

**ANALYSE FRÉQUENTIELLE DES
PRÉCIPITATIONS DE LA STATION
ASSOMPTION**

**Chaire industrielle Hydro-Québec / CRSNG
en Hydrologie statistique**

Rapport de recherche R937

Mars 2007

**ANALYSE FRÉQUENTIELLE DES PRÉCIPITATIONS DE LA STATION
ASSOMPTION**

par

André St-Hilaire

**Chaire en hydrologie statistique
Institut national de la recherche scientifique (INRS)
Centre Eau, Terre et Environnement
490 de la Couronne, Québec, (Québec) G1K 9A9**

Rapport de recherche

Mars 2007

Référence

ST-HILAIRE A. 2007. Analyse fréquentielle des précipitations de la station Assomption.
Rapport de recherche R937 de l'INRS-ETE. v + 12 pages.

TABLE DES MATIÈRES

TABLE DES MATIÈRES	III
LISTE DES TABLEAUX	IV
LISTE DES FIGURES	V
1. INTRODUCTION	1
2. ANALYSE FRÉQUENTIELLE LOCALE.....	2
2. DONNÉES UTILISÉES ET HYPOTHÈSES DE BASE.....	3
3. RÉSULTATS	7
3.1 ANALYSE FRÉQUENTIELLE DES PLUIES JOURNALIÈRES	7
3.2 COURBES INTENSITÉ-DURÉE-FRÉQUENCE.....	9
7. CONCLUSION.....	11
8. RÉFÉRENCES.....	12

LISTE DES TABLEAUX

TABLEAU 1. STATISTIQUES DESCRIPTIVES DES MAXIMUMS SAISONNIERS POUR DIFFÉRENTES DURÉES ($N=43$).	3
TABLEAU 2. STATISTIQUES DES TESTS POUR DIFFÉRENTES DURÉES. LES VALEURS EN GRAS INDIQUENT LES TESTS ÉCHOUÉS.	4
TABLEAU 3. CRITÈRES DE SÉLECTION DES DISTRIBUTIONS STATISTIQUES AJUSTÉES AUX DONNÉES.	8

LISTE DES FIGURES

FIGURE 1. SÉRIE CHRONOLOGIQUE DES MAXIMUMS SAISONNIERS DE DURÉE 2 HEURES.	5
FIGURE 2. FONCTION D'AUTOCORRÉLATION DES MAXIMUMS SAISONNIERS DE DURÉE 24 HEURES.....	6
FIGURE 3. COMPARAISON DE L'AJUSTEMENT DES MAXIMUMS SAISONNIERS DE DURÉE 4 HEURES.....	9
FIGURE 4. COURBE INTENSITÉ-DURÉE-FRÉQUENCE POUR LA STATION ASSOMPTION.	10

1. INTRODUCTION

Le présent rapport traite de l'analyse de données de hauteur de pluie totale de la station Assomption (Environnement Canada, # 7014160) située dans une zone ayant reçu de fortes précipitations durant l'été 2006.

L'objectif de ce projet consiste à estimer les quantiles d'une période de retour donnée (e.g. 2, 5, 10, 20 100 ans) à l'aide de méthodes d'analyse fréquentielle locale et ce, pour différentes durées. Une fois les quantiles estimés, ces valeurs sont comparées aux précipitations mesurées lors de l'événement du 19 août 2006.

La suite du rapport est structurée comme suit : La section 2 décrit brièvement les fondements de l'analyse fréquentielle locale. Les données utilisées sont décrites dans la section 3 et les résultats sont présentés dans la section 4. Les analyses graphiques détaillées sont présentées en annexes.

2. ANALYSE FRÉQUENTIELLE LOCALE

Un des outils de base de l'analyse de l'occurrence d'événements extrêmes est l'analyse fréquentielle. Cette méthode statistique a comme objectif principal d'utiliser des mesures d'événements passés pour estimer les probabilités futures d'occurrence. L'analyse fréquentielle est utilisée, en particulier, pour estimer l'ampleur de l'événement temporel x_T auquel est associé un période de retour T (quantile de période de retour T ou de probabilité au dépassement $p = 1/T$). L'estimation \hat{x}_T de la valeur du quantile s'obtient en ajustant une loi de probabilité $F(x; \theta)$ à un échantillon de n observations $x = \{x_1, \dots, x_n\}$, où θ représente le vecteur de paramètres associé à la distribution de probabilité F . Les étapes principales de l'analyse fréquentielle ont été présentées dans le rapport de St-Hilaire (2005) et ne seront pas répétées ici. L'ensemble des étapes méthodologiques ont été accomplies à l'aide du logiciel HYFRAN (2003).

