

UNIVERSITÉ DU QUÉBEC

Mémoire présenté à

l'INSTITUT NATIONAL DE LA RECHERCHE SCIENTIFIQUE

comme exigence partielle
de la
maîtrise ès sciences (eau)

par

David Berryman

B. Sc. Biologie

La détection des tendances dans les séries temporelles de paramètres
de la qualité de l'eau à l'aide de tests non paramétriques

Août 1984

À Edward et Irène Berryman

J'ai commencé à m'intéresser aux nombres pour me sentir moins seul. À quatorze ans, je passais des nuits blanches à compter jusqu'à des millions, dans l'espoir de rencontrer quelqu'un dans le tas. J'ai fini dans les statistiques.

Romain Gary dans Gros-Câlin

REMERCIEMENTS

Mes remerciements vont à:

Bernard Bobée (INRS-Eau), directeur de mémoire, qui a proposé le sujet et dirigé cette recherche;

Daniel Cluis (INRS-Eau), conseiller, pour l'éclairage sur la notion d'autocorrélation et pour l'aide apportée à la réalisation du quatrième chapitre.

John Haemmerli (Environnement Canada), Michel Goulet et Paul Potvin (Environnement Québec) pour l'intérêt manifesté à l'égard de ce sujet.

Anne Provencher et Johanne Parrot pour la dactylographie.

André Parent pour le graphisme.

RÉSUMÉ

En étudiant les caractéristiques des séries temporelles de paramètres de la qualité de l'eau et l'ensemble des méthodes mathématiques et statistiques utilisables pour analyser ces séries, on est amené à choisir les tests non paramétriques pour le développement d'un outil de détection de tendances. Une revue des tests non paramétriques pour détecter les tendances permet d'identifier les tests de Mann-Whitney, Spearman et Kendall comme étant les plus performants. Récemment, ces mêmes tests ont été adaptés pour tenir compte des problèmes de persistance et de périodicité rencontrés dans les séries de qualité de l'eau. Pour les tendances monotones, une procédure permettant de choisir un test en fonction des caractéristiques des séries et des limitations des tests, est proposée. Une adaptation récente du test de Kendall, est étendue au test de Mann-Whitney pour tenir compte de la périodicité dans les séries présentant un saut. Pour ce type de tendances, une procédure analogue à celle utilisée pour les tendances monotones est proposée. Des vérifications de robustesse et de puissance restent à effectuer avant que cette procédure puisse être utilisée de manière systématique.

SUMMARY

After considering the characteristics of water quality time series and the available statistical and mathematical methods to analyse these series, non parametric tests are chosen for the development of a trend detection procedure. A review of non parametric tests for trend leads to the evidence that Mann-Whitney, Spearman and Kendall tests are the best fit for trend detection. Recently these same tests have been adapted to account for dependence and seasonality: common features of water quality time series. For monotone trends a procedure allowing to select the pertinent test considering the time series characteristics and the practical limitations of the tests is proposed. A recent adaptation of Kendall's test is extended to Mann-Whitney's to account for seasonality in time series with a step trend. A procedure, similar to the one used for monotone trends, is proposed for step trends. Robustness and power remain to be checked before this last procedure can be used on a general basis.

TABLE DES MATIÈRES

	Page
Remerciements	i
Résumé	ii
Summary	iii
Table des matières	iv
Liste des tableaux	viii
Liste des figures	x
 1. Introduction	 1
1.1 Problématique et objectif	1
1.2 Cadre scientifique	3
1.3 Contenu	8
 2. Revue, classification et comparaison des tests de détection des tendances	 11
2.1 Introduction	11
2.2 Définitions	13
2.2.1 Hypothèses, puissance, efficacité asymptotique relative des tests statistiques	 13
2.2.2 Les séries	16
2.3 Les tendances monotones	18
2.3.1 La définition et les bases de la détection d'une tendance monotone	 18
a) Hypothèses de base	21
2.3.2 Test du nombre de séquences homogènes et test du nombre de points critiques	 22
a) Les hypothèses statistiques	22
b) La statistique	25
c) Les autres caractéristiques	27
2.3.3 Les tests de Irwin-Fischer, Foster-Stuart, des différences successives de Cox-Stuart, Mann et Lehman	 28
a) Les hypothèses statistiques	28
b) La statistique	28
c) Les autres caractéristiques	30

2.3.4	Les tests de Kendall et de Spearman	31
	a) Les hypothèses statistiques	31
	b) La statistique	32
	c) Les autres caractéristiques	33
2.3.5	Discussion	34
2.4	Les tendances en saut	36
2.4.1	La définition et les bases de la détection d'une tendance en saut	36
	a) Les hypothèses de base	37
2.4.2	Les tests de Wald-Wolfowitz et de Kolmogorov- Smirnov	38
	a) Les hypothèses statistiques	38
	b) La statistique	39
	c) Les autres caractéristiques	49
2.4.3	Les tests de la médiane et de Mann-Whitney	42
	a) Les hypothèses statistiques	42
	b) La statistique	42
	c) Les autres caractéristiques	42
2.4.4	Les tests basés sur les variables normalisées: Terry-Hoeffding, Van der Waerden, Bell-Doksum .	44
	a) Les hypothèses statistiques	44
	b) La statistique	45
	c) Les autres caractéristiques	45
2.4.5	Discussion	46
2.5	Tendances en escalier	48
2.5.1	La définition et les bases de la détection d'une tendance en escalier	48
2.5.2	Les tests de Kruskal-Wallis, Terpstra- Jonckheere et Puri	49
	a) Les hypothèses statistiques	49
	b) La statistique	50
	c) Les autres caractéristiques	50
2.5.3	Discussion	50

3. Application des tests aux séries temporelles	55
3.1 Exigences théoriques des tests : les hypothèses de base	56
3.1.1 Hypothèse de l'égalité de la distribution des erreurs	58
3.1.2 Hypothèses de l'indépendance des erreurs	61
3.2 Caractères pratiques des séries observées : la persistance et la périodicité	66
3.2.1 La persistance	66
3.2.2 La périodicité	68
3.2.3 La persistance et la périodicité	69
3.3 Adaptation et application des tests aux séries dans le cas des tendances monotones	71
3.3.1 Séries avec persistance	71
a) Résumé de la méthode (Lettenmaier, 1976) ..	71
b) Commentaires	73
3.3.2 Séries avec périodicité	74
a) Résumé de la méthode (Hirsch <u>et al.</u> , 1982) .	74
b) Commentaires	82
3.3.3 Séries avec persistance et périodicité	84
a) Résumé de la méthode (Hirsch et Slack, 1983)	84
b) Commentaires	91
3.4 Adaptation et application des tests aux séries dans le cas des tendances en saut	92
3.4.1 Séries avec persistance	93
a) Résumé de la méthode (Lettenmaier, 1976) ..	93
b) Commentaires	94
3.4.2 Séries avec périodicité	94
a) Adaptation de la méthode de Hirsch <u>et al.</u> (1982)	94
b) Commentaires	97

3.4.3	Séries avec persistance et périodicité	98
a)	Adaptation de la méthode de Hirsch et Slack, (1983)	98
b)	Commentaires	99
3.5	Compléments	100
3.5.1	Cas à plus d'une observation par période	100
3.5.2	Effet d'une variable externe	102
a)	Fonction double-masse	103
b)	Calcul des résidus	103
4.	Procédure de détection des tendances monotones	107
4.1	Présentation	107
4.2	Description	110
4.2.1	Opérations de base	110
4.2.2	Décisions d'orientation	111
4.2.3	Traitement; séries avec périodicité.....	115
4.2.4	Traitement; séries avec périodicité et dépendance	115
4.2.5	Traitement; séries avec dépendance	116
4.2.6	Traitement; séries sans dépendance ni périodicité	117
4.3	Discussion	117
4.3.1	Les boîtes 13, 22, 23, 24, 25 et 26	117
4.3.2	Simplification de la procédure	120
5.	Conclusion	123
	Bibliographie	125
Annexe 1	Fonctions Cusum	1.1
Annexe 2	Test de normalité (Lilliefors, 1967)	2.1
Annexe 3	N et K minimaux pour l'approximation normale du test de Mann-Whitney	3.1
Annexe 4	Test d'égalité des variances	4.1
Annexe 5	Estimé de la pente de la tendance	5.1

LISTE DES TABLEAUX

	Page
1.1 Critères de sélection de l'outil de détection et caractéristiques des séries temporelles	4
1.2 Critères de sélection non respectés et caractéristiques des séries non adéquates aux approches mathématiques	4
2.1 Classification des tests de tendance monotone	23
2.2 Caractéristiques des tests de tendance monotone	24
2.3 EAR du test de Kendall par rapport au test paramétrique (Pearson)	33
2.4 Classification des tests de tendance en saut	39
2.5 Caractéristiques des tests de tendance en saut	41
2.6 EAR du test de Mann-Whitney par rapport au test de Student	43
2.7 Classification des tests de tendance en escalier	51
2.8 Caractéristiques des tests de tendance en escalier	52
3.1 Données mensuelles fictives sur six ans	76
3.2 Valeurs de la statistique de Kendall pour chaque mois et le total	78
3.3 Variance des S_k	80
3.4 Données brutes et leur rang, mois 1 et 2	87
A3.1 Permutation des rangs et statistique de Mann-Whitney pour une série de quatre observations	3.9
A3.2 Distribution de U pour une sous-série de 4 observations dont 2 avant le saut	3.9
A3.3 Calcul de la distribution exacte de la statistique totale U	3.10
A3.4 Distribution théorique (approximation normale de U)	3.11
A3.5 Distribution exacte ($P(u)$) et théorique ($G(u)$) pour $N = 4$ et $K = 4$	3.12

A3.6	Distribution exacte ($P(u)$) et théorique ($G(u)$) pour $N = 2$ et $K = 4$	3.13
A3.7	Effet de K sur l'ajustement, $N = 4$	3.14
A3.8	Effet de N sur l'ajustement, $K = 2$	3.14
A3.9	Erreur maximale, $N = 2$ et $K = 12$	3.14

LISTE DES FIGURES

	Page
2.1 Exemples de figures monotones et non-monotones	19
3.1 Graphe des résidus en fonction du temps	59
3.2 Valeurs de r_1 significatives aux seuils $\alpha = 0,05$ et $0,01$, en fonction de N	63
3.3 Corrélogramme d'un processus autorégressif, Markovien d'ordre 1	67
3.4 Corrélogramme d'un processus autorégressif avec périodicité	70
3.5 Puissance du test de Spearman en fonction de N't au seuil $\alpha = 0,05$	73
3.6 Graphique des données mensuelles fictives	76
3.7 Puissance du test de Mann-Whitney en fonction de Nt à $\alpha = 0,05$	94
4.1 Système de détection des tendances monotones	
A1.1 Évolution de la fonction cusum pour différentes valeurs de M dans le cas d'un changement brusque	1.2
A1.2 Évolution de la fonction cusum pour différentes valeurs de M dans le cas d'une tendance continue décroissante	1.3
A3.1 Distribution exacte de $(P(u))$ et théorique $(G(u))$, $K = 2$, $N = 4$	3.11
A3.2 Distribution exacte de $(P(u))$ et théorique $(G(u))$, $K = 4$, $N = 4$	3.12
A3.3 Distribution exacte de $(P(u))$ et théorique $(G(u))$, $K = 12$, $N = 2$	3.13

1. INTRODUCTION

1.1 Problématique et objectifs

La détérioration progressive de la qualité des eaux a amené, il y a quelques années déjà, l'injection de fonds publics dans l'implantation de réseaux de mesure et de surveillance. À partir de mesures de différents paramètres, le gestionnaire peut connaître l'état de cette ressource et proposer des mesures correctrices si cet état n'est pas jugé satisfaisant. Dans un tel cas, les actions entreprises conduiront à une modification de l'état de la ressource. Il sera alors nécessaire d'identifier les effets de l'action entreprise, peut-être de réajuster l'action, revérifier et ainsi de suite: un processus itératif qui peut être long et onéreux.

Le gestionnaire peut aussi utiliser ces mesures pour prédire l'état de la ressource à différents horizons temporels. Cette approche lui permet une action fondamentalement différente: il remplace l'action corrective par des mesures préventives avant que cette ressource n'atteigne le niveau de dégradation jugé insatisfaisant (Sherwani, 1975). Cette approche plus souple s'intégrera beaucoup plus facilement dans une politique d'ensemble de gestion de l'eau.

Ce constat nous amène à reconnaître la nécessité d'un outil permettant:

- de détecter les éventuels changements (tendances) de paramètres représentatifs de la qualité de l'eau;
- de prédire l'état de la ressource à différents horizons temporels.

Les méthodes proposées doivent être objectives, c'est-à-dire exemptes de la subjectivité du gestionnaire; et quantitatives de façon à permettre un dosage des mesures préventives. Les méthodes proposées doivent être considérées comme un support objectif pour le gestionnaire qui doit, avec son expérience et son jugement, rester critique dans l'interprétation des résultats.

Des rencontres, en février 1983, avec les principaux intervenants dans la surveillance de la qualité des eaux au Québec ont:

- confirmé la pertinence et la nécessité de telles méthodes objectives;
- permis de mieux saisir la réalité quotidienne de ces gestionnaires. Ceci nous a amené à définir trois critères de l'outil à développer: facilité et rapidité d'utilisation, simplicité des fondements théoriques.

En effet, les méthodes proposées devront être adaptées à des utilisateurs qui:

- ont des connaissances de base en statistique sans pour autant être des spécialistes (facilité d'utilisation);
- ne peuvent consacrer qu'une partie de leur temps et de leurs ressources à la recherche de tendances (rapidité d'utilisation);
- recherchent des méthodes dont ils comprennent bien les limites et les bases théoriques de façon à les utiliser avec souplesse, adaptabilité et sens critique (simplicité des fondements théoriques).

De cet exposé de la problématique, un premier objectif se dessine: développer un outil, c'est-à-dire un ensemble de méthodes permettant de détecter les tendances dans les séries temporelles de la qualité de l'eau. Cet outil doit donc être OBJECTIF, QUANTITATIF, D'UTILISATION SIMPLE ET RAPIDE et SIMPLE DANS SES FONDEMENTS THÉORIQUES.

Un deuxième objectif, en dehors du champ de cette recherche, concerne la prédiction de l'état de la ressource par modélisation de la tendance détectée.

1.2 Cadre scientifique

L'objectif et les critères de choix étant maintenant précisés, il est nécessaire d'identifier l'ensemble des méthodes statistiques et mathématiques susceptibles de constituer l'outil recherché.

Tableau 1.1: Critères de sélection de l'outil et caractéristiques des séries (d'après Sherwani, 1975; Lettenmaier, 1976; Hirsch et al., 1982)

Critères de l'outil	Caractéristiques des séries généralement observées en qualité de l'eau
1. Quantitatif 2. Objectif 3. Simple d'utilisation 4. Rapide d'utilisation 5. Simple dans les bases théoriques	A. Données généralement mutuellement dépendantes B. Données ne provenant pas d'une loi normale C. Données pouvant ne pas provenir d'une même loi de distribution D. Données éventuellement manquantes (pas à intervalles réguliers) E. Données parfois non quantifiées ou tronquées (ex: sous le seuil de détection)

Tableau 1.2: Critères de sélection non respectés et caractéristiques des séries non adéquates aux approches mathématiques

Approche mathématique	Critère non respectée	Caractéristique non respectée
- méthodes en série de temps	3 - 4 - 5	D - E
- méthodes graphiques (double masse, CUSUM)	1 - 2	A - C - D
- tests paramétriques		A - B - E
- tests non-paramétriques		A - C

Le tableau 1.1 reprend les critères de sélection de l'outil recherché et énumère les caractéristiques des séries temporelles généralement observées pour les paramètres de la qualité de l'eau. En plus de répondre aux cinq critères qui lui sont propres, la procédure à développer devra être applicable aux séries possédant les caractéristiques du tableau 1.1.

En effet, les séries temporelles des paramètres de la qualité de l'eau sont souvent constituées de données:

- corrélées entre elles: Par exemple, la température d'un lac mesurée aujourd'hui ne sera pas indépendante de celle mesurée hier. Par contre, elle pourrait être à peu près indépendante de celle mesurée il y a deux mois. Le temps séparant les observations aura donc beaucoup d'influence sur la corrélation (dépendance) entre les observations.
- qui ne proviennent pas d'une loi normale: La distribution des fréquences des différentes valeurs d'un paramètre de la qualité autour de la moyenne peut être asymétrique, surtout si ce paramètre est relié au débit.
- qui ne proviennent pas d'une même loi de distribution: Le débit est une des variables déterministes les plus importantes pour la qualité de l'eau en rivière. Ainsi, les mesures de qualité mesurée en crue

et en étiage peuvent constituer des populations statistiquement différentes.

- manquantes: Du travail de terrain à l'analyse des résultats, un grand nombre de facteurs (climat, transport, erreur au laboratoire) peuvent éliminer des mesures. Il en résulte une série où le pas de temps entre les observations n'est pas constant.
- non quantifiées ou tronquées: Parfois, comme dans le cas des métaux traces ou des pesticides, la valeur d'un paramètre peut être en deçà de la capacité de détection de l'appareil de mesure. La valeur peut alors être exprimée en termes tels que "inférieure à ...".

Ces caractéristiques peuvent empêcher ou limiter l'utilisation de certaines méthodes en détection de tendances. Par exemple:

- de façon rigoureuse, les tests paramétriques ne peuvent être utilisés si les données sans tendance ne sont pas indépendantes ou si elles ne proviennent pas d'une population suivant une loi normale (Alvo, 1981);
- les méthodes en série de temps sont "techniquement très difficiles à utiliser" lorsque le pas de temps entre les observations n'est pas régulier (Alvo, 1981).

Pour résumer cette problématique le tableau 1.2 présente les approches mathématiques et statistiques permettant l'analyse des séries

temporelles. Il présente aussi les critères de l'outil auquel ces approches ne répondent pas ou peu, de même que les caractéristiques de séries qui peuvent limiter ou empêcher l'emploi d'une méthode.

Commentaires du tableau 1.2

Les méthodes en série de temps sont basées sur la modélisation des composantes (tendance, cycles) de la série. Ce sont des méthodes raffinées mais complexes, qui demandent que:

- l'utilisateur ait une bonne expérience de l'analyse des séries temporelles;
- les données soient prises à intervalles réguliers et qu'elles soient quantifiées (pas sous un seuil de détection).

Les méthodes graphiques ne sont pas quantitatives et objectives car la tendance est décelée par un examen visuel d'un graphique et non par l'application d'une règle de décision objective comme dans le cas des tests. Elles ne peuvent être utilisées avec les données non quantifiées. Cependant, ces techniques sont simples et rapides. Une méthode utilisant les fonctions CUSUM est présentée à l'annexe 1. Les fonctions double-masse et CUSUM ont déjà été utilisées en qualité de l'eau (Cluis, 1981).

Les tests paramétriques exigent que les données soient indépendantes, quantifiées et issues d'une population normale.

Comme les précédents, les tests non-paramétriques sont conformes aux critères de sélection de l'outil. Par contre, il sont moins exigeants pour les caractéristiques des séries. Les données n'ont pas besoin de parvenir d'une distribution normale et des données non quantifiées (sous le seuil de détection par exemple) peuvent être utilisées.

Cette analyse nous amène à choisir les tests non-paramétriques comme champ d'investigation pour les méthodes devant constituer l'outil de détection des tendances.

1.3 Contenu

L'objectif formulé (recherche d'un outil de détection) et le champ de recherche précisé (tests non-paramétriques), l'étape suivante est le survol de ce champ pour y recenser les méthodes applicables à la détection de tendances. Les résultats de cette revue font l'objet du chapitre 2. Les tests applicables à la détection des tendances y sont classifiées, leur efficacité est discutée, et les tests les plus performants sont retenus. La lecture de ce chapitre n'est pas nécessaire à la compréhension du reste du texte, sauf pour les définitions en 2.2.

À l'intérieur d'une série de données en qualité de l'eau la valeur de chaque donnée peut être composée de l'addition de plusieurs éléments:

$$y_i = T_i + C_i + e_i$$

- 1- la moyenne et sa tendance (T_i);
- 2- les cycles (C_i);
- 3- la composante aléatoire (e_i), (Sherwani, 1975).

La tendance, les cycles de même que la persistance naturelle du paramètre dans le temps (ex. de la température du lac) créent de la corrélation ou dépendance entre les observations d'une série. Tous les tests de détection exigent qu'entre les observations il n'y ait pas de dépendance autre que celle due à la tendance.

Lettenmaier en 1976, Hirsch et al. en 1982 et Hirsch et Slack en 1984 ont choisi les mêmes tests que ceux retenus au chapitre 2 et les ont adaptés pour tenir compte de la dépendance due à la persistance et à l'effet des cycles. Ces adaptations sont résumées au troisième chapitre. Une méthode permettant de détecter une tendance par saut dans une série à composante périodique a été développée à partir des résultats de Hirsch.

Toutes ces méthodes ont leurs limites d'application, leurs exigences face aux séries. L'approche consiste donc en:

- un examen préliminaire de la série;
- l'examen des caractéristiques des séries permettant de vérifier si les exigences des méthodes sont respectées;

- le choix d'une méthode et son application sur la série.

Ces trois points constituent l'essentiel de l'outil de détection des tendances progressives (monotones) présenté au dernier chapitre.

2. REVUE, CLASSIFICATION ET COMPARAISON DES TESTS DE DÉTECTION DES TENDANCES

2.1 Introduction

Il existe de nombreux tests applicables à la détection des tendances. Plusieurs de ces tests ont des points communs, certains sont équivalents. On peut trouver des tests différents connus sous le même nom, et le même test peut avoir des noms différents. Il n'existe pas de synthèse portant sur l'ensemble de ces tests. Un effort de comparaison et de classification est donc nécessaire. C'est le résultat de ce travail de synthèse qui est présenté ici.

Une revue des articles relatifs à l'application des tests à la détection de tendance a permis d'identifier plus de 20 tests applicables. Ce chapitre a comme objectif de comparer ces tests de façon à permettre une classification, puis une pré-sélection des meilleurs tests de détection de tendance. Cette sélection est faite à partir des qualités intrinsèques des tests.

Les tests sont classés en fonction de plusieurs critères dont:

- le type de tendance pour lequel le test s'applique;
- les hypothèses statistiques que l'on veut examiner par le test (H_0 et H_1);

- la statistique sur laquelle le test se base. On identifie le principe du test et la formulation mathématique de la statistique.

Les résultats de cette classification sont résumés aux tableaux 2.1 pour les tendances monotones, 2.4 pour les tendances en saut et 2.7 pour les tendances en escalier. Cette classification des tests est complétée par une description (tableaux 2.2, 2.5 et 2.8) portant sur:

- les références pour la détermination des valeurs critiques dans le cas des petits échantillons;
- la formule d'approximation par une loi statistique connue dans le cas des grands échantillons;
- la correction pour la continuité;
- la correction pour le cas où l'on dispose de plusieurs observations égales;
- la puissance des tests exprimée lorsque cela est possible par l'efficacité asymptotique relative.

Même si dans les tableaux la classification et la description sont des thèmes séparés, il est impossible de classer sans décrire, de décrire sans discuter. C'est pourquoi dans le texte les deux thèmes se

mêlant, se recoupent. Le texte assure la liaison entre les tableaux et en précise le contenu en mettant l'accent sur les différences et les points communs entre les tests, ainsi que sur l'évolution de l'un à l'autre. La description sera plus exhaustive pour certains tests, moins pour d'autres en fonction de leur intérêt pratique pour la détection des tendances. Un examen des tableaux et une lecture du texte en parallèle plutôt que successive facilitera la compréhension de l'ensemble. La partie 2.2 rappelle les définitions des termes utilisés dans les tests statistiques.

2.2 Définitions

2.2.1 Hypothèses, puissance, efficacité asymptotique relative des tests statistiques

Sans reprendre la théorie des tests qui est décrite de manière détaillée par de nombreux auteurs (par exemple, Baillargeon et Rainville, 1977), rappelons que les tests permettent d'infirmer ou non une hypothèse nulle et que le rejet (non-rejet) de l'hypothèse nulle implique l'acceptation (non-acceptation) d'une hypothèse alternative.

