REVUE BIBLIOGRAPHIQUE
DES TESTS DE STATIONNARITÉ

REVUE BIBLIOGRAPHIQUE DES TESTS DE STATIONNARITÉ

Rapport préparé pour

Environnement Canada
Division Hydrologie
373 Sussex Drive, Block E
Ottawa, Ontario K1A 0H3

par

Dany Faucher
Taha B.M.J. Ouarda
Bernard Bobée

Chaire en Hydrologie statistique
Institut national de la Recherche scientifique, INRS-Eau
2800, rue Einstein, C.P. 7500, Sainte-Foy (Québec) G1V 4C7

Rapport de recherche N° R-499

Juin 1997

ISBN 2-89146-393-5

ÉQUIPE DE RECHERCHE

Ont participé à la réalisation de cette étude:

Environnement Canada
Paul Pilon
Klaus Wiebe
Bob Hale

Chaire en Hydrologie statistique
Institut national de la Recherche scientifique, INRS-Eau

Dany Faucher
Taha B.M.J. Ouarda
Bernard Bobée

TABLE DES MATIÈRES

AVANT-PROPOS

Ce rapport a été réalisé dans le cadre d'un contrat accordé à la Chaire en Hydrologie statistique à l'INRS-Eau par la Division Hydrologie d'Environnement Canada. Ce volet présente une revue de littérature des tests de stationnarité. Les auteurs du rapport tiennent à exprimer leur reconnaissance à MM. Paul Pilon, Klaus Wiebe et Bob Hale pour leurs commentaires et leurs réflexions. Les auteurs tiennent spécialement à remercier M. Paul Pilon pour avoir suivi de très prés le déroulement des travaux. Les auteurs veulent aussi remercier leurs collègues Luc Perreault et Peter Rasmussen de la Chaire en Hydrologie statistique à l'INRS-Eau, ainsi que le Professeur Michel Slivitzky pour leurs précieuses suggestions.

1 TEST DE STATIONNARITÉ

1.1 Définition de la notion de stationnarité

Une série de données est dite stationnaire, lorsque l'on ne retrouve pas de variations temporelles significatives, autres que les fluctuations aléatoires dans les valeurs classées chronologiquement. On peut détecter une non-stationnarité en examinant la moyenne ou la variance de la série et en évaluant s'il y a eu un changement significatif à une date donnée. De façon générale, on peut identifier quatre types de non-stationnarité : le saut de la moyenne, la tendance de la moyenne, la tendance en escalier de la moyenne et les changements de variabilité.

1.2 Saut de la moyenne (*jump*)

Le changement de la moyenne peut se faire assez rapidement, il y a alors une tendance avec saut de la moyenne, c'est-à-dire qu'il survient un changement brusque de la moyenne de la série à un certain moment dans le temps (figure 1). Une série où il y a une tendance en saut est définie par:

$$x_i = \begin{cases} \mu + \varepsilon_i & pour\ i = 1, 2, \ldots, \tau; \\ \mu + \delta + \varepsilon_i & pour\ i = \tau + 1,\ \tau + 2, \ldots, n; \end{cases} \tag{1}$$

où μ représente la moyenne de la série avant le saut, $\tau + 1$ est le moment où a lieu le saut dans la série, δ est l'amplitude du saut ($\mu + \delta$ est la moyenne après le saut) et ε_i est une variable aléatoire indépendante et normalement distribuée de moyenne 0 et de variance σ^2.

La plupart des tests pour identifier ce type de non-stationnarité sont des tests de comparaison de moyennes. On doit alors diviser la série en deux sous-échantillons, correspondant aux données situées avant et après le moment τ où l'on croit qu'il y a un saut de la moyenne et calculer la moyenne de ces deux sous-périodes.

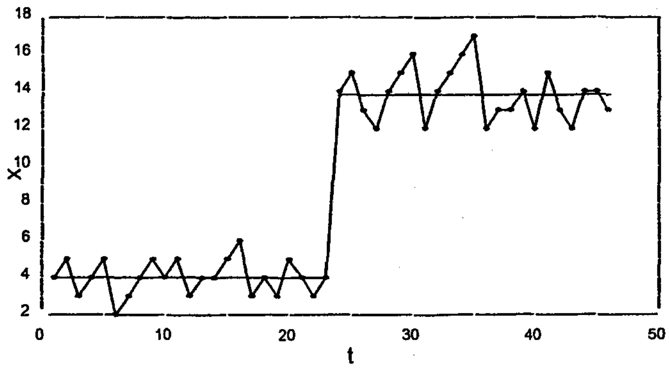

Figure 1. Non-stationnarité avec saut

1.3 Tendance monotone (*trend*)

Le changement de moyenne peut aussi se faire graduellement, c'est-à-dire qu'il y a présence d'une tendance monotone croissante (ou décroissante) dans la série (figure 2). La fonction décrivant une série ayant une tendance monotone peut être définie par:

$$x_i = T_i + \varepsilon_i$$

où x_i représente l'observation i de la série, T_i représente le paramètre de la tendance et ε_i est une variable aléatoire indépendante et normalement distribuée de moyenne 0 et de variance σ^2.

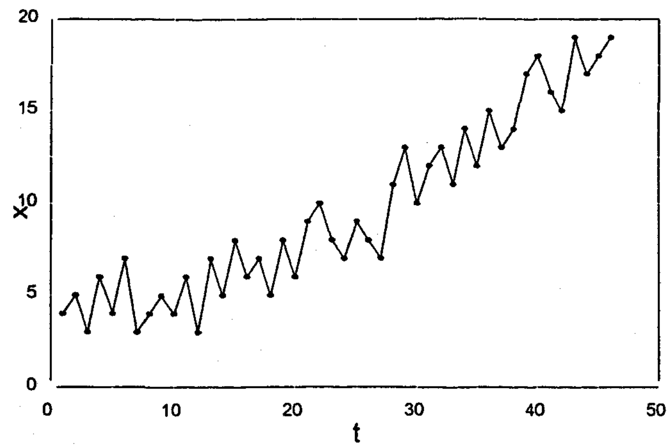

Figure 2. Non-stationnarité avec tendance

1.4 Tendance en escalier (*stepping levels* ou *shifting levels*)

Il existe aussi un type de non-stationnarité que l'on appelle tendance en escalier, qui se caractérise par une succession de changements dans la moyenne (figure 3). Une série où il y a une tendance escalier, avec k sauts dans la moyenne, a la structure suivante:

$$
\begin{aligned}
x_i &= \mu + \varepsilon_i & &\textit{pour } i = 1, 2, \ldots, \tau_1; \\
x_i &= \mu + \delta_1 + \varepsilon_{1i} & &\textit{pour } i = \tau_1 + 1,\ \tau_1 + 2, \ldots, n - \tau_1; \\
&\ldots & & \\
x_i &= \mu + \delta_k + \varepsilon_{ki} & &\textit{pour } i = \tau_k + 1,\ \tau_k + 2, \ldots, n;
\end{aligned}
$$

où μ représente la moyenne de la série avant le premier saut, $\tau_i + 1$ est le moment où a lieu le i^e saut dans la série, δ_i est l'amplitude du i^e saut (μ + δ_i est la moyenne après le saut) et ε_i est une variable aléatoire indépendante et normalement distribuée de moyenne 0 et de variance σ^2.

Les différents tests paramétriques et non paramétriques utilisés pour vérifier la stationnarité en saut, en tendance ou en escalier, sont présentés dans la section 2 du présent rapport.

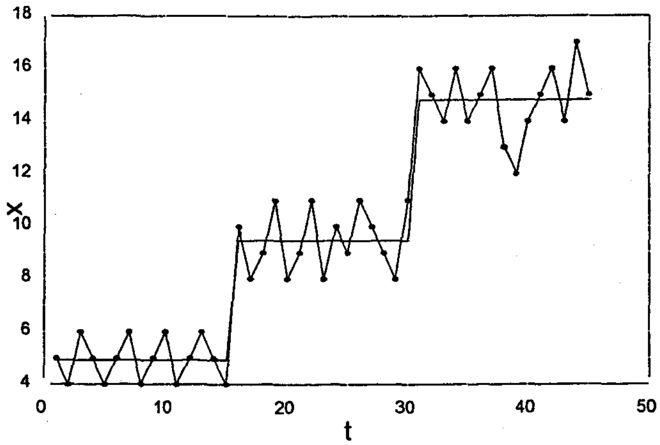

Figure 3. Non-stationnarité avec tendance en escalier

1.5 Changements dans la variabilité

Finalement, on peut trouver des changements dans la variabilité d'une série de données temporelles, c'est-à-dire, qu'il peut y avoir présence de sauts dans la variance de la série d'observation. Ce phénomène se manifeste par une augmentation ou une diminution des fluctuations des données dans le temps. La figure 4 illustre le cas d'une série où il y a eu un saut positif de la variance aux environs du temps $t = 25$. Dans ce cas, la variance après le saut semble plus élevée qu'avant le saut. Les tests relatifs aux changements dans la variabilité sont présentés à la section 3 du présent rapport.

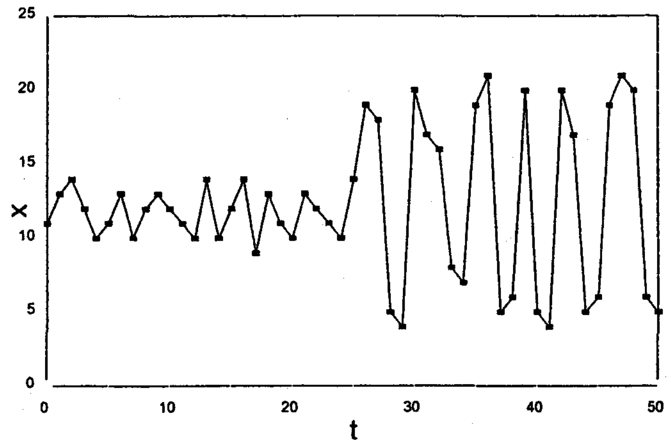

Figure 4. Non-stationnarité de la variance

2 TESTS DE STATIONNARITÉ DE LA MOYENNE

Afin de tester la stationnarité de la moyenne d'une série de temps, on a le choix d'utiliser deux types de tests: les tests paramétriques et les tests non paramétriques. Les tests paramétriques sont effectués en considérant que la statistique du test est distribuée selon une fonction de distribution bien définie, comme par exemple, la loi de Student. Dans le cas des tests paramétriques faisant intervenir la distribution normale, la loi de Student et la loi de Fisher, on doit préalablement s'assurer que les données standardisées sont indépendamment et identiquement distribuées selon la loi normale centrée réduite ($N(0,1)$). Mais lorsque l'on rejette l'hypothèse de normalité de la série ou que l'on ne dispose pas de suffisamment de données pour vérifier cette hypothèse, ces tests ne peuvent être utilisés. On peut alors avoir recours à des tests non paramétriques, qui permettent de s'affranchir de l'hypothèse de normalité, en fondant l'analyse exclusivement sur les rangs des observations. L'avantage de ces tests est que l'on ne doit poser aucune hypothèse sur la distribution des données et que l'on peut très souvent effectuer quand même le test même avec un nombre limité de données.

Cette section sera donc divisée d'abord en deux parties : les tests paramétriques de stationnarité de la moyenne seront abordés dans la section 2.1 et les tests non paramétriques de stationnarité de la moyenne seront présentés à la section 2.2. Dans chacune des sections, on retrouve la description de différents tests qui peuvent être utilisés dans les 3 cas de non-stationnarité de la moyenne discutés à la section 1. Pour chaque test, on commencera d'abord par mentionner pour quel type de non-stationnarité il est applicable. Ensuite, on présentera la statistique du test et la distribution de celle-ci en vue de déterminer la région critique du test. Finalement, pour la plupart des tests, les différentes étapes à suivre seront indiquées de façon à pouvoir utiliser ces tests facilement.

La section 2.3 présente un tableau qui résume les conditions d'application des différents tests de stationnarité de la moyenne présentés aux sections 2.1 et 2.2. On peut s'y référer pour savoir rapidement à quel type de non-stationnarité correspond chaque test.