2. DONNÉES UTILISÉES ET HYPOTHÈSES DE BASE

Les données de pluie horaires de la station Assomption ont été obtenues de la base de données d'Environnement Canada. Ce pas de temps permet d'estimer les quantiles pour différentes durées en sommant les précipitations horaires mesurées pour la durée d'intérêt. Les durées analysées dans le cadre du présent projet sont 1 heure, 2 heures, 4 heures, 6 heures et 24 heures. Seules les données provenant de la période estivale (juin à septembre) ont été utilisées dans cette analyse. On dispose à cette station de 43 années de données, débutant en 1964 et se terminant en 2006 (1978 est une année manquante). Les statistiques descriptives des maximums saisonniers des différentes durées sont présentées au tableau 1.

Tableau 1. Statistiques descriptives des maximums saisonniers pour différentes durées ($n=43$).

Durée	1 heure	2 heures	4 heures	6 heures	12 heures	24 heures
Moyenne (mm)	20,1	27,0	35,0	40,0	46,0	53,0
Minimum (mm)	10,0	12,7	16,0	19,0	21,0	11,0
Maximum (mm)	70,0	77,0	87,0	89,0	98,0	110,0
Coefficient de variation (Cv)	0,51	0,49	0,43	0,40	0,42	0,39
Coefficient d'asymétrie (Cs)	3,18	2,06	1,50	1,30	1,30	1,50
Coefficient d'aplatissement (Ck)	14,0	6,9	4,90	3,80	3,50	4,50

Par la suite, les tests d'indépendance, de stationnarité et d'homogénéité (voir St-Hilaire, 2005) ont été validés pour les échantillons de différentes durées. Les statistiques des tests sont résumées dans le tableau 2. Fait important à noter, les échantillons de durée inférieure à deux

heures échouent le test de stationnarité (Kendal, 1975) et ceux de durée supérieure à deux heures échouent le test d'indépendance (Wald et Wolfowitz, 1943 ; tableau 2).

Tableau 2. Statistiques des tests pour différentes durées. Les valeurs en gras indiquent les tests échoués.

Durée		1 heure	2 heures	4 heures	6 heures	12 heures	24 heures
Indépendance	U	0,57	0,69	2,1	2,8	2,6	2,6
	p	0,57	0,49	0,04	0,006	0,009	0,009
Stationnarité	K	2,7	1,96	2,0	1,3	0,043	0,64
	p	0,070	0,05	0,04	0,20	0,97	0,52
Homogénéité	W	1,31	1,04	2,2	0,051	1,1	0,72
	p	0,19	0,30	0,031	0,96	0,28	0,47

La stationnarité (pas de tendances dans le temps ou de sauts dans la moyenne) est invalidée pour les événements de courte durée. La figure 1 montre la série chronologique des maximums saisonniers de durée deux heures. On constate que les trois valeurs les plus élevées sont toutes situées vers la fin de la période d'observation, ce qui a pour effet de rendre la série non stationnaire avec une tendance à la hausse.