Par exemple :

$$H_0 : \mu = 0 \quad (\text{hypothèse nulle})$$

$$H_1 : \mu \neq 0 \quad (\text{hypothèse alternative})$$

ou encore:

$$H_0 : \mu \geq 0 \quad (\text{hypothèse nulle})$$

$$H_1 : \mu < 0 \quad (\text{hypothèse alternative})$$

Tirer une conclusion sur une population à partir d'un échantillon, implique un risque d'erreur dû au caractère aléatoire de cet échantillon. Il y a deux types d'erreurs possibles:

- Erreur de type I (de probabilité α): on rejette l'hypothèse nulle alors qu'elle est vraie;
- Erreur de type II (de probabilité β): on ne rejette pas l'hypothèse nulle (H_0) alors qu'elle est fausse et que H_1 est vraie (ou on rejette H_1 alors qu'elle est vraie)

Lors de l'application d'un test, la probabilité α est habituellement fixée à 1 % ou 5 %, c'est le niveau ou seuil de signification du test.

On appelle puissance, la quantité $(1 - \beta)$ qui est en fait la proportion de cas où l'on va rejeter l'hypothèse nulle alors qu'elle est fausse (ou accepter H_1 lorsqu'elle est vraie).

La mesure de l'efficacité relative locale d'un test A par rapport à un test B est donnée par :

$$E_{AB} = \frac{N_B}{N_A}$$

où

N_A = nombre d'observations nécessaires au test A;

N_B = nombre d'observations nécessaires au test B.

pour que les deux tests aient la même puissance. Elle est dite relative car c'est la mesure d'un test (A) par rapport à un autre (B). Elle est dite locale car cette mesure est spécifique à un seuil α et a une valeur spécifique du paramètre mesuré.

On définit aussi l'efficacité asymptotique relative :

$$EAR = \limite \left(\frac{N_B}{N_A} \right), \text{ lorsque } N_A \rightarrow \infty$$

Contrairement à l'efficacité, cette quantité ne dépend pas de α et de la valeur de la variable. C'est ce qui en fait une bonne base de comparaison de puissance entre les deux tests.

Ajoutons que l'EAR ne dépend que des populations d'origine des échantillons, qu'elle peut être déterminée de façon théorique et que dans les cas où les comparaisons ont été faites, les efficacités locales correspondaient étroitement à l'efficacité asymptotique (Marascuilo, 1977).

2.2.2 Les séries

Les variables utilisées dans la formulation mathématique des séries se réfèrent aux définitions suivantes:

Soient les observations: $y_1, \dots, y_i, \dots, y_N$

mesurées respectivement aux temps: $x_1, \dots, x_i, \dots, x_N$

L'indice $i = 1, \dots, N$ représente l'ordre chronologique des observations. C'est aussi le rang de la variable temps (x).

$y_1, \dots, y_i, \dots, y_N$ représentent les rangs des observations de la variable dépendante (y_i) classée par ordre croissant. On donne le rang 1 à la plus petite valeur et le rang N à la plus grande.

Exemple

Soit une série de mesures quotidiennes de pH durant 10 jours :

	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	y_{10}
y_i	5,6	5,8	5,4	6,0	7,0	7,0	6,3	6,8	7,5	8,0
x_i et i	1	2	3	4	5	6	7	8	9	10
$y_i =$	2	3	1	4	7,5	7,5	5	6	9	10

Les deux valeurs de $pH = 7,0$ occuperaient les rangs sept et huit.
 Dans un tel cas, on leur donne le rang moyen calculé par:

$$\frac{\sum T_i}{t_i}$$

où:

$\sum T_i$: somme des rangs des valeurs égales dans le cas d'égalité j ;

t_i : nombre de valeurs égales dans le cas d'égalité j .

Dans notre cas, on a :

$$\frac{7 + 8}{2} = 7,5$$

- positive et peut être nulle sur une partie du domaine,

ou

- négative et peut être nulle sur une partie du domaine.

Des exemples de ces fonctions sont illustrés à la figure 2.1

Exemple :

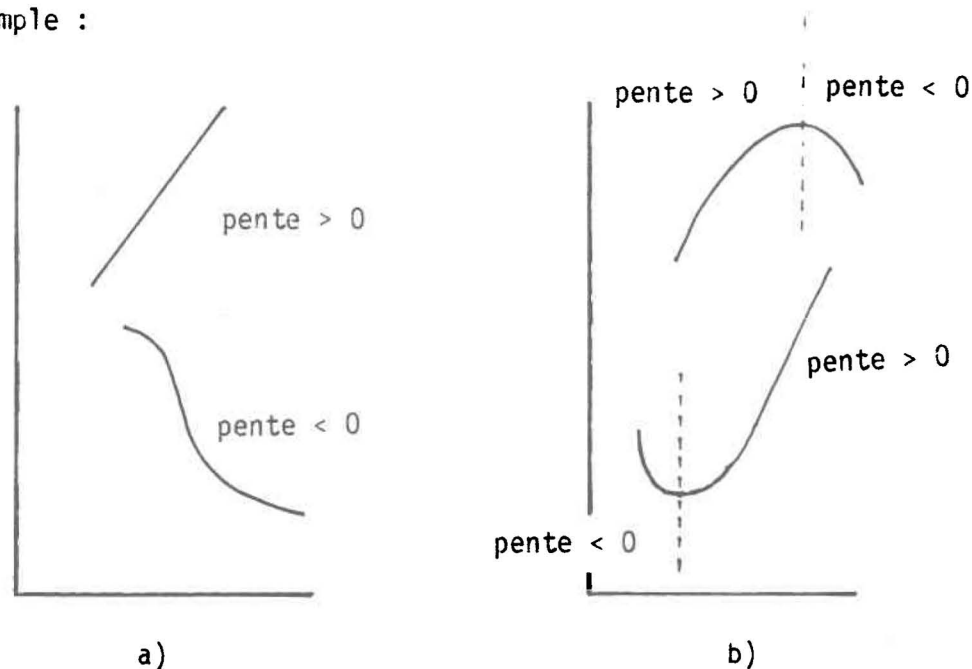


Fig. 2.1 : Exemples de fonctions monotones (a) et non-monotones (b)

La fonction décrivant la série ayant une tendance monotone peut être définie par :

$$y_i = u_i + \varepsilon_i \quad \text{où} \quad \varepsilon_i \text{ est une variable aléatoire indépendante distribuée normalement avec une}$$

moyenne 0 et une variance σ^2

et où u_i est la variable de la tendance

Il y a 2 façons de détecter la tendance, c'est-à-dire prouver que u_i est une constante.

- 1) On peut se servir de tests qui vérifient si la série des y_i est aléatoire, c'est-à-dire si elle ne dépend que des ε_i . Ceci revient à prouver l'absence de tendance, c'est-à-dire à l'examen de l'hypothèse $u_i = C$. C'est sur ce principe que se basent les six premiers tests du tableau 2.1.

- 2) Pour une tendance linéaire sur un domaine des x_i , on aura:

$$y_i = a + b x_i + \varepsilon_i$$

où a est le niveau de base au début du domaine

et b est la pente de la tendance dans ce domaine.

Le test peut alors porter directement ou indirectement sur b . Si $b \neq 0$ il y a tendance, il y a une relation entre y_i (paramètre dépendant) et x_i (le temps). Pour vérifier si $b \neq 0$, on peut donc vérifier si les y_i dépendent des x_i . C'est ce que font les quatre derniers tests du tableau 2.1. Ces tests peuvent s'appliquer sur un do-

maine ou sur l'ensemble de la série, même si la tendance n'est pas linéaire.

Le champ d'application des tests qui suivent est plus vaste que celui de leurs équivalents paramétriques. Les tests non paramétriques s'appliquent aux tendances monotones alors que les tests paramétriques sont plus stricts: ils vérifient la présence ou l'absence d'une tendance linéaire (Marascuilo, 1977), et leur efficacité diminue lorsque la tendance n'est pas linéaire.

a) Hypothèses de base

L'utilisation de tous ces tests pose les mêmes conditions d'application sur les données d'une série:

- LES ε_i DOIVENT ÊTRE MUTUELLEMENT INDÉPENDANTS;
- LES ε_i DOIVENT PROVENIR D'UNE MÊME POPULATION CONTINUE (Conover, 1971).

L'univers auquel ces tests s'appliquent est constitué des séries dont les observations sont indépendantes s'il n'y a pas de tendance. Les tests ne peuvent être utilisés lorsque des facteurs autres que la tendance, comme la persistance ou la périodicité, créent de la dépendance entre les observations.

On trouve au paragraphe 3.1 les méthodes permettant de vérifier si la série respecte ces hypothèses.

Le tableau 2.1 présente dix tests de détection des tendances monotones que l'on retrouve dans la littérature. Un examen attentif des subdivisions du tableau 1 permet de saisir rapidement les points communs et les différences entre ces tests. On remarque, du haut vers le bas, une progression croissante dans:

- la précision des hypothèses;
- le contenu en information de la statistique;
- la puissance du test (tableau 2.2)

2.3.2 Test du nombre de séquences homogènes et test du nombre de points critiques

a) Les hypothèses statistiques (cf. tableau 2.1)

Pour ces tests, les hypothèses portent sur l'indépendance des observations de la série (Siegel, 1956):

H_0 : on ne constate pas de dépendance entre les observations;

H_1 : les observations sont mutuellement dépendantes.

Hypothèse	Statistique		Noms									
	Base	Formulation mathématique										
H_0 : on ne constate pas de dépendance entre les observations H_1 : les observations sont mutuellement dépendantes	le nombre de séquences homogènes* par rapport à un critère ex.: • supérieur, inférieur à la médiane • croissante ou décroissante nombre de maximums et minimums locaux	$T = n + m$ où: n : nombre de séquences homogènes de $y_i > \text{médiane}$ ou $y_i > y_{i+1}$ m : nombre de séquences homogènes de $y_i < \text{médiane}$ ou $y_i < y_{i+1}$ ou n : nombre de maximums locaux m : nombre de minimums locaux	Test du nombre de séquences homogènes Test du nb. de points critiques									
	le nombre d'observations : - supérieure à la médiane générale - inférieure à la médiane générale dans chacune des moitiés de l'échantillon	<table border="1"><tr><td>1e demi</td><td>2e demi</td><td>Total</td></tr><tr><td>A_1 B_1</td><td>A_2 B_2</td><td>T T_a b</td></tr><tr><td>T_1</td><td>T_2</td><td>T</td></tr></table> A = nb obs. > médiane gén. B = nb obs. < médiane gén. On calcule la pb. exacte de cette distr (Marascuilo, 1977)	1e demi	2e demi	Total	A_1 B_1	A_2 B_2	T T_a b	T_1	T_2	T	Irwin-Fisher
1e demi	2e demi	Total										
A_1 B_1	A_2 B_2	T T_a b										
T_1	T_2	T										
H_0 : on ne constate pas de dépendance entre les observations H_1 : les observations sont mutuellement dépendantes	le nombre d'observations supérieures à toutes les précédentes et le nombre d'observations inférieures à toutes les précédentes	$D = \sum_{i=1}^N dr$ où $dr = ur + lr$ où ur = nombre de $y_i > y_j$ lr = nombre de $y_i < y_j$ ou $j < i$	Foster et Stuart									
ou il y a une tendance décroissante	le nombre de comparaisons discordantes*, le nombre total de comparaisons est de l'ordre de $N/2$.	$S = \sum_{i=1}^{N-1} S_i$ où $S_i = 1$ si $(y_i - y_{i+1}) > 0$ autrement $S_i = 0$ $S = \sum_{i=1}^{N/2} S_i$ où $S_i = 1$ si $(y_i - y_{N/2+i}) > 0$ autrement $S_i = 0$	Test des différences successives Cox-Stuart									
H_0 : les observations de la série y sont indépendantes des observations de la série x	le nombre de comparaisons discordantes le nombre total de comparaisons est de l'ordre de $N!$	$Q = \sum_{i=1}^N S_i$ où $S_i = 1$ si $(y_i - y_j) > 0$ quand $i < j$ autrement $S_i = 0$ $D_2 = \sum_{i=1}^N (j-i) S_i$ où $S_i =$ même que test de Mann $\tau = \frac{S}{\frac{1}{2} N(N-1)}$ où $S = P - Q$ ou $P = 1$ si $y_i < y_j$, $Q = 1$ si $y_i > y_j$, pour $i < j$ autrement P et $Q = 0$	Mann Lehman Kendall									
H_1 : les observations de la série y ne sont pas indépendantes des observations de la série x	pour chaque observation, la différence de rang dans la série y et x	$r_s = 1 - \frac{6 \sum_{i=1}^N d_i^2}{N^3 - N}$ où $d_i = i - Y_i$	Spearman									

* séquences homogène : suite d'observations de même caractère dans la série transformée à caractère binaire
 * comparaisons discordantes : cas où $y_i > x_j$ et $i < j$

Important: voir les conditions d'application de ces tests à la page 21.

Tableau 2.1: Classification des tests de tendance monotone.

Nom	Correction pour les cas de valeurs égales	Valeurs critiques			ERA
		Petits échantillons (table)	Grands échantillons		
			Approximation ¹	Correction Continuité	
Nombre de séquences homogènes	ne s'applique pas	Conover 1971 (p. 414)	$E(T) = \frac{2mn}{m+n} + 1$ $Var(T) = \frac{2mn(2mn-m-n)}{(m+n)^2(m+n-1)}$	$T + \frac{1}{2} \leq T < E(T)$ $T - \frac{1}{2} \leq T > E(T)$	<< Kendall < que Stuart
Nombre de points critiques	ne s'applique pas	Conover 1971 (p. 414)	$E(T) = \frac{2(N-2)}{90}$ $Var(T) = \frac{16N-29}{90}$	"	0,02(2) là où 0,96 à Kendall 0,0 (5) là où 0,95 à Kendall
Irwin-Fisher	les observations = à la médiane sont distribuées alternativement dans les classes < et > que la médiane	Marascuilo, 1977 (p. 494)	$X^2 = \sum_{i=1}^2 \sum_{j=1}^2 \frac{X_{ij} - N/4}{N/4}$ i=indice pour moitiés inf. et sup. à la médiane X_{ij}=nb. de valeurs dans le groupe tel qu'indice X² suit loi X² d.l.=1	$z_{1/2} \leq X_{1/2}$ de façon à minimiser (X _{1/2} -N/4)	0,78 (3)
Foster-Stuart	le rang accordé aux valeurs égales se fait au hasard	Foster-Stuart, 1954 (p. 6)	$E(d_r) = 0$ $Var(d_r) = 2 \log N - 0,8756$	$dr + \frac{1}{2} \leq dr < E(dr)$ $dr - \frac{1}{2} \leq dr > E(dr)$	0,18(2) là où 0,96 à Kendall
Différences successives	On élimine une des deux valeurs. N est le nombre de données après élimination	Marascuilo, 1977 (p. 493)	$E(S) = \frac{N-1}{2}$ $Var(S) = \frac{N+1}{12}$	" avec S	0,0 (5) là où 0,95 à Kendall
Cox-Stuart	on compte 1/2 pour une comparaison où y ₁ = y ₂	Marascuilo, 1977 (p. 493)	$E(S) = N/2$ $Var(S) = N/4$	"	0,78 (3)
Mann		Marascuilo, 1977 (p. 524)	$E(Q) = \frac{1}{2} N(N-1)$ $Var(Q) = \frac{1}{12} N(N-1)(2 \text{ ou } 5)$	" avec Q	0,98 (4)
Lehman	deux autres formules sont alors proposées pour E(D ₂) et Var(D ₂)	Marascuilo, 1977 (p. 523)	$E(D_2) = \frac{N^3 - N}{6}$ $Var(D_2) = \frac{N^2(N+1)^2(N-1)}{36}$	" avec D ₂	0,98 (4)
Kendall	$T = \frac{S}{\sqrt{N(N-1) - T_x \sqrt{N(N-1)} - T_y}}$ $T_x = \frac{1}{2} \sum (t_x - 1)$ $T_y = \frac{1}{2} \sum (t_y - 1)$ t_x = nombre de x_i de chaque cas d'égalité t_y = nombre de y_i de chaque cas d'égalité	Marascuilo 1977 (p. 524)	$E(\tau) = 0$ $Var(\tau) = \frac{2(2N+5)}{9N(N-1)}$	$S+1 \leq T$ $S < 0$ $S-1 \leq T$ $S > 0$	0,99 (3)
Spearman	$r_s = \frac{1}{6} \frac{N^3 - N - \sum d_i^2 - T}{\sqrt{1(N^3 - N) - 2T} \sqrt{1(N^3 - N) - 2U}}$ $U = \frac{1}{12} \sum (u_i^2 - u_i)$ $T = \frac{1}{12} \sum (t_i^2 - t_i)$ u_i = nombre de y_i de chaque cas d'égalité t_i = nombre de x_i de chaque cas d'égalité	Marascuilo, 1977 (p. 523)	$t = r_s \frac{N-2}{1-r_s^2}$ t suit Student d.l.=2	" même que précédent avec d _i	0,98 (3)

Tableau 2.2 : Caractéristiques des tests pour tendance monotone.

Note : 1- On a, $Z = \frac{\text{Stat calculé} - E(\text{Stat})}{\sqrt{\text{Var}(\text{Stat})}}$ qui suit une loi normale (0, 1)

Pour ERA : 2- Puissance pour N = 100, α = 0,05 et Δ = 0,01, réf. Foster et Stuart, 1955

3- Réf. Marascuilo, 1977

4- Réf. Lehman, 1975

5- Réf. Cochet, 1981

6- Réf. Stuart, 1957

Une hypothèse alternative un peu plus précise a été proposée:

H_1 : les observations de valeur voisine ont tendance à être groupées dans le temps (Lehman 1975).

Cette dernière hypothèse se rapproche de l'idée de base du test: vérifier si les observations qui se rapprochent par leur valeur sont proches dans la séquence d'échantillonnage.

Les causes du regroupement chronologique des observations de valeur proche sont la tendance, l'autocorrélation et les cycles (Lehman, 1975). Etant donné ces trois causes et la possibilité d'un effet combiné, l'interprétation du rejet de H_0 peut être difficile, surtout si la tendance n'est pas évidente sur le graphique des données.

b) La statistique

La série originale est transformée en une autre qui ne peut prendre que deux valeurs, en fonction d'un critère. Lorsque l'on présente ce test comme un test de tendance on transforme notre série en une suite de A et de B,

$$y_i = A \text{ si } y_i > \text{médiane}$$

$$\text{et } y_i = B \text{ si } y_i < \text{médiane.}$$

À partir de cette nouvelle série, le test porte sur le nombre de séquences homogènes (cf. 2.2.2b) de A et B (Conover, 1971).

Le codage dichotomique peut aussi s'effectuer en fonction de la pente de la séquence. On a alors :

$y_i = A$ si y_i fait partie d'une séquence croissante

$y_i = B$ si y_i fait partie d'une séquence décroissante.

et de manière équivalente:

alors $y_i = A$ si $y_i < y_{i+1}$

et $y_i = B$ si $y_i > y_{i+1}$

Les sections à pente nulle (c'est-à-dire $y_i = y_{i+1}$) sont intégrées à la séquence qui les précèdent. C'est encore le nombre de séquences qui est testé. Si ce nombre est trop petit (c'est-à-dire si il existe des séquences longues), la série n'est pas aléatoire. Si de plus les séries les plus longues sont de type A, le rejet de H_0 peut être interprété comme l'acceptation d'une hypothèse de tendance croissante.

Les tests du nombre de points critiques et du nombre de séquences croissantes et décroissantes sont équivalents, ils calculent exactement la même statistique car le nombre de séquences croissantes et décroissantes est égal au nombre de points critiques (Lehman, 1975).

c) Les autres caractéristiques (cf. tableau 2.2)

La transformation d'une série de valeurs numériques qui peuvent être assez précises en une série binaire implique une perte d'information et peut expliquer la faible puissance de ces tests (Siegel, 1956).

La notion de puissance est assez ambiguë pour les tests portant sur l'indépendance entre les observations d'une série car cette puissance est relative à l'hypothèse alternative (stationarité, persistance, périodicité) et au test utilisé pour faire la comparaison. Cependant, on peut montrer, (Cox-Stuart, 1954) pour une population normale ayant une tendance dont l'accroissement $\Delta = 0,01$ entre les observations successives, que la puissance du test du nombre de points critiques pour des échantillons de taille 100 est de 0,02 alors que celle du test de Kendall est de 0,96.

Le test des séquences homogènes basé sur la médiane a une puissance inférieure à celle du test de Cox et Stuart, elle-même inférieure à celle du test de Kendall (Cochet 1981).

Les tests des séquences homogènes et du nombre de points critiques se caractérisent par :

- des hypothèses peu précises;
- une statistique à faible contenu en information;

- une faible puissance.

2.3.3 Tests de Irwin-Fisher Foster-Stuart, des différences successives, Cox-Sturat, Mann et Lehman

a) Les hypothèses statistiques

Pour les quatre premiers tests, l'hypothèse nulle (H_0) de l'indépendance des observations est testée contre l'hypothèse alternative de l'existence d'une dépendance. Dans les conditions d'application des tests (p. 21), l'indépendance des observations implique la stationnarité de la série, et la dépendance est le résultat d'une tendance. Pour les deux autres, les hypothèses sont discutées en 2.3.4.

b) La statistique

Pour la statistique de Irwin-Fisher on calcule la probabilité associée à la distribution exacte du tableau 2.1. Pour le test de Foster et Stuart (1954) la statistique est basée sur le nombre de maximums et de minimums depuis le début de la série, c'est-à-dire le nombre de valeurs supérieures à toutes les précédentes et le nombre de fois qu'une observation est inférieure à toutes les précédentes. La statistique peut aussi être calculée en faisant le comptage des maximums et minimums deux fois: une fois en commençant à compter par le début de la série, une autre fois en commençant par la fin.

Pour les tests des différences successives et celui de Cox-Stuart (Cochet, 1980; Cox-Stuart, 1955) on compare les valeurs de la série deux à deux, et la statistique est le nombre de comparaisons discordantes, c'est-à-dire celles où $Y_i > Y_j$ pour $i < j$. Dans le cas des différences successives, chaque valeur est comparée avec la suivante. Dans le cas du test de Cox-Stuart, pour augmenter la différence entre les valeurs comparées, on divise l'échantillon en deux et l'on compare y_i avec $y_{\frac{N}{2} + i}$. Cox et Stuart (1955) proposent même de diviser l'échantillon en trois, d'abandonner le tiers central et de faire les comparaisons entre y_i et $y_{\frac{2N}{3} + i}$. Pour les trois tests, on a donc respectivement $N-1$, $N/2$ et $N/3$ comparaisons.

Pour les tests de Mann et Lehman (Lehman, 1975), la statistique est encore le nombre de comparaisons discordantes, mais on ne compare pas les observations deux à deux. Chaque observation est comparée à toutes les autres. Le nombre de comparaisons passe alors à $N!$ Dans la statistique de Lehman, on a voulu donner plus de poids aux comparaisons pour lesquelles les observations sont espacées. La cote 1 accordée aux comparaisons discordantes est donc multipliée par $(j - i)$, la distance entre les observations comparées.

c) Les autres caractéristiques

Elles sont résumées au tableau 2.2. L'EAR du test de Irwin-Fisher comparé au test de Student est de 0,78 (Marascuilo, 1977). Le test de Foster et Stuart, avec une statistique qui lui est bien particulière a une puissance de 0,18 dans des conditions (décrites au bas du tableau 2.2) où le test de Kendall a une puissance de 0,96, et en faisant le comptage deux fois, une fois dans chaque sens, la puissance passe à 0,26 (Foster et Stuart, 1954).

L'EAR des tests des différences successives (Cochet, 1980) et du test du nombre de points critiques (Kendall, 1962) est de zéro comparé au test de Foster et Stuart (Stuart 1957). Par rapport au test des différences successives, le test de Cox-Stuart compare des valeurs plus éloignées l'une de l'autre, pour lesquelles la différence est plus importante s'il y a tendance. Même si le nombre de comparaisons diminue de $N-1$ à $N/2$, l'EAR atteint 0,78 comparé au test de Student (Stuart, 1957 et Conover, 1971). En éliminant le tiers central et en comparant deux à deux les observations des premier et troisième tiers, l'EAR passe à 0,83 (Stuart, 1957).

Lorsque le nombre de comparaisons passe de N à $N!$ l'EAR atteint 0,98. Les tests de Mann et Lehman ont la même EAR (Lehman, 1975).