Finalement, à la section 2.4, une comparaison des différents tests de stationnarité de la moyenne sur des bases théoriques, sera présentée.

2.1 Tests paramétriques de stationnarité de la moyenne

- **Test de Student**

Le test de Student relatif à la comparaison de la moyenne de deux populations peut être utilisé pour évaluer l'existence de sauts dans la moyenne d'une population (Lettenmaier, 1976). Ce test consiste à vérifier si la moyenne des données situées avant un saut est significativement différente de la moyenne des données situées après ce saut. Notons μ_1 la moyenne des n_1 données situées avant le saut et μ_2, la moyenne des n_2 observations situées après le saut. Le test consiste alors à examiner l'hypothèse suivante:

$$H_o: \ \mu_1 = \mu_2$$

avec l'hypothèse alternative :

$$H_1: \ \mu_1 \neq \mu_2$$

On peut facilement montrer que la statistique du test, t_0, est distribuée selon une loi de Student à $n_1 + n_2 - 2$ degrés de liberté (Montgomery, 1991):

$$t_0 = \frac{\bar{x}_1 - \bar{x}_2}{S_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} \approx t_{\alpha/2, n_1+n_2-2}$$

où

$$S_p = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

où s_1^2 et s_2^2 représentent les variances non biaisées des données situées avant et après le saut et S_p représente la variance combinée des deux sous-échantillons.

On peut résumer les étapes de ce test de la manière suivante:

Étape 1

Identifier la date à partir de laquelle on veut examiner l'hypothèse d'un saut dans la moyenne et calculer les moyennes et écarts-types des données situées avant et après celui-ci.

Étape 2

Tester l'hypothèse nulle d'égalité des moyennes en calculant la statistique t_0.

Étape 3

Au niveau de signification de α, rejeter H_0 si $t_0 \notin \left[-t_{\alpha/2,n_1+n_2-2}, t_{\alpha/2,n_1+n_2-2}\right]$.

- **Test des déviations cumulées**

Le test des déviations cumulées pour la vérification de l'hypothèse d'homogénéité peut être utilisé pour détecter la présence d'un saut dans la moyenne d'une série de données (Buishand, 1982; WMO, 1988). Ce test ainsi que celui du rapport de vraisemblance de Worsley (1979) présenté dans ce qui suit, s'appuient sur l'hypothèse que les données sont indépendantes et normalement distribuées.

Soit la variable S_k^* qui correspond à la somme des écarts par rapport à la moyenne $\bar{x}$ définie par :

$$S_k^* = \sum_{i=1}^{k}(x_i - \bar{x}) \qquad \text{pour } k = 1,2,\ldots,n\text{-}1 \,;$$

où

$$S_0^* = S_n^* = 0$$

Pour une série où il n'y a pas de tendance monotone ou de saut dans la moyenne, on devrait s'attendre à ce que les valeurs de S_k^* fluctuent autour de 0. Par contre en présence d'une tendance décroissante ou d'une diminution brusque de la moyenne, la plupart des valeurs de S_k^* devraient être positives. De façon analogue, pour une croissance de la moyenne, il devrait y avoir davantage de S_k^* négatifs que positifs.

On obtient les sommes partielles ajustées en divisant chacun des S_k^* par l'écart-type de l'ensemble des données :

$$S_k^{**} = \frac{S_k^*}{S_x} \qquad \text{pour } k = 0,1,\ldots,n \ ;$$

où

$$S_x^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2$$

On utilise comme statistique de ce test, la valeur maximale des S_k^{**} en valeur absolue:

$$Q = \max_{0 \le k \le n}\left|S_k^{**}\right|$$

On peut trouver les valeurs critiques de ce test dans Buishand (1982). Lorsque n tend vers l'infini, les valeurs critiques de Q peuvent être obtenues à partir de la table de la statistique de Kolmogorov-Smirnov (Buishand, 1982).

- **Test du rapport de vraisemblance de Worsley (1979)**

Le test de Worsley peut être utilisé pour détecter la présence d'un saut dans la moyenne d'une série de données (Buishand, 1982; WMO, 1988). On considère $\bar{x}$ et S_x comme étant la moyenne et l'écart-type de l'ensemble de la série.

Soit la variable S_k^*, la somme des écarts à la moyenne, telle que:

$$S_k^* = \sum_{i=1}^{k} (x_i - \bar{x}) \qquad \text{pour } k = 1, 2, \ldots, n;$$

où

$$S_0^* = 0 \qquad \text{et} \qquad S_n^* = 0$$

et les variables Z_k^* et V telles que:

$$Z_k^* = \frac{\sqrt{k(n-k)}S_k^*}{S_x}$$

$$V = \max_{1 \le k \le n-1} |Z_k^*|$$

où Z_k^* représente une statistique de Student pour tester la différence entre la moyenne des k premières observations et la moyenne des $(n-k)$ dernières observations de la série de données.

La statistique du test, W, est donnée par Worsley (1979):

$$W = \frac{V\sqrt{n-2}}{\sqrt{1-V^2}}$$

Les valeurs critiques de ce test sont également présentées dans Worsley (1979) ou dans Buishand (1982). Comme dans le cas du test des déviations cumulées, la position du maximum de Z_k^*, peut être considérée comme étant une estimation de la date de changement de la moyenne de la série.

- **Test de Hinkley (1971)**

On définit les cumuls suivants (Bernier, 1977):

$$S_0 = 0; \; S_n = \sum_{j=1}^{n}(x_j - \mu + \delta\sigma)$$

où μ et σ représentent la moyenne et l'écart-type de l'échantillon $x_1, x_2, \ldots, x_N$. Le moment τ où un changement apparaît dans la moyenne, peut être estimé par le premier indice n où S_n atteint sa valeur maximale :

$$\hat{\tau} = \textit{minimum} \; (\,n \; \textit{tel que} \; S \;\geq S_u)$$

$$\textit{pour}\; u = 1, 2, \ldots, N$$

Le test consiste à rejeter l'hypothèse d'absence de changement si :

$$S_N - \max_{n \leq N} S_n < -h$$

Hinkley (1971) présente une description des méthodes de choix des paramètres h et δ.

- **Test de Cramer (1946)**

Le test de Cramer peut être utilisé pour comparer la moyenne de sous-échantillons d'une série de données, avec la moyenne générale de la série entière (Cramer, 1946; WMO, 1966). Ce test consiste à déterminer si les différences de moyennes sont assez grandes pour rejeter l'hypothèse du caractère aléatoire de la série d'observations, c'est-à-dire pour rejeter l'hypothèse d'absence de sauts ou de tendance de la moyenne.

Soit $\bar{x}$ et s, la moyenne et l'écart-type de la série de taille N et soit $\bar{x}_i$, la moyenne du i^e sous-échantillon de taille n_i de la série, telle que:

$$\bar{x}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} x_{ij}$$

Soit la variable t_i telle que:

$$t_i = r_i\sqrt{\frac{n_i(N-2)}{N-n_i\left(1+r_i^2\right)}}$$

où

$$r_i = \frac{\left(\bar{x}_i - \bar{x}\right)}{s}$$

On peut montrer que la variable t_i suit approximativement une distribution de Student avec N-2 degrés de liberté (Cramer, 1946).

On peut résumer les étapes de ce test de la manière suivante:

Étape 1

Identifier le moment des sauts potentiels dans la moyenne de la série et diviser cette dernière en k sous-échantillons.

Étape 2

Calculer la moyenne $\bar{x}_i$ de la sous-période i, le coefficient r_i et en déduire la valeur de la statistique t_i.

Étape 3

Au niveau de signification α, rejeter H_0 (égalité de la moyenne du sous-échantillon i et de la moyenne générale) si: $t_i \notin \left[-t_{\alpha/2,N-2};t_{\alpha/2,N-2}\right]$.

Étape 4

Effectuer le test de la même façon pour les k-1 autres sous-périodes identifiées.

- **Régression**

Une méthode proposée par Woodward et Gray (1993), utilise un modèle de régression pour détecter la présence d'une tendance dans une série de N observations. L'équation suivante représente le modèle où:

$$Y_t = \beta_0 + \beta_1 t + e_t$$

où:

t, la variable indépendante du modèle, représente le temps ;

Y_t, la variable dépendante du modèle, représente la caractéristique d'intérêt (température, débit,...)

e_t représente les résidus ;

β_0, représente l'ordonnée à l'origine de la droite de régression; et

β_1, représente la pente de la droite de régression.

L'approche consiste donc à tester l'hypothèse suivante:

H_0: $\beta_1 = 0$ *(la pente de la droite de régression est nulle)*
(absence de tendance dans la série)

avec l'hypothèse alternative:

H_1: $\beta_1 \neq 0$ *(présence d'une tendance croissante ou décroissante)*

Si l'on a une idée *a priori* du sens de la tendance, c'est-à-dire si l'on sait qu'elle est soit croissante ou décroissante, on peut effectuer un test unilatéral (H_1: $\beta_1 > 0$ ou H_1: $\beta_1 < 0$).

Si l'on suppose que les observations sont recueillies à intervalles de temps constants, de façon à ce que $t = 1, 2, 3, \ldots, n$, on a l'estimation suivante de la pente de la droite de régression par la méthode des moindres carrés:

$$\hat{\beta}_1 = \frac{\sum_{t=1}^{n}(t-\bar{t})Y_t}{\sum_{t=1}^{n}(t-\bar{t})^2}$$

Sous l'hypothèse que les résidus sont indépendants et normalement distribués $N(0,\sigma^2)$, on peut estimer la variance de $\hat{\beta}_1$ de la façon suivante:

$$\hat{\sigma}_{\beta_1} = \sqrt{\frac{\sum_{t=1}^{n}\left(Y_t - \hat{\beta}_0 - \hat{\beta}_1 t\right)^2}{(n-2)\sum_{t=1}^{n}(t-\bar{t})^2}}$$

En utilisant l'identité

$$\sum_{t=1}^{n}(t-\bar{t})^2 = \frac{n\left(n^2-1\right)}{12}$$

on obtient:

$$\hat{\sigma}_{\beta_1} = \sqrt{\frac{12\sum_{t=1}^{n}\left(Y_t - \hat{\beta}_0 - \hat{\beta}_1 t\right)^2}{n(n-2)\left(n^2-1\right)}}$$

où $\hat{\beta}_0 = \overline{Y} - \hat{\beta}_1 \bar{t}$.

On peut montrer que, sous l'hypothèse H_0 (pente nulle), la statistique $t_0 = \hat{\beta}_1 / \hat{\sigma}_{\beta_1}$ suit approximativement une distribution de Student avec N-2 degrés de liberté.

On peut résumer les étapes de ce test de la manière suivante:

Étape 1

Vérifier les hypothèses sur les erreurs (indépendance et normalité).

Étape 2

Calculer l'estimation de la pente $\hat{\beta}_1$ et estimer sa variance $\hat{\sigma}_{\beta_1}$.

Étape 3

Au niveau de signification α, rejeter H_0 (absence de tendance) si:

$t_0 \notin \left[-t_{\alpha/2,N-2} ; t_{\alpha/2,N-2}\right]$

- **Méthode de Kite (1989)**

Kite (1989) propose d'utiliser un modèle additif linéaire pour représenter une série de données x_t. Le modèle est le suivant:

$$x_t = P_t + T_t + R_t$$

où P_t est une composante de périodicité, T_t est une composante de tendance et R_t est une composante stochastique. Kite propose d'effectuer une analyse spectrale afin de détecter la présence éventuelle d'une de ces composantes dans les données et de pouvoir l'éliminer.

On peut détecter et corriger la composante de périodicité en utilisant le périodogramme de Shustler (Matalas, 1967), tandis que la composante stochastique peut être représentée par un modèle autorégressif. Pour ce qui est de la composante de tendance, elle peut être analysée et supprimée en utilisant une régression polynomiale telle que:

$$T_t = a_0 + a_1 t + a_2 t^2 + \ldots + a_p t^p$$

Pour plus de détails sur la méthode elle-même ainsi que sur l'analyse spectrale, on peut se référer à Kite (1989).