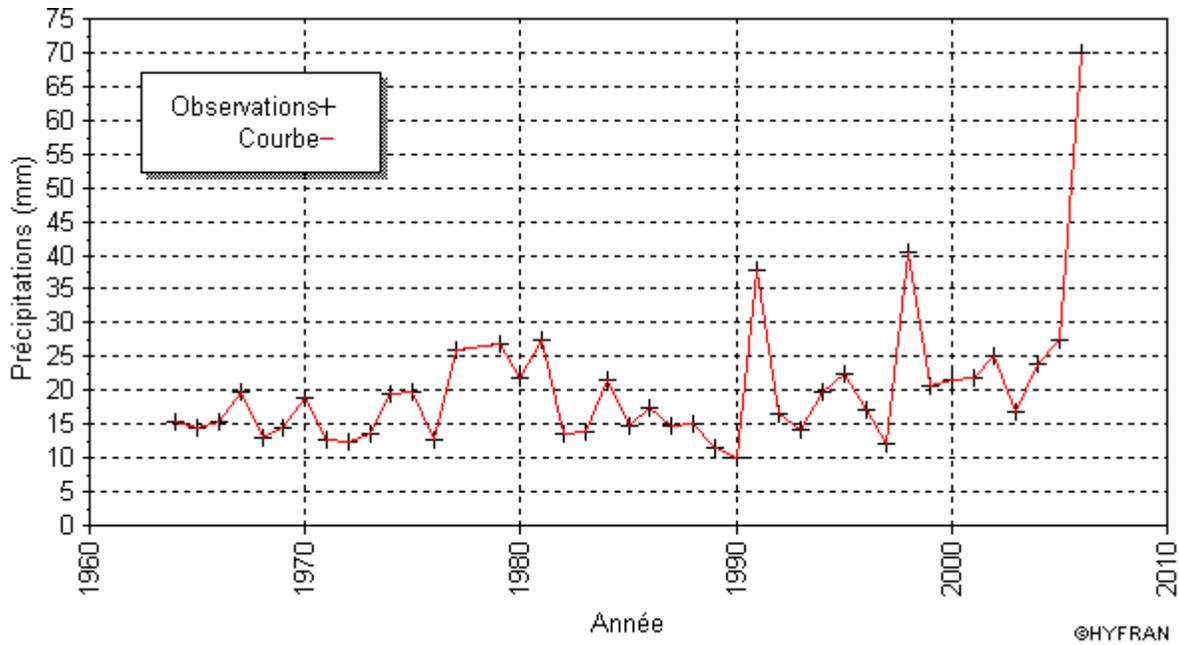


Figure 1. Série chronologique des maximums saisonniers de durée 2 heures.

Les séries de durée supérieure à deux heures sont caractérisées par la présence d'une autocorrélation (corrélation de la mesure présente avec celle(s) du passé) dans la série. À titre d'exemple, la figure 2 montre la fonction d'autocorrélation pour les maximums saisonniers de durée 24 heures. On constate que l'autocorrélation d'ordre 1 (corrélation avec la mesure du pas de temps précédent) est supérieure à 0,3 et est significative. Une analyse fréquentielle ne peut théoriquement être pratiquée sur un échantillon autocorrélé.

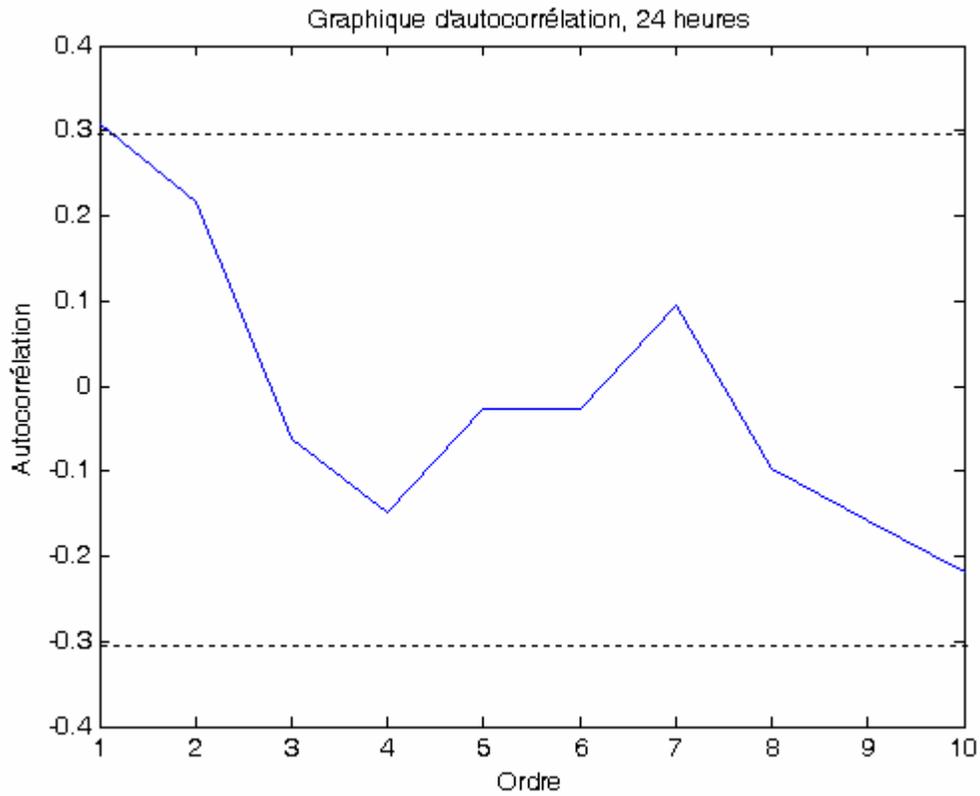


Figure 2. Fonction d'autocorrélation des maximums saisonniers de durée 24 heures.

