecter des inhomogénéités multiples. Par la suite, elles ont été appliquées aux données réelles pour comparer les résultats avec l'information historique disponible. Les grandes conclusions de ce travail sont:

- Toutes ces méthodes se sont montrées efficaces pour détecter un saut d'amplitude 2σ .
- La régression à deux phases semble plus appropriée pour détecter des tendances que des sauts.
- Pour identifier un saut de faible amplitude, la méthode d'Alexandersson, le test bivarié et le test de Student ont offert la meilleure performance.
- Le test de Student est moins approprié pour détecter des sauts multiples. En effet, les valeurs critiques augmentent considérablement vers la fin de la série et ainsi le test est inapproprié pour examiner les segments plus courts de la série de base.
- La technique d'Alexandersson et le test bivarié ont été retenus pour détecter des sauts multiples. Lorsque le nombre de discontinuités augmente, la méthode d'Alexandersson fournit une meilleure performance.
- Par ailleurs, l'efficacité de la méthode d'Alexandersson diminue lorsque les séries de référence ont des écarts-type de 0.5 à 1.5 fois celui de la série de base.
- Pour les changements plus graduels, le test bivarié et la méthode d'Alexandersson tendent à identifier un saut dans le milieu de cette tendance.
- Les deux méthodes précédentes ont été appliquées à des stations possédant de l'information historique et ont identifié des discontinuités non documentées.

Enfin, Zurbenko *et al.* (1996) ont comparé le filtre Kolmogorov-Zurbenko avec le

critère de Schwarz pour détecter des discontinuités dans des séries de données en altitude. Ils ont appliqué les techniques sur des séries simulées contenant des patrons saisonniers. Les deux méthodes ont la propriété d'estimer directement la dimension du modèle sans passer par des tests d'hypothèse. Le filtre a offert une meilleure performance que le critère de Schwarz sur les séries avec des tendances cycliques. Par contre, le critère de Schwarz offre une performance supérieure lorsque les données sont distribuées normalement. Les deux méthodes étaient moins performantes pour détecter de petites discontinuités. Zurbenko *et al.* (1996) suggèrent d'utiliser le critère de Schwarz pour faire une analyse exploratoire en support à d'autres méthodes et à l'information historique.

ANALYSE CRITIQUE DES MÉTHODES

Les forces et les faiblesses des méthodes d'homogénéisation dépendent de plusieurs facteurs et plusieurs regroupements (qui se recoupent) sont possibles. Ils seront organisés de trois façons: (a) leur caractère objectif ou subjectif; (b) l'utilisation de l'information des stations voisines et des métadonnées; et (c) le nombre et le type des inhomogénéités pour lesquels leur théorie est conçue.

Classification selon leur caractère objectif ou subjectif

Méthodes objectives Les méthodes dites objectives se basent sur des considérations mathématiques pour repérer les inhomogénéités et peuvent ne pas les détecter si certaines de leurs hypothèses sont violées. Certaines méthodes (approche bivariable, méthode d'Alexandersson, test de Student séquentiel, les méthodes basées sur la régression) requièrent que la variance de la série soit constante. Les méthodes non paramétriques comme le test de Wilcoxon séquentiel ont l'avantage d'être basées sur des concepts de statistique non paramétrique robustes et indépendants de toute distribution. De plus, les observations extrêmes ont moins d'effet sur le résultat du test. Cependant, il peut arriver qu'on perde en puissance ce qu'on gagne en robustesse. Une étude de robustesse quant aux postulats exigés pourrait être effectuée pour vérifier jusqu'à quel point leur non-respect affecte les résultats. Ensuite, l'application séquentielle des tests de Wilcoxon et de Student peut modifier l'erreur de type 1, mais ce problème peut se régler en diminuant le seuil critique de chaque méthode. Enfin, la méthode d'Alexandersson et le test de Student séquentiel sont sensibles au coefficient de variation de la série (écart-type/moyenne). Le test d'Alexandersson se montre plus ou moins sensible selon que ce coefficient est faible ou élevé (Alexandersson & Moberg, 1997). Une variance forte fera en sorte que la statistique de Student soit faible et ainsi la série est déclarée homogène sans l'être nécessairement.