En développant algébriquement la statistique pour qu'elle soit plus facile à compiler, Lehman arrive à:

$$D = \sum Y_i^2 - 2 \sum iY_i + \sum i$$

On voit que ce D est le développement de $\sum_{i=1}^N (Y_i - i)^2$, qui est la statistique de Spearman (voir tableau 1). D'ailleurs Lehman utilise la table de Spearman $D = \sum (Y_i - i)^2$ pour les valeurs critiques de sa statistique. Le test de Lehman est donc le même que celui de Spearman. De même, le test de Mann peut être développé et aboutit au test de Kendall (Lehman, 1975).

Lehman (1975) et Marascuilo (1977) présentent ces quatre tests comme des tests de dépendance. Tel que présenté par Lehman, les tests de Mann et de Lehman sont des formes moins élaborées des tests de Kendall et Spearman respectivement. Ces deux derniers aussi sont des tests d'indépendance basés sur la notion de corrélation non-paramétrique. L'absence de corrélation entre deux variables étant interprétée comme l'indépendance de ces deux variables.

2.3.4 Les tests de Kendall et de Spearman

a) Les hypothèses statistiques

Les hypothèses H_0 et H_1 (tableau 2.1) s'expriment sous la forme d'une dépendance ou indépendance entre la série de la variable y et le temps. Evidemment, le temps n'est pas une variable aléatoire, il augmente de mesure en mesure. Si la série de la variable " y " croît avec

le temps, il y a tendance croissante, si au contraire, y décroît avec le temps, il y a tendance décroissante.

b) La statistique

Pour le test de Kendall (Lehman, 1975), la statistique est basée sur la même quantité que pour le test de Mann et Lehman, soit le nombre de comparaisons discordantes sur $N!$ possibilités. Dans le tableau, la statistique τ est exprimée en fonction de S qui est la différence entre les nombres de couples concordants (P) et discordants (Q) (déf. au bas du tableau 1). τ peut aussi s'exprimer en fonction de Q seulement:

$$\tau = 1 - \frac{4Q}{N(N-1)}$$

La formulation en fonction de S (cf. tableau 2.1) a été conservée car c'est en fonction de S que s'exprime la formule de τ pour les cas où il y a des égalités.

Pour le test de Spearman, la statistique est basé sur la somme des différences d_i entre le rang (i) dans la série du temps et celui dans la série de la variable y (Y_i).

c) Les autres caractéristiques

Elles sont résumées au tableau 2.2. Les statistiques de ces deux tests ont des distributions telles que H_0 est rejeté avec la même efficacité (Marascuilo, 1977). La différence dans le coefficient obtenu n'est due qu'à une différence d'échelle.

L'EAR des tests de Kendall et de Spearman par rapport à celle du test basé sur le coefficient de corrélation paramétrique (Pearson) lorsque les deux variables sont normales et que la relation est linéaire est de 0,91. Si les x augmentent de façon unitaire, comme ce peut être le cas dans les séries chronologiques, l'EAR de ces tests non-paramétriques est de 0,98 (Marascuilo, 1977).

Si la population d'origine n'est pas normale, l'EAR du test de Kendall peut dépasser celle du coefficient paramétrique de Pearson (tableau 2.3).

Distribution	EAR
normale	0,912
uniforme	1
double exponentielle	1,266

Tableau 2.3: EAR du test de Kendall par rapport au test paramétrique (Pearson) (tiré de Hollander, 1973).

Les tests de Kendall et Spearman se caractérisent donc par :

- une hypothèse précise (corrélation entre x et y);
- des statistiques utilisant beaucoup d'information ($N!$ comparaisons);
- une puissance élevée;
- des formules précises pour le cas des égalités.
- le test de Kendall se prête bien à un calcul progressif permettant de situer le début de la tendance (Sneyers 1970).
- une approximation à la loi normale valable pour les petits échantillons (cf. Siegel, 1956);

2.3.5 Discussion

Certains test applicables à la détection de tendances monotones n'ont pas été présentés ici. Ces tests sont utilisés pour les données à échelle nominale. Leur utilisation implique un regroupement des données en classes, et leur puissance est faible comparée aux tests basés sur la corrélation de rang (Kendall et Spearman) lorsque les données sont disponibles sur une échelle ordinale ou numérique (Marascuilo, 1977). C'est le cas des tests de signe (Siegel, 1956).

Hollander (1971) présente le test de Kendall sous une forme généralisée au cas où l'on peut avoir $x_j <, >, \text{ ou } = x_i$ pour $i < j$. La statistique et les tables de valeurs critiques sont les mêmes que pour le test de Kendall tel que présenté ici. En fait, Mann (1945) est le premier à avoir utilisé les couples discordants pour un test de tendance. Son test est le même que le test de Kendall tel que présenté ici (Hollander, 1973; Cochet, 1980). Dans la même publication Mann (1945) propose un test pouvant être utilisé lorsque la tendance semble forte.

Dans le cas plus précis de l'hypothèse d'une tendance linéaire plutôt que simplement monotone, le modèle est:

$$y_i = a + b x_i + \epsilon_i \quad i = 1 \dots N$$

Une estimation non paramétrique et un intervalle de confiance non paramétrique de b ainsi qu'un test non paramétrique pour examiner l'hypothèse que le b de la population est égale à un b_0 donné ont été proposés (Theil, 1950). Toutes ces procédures ont été généralisées au cas où il y a des égalités dans les x_i (Sen 1968). Le test de b lorsque l'on pose $b_0 = 0$ revient au test de Mann (Hollander 1973).

On peut classer les dix tests considérés sur une échelle d'efficacité asymptotique relative:

0	10	15	30	75	85	90	100%
Points critiques	Foster-Stuart			Irwin-Fisher Cox-Stuart		Kendall Spearman Mann Lehman	
Suites homogènes							
Différences successives							

Les tests les plus puissants sont les tests de corrélation de rang à $N!$ comparaisons: Kendall, Spearman, Mann et Lehman. Les deux derniers sont équivalents mais moins développés que les deux premiers. La puissance des tests de Kendall et Spearman est presque égale à celle des tests paramétriques, dans les conditions d'application optimales de ces derniers (EAR = 0,91 à 0,98). Dans d'autres conditions d'application, leur puissance peut dépasser celle des tests paramétriques (tableau 2.3). Ces tests ont aussi une formulation précise pour les cas où il y a des égalités dans les données (tableau 2.2). Les tests de Kendall et de Spearman apparaissent donc comme les meilleurs tests pour la détection d'une tendance monotone dans des conditions générales.

2.4 Les tendances en saut

2.4.1 La définition et les bases de la détection d'une tendance en saut

Une série où il y a une tendance en saut peut être modélisée par (Lettenmaier, 1976) :

$$y_i = \mu_1 + \varepsilon_i + (\mu_2 - \mu_1)_k$$

avec:

$$(\mu_2 - \mu_1)_k = 0 \text{ pour } i < k, \text{ et}$$

$$(\mu_2 - \mu_1)_k = \mu_2 - \mu_1 \text{ pour } i > k, \text{ et où}$$

μ_1 : moyenne après le saut

μ_2 : moyenne avant le saut

ε_i : variable aléatoire, indépendante, distribuée normalement, de moyenne 0 et de variance σ^2

k : rang temporel de la variable pour laquelle il y a le saut.

Le test permet d'examiner l'hypothèse $H_0: (\mu_2 - \mu_1) \neq 0$, c'est-à-dire de vérifier si les moyennes avant et après le saut sont significativement différentes. Ces tests reviennent donc à la comparaison des moyennes de deux populations.

a) Hypothèses de base

Ces tests posent les mêmes hypothèses de base (conditions d'application) que les tests de détection de tendances monotones. Ces conditions sont énumérées à la section 2.3.1 a). De plus, si l'on décompose l'échantillon global en deux sous-échantillons, il doit y avoir indépendances des résidus (ε_i) à l'intérieur des sous-échantillons et entre les sous-échantillons.

2.4.2 Les tests de Wald-Wolfowitz et de Kolmogorov-Smirnov

a) Les hypothèses statistiques

L'hypothèse H_0 porte sur l'égalité des distributions. Le test de Wald-Wolfowitz est basé sur une statistique similaire à celle du test des séquences homogènes pour les tendances monotones. En plus d'être sensibles aux différences de moyennes, ce test permet de détecter des inégalités d'autres caractéristiques de la distribution, comme la variance (Siegel, 1956; Conover, 1971). La cause du rejet de H_0 peut donc être difficile à identifier et rend malaisé l'interprétation des résultats. Le test de Klotz permettant de vérifier l'égalité des variances peut être utilisé avant de tester l'égalité des moyennes. Ce test est présenté à l'annexe 4.

Le test de Komogorov-Smirnov vérifie l'ajustement des distributions des deux échantillons. Comme le test précédent, il compare les distributions par l'ensemble de leurs caractéristiques.

Lorsque la variance et la symétrie des distributions ont été testées et que l'hypothèse de leur égalité n'a pas été rejetée, on pourrait considérer l'application de ces tests pour vérifier l'égalité des valeurs centrales, mais pour une telle fonction, le test de Mann-Whitney est plus puissant. (Conover, 1971).

b) La statistique

Les statistiques de ces tests sont décrites au tableau 2.4. Cramer-Von Mises (1946) ont proposé un test de même type que celui de Kolmogorov-Smirnov. Si le saut a lieu à l'observation k , la statistique est la somme des différences $(y_i - y_{k+i})$. L'efficacité de ce test n'est que très légèrement supérieure à celle du test de Kolmogorov-Smirnov (Conover, 1971).

c) Les autres caractéristiques

L'utilisation de ces tests dans le cas où il y a des égalités pose des problèmes (voir tableau 2.5).

Comme dans le cas du test du nombre de séquences homogènes pour les tendances monotones, le calcul de la puissance de ces tests dépend de l'hypothèse alternative et du test utilisé pour la comparaison. Le rejet de H_0 peut dépendre de plusieurs causes ce qui peut rendre ambigu le choix d'un test pour faire la comparaison de puissance. Toutefois, pour rejeter H_0 dans le cas où les distributions sont normales et ne diffèrent que par leur moyenne, le test de Wald-Wolfowitz est faible comparé au test de Student (Siegel, 1956). Dans la même situation, l'EAR du test de Komogorov-Smirnov se situe entre 0,64 et 0,75 (Lehman, 1975). Sa limite inférieure serait de 0,63 (Hollander, 1973). Ce test serait plus puissant que celui de Mann-Witney pour les échantillons très petits (Marascuilo, 1977).

Hypothèse	Statistique		Noms
	Base	Formulation mathématique	
H_0 : Les deux échantillons ont la même fonction de distribution H_1 : Les deux échantillons n'ont pas la même fonction de distribution ou Les valeurs d'un échantillon ont tendance à être plus petites que celles de l'autre (test unilatéral)	À partir de l'échantillon total, le nombre de séquences homogènes* provenant d'un des deux échantillons d'origine La plus grande différence sur l'axe y entre les fonctions de distribution cumulées des deux échantillons	T = nombre séquences homogènes d'un des échantillons Bilatéral $T = \max (S_1(x) - S_2(x))$ Unilatéral $T = \max (S_1(x) - S_2(x))$ où $S_1(x)$ = valeur de la F_1^* au point x $S_2(x)$ = valeur de la F_2 au point x	Wald-Wolfwitz Kolmogorov-Smirnov
H_0 Les deux échantillons proviennent de population ayant la même médiane H_1 Les deux échantillons ne proviennent pas de population ayant la même médiane ou La médiane de la population d'origine d'un échantillon est supérieure à celle de l'autre (test unilatéral)	Pour chaque échantillon, le nombre d'observations < et le nombre d'observations > que la médiane générale	$P = \frac{(A+B)!(C+D)!(A+C)!(B+D)!}{N!A!B!C!D!}$ où A: nb. obs. échan. 1 > médiane B: nb. obs. échan. 1 < médiane C: nb. obs. échan. 2 > médiane D: nb. obs. échan. 2 < médiane P: probabilité de cette distribution sous H_0	De la médiane
H_0 Les deux échantillons proviennent de population ayant la même moyenne H_1 Les deux échantillons ne proviennent pas de population de même moyenne ou La moyenne de la population d'origine d'un échantillon est supérieure à celle de l'autre. (test unilatéral)	Le nombre de fois qu'une observation du 1er échantillon est < qu'une observation du 2e échantillon On calcule la somme des valeurs de Z* associées aux rangs d'un des échantillons dans l'échantillon total. Somme des valeurs Z correspondant aux probabilités empiriques associées aux rangs des observations d'un des échantillons dans l'échantillon total On tire de façon aléatoire N valeurs de Z que l'on ordonne par rang. La statistique est la somme des Z de même rang que les observations d'un des échantillons dans l'échantillon total	$U = \sum_{i=1}^{n_1} R_i - \frac{n_1(n_1+1)}{2}$ R_i = pour chaque observation du groupe 1, le nombre d'observation du groupe 2 qui lui est supérieure $T_1 = \sum_{i=1}^{n_1} Z(Y_j)$ où $Z(Y_j)$ = valeur de Z associé aux Y_j (Table A-15 Marascuilo) n_1 = effectif de l'échantillon 1 $T_1 = \sum_{i=1}^n Z_i$ Z_i = valeur de Z associé à P_j $P_j = \frac{Y_j}{N+1}$ = prob. empirique associé au rang Y_j $T_1 = \sum_{i=1}^n Z_i$ Z_i = valeur de Z pris au hasard de même rang que le Y_j T_2 = même pour échantillon 2	Mann-Whitney Terry-Hoeffding Van der Waerden Bell-Doksum

Note : * Séquence homogène : voir définition au bas du tableau 1
 * F_1 fonction de distribution cumulée de l'échantillon 1 (F_2 = échantillon 2)
 * Z variable normale, centrée, réduite
 * Y_j rang des observations provenant d'un des échantillons dans l'échantillon total ordonné

Tableau 2.4 : Classification des tests pour les tendances en saut

Nom	Correction pour les cas de valeurs égales	Valeurs critiques		ERA
		Petits échantillons (table)	Grands échantillons	
			Approximation ¹ Correction Continuité	
Wald-Wolfowitz	Si des valeurs d'un échantillon sont = à celles d'un autre, il faut essayer les valeurs ≠ de T qui peuvent en résulter. Si on arrive à ≠ décision pour ce qui est du rejet de H ₀ on ne peut utiliser le test	Conover, 1971 (p. 414)	n_1 ou $n_2 > 20$ $E(T) = \frac{2n_1n_2}{(n_1 + n_2)} + 1$ $Var(T) = \frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2 (n_1 + n_2 - 1)}$	T > E(T) alors T - ½ T < E(T) alors T + ½ faible
Kolmogorov-Smirnov	On ne s'en occupe pas, et le test est conservatif (voir note 4)	Conover, 1971 (p. 399-400)	$n_1, n_2 > 20$ $T = C \sqrt{\frac{n_1 + n_2}{n_1 n_2}}$, si $n_1 \neq n_2$ $T = \frac{C}{\sqrt{n}}$, si $n_1 = n_2 = n$ C dépend du seuil α, il est disponible dans les tables (Conover, 1971)	1-2 0,64
De la médiane	pas de correction nécessaire	Siegel, 1956 (p. 256)	$n_1 + n_2 > 40$ $X^2 = \frac{(AD-BC - \frac{N}{2})^2}{(A+B)(C+D)(A+C)(B+D)}$ X ² : suit loi chi-deux d. l. = 1	Ne change rien 1-63
Mann-Whitney	$Var(u) = \frac{n_1 n_2 (n_1 + n_2 + 1)}{12}$ $1 - \sum_{s=1}^S \frac{ts^3 - ts}{n^3 - n}$ S : nombre de groupes où il y a des rangs égaux ts : nombre de valeurs égales par groupes	Marascuilo (p.504) Bobée 1981	$n_1, n_2 > 20$ $E(u) = \frac{n_1 n_2}{2}$ $Var(u) = \frac{n_1 n_2 (n_1 + n_2 + 1)}{12}$	± ½ à u aux limites de la zone critique 0,955
Terry-Hoeffding	On donne un même rang moyen aux valeurs égales, ou on leur donne leur rang au hasard	Calcul de la probabilité exacte de T ₁ : Marascuilo, (p. 281 - 282)	$n_1, n_2 > 9$ $E(T_1) = 0$ $Var T_1 = \frac{n_1 \cdot n_2}{N-1} \sum_{i=1}^N \frac{Z^2(Y_j)^2}{N}$, où $Z(Y_j) = z$ associé au Y _j	T ₁ < 0, alors T ₁ + ½ T ₁ > 0, alors T ₁ - ½ 1,00
Van der Waerden	"	Calcul de la probabilité exacte Marascuilo, (p. 283-284)	$n_1, n_2 > 8$ $E(T_1) = 0$ $Var T_1 = \frac{n_1 n_2}{N-1} \sum_{i=1}^N \frac{Z^2}{N}$ Z _i = déf. au tableau 4 section Van der Waerden	T ₁ < 0 alors T ₁ + ½ T ₁ > 0 alors T ₁ - ½ 2-3
Bell-Doksum	"	Calcul de la probabilité exacte Marascuilo, (p. 285)	$n_1, n_2 > 8$ $Z = \frac{T_1/n_1 - T_2/n_2}{\sqrt{\frac{1}{n_1} - \frac{1}{n_2}}}$ Z suit normale (0,1)	T ₁ < T ₂ alors T ₁ + ½ T ₁ > T ₂ alors T ₁ - ½ 2,3 1,00

NOTES:

- 1: voir Note 1, Tableau 2
Pour EAR = 1: réf. Siegel 1956
2: réf. Conover 1971
3: réf. Marascuilo 1977
4: Solution proposée: Hajek 1969

Tableau 2.5 Caractéristiques des tests de tendance en saut

2.4.3 Tests de la médiane et de Mann-Whitney

a) Les hypothèses statistiques

Ces tests servent à vérifier si les deux échantillons diffèrent par leur valeur centrale. D'une façon rigoureuse, on dira que Mann-Whitney teste l'hypothèse de l'égalité de l'espérance des moyennes si les variances des deux distributions sont égales. Sinon, il teste l'égalité des médianes (Marascuilo, 1977).

Ces tests sont plus efficaces que les tests précédents pour détecter des divergences de la valeur centrale.

b) La statistique

Les statistiques de ces deux test sont présentées au tableau 2.4. Dans le test de la médiane, toutes les observations sont comparées à la médiane générale. Dans le cas du test de Mann-Whitney, chaque observation d'un échantillon est comparée à toutes celles de l'autre. Il y a donc un plus grand nombre de comparaisons, donc d'information qui entre dans la statistique de Mann-Whitney.

c) Les autres caractéristiques

Ces caractéristiques sont résumées au tableau 2.5. L'EAR du test de Mann-Whitney (0,955) est supérieure à celle du test de la médiane.

Cette différence est vraisemblablement due à la plus grande quantité d'information prise en compte dans le test de Mann-Whitney que dans celui de la médiane.

Le test de Mann-Whitney est un des tests non-paramétriques les plus utilisés en raison de sa puissance. L'EAR de ce test est toujours supérieure à 0,864 (Marascuilo 1977), alors que dans certaines conditions, celle du test du t de Student peut être de zéro; en effet, pour plusieurs distributions l'EAR du test non-paramétrique est supérieur à celle du test de Student (voir tableau 2.6).

DISTRIBUTION	EAR
Normale	0,955
Uniforme	1,0
Logistique	1,097
Double exponentielle	1,5
Exponentielle	3,0

Tableau 2.6 : EAR du test de Mann-Whitney par rapport à celle du test de Student (tiré de Hollander, 1973)

On trouve au tableau 2.5 une formule modifiée de la variance, que l'on doit considérer lorsqu'il y a des égalités entre les observations. L'effet des égalités est mineur; il est plus marqué dans les cas où plusieurs observations ont la même valeur que dans celui où il y a plusieurs cas d'égalité entre seulement deux d'observations. Le test utilisé sans correction est conservateur (Siegel, 1956).

Hollander (1973) présente une généralisation du test pour H_0 telle que:

$$H_0 : \mu_1 - \mu_2 = \Delta$$

Dans le cas du test de Mann-Whitney on a $\Delta = 0$ sous l'hypothèse H_0 . Il est également possible de déterminer un intervalle de confiance pour Δ . Le test de Mann-Whitney peut également être modifié pour tester l'égalité des variances (Conover, 1971).

Le test de Mann-Whitney peut être également modifié dans le cas où les mesures sont bornées par un maximum, un minimum ou les deux, dus à une cause externe : les limites d'un appareil de mesures par exemple. Hollander (1973) et Lehman (1975) donnent plusieurs références à ce sujet.

2.4.4 Les tests basés sur les variables normalisées: Terry-Hoeffding, Van Der Waerden, Bell-Doksum

a) Les hypothèses statistiques

Dans le cas de ces trois tests, les hypothèses sont les mêmes que celle du test de Mann-Whitney.

b) La statistique

Les statistiques relatives à ces tests sont présentées dans le tableau 2.4. Pour ces trois tests, les données sont remplacées par des nombres dont la distribution est basée sur la loi normale. Les rangs sont remplacés par la valeur correspondante de la variable normale centrée réduite (Z) dans le test de Terry Hoeffding. Dans le test de Van der Waerden, c'est la probabilité empirique associée au rang qui est remplacée par le Z correspondant. Pour le test de Bell Doksum, on tire au hasard N valeurs de Z ; on fait ensuite la somme des N valeurs de Z de même rang que les n observations d'un échantillon parmi toutes les N observations. L'utilisation de ce test impose l'entrée d'une autre composante aléatoire dans la variable, par la sélection au hasard des valeurs de Z . Ainsi on peut arriver à deux valeurs différentes de la statistique en faisant le test deux fois sur les mêmes données.

c) Autres caractéristiques

Elles sont résumées au tableau 2.5. L'EAR du test basé sur les valeurs normalisées, par rapport au test t , lorsque la différence entre les distributions n'est due qu'à une différence de moyenne est toujours supérieure à 1, quelque soit la distribution (Conover, 1971; Marascuilo, 1977). Ces trois tests ont la même fonction de puissance (Marascuilo, 1977). Dans les mêmes conditions, ces fonctions diffèrent peu de celles de Mann-Whitney (Lehman, 1975). Lorsque les populations d'origine ont une composante asymptotique réduite, par exemple pour

les lois normales et exponentielles, l'efficacité de ces tests est supérieure à celle du test de Mann-Whitney. C'est l'inverse si la composante asymptotique des distributions est grande comme dans le cas des distributions uniforme, logistique, double exponentielle, et Cauchy (Marascuilo, 1977).

2.4.5 Discussion

Certains tests applicables à la détection de tendances en saut ont été omis :

Le test de Tukey : un test prenant en compte peu d'information, dont la puissance n'est pas connue (Conover, 1971)

Le test de "randomization" : ce test ne s'applique que pour les petits échantillons. Les tests de Mann-Whitney et les tests de valeurs normalisées sont des adaptations plus fonctionnelles de ce test.

Les tests pour les mesures à échelle nominale : Ces tests sont moins puissants que ceux décrits ici lorsque les données sont disponibles sur une échelle ordinale ou numérique.

Une fois l'existence d'une différence entre les moyennes confirmée par ces tests, on peut utiliser d'autres procédures non-paramétriques pour estimer cette différence de même que l'intervalle de confiance

de la différence théorique (Hollander, 1973; Lehman, 1975; Marascuilo, 1977).

Les tests les plus puissants pour la détection d'une tendance par saut sont le test de Mann-Whitney et les tests basés sur les valeurs normalisées. Le choix entre les deux dépend des caractéristiques de la population d'origine. En général, le choix de l'un ou l'autre est sans grande incidence sur la puissance (Marascuilo, 1977). Si l'on considère de plus que le test de Mann-Whitney :

- a presque la même puissance (0,955) que son équivalent paramétrique (test t) dans les conditions d'applications optimales de ce dernier;
- a une puissance qui n'est jamais inférieure à 0,84 et qui reste supérieure à celle du test paramétrique (test t) lorsque la population n'est pas normale;
- a une statistique d'un calcul plus simple que les tests de valeurs normalisées.

On en conclut qu'en général, le test de Mann-Whitney est préférable pour détecter une tendance en saut.