Compte tenu des nombreux tests d'hypothèse échoués, il ne convient pas d'utiliser la série chronologique de la station Assomption en analyse fréquentielle. L'estimation des quantiles à quand même été réalisée à titre indicatif, mais elle ne devrait pas être sujette à interprétation.

3. RÉSULTATS

3.1 Analyse fréquentielle des pluies journalières

Trois lois statistiques utilisées fréquemment en analyse fréquentielle ont été ajustées aux échantillons. Une loi à deux paramètres, la loi lognormale a été ajustée et comparée à deux lois à trois paramètres. Il s'agit de la distribution GEV (« Generalized Extreme Value ») et de la loi log-normale III. Différentes méthodes existent pour ajuster les distributions statistiques aux échantillons. Les trois lois utilisées ont été ajustées par la méthode du maximum de vraisemblance (Bobée et Ashkar, 1991).

La sélection de la distribution statistique la mieux ajustée aux échantillons est faite à l'aide de deux critères, soit le critère d'Akaike et le critère d'information bayésien (BIC). Ces deux critères permettent de choisir la loi la mieux ajustée en tenant compte de l'erreur d'estimation et de la parcimonie (nombre de paramètres à ajuster). La distribution pour laquelle les valeurs des deux critères sont les plus faibles est celle qui est sélectionnée. Le tableau 2 montre les résultats de cette analyse comparative de l'ajustement des lois statistiques. On y constate que les valeurs des deux critères sont assez approchées pour les trois lois. De plus, dans certains cas, les deux critères ne choisissent pas nécessairement la même loi (e.g. durée 12 heures). Les lois GEV et lognormales à deux paramètres ont été sélectionnées à deux reprises sur la base de ces critères.

Tableau 3. Critères de sélection des distributions statistiques ajustées aux données.

Durée	Loi statistique	Akaike	BIC
1	Log Normale	289,1	285,6
	Log Normale III	284,1	278,9
	GEV	283,5	278,3
2	Log Normale	317,8	314,3
	Log Normale III	315,9	310,8
	GEV	315,7	310,5
4	Log Normale	337,7	334,2
	Log Normale III	339,4	334,2
	GEV	339,7	334,5
6	Log Normale	344,9	341,5
	Log Normale III	346,3	341,1
	GEV	346,7	341,5
12	Log Normale	359,3	355,8
	Log Normale III	359,7	354,4
	GEV	359,7	354,5
24	Log Normale	362,6	359,1
	Log Normale III	360,3	355,1
	GEV	360,6	355,3

Une comparaison visuelle des ajustements aide à procéder à un choix final. La figure 3 compare l'ajustement des maximums saisonniers de durée 4 heures pour les trois lois. On constate que les valeurs plus élevées de précipitation totales sont mieux ajustées par les lois à trois paramètres.

Pour cette raison, la loi GEV a été sélectionnée pour construire les courbes d'intensité-durée-fréquence.