Autant les approches basées sur la régression que celles basées sur les méthodes statistiques classiques (sauf les tests non-paramétriques) reposent sur la normalité des données. De plus, la segmentation des séries sous-entend la normalité des données. Cependant, même si les données ne sont pas normales, les séries de différences ou de rapports peuvent tout de même respecter la normalité. Par ailleurs, la méthode Bayésienne multi-sauts possède l'avantage de pouvoir être adaptée sans difficulté à une distribution non-normale. De plus, les codes mettant en œuvre toutes les méthodes

statistiques classiques s'exécutent rapidement, tandis que le temps de calcul des méthodes Bayésiennes est considérablement plus élevé. Il demeure néanmoins raisonnable pour une application sur un petit ensemble de séries. Enfin, plusieurs techniques tendent à détecter des inhomogénéités au début et à la fin des séries. Pour contrer ce problème, les inhomogénéités inexplicables identifiées en début ou en fin de série peuvent être ignorées.

Méthodes subjectives Le fait qu'elles soient basées sur le jugement de l'utilisateur constitue à la fois un point fort et un point faible. La personne qui utilise une technique subjective peut évaluer le poids à donner à certains éléments selon leur importance (la qualité des métadonnées et des séries voisines ou l'amplitude d'une rupture par rapport à la variance de la série). Par ailleurs, une personne non experte risque de commettre des erreurs de jugement. Néanmoins, ces méthodes constituent une bonne façon de s'approprier les données à titre exploratoire.

Classification selon l'utilisation de l'information aux stations voisines et des métadonnées

Méthodes utilisant l'information d'une ou plusieurs séries voisines L'efficacité des méthodes qui utilisent une ou plusieurs séries voisines dépend de la quantité de stations voisines et de la qualité de leurs données. Une inhomogénéité éventuelle dans une série voisine aura moins d'impact sur l'ajustement du modèle lorsque les séries voisines sont nombreuses. Par exemple, lorsqu'il y a une seule série voisine (e.g. analyse des doubles accumulations ou analyse graphique des rapports), il est impossible de savoir laquelle de la série de base ou de la série de référence contient une inhomogénéité. De ce fait, ces méthodes se montrent efficaces lorsque les dates de changements sont connues (Kohler, 1949). De plus, si la série de référence est mal choisie, des variations naturelles causant des changements de pente peuvent être interprétées comme des inhomogénéités. La méthode de Tayanc est également adéquate lorsque les métadonnées sont exhaustives. En effet, lorsque la station visée n'est pas documentée, cette approche permet seulement de classer la station comme douteuse. Les méthodes graphiques incluant plusieurs séries voisines seront beaucoup plus appropriées si les métadonnées sont incomplètes (sommes cumulatives parallèles, analyse graphique des séries de différence).

Quoiqu'il soit quasi impossible de s'assurer de l'homogénéité d'une station voisine, il est important d'en tester l'hypothèse pour éviter d'interpréter un changement climatique comme une inhomogénéité. Par contre, lorsque toutes les stations du réseau étudié ont des changements d'instruments simultanés, les tests d'homogénéité absolus sont préférables aux tests relatifs qui se basent sur des stations voisines. Lorsque le nombre de stations à homogénéiser est élevé, on peut considérer des méthodes comme celles de Szentimrey ou celle de Caussinus & Mestre qui, tout en homogénéisant simultanément un ensemble de stations, ne se basent pas sur l'hypothèse que la ou les séries voisines utilisées sont homogènes.