- **Approche Bayésienne**

Lee et Heghinian (1977) ont proposé une méthode basée sur l'approche bayésienne pour calculer les distributions marginales et jointes *a posteriori* des paramètres τ et δ, correspondant respectivement au moment et à l'amplitude du saut (cf. eq.1). Cette procédure permet ainsi d'associer une probabilité à chaque valeur des paramètres τ et δ.

On suppose *a priori* l'existence d'un saut dans la série de données. Sans information additionnelle, la probabilité *a priori* que le saut survienne entre τ et $\tau + 1$ est:

$$f(\tau) = \frac{1}{n-1} \qquad \tau = 1, 2, \ldots, n\text{-}1$$

De plus, on fait l'hypothèse des distributions *a priori* de δ, μ et σ suivantes:

$f(\delta)$ *est normale de moyenne 0 et de variance* σ_δ^2 ;

$f(\mu)$ *est normale de moyenne 0 et de variance* σ_μ^2 ;

$f(\sigma)$ *est gamma inverse et sa densité est proportionnelle à* $1/\sigma$.

Ainsi, la distribution *a posteriori* de τ donne, pour chaque date, la probabilité qu'il y ait un saut dans la moyenne:

$$f(\tau|x_1,\ldots,x_n) \propto \left[\frac{n}{\tau(n-\tau)}\right]^{1/2} [R(\tau)]^{-(n-2)/2} \qquad 0 \le \tau \le n-1$$

où

$$R(\tau)=\left[\sum_{i=1}^{\tau}\left(x_i-\bar{x}_\tau\right)^2+\sum_{i=\tau+1}^{n}\left(x_i-\bar{x}_{n-r}\right)^2\right]\left[\sum_{i=1}^{n}\left(x_i-\bar{x}_n\right)^2\right]^{-1}$$

où $\bar{x}_\tau$, représente la moyenne de l'échantillon des τ premières observations; $\bar{x}_{n-\tau}$, la moyenne des $n-\tau$ dernières observations et $\bar{x}_n$, la moyenne des n observations de la série.

Pour obtenir la distribution marginale *a posteriori* de δ, on doit intégrer la distribution jointe suivante, sur tout le domaine de τ:

$$f\left(\delta\middle|x_1,\ldots,x_n\right)=\sum_{\tau=1}^{n-1}f\left(\delta\middle|\tau,x_1,\ldots,x_n\right)f\left(\tau\middle|x_1,\ldots,x_n\right)$$

où

$$f\left(\delta\middle|\tau,x_1,\ldots,x_n\right)=\frac{(n-4)^{-1/2}}{\sigma_t(\delta)B\left(\frac{1}{2},\frac{n-2}{2}\right)}\left[1+\frac{\left(\delta-\mu_\tau(\delta)\right)^2}{(n-4)\sigma_\tau^2(\delta)}\right]^{-\frac{n-1}{2}}$$

Par ailleurs, un intervalle symétrique de crédibilité au niveau 100(1-α)% pour δ , $C_{1-\alpha}[a,b]$, a été proposé par Berger (1985) tel que (où α représente le niveau de confiance):

$$P\{a\le\delta\le b\}=\int_a^b f\left(\delta\middle|x_1,\ldots,x_n\right)d\delta=1-\alpha$$

où

$$P\{\delta\le a\}=\alpha/2 \quad \text{et} \quad P\{\delta\le b\}=1-\alpha/2$$

Une valeur particulière d'amplitude du saut, δ_0, peut être jugée comme non plausible si elle se trouve à l'extérieur de l'intervalle de crédibilité, $\delta_0\notin C$. Par contre, cela ne nous permet pas pour autant de rejeter l'hypothèse pour laquelle $\delta=\delta_0$ (Perreault et *al.*, 1996). Pour ce faire, il faudrait plutôt utiliser un test approprié.

Bernier (1994) propose un test basé sur le rapport des probabilités de changement ou de non-changement dans le moyenne. Ces probabilités sont définies de la façon suivante:

$$P\left\{H_{nc}\middle|x_1,\ldots,x_n\right\}=\sum_{\tau\in H_{nc}} f\left(\tau\middle|x_1,\ldots,x_n\right)$$

$$P\left\{H_{c}\middle|x_1,\ldots,x_n\right\}=\sum_{\tau\in H_{c}} f\left(\tau\middle|x_1,\ldots,x_n\right)$$

où H_{nc} et H_c, représentent les sous-ensembles des valeurs de τ caractérisant les hypothèses de non-changement et de changement, respectivement. En appliquant le test, on rejette l'hypothèse de non-changement dans la série si:

$$\frac{P\left\{H_{c}\middle|x_1,\ldots,x_n\right\}}{P\left\{H_{nc}\middle|x_1,\ldots,x_n\right\}}\geq\frac{\alpha_{nc}}{\alpha_c}$$

où α_{nc} et α_c représentent respectivement, les coûts reliés au rejet de l'hypothèse de non-changement et de l'hypothèse de changement, alors qu'elle est vraie, c'est-à-dire l'erreur de type I pour chacune des hypothèses.

Plusieurs applications de cette approche bayésienne ont été publiées dont celles de Bruneau et Rassam (1983) et de Solow (1988).

- **Test de Hubert (1989)**

Hubert et *al.* (1989) ont proposé une procédure originale de segmentation des séries de données. Cette segmentation peut éventuellement mettre en évidence une évolution au sein de la série, sous forme de séquences successives. Seulement une brève description de la méthode est présentée ici; pour plus de détails, on peut consulter Hubert et *al.* (1989).

On définit un segment comme étant une séquence des données issues d'une série chronologique de taille N. Toute partition de la série initiale en m segments constitue une segmentation d'ordre m de cette série. Pour tout ordre de segmentation m compris entre 1 et N, il existe plusieurs segmentations possibles. Il y a donc une étape dans cette méthode qui a la fonction de dénombrer les segmentations possibles pour un ordre donné. Ensuite, il s'agit de déterminer parmi toutes ces possibilités, quelle segmentation est la plus appropriée pour la série. Cette segmentation optimale, déterminée à l'aide d'un algorithme, est celle minimisant l'écart quadratique entre la série et la segmentation considérée. Une fois la segmentation optimale déterminée, une procédure doit être complétée par l'introduction d'une contrainte s'appliquant aux segmentations produites. Celles-ci ne seront acceptables que si les moyennes de deux segments contigus sont significativement différentes. Le test incorporé à l'algorithme de Hubert, est celui des contrastes de Scheffé (1959). Donc, dans l'algorithme, la segmentation qui sera retenue comme étant la segmentation optimale, devra nécessairement être valide au sens du test de Scheffé.

L'hypothèse nulle de ce test (la série étudiée est stationnaire) n'est pas rejetée si la procédure ne produit pas segmentation acceptable d'ordre supérieur ou égal à 2. Elle sera rejetée dans le cas contraire.

2.2 Tests non paramétriques de stationnarité de la moyenne

- **Test de Wilcoxon (1945)**

On peut utiliser le test de Wilcoxon (1945) pour évaluer s'il existe une différence significative entre la moyenne d'une série avant le saut et la moyenne de la série après le saut. Évidemment, il faut connaître *a priori* le moment du saut pour être en mesure d'effectuer ce test.

On divise alors la série de données en deux sous-échantillons de taille n_1 et n_2, regroupant respectivement les observations situées avant et après le saut de la moyenne respectivement, de sorte que $N = n_1 + n_2$. On cherche donc à tester l'hypothèse suivante:

H_0: $\mu_1 = \mu_2$ *les moyennes avant et après le saut sont égales*

avec l'hypothèse alternative (dans le cas d'un test bilatéral):

H_1: $\mu_1 \neq \mu_2$ *les moyennes avant et après le saut sont différentes*

Tel que déjà mentionné dans la section 2, les tests non paramétriques sont effectués non pas sur les données elles-mêmes, mais sur les rangs associés à ces observations. On range donc la série en ordre croissant et on associe un rang de classement à chacune des observations. On obtient alors:

$R_1, R_2, \ldots, R_{n_1}$: *les rangs des observations situées avant le saut dans l'échantillon total;*

$S_1, S_2, \ldots, S_{n_2}$: *les rangs des observations situées après le saut dans l'échantillon total.*

Pour tous les tests non paramétriques qui s'appuient sur la notion de rang de classement, à moins d'avis contraire, les rangs des données situées dans *k* sous-groupes sont assignés selon leur position dans l'échantillon total. L'exemple suivant permettra de mieux comprendre la manière dont les rangs doivent être assignés dans une série de données.

Exemple 1 :

Prenons une série de données x_i (représentant une caractéristique quelconque) de taille $N = 10$ (tableau 2.1). On désire vérifier l'existence d'un saut au temps $t = 7$. On range d'abord les données en ordre croissant et on leur attribue un rang de classement (tableau 2.1).

Tableau 2.1 Exemple d'assignation des rangs de classement pour une série de taille $N = 10$.

Temps	1	2	3	4	5	6	7	8	9	10
Observations	100	75	85	74	92	88	60	65	55	45
Rangs	10	6	7	5	9	8	3	4	2	1

Par la suite, la série peut être divisée en deux sous-échantillons, en ne conservant que les rangs de classement des observations dans l'échantillon total.

Rangs des données situées avant le saut de la moyenne (R_i):

$$R_1 = 10 \ ; \ R_2 = 6 \ ; \ R_3 = 7 \ ; \ R_4 = 5 \ ; \ R_5 = 9 \ ; \ R_6 = 8$$

Rangs des données situées après le saut de la moyenne (S_j):

$$S_1 = 3 \ ; \ S_2 = 4 \ ; \ S_3 = 2 \ ; \ S_4 = 1$$

La statistique du test de Wilcoxon, est la somme des rangs des observations situées après le saut de la moyenne:

$$W_s(n_1, n_2) = \sum_{j=1}^{n_2} S_j$$

On pourrait aussi utiliser la somme des rangs des observations situées avant le saut comme valeur de la statistique du test:

$$W_r(n_1, n_2) = \sum_{i=1}^{n_1} R_i$$

Ces deux statistiques sont liées par la relation suivante:

$$W_s(n_1, n_2) + W_r(n_1, n_2) = \frac{N(N+1)}{2}$$

On rejette H_0:, l'hypothèse d'égalité des moyennes, lorsque:

$$W_s(n_1, n_2) > c_1 \qquad \text{ou} \qquad W_s(n_1, n_2) < c_2$$

Les valeurs critiques c_1 et c_2 sont choisies de telle manière que:

$$P_{H_0}\{W_s(n_1, n_2) \le c_1\} + P_{H_0}\{W_s(n_1, n_2) \ge c_2\} = \alpha$$

On choisit généralement c_1 et c_2 tels que:

$$P_{H_0}\{W_s \le c_1\} \approx \alpha/2 \qquad \text{et} \qquad P_{H_0}\{W_s \ge c_2\} \approx \alpha/2$$

La loi exacte de $W_s(n_1, n_2)$ se calcule par dénombrement et elle est tabulée dans Lehmann (1975).

On peut résumer les étapes de ce test de la manière suivante:

Étape 1

Ranger les données en ordre croissant et assigner un rang à chacune des observations.

Étape 2

Calculer $W_s(n_1, n_2)$, la somme des rangs des observations situées après le saut.

Étape 3

Trouver les valeurs critiques c_1 et c_2 à partir des tables et du niveau de signification fixé α.

Étape 4

Au niveau de signification α, rejeter H_0 si: $W_s(n_1, n_2) > c_1$ ou $W_s(n_1, n_2) < c_2$.

- **Test de Mann-Whitney (1947)**

On peut utiliser le test de Mann-Whitney (1947) pour évaluer s'il existe une différence significative entre la moyenne d'une série avant le saut et la moyenne de la série après le saut. On doit cependant connaître *a priori* le moment du saut pour être en mesure d'effectuer ce test.

Soit n_1 et n_2, le nombre d'observations situées avant et après le saut de la moyenne respectivement, de sorte que $N = n_1 + n_2$. On range les données en ordre croissant et on peut ainsi calculer les valeurs suivantes:

$$V = W_r(n_1, n_2) - \frac{n_1(n_1 + 1)}{2}$$

$$W = n_1 n_2 - V$$

où $W_r(n_1, n_2)$ est la statistique du test de Wilcoxon, la somme des rangs des données situées avant le saut de la moyenne. V représente le nombre de fois qu'une observation située avant le saut suit une observation située après le saut dans l'échantillon global selon les rangs qui leur ont été assignés et W représente le nombre de fois qu'une observation située après le saut suit une observation située avant le saut dans l'échantillon global (Bobée et Ashkar, 1991).