2.5 Tendance en escalier

2.5.1 La définition et les bases de la détection d'une tendance en escalier

Une série dont la représentation graphique a la forme d'un escalier peut être divisé en plusieurs échantillons correspondant à chacun des paliers. Le modèle de cette série est une répétition du modèle proposé pour les tendances en saut. On a alors :

$$y_i = \mu_j + \varepsilon_i + (\mu_j - \mu_m)_k, \text{ où}$$

$$(\mu_j - \mu_m)_k = 0 \text{ pour } i \leq k$$

$$(\mu_j - \mu_m)_k = \mu_j - \mu_m \text{ pour } i > k$$

μ_j = moyenne du palier avant le saut

μ_m = moyenne du palier après le saut

k = rang temporel des variables où il y a un saut entre le palier j et m .

ε_i = variable aléatoire indépendante distribuée normalement de moyenne zéro et variance σ^2

Il s'agit donc de vérifier l'hypothèse nulle :

$$H_0 : \mu_1 = \mu_j = \mu_m, \text{ pour tout palier } j, m.$$

On constate une forte analogie du modèle et des hypothèses de ces tests avec ceux des tendances en saut. Cette ressemblance n'est pas fortuite car les tests présentés ici sont des formes généralisées des tests de tendance en saut. Les tests de Kruskal-Wallis et de Terpstra-Jonckheere sont des versions à K échantillons du test Mann-Whitney (Conover, 1971; Hollander, 1973). Le test de Puri est relié au test de Van der Waerden de la même façon (Puri, 1965).

a) Hypothèses de base

Les hypothèses de base que les séries doivent respecter sont les mêmes que celles énoncées pour les tests de tendance en saut (section 2.4.1 a), sauf qu'elles s'appliquent à K échantillons au lieu de deux.

2.5.2 Les tests de Kruskal-Wallis, Terpstra-Jonckheere et Puri

a) Les hypothèses statistiques

L'hypothèse nulle de l'égalité des valeurs centrales des distributions des populations dont proviennent les K échantillons est la même pour les trois tests:

$$H_0 : \mu_1 = \dots \mu_j = \mu_m$$

Dans le cas de Kruskal-Wallis, l'hypothèse alternative est :

H_1 : au moins 1 des μ_j différent des autres.

Pour les deux autres tests, cette hypothèse est précisée:

H_1 : $\mu_1 \leq \dots \leq \mu_j \leq \mu_m$, avec au moins une inégalité stricte.

b) La statistique

Les statistiques de ces tests sont présentées au tableau 2.7. On note que le test de Puri est basé sur une variable normalisée.

c) Les autres caractéristiques

Elles sont résumées au tableau 2.8. L'EAR des tests de Kruskal-Wallis et de Terpstra-Jonckheere sont les mêmes que celle du test de Mann-Whitney (cf. 2.4.3 c) (Hollander, 1973). Dans les mêmes conditions, le test de Puri a les mêmes propriétés asymptotiques que le test de Van der Waerden (cf. 2.4.4 c) (Puri, 1965).

2.5.3 Discussion

Les tests présentés ici sont des généralisations pour K échantillons des tests de tendance en saut présenté en 2.4.

Hypothèse	Statistique		Noms
	Base	Formulation mathématique	
H_0 = les k échantillons ont la même fonction de distribution H_1 = les k échantillons n'ont pas la même fonction de distribution où $H_0 : \mu_1 = \mu_j = \mu_m$ $H_1 : \mu_1 \neq \mu_j \neq \mu_m$ L'égalité des variances est supposée ou confirmée	La somme des rangs pour chaque échantillon (rang dans l'échantillon total)	$H = \frac{12}{N(N-1)} \cdot \sum_{j=1}^K \frac{R_j^2}{n_j} - 3(N+1)$ où K = nombre d'échantillons n_j = effectif de l'échantillon j R_j = somme des rangs de l'échantillon j	Kruskal-Wallis
H_0 = les k échantillons ont la même moyenne H_1 = les moyennes des échantillons sont ordonnées : $\mu_1 < \mu_j < \dots \mu_m$ avec au moins 1 inégalité stricte	Chaque échantillon est comparé au suivant. On apparie les observations de même rang, et l'on compte les couples concordants: là où la valeur y du 1er échantillon est inférieure à celle du 2e La somme des valeurs normalisées des rangs de l'échantillon, dans l'échantillon total	$J = \sum_{u=1}^{K-1} \sum_{v=u+1}^K U_{uv}, \text{ où}$ $U_{uv} = \sum_{i=1}^{n_u} \sum_{i=1}^{n_v} \phi(y_{iu}, y_{iv}), \text{ où}$ K = nb. d'échantillons u = indice du 1er échantillon de la comparaison v = indice du 2e échantillon de la comparaison n_u = effectif de u n_v = effectif de v $\phi(y_{iu}, y_{iv}) = 1$ si $y_{iu} < y_{iv}$ $\phi(y_{iu}, y_{iv}) = 0$ autrement $P = \sum_{k=1}^K (L_k \cdot E_k), \text{ où}$ K = nb. d'échantillons $L_k = nc$ si n'est pair $2nc$ si n'est impair n = effectif de l'échantillon c = coefficient de contraste linéaire, dépend de n Table : (Berenson, 1978). E_k : Somme des valeurs normalisées des rangs d'observation de l'échantillon k dans l'échantillon total (Table A-15, Marascuilo, 1977)	Terpstra-Jonckheere Puri

Tableau 2.7 : Classification des tests de tendance en escalier

Nom	Correction pour les cas de valeurs égales	Valeurs critiques			ERA
		Petits échantillons (table)	Grands échantillons		
			Approximation	Correction Continuité	
Kruskal-Wallis	$H_1 = \frac{H}{1 - \sum_{s=1}^S \frac{(t_s^3 - t_s)}{N^3 - N}}$ S = nb de groupes de rang égaux t_s = nb de valeurs dans un groupe de rang égal	Conover, 1971 pour K = 3 et et n _k < 5, (p. 393.) Si non on utilise l'approximation χ ²	H < X ² _{k-1;α} → H ₀ pas rejeté H > X ² _{k-1;α} → H ₀ rejeté, où X ² _{k-1;α} = valeur du chi-carré à K-1 d.l. et seuil α	H > X ² _{k-1;α} alors : H - ½ H < X ² _{k-1;α} alors : H + ½	0,955
Terpstra-Jonckheere	Si y _{iu} = y _{iv} alors Ø (y _{iu} , y _{iv}) = ½	Hollander (p. 311) Odeh (p. 919)	$E(J) = \frac{N^2 - \sum_{j=1}^k n_j^2}{4}$ Var (J) = [N ² (2N+3) - $\sum_{j=1}^k N_j^2 (2N+3)$] / 72	J > E(J), alors : J - ½ J < E(J) alors : J + ½	0,955
Puri	Si plusieurs valeurs ont la même valeur, on leur donne le même rang : le rang moyen		Valeur critique pour P au seuil α : P _{cr} = Z _{1-α} • σ _p où $\sigma_p = \frac{n \sum_{i=1}^n \sum_{k=1}^K E^2}{N-1} \sum_{k=1}^K L_k^2$ Voir Berenson 1978 pour tables de σ _p pour : 5 < n < 20 3 < k < 7		1,00

Tableau 2.8 : Caractéristiques des tests de tendance en escalier

Le test de Kruskal-Wallis n'en est pas un de détection de tendance à proprement parler. Il vérifie l'égalité des distributions des paliers. Le rejet de H_0 n'implique pas l'existence d'une tendance. Par exemple, l'hypothèse H_0 pourrait être rejeté dans le cas où on aurait trois paliers, successivement très haut, très bas et très haut. Ce test est robuste aux différences des autres paramètres (écart type, asymétrie) des distributions (Marascuilo, 1977).

Si le niveau des paliers croît ou décroît dans le temps, on utilisera un des deux autres tests. La raison est bien simple : pour ces deux tests le rejet de H_0 implique l'acceptation d'une hypothèse alternative qui précise le sens de la tendance. Le choix à faire entre ces deux tests n'est pas évident. Il repose sur les mêmes considérations par rapport à la population d'origine que celles détaillées en 2.4.3.c. pour les tests de tendance en saut. Ainsi, on utilisera le test de Puri dans des conditions similaires à celle du test de Van der Waerden et le test de Terpstra-Jonckheere, dans les conditions où l'on choisit plutôt le test de Mann-Whitney. Étant donné le parallèle avec les tests à 2 échantillons, on s'attend à ce qu'il n'y ait pas une grande différence de puissance dans l'utilisation des tests de Puri et de Terpstra-Jonckheere.

3. APPLICATION DES TESTS AUX SÉRIES TEMPORELLES

Ce chapitre expose la problématique de l'utilisation des tests sur les séries temporelles en insistant particulièrement sur l'inadéquation entre les exigences théoriques des tests (3.1) et certaines caractéristiques des séries observées (3.2).

En 3.3 et 3.4, on résume et commente les façons d'adapter les tests pour prendre en compte les particularités des séries.

Les séries inadéquates pour les tests sont caractérisées par un effet de persistance (3.2.1), une périodicité (3.2.2), ou les deux (3.2.3). À ces trois types de séries, correspondent respectivement trois types d'adaptation des tests : la méthode de Lettenmaier (1976) (3.3.1); de Hirsch et al. (1982) (3.3.2) et de Hirsch et Slack (1983) (3.3.3). Pour ces trois mêmes types de série, mais caractérisées par des tendances en saut plutôt que monotone, des éléments de solution sont proposés en 3.4.1, 3.4.2 et 3.4.3.

Les auteurs cités précédemment ont choisi pour l'étude de tendances monotones les tests de Kendall, de Spearman et de Mann-Whitney. La revue des tests effectuée au chapitre précédent a d'ailleurs démontré que ces trois tests sont, de manière générale, les plus performants. Ce sont donc eux que nous retiendrons dans la suite.

3.1 Exigences théoriques des tests : les hypothèses de base

Les trois types de tendance présentés au chapitre 2 peuvent être décrites par un modèle:

$$Y_i = \tau(x_i) + \epsilon_i \quad (\text{éq. 3.1})$$

où

$\tau(x_i)$ une fonction de x_i particulière au type de tendance;

ϵ_i est la partie de y_i non expliquée par la tendance: l'erreur.

Les tests de détection de tendances nécessitent deux conditions d'application sur les erreurs ϵ_i .

- 1- les ϵ_i doivent provenir d'une même population continue de moyenne nulle et de variance σ_{ϵ_i} ;
- 2- les ϵ_i doivent être mutuellement indépendants.

Lorsque les tests sont appliqués à des séries ne respectant pas ces conditions, les erreurs de type 1 et 2 (α et β) déduites de la théorie peuvent changer. La conclusion du test peut alors être erronée. Avant d'utiliser un test il est préférable de vérifier si les observations respectent ces deux hypothèses de base.

Cette vérification ne peut se faire à partir des erreurs ε_i puisqu'elles ne sont pas connues. On utilise alors les résidus e_i qui sont des approximations des erreurs (Brown 1976; LLamas et Triboulet 1976).

Dans le modèle général présenté plus haut, s'il n'y a pas de tendance:

$$\tau(x) = 0 \quad \text{donc } y_i = \varepsilon_i \quad (\text{éq. 3.2})$$

On peut alors utiliser les valeurs de y , pour vérifier les conditions d'application (ou hypothèses de base). Par contre si $\tau(x_i) \neq 0$, il faut soustraire la composante expliquée par la tendance. On se retrouve alors dans un cercle vicieux: il faut connaître la tendance pour vérifier si les conditions d'application des tests de détection de cette tendance sont respectées. La solution est de faire une approximation de la tendance. Cette approximation se fait par la méthode usuelle d'ajustement d'un polynome dont les coefficients sont déterminés par la méthode des moindres carrés. Fréquemment l'approximation linéaire est suffisante (LLamas et Triboulet 1976). On obtient ainsi $T(x_i)$ une approximation de $\tau(x_i)$. On a donc:

$$y_i = T(x_i) + e_i \quad (\text{éq. 3.2})$$

où

$T(x_i)$ = composante de y_i expliqué par l'approximation linéaire de la tendance;

$e_i = y_i - T_i$ est le résidu, c'est-à-dire la composante de y_i non prise en compte $T(x_i)$. C'est une approximation de l'erreur ε_i .

La vérification des hypothèses de base se fait alors à partir des résidus e_i .

3.1.1 Hypothèse de l'égalité des distributions des erreurs

La méthode des moindres carrés à partir de laquelle on obtient les résidus consiste à ajuster une droite de façon à minimiser la somme des carrés des résidus. Puisque les e_i sont des estimations des ε_i tels que $E(\varepsilon_i) = 0$, le modèle ajusté est tel que la moyenne $\bar{e}$ des e_i est nulle. On a donc:

$$\frac{1}{N} \sum_{i=1}^N e_i = 0 \quad \text{et} \quad \sum_{i=1}^N e_i = 0$$

Il reste donc à vérifier si les e_i ont la même variance autour de zéro puisque les ε_i proviennent d'une même population donc $\text{Var } \varepsilon_i = \sigma_\varepsilon$. Une première approche pour cette vérification consiste à examiner le graphique des e_i en fonction du temps.

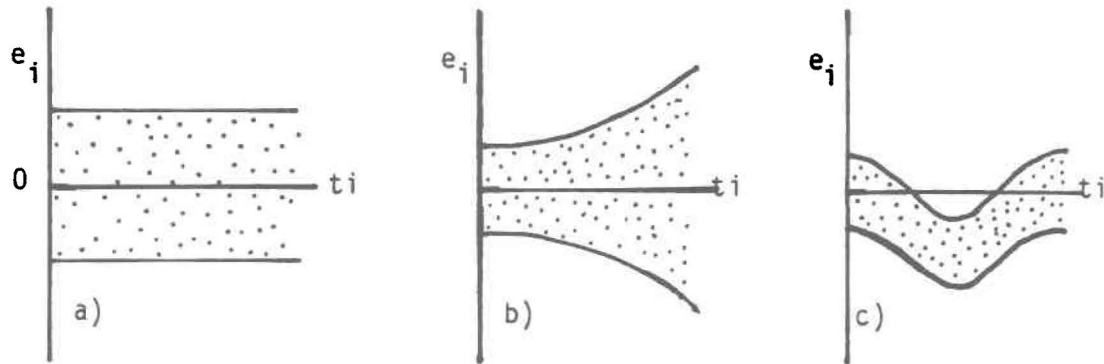


FIGURE 3.1 : Graphe des résidus en fonction du temps.

La figure 3.1a illustre le cas où les e_i suivent une même loi de distribution. La variance autour de zéro est constante et il y a homoscédasticité.

En 3.1b la variance augmente avec le temps, elle n'est pas constante, il y a hétéroscédasticité. Dans ce cas, la première hypothèse base n'est pas respectée.

D'une façon rigoureuse, on peut tester si la valeur absolue des résidus est relié au temps de façon significative par le test de Spearman (Brown, 1975).

La statistique utilisée est:

$$r_s = \frac{1 - 6 \sum_{i=1}^n D_i^2}{N(N^2 - 1)} \quad (\text{éq. 3.3})$$

où $D_i = (i - Y_i)$ avec

i : rang temporel de l'observation y_i ; la première observation a le rang 1 et la dernière, le rang N ;

Y_i : rang de $|e_i|$, la plus petite valeur de $|e_i|$ a le rang 1 et la plus grande, le rang N .

Le tableau 2.2 indique la manière de déterminer les valeurs critiques du test.

Si r_s est différent de zéro pour le niveau de signification choisi, on conclut à l'hétéroscédasticité c'est-à-dire au rejet de l'hypothèse de base d'égalité des variances de e_i . Ce type d'hétéroscédasticité (croissance monotone) n'est qu'une des façons dont la variance peut changer. Dans le cas où elle changerait de façon brusque, on pourrait utiliser un des tests de comparaison de la variance (Marascuilo, 1977). En fait, le test à utiliser dépend du graphique et du type d'hétéroscédasticité soupçonnée.

Dans les cas de périodicité, le graphique des résidus sera d'un troisième type, illustré à la figure 3.1c. La composante déterministe cyclique fait varier la valeur des résidus de l'approximation linéaire de façon non aléatoire. On voit que la distribution des résidus n'est donc pas la même pour tout y_i .

Une périodicité dans les données aura aussi pour effet d'infirmier l'hypothèse de base d'indépendance mutuelle de résidus.

3.1.2 Hypothèse de l'indépendance des erreurs

Si une série ne comporte pas de tendance, on aura:

$$P [y_i < y_j] = P [y_i > y_j] = 1/2$$

où, $i < j$.

Par contre, si il y a une tendance croissante, on remarquera que:

$$P [y_i < y_j] > P [y_i > y_j] \quad \text{pour} \quad i < j$$

Dans ce cas, on constate que l'observation y_j n'est pas indépendante de y_i . Quel que soit y_i , y_j aura tendance à être plus élevé. Les tests de détection de tendance procèdent en mesurant cette dépendance entre les observations.

Si les observations sont corrélées (dépendantes) mais que cette dépendance est due à un phénomène autre qu'une tendance, on comprend que le test pourrait conclure à l'existence d'une tendance alors qu'il n'y en a pas. Pour cette raison, le test pose l'hypothèse de base que les observations sont indépendantes, sauf pour l'effet de la tendance. C'est-à-dire que les résidus e_i obtenus par soustraction de la tendance sont indépendants.

Dans les séries temporelles de variables de la qualité de l'eau, en plus de la tendance, deux phénomènes peuvent créer de la dépendance entre les observations. Il s'agit de la périodicité et de la persistance. Pour que le test ne conclue pas à l'existence d'une tendance lorsqu'il n'y en a pas (erreur de type 1), il faut éliminer l'effet de ces phénomènes. C'est l'objectif des méthodes proposées en 3.3 et 3.4. La section 3.2 présente le sens physique de la périodicité et de la persistance. Les méthodes permettant de vérifier l'indépendance sont présentées ici.

La dépendance mutuelle entre les e_i peut être vérifiée par le test d'Anderson (1942). On calcule le coefficient d'autocorrélation d'ordre 1, c'est-à-dire la corrélation entre un résidu (e_i) et le suivant (e_{i+1}).

Le coefficient r_1 est défini par:

$$r_1 = \frac{(e_1 e_2 + \dots + e_i e_{i+1} + e_n e_1) - \sum_{i=1}^N e_i^2}{\sum_{i=1}^N e_i^2 - (\sum_{i=1}^N e_i)^2 / N} \quad (\text{éq. 3.4})$$

Il varie entre -1 et 1. Si:

$r_1 \approx 1$: il y a autocorrélation positive ou dépendance directe entre e_i et e_{i+1} . Par exemple si e_i est grand, e_{i+1} aura tendance à être grand;

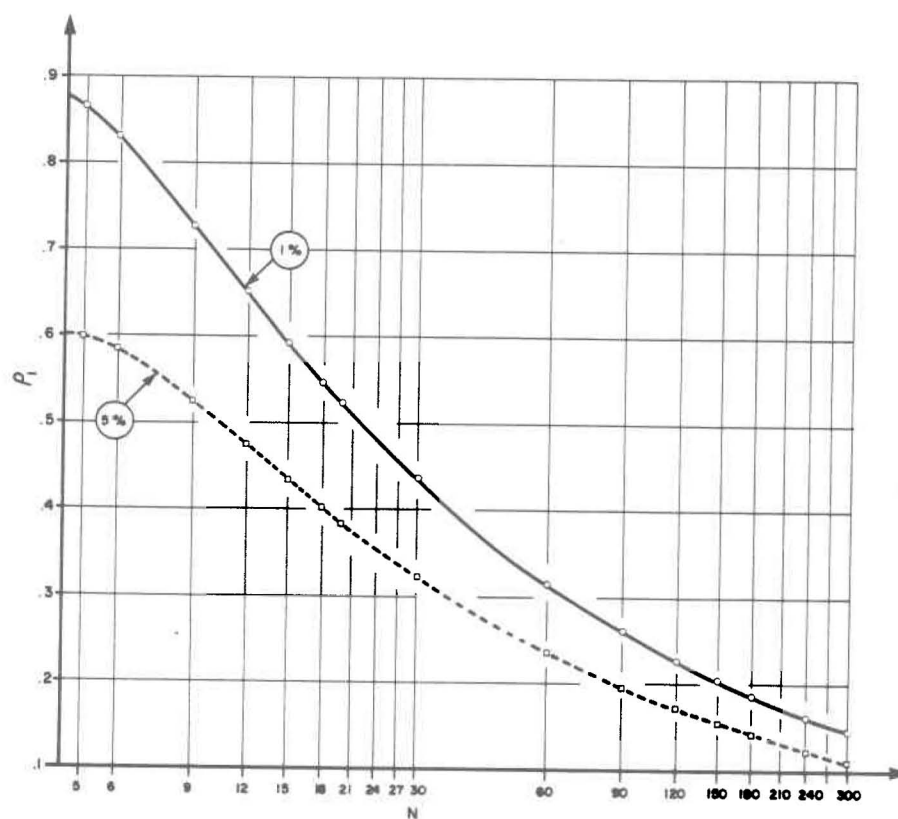


Figure 3.2 : Valeur de r_1 significatives aux seuils 5 et 1 % en fonction de N (tiré de Bobée et Cluis, 1979).

$r_1 \approx 0$: il n'y a pas autocorrélation ou dépendance entre e_i et e_{i+1} ;

$r_1 \approx -1$: il y a autocorrélation négative ou dépendance inverse entre e_i et e_{i+1} . Par exemple si e_i est petit, e_{i+1} aura tendance à être grand.

La figure 3.2 donne les valeurs significatives de r_1 en fonction de la taille de la série et du niveau de signification α . On trouve les formules exactes dans Bobée et Cluis, 1979.

Le test d'Anderson examine l'hypothèse $H_0: \rho_i = 0$ où ρ_1 est le coefficient théorique estimé par r_1 .

Par exemple, pour une série de 60 observations ayant un $r_1 = 0,4$, on dira, au niveau de 1 %, que les e_i sont dépendants, que l'hypothèse de base n'est pas respectée puisque la valeur critique pour $\alpha = 0,01$ donné par la figure 3.2 est $r_1 = 0,32$.

L'étude de la dépendance entre les valeurs de la série sera plus complète si le coefficient d'autocorrélation est calculé pour différents pas de temps. Anderson (1942) a généralisé son test aux coefficients d'autocorrélation d'ordre 2 (entre e_i et e_{i+2}), d'ordre 3 (entre e_i et e_{i+3}) etc. L'équation qui suit donne le coefficient d'autocorrélation pour un pas (ordre) de temps (p) choisi.

$$r_p = \frac{\frac{1}{N-P} \sum_{i=1}^{N-P} e_i e_{i+p} - \frac{1}{(N-P)^2} \sum_{i=1}^{N-P} e_i \sum_{i=1}^{N-P} e_{i+p}}{\left[\frac{1}{N-P} \sum_{i=1}^{N-P} e_i^2 - \frac{1}{(N-P)^2} \left[\sum_{i=1}^{N-P} e_i \right]^2 \right]^{\frac{1}{2}} \left[\frac{1}{N-P} \sum_{i=1}^{N-P} e_{i+p}^2 - \frac{1}{(N-P)^2} \left[\sum_{i=1}^{N-P} e_{i+p} \right]^2 \right]^{\frac{1}{2}}}$$

(éq. 3.5)

(Llamas et Triboulet, 1976)

Pour une série de taille N, plus l'ordre p est grand, moins il y a de valeurs pour calculer le coefficient. Les coefficients pour les ordres élevées sont donc sensibles aux variations aléatoires de la série. Pour cette raison, il est d'usage de limiter le calcul des coefficients aux ordres 1 à p tel que $\frac{p}{N} \leq 0,10$ (Chow, 1964).

Le calcul des coefficients r_p est important dans le cas des séries périodiques. En effet, on comprend que pour une périodicité d'un pas de temps p il y aura corrélation positive entre les observations séparées d'un temps de p (donc $r_p > 0$) et corrélation négative entre les observations séparées d'un temps de P/2.

À moins qu'il n'y ait périodicité, la dépendance entre les observations diminue lorsque augmente le temps qui les sépare. Le coefficient le plus élevé est alors r_1 . Le test doit porter sur le plus grand coefficient, le plus différent de zéro. Si ce coefficient n'est pas r_1 , on trouvera dans Anderson (1942) les procédures pour évaluer

les niveaux significatifs des coefficients autre que r_1 . On appelle corrélogramme, le graphique de r_p en fonction de p .