Méthodes utilisant les métadonnées Ces méthodes comprennent les approches subjectives, mais aussi les approches objectives dont les résultats doivent être corroborés par les métadonnées avant qu'une correction ne soit effectuée. Les points

forts et les points faibles de ces méthodes sont les mêmes que ceux des approches subjectives.

Classification selon le nombre et le type des inhomogénéités pour lesquels leur théorie est conçue

Méthodes conçues à l'origine pour détecter une seule inhomogénéité à la fois

Les mathématiques de la plupart de ces méthodes visent à détecter un seul saut de moyenne dans la série. Il est possible que ces méthodes soient moins performantes à détecter des sauts multiples plutôt qu'un seul saut de moyenne. En présence de sauts multiples, la correction des données peut être biaisée. L'estimation d'un saut peut être faussée par la présence d'autres sauts dans la série. Lorsqu'une méthode développée pour détecter un changement de moyenne est appliquée séquentiellement pour détecter des sauts multiples, une attention particulière doit être accordée à la façon d'estimer l'amplitude des sauts pour ne pas laisser d'autres sauts éventuels interférer dans les calculs. En effet, lorsque les facteurs de correction sont calculés pendant la segmentation de la série, ils ne sont pas nécessairement déterminés seulement avec des segments homogènes.

Méthodes conçues à l'origine pour détecter plusieurs inhomogénéités simultanément

Les méthodes qui peuvent détecter des sauts multiples en une application (méthode Bayésienne multivariée multi-sauts, filtre de Kolmogorov-Zurbenko, critère de Schwarz, segmentation des séries, méthode de Szentimrey, méthode de Caussinus & Mestre) ont l'avantage de fournir des estimations non biaisées par d'autres discontinuités. En effet, elles permettent d'éviter l'interférence d'éventuels sauts non encore détectés.

Méthodes conçues pour détecter plusieurs types d'inhomogénéités

Certaines méthodes ont la capacité de détecter à la fois des sauts et des tendances (méthode d'Alexandersson, régression multiple, régression à deux phases). Cependant, le test d'Alexandersson pour les sauts semble détecter un saut près du milieu d'un segment en réalité tendancieux. De plus, le test d'Alexandersson destiné aux tendances conduit à une classification de plusieurs sauts consécutifs comme une tendance. La régression à deux phases se montre plus adéquate pour détecter des tendances que des sauts (Easterling & Peterson, 1992). Par ailleurs, les méthodes possédant la capacité de déceler une tendance sont portées à mal classifier les petits sauts (Ducré-Robitaille *et al.*, 2003). La détection de sauts de faibles amplitudes est problématique dans la plupart des cas. Par contre, en travaillant avec les déviations cumulées, un changement de moyenne est plus prononcé que dans une série d'observations brutes. Par conséquent, les tests sur les déviations cumulées permettent de détecter plus facilement des inhomogénéités de faible amplitude.

ANALYSE D'APPLICABILITE AUX SERIES DE PRECIPITATIONS TOTALES ANNUELLES

Cette section vise à identifier les méthodes qui peuvent homogénéiser des séries de précipitations totales annuelles. Premièrement, certaines d'entre elles comme l'analyse