Soit la variable U telle que :

$U =$ *la plus faible des valeurs entre V et W.*

$U = minimum(V, W)$

Pour un échantillon de taille N supérieure à 20 et pour des tailles de groupes n_1 et n_2 supérieures à 3, on peut montrer que la variable U suit approximativement une loi normale de moyenne $E[U]$ et de variance $Var[U]$, données respectivement par:

$$E[U] = \frac{n_1 n_2}{2}$$

$$Var[U] = \frac{n_1 n_2}{N(N-1)} \left(\frac{N^3 - N}{12} - \sum T \right)$$

avec $T = (J^3 - J)/12$, où J est le nombre d'égalités à un certain rang. La somme $\sum T$ se fait pour toutes les égalités de rangs dans les deux groupes de données de taille n_1 et n_2.

La statistique du test sous l'hypothèse nulle (égalité des moyennes) est alors:

$$Z = \frac{U - E[U]}{\sqrt{Var(U)}} \approx N(0,1)$$

où $N(0,1)$ représente la distribution normale de moyenne nulle et de variance égale à 1.

On peut résumer les étapes de ce test de la manière suivante:

Étape 1

Ranger les données en ordre croissant et assigner un rang à chacune de celles-ci.

Étape 2

Calculer les rangs des données. Calculer V, W et en déduire la valeur de U.

Étape 3

Tester l'hypothèse nulle d'égalité des moyennes avec la statistique Z en calculant sa valeur observée comme suit:

$$Z_{obs} = \frac{U - \overline{U}}{\sigma_U}$$

Étape 4

Au niveau de signification α, rejeter H_0 si : $|Z_{obs}| \notin [z_{1-\alpha/2}, z_{\alpha/2}]$.

- **Test de Lepage (1991)**

Le test de Lepage (Yonetani, 1991) peut être utilisé pour tester la présence d'un saut dans une série. Il faut donc considérer la série d'observations sous forme de deux sous-échantillons. La statistique du test de Lepage est en fait une somme des carrés de la statistique de Wilcoxon (1945) et de la statistique de Ansari-Bradley qui est présentée à la section 3.2 (Yonetani, 1991):

$$HK = \frac{(W - E[W])^2}{Var[W]} + \frac{(A - E[A])^2}{Var[A]}$$

Le calcul de la statistique du test doit donc être effectué sur les données rangées en ordre croissant. Considérons $u_i = 1$ lorsque la i^{e} plus petite observation appartient aux données situées avant le saut de la moyenne et que $u_i = 0$ lorsque celle-ci est située après le saut. On peut donc en déduire les statistiques de Wilcoxon présenté précédemment et de Ansari-Bradley:

$$W = \sum_{i=1}^{n_1+n_2} i u_i$$

$$A = \sum_{i=1}^{n_1} i u_i + \sum_{i=n_1+1}^{n_1+n_2} (n_1 + n_2 - i + 1) u_i$$

Pour un échantillon de taille suffisamment grande, on peut montrer que W et A suivent approximativement des distributions normales de moyennes respectivement $E[W]$ et $E[A]$ et de variances $Var[W]$ et $Var[A]$.

$$E[W] = \frac{1}{2} n_1 (n_1 + n_2 + 1); \qquad Var[W] = \frac{1}{12} n_1 n_2 (n_1 + n_2 + 1)$$

$$E[A] = \frac{1}{4} n_1 (n_1 + n_2 + 2); \qquad Var[A] = \frac{n_1 n_2 (n_1 + n_2 - 2)(n_1 + n_2 + 2)}{48(n_1 + n_2 - 1)}$$

Si l'on soustrait la moyenne aux statistiques W et A et qu'on les divise par leurs écarts-types, on se retrouve avec deux statistiques qui suivent une distribution normale centrée réduite. Puisque la somme des carrés de lois normales centrées réduites est équivalente à une loi du khi-deux, on peut dire que la statistique HK suit une loi du khi-deux avec deux degrés de liberté.

On peut résumer les étapes de ce test de la manière suivante:

Étape 1

On range les données en ordre croissant et on en déduit les valeurs u_i.

Étape 2

On calcule la statistique du test HK.

Étape 3

Au niveau de signification α, rejeter l'hypothèse H_0 (absence de différence entre les données avant et après le saut) si: $HK > \chi^2_{\alpha,(2)}$.

- **Test de Terry (1952) - Hoeffding (1950)**

Ce test consiste à comparer la moyenne de deux sous-échantillons d'une série de données. On doit d'abord ranger les données en ordre croissant et leur associer un rang comme dans le cas du test de Mann-Whitney (1947) présenté précédemment. Par la suite, les rangs sont remplacés par la valeur correspondante de la variable normale centrée réduite (Marascuilo et McSweeney, 1977). Soit la variable T_1 définie par:

T_1 = *la somme des valeurs de Z (statistiques normales centrée réduites N(0,1)) associées aux rangs du premier sous-échantillon dans la série complète.*

Pour un échantillon suffisamment grand, c'est-à-dire où les deux sous-échantillons d'intérêt ont une taille supérieure à 8, on peut montrer que la variable T_1 suit approximativement une distribution normale de moyenne $E[T_1]$ et de variance $Var[T_1]$. On calcule la moyenne et la variance de la variable T_1 de la manière suivante:

$$E[T_1] = 0$$

$$Var[T_1] = \frac{n_1 n_2}{(n_1 + n_2 - 1)} \sum_{i=1}^{n_1+n_2} \frac{Z(x_i)^2}{n_1 + n_2}$$

où

$$Z(x_i) = \textit{valeur de Z associée aux } x_i.$$

La statistique du test sous l'hypothèse nulle (égalité des moyennes des sous-échantillons de taille n_1 et n_2) est alors:

$$Z = \frac{T_1 - E[T_1]}{\sqrt{Var[T_1]}} \approx N(0,1)$$

On peut résumer les étapes de ce test de la manière suivante:

Étape 1

Identifier le moment du saut potentiel dans la moyenne de la série et diviser cette dernière en deux sous-échantillons contenant les données avant et après le saut.

Étape 2

Calculer T_1, la somme des valeurs de Z associées aux rangs du premier échantillon dans la série complète:

$$T_1 = \sum_{i=1}^{n_1} Z(x_i)$$

Étape 3

Tester l'hypothèse nulle d'égalité des moyennes avec la statistique Z en calculant sa valeur observée $|Z_{obs}|$.

Étape 4

Au niveau de signification α, rejeter H_0 si: $|Z_{obs}| \notin \left[z_{1-\alpha/2}, z_{\alpha/2}\right]$.

- **Test de Van der Waerden (1952/1953)**

Ce test consiste à comparer la moyenne de deux sous-échantillons d'une série de données en remplaçant les observations par les valeurs de la distribution normale centrée réduite associées à la valeur de leur rang (Marascuilo et McSweeney, 1977; Hajek et Sidak, 1967). Soit les variables suivantes:

$$P_i = \frac{R_i}{(n_1 + n_2 + 1)}$$

P_i représente la probabilité empirique associée au rang R_i.

Z_i = *valeur de Z (variable N(0,1)) associée à* P_i.

$$T_1 = \sum_{i=1}^{n_1} Z_i$$

T_1 représente la somme des valeurs Z correspondant aux probabilités empiriques associées aux rangs des observations du sous-échantillon de taille n_1.

La différence majeure entre ce test et le test de Terry-Hoeffding présenté précédemment, réside au niveau de la détermination des valeurs de Z. Dans le cas du test de Terry-Hoeffding, les valeurs de Z sont déterminées directement à partir des rangs des observations tandis que dans le cas de Van der Waerden, ces valeurs sont déterminées à partir des probabilités empiriques associées aux rangs des observations.

Pour un échantillon suffisamment grand, c'est-à-dire où les deux sous-échantillons d'intérêt ont une taille supérieure à 8, on peut montrer que la variable T_1 suit approximativement une distribution normale de moyenne $E[T_1]$ et de variance $Var[T_1]$. On calcule ainsi la moyenne et la variance de la variable T_1 comme suit:

$$E[T_1] = 0$$

$$Var[T_1] = \frac{n_1 n_2}{(n_1 + n_2 - 1)} \frac{\sum_{i=1}^{n_1+n_2} Z_i^2}{(n_1 + n_2)}$$

La statistique du test sous l'hypothèse nulle (égalité des moyennes des sous-échantillons de taille n_1 et n_2) est alors:

$$Z = \frac{T_1 - E[T_1]}{\sqrt{Var[T_1]}} \approx N(0,1)$$

On peut résumer les étapes de ce test de la manière suivante:

Étape 1

Identifier le moment du saut potentiel dans la moyenne de la série et diviser cette dernière en deux sous-échantillons contenant les données avant et après le saut.

Étape 2

Calculer T_1, la somme des valeurs de Z associées aux probabilités empiriques des rangs des observations du premier échantillon de la série:

$$T_1 = \sum_{i=1}^{n_1} Z_i$$

Étape 3

Tester l'hypothèse nulle d'égalité des moyennes avec la statistique Z en calculant sa valeur observée $|Z_{obs}|$.

Étape 4

Au niveau de signification α, rejeter H_0 si: $|Z_{obs}| \notin [z_{1-\alpha/2}, z_{\alpha/2}]$.

Des applications de ce test sont présentées dans Berryman (1984) et Berryman et *al.* (1988).

- **Test de Bell-Doksum**

Ce test consiste à comparer la moyenne de deux sous-échantillons d'une série de données. On doit tirer de façon aléatoire un nombre n de valeurs de Z (variables $N(0,1)$). Ce nombre n doit correspondre à la taille de la série de données. On ordonne ensuite ces valeurs de Z par ordre croissant en leur assignant un rang (Marascuilo et McSweeney, 1977).

Soit la variable T_1 telle que:

T_1 = *la somme des Z de même rang que les observations du premier sous-échantillon.*

$$T_1 = \sum_{i=1}^{n_1} Z_i$$

où

Z_i = *valeur de Z prise au hasard de même rang que l'observation* x_i.

De la même façon, on calcule la variable T_2 pour le deuxième sous-échantillon. Pour un échantillon suffisamment grand, c'est-à-dire où les deux sous-échantillons d'intérêt ont une taille supérieure à 8, on peut montrer que la statistique Z suivante suit une distribution normale centrée réduite $N(0,1)$:

$$Z = \frac{\left(\frac{T_1}{n_1}\right) - \left(\frac{T_2}{n_2}\right)}{\sqrt{\frac{1}{n_1} - \frac{1}{n_2}}} \approx N(0,1)$$

On peut résumer les étapes de ce test de la manière suivante:

Étape 1

Identifier le moment du saut potentiel dans la moyenne de la série et diviser cette dernière en deux sous-échantillons contenants les données avant et après le saut.

Étape 2

Calculer T_1 et T_2, la somme des valeurs de Z associées aux probabilités empiriques associées aux rangs des observations du premier et du second échantillon de la série, respectivement:

$$T_1 = \sum_{i=1}^{n_1} Z_i \,; \qquad T_2 = \sum_{i=n_1+1}^{n_2} Z_i$$

Étape 3

Tester l'hypothèse nulle d'égalité des moyennes avec la statistique Z en calculant sa valeur observée $|Z_{obs}|$.

Étape 4

Au niveau de signification α, rejeter H_0 si: $|Z_{obs}| \notin \left[z_{1-\alpha/2}, z_{\alpha/2}\right]$.

Des applications de ce test sont présentées dans Berryman (1984) et Berryman et *al.* (1988).