3.2 Caractéristiques pratiques des séries observées: la persistance et la périodicité

Cette section introduit deux caractéristiques très fréquentes des séries temporelles observées en pratique et ayant pour effet d'infirmier les hypothèses de base (conditions d'application) des tests.

3.2.1 La persistance

Les facteurs, naturels ou non, déterminant la valeur d'une mesure ne sont pas des phénomènes instantanés. Ils agissent selon des échelles de temps très variables. Certains facteurs influençant une variable peuvent être assez stables et font que la valeur a tendance à se maintenir dans le temps. On peut constater cette persistance si la fréquence d'échantillonnage est assez élevée par rapport aux taux de variation de la variable. Dans un tel cas, les valeurs qui se suivent dans la série auront des valeurs voisines. Par exemple, les petites valeurs seront plus fréquemment suivies par une petite valeur que par une grande. Les observations ne sont donc pas indépendantes de celles qui les précèdent. Dans une telle série, il y a autocorrélation ou dépendance entre les observations. Le processus générant cette série est dit autorégressif.

Cette autocorrélation peut-être mesurée par les coefficients d'autocorrélation et la signification de ces coefficients peut être testée. Si le coefficient est significatif c'est que les observations ne sont pas indépendantes et que la deuxième hypothèse de base des tests n'est pas respectée (section 3.1.2).

Dans plusieurs cas, la dépendance (mesurée par r_1 , (3.4) et r_p , (3.5)) diminue avec le temps séparant les observations. Le processus est dit markovien d'ordre 1 si la diminution est telle que :

$$r_p = r_1^p \quad (\text{éq. 3.6})$$

La figure 3.3 illustre la forme du corrélogramme obtenu dans ce cas.

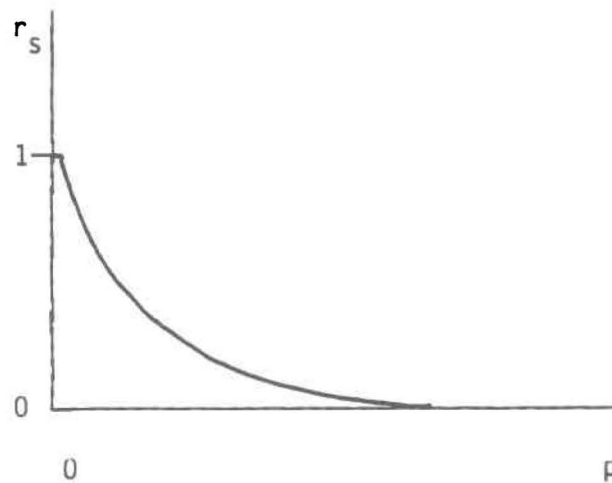


Figure 3.3: Corrélogramme d'un processus autorégressif, Markovien d'ordre 1

Si une observation est corrélée à la précédente, une partie de l'information qu'elle contient était déjà contenue dans l'observation précédente. Ainsi, par extension, on comprend que N observations corrélées entre elles n'apportent pas autant d'information que N obser-

vations indépendantes. C'est pour cette raison que les tests, qui considèrent toujours le nombre N , ne peuvent être utilisés rigoureusement si les observations sont dépendantes.

Une solution à ce problème est de déterminer le nombre d'observations indépendantes N' équivalent aux N observations corrélées en termes de contenu en information. C'est l'approche de Lettenmaier (1976) résumée et commentée en 3.3.1

3.2.2 La périodicité

Plusieurs variables sont reliées à des cycles, par exemple : le cycle des saisons. Ces variables ont un comportement périodique évident. La température de l'eau en est un exemple.

Dans une série de température de l'eau, on constatera que les observations de janvier s'éloignent plus de la moyenne annuelle que celle de mai. Les écarts par rapport à cette moyenne ne proviennent donc pas d'une même loi de distribution. La première hypothèse de base des tests ne sera pas respectée si les résidus qui ont servi à la vérifier sont ceux qui restent d'une approximation linéaire de la tendance de la moyenne. De plus, tel qu'expliqué en 3.1.2, la périodicité créera de l'autocorrélation significative et entraînera le rejet de la deuxième hypothèse de base.

Une façon de contourner ce problème est de diviser la série originale en sous-séries. Chaque sous-série comprend les observations qui sont affectés de la même façon par la composante cyclique. Par exemple, pour une série de données mensuelles d'une durée de cinq ans, on pourrait faire 12 classes correspondant aux douze mois de l'année. Chaque sous-série comprendrait les cinq observations d'un même mois, affectées de la même façon par le cycle annuel. La subdivision pourrait aussi se faire en quatre sous-séries correspondant aux saisons. En fait, le nombre de sous-séries importe peu, l'essentiel étant que les observations d'une sous-série soient affectées de façon homogène par la composante cyclique et qu'ainsi les résidus des observations d'une sous-série proviennent d'une même loi de distribution. La première hypothèse de base est alors respectée à l'intérieur des sous-séries. C'est sur cette approche que la méthode de Hirsch et al. (1982) présentée en 3.3.2 est basée.

3.2.3 La persistance et la périodicité

Plusieurs séries présentent à la fois les phénomènes de persistance et de périodicité. Dans la série des températures, en plus de la périodicité, on remarquera une persistance : les mesures diffèrent peu si elles sont rapprochées dans le temps.

La figure 3.4 illustre le corrélogramme obtenu dans ce cas. Pour des observations mensuelles r_p passe par un maximum (corrélation positive) au pas de temps de 1 an et ses multiples. Ils passe par un mini-

mum (corrélation négative) au pas de $\frac{1}{2}$ an et ses multiples, mais plus les observations sont éloignées, plus l'autocorrélation diminue en valeur absolue.

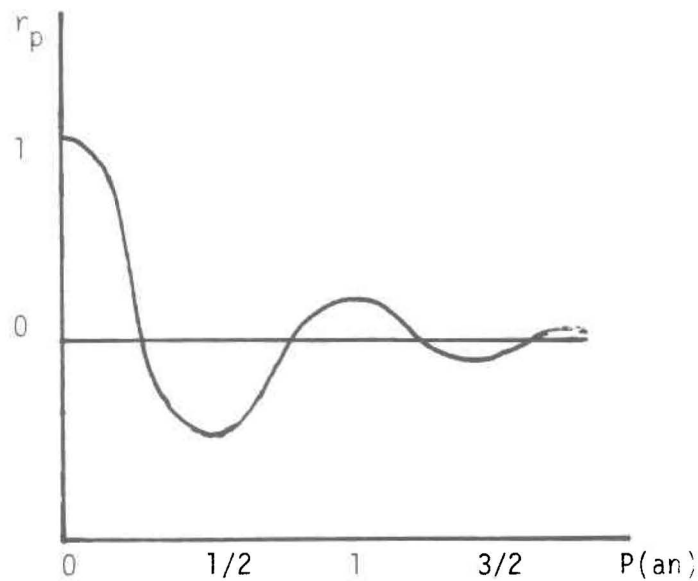


Figure 3.4 : Corrélogramme d'un processus autorégressif avec périodicité

La méthode de Hirsch et al. (1982) basée sur la division de la série en sous-séries homogènes tant qu'à l'effet de la composante cyclique suppose qu'il n'y a pas de dépendance entre les observations sauf pour celle due à la tendance ou à la périodicité. Pour les séries ayant une dépendance résiduelle due à la persistance, les tests, basés sur la mesure de la dépendance, pourraient détecter une tendance alors qu'il n'y en a pas. Hirsch et Slack (1984) ont étendu la méthode de

Hirsch et al. (1982) pour tenir compte de ces séries à périodicité et persistance. Cette méthode est présentée en 3.3.3.

3.3 Adaptation et application des tests aux séries dans le cas des tendances monotones

3.3.1. Séries avec persistance

a) Résumé de la méthode (Lettenmaier, 1976)

La section 3.2.1 se termine en présentant le concept des N' observations non corrélées contenant la même quantité d'information que N observations corrélées.

Pour une série markovienne (voir en B) ce N équivalent à utiliser dans les tests est donné par:

$$\frac{1}{N'} = \frac{1}{N} + \frac{2}{N^2} \frac{r_1^{(N-1)t} - Nr_1^{2t} + (N-1)r_1^t}{(r_1^t - 1)^2} \quad (\text{éq. 3.7})$$

où:

r_1 est le coefficient d'autocorrélation d'ordre 1

N est l'effectif de l'échantillon

N' est le nombre équivalent d'observations non corrélées

t est l'intervalle d'échantillonnage

La puissance d'un test de détection de tendance dépend de l'amplitude de la tendance, de la taille de l'échantillon et de la variance de cet échantillon. La puissance du test est directement proportionnelle aux deux premiers et inversement proportionnelle au troisième, la variabilité ayant pour effet de masquer la tendance. Lettenmaier résume ces trois facteurs en un coefficient : N'_t , dont la définition dépend du type de tendance étudiée. Si les observations sont corrélées et que l'on a calculé N' , on a N'_t :

$$N'_t = \frac{\text{Tr} [N' (N' + 1) (N' - 1)]^{1/2}}{N' (12)^{1/2} s} \quad (\text{éq. 3.8})$$

où $\text{Tr} = \tau N$ et où τ = est la pente de la tendance estimée par:

$$\tau = (\sum \hat{x}_i \hat{y}_i) / (\sum \hat{x}_i^2)$$

avec

$$\hat{x}_i = x_i - \bar{x}_i$$

$$\hat{y}_i = y_i - \bar{y}_i$$

$\bar{x}_i$ et $\bar{y}_i$ sont les moyennes des x et des y

N est la taille de l'échantillon

s est l'écart-type de l'échantillon

Le N'_t calculé, on obtient la puissance du test au seuil $\alpha = 0,05$ par la figure 3.5.

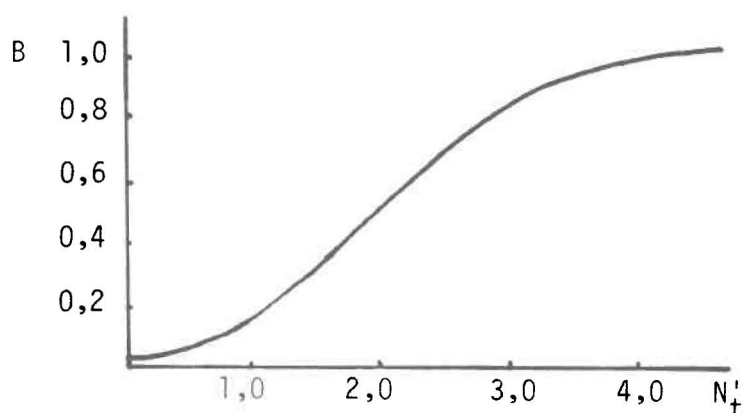


Figure 3.5 : Puissance (β_3) du test de Spearman en fonction de N'_t au seuil $\alpha = 0,05$ (tiré de Lettermaier, 1976).

b) Commentaires

On peut donc évaluer la puissance (au seuil $\alpha = 0,05$) du test de détection des tendances lorsque les données sont corrélées. Il faut cependant remarquer que cette méthode s'applique mal lorsque le processus n'est pas strictement Markovien. En effet, il est démontré que l'estimation du nombre équivalent d'observations indépendantes N' (par la relation 3.8) est très sensible aux déviations du modèle Markovien (Cluis, 1982).

Il devient donc essentiel de vérifier le caractère Markovien d'ordre 1 de la série. Pour faire cette vérification il faut:

- faire une approximation de la tendance et calculer les résidus;
- ajuster un modèle markovien d'ordre 1 à ces résidus;
- calculer les résidus du modèle markovien ajusté et en vérifier la normalité.

Cette vérification allourdie l'application de la méthode. Pour s'assurer du caractère Markovien très strict, il faudrait imposer un niveau α très exigeant pour le test de normalité des résidus. On trouve en 4.3.2 le type de modèle à ajuster et en annexe 2, un test de normalité à utiliser sur les résidus. Ce test, de même que son équivalent paramétrique ne sont pas puissants. L'examen visuel de l'ajustement entre l'histogramme des résidus et une courbe normale ayant les mêmes paramètres serait plus révélateur.

3.3.2 Séries avec périodicité

a) Résumé de la méthode (Hirsch et al. 1982)

Après avoir subdivisé la série de la manière décrite au paragraphe 3.2.2, Hirsch et al. (1982) démontre que l'on peut calculer la statistique du test de Kendall pour chacune des sous-séries, faire la somme de ces statistiques et utiliser cette somme pour tirer la conclusion du test.

La méthode de Hirsch est illustrée dans l'exemple qui suit. La figure 3.5 et le tableau 3.1 donnent les valeurs d'une série fictive de données mensuelles d'une durée de 6 ans. Le coefficient d'autocorrélation d'ordre 1 de la série estimé par $r_1 = 0,62$ (éq. 3.4) est très significativement différent de zéro. La méthode proposée ici ne s'applique pas à ce cas de série à dépendance. Toutefois, la méthode est utilisée sur cette série à titre d'exemple de calcul et pour illustrer, par comparaison avec la méthode utilisée en 3.3.3, l'effet de la

dépendance sur le résultat du test. La série est subdivisée en 12 sous-séries correspondant aux 12 mois.

On admet que pour le même mois K , ($K = 1$ à 12) les observations proviennent d'une même population.

On définit la statistique mensuelle S_k de Kendall par:

$$S_k = \sum_{\substack{i=1 \\ i < j}}^{N_k-1} \sum_{j=2}^{N_k} B [x_i - x_j] \quad (\text{éq. 3.9})$$

où $B [\theta] = +1$ si $\theta > 0$

0 si $\theta = 0$

-1 si $\theta < 0$

$K = 12$ est le nombre de sous-séries

k est l'indice de la sous-série; $k = 1$ à 12 pour le cas d'une subdivision mensuelle

i et j = indices des observations des années i et j dans les sous-séries.

N_k = nombre d'observation dans la sous-série k . Dans cet exemple $N_k = 6$.

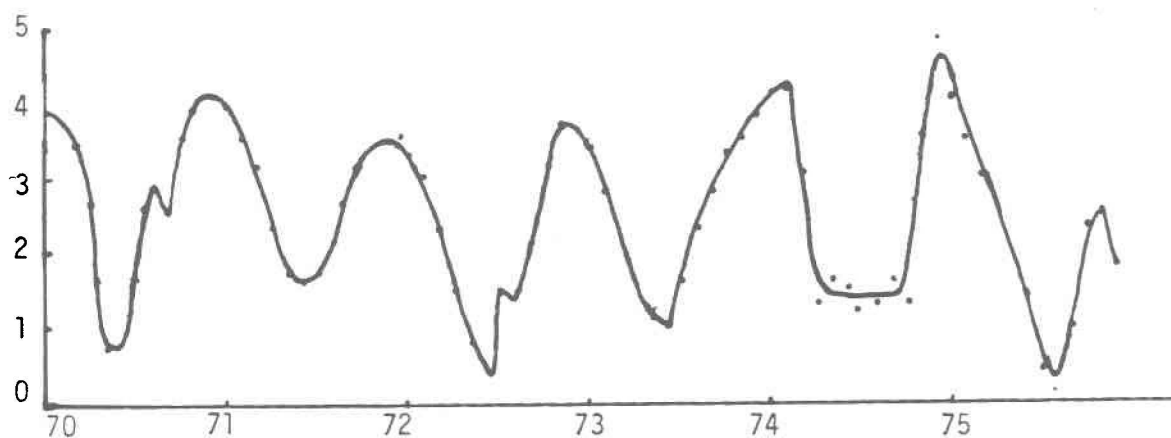
Pour $k = 1$ (janvier) on a les observations:

4,0 - 3,9 - 3,3 - 3,4 - 4,0 - 4,0. Pour

Tableau 3.1 : Données mensuelles fictives sur six ans.

	JAN.	FEV.	MAR.	AVR.	MAI	JUI.	JUIL.	AOU.	SEPT.	OCT.	NOV.	DÉC.
k an	1	2	3	4	5	6	7	8	9	10	11	12
70	4,0	3,8	3,4	2,6	0,7	0,7	1,1	2,8	2,5	3,5	4,0	4,1
71	3,9	3,5	3,1	2,3	1,7	1,6	1,7	2,1	3,0	3,4	3,4	3,5
72	3,3	3,0	2,3	1,5	0,8	0,4	1,5	1,4	2,1	2,7	3,7	4,5
73	3,4	2,8	3,0	1,4	1,1	1,0	1,6	2,3	2,8	3,3	3,5	3,7
74	4,0	4,2	3,0	1,3	1,6	1,5	1,2	1,3	1,6	1,3	3,0	4,7
75	4,0	3,5	3,0	2,5	2,	1,4	0,4	0,2	1,0	2,8	3,0	1,8

Figure 3.6 Graphique des données mensuelles fictives



$i = 1$ et $j = 2$, on a $(4,0 - 3,9) > 0$, donc $B(\theta)$ vaut 1

" " " $j = 3$, on a $(4,0 - 3,3) > 0$, donc $B(\theta)$ vaut 1

" " " $j = 4$, on a $(4,0 - 3,4) > 0$, donc $B(\theta)$ vaut 1

" " " $j = 5$, on a $(4,0 - 4,0) = 0$, donc $B(\theta)$ vaut 0

" " " $j = 6$, on a $(4,0 - 4,0) = 0$, donc $B(\theta)$ vaut 0

$i = 2$ et $j = 3$, on a $(3,9 - 3,3) > 0$, donc $B(\theta)$ vaut 1

" " " $j = 4$, on a $(3,9 - 3,4) > 0$, donc $B(\theta)$ vaut 1

" " " $j = 5$, on a $(3,9 - 4,0) < 0$, donc $B(\theta)$ vaut -1

" " " $j = 6$, on a $(4,0 - 4,0) < 0$, donc $B(\theta)$ vaut -1

$i = 3$ et $j = 4$, on a $(3,3 - 3,4) < 0$, donc $B(\theta)$ vaut -1

" " " $j = 5$, on a $(3,3 - 4,0) < 0$, donc $B(\theta)$ vaut -1

" " " $j = 6$, on a $(3,3 - 4,0) < 0$, donc $B(\theta)$ vaut -1

$i = 4$ et $j = 5$, on a $(3,4 - 4,0) < 0$, donc $B(\theta)$ vaut -1

" " " $j = 6$, on a $(3,4 - 4,0) < 0$, donc $B(\theta)$ vaut -1

$i = 5$ et $j = 6$, on a $(4,0 - 4,0) = 0$, donc $B(\theta)$ vaut 0

-2

Donc pour janvier ($k = 1$): $S_k = S_1 = -2$.

Le tableau 3.2 donne les valeurs de S_k pour tous les mois, calculées de la même façon

k	1	2	3	4	5	6	7	8	9	10	11	12	$\sum S_k = S$
S_k	-2	+2	+6	+3	+9	-3	+3	+11	+9	+11	+4	-3	+31

Tableau 3.2 : Valeur de la statistique de Kendall pour chaque mois et valeur totale

$$\text{La statistique totale} = S = \sum_{k=1}^K S_k = + 31 \quad (\text{éq. 3.10})$$

Pour calculer la probabilité que $S = 31$ lorsqu'il n'y a pas de tendance on utilise une approximation de la loi normale. On se sert de l'espérance mathématique de S définie par:

$$E(S) = \sum_{k=1}^{K=12} E(S_k) = 0 \quad (\text{éq. 3.11})$$

puisque $E(S_k) = 0$

- La variance de S est telle que:

$$\text{VAR}(S) = \sum_{k=1}^{K=12} \text{Var}(S_k) + \sum_{g=1}^{K=12} \sum_{\substack{k=1 \\ g \neq k}}^{K=12} \text{Cov}(S_g, S_k) \quad (\text{éq. 3.12})$$

Dans les séries où il y a indépendance des observations, la statistique calculée pour un mois sera indépendante de celle des autres mois. Donc, le terme $\text{Cov} (S_k, S_g) = 0$, la variance de S se réduit au premier terme de l'équation.

$$\text{Var} (S) = \sum_{k=1}^{K=12} \text{Var} (S_k) \quad (\text{éq. 3.13})$$

avec:

$$\text{Var} (S_k) = \frac{N_k (N_k - 1) (2N_k + 5) - \sum^T (t_i - 1) (2t_i - 5)}{18} \quad (\text{éq. 3.14})$$

où

N_k est l'effectif de la sous-série k

t_i est le nombre d'observation dans chaque cas d'égalité

T est le nombre de cas d'égalité

S'il n'y a pas de valeurs égales:

$$\text{VAR} (S_k) = \frac{N_k (N_k - 1) (2 N_k + 5)}{18} \quad (\text{éq. 3.15})$$

Par exemple, pour janvier ($k = 1$), il y a trois observations de valeur 4,0. On a donc $N_k = 6$ et $t_i = 3$

$$\text{Var}(S_1) = \frac{6(6-1)(2-6+5) - (3-1)[(2 \cdot 3) + 5]}{18} = 24,7$$

En calculant de la même façon pour les autres mois on obtient les résultats du tableau 3.3.

k	1	2	3	4	5	6	7	8	9	10	11	12	Σ
$\text{VAR}(S_k)$	24,7	27,8	28,3	28,3	28,3	28,3	28,3	28,3	28,3	28,3	27,8	28,3	335

Tableau 3.3 : Variance des S_k

Donc, d'après l'équation (3.13), on obtient:

$$\text{Var}(S) = 335$$

En corrigeant pour la continuité, on obtient la variable Z normale centrée et réduite.

$$Z = \frac{S - 1}{\text{Var}(S)^{1/2}} \quad \text{si } S > 0$$

$$Z = 0 \quad \text{si } S = 0 \quad (\text{éq. 3.16})$$

$$Z = \frac{S + 1}{\text{VAR}(S)^{1/2}} \quad \text{si } S < 0$$

Dans cet exemple on obtient:

$$Z = \frac{31 - 1}{\sqrt{335}} = 1,63$$

La table de la loi normale centrée, réduite donne $P[Z > 1,63] = 5,16 \%$. On peut donc rejeter l'hypothèse de l'absence d'une tendance, à 5,16% d'erreur. S est positif ce qui veut dire que les valeurs du début de la série sont supérieures à celles de la fin. La tendance est donc décroissante. On verra que la méthode de Hirsch et Slack (1983), tenant compte de l'autocorrélation n'arrive pas à la même conclusion.

Remarque: L'examen du tableau 3.2 montre que la valeur de S_k des mois 8, 9 et 10 contribuent largement à la valeur de S . En refaisant le calcul après avoir éliminer ces mois on a S' .

$$S' = \sum_{k=1,11}^{7,12} S_k = 1$$

et

$$\text{VAR} (S') = 250,1$$

On a alors

$$Z' = \frac{1 - 1}{\sqrt{250,1}} = 0$$

Dans ce cas, il n'y a pas de tendance. La tendance décelée au premier examen serait due à l'effet des mois d'août, septembre et octobre.

b) Commentaires

Les résultats de Hirsch et al. (1982) démontrent que cette méthode est robuste dans les cas de périodicité et de non-normalité des séries. Cependant, elle ne peut être utilisée que sur les séries dont la dépendance résiduelle, celle obtenue après avoir enlevé la tendance et la périodicité est nulle. Il ne doit y avoir de dépendance autre que celle due à la périodicité ou à la tendance.

Cette méthode est sensible, même à des niveaux de dépendance résiduelle considérées non-significatifs par le test d'Anderson (1942). Par exemple, pour des données mensuelles sans tendance observées durant 10 ans ($N = 120$), le seuil significatif de r_1 à 5 % est 0,17. Hors, à

$r_1 = 0,17$ et $N = 120$, les simulations de Hirsch et Slack (1984) démontrent que le niveau α du test, fixé à 0,05, passe en fait à 0,082, au-delà de l'intervalle de confiance autour de α : 0,04 à 0,06.

On en conclut que :

- si r_1 (éq. 3.4) calculé après avoir retiré la tendance, est significatif au niveau de 5 % (fig. 3.2), on ne peut utiliser la méthode proposée;
- il y a une zone grise où le test d'Anderson (1982) ne détecte pas un r_1 significatif mais où la méthode exposée ici ne peut être appliquée de façon exacte. Pour $N = 120$, cette zone est de $r_1 = 0,10$ à 0,17.
- pour des séries dont la tendance a été retirée de $N < 120$, $r_1 < 0,10$ est un seuil conservateur et la méthode peut être utilisée.