des doubles accumulations, la méthode du vecteur régional, l'analyse graphique des rapports, l'analyse graphique des déviations cumulées, la méthode d'Alexandersson, l'approche bivariée et les tests sur les déviations cumulées ont été développées spécialement pour l'homogénéisation des précipitations totales annuelles. De plus, les méthodes de Szentimrey, de Jaruskova et de Caussinus & Mestre permettent d'homogénéiser plusieurs types de variables incluant les précipitations. Plusieurs applications ont été faites sur des données de précipitations. Par exemple, la méthode d'Alexandersson a servi à homogénéiser des séries de précipitations en Norvège (Hanssen-Bauer & Forland, 1994), en France (Lamarque & Jourdain, 1994), en Suède et en Finlande (Peterson *et al.*, 1998). La procédure de Boroneant & Tomozeiu (1999) a permis d'homogénéiser des séries de précipitations en Italie (Tomozeiu *et al.*, 2000). Wijngaard *et al.* (2003) ont appliqué leur méthode sur des séries de précipitations localisées en Europe. Puisque les précipitations sont très variables, Wijngaard *et al.* (2003) ont extrait le nombre de jours humides par année (précipitations de plus de 1 mm) et appliqué leur méthode sur cette variable. Cette variable a pour effet de réduire la variabilité et ainsi, seuls les sauts assez grands peuvent être détectés. De plus, cette approche a été choisie pour tester l'homogénéité de séries de pluie en Éthiopie (Conway *et al.*, 2004). Chez Météo-France, les séries de précipitations sont homogénéisées selon la méthode de Caussinus & Mestre. En Autriche, l'homogénéité du réseau de mesure de précipitations a été vérifiée au moyen du test de Craddock sur les déviations cumulées combiné à une analyse des métadonnées (Peterson *et al.*, 1998). En Suisse, les séries de précipitations ont été homogénéisées en combinant 12 méthodes et la consultation des métadonnées en utilisant le logiciel THOMAS (Begert *et al.*, 2005). Au Canada, les précipitations ont été réhabilitées pour des changements d'instruments systématiques à tout le réseau et pour des modifications dans la façon de prendre les observations (Mekis & Hogg, 1999). Aux États-Unis, les métadonnées ainsi que la méthode de Karl & Williams sont combinées. La méthode de Rhoades & Salinger (1993) est appliquée en Nouvelle-Zélande et dans les îles du Pacifique. Enfin, la régression à deux-phases selon le modèle de Wang (2003) a été utilisée par un regroupement de scientifiques pour homogénéiser des extrêmes de température et de précipitations en Amérique Centrale et au nord de l'Amérique du Sud (Aguilar *et al.*, 2005).

Certaines techniques développées afin d'homogénéiser des séries de températures moyennes annuelles pourraient performer également pour des séries de précipitations totales annuelles. Cela va dépendre de la distribution des données (normale ou non-normale). Par exemple, l'applicabilité de la régression multiple, de la méthode Bayésienne multivariée multi-sauts, du filtre Kolmogorov-Zurbenko, du critère de Schwarz, du test de Student et de Wilcoxon séquentiel n'a pas encore été testée. Les méthodes citées dans cet article ont été développées dans le but précis de tester ou de corriger l'homogénéité de séries climatiques. Par ailleurs, d'autres techniques statistiques qui servent à analyser les ruptures dans les séries chronologiques pourraient également accomplir cette tâche (Chen & Gupta, 2000, 2001).

CONCLUSIONS

La correction des inhomogénéités dans les séries climatiques est cruciale pour toutes les études qui portent sur des variables climatiques. Plus particulièrement, des séries de

précipitations de qualité permettront de prendre des décisions basées sur des données fiables dans différentes recherches sur les changements climatiques ou encore dans d'autres domaines comme l'irrigation ou la production d'énergie.

Un grand nombre de méthodes d'homogénéisation ont été présentées dans cette revue de littérature. Les méthodes objectives ont été approfondies parce qu'elles sont plus complètes. Néanmoins, il ne faut pas négliger la possibilité d'utiliser une approche subjective pour bien s'approprier les données et éviter de commettre des erreurs. La tendance dans les dernières méthodes développées consiste à combiner plusieurs tests très différents et de déclarer une série inhomogène seulement lorsque plusieurs tests sont significatifs. Cela évite de réduire le nombre de fausses alarmes. De plus, on insiste sur le fait que la consultation des métadonnées est essentielle malgré l'utilisation d'une méthode sophistiquée. En effet, il est important de connaître la cause des ruptures artificielles qui sont corrigées.