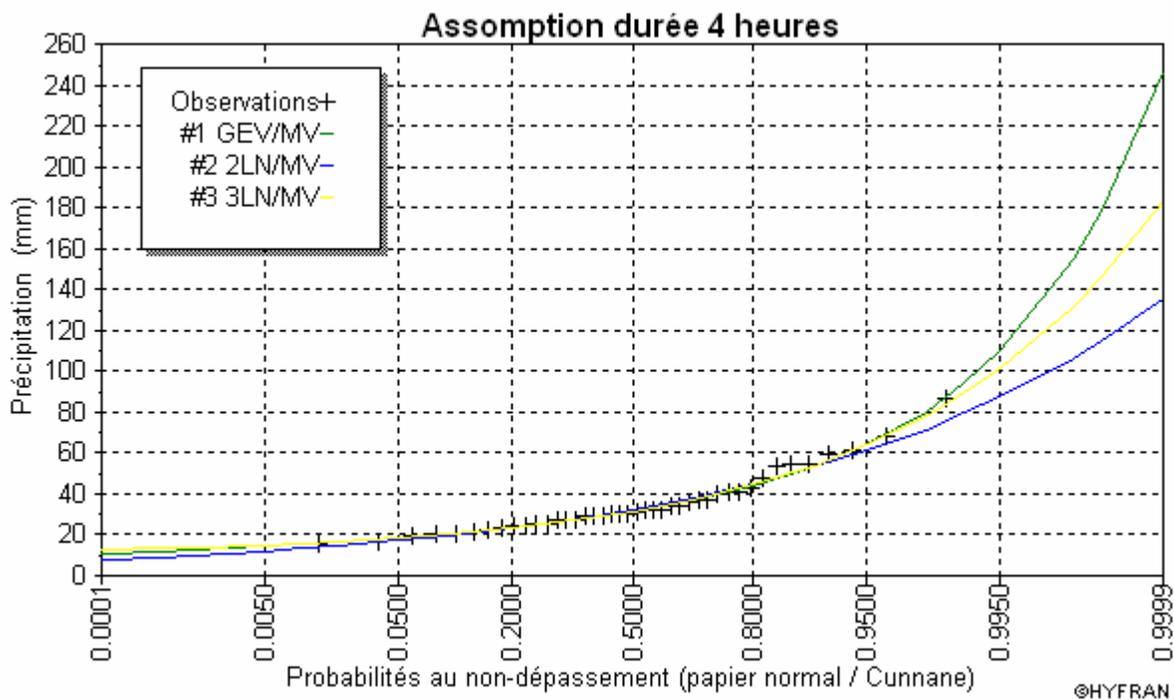


Figure 3. Comparaison de l'ajustement des maximums saisonniers de durée 4 heures.

3.2 Courbes Intensité-Durée-Fréquence

La loi GEV a donc été ajustée à chacune des séries de maximums saisonniers de différentes durées. La figure 4 montre la courbe intensité-durée-fréquence qui en découle. L'événement qui s'est produit le 19 août 2006 peut être comparé à ces courbes. La précipitation total horaire maximum enregistrée ce jour là est de 70 mm. En transposant cette valeur sur la courbe de durée 1 heure, on obtient une période de retour de l'ordre de 140 ans. Si on ajoute les précipitations mesurées durant les 5 heures suivantes, on obtient un total de 85 mm pour une durée 6 heures. Un tel événement aurait une récurrence inférieure à 50 ans.