- **Test de Terpstra (1952) - Jonckheere (1954)**

Ce test a été développé indépendamment par Terpstra (1952) et Jonckheere (1954) (Neave et Worthington, 1988). Il sert à vérifier l'existence d'une tendance en escalier dans une série de données (cf. section 1). Il consiste à diviser la série de N observations en k sous-échantillons où chacun de ceux-ci est comparé au suivant. Les hypothèses que l'on peut tester sont définies de la manière suivante (Hollander et Wolfe, 1973):

H_0 : *les k échantillons ont la même moyenne;* $\mu_1 = \mu_2 = \ldots = \mu_k$;

avec l'hypothèse alternative:

H_1 : *les moyennes des échantillons sont ordonnées;* $\mu_1 \leq \mu_2 \leq \ldots \leq \mu_k$;
avec au moins une égalité stricte.

Soit la variable J telle que:

J = *nombre de couples concordants*

c'est-à-dire le nombre de couples où la valeur de l'observation d'un échantillon est inférieure à celle de l'échantillon suivant.

Pour pouvoir calculer la statistique du test J, on doit d'abord définir la variable U_{ab}, en considérant que a représente l'indice du premier échantillon de la comparaison et b, l'indice du deuxième échantillon de la comparaison:

$$U_{ab} = \sum_{i=1}^{n_a} \sum_{i=1}^{n_b} \Phi(x_{ia}, x_{ib})$$

où n_a est la taille du premier échantillon de la comparaison, n_b est la taille du deuxième échantillon de la comparaison et où:

$$\Phi(x_{ia}, x_{ib}) = 1 \quad si\ x_{ia} < x_{ib}$$

$$\Phi(x_{ia}, x_{ib}) = 1/2 \quad si\ x_{ia} = x_{ib}$$

$$\Phi(x_{ia}, x_{ib}) = 0 \quad si\ x_{ia} > x_{ib}$$

On peut donc exprimer le nombre total de couples concordants de la façon suivante:

$$J = \sum_{a=1}^{k-1} \sum_{b=a+1}^{k} U_{ab}$$

Pour un échantillon suffisamment grand, on peut montrer que la variable J suit approximativement une distribution normale de moyenne $E[J]$ et de variance $Var[J]$, données par :

$$E[J] = \frac{1}{4}\left(N^2 - \sum_{j=1}^{k} n_j^2\right)$$

$$Var[J] = \frac{1}{72}\left[N^2(2N+3) - \sum_{j=1}^{k} n_j^2(2n_j + 3)\right]$$

La statistique du test sous l'hypothèse nulle (égalité des moyennes des k sous-échantillons) est alors:

$$Z = \frac{J - E[J]}{\sqrt{Var[J]}} \approx N(0,1)$$

On peut résumer les étapes de ce test de la manière suivante:

Étape 1

Identifier la date des sauts potentiels dans la moyenne de la série et diviser cette dernière en k sous-échantillons.

Étape 2

Évaluer le nombre de couples concordants J.

Étape 3

Tester l'hypothèse nulle d'égalité des moyennes avec la statistique Z en calculant sa valeur observée $|Z_{obs}|$.

Étape 4

Au niveau de signification α, rejeter H_0 si: $|Z_{obs}| \notin [z_{1-\alpha/2}, z_{\alpha/2}]$.

Des applications de ce test sont présentées dans Berryman (1984) et Berryman et *al.* (1988).

- **Test de Kruskall-Wallis (1952)**

Ce test très utilisé en pratique, sert à vérifier la tendance en escalier d'une série de données, en vérifiant l'égalité des moyennes de k sous-échantillons (Kruskall et Wallis, 1952). Contrairement au test de Terpstra-Jonckheere, l'hypothèse alternative est qu'au moins une des moyennes est différente des autres (Cochet et Bobée, 1980).

Soit la variable K telle que:

$$K = \frac{12}{N(N-1)}\left(\sum_{i=1}^{k}\frac{R_i^2}{n_i}\right) - 3(N+1)$$

où R_i est la somme des rangs de l'échantillon i de taille n_i.

On peut montrer que pour un échantillon contenant plus de deux sous-échantillons (k>2), chacun de taille supérieure à 5 (n_i>5), la statistique K suit approximativement une loi du khi-deux à k-1 degrés de liberté.

On peut résumer les étapes de ce test de la manière suivante:

Étape 1

Identifier la date des sauts potentiels dans la moyenne de la série et diviser cette dernière en k sous-échantillons.

Étape 2

Ranger les données en ordre croissant et assigner un rang à chacune d'elles.

Étape 3

Calculer la valeur de la statistique du test K.

Étape 4

Au niveau de signification α, rejeter l'hypothèse H_0 (égalité des moyennes des k sous-échantillons) si: $K > \chi^2_{\alpha,(k-1)}$.

- **Test de Mann (1945)**

Le test de Mann (1945) est utilisé pour tester si les données d'une série sont ordonnées aléatoirement contre l'alternative d'une tendance monotone dans le temps (WMO, 1988). Considérons les rangs R_i des N observations rangées en ordre chronologique.

Soit la variable S telle que:

$$M = \sum_{i<j}^{N} S_{ij}$$

où

$$S_{ij} = 1 \qquad si\ \left(R_j - R_i\right) > 0$$

$$S_{ij} = 0 \qquad si\ \left(R_j - R_i\right) = 0$$

$$S_{ij} = -1 \qquad si\ \left(R_j - R_i\right) < 0$$

On peut montrer que pour un échantillon de grande taille ($N \geq 10$), la statistique M suit approximativement une distribution normale de moyenne $E[M]$ et de variance $Var[M]$. On calcule ainsi la moyenne et la variance de la variable M de la façon suivante:

$$E[M] = 0$$

$$Var[M] = \frac{N(N-1)(2N+5)}{18}$$

La statistique du test sous l'hypothèse nulle (données distribuées aléatoirement) est alors:

$$Z = \frac{M - E[M]}{\sqrt{Var[M]}} \approx N(0,1)$$

On peut résumer les étapes de ce test de la manière suivante:

Étape 1

Ranger les données de la série en ordre croissant et leur assigner un rang.

Étape 2

Calculer la valeur de S_{ij} pour chaque paire de données et en faire la somme.

Étape 3

Tester l'hypothèse nulle avec la statistique Z.

Étape 4

Au niveau de signification α, rejeter H_0 si: $|Z_{obs}| \notin [z_{1-\alpha/2}, z_{\alpha/2}]$.

- **Test de Kendall (*turning points test*) (1976)**

Le test d'indépendance de Kendall (1976) peut être utilisé pour tester la stationnarité d'une série de données (Ouarda et *al.*, 1996). Il consiste à assigner une valeur de 1 à chacune des observations x_i supérieure ou inférieure à l'observation la précédant et à l'observation la suivant (Srikanthan et *al.*, 1983):

Si $\{x_i > x_{i-1}; x_i > x_{i+1}\}$ ou $\{x_i < x_{i-1}; x_i < x_{i+1}\}$; alors on assigne une valeur de 1 à x_i;
Sinon, on assigne une valeur de 0 à x_i.

Soit la variable K telle que:

$$K = \textit{nombre de valeurs de } 1 \textit{ obtenues par le critère ci-haut}$$

Pour un échantillon de taille N suffisamment grand, on peut montrer que la variable K suit approximativement une distribution normale de moyenne $E[K]$ et de variance $Var[K]$, données par:

$$E[K] = \frac{2(N-2)}{3}$$

$$Var[K] = \frac{16N-29}{90}$$

La statistique du test sous l'hypothèse nulle (absence de tendance dans la série) est alors:

$$Z = \frac{K - E[K]}{\sqrt{Var[K]}} \approx N(0,1)$$

On peut résumer les étapes de ce test de la manière suivante:

Étape 1

Assigner les valeurs de 0 ou de 1 aux données de la série selon le critère mentionné précédemment.

Étape 2

Dénombrer la quantité de valeurs de 1 et calculer ensuite la moyenne et la variance de la statistique K.

Étape 3

Tester l'hypothèse nulle d'absence de tendance, avec la statistique Z en calculant sa valeur observée $|Z_{obs}|$.

Étape 4

Au niveau de signification α, rejeter H_0 si: $|Z_{obs}| \notin [z_{1-\alpha/2}; z_{\alpha/2}]$.

- **Test de Gold (1929)**

Le test d'indépendance de Gold (1929) peut être utilisé pour tester la stationnarité d'une série de données (Ouarda et *al.*, 1996). Il consiste à évaluer le nombre de séquences de données inférieures ou supérieures à la médiane.

Soit les variables S et G telles que:

s = *le nombre d'observations d'une série de données consécutives inférieures ou supérieures à la médiane*

G(s) = le nombre total de séries de taille s, étant inférieures ou supérieures à la médiane

L'espérance de la variable *G(s)* est donnée par:

$$E[G(s)] = \frac{(N+3-s)}{2^{s+1}}$$

où N est la taille de la série de données. On peut montrer que sous l'hypothèse nulle d'absence de tendance, la statistique Q suivante suit approximativement une loi du khi-deux avec $(s'-1)$ degrés de liberté:

$$Q = \sum_{s=1}^{s'} \left(\frac{\left(G(s) - E[G(s)]\right)^2}{E[G(s)]} \right) \approx \chi^2_{\alpha,(s'-1)}$$

où s' représente la longueur maximale d'une séquence de données étant inférieures ou supérieures à la médiane.

On peut résumer les étapes de ce test de la manière suivante:

Étape 1

Identifier toutes les séquences de données consécutives de différentes longueurs.

Étape 2

Déduire le nombre total de séries pour chaque longueur et calculer les espérances $E[G(s)]$.

Étape 4

Calculer la statistique Q en faisant la somme pour chaque taille de séquence s.

Étape 5

Au niveau de signification α, rejeter l'hypothèse H_0 (absence de tendance) si:

$Q > \chi^2_{\alpha,(s'-1)}$.

- **Test de Pettitt (1979)**

Pettitt (1979) a proposé un test qui est basé sur la statistique de Mann-Whitney (1947) présenté précédemment. Ce test permet de déterminer à quelle date un saut a pu se produire dans la moyenne d'une série de données chronologiques de taille N (Slivitzky et Mathier, 1994).

Soit la variable X_k telle que :

$$X_k = 2R_k - k(N+1)$$

où

$$R_k = \sum_{i=1}^{k} r_i$$

Le moment du saut est représenté par la variable k qui peut prendre les valeurs de 1, ..., N. La variable r_i représente le rang de l'observation x_i dans la série entière rangée en ordre croissant et ainsi, R_k représente la somme des rangs des observations situées avant le saut, c'est-à-dire, la somme des rangs des k premières observations de la série chronologique.

Le saut potentiel de la moyenne est déterminé en calculant la statistique X_k pour tout k. La date du saut potentiel est alors déterminé en évaluant pour quelle valeur de k la statistique X_k est maximale ou minimale, dépendant du fait que le changement est une diminution ou une augmentation de la moyenne.

La distribution de cette statistique (pour $N > 30$) est basée sur une expression asymptotique de la distribution Kolmogorov-Smirnov (Pettitt, 1979).

2.3 Conditions d'application des tests de stationnarité de la moyenne

Le tableau récapitulatif suivant résume les conditions d'application des différents tests paramétriques et non paramétriques de stationnarité de la moyenne qui ont été décrits dans les sections 2.1 et 2.2. Pour chacun des tests, on note s'il peut être appliqué dans le cas de détection de saut, de tendance monotone ou de tendance en escalier de la moyenne.

Tableau 2.2 Conditions d'application des tests de stationnarité de la moyenne

Tests	Saut	Tendance monotone	Tendance en escalier
Paramétriques			
Student	•		
Déviations cumulées	•		
Worsley	•		
Hinkley	•		
Cramer	•		•
Régression		•	
Kite		•	
Bayes	•		
Hubert			•
Non paramétriques			
Wilcoxon	•		
Mann-Whitney	•		
Lepage	•		
Terry-Hoeffding	•		
Van der Waerden	•		
Bell-Doksum	•		
Terpstra-Jonckheere	•		•
Kruskall-Wallis	•		•
Mann		•	
Kendall *turning point*		•	
Gold		•	
Pettitt	•		

2.4 Comparaisons des tests de stationnarité de la moyenne

Dans la section précédente, un certain nombre de tests servant à tester la non-stationnarité d'une série de données ont été présentés. Ces tests possèdent chacun leurs avantages et leurs inconvénients et certains sont mieux adaptés que d'autres dans des situations particulières. On doit d'abord mentionner que de manière générale les tests paramétriques sont plus puissants que les tests non paramétriques lorsque les données sont indépendamment et identiquement distribuées selon la loi normale. Mais dans le cas où la distribution des données est inconnue, l'hypothèse de normalité a été rejetée ou bien que l'on ne dispose de pas suffisamment de données pour s'en assurer, les tests non paramétriques sont beaucoup plus puissants que les tests paramétriques. Ces tests présentent donc l'avantage d'être indépendants de la distribution des données et d'être applicables pour des échantillons de faible taille.