Le choix de la méthode optimale à utiliser ou même la combinaison des méthodes à utiliser dépend fortement des caractéristiques statistiques de la station. Pour aider les utilisateurs dans le choix d'une méthode, des études comparatives de la robustesse et de la puissance des méthodes sur des données de précipitations totales synthétiques seraient grandement appréciées. Par exemple, on pourrait évaluer la robustesse des méthodes lorsque l'hypothèse de normalité des données (ou des séries de différences/rapports, selon la méthode) n'est pas vérifiée. Cependant, il est fort possible que ce point ne soit pas crucial. Les données synthétiques pourraient également contenir des changements de variance et/ou des inhomogénéités dans les séries voisines. On pourrait ainsi avoir une idée de l'impact du non-respect des postulats propres à chaque méthode. L'aptitude de différentes méthodes d'homogénéisation des précipitations totales annuelles à identifier des séries homogènes et inhomogènes n'a pas encore été analysée.

Remerciements Les auteurs désirent remercier le Consortium OURANOS, le Conseil de Recherche en Sciences Naturelles et Génie du Canada (CRSNG) et le Centre d'Études Nordiques (CEN) pour avoir financé cette recherche. Les auteurs aimeraient remercier MM. Abderrahmane Yagouti et Gilles Boulet du Ministère du Développement durable, de l'Environnement et des Parcs du Québec pour leur collaboration. Les auteurs désirent également remercier Olivier Mestre et un autre réviseur anonyme pour leurs commentaires et suggestions qui ont contribué à améliorer la qualité de l'article.

REFERENCES

- Aguilar, E., Auer, I., Brunet, M., Peterson, T. C. & Wieringa, J. (2003) Guidelines on climate metadata and homogenization. Report WMO-TD 1186, World Meteorological Organization, Geneva, Switzerland.
- Aguilar, E., Peterson, T. C., Obando, P. R., Frutos, R., Retana, J. A., Solera, M., Soley, J., Garcia, I. G., Araujo, R. M., Santos, A. R., Valle, V. E., Brunet, M., Aguilar, L., Alvarez, L., Bautista, M., Castanon, C., Herrera, L., Ruano, E., Sinay, J. J., Sanchez, E., Oviedo, G. I. H., Obed, F., Salgado, J. E., Vazquez, J. L., Baca, M., Gutierrez, M., Centella, C., Espinosa, J., Martinez, D., Olmedo, B., Espinoza, C. E. O., Nunez, R., Haylock, M., Benavides, H. & Mayorga, R. (2005) Changes in precipitation and temperature extremes in Central America and northern South America, 1961–2003. *J. Geophys. Res. Atmospheres* **110**(D23), 23107, doi:10.1029/2005JD006119.
- Akaïke, H. (1973) Information theory and an extension of the maximum likelihood principle. *Proc. Second Int. Symp. on Information Theory* (ed. by B. N. Petrov & F. Csaki) (Budapest, Hungary), 267–281.
- Alexandersson, H. (1986) A homogeneity test applied to precipitation data. *J. Climatol.* **6**, 661–675.
- Alexandersson, H. & Moberg, A. (1997) Homogenization of Swedish temperature data. Part 1: Homogeneity test for linear trends. *Int. J. Climatol.* **17**, 25–34.