On utilise le coefficient r_1 pour vérifier l'applicabilité de la méthode car s'il n'y a pas de persistance, r_1 ne devrait pas être différent de zéro, alors que le r_p correspondant au pas p du cycle pourrait être différent de zéro.

En terme de puissance, cette méthode est supérieure aux équivalents paramétriques avec ou sans correction pour la périodicité sauf si la série est courte, normale et qu'il n'y a pas de périodicité. L'utilisation du test de Kendall, tel que proposé ici, peut donc être géné-

ralisée car le cas décrit plus haut est rare et parce que la méthode de Kendall peut être supérieure aux autres tests même pour des cas de non-normalité non détectable par les méthodes usuelles.

La méthode arrive à détecter des tendances non visibles par un examen visuel des graphique de données (Hirsch et al., 1982).

3.3.3 Séries avec persistance et périodicité

a) Résumé de la méthode (Hirsch et Slack, 1984)

Lorsque entre les observations il y a de la dépendance autre que celle due à la périodicité ou à la tendance, c'est-à-dire la persistance, la méthode exposée en 3.3.2 ne peut être utilisée. Hirsch et Slack (1984) proposent un test identique sauf pour l'expression de la variance (éq. 3.12). Dans le cas d'une série sans tendance d'observations indépendantes, le terme $\text{Cov}(S_g, S_k)$ est nul pour toutes les sous-séries g, k . Lorsqu'il y a dépendance entre les observations les statistiques calculées pour les sous-séries ne sont plus indépendantes, donc $\text{Cov}(S_g, S_k) \neq 0$ et ce terme doit être estimé.

Hirsch et Slack (1984) se basent sur une formule non-paramétrique pour estimer cette covariance. Elle est tirée de Dietz et Killen (1981).

$$\text{Cov}(S_g, S_k) = K_{gk} / 3 + (N^3 - N) r_{gk} / 9 \quad (\text{éq. 3.17})$$

où :

N : effectif des sous-séries g et k ;

K_{gk} : coefficient de corrélation de Kendall entre les sous-séries g et k

r_{gk} : coefficient de corrélation de Spearman entre les sous-séries g et k ;

$$K_{gk} = \frac{1}{N} \sum_{i < j} \text{sign} [(y_{jg} - y_{ig}) (y_{jk} - y_{ik})] \quad (\text{éq. 3.18})$$

où :

$\text{sign} [A] : + 1 \quad \text{si } A > 0;$

$\text{sign} [A] : 0 \quad \text{si } A = 0;$

$\text{sign} [A] : - 1 \quad \text{si } A < 0.$

$$r_{gk} = \frac{1/6 (N^3 - N) \sum_{i=1}^N d_i^2 - T - U}{\sqrt{1/6 (N^3 - N) - 2T} \sqrt{1/6 (N^3 - N) - 2U}} \quad (\text{éq. 3.19})$$

d_i : est la différence entre les rangs dans les deux sous-séries g et k pour les deux observations de l'année i ;

$$T = \frac{1}{12} \sum_{i=1}^T (t_i^3 - t_i)$$

et

$$U = \frac{1}{12} \sum^U (u_i^3 - u_i)$$

t_i et u_i : nombres d'observations dans chaque cas d'égalité dans les sous-séries g et k, respectivement;

N : effectif des deux sous-séries.

Exemple :

Calcul de la covariance entre les sous-séries 1(indice g) et 2(indice h) de la série au tableau 3.1. Ces données et leur rang sont reportés au tableau 3.4.

Note: pour les valeurs y_i égales on donne le rang Y_i moyen. Par exemple, pour $k = 1$, les rangs 4, 5 et 6 sont égaux, on leur donne la valeur $\frac{4 + 5 + 6}{3} = 5$.

Mois	1 = g		2 = k	
Année	y_i	Y_i	y_i	Y_i
1	4,0	5	3,8	5
2	3,9	3	3,5	3,5
3	3,3	1	3,0	2
4	3,4	2	2,6	1
5	4,0	5	4,2	6
6	4,0	5	3,5	3,5

Tableau 3.4 : Données brutes et leurs rangs, mois 1 et 2

Calcul de K_{gh}

La fonction,

$$\text{Sign} [(y_{jg} - y_{ig}) (y_{jk} - y_{ik})]_{i < j}$$

donnera :

- + 1 si $(y_{jg} - y_{ig})$ et $(y_{jk} - y_{ik})$ sont de même signe;
- 1 si $(y_{jg} - y_{ig}) > 0$ et $(y_{jk} - y_{ik})$ sont de signe contraire;
- 0 si $(y_{jg} - y_{ig})$ et / ou $(y_{jk} - y_{ik})$ sont nuls.

Dans notre exemple on obtient donc :

Ind. des obs.	$(y_{jg} - y_{ig})$	$(y_{jk} - y_{ik})$	sign [A]
i = 1 et j = 2	< 0	< 0	+ 1
3	< 0	< 0	+ 1
4	< 0	< 0	+ 1
5	0	> 0	0
6	0	> 0	0
i = 2 et j = 3	< 0	< 0	+ 1
4	< 0	< 0	+ 1
5	> 0	> 0	+ 1
6	> 0	> 0	+ 1
i = 3 et j = 4	> 0	< 0	- 1
5	> 0	> 0	+ 1
6	> 0	> 0	+ 1
i = 4 et j = 5	> 0	> 0	+ 1
6	> 0	> 0	+ 1
i = 5 et j = 6	0	< 0	0

$$\sum = 10$$

Pour $g = 1$ et $k = 2$, on a :

$$\sum_{\substack{i=1 \\ j>i}}^{N=6} \text{sign} [(y_{jg} - y_{ig}) (y_{jk} - y_{ik})] = 10$$

Calcul de r_{gh}

À partir des données du tableau 3.4 et de l'équation 3.14, on calcule:

$$\sum_{i=1}^{N=6} d_i^2 = (5-5)^2 + (3-3,5)^2 + (1-2)^2 + (2-1)^2 + (5-6)^2 + (5-3,5)^2 = 5,5$$

$$U : 1/12 (3^3 - 3) = 2,0$$

$$T : 1/12 (2^3 - 2) = 0,5$$

$$r_{gk} = \frac{1/6(6^3 - 6) - 5,5 - 0,5 - 2,0}{\sqrt{1/6 (6^3 - 6) - 2 \cdot 0,5} \sqrt{1/6 (6^3 - 6) - 2 \cdot 2,0}} = 0,832$$

Par l'équation 3.17 on obtient:

$$\text{Cov} (S_1, S_2) = \frac{10}{3} + (6^3 - 6) \frac{0,832}{9} = 22,75$$

Pour calculer le terme :

$$\sum_{\substack{g=1 \\ g \neq k}}^{K=12} \sum_{k=1}^{K=12} \text{Cov} (S_g, S_k)$$

de l'équation de la variance (3.12), il faut faire le même calcul pour toutes les combinaisons des 12 mois, soit 66 fois. Pour simplifier, posons la covariance entre toutes les autres combinaisons de mois à la moitié de celle calculée ici pour le mois 1 et 2. On peut s'attendre à ce que la covariance réelle soit supérieure à cet estimé. On aura donc :

$$\sum_{\substack{g=1 \\ g \neq k}}^{K=12} \sum_{k=1}^{K=12} \text{Cov} (S_g, S_k) \approx 22,75 + 65 (11,4) \approx 762,0$$

La variance (éq. 3.12) est obtenue en ajoutant au résultat obtenu en 3.3.2 (cas où $\text{Cov}(S_g, S_k) = 0$) la somme des covariances qui vient d'être calculée:

$$\text{Var}(S) \approx 335 + 762,0 \approx 1097,0$$

On calcule la valeur de Z par l'équation 3.16 :

$$Z = \frac{31 - 1}{\sqrt{1097,0}} \approx 0,97$$

0,97 est inférieur à 1,65, qui représente le seuil pour rejeter l'hypothèse nulle de l'inexistence d'une tendance au niveau de signification $\alpha = 5\%$. On conclut qu'il n'y a pas de tendance. Avec la covariance réelle, supérieure à cet estimé conservateur, on aurait obtenu une variance supérieure donc un Z encore plus petit. Le non-rejet de H_0 aurait donc été confirmé.

Pour la même série, les deux méthodes arrivent à des résultats nettement différents. Par la procédure 3.3.2, H_0 est tout juste acceptable ($Z = 1,63$ et $\alpha = 0,0516$). L'acceptation de H_0 est beaucoup plus sûre par la méthode avec estimé de la covariance ($Z = 0,91$ et $\alpha = 0,18$). Ceci est en accord avec les résultats de Hirsch et Slack 1984 qui démontrent que pour les séries avec persistance, l'application de la méthode sans estimé de la covariance implique une augmentation de l'erreur de type 1 (rejeter H_0 alors qu'elle est vraie).

b) Commentaires

L'approximation par la loi normale dans cette méthode est bonne pour $N = 10$, mais non à $N = 5$. À $N = 6$ (cas de l'exemple) le test n'est pas tout à fait exact au seuil $\alpha = 0,5$. Le seuil réel est $\alpha = 0,02$. L'autocorrélation résiduelle (série sans tendance) ne doit pas être trop forte, les valeurs limites données par Hirsch et Slack (1984) sont: $r_1 < 0,65$ pour $N = 10$ et $r_1 < 0,5$ pour $N = 20$.

Lorsque les observations sont indépendantes, la procédure suppose que $\text{Cov}(S_g, S_k) = 0$ pour tout g et k est d'une puissance supérieure. Cette supériorité est de 0 à 25 % dépendamment du rapport tendance sur écart-type (Hirsch et Slack, 1984).

La formulation de r_{gk} (éq. 3.19) implique que les sous-séries g et k aient le même nombre d'observations. Dans le cas de données manquantes dans une série, on met de côté, pour ce calcul de r_{gk} , la valeur correspondante dans l'autre sous-série.

Lorsqu'il y a plusieurs valeurs égales ou situées en deça d'une limite de détection, la formule 3.17 surestime la covariance.

On utilise alors plutôt la relation:

$$\text{Cov}(S_g, S_k) = K_{gk} / + \frac{4}{3} T_{gk} - \frac{N(N+1)^3}{9} \quad (\text{éq. 3.20})$$

où

$$T_{g_k} = \sum_{i=1}^N (Y_{ig} \cdot Y_{ik})$$

Pour $Cov(S_1, S_2)$, on obtient 21,3 par la relation 3.20 et 22,7 par la relation 3.19.

3.4 Adaptation et application des tests aux séries dans le cas des tendances en saut

L'adaptation du test de Mann-Whitney aux séries à composante cyclique n'a pas été effectuée. Ce chapitre propose l'extension des méthodes de Hirsch et al. (1982) et de Hirsch et Slack (1984) au test de Mann-Whitney. La formulation des tests est proposée et les limites théoriques dues à la taille de la série ont été déterminées. Par contre, la vérification de la robustesse (respect du niveau α) et de la puissance du test nécessitent un grand nombre de simulations et n'ont pu être effectuées dans le cadre de cette étude. Pour cette raison, les adaptations présentées en 3.4.2 et 3.4.3 doivent être vues comme préliminaires et indicatives, et elles ne peuvent être considérées comme des résultats rigoureusement applicables en pratique.

3.4.1 Séries avec persistance

a) Résumé de la méthode (Lettenmaier, 1976)

On évalue le nombre d'observations indépendantes N' équivalent au nombre d'observations corrélées N par la même formule (3.7) que pour les tendances monotones. Ce nombre N' calculé, on peut calculer le rapport N'_t :

$$N'_t = \frac{T_r \sqrt{N'}}{2 S} \quad (\text{éq. 3.21})$$

où

S : écart-type de l'échantillon;

N' : nombre équivalent d'observations indépendantes;

T_r : fonction de la tendance = $\bar{y}_1 - \bar{y}_2$;

$\bar{y}_1, \bar{y}_2$: moyennes avant et après le saut.

Avec N'_t on obtient la puissance du test pour $\alpha = 0,05$, par la figure 3.7.

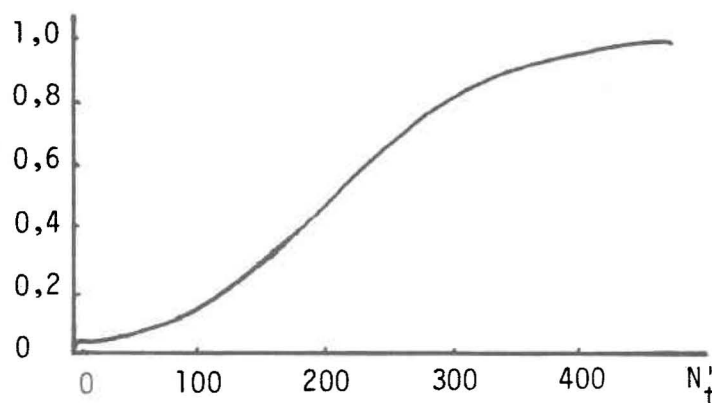


Figure 3.7 : Puissance du test de Mann-Whitney en fonction de N'_t à $\alpha = 0,05$ (tiré de Lettenmaier, 1976)

b) Commentaires

La méthode étant basée sur le nombre équivalent d'observations indépendantes N' , les considérations de la section 3.3.1 B sur la précision du modèle markovien, s'appliquent ici aussi.

3.4.2 Séries avec périodicité

a) Adaptation de la méthode de Hirsch et al., 1982

Encore une fois, les observations sont divisées en sous-séries. On définit les statistiques de la sous série par 3.22 :

$$U_k = \sum_{i=1}^{N_k} R_i - \frac{N_1 (N_1 + 1)}{2} \quad (\text{éq. 3.22})$$

où N_k est l'effectif de la sous-série k , avant le saut;

R_i est le nombre d'observations après le saut, supérieures à l'observation i avant le saut;

N.B. : Si N_k , le nombre d'observations après le saut est inférieur à N_1 , on remplace N_1 par N_2 dans la formule.

On calcul ainsi la statistique de chacune des sous-séries.

La moyenne et la variance de la statistique de la sous-série d'indices K sont données par:

$$E(U_k) = \frac{N_1 N_2}{2} \quad (\text{éq. 3.23})$$

$$\text{Var}(U_k) = \frac{N_1 N_2 (N_1 + N_2 + 1)}{12} \quad (\text{éq. 3.24})$$

S'il y a des égalités entre les observations, on a:

$$\text{Var}(U_k) = \frac{N_1 N_2 (N_1 + N_2 + 1)}{12} \left(1 - \sum_{s=1}^S \frac{t_s^3 - t_s}{N^3 - N} \right) \quad (\text{éq. 3.25})$$

où S est le nombre de cas d'égalité;

t_s est le nombre d'observations dans un cas d'égalité s .

Pour la statistique totale, on a :

$$U = \sum_{k=1}^K U_k \quad (\text{éq. 3.26})$$

$$E(U) = \sum_{k=1}^K E(U_k) \quad (\text{éq. 3.27})$$

$$\text{Var}(U) = \sum_{k=1}^K \text{Var}(U_k) + \sum_{g=1}^K \sum_{\substack{k=1 \\ g \neq k}}^K \text{Cov}(U_g, U_k) \quad (\text{éq. 3.28})$$

Dans ce cas comme en 3.3.2, l'indépendance des observations, implique que le terme de la covariance est nul. Donc :

$$\text{Var}(U) = \sum_{k=1}^K \text{Var}(U_k) \quad (\text{éq. 3.29})$$

Pour le test de Mann-Whitney, on a une approximation par la loi normale corrigée par continuité :

$$Z = \frac{U - \frac{1}{2} - E(U)}{\text{Var}(U)^{\frac{1}{2}}} \quad \text{si } U > E(U)$$

$$Z = 0 \quad \text{si } U = E(U) \quad (\text{éq. 3.30})$$

$$Z = \frac{U + \frac{1}{2} - E(U)}{\text{Var}(U)^{\frac{1}{2}}} \quad \text{si } U < E(U)$$

b) Commentaires

Le calcul de la distribution exacte de U sous l'hypothèse de l'absence d'une tendance démontre (Annexe 3) que l'approximation de la loi normale (3.30) est satisfaisante pour un nombre minimum d'observations par sous-séries :

$$N_1 = N_2 = 1 \text{ lorsque } K = 12$$

$$N_1 = N_2 = 2 \text{ lorsque } K = 4$$

$$N_1 = N_2 = 3 \text{ lorsque } K = 2$$

À l'aide de simulations, il reste à vérifier la robustesse (maintien de α) face à la longueur, la non-normalité, la périodicité et la dépendance comme cela a été fait pour le test de Kendall (Hirsch et al., 1982).

De plus, la puissance de cette procédure pourrait être comparée à celle du test de Student avec et sans désaisonnalisation en fonction de la distribution des observations.

Étant donné le problème fondamental du contenu en information lorsque les mesures sont corrélées, on peut s'attendre à ce que ce test ne soit pas exact (augmentation de α) lorsqu'il y a dépendance.

3.4.3 Séries avec persistance et périodicité

a) Adaptation de la méthode de Hirsch et Slack (1984)

Dans le cas de séries où il y a dépendance, le terme :

$$\sum_{g=1}^K \sum_{\substack{k=1 \\ g \neq k}}^K \text{Cov} (U_g, U_k)$$

de la formule 3.28, n'est pas nul et doit être estimé. L'équation permettant d'estimer cette variance est la même que celle utilisée en 3.3.3.

Le saut a pour effet d'accroître les valeurs de K_{gh} et r_{gh} de cette équation. Pour annuler cet effet, il faut enlever l'effet du saut. On calcule la moyenne avant et après le saut et on soustrait cette moyenne aux observations correspondantes. Les résidus obtenus ont alors une moyenne nulle. C'est à ces résidus que l'on donne les rangs pour le calcul de K_{gh} et r_{gh} . La transformation :

$$y'_{i1} = y_{i1} - \bar{y}_1$$

(éq. 3.33)

$$y'_{i2} = y_{i2} - \bar{y}_2$$

où :

$\bar{y}_1$ et $\bar{y}_2$ = moyenne avant et après le saut;

y_{i1} et y_{i2} = observation avant et après le saut;

y'_{i1} et y'_{i2} = résidus de moyenne 0.

B. Commentaires

Il resterait à évaluer la longueur minimale des séries et la dépendance maximale pour laquelle cette procédure peut s'appliquer avec robustesse. La puissance pourrait être comparée à celle de la méthode précédente (3.4.2) dans les cas de séries indépendantes.

Les mêmes considérations qu'en 3.3.3 B concernant les données manquantes, tronquées et les cas d'égalités s'appliquent ici. La formule 3.20 doit alors être utilisée avec les rangs obtenus après la soustraction de la moyenne, comme en 3.4.3 A.

3.5 Compléments

3.5.1 Cas à plus d'une observation par période

Si la série des données comprend plus d'une observation par période, la méthode proposée doit être modifiée. Prenons le cas d'une série de 2 observations par mois durant 6 ans. Les sous-séries mensuelles sont maintenant composées de 12 observations plutôt que 6, et le temps séparant ces observations n'est pas constant. Il y a 2 semaines entre la première et la deuxième, près d'un an entre la deuxième et la troisième, 2 semaines entre la troisième et la quatrième et ainsi de suite.

Dans ce cas, la procédure de calcul doit être modifiée. On peut utiliser deux approches :

- 1) Hirsch et al. (1982) proposent d'utiliser la formule de Kendall pour les cas d'égalités. Les données d'un même mois sont considérées comme un cas d'égalité dans le temps.

Au lieu de $\tau = \frac{S}{\frac{1}{2} N (N - 1)}$ on utilise :

$$\tau = \frac{S}{\sqrt{\frac{1}{2} N (N - 1) - T_x} \sqrt{\frac{1}{2} N (N - 1) - T_y}}$$

avec:

$$T_x = \sum t_x (t_x - 1)$$

où, t_x est le nombre d'observations dans chaque cas d'égalité de la série du temps.

$$T_y = \sum T_y (t_y - 1)$$

où, t_y est le nombre d'observations de la série du paramètre mesuré.

Pour chaque mois, on a deux valeurs "égales" par leur position dans le temps. On a donc $t_x = 2$, et ce pour les 12 mois de chaque année pendant 6 ans :

$$T_x = 6 \sum_{i=1}^{12} 2 (2 - 1) = 144$$

Les méthodes de Hirsch, en composant des sous-séries avec un pas d'un an entre les observations à l'avantage de remplacer la série entière par plusieurs sous-séries à l'intérieur desquelles il y a peu d'autocorrélation. Si les sous-séries comprennent des observations séparées de quelques semaines seulement, l'autocorrélation des sous-séries peut demeurer assez élevé. La signification et la puissance des test n'ont pas été étudiés dans ce cas.

- 2) Une méthode alternative consiste à n'utiliser que l'observation la plus près du centre du mois (ou de la période) qu'elle représente (Lettenmaier, 1982).

La médiane des deux observations mensuelles peut être utilisée pour les graphiques de la valeur du paramètre en fonction du temps; mais pour les tests, cette méthode pose le problème de l'égalité de la variance de ces médianes (Lettenmaier, 1982).

3.5.2 Effet d'une variable externe

Lorsqu'une tendance a été détectée, on peut être intéressé à vérifier si elle est due à l'influence d'une troisième variable. Dans l'étude de la qualité de l'eau en rivière par exemple, plusieurs variables sont influencées par le débit. Le débit étant variable, il est important de vérifier si la tendance détectée n'est pas due à une variation du débit.

Une première approche, les fonctions double-masse, permet de chercher les ruptures dans la relation débit-variable de qualité. Ces ruptures représentent des effets dus à d'autres variables que le débit.

Une deuxième méthode basée sur le calcul des résidus de la régression entre le débit et la variable de qualité permet de considérer une nouvelle variable où l'effet de débit est éliminé.

a) Fonction double-masse

La fonction double-masse est une méthode permettant de détecter des changements dans le rapport entre deux variables très corrélées. Cette méthode est simple et efficace mais n'est pas objective. Le résultat du test dépend de l'examen visuel d'un graphique au lieu de l'application objective d'une règle de décision.

Elle est limitée aux cas où il y a une relation directement proportionnelle significative entre les deux variables et nécessite que les observations de la variable explicative proviennent d'une même population (Cluis, 1981). Dans l'exemple de la relation variable de qualité - débit, étant donné la 1^{ère} condition, ces méthodes ne pourront être appliquées si les sources ponctuelles ont un effet important sur le paramètre de qualité. On trouve dans Cluis (1981) un exemple intéressant de l'utilisation des fonctions double-masse dans la recherche d'une modification de la relation débit-qualité. On y utilise les débits massiques plutôt que les concentrations.

b) Calcul des résidus

Hirsch et al. (1982) proposent de soustraire aux observations l'effet de débit et de vérifier s'il y a une tendance dans les résidus. Ces auteurs considèrent d'abord des fonctions de la variable de qualité (y) en fonction du débit (Q):

$$\hat{y} = \lambda_1 + \frac{\lambda_2}{Q} + \varepsilon \quad (\text{éq. 3.34})$$

ou encore

$$\hat{y} = \lambda_1 + \frac{\lambda_2}{1 + \lambda_3 Q} + \varepsilon \quad (\text{éq. 3.35})$$

pour les cas où l'augmentation du débit a un effet de dilution; et :

$$\hat{y} = \lambda_1 + \lambda_2 Q + \lambda_3 Q^2 + \varepsilon \quad (\text{éq. 3.36})$$

pour les cas où l'augmentation du débit a un effet de lessivage.

λ_i = constante

Q = débit

Y = variable de qualité estimée

ε = variable aléatoire, normale et centrée.

Il faut ajuster aux observations le modèle qui décrit le mieux le phénomène. Les estimés de la valeur de la variable y_{ij} sont calculés à partir des débits mesurés Q_{ij} . Les résidus sont calculés par :

$$y'_{ij} = y_{ij} - \hat{y}_{ij} \quad (\text{éq. 3.37})$$

y'_{ij} = résidu sans l'effet de débit

y_{ij} = mesure de la variable

y_{ij} = composante de y_{ij} expliquée par le débit

Les tests de détection sont appliquées aux y'_{ij} .