Dans le cas où l'on utilise un test non paramétrique, on peut déterminer les valeurs critiques du test de deux façons, selon la longueur de la série de données. D'abord, pour de petits échantillons, le calcul de probabilité exacte peut être effectué théoriquement par dénombrement. Des tables ont été construites afin d'obtenir les valeurs critiques de chacun de ces tests et en faciliter l'utilisation. Ensuite, pour de grands échantillons, on utilise l'approximation normale dans le but d'accroître l'efficacité de ces tests. Le terme " grand échantillon " a une signification différente pour chaque test non paramétrique. Par exemple, pour le test de Mann-Whitney, il est nécessaire d'avoir un échantillon de taille $N = 20$ et des tailles de sous-échantillons n_1 et n_2 supérieures à 3, pour que l'approximation normale soit valable; tandis que dans le cas du test de Terry-Hoeffding, il est nécessaire d'avoir des sous-échantillons de plus de 8 observations.

Dans le but d'augmenter l'efficacité des tests, il est important d'effectuer une analyse préliminaire des données, afin de déterminer le type de non-stationnarité que la série peut présenter. On peut d'abord faire un examen graphique des données, en examinant l'évolution de la série dans le temps, mais cette méthode requiert généralement une très bonne connaissance physique du domaine en question. Une alternative intéressante est de construire un graphique de la fonction CUSUM et de la fonction double-masse, afin de déceler une cassure (saut), une forme parabolique (tendance monotone) ou plusieurs cassures (tendance en escalier). Ces deux méthodes d'analyse graphique sont décrites brièvement dans le travail de Jean-Cléophas Ondo et on peut aussi trouver une bonne description dans Cluis et *al.* (1987). Il est donc important d'effectuer une analyse préalable de la série.

Pour les tests de non-stationnarité par saut de la moyenne, c'est le test de Mann-Whitney qui est généralement le plus efficace. C'est le test non paramétrique le plus utilisé en raison de sa puissance. En effet, l'efficacité asymptotique relative de ce test est toujours supérieure à 0,864 (Marascuilo, 1977), alors que dans certaines conditions, celle du test de Student peut être nulle. En fait, pour plusieurs distributions l'efficacité asymptotique relative est supérieure à celle du test de Student (Hollander, 1973). De plus, dans les conditions optimales d'application du test paramétrique de Student, c'est-à-dire dans le cas où les données sont indépendantes et identiquement distribuées normalement, le test de Mann-Whitney a presque la même puissance (0,955) que pour le test de Student (1,000). D'autres tests comme ceux de Terry-Hoeffding, de Van der Waerden et de Bell-Doksum, possèdent une efficacité asymptotique relative comparable à celle du test de Mann-Whitney. En effet, lorsque les populations d'origine ont une composante asymptotique réduite, comme dans le cas de la distribution normale et de la loi exponentielle, l'efficacité de ces tests est supérieure à celle du test de Mann-Whitney. Par contre, dans le cas où la composante asymptotique est grande, par exemple pour les distributions uniforme, double exponentielle et Cauchy, l'efficacité du test de Mann-Whitney est supérieure (Marascuilo, 1977). Par ailleurs, une approche intéressante à été proposée par Lee et Heghinian (1977), c'est un test basé sur l'approche bayésienne. Cette méthode permet d'associer une distribution de probabilité aux paramètres τ et δ, le moment et l'amplitude du saut. On peut toujours par la suite, effectuer un test de comparaison de moyennes après avoir déterminé le moment du saut le plus probable avec la fonction de distribution *a posteriori*.

Pour ce qui est des tests de détection de tendance monotone, c'est le test de Mann qui semble être le plus indiqué. D'abord il possède tous les avantages d'un test non paramétrique et ensuite, il a la qualité d'être relativement simple d'utilisation. La méthode de détection de tendance monotone basée sur la régression est elle aussi assez simple à utiliser, mais elle possède l'inconvénient d'avoir à vérifier les hypothèses classiques sur les résidus; ils doivent être indépendants et identiquement distribués selon une loi normale de moyenne nulle et de variance σ^2. De plus, ce test est inutilisable lorsqu'il y a présence d'autocorrélation dans la série, ce qui est fréquent dans le cas de données climatiques (Woodward et Gray, 1993). Finalement, le test de Gold et le test de Kendall (*turning point*) ont une puissance à peu près équivalente (Srikanthan et *al.*, 1983), mais celle-ci demeure relativement faible.

Finalement, pour les tests de stationnarité appliqués à une tendance en escalier, les tests de Kruskall-Wallis et de Terpstra-Jonckeere possèdent la même fonction de puissance, laquelle est comparable à celle du test de non-stationnarité par saut de la moyenne de Mann-Whitney, puisqu'ils représentent tous deux, la généralisation à k sous-échantillons du test de Mann-Whitney utilisé pour comparer la moyenne de deux sous-échantillons (Cochet et Bobée, 1980; Hollander et Wolfe, 1973). Par contre, le test de Terpstra-Jonckeere est mieux adapté à la détection de tendance que le test de Kruskall-Wallis, puisqu'il évalue la croissance ou la décroissance de la moyenne à travers les k sous-échantillons, tandis que le test de Kruskall-Wallis se limite à vérifier les différences entre les k moyennes. Il y a aussi la procédure proposée par Hubert et *al.* qui permet de déterminer la segmentation optimale d'une série et d'en déduire éventuellement une tendance croissante ou décroissante. Cette méthode assez originale permet non seulement de tester l'existence d'une certaine tendance en escalier, mais elle permet aussi d'évaluer toutes les possibilités et d'en déduire la segmentation optimale significative ou non.

3 TESTS DE CHANGEMENTS DANS LA VARIABILITÉ

Comme pour la section 2, cette section sera divisée en deux parties: dans la section 3.1, les tests paramétriques de stationnarité de la variance seront traités et les tests non paramétriques de stationnarité de la variance seront décrits dans la section 3.2. Un tableau synthèse et la comparaison des différentes méthodes seront présentés aux sections 3.3 et 3.4, respectivement.

3.1 Tests paramétriques de stationnarité de la variance

- **Test de Fisher**

Ce test classique est l'équivalent pour la comparaison des variances, du test de Student pour la comparaison de moyennes. Comme dans le cas des tests sur la moyenne, le test de Fisher s'effectue sur la comparaison de deux sous-échantillons d'une série de données (Montgomery, 1991). On doit identifier la date d'un éventuel saut dans la variance et on divise l'échantillon en deux groupes. L'hypothèse à tester est donc la suivante :

$$H_0: \quad \sigma_1^2 = \sigma_2^2$$

avec l'alternative (dans le cas d'un test bilatéral) :

$$H_1: \quad \sigma_1^2 \neq \sigma_2^2$$

Soit la variable F telle que:

$$F = \frac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2}$$

où s_1^2 est la variance du sous-échantillon situé avant le saut ; s_2^2 est la variance du sous-échantillon situé après le saut ; σ_1^2 est la variance de la population qui correspond aux données situées avant le saut et σ_2^2 est la variance de la population qui correspond aux données situées après le saut.

Sous l'hypothèse H_0 d'égalité des variances $\left(\sigma_1^2 = \sigma_2^2\right)$, la statistique F suit une distribution de Fisher avec $n_1 - 1$ et $n_2 - 1$ degrés de liberté:

$$F_{obs} = \frac{s_1^2}{s_2^2} \approx F_{n_1-1,n_2-1}$$

On peut résumer les étapes de ce test de la manière suivante:

Étape 1

Vérifier que les données sont indépendantes et qu'elles proviennent d'une population normale.

Étape 2

Calculer la statistique du test F_{obs}.

Étape 3

Au niveau de signification α, rejeter l'hypothèse H_0: (égalité des variances des deux sous-échantillons) si: $F_{obs} > F_{n_1-1,n_2-1;1-\alpha/2}$.

- **Test de Bartlett (1947)**

Le test de Bartlett (1947) est utilisé pour vérifier l'hypothèse d'égalité des variances de k sous-échantillons d'une série de données de taille N, contre l'hypothèse alternative qu'au moins une des variances est différente des autres. Avant effectuer le test de Bartlett, il faut d'abord s'assurer que les observations sont indépendantes et qu'elles proviennent d'une population normale (Montgomery, 1991).

Soit la variable B telle que:

$$B = 2{,}3026\frac{q}{c}$$

où

$$q = (N-k)\log_{10}\left(s_p^2\right) - \sum_{i=1}^{k}(n_i - 1)\log_{10}\left(s_i^2\right)$$

$$c = 1 + \frac{1}{3(k-1)}\left(\sum_{i=1}^{k}\frac{1}{(n_i-1)} - \frac{1}{(N-k)}\right)$$

$$s_p^2 = \frac{\sum_{i=1}^{k}(n_i - 1)s_i^2}{N-k}$$

et où s_i^2 est la variance du i^e sous-échantillon et s_p^2 est la variance combinée des k sous-échantillons.

Plus les k variances s_i^2 sont différentes, plus la valeur de la variable q est élevée. À l'opposé, si elles sont toutes égales, la valeur de q sera nulle. L'hypothèse H_0 doit donc être rejetée pour de grandes valeurs de B.

On peut montrer que la variable B suit approximativement une distribution du khi-deux à k-1 degrés de liberté ($\chi^2_{\alpha,(k-1)}$), lorsque la taille des k sous-échantillons est supérieure ou égale à 5.

On peut résumer les étapes de ce test de la manière suivante:

Étape 1

Vérifier que les données sont indépendantes et qu'elles proviennent d'un population normale.

Étape 2

Identifier les différents sauts potentiels dans la variance et diviser l'échantillon en k sous-échantillon.

Étape 3

Calculer la variance de chacun des sous-échantillons, en déduire s_p^2, q, c et la valeur de la statistique B.

Étape 4

Au niveau de signification α, rejeter l'hypothèse H_0: (égalité des variances des k sous-échantillons) si: $B > \chi^2_{\alpha,(k-1)}$.

- **Test de Hartley**

Le test de Hartley (Neter et Wasserman, 1974) peut être utilisé pour comparer k variances, lorsque les échantillons proviennent de populations normales et qu'ils ont des tailles identiques $(n_1 = n_2 = \ldots = n_i = \ldots = n)$. Une condition supplémentaire est que la taille des sous-échantillons doit être supérieure à 3 $(n \geq 3)$.

Soit la variable H telle que:

$$H = \frac{\max_i (s_i^2)}{\min_i (s_i^2)}$$

où s_i^2 représente la variance non biaisée de l'échantillon i.

Lorsque la valeur de H est voisine de 1, l'hypothèse d'égalité des k variances est acceptée. Par contre, lorsque H est beaucoup plus élevé que 1, on rejette l'hypothèse H_0. Il existe des tables dans lesquelles ont peut retrouver des valeurs critiques pour la statistique H (Neter et Wasserman, 1974). Ce test a l'inconvénient d'être sensible au non-respect de l'hypothèse de normalité de la population dont proviennent les échantillons et il nécessite l'égalité des tailles des k sous-échantillons.

3.2 Tests non paramétriques de stationnarité de la variance

- **Test de Mood (1954)**

Le test de Mood (1954) est l'équivalent non paramétrique du test de Fisher d'égalité des variances de deux populations. On doit donc diviser l'échantillon en deux sous-échantillons.