- Asselin, J., Ouarda, T. B. M. J., Fortin, V. & Bobée, B. (1999) Une procédure Bayésienne bivariée pour détecter un décalage de la moyenne. Rapport de recherche R-528, INRS-Eau, Chaire en hydrologie statistique, Québec, Canada.
- Begert, M., Schlegel, T. & Kirchhofer, W. (2005) Homogeneous temperature and precipitation series of Switzerland from 1864 to 2000. *Int. J. Climatol.* **25**, 65–80.
- Bois, P. (1986) Contrôle de séries chronologiques corrélées par étude du cumul des résidus de la corrélation. *Deuxièmes Journées Hydrologiques de l'ORSTOM*, Montpellier, France.
- Boroneant, C. & Tomozeiu, R. (1999) Experience with homogeneity testing of temperature data at Bucuresti Filaret station. *Proc. Second Seminar for Homogenization of Surface Climatological Data* (Budapest, Hungary). WMO-TD no. 962, 113–123. World Meteorological Organization, Geneva, Switzerland.
- Brunet-Moret, Y. (1979) Homogénéisation des précipitations. *Cah. de L'ORSTOM Série Hydrologie* **16**, 147–170.
- Buishand, T. A. (1982) Some methods for testing the homogeneity of rainfall records. *J. Hydrol.* **58**, 11–27.
- Buishand, T. A. (1984) Tests for detecting a shift in the mean of hydrological time series. *J. Hydrol.* **73**, 51–69.
- Caussinus, H. & Lyazrhi, F. (1997) Choosing a linear model with a random number of change-points and outliers. *Ann. Inst. Statist. Math.* **49**, 761–775.
- Caussinus, H. & Mestre, O. (2004) Detection and correction of artificial shifts in climate series. *J. Roy. Statist. Soc. Series C: Applied Statistics* **53**, 405–425.
- Chen, J. & Gupta, A. K. (2000) *Parametric Statistical Change-point Analysis*. Birkhauser, Boston, USA.
- Chen, J. & Gupta, A. K. (2001) On change-point detection and estimation. *Communications in Statistics-Simulation and Computation* **30**, 665–697.
- Conway, D., Mould, C. & Bewket, W. (2004) Over one century of rainfall and temperature observations in Addis Ababa, Ethiopia. *Int. J. Climatol.* **24**, 77–91.
- Craddock, J. M. (1979) Methods of comparing annual rainfall records for climatic purposes. *Weather* **34**, 332–346.
- Datsenko, N. M., Moberg, A. & Sonechkin, D. M. (2002) Objective time-scale-dependent homogenization of early instrumental temperature series. *Theor. Appl. Climatol.* **72**, 103–126.
- Ducré-Robitaille, J. F., Boulet, G. & Vincent, L. A. (2003) Comparison of techniques for detection of discontinuities in temperature series. *Int. J. Climatol.* **23**, 1087–2003.
- Easterling, D. R. & Peterson, T. C. (1992) Techniques for detecting and adjusting for artificial discontinuities in climatological time series: a review. *Proc. Fifth Int. Meeting on Statistical Climatology* (Toronto, Ontario, Canada), J28–J32. Steering Committee for International Meetings on Statistical Climatology.
- Easterling, D. R. & Peterson, T. C. (1995) A new method for detecting undocumented discontinuities in climatological time series. *Int. J. Climatol.* **15**, 369–377.
- Fearnhead, P. (2006) Exact and efficient Bayesian inference for multiple changepoint. *Statistics and Computing*, **16**(2), 203–213.