4. PROCÉDURE DE DÉTECTION DES TENDANCES MONOTONES

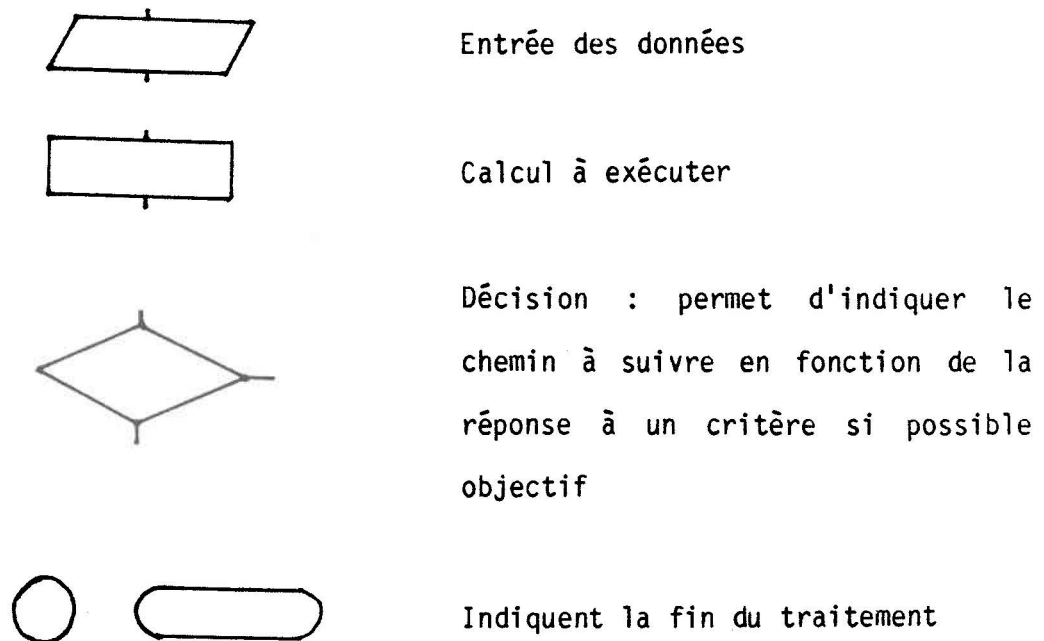
4.1 Présentation

Ce chapitre est une proposition de procédure globale de détection des tendances monotones dans les séries temporelles. Cet outil a été structuré à partir d'une synthèse des méthodes proposées dans la littérature et décrites dans ce travail. Ces méthodes ont été mises au point pour traiter des cas réels. La procédure proposée devrait donc s'appliquer à plusieurs types de séries rencontrées dans la réalité. Par contre, ce système n'a pas été éprouvé comme tel, et il serait intéressant de le confronter à des séries dont les caractéristiques de taille, la structure de la persistance, la forme et l'amplitude de la tendance sont connues. La procédure est voulu exacte, quantitative, objective et simple d'utilisation. Le critère de simplicité implique qu'en certains endroits les techniques mathématiques nécessaires à l'exactitude et à l'objectivité ont du être remplacées par le jugement de l'utilisateur. Ce compromis est nécessaire pour alléger le système et le rendre utilisable par le gestionnaire. Ce système est proposé en tant qu'outil de gestion, et non comme une procédure sophistiquée, qui ne serait utilisable que par des spécialistes en statistiques.

Des méthodes de détection des tendances par saut ont été proposées à la section 3.4.1. Elles sont calquées sur les formules utilisées pour les tendances monotones, mais les simulations permettant de connaître leurs limitations n'ont pas été effectuées. Ces limitations,

une fois connues, on conçoit qu'un système de détection des tendances par saut, calqué sur celui des tendances monotones (Figure 4.1), puisse être proposé. Dans un tel système, l'ensemble 3 de la figure 4.1 serait remplacé par la démarche décrite en 3.4.2, l'ensemble 4 par la méthode présentée en 3.4.3, la partie 5 par les formules données en 3.4.1 et l'exemple 6 par un test de Mann-Whitney.

La procédure pour les tendances monotones est représentée à la figure 4.1. Les opérations sont décrites en 4.2 et discutées en 4.3. La symbolique de la figure suit les conventions usuelles :



L'approche peut être résumée en trois points:

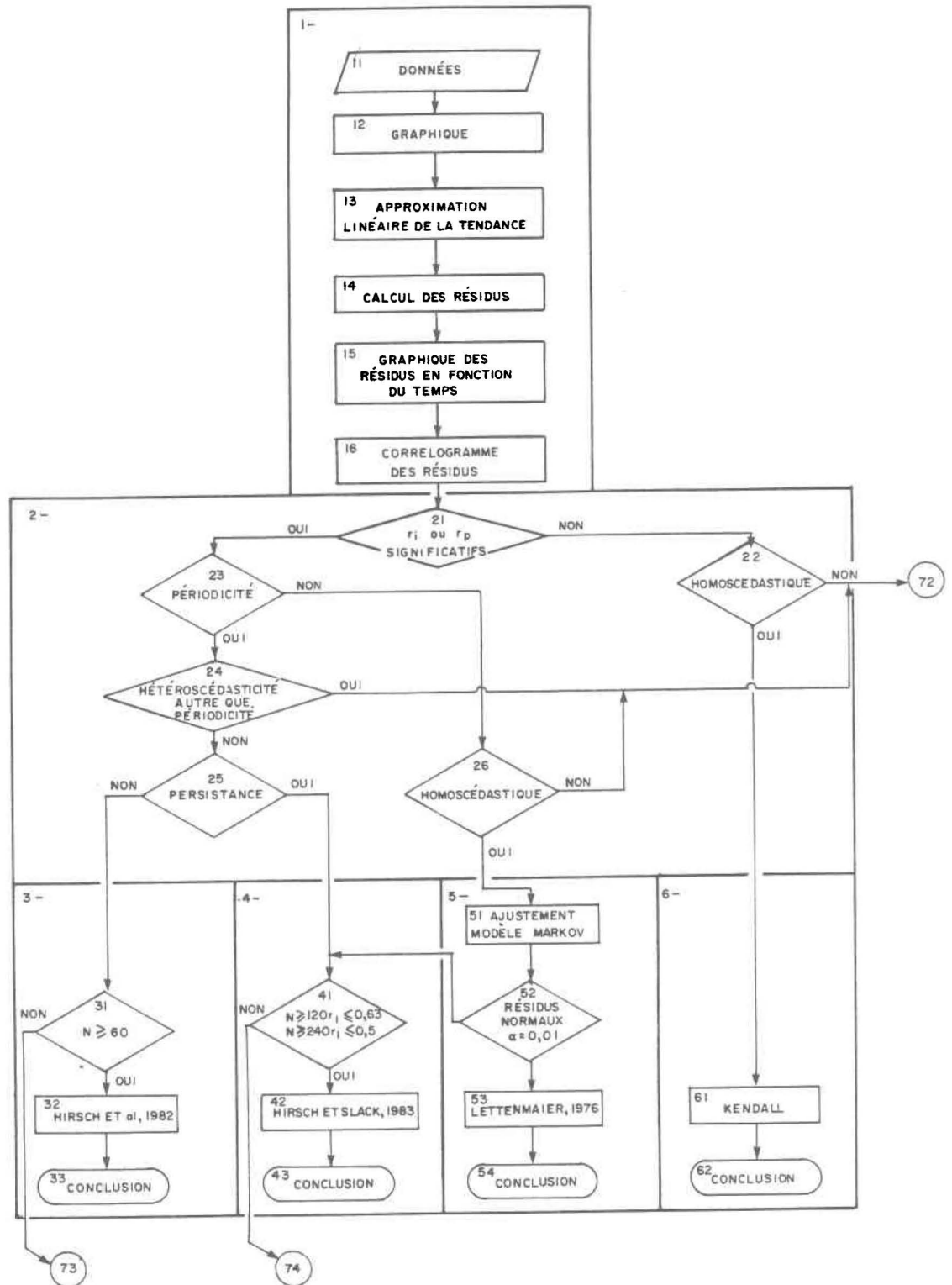


Figure 4.1: Procédure de détection des tendances monotones

- 1- Approximation linéaire de la tendance (rectangle 1, figure 4.1).
- 2- Les résidus de l'approximation linéaire servent à vérifier les hypothèses de base d'homoscédasticité et d'indépendance des erreurs. On vérifie si la série a un caractère périodique. La série est acheminée vers la méthode de traitement pertinente (rectangle 2, figure 4.1).
- 3- La méthode de traitement appropriée est appliquée sur la série de données brutes (rectangle 3, 4, 5, 6, figure 4.1).

4.2 Description¹

4.2.1 Opérations de base

Les données sont lues (11) et un graphique du paramètre en fonction du temps est tracé (12). En 13, la méthode des moindres carrés est utilisée pour ajuster une droite sur la série. De cette approximation linéaire de la tendance, on calcule les résidus (14) et on trace le graphique des résidus en fonction du temps (15). On utilise les formules 3.4 et 3.5 pour calculer r_1 et r_p pour $p < N/10$. Le corrélogramme peut alors être tracé (16).

¹ Dans cette section, sauf lorsque cela est précisé, les nombres réfèrent aux numéros des cases de la figure 4.1. Le mot résidu désigne les résidus de l'approximation linéaire de la tendance.

4.2.2 Décisions d'orientation

Toutes les décisions quant au choix de la méthode de traitement sont reliées à la vérification d'hypothèses de base à l'aide des résidus de l'approximation de la tendance. En 21 on vérifie l'indépendance des résidus. Le coefficient r_1 calculé en 16 est comparé à celui de la figure 3.2 au niveau $\alpha = 0,05$ et pour le nombre d'observations de la série. Il est conseillé de limiter l'étude de corrélation aux pas $p = 1$ à p tel que $\frac{p}{N} \leq 0,10$ (Chow, 1969). Pour de tels pas de temps la figure 3.2 peut être utilisée pour les seuils significatifs (Anderson, 1942).

Si cette vérification confirme l'hypothèse d'indépendance, la série des résidus passe en 22 où l'homoscédasticité est vérifiée à l'aide du graphique tracé en 15. L'utilisation de ce graphique et des méthodes exactes (tests) sont expliqués en 3.1.1. Pour ce qui est des méthodes exactes, le choix dépend de l'hétéroscédasticité soupçonnée. Si les résidus sont homoscédastiques, la série passe à la procédure de traitement 6 pour séries sans périodicité ni persistance.

Dans l'éventualité du non-respect de l'hypothèse d'homoscédasticité, la série sort du système de traitement (72): l'utilisateur doit alors chercher la cause de l'hétéroscédasticité. Il trouvera peut-être l'effet d'une cyclicité non détectée en 2.1. Dans un tel cas, il n'a

qu'à reprendre l'organigramme en 22. L'hétéroscédasticité peut aussi être d'un autre type (ex. : variation monotone ou en saut). L'utilisateur peut vérifier si une troisième variable explique cette hétéroscédasticité. Dans ce cas, il sera peut-être possible d'éliminer l'effet de cette variable en utilisant la méthode proposée en 2.3.5 b.

Si les résidus sont autocorrélés de façon significative, la série est acheminée en 23. Il n'y a pas de critère quantitatif et objectif pour la décision sur le caractère périodique de la série (21). La décision est prise par l'utilisateur, après l'examen visuel du graphique des données et du corrélogramme. Une périodicité peut ne pas être visible sur le graphique des données, mais l'être sur le corrélogramme.

Un mauvais jugement de la part de l'utilisateur quant à la périodicité est cependant sans grandes incidences sur le résultat final, comme on le verra en 4.3.

Si la périodicité n'est pas évidente, on vérifie l'hypothèse d'homoscédasticité telle que décrite précédemment (26). Si la série des résidus n'est pas homoscédastique, elle sort en 72. Ce cas a été décrit plus haut. Si l'hypothèse d'homogénéité de la variance est confirmée, la série passe à l'ensemble 5, adapté aux séries à persistance sans périodicité.

En 23, si une périodicité est évidente dans la série des résidus, elle sera acheminée en 24 où l'homoscédasticité est examinée. Dans ce

cas, la série ne doit pas être rejetée en 72 s'il y a hétéroscédasticité à cause du caractère périodique comme dans le cas de la figure 3.1 b. La série est dirigée vers les méthodes de traitement en 3 et 4 qui tiennent compte de la périodicité.

À la boîte 25 se pose le choix entre les méthodes adaptées au traitement des séries à périodicité. L'approche 4 est:

- exacte même pour des séries dont la persistance détermine un coefficient r_1 significatif; ce r_1 peut aller jusqu'à 0,65 pour une série de 120 observations, jusqu'à 0,5 pour une série de 240 observations;
- d'application plus lourde que la méthode en 3;
- peut être jusqu'à 25 % moins puissante (dépendamment de l'intensité de la tendance) que la méthode en 3 s'il n'y a pas de persistance dans la série, autrement dit si la dépendance mesurée dans la série des résidus provenant de la soustraction de la tendance (16) n'est due qu'à l'effet de la périodicité.

La méthode en 3 est:

- simple et rapide d'utilisation;
- mais n'est pas exacte pour les séries à persistance significative au niveau $\alpha = 0,05$.

Le corrélogramme peut être utile ici. Par exemple, si le seul coefficient significatif est celui correspondant au pas de temps du cycle, on peut s'attendre à ce qu'il n'y ait pas de dépendance autre que celle causée par la périodicité. L'utilisation de la méthode 3 serait exacte. Par contre, si le corrélogramme ressemble à celui de la figure 3.4 on s'attend à ce qu'il y ait persistance en plus de la périodicité.

De façon rigoureuse, pour s'assurer l'exactitude de la méthode choisie, il faudrait:

- 1- ajuster un modèle cyclique à la série des résidus de la tendance;
- 2- calculer les résidus de cet ajustement;
- 3- vérifier l'indépendance de ces résidus.

S'ils sont indépendants, on utilise la méthode en 3 car elle est plus puissante, sinon on doit utiliser l'approche 4 car l'autre n'est pas exacte. Si on ne veut pas faire la modélisation et si l'on veut être conforme à l'usage voulant que l'erreur α soit connue et contrôlée, alors il faut utiliser la méthode 4.

4.2.3 Traitement; séries avec périodicité

Les séries de résidus indépendants sauf pour l'effet de la composante périodique ont été amenées à l'ensemble 3. Si la série est suffisamment longue : 5 ans d'observations mensuelles (31) on utilise la méthode de Hirsch et al. (1982) (32) sur la série des données brutes et l'on conclut (33). Sinon on sort du système en 73.

Théoriquement, la méthode de Hirsch et al. (1982) est valable pour des séries aussi courtes que 2 ans. Par contre sa robustesse et sa puissance n'ont été vérifiées que pour un minimum de 5 ans.

4.2.4 Traitement; séries avec périodicité et dépendance

Il faut vérifier si la série est assez longue (minimum 120 observations) et si l'autocorrélation n'est pas trop forte (41). Si ces conditions sont respectées, on applique la méthode de Hirsch et Slack (1984) et on conclut (42 et 43). Si elles ne le sont pas, on sort du système en 74.

Dans ce cas, il reste une possibilité pour avoir une fréquence de mesure plus faible: éliminer des observations, ce qui conduit à une diminution de l'autocorrélation dans la série des résidus, mais aussi de la longueur de cette série. On recherche alors un pas de temps pour lequel l'autocorrélation sera suffisamment diminuée mais la taille

encore suffisamment grande pour que l'une des méthodes de Hirsch soit applicable.

Les séries rejetées du système en 74, à cause d'une dépendance trop forte méritent une attention spéciale. Une part de cette autocorrélation peut en effet être due à une tendance. Dans un tel cas, on a un corrélogramme où les r_p restent élevés; ils ne diminuent pas à zéro lorsque p augmente, contrairement aux cas des figures 3.3 et 3.4 (Yevjevich, 1969).

4.2.5 Traitement; séries avec dépendance

La méthode proposée par Lettenmaier (1976) lorsqu'il y a persistance s'applique aux cas de dépendance Markovienne d'ordre 1 ($r_p = r_1^p$). Pour vérifier cette condition, il faut ajuster un modèle Markovien d'ordre 1 à la série des résidus (51) (voir 4.3.2). On vérifie la normalité des résidus des observations par rapport au modèle (52) par l'examen visuel de l'ajustement de l'histogramme des résidus par rapport à une courbe normale ayant les mêmes paramètres. La méthode est sensible à des déviations, même très légères, du modèle Markovien d'ordre 1 (Cluis, 1983).

Si cette condition est respectée, on applique la méthode sur la série des données brutes et l'on conclut (53 et 54).

Si la persistance ne peut être modélisée par un processus Markovien d'ordre 1, la dépendance mutuelle entre les résidus peut être due à une périodicité non détectée par l'utilisateur (en 21), ou à une persistance non-markovienne, ou à une combinaison des deux. La série des résidus est alors transférée en 41 car la méthode 4 est robuste dans ces cas.

4.2.6 Traitement; séries sans dépendance ni périodicité

Si les hypothèses de base sont respectées (23 et 24) le test de Kendall peut être appliqué sur les séries sans subdivision en sous-séries et sans calcul du nombre équivalent d'observations indépendantes. On utilise le test sur la série entière et l'on conclut (61 et 62).

4.3 Discussion

4.3.1 Les boîtes 13 - 22 - 23 - 24 - 25 - 26

Sauf pour la première, ces boîtes font ou peuvent faire appel au jugement de l'utilisateur. En 13 on calcule les résidus de l'approximation linéaire de la tendance. Qu'arrive-t-il lorsque la tendance monotone n'est pas linéaire ? On obtient alors un graphique des résidus en fonction du temps tel que celui de la figure 4.2. On s'aperçoit que l'autocorrélation est négative pour la première partie alors qu'elle est positive pour la deuxième. Peut-on s'attendre à ce que ces deux autocorrélations s'annulent ? À ce que l'approximation linéaire élimi-

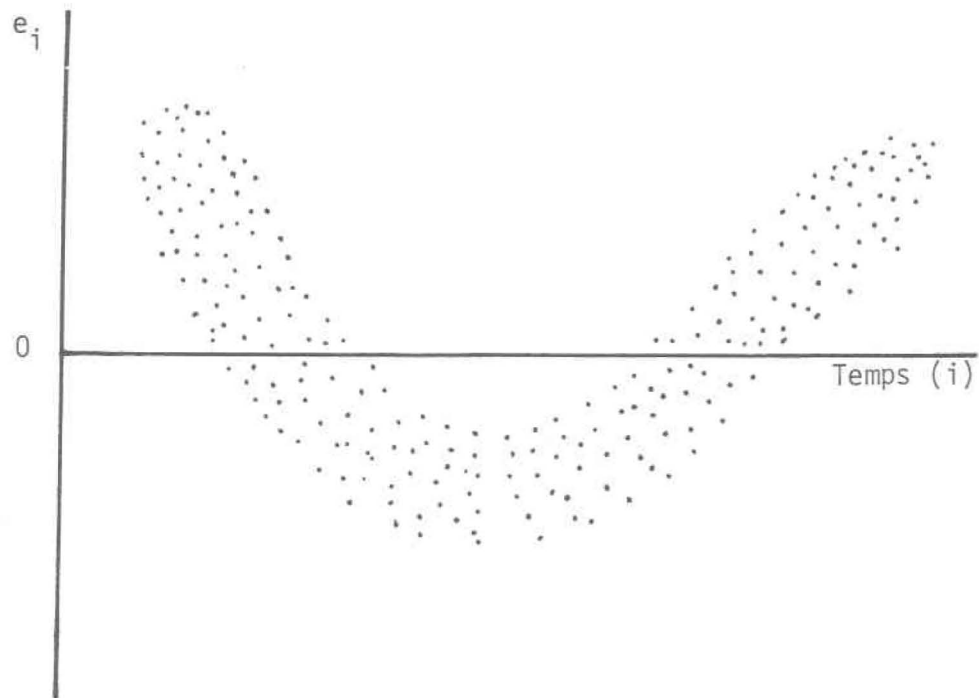


Figure 4.2: Résidus de l'approximation linéaire d'une tendance monotone en courbe

ne la dépendance due à la tendance pour les tendances non linéaires ? Des études plus approfondies seraient nécessaires pour répondre à ces questions. Si la réponse est non, alors certaines séries pourraient être acheminées vers la méthode 4 au lieu de 3, ou la méthode 5 au lieu de 6. Il en résulterait une perte de puissance.

La boîte 23 vérifie le caractère périodique de la série. Remplacer cette boîte subjective par une opération objective impliquerait l'introduction d'un processus itératif (boucle) de modélisation du cycle jusqu'à l'obtention de résidus normaux; une procédure beaucoup trop lourde en pratique pour des études simplifiées.

La boîte 23 telle que formulée est un raccourci. Une erreur de jugement est sans conséquence en effet:

- Si l'on juge à tort qu'il n'y a pas de périodicité :
 - la série reviendra dans l'ensemble 4 via le traitement 5 car on ne pourra y ajuster un modèle markovien.
- Si l'on juge à tort qu'il y a périodicité :
 - de toute façon, il y a peu de chance qu'un périodicité semble évidente et qu'un modèle markovien puisse être ajusté à la série des résidus de l'approximation linéaire;

- les méthodes 3 et 4 sont exactes dans le cas sans périodicité, il reste à savoir s'il y a perte de puissance par rapport aux méthodes 6 et 5.

On se souvient qu'il y a une "zone grise" où le coefficient d'autocorrélation d'ordre 1 n'est pas significatif à $\alpha = 0,05$ mais où la méthode 3 n'est pas appliquée avec exactitude. Si un modèle cyclique a été ajusté aux résidus et que des nouveaux résidus calculés ont une autocorrélation nettement inférieure au r_1 significatif à $\alpha = 0,05$ (figure 3.2), on peut utiliser la méthode 3. Sinon, pour avoir une procédure exacte, il faut effectuer le traitement en 4.

La façon de vérifier l'homoscédasticité (23 - 24 - 25 - 26) est laissée à l'usager car elle dépend de la cause soupçonnée de l'hétéroscédasticité. Des méthodes ont été proposées en 3.1.1.

4.3.2 Simplification de la procédure

Il serait intéressant de comparer la puissance des méthodes 3 et 6 d'une part et 4 et 5 d'autre part, dans les conditions d'application des méthodes 6 et 5, c'est-à-dire sans périodicité. Si la différence est négligeable, les ensembles 5 et 6 de même que les boîtes 23, 24 et 72 pourraient être délaissés. La vérification de l'effet de subdivision de la série lorsque ce n'est pas nécessaire serait intéressante surtout pour comparer les méthodes 4 et 5. En effet, la méthode 5 est particulièrement lourde et d'application restreinte. En 5.1 elle de-

mande l'ajustement d'un modèle Markovien d'ordre 1. Ce modèle est de la forme:

$$y_i = r_1 y_{i+1} + \xi_i \sqrt{1 - r_1^2} + A \quad (\text{éq. 41})$$

où :

r_1 : coefficient d'autocorrélation d'ordre 1

ξ_i : variable aléatoire normale centrée

A : niveau de base

Les résidus des observations par rapport aux estimés du modèle sont testés pour vérifier leur normalité au seuil $\alpha = 0,01$ (Annexe 2). On trouve à l'annexe 2 un test non-paramétrique permettant de vérifier la normalité. Ce test et ses équivalents paramétriques ne sont pas très puissants. L'examen visuel de l'ajustement de l'histogramme des résidus par rapport à une courbe normale ayant les mêmes paramètres pourrait être plus révélateur.

Un usager que l'ajustement d'un tel modèle rebute peut toujours utiliser la méthode 4. La méthode est exacte même s'il n'y a pas de périodicité, mais on ne connaît pas les implications en ce qui concerne la puissance.

CONCLUSION

L'objectif de développement d'un outil permettant de détecter les tendances en qualité de l'eau qui soit quantitatif, objectif, simple et rapide d'utilisation et accessible dans ses fondements théoriques, a été réalisé à partir d'une synthèse des méthodes présentées dans la littérature. Cet instrument est présenté au chapitre 4, des simplifications sont proposées en 4.3.2. Au sens strict, l'outil ne respecte pas les critères de sélection à cause de la nécessité d'un compromis entre l'objectivité et la simplicité.