Soit n_1 et n_2, le nombre d'observations situées avant et après le saut de la moyenne respectivement, de sorte que $N = n_1 + n_2$. Ainsi, on a:

$x_1, x_2, \ldots, x_{n_1}$ 1[er] sous-échantillon (situé avant le saut);

$x_{n_1+1}, x_{n_1+2}, \ldots, x_N$ 2[e] sous-échantillon (situé après le saut)

Comme dans le cas du test de Mann-Whitney pour la comparaison de deux moyennes, on doit ranger les données en ordre croissant et assigner un rang à chacune d'elles.

Soit la variable M telle que:

$$M = \sum_{i=1}^{n_1} \left[R_i - \frac{N+1}{2} \right]^2$$

où R_i est le rang de l'observation i dans l'échantillon total.

On utilise ce test pour des échantillons assez grands, lorsque les distributions des populations sont éloignées de la normalité. Pour un échantillon de grande taille, on peut montrer que la variable M suit approximativement une distribution normale de moyenne $E[M]$ et de variance $Var[M]$, données par:

$$E[M] = n_1 \frac{N^2 - 1}{12}$$

$$Var[M] = \frac{n_1 n_2 (N+1)(N^2 - 4)}{180}$$

La statistique du test sous l'hypothèse nulle (égalité des variances) est donc:

$$Z = \left| \frac{M - E[M]}{\sqrt{Var[M]}} \right|$$

On peut résumer les étapes de ce test de la manière suivante:

Étape 1

Ranger les données en ordre croissant et associer les rangs pour les données situées avant le saut de la moyenne. Calculer la valeur de M.

Étape 2

Tester l'hypothèse nulle d'égalité des moyennes avec la statistique Z en calculant sa valeur observée comme suit:

$$Z_{obs} = \frac{M - \overline{M}}{\sigma_M}$$

où $\overline{M}$ et σ_M représentent respectivement la moyenne et la variance de la statistique M, telles que définies précédemment.

Étape 3

Au niveau de signification α, rejeter H_0 (égalité des variances) si : $|Z_{obs}| \notin [z_{1-\alpha/2}, z_{\alpha/2}]$.

- **Test de Ansari-Bradley (1960)**

Le test de Ansari-Bradley (1960) est utilisé pour la comparaison de la variance de deux sous-échantillons. Il s'applique uniquement dans le cas où les moyennes des deux sous-échantillons sont supposées égales (Hollander et Wolfe, 1973). Pour représenter les données des deux sous-échantillons, on utilise les modèles suivants:

$$x_i = \mu + \sigma_1 e_i \qquad i = 1, \ldots, n_1$$

$$x_{n_1+j} = \mu + \sigma_2 e_{n_1+j} \qquad j = 1, \ldots, n_2$$

où $e_1, \ldots, e_N$ sont des variables aléatoires indépendantes provenant de la même population ayant une moyenne de 0 et une variance σ_e^2.

Soit la variable A telle que:

$$A = \textit{somme des rangs } R_i \textit{ du premier sous-échantillon}$$
$$A = \sum_{i=1}^{n_1} R_i$$

Les rangs R_i sont assignés de la façon suivante pour les deux sous-échantillons combinés: on attribue la valeur de 1 à la plus petite et à la plus grande valeur de la série. Ensuite, on attribue la valeur 2 à la deuxième plus petite et à la deuxième plus grande valeur, et ainsi de suite. Selon que la taille de la série N soit paire ou impaire, on associe les rangs de la façon suivante:

N pair: $\quad 1, 2, 3, \ldots, N/2, N/2, \ldots, 3, 2, 1$

N *im*pair: $1,2,3,\ldots,(N-1)/2,(N+1)/2,(N-1)/2,\ldots,3,2,1$

On peut trouver une table des valeurs critiques $\omega_1(\alpha_1, n_1, n_2)$ et $\omega_2(\alpha_2, n_1, n_2)$ pour la statistique A dans Hollander et Wolfe (1973) où α_1 et α_2 représentent les seuils.

Dans le cas d'un échantillon de grande taille, on peut montrer que la variable A suit approximativement une distribution normale de moyenne $E[A]$ et de variance $Var[A]$, données par:

$$\begin{cases} E[A] = \dfrac{n_1(n_1+n_2+2)}{4} \\ Var[A] = \dfrac{n_1 n_2 (n_1+n_2+2)(n_1+n_2-2)}{48(n_1+n_2-1)^2} \end{cases} \quad \text{si } N \text{ est pair}$$

$$\begin{cases} E[A] = \dfrac{n_1(n_1+n_2+1)^2}{4(n_1+n_2)} \\ Var[A] = \dfrac{n_1 n_2 (n_1+n_2+1)\left(3+(n_1+n_2)^2\right)}{48(n_1+n_2)^2} \end{cases} \quad \text{si } N \text{ est impair}$$

La statistique du test sous l'hypothèse nulle (égalité des variances) est alors:

$$Z = \frac{A - E[A]}{\sqrt{Var[A]}} \approx N(0,1)$$

On peut résumer les étapes de ce test de la manière suivante:

Étape 1

Vérifier l'égalité des moyennes de deux sous-échantillons.

Étape 2

Ranger les observations en ordre croissant et assigner un rang à chacune d'elles.

Étape 3

Calculer la somme des rangs pour le premier des deux sous-échantillons.

Étape 4a

Pour un échantillon de petite taille, dans le cas d'un test bilatéral, rejeter H_0 (égalité des variances) si:

$$A \geq \omega_2(\alpha_2, n_1, n_2) \qquad \text{ou} \qquad A \leq \omega_1(\alpha_1, n_1, n_2)$$

où $\alpha = \alpha_1 + \alpha_2$.

Étape 4b

Pour un échantillon de grande taille, au niveau de signification α, rejeter H_0 si: $|Z_{obs}| \notin [z_{1-\alpha/2}, z_{\alpha/2}]$.

- **Test de Moses (1963)**

Le test de Moses (1963) s'applique pour comparer les variances de deux parties d'une série de données dans le cas où leurs moyennes ont été jugées significativement différentes (Hollander et Wolfe, 1973). Pour représenter les données des deux sous-échantillons, de taille n_1 et n_2 respectivement, on utilise les modèles suivants:

$$x_i = \mu_1 + \sigma_1 e_i \qquad i = 1, \ldots, n_1$$

$$x_{n_1+j} = \mu_2 + \sigma_2 e_{n_1+j} \qquad j = 1, \ldots, n_2$$

où $e_1,\ldots,e_N$ sont des variables aléatoires indépendantes provenant de la même population ayant une moyenne de 0 et une variance σ_e^2.

On doit choisir une valeur $m \geq 2$ et diviser les séries en n_1' et n_2' sous-groupes de taille m, respectivement. On élimine les observations en trop.

Pour $i = 1,\ldots,n_1'$, les valeurs $x_{i1},\ldots,x_{im}$ représentent le i^e sous-groupe de m observations. De la même façon, pour $j = 1,\ldots,n_2'$, les valeurs $y_{j1},\ldots,y_{jm}$ représentent le j^e sous-groupe de m observations.

Pour la série située avant le saut de la variance, soit les valeurs $C_1,\ldots,C_{n_1'}$ telles que:

$$C_i = \sum_{p=1}^{m}\left(x_{ip} - \bar{x}_i\right)^2 \qquad \text{pour } i = 1,\ldots,n_1'$$

où

$$\bar{x}_i = \frac{1}{m}\sum_{p=1}^{m} x_{ip}$$

Pour la série située après le saut de la variance, soit les valeurs $D_1,\ldots,D_{n_2'}$ telles que:

$$D_j = \sum_{q=1}^{m}\left(y_{jq} - \bar{y}_j\right)^2 \qquad \text{pour } j = 1,\ldots,n_2'.$$

où

$$\bar{y}_j = \frac{1}{m}\sum_{q=1}^{m} y_{jq}$$

Par la suite, le test consiste a utiliser la même procédure que dans le cas du test de Wilcoxon pour la comparaison de moyennes. On range les C_i et les D_j en ordre croissant et on associe un rang R_t à chacune de ces valeurs.

Soit la variable W telle que:

$$W = \textit{la somme des rangs correspondant aux } C_i$$

$$W = \sum_{t=1}^{n_1'} R_t$$

Pour un échantillon de grande taille, on peut montrer que la variable W suit approximativement une distribution normale de moyenne $E[W]$ et de variance $Var[W]$. La moyenne et la variance de la variable W sont données par:

$$E[W] = \frac{n_1'(n_1' + n_2' + 1)}{2}$$
$$Var[W] = \frac{n_1' n_2'(n_1' + n_2' + 1)}{12}$$

La statistique du test sous l'hypothèse nulle (égalité des variances) est alors:

$$Z = \frac{W - E[W]}{\sqrt{Var[W]}} \approx N(0,1)$$

On peut résumer les étapes de ce test de la manière suivante:

Étape 1

Vérifier que les moyennes des deux sous-échantillons sont différentes.

Étape 2

Diviser les séries s_1 et s_2 en n_1' et n_2' sous-groupes de taille m, respectivement et calculer les C_i et les D_j.

Étape 3

Ranger les C_i et les D_j en ordre croissant et leur assigner un rang.

Étape 4

Calculer la somme des rangs pour le premier des deux sous-échantillons.

Étape 5

Pour un échantillon de grande taille, au niveau de signification α, rejeter H_0 (égalité des variances) si: $|Z_{obs}| \notin [z_{1-\alpha/2}, z_{\alpha/2}]$.

- **Test de Kruskal-Wallis**

Le test de Kruskal-Wallis de comparaison de k moyennes peut être utilisé pour tester l'égalité des variances de k séquences de données d'une série, contre l'hypothèse alternative qu'au moins une des variances est différente des autres (WMO, 1988; WMO, 1992). Par contre, dans le cas de comparaison de variances, les rangs sont associés à la valeur absolue de l'écart de chaque observation à la moyenne générale (Sneyers, 1975):

$$D_{ji} = |x_{ji} - \bar{x}| \qquad \text{pour } i = 1,\ldots,k \text{ et } j = 1,\ldots,n_i$$

où x_{ji} représente l'observation j du sous-échantillon i de taille n_i.

Soit la variable K telle que:

$$K = \frac{12}{N(N-1)}\left(\sum_{i=1}^{k}\frac{R_i^2}{n_i}\right) - 3(N+1)$$

où R_i est la somme des rangs de l'échantillon i de taille n_i.

On peut montrer que pour un échantillon contenant plus de deux sous-échantillons (k>2), chacun de taille supérieure à 5 ($n_i > 5$), la statistique K suit approximativement une loi du khi-deux à k-1 degrés de liberté.

On peut résumer les étapes de ce test de la manière suivante:

Étape 1

Identifier le moment des sauts potentiels dans la variance de la série et diviser cette dernière en k sous-échantillons.

Étape 2

Calculer les D_{ji}, les ranger en ordre croissant et leur associer un rang.

Étape 3

Calculer la valeur de la statistique du test K.

Étape 4

Au niveau de signification α, rejeter l'hypothèse H_0 (égalité des variances des k sous-échantillons) si: $K > \chi^2_{\alpha,(k-1)}$

3.3 Conditions d'application des tests de stationnarité de la variance

Le tableau récapitulatif suivant résume les conditions d'application des différents tests paramétriques et non paramétriques de stationnarité de la variance qui ont été mentionnés dans les sections 3.1 et 3.2. Pour chacun des tests, on note s'il peut être appliqué dans le cas de comparaison de la variance de deux parties de la série de données ou dans le cas de comparaison de la variance de k parties de la série de données.

Tableau 3.1 Conditions d'application des tests de stationnarité de la variance

Tests	Comparaison de 2 variances	Comparaison de k variances
Paramétriques		
Fisher	•	
Bartlett		•
Hartley		•
Non paramétriques		
Mood	•	
Ansari-Bradley	•	
Moses	•	
Kruskall-Wallis		•

3.4 Comparaisons des tests de stationnarité de la variance

Dans les sections 3.1 et 3.2, 7 tests de stationnarité de la variance ont été présentés. Trois d'entre eux sont des tests paramétriques et les quatre autres sont non paramétriques. Le test de Hartley est relativement simple à utiliser, mais il n'est pas très puissant. Les tests de Fisher et de Bartlett sont des tests qui sont couramment utilisés dans plusieurs domaines d'étude. Ce sont habituellement ces deux tests qui sont incorporés dans les logiciels d'analyses statistiques, pour vérifier l'homogénéité de la variance d'une série de données. En plus, ils ont l'avantage d'être très simples d'utilisation. Mais, comme la plupart des tests paramétriques, les erreurs doivent être indépendantes et identiquement distribuées selon une loi normale de moyenne nulle et de variance σ^2. De plus, ces deux tests ont la particularité d'être très sensibles au non respect de la normalité.