- Gullett, D. W., Vincent, L. & Malone, L. H. (1991) Homogeneity testing of monthly temperature series. Application of multiphase regression models with mathematical changepoints. Report 90-4, Canadian Climate Center, Downsview, Ontario, Canada.
- Gullett, D. W., Vincent, L. A. & Sajecki, P. J. F. (1990) Testing for homogeneity in temperature time series at Canadian climate stations. Report 91-10, Canadian Climate Center, Downsview, Ontario, Canada.
- Hanssen-Bauer, I. & Forland, E. J. (1994) Homogenizing long Norwegian precipitation series. *J. Climate*, **7**, 1001–1013.
- Hawkins, D. M. (2001) Fitting multiple change-point models to data. *Computational Statistics & Data Analysis* **37**, 323–341.
- Hubert, P., Carbonnel, J. P. & Chaouche, A. (1989) Segmentation des séries hydrométéorologiques—application à des séries de précipitations et de débits de l'Afrique de l'Ouest. *J. Hydrol.* **110**, 349–367.
- Jaruskova, D. (1996) Change-point detection in meteorological measurement. *Mon Weath. Rev.* **124**, 1535–1543.
- Jones, P. D., Raper, R. S., Diaz, H. F., Kelly, P. M. & Wigley, T. M. L. (1986) Northern hemisphere surface air temperature variations: 1851–1984. *J. Clim. Appl. Met.* **25**, 161–179.
- Karl, T. R. & Williams, C. N., Jr (1987) An approach to adjusting climatological time series for discontinuous inhomogeneities. *J. Clim. Appl. Met.* **26**, 1744–1763.
- Kehagias, A., Nidelkou, E. & Petridis, V. (2006) A dynamic programming segmentation procedure for hydrological and environmental time series. *Stochast. Environ. Res. Risk Assessment* **20**, 77–94.
- Khalik, M. N. & Ouarda, T. B. M. J. (2006) A note on the critical values of the Standard Normal Homogeneity Test (SNHT). *Int. J. Climatol.* (in press), doi:10.1002/joc.1438.
- Kohler, M. A. (1949) On the use of double-mass analysis for testing the consistency of meteorological records and for making required adjustments. *Bull. Am. Met. Soc.* **30**, 188–189.
- Lamarque, P. & Jourdain, S. (1994) Elaboration de longues séries climatologiques homogènes pour l'étude de l'évolution climatique. *La Météorologie* **8**, 61–68.
- Lanzante, J. R. (1996) Resistant, robust and non-parametric techniques for the analysis of climate data: theory and examples, including applications to historical radiosonde station data. *Int. J. Climatol.* **16**, 1197–1226.
- Lee, A. F. S. & Heghinian, S. M. (1977) A shift of the mean level in a sequence of independent normal random variables: a Bayesian approach. *Technometrics* **19**, 503–506.
- Lubes-Niel, H., Masson, J. M., Paturel, J. E. & Servat, E. (1998) Variabilité climatique et statistiques. Étude par simulation de la puissance et de la robustesse de quelques tests utilisés pour vérifier l'homogénéité de chroniques. *Rev. Sci. Eau* **11**, 383–408.
- Lund, R. & Reeves, J. (2002) Detection of undocumented changepoints: a revision of the two-phase regression model. *J. Climate* **15**, 2547–2554.
- Maronna, R. & Yohai, V. J. (1978) A bivariate test for the detection of a systematic change in mean. *J. Am. Statist. Assoc.* **73**, 640–645.