Pour les tendances monotones, rappelons que, même s'il y a asymétrie dans la distribution,

- les séries sans persistance ni périodicité peuvent être traitées par le test de Kendall;
- les séries à persistance sans périodicité sont traitées par la méthode de Lettenmaier si la persistance est markovienne, sinon elles peuvent être traitées par la méthode applicable aux séries à périodicité et persistance;
- les séries à périodicité peuvent être traitées si elles ont un minimum de 5 cycles.

- les séries à persistance et périodicité peuvent être traitées si elles sont suffisamment longues et si l'autocorrélation n'est pas trop forte (N minimum: 120; $r_1 < 0,65$ si $N = 120$; $r_1 < 0,5$ si $N = 240$). En dehors de ces limites, les méthodes ne sont pas appliquées avec exactitude. Ces méthodes sont déjà utilisées par les auteurs qui les proposent et par d'autres (Hirsch et al., 1982; Lettenmaier, 1982).

Dans le cas des tendances par saut, la formulation mathématique des tests a été proposée et les limites théoriques, quant à la taille des échantillons, ont été évaluées. Il reste à étudier l'effet de la taille et de la persistance sur les erreurs de type 1 et 2, ce qui permettra de poser les limites pratiques des tests proposés.

BIBLIOGRAPHIE

- ANDERSON, O.D. (1976).
Time series analysis and forecasting, Butterworths, 182 p.
- ANDERSON, R.L. (1942).
Distribution of the serial correlation coefficient, Ann. Math. Stat, 13,1.
- ALVO, M. (1981).
Litterature review of trend analysis (1970-1980), Rapport SCS Consultants, 60 p.
- BAILLARGEON, G. et J. RAINVILLE (1977).
Statistique appliquée, tome 2. 5^e édition, Ed. SMG.
- BERENSON, M.L. (1978).
A simple distribution free test for trend in one-way layouts, Edu. Psy. Measur., 38.
- BERNIER, J. (1977).
Etude de la stationarité des séries hydrométéorologiques, La houille blanche, 32, 4 313-319.
- BOBÉE, B. (1983).
Note de cours, 2 année de maîtrise, INRS-Eau.
- BOBÉE, B. (1980).
Notes sur les tests non paramétriques, non publié.
- BOBÉE, B. et D. CLUIS (1979).
Etude de la persistance de quelques paramètres de la qualité de l'eau, J. of Hydro., 44, 17-36.
- BROWN, A. (1975).
Régression et corrélation, Notes de cours MQT-11947, Sciences de l'administration, Univ. Laval, 150 p.
- BRUNET-MORET, Y. (1971).
Etude de l'homogénéité des séries chronologiques de précipitations annuelles par la méthode des doubles-masses, Cahier de l'ORSTOM, Série Hydrol. VIII, 4, 3-31.
- CHOW, V.T. (1964).
Handbook of applied hydrology, McGraw-Hill Book Co.
- CLUIS, D. (1981).
Detection of seasonal and annual water quality trends using visual techniques : double-mass curves and cusum fonctions, International conference on time-series methods in hydrosiences, CCIW.

- CLUIS, D.A. (1983).
Information content of the mean and rescaled range of short samples derives from AR2 processes.
- COCHET, Y. (1981).
Méthode statistiques applicables à la recherche de l'évolution des paramètres de qualité des eaux superficielles, thèse de doctorat, Université de Montpellier.
- COCHET, Y. et B. BOBÉE (1980).
Etude des tendances à l'aide de tests non-paramétriques INRS-Eau, rapport interne n° 71.
- CONOVER, W.J. (1971).
Practical non-parametric statistics, John Wiley Pub., 462 p.
- COX, D. R. et A. STUART (1955).
Some quick tests for trend in location and dispersion, Biometrika, 42, 80-95.
- DIETZ, E.J. et T.J. KILLEEN (1981).
A nonparametric multivariate test for monotone trend with pharmaceutical applications. J. Am. Stat. Ass. 76-373.
- FERGUSON, G.A. (1965).
Non-parametric trend analysis, McGill Uni. Press.
- FIERING, M.B. et B. JACKSON (1971).
Synthetic streamflows, Wat. Res. Monograph Series.
- FIERING, M.G. et al. (1971).
Water resources systems analysis, Res. paper n° 3, Environnement Canada, Ottawa.
- FOSTER, F.G. et A. STUART (1954).
Distribution-free tests in time-series based on the breaking of records, Jour. Roy. Stat. Soc. série B, vol. XVI, n° 1.
- HALD, A. (1982).
Statistical tables and formules. Wiley et Sons.
- HIRSCH, R.M. et al. (1982).
Techniques of trend analysis for monthly water quality data, Wat. Res. Res. 18, 1, 107-121.
- HIRSCH, R.M. et J.R. SLACK (1984).
A nonparametric trend test for seasonal data with serial dependence. W. Res. Res. 20 (6), pp. 727-732.

- HOLLANDER, M. et D.A. WOLFE (1973).
Non-parametric statistical methods, John Wiley & Sons.
- KENDALL, M.G. (1962).
Rank-correlation methods, Charles Griffin and Co., 3^e ed.
- KOZIOL, J.A. et al. (1978).
A simple non-parametric test for trend : an example based on a biological model for cancer of the pancreas, *Biom. J.* vol. 20, n° 5.
- KLOTZ, J. (1962).
Non-parametric test for scale. *Ann. Mat. Stat.* 33, pp. 498-512.
- KRUSKAL, W.H. et W.A. WALLIS (1952).
Use of ranks in one criterion variance analysis, *J. Am. Stat. Ass.*, 47, 583-621.
- LALOIRE, J.C. (1972).
Méthodes du traitement des chroniques, Dunod, 194 p.
- LAZARO, T.R. (1976).
Non-parametric statistical analysis of annual peak flow data from a recently urbanized watershed, *Wat. Res. Bul.* 12, 1, 101-107.
- LEHMAN, E.L. (1953).
The power of rank tests, *Ann. of Math. Stat.*, 24.
- LEHMAN, E.L. et H.J.M. D'ABRERA (1975).
Non-parametrics : statistical methods based on ranks, Holden-Day inc.
- LEHNER, P. (1978).
The handbook of ethological measurements, Garland STMP. Press. New York.
- LETTENMAIER, D.P. (1976).
Detection of trend in water quality data from records with dependent observations, *Wat. Res. Research*, 12 (5), 1037-1046.
- LETTENMAIER, D.P. et al. (1982).
Routine streams and rivers quality trend monitoring review, Technical report n° 75, Uni. of Washington, Seattle.
- LILLIEFORS, H.W. (1967).
One the Kolmogorov-Smirnov test for normality with mean and variance unknown. *J. Am. Stat. Ass.*, 62, pp. 399-402.
- LLAMAS, J. et TRIBOULET (1976).
Les cahiers du Centreau, vol. I, n° 3.
- MANN, H.B. (1945).
Non-parametric test against trend, *Econometrica*, 13, 245-269.

MANN, H.B. et D.R. WHITNEY (1947).

On a test wheter one of two random variables is stochastically larger than the other, Ann. Math. Stat., 18, 50-60.

MARASCUILO, L.A. et M. MCSWEENEY (1967).

Non-parametric post-hoc comparisons for trend, Psych. Bull., 67-6, 401-412.

MARASCUILO, L.A. et M. MCSWEENEY (1977).

Non-parametric and distribution-free methods for the social sciences, Brooks/Cole Pub. Co. 556 p.

McCuen, R.H. et D. JAMES (1972).

Non-parametrics statistical methods in urban hydrologic research, Wat. Res. Bul. 8, 5.

ODEH, R.E. (1971).

On Jonckheere's K-sample test against ordered alternatives, Technometrics, 13, 4.

OPRIAN, C.A. (1976).

On the exact power of rank test for a Lehman-type trend alternative, Metron, 34, 191-200.

PETERS, N.E. et al. (1982).

Temporal trends in the acidity of precipitation and surface waters of New York, U.S. Geo. Survey Wat. Sup., paper 2188.

POTTER, R.W. et G.W. STURM (1981).

The power of Jonckheere's test, American Statistician, 35, 4.

PURI, M.L. (1965).

Some distribution-free K-sample rank tests of homogeneity against alternatives, Commu. Pure. Appl. Math., vol XVIII.

SEN, P.K. (1968).

Estimates of the regression coefficient based on Kendell's Tau. J. Am. Statist. Assoc. 63 : 1379-1389.

SHERWANI, J.K. et D.H. MOREAU (1975).

Strategies for water quality monitoring, Water Research Institute of the University of North Carolina, rapport n° 107, 86-92.

SIEGEL, S. (1956).

Non-parametric statistics for the behavioral sciences, McGraw-Hill.

SNEYERS, R. (1975).

Sur l'analyse statistique des séries d'observation, notes techniques, n° 143, OMM.

SNYDER, M. (1971).

Summary and evaluation of methods for detecting hydrologic changes, Jour. Hydro. 12, 311-338.

ANNEXE 1¹

FONCTIONS CUSUM

On considère la fonction CUSUM (j) associée au temps j telle que pour une série ($x_1, x_2, \dots, x_i, \dots, x_N$) :

$$\text{CUSUM}(j) = \sum_{i=1}^j x_i - (j/m) \sum_{i=1}^m x_i$$

m est une valeur comprise entre 2 et N.

Les différentes courbes obtenues en faisant varier m sont sensiblement de même forme et leur analyse conduit à des conclusions identiques. Le choix de m peut donc être arbitraire, par exemple N/2.

Si la moyenne $\bar{x}_j$ de la sous-série ($x_1 \dots x_j$) est constante pour tout $j < N$, la fonction CUSUM est nulle. Si la moyenne x_j augmente (diminue) brusquement pour $j = j_0$, la fonction CUSUM devient linéairement croissante (décroissante) pour $j > j_0$ (Figure 1). Si la moyenne x_j augmente (diminue) progressivement à partir de $j = j_0$, la fonction CUSUM devient quadratique croissante (décroissante) pour $j > j_0$ (Figure 2). Les figures 1 et 2 sont relatives à des exemples hypothétiques dont les données sont fournies en annexe 2.

¹ Tiré de Cochet et Bobée, 1980

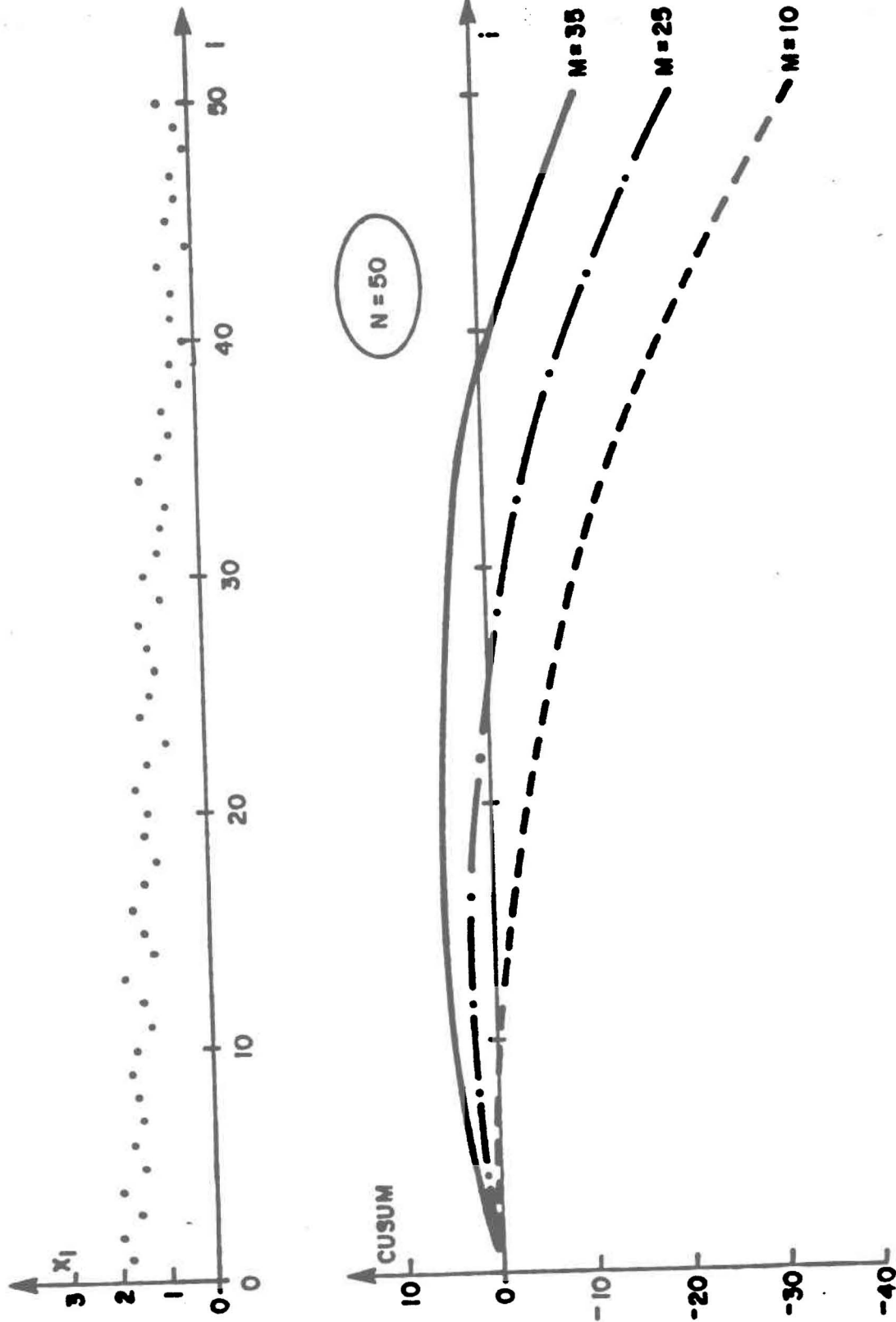


Figure 2 . Évolution de la fonction CUSUM pour différentes valeurs de M , dans le cas d'une tendance continue décroissante .

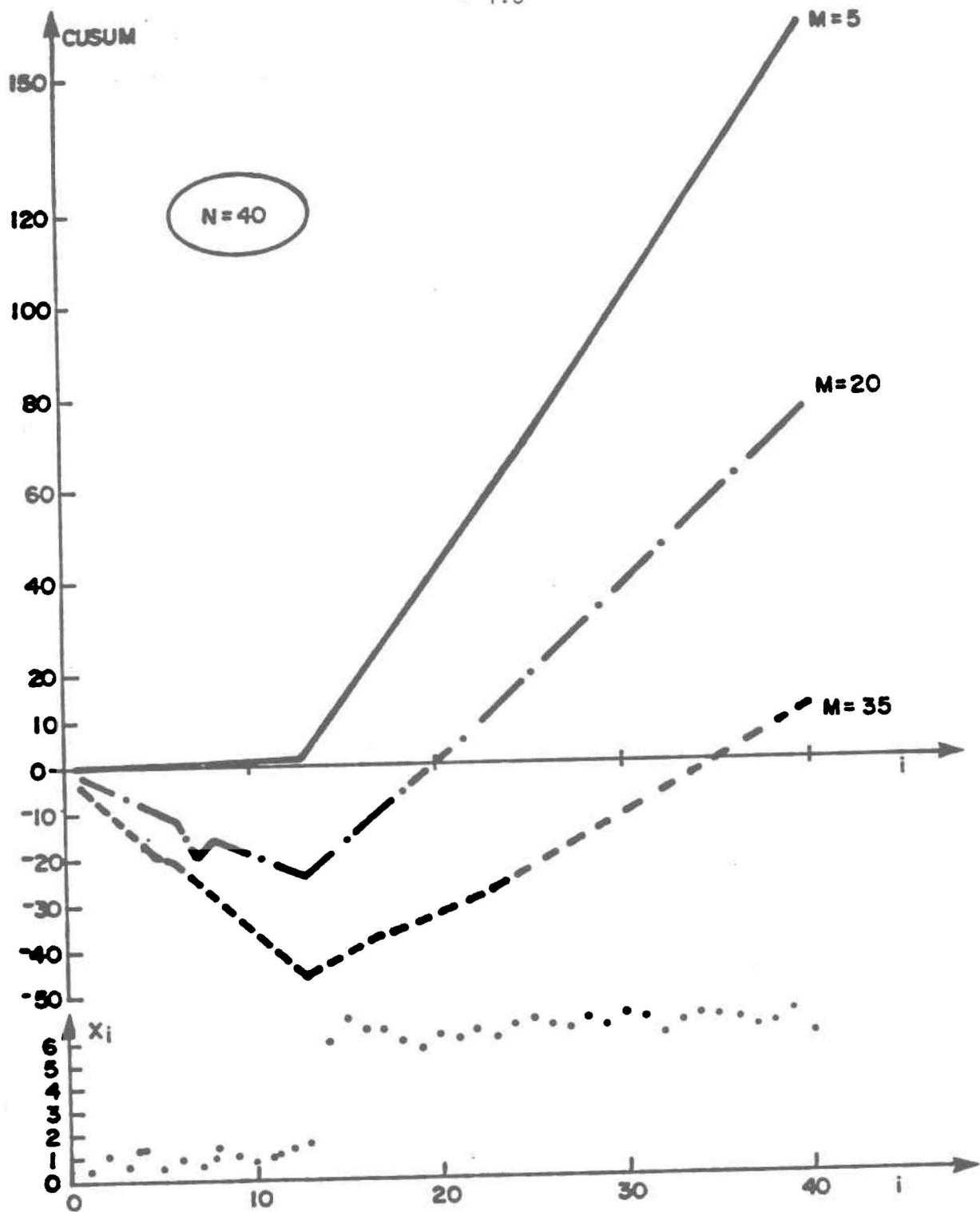


Figure 1. Évolution de la fonction CUSUM pour différentes valeurs de M dans le cas d'un changement brusque.

Le fait de cumuler les valeurs x_i augmente la sensibilité du CUSUM par rapport à celle de la simple représentation graphique. Il faut noter en particulier la différence d'échelle entre les deux représentations . Cette méthode empirique permet de visualiser de façon plus nette une éventuelle tendance dans une série. Il est également possible, à partir de l'étude des graphiques, dans le cas de l'hypothèse d'une tendance, d'en préciser le type.

Cette mise en évidence graphique peut alors être suivie de l'application de tests non-paramétriques. Le choix des tests dépend du type supposé de tendances. On pourra alors conclure à l'existence ou la non existence d'une tendance pour un niveau de signification donné.

ANNEXE 2
TEST DE NORMALITÉ (LILLIEFORS, 1967)

Ce test est une adaptation du test de Kolmogorov-Smirnov. Il vérifie la normalité d'une distribution en examinant la qualité de l'ajustement entre cette distribution et une distribution normale ayant les mêmes paramètres (moyenne et variance).

Procédure :

- 1) Classer les observations par rang croissant.
- 2) Standardiser les observations. On utilise ici la formule usuelle

$$Z_i = \frac{y_i - \bar{y}}{S}, \text{ où :}$$

$\bar{y}$ = moyenne de l'échantillon;

S = écart-type de l'échantillon;

Z_i = observation standardisée.

- 3) Calcul de la distribution cumulée exacte. La fréquence au non-dépassement de chaque observation y_i est donnée par :

$$P_i (Z < Z_i) = \frac{y_i}{N} \text{ où :}$$

y_i = rang de l'observation i .

- 4) Calcul de la distribution normale cumulée. On trouve, par la table A-2 (Marascuilo, 1967) les fréquences au non-dépassement des Z_i calculés au point 2. On appelle cette quantité :

$$P_i (Z < Z_i)$$

- 5) Calcul de la statistique. La statistique est la distance séparant les deux distributions (des points 3 et 4) au point où elles sont les plus éloignées, c'est-à-dire là où l'ajustement est le moins bon. Ce point n'a pas besoin de faire partie de l'échantillon. Cette distance est :

$$KS = \text{Max } |P_i (Z < Z_i) - P_i (Z - Z_i)|$$

- 6) Décision. La valeur de KS est confrontée aux KS théoriques de la table A-12 (Marascuilo, 1967) au seuil α et pour le nombre d'observations de l'échantillon. On rejette l'hypothèse de la normalité si $KS > KS \text{ théorique}$.

ANNEXE 3

N ET K MINIMAUX POUR L'APPROXIMATION NORMALE DU TEST DE MANN-WHITNEY

Il n'y a pas de table donnant la distribution exacte de la statistique d'un test lorsqu'elle provient de la sommation des statistiques de plusieurs sous-séries.

Dans un tel cas on ne peut utiliser que l'approximation à la loi normale pour les grands échantillons. Il faut donc vérifier les nombres d'observations et de sous-séries minimaux pour que la statistique suive une loi normale. C'est ce qu'ont fait Hirsch et al., 1982 pour le test de Kendall et c'est ce qui est fait ici pour le test de Mann-Whitney.

Pour cette vérification on compare la distribution exacte de la statistique et celle obtenue par l'approximation normale. L'exemple suivant présente les calculs permettant d'obtenir les deux distributions. Exemple : nombre d'observations par sous-série : 4, dont 2 avant le saut; nombre de sous-séries : 2.

Calcul de la distribution exacte

On commence par calculer la distribution d'une sous-série. Le tableau A.3.1 donne les permutations possibles des rangs pour une série de quatre observations. Il donne aussi la statistique de Mann-Whitney

dans chaque cas. Rappelons que cette quantité est basée sur :

$$U = \sum_{i=1}^N R_i$$

R_i : nombre d'observations du groupe 2, supérieures à l'observation i du groupe 1.

On obtient par le tableau A3.1 la distribution exacte de U sous l'hypothèse de l'absence de tendance; toutes les permutations sont équi-probables. Cette distribution est résumée au tableau A3.2.

Dans le cas où on a deux sous-séries de 4 observations, la statistique totale est la somme des statistiques des deux sous-séries, mais la probabilité d'avoir cet total est le produit des probabilités des statistiques des sous-séries. En effet :

$$P(A + B) = P(A) \cdot P(B)$$

où $P(A)$: probabilité d'avoir A .

C'est en suivant cette règle que l'on compose le tableau A3.3 qui donne la distribution de U obtenue par sommation de deux sous-séries de 4 observations ($K = 2$, $N_1 = N_2 = 2$). Cette distribution est présentée à la figure A3.1.

Calcul de la distribution théorique

Pour calculer la distribution théorique de U basée sur l'approximation à la loi normale on a :

$$Z^* = \frac{U - E(U)}{\sqrt{\text{Var}(U)}} \quad (\text{éq. A3.1})$$

où

$$E(U) = \sum_k^k E(U_k) \quad (\text{éq. A3.2})$$

et

$$\text{Var}(U) = \sum^K \text{Var}(U_k) \quad (\text{éq. A3.3})$$

$$E(U_k) = \frac{n_1 + n_2}{2} \quad (\text{éq. A3.4})$$

$$\text{Var}(U_k) = \frac{n_1 n_2 (n_1 + n_2 + 1)}{12} \quad (\text{éq. A3.5})$$

* On a obtenu un meilleur ajustement avec cette formule que celle corrigée par continuité.

Dans notre cas : $n_1 = n_2 = 2$. On a donc :

$$E(U_k) = 2$$

$$\text{Var}(U_k) = 1,167$$

$$E(u) = 4$$

$$\text{Var}(u) = 3,334$$

Le tableau A3.4 donne la valeur de Z calculée par A3.1 pour chaque valeur de U . À partir de Z on obtient $F(Z)$ à l'aide d'une table (Hall, 1952). $F(Z)$ est l'ordonnée à l'origine de Z pour une distribution normale de moyenne zéro et de d'écart-type 1.

On sait (Bobée, 1983) que :

$$G(U) = F(Z) \left| \frac{dz}{dU} \right| = F(Z) \cdot \frac{1}{\sqrt{\text{Var}(U)}}$$

On obtient donc $G(U)$ en divisant $F(Z)$ par $\sqrt{\text{Var}(U)}$. $G(U)$ est l'ordonnée à l'origine d'une loi normale d'écart-type $\sqrt{\text{Var}(U)}$, la distribution théorique que l'on cherchait. Les $G(U)$ du tableau A3.4 sont portés à la figure A3.1 pour fin de comparaison.

À l'examen de la figure A3.1 on s'aperçoit que l'ajustement n'est pas très bon.

En reprennant le même calcul mais pour le cas de 4 sous-séries de 4 observations, on arrive aux résultats du tableau A3.5 et de la figure A3.2. L'ajustement est amélioré de beaucoup.

Cas où K