Le test de Mood, qui est en fait l'équivalent non paramétrique du test de Fisher, représente une bonne alternative à ces deux tests dans le cas où les données ne proviennent pas de populations normales. Finalement, le test de Kruskall-Wallis est un test très efficace qui a l'avantage de pouvoir être utilisé pour tester à l'hypothèse d'égalité de k moyennes et de k variances.

4 CONCLUSIONS

On dit qu'une série de données rangée dans le temps est stationnaire, lorsque l'on ne retrouve pas de variations temporelles significatives, en excluant les fluctuations aléatoires. On peut détecter une non-stationnarité en examinant la moyenne ou la variance de la série et en évaluant s'il y a eu un changement significatif de ceux-ci à travers le temps. De façon générale, on peut identifier quatre types de non-stationnarité: le saut de la moyenne, la tendance de la moyenne, la tendance en escalier de la moyenne et les changements de variabilité.

Les tests traditionnellement utilisés pour tester la stationnarité d'une série de données sont des tests paramétriques, c'est-à-dire, des tests qui sont liés à une certaine fonction de distribution statistique, qui est généralement la loi normale. Pour utiliser ces tests, on doit donc vérifier au préalable certaines hypothèses concernant les données de la série. On doit entre autres vérifier si les observations sont indépendantes et identiquement distribuées selon une loi normale. Cette contrainte peut être assez gênante, en effet lorsque la normalité d'une population est rejetée, il est impossible d'utiliser ce test paramétrique. C'est dans des situations semblables que l'on a recours aux tests non paramétriques qui, eux, ne font aucune hypothèse sur la distribution des observations. Cet avantage a fait en sorte que ces tests ont connu un regain de popularité au cours des dernières années et ce, dans presque tous les domaines de recherche. Les tests non paramétriques sont généralement plus puissants que les tests paramétriques équivalents, lorsqu'il y a absence de normalité. De plus, dans les situations optimales d'application des tests paramétriques, les tests non paramétriques sont dans bien des cas presque aussi puissants que les tests paramétriques.

D'abord, dans le but d'accroître l'efficacité d'un test, on peut d'abord passer à un examen graphique des données à l'aide de la fonction CUSUM et de la fonction double-masse. Ces méthodes permettent de déterminer le type de non stationnarité qui peut être présent dans les données et donc de faire un bon choix de test. Pour les tests de non stationnarité par saut de la moyenne, ce sont les tests de Mann-Whitney, de Terry-Hoeffding, de Van der Waerden et de

Bell-Doksum qui sont les plus efficaces. Pour ce qui est des tendances monotones, c'est le test de Mann qui semble le plus approprié. Finalement, pour les tests de tendance par escalier, le test de Kruskall-Wallis possède une puissance comparable à celle du test de Terpstra-Jonckeere, mais il possède de plus l'avantage de pouvoir être utilisé pour tester la stationnarité de la variance.

Il existe une très grande variété de tests de stationnarité de la moyenne ou de la variance. Dans ce rapport, seulement les tests les plus puissants et les plus utilisés ont été décrits. On peut se référer au rapport indépendance et homogénéité pour la description de tests d'homogénéité qui peuvent être utilisés dans le cas de détection de la stationnarité. En fait tous les tests utilisés pour tester l'homogénéité contre l'hypothèse de tendance par saut ou de tendance en escalier, peuvent aussi être utilisés pour tester la stationnarité.

5 RÉFÉRENCES

Ansari, A.R. et R.A. Bradley (1960). Rank-sum tests for dispersions. Ann. Math. Statist., 33: 1148-1159.

Bartlett, M.S. (1947). The use of transformations. Biometrics, 3: 39-52.

Berger, J.O. (1985). *Statistical decision theory and Bayesian analysis*. Berlin, Springer, 617 pp.

Bernier, J. (1977). Étude de la stationnarité des séries hydrométéorologiques. La Houille Blanche, 4(32): 313-319.

Bernier, J. (1994). Statistical detection of changes in geophysical series. Engineering Risk in Natural Resources Management, Duckstein L. and E. Parent, NATO Advanced Studies Institute Series, Series E-Vol.275, Kluwer, 159-176.

Berryman, D. (1984). La détection des tendances dans les séries temporelles de paramètres de la qualtié de l'eau à l'aide de tests non-paramétriques. M.Sc. Thesis, Institut national de la recherche scientifique (INRS-Eau), Sainte-Foy, Québec.

Berryman, D., B. Bobée, D. Cluis et J. Haemmerli (1988). Nonparametric tests for trend detection in water quality time series. Water Resourc. Res., 24(3): 545-555.

Bobée, B. et F. Ashkar (1991). *The gamma family and derived distributions applied in hydrology*. Water Resources Publications. Littleton, Colorado, 203 pp.

Bruneau, P. et J.-C. Rassam (1983). Application d'un modèle bayésien de détection de changements de moyennes dans une série. Journal des Sciences Hydrologiques, 28: 341-354.

Buishand, T.A. (1982). Some methods for testing the homogeneity of rainfall records. J. Hydrol., 58: 11-27.

Cluis, D., C. Laberge et C. Houle (1987). Détection de tendance en qualité de l'eau. Rapport méthodologique. Institut national de la recherche scientifique (INRS-Eau), Sainte-Foy, Québec, rapport scientifique no. 229, 121pp.

Cochet, Y. et B. Bobée (1980). Études des tendances à l'aide de tests non-paramétriques. Institut national de la recherche scientifique (INRS-Eau), Sainte-Foy, Québec, Rapport interne no. 71.

Cramer, H. (1946). Mathematical methods of statistics. Princeton University Press, N.J., 575 pp.

Gold, E. (1929). Notes on the frequency of occurence of sequence in a series of events of two types. Q.J.R. Meteorol. Soc., 55 : 307-309.

Hajek, J. et Z.S. Sidak (1967). *Theory of rank tests*. Academic Press, New York, 297 pp.

Hinkley, D.V. (1971). Inference about the change-point from cumulative sum tests. Biometrica, 58(3).

Hoeffding, W. (1950). « Optimum » non-parametric tests. Proc. IInd Berkeley Symposium, 83-92.

Hollander, M. et D.A. Wolfe (1973). *Non-parametric statistical methods*. John Wiley, New-York, 503 pp.

Hubert, P., J.P. Carbonnel et A. Chaouche (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.

Jonckheere, A.R. (1954). A distribution-free *k*-sample test against ordered alternatives. Biometrika, 41:133-145.

Kendall, M. et A. Stuart (1976). The advanced theory of statistics. Vol. 3. Charles Griffin, London.

Kite, G. (1989). Use of time series analysis to detect climatic change. J. Hydrol., 111: 259-279.

Kruskall, W.H. et W.A. Wallis (1952). Use of ranks in one-criterion variance analysis. J. Am. Statist. Assoc., 47: 583-612.

Kruskall, W.H. et W.A. Wallis (1953). Use of ranks in one-criterion variance analysis. J. Am. Statist. Assoc., 48: 907-911.

Lee, A.S.F. et S.M. Heghinian (1977). A shift of the mean level in a sequence of independant normal random variables - A Bayesian approach, Technometrics, 19(4): 503-506.

Lehmann, E.L. (1975). *Nonparametrics: statistical methods based on ranks.* Holden-Day, Oakland, California, 457 pp.

Lettenmaier, D.P. (1976). Detection of trends in water quality data from records with dependant observations. Water Resourc. Res., 12(5):1037-1046.

Mann, H.B. (1945). Nonparametric tests against trend. Econometrika, 13: 245-259.

Mann, H.B. et D.R. Whitney (1947). On the test of whether one of two random variables is stochastically larger than the other. Ann. Math. Statist., 18: 50-60.

Marascuilo, L.A. et M. McSweeney (1977). *Non-parametric and distribution-free methods for the social sciences*. Brooks/Cole Pub. Co., 556 pp.

Matalas, N.C. (1967). Time series analysis. Water Resourc. Res., 3: 817-829.

Montgomery, D.C. (1991). *Design and analysis of experiments*. John Wiley & Sons, 649 pp.

Mood, A.M. (1954). On the asymptotic efficiency of certain nonparametric two-sample tests. Ann. Math. Statist., 25: 514-522.

Moses, L.E. (1963). Rank tests of dispersion. Ann. Math. Statist., 34: 972-983.

Neave, H.R. et P.L. Worthington (1988). *Distribution-free tests*. Unwin Hyman ltd., London, U.K., 430 pp.

Neter, J. et W. Wasserman (1974). Applied linear statistical models; regression, analysis of variance and experimental designs. Homewood, Illinois, 842 pp.

Ouarda, T. B.M.J., P.F. Rasmussen et B.Bobée (1996). Rationalisation du réseau hydrométrique de la province de Québec pour le suivi des changements climatiques. Chaire industrielle en hydrologie statistique, INRS-Eau rapport de recherche no R-476, 71 pp.

Perreault, L., M., Haché et M. Slivitzky (1996). Detection of changes in the mean level of precipitation and runoff over Eastern Canada and U.S. using a Bayesian approach. Article soumis à "Climate Change".

Pettitt, A.N. (1979). A nonparametric approach to the change point problem. Applied Statistics, 28(2): 126-135.

Scheffé, M. (1959). The analysis of variance. Wiley, New York, 477 pp.

Slivitzky, A.N. et L. Mathier (1994). Climatic changes during the 20th century on the Laurentian Great Lakes and their impacts on hydrologic regime. Engineering Risk in Natural Resources Management, Duckstein L. and E. Parent, NATO Advanced Studies Institute Series, Series E-Vol.275, Kluwer, 235-251.

Sneyers, R. (1975). Sur l'analyse statistique des séries d'observations. Note technique no. 143, Organisation météorologique mondiale, 192 pp.

Solow, A.R. (1988). A Bayesian approach to statistical inference about climate change. J. Climate, 1(5): 1401-1405.

Srikanthan, R., T.A. McMahon et J.L. Irish (1983). Time series analyses of annual flows of Australian streams. J. of Hydrology, 66: 213-226.

Terpstra, T.J. (1952). The asymptotic normality and consistency of Kendall's test against trend when ties are present in one ranking. Indag. Math., 14: 327-333.

Terry, M.E. (1952). Some rank order tests which are most powerful against specific parametric alternatives. Ann. Math. Statist., 23: 346-366.

Van der Waerden, B.L. (1952/1953). Order tests for the two-sample problem and their power. I, II, III. Indagationes math. 14 (Proc. Kon. Nederl. Akad. Wet. 55), 453-458; Indagationes math. 15 (Proc. 56), 303-310, 311-316; Correction : Indagationes math. 15 (Proc. 56), 80.

Wilcoxon, F. (1945). Individual comparisons by ranking methods. Biometrics 1: 80-83.

W.M.O. (World Climate Programme) (1966). Climatic change. Technical Note, no. 79, WMO no. 195, TP 100.

W.M.O. (World Climate Programme) (1988). Analyzing long Time Series of hydrological data with respect to climate variability and change WCAP-3, WMO/TD - No. 224, World Meteorological Organization.

W.M.O. (World Climate Programme) (1992). A survey of current approaches to modeling of hydrological time series with respect to climate variability and change WCASP-23, WMO/TD - No. 534, World Meteorological Organization.

Worsley, K.J. (1979). On the likelihood ratio test for shift in location of normal populations. J. Am. Stat. Assoc., 74: 365-367.

Woodward, W.A. et H.L. Gray (1993). Global warming and the problem of testing for trend in time series data. J. Climate, 6: 953-962.

Yonetani, T. (1991). Discontinuous changes of precipitation in Japan after 1900 detected by the Lepage test. J. Meteor. Soc. Japan, 70(1): 95-104.