- Mekis, E. & Hogg, W. D. (1999) Rehabilitation and analysis of Canadian daily precipitation series. *Atmos. Ocean* **37**, 53–85.
- Mielke, P. W., Berry, K. J. & Brier, G. W. (1981) Application of multi-response permutation procedures for examining seasonal changes in monthly mean sea-level pressure patterns. *Mon. Weath. Rev.* **109**, 120–126.
- Ouarda, T. B. M. J., Rasmussen, P. F., Cantin, J.-F., Bobée, B., Laurence, R., Hoang, V. D. & Barabé, G. (1999) Identification d'un réseau hydrométrique pour le suivi des modifications climatiques dans la province de Québec. *Rev. Sci. Eau* **12**, 425–448.
- Perreault, L., Haché, M., Slivitzky, M. & Bobée, B. (1999) Detection of changes in precipitation and runoff over eastern Canada and US using a Bayesian approach. *Stochast. Environ. Res. Risk Assessment* **13**, 201–216.
- Perreault, L., Parent É., Bernier J., Bobée B. & Slivitzky M. (2000) Retrospective multivariate Bayesian change-point analysis: a simultaneous single change in the mean of several hydrological sequences. *Stochast. Environ. Res. Risk Assessment* **14**, 243–261.
- Peterson, T. C., Easterling, D. R., Karl, T. R., Groisman, P., Nicholls, N., Plummer, N., Torok, S., Auer, I., Boehm, R., Gullett, D., Vincent, L., Heino, R., Tuomenvirta, H., Mestre, O., Szentimrey, T., Salinger, J., Forland, E. J., Hanssen-Bauer, I., Alexandersson, H., Jones, P. & Parker, D. (1998) Homogeneity adjustments of *in situ* atmospheric climate data: a review. *Int. J. Climatol.* **18**, 1493–1517.
- Pettitt, A. N. (1979) A non-parametric approach to the change-point problem. *Appl. Statist.* **28**, 126–135.
- Potter, K. W. (1981) Illustration of a new test for detecting a shift in mean in precipitation series. *Mon. Weath. Rev.* **109**, 2040–2045.
- Rhoades, D. A. & Salinger, M. J. (1993) Adjustment of temperature and rainfall records for site change. *Int. J. Climatol.* **13**, 899–913.
- Schwarz, G. (1978) Estimating the dimension of a model. *Ann. Statist.* **6**, 461–464.
- Seidou, O. & Ouarda, T. B. M. J. (2007) Recursion-based multiple changepoint detection in multivariate linear regression and application to river streamflow. *Water Resour. Res.* (accepted).
- Seidou, O., Asselin, J. J. & Ouarda, T. B. M. J. (2007) Bayesian multivariate linear regression with application to changepoint models in hydrometeorological variables. *Water Resour. Res.* (accepted).
- Slonosky, V. C., Jones, P. D. & Davies, T. D. (1999) Homogenization techniques for European monthly mean surface pressure series. *J. Climate* **12**, 2658–2672.
- Solow, A. R. (1987) Testing for climate change: an application of the two-phase regression model. *J. Clim. Appl. Met.* **26**, 1401–1405.
- Szentimrey, T. (1997) Statistical procedure for joint homogenization of climatic time series, *Proc. First Seminar for Homogenization of Surface Climatological Data* (Budapest, Hungary), 47–62. Hungarian Meteorological Service.
- Szentimrey, T. (1999) Multiple analysis of series for homogenization (MASH). *Proc. Second Seminar for Homogenization of Surface Climatological Data* (Budapest, Hungary), 27–46. WMO-TD no. 962. World Meteorological Organization, Geneva, Switzerland.
- Tayanc, M., Dalfes, H. N., Karaca, M. & Yenigun, O. (1998) A comparative assessment of different methods for detecting inhomogeneities in Turkish temperature data sets. *Int. J. Climatol.* **18**, 561–578.
- Thompson, C. S. (1984) Homogeneity analysis of rainfall series: an application of the use of a realistic rainfall model. *J. Climatol.* **4**, 609–619.
- Tomozeiu, R., Busuioc, A., Marletto, V., Zinoni, F. & Cacciamani, C. (2000) Detection of changes in the summer precipitation time series of the region Emilia-Romagna, Italy. *Theor. Appl. Climatol.* **67**, 193–200.
- Vincent, L. A. (1998) A technique for the identification of inhomogeneities in Canadian temperature series. *J. Climate* **11**, 1094–1105.
- Wang, X. L. (2003) Comments on “Detection of undocumented changepoints: a revision of the two-phase regression model”. *J. Climate* **16**, 3383–3385.
- Wijngaard, J. B., Klein Tank, A. M. G. & Können, G. P. (2003) Homogeneity of 20th century European daily temperature and precipitation series. *Int. J. Climatol.* **23**, 679–692.
- Young, K. C. (1993) Detecting and removing inhomogeneities from long-term monthly sea level pressure time series. *J. Climate* **6**, 1205–1220.
- Zurbenko, I., Porter, P. S., Rao, S. T., Ku, J. Y., Gui, R. & Eskridge, R. E. (1996) Detecting discontinuities in time series of upper-air data: development and demonstration of an adaptive filter technique. *J. Climate* **9**, 3548–3560.

Reçu le 18 juillet 2005; accepté le 16 octobre 2006