Courbes Intensité-Durée-Fréquence

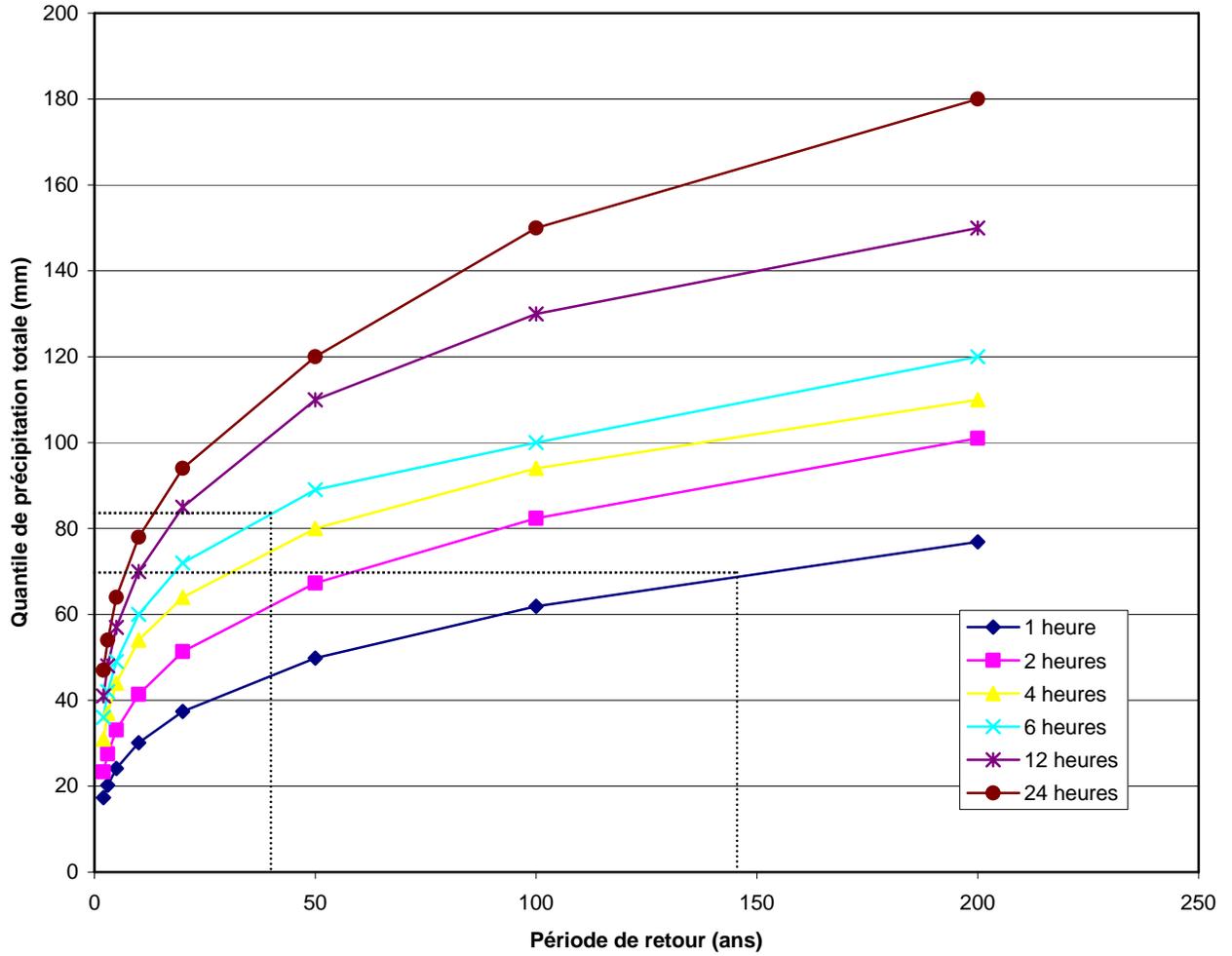


Figure 4. Courbe Intensité-Durée-Fréquence pour la station Assomption.

7. CONCLUSION

L'analyse fréquentielle locale des données de précipitation de la station Assomption a été réalisée à partir de données horaires couvrant la période 1964-2006. Les hauteurs de pluie cumulatives pour les durées de 1, 2, 4, 6, 12 et 24 heures ont été calculées et les maximums saisonniers (juin à septembre) de chaque année ont été extraits. Un problème a été rencontré en ce qui a trait au respect des hypothèses de stationnarité et d'indépendance. En effet, certaines séries de maximums saisonniers ont échoué ces tests.

Les quantiles ont quand même été estimés à titre indicateur. On y remarque que la période de retour de la hauteur de pluie horaire maximale du 19 août 2006 est relativement élevée (140 ans). Par contre, les précipitations totales de cet événement, qui se sont échelonnées sur 6 heures, ont eu une récurrence plus faible (< 50 ans).

Il faut cependant se rappeler que ces conclusions ne peuvent théoriquement être utilisées dans la mesure où les hypothèses de base de l'analyse fréquentielle n'ont pas été respectées. Il est recommandé de procéder à des analyses supplémentaires sur des stations avoisinantes.

8. RÉFÉRENCES

Bobée, B., F. Ashkar, 1991. The gamma family and derived distributions applied in hydrology. *Water Resources Publications* 203 pages.

HYFRAN. 2003. A software for statistical modeling. INRS-Eau, University of Québec.

Kendall, M.G., 1975. Rank Correlation Methods. *Charles Griffin*, London.

St-Hilaire A. 2005. Analyse fréquentielle des précipitations des stations de Montréal et Shawinigan. Rapport de recherche R832 de l'INRS-ETE. v + 23 pages.

Wald, A., J. Wolfowitz, 1943. An exact test for randomness in the nonparametric case based on serial correlation. *Annals of Mathematical Statistics.*, **14**, 378-388.

Wilcoxon, F., 1945. Individual comparisons by ranking methods. *Biometrics.*, **1**, 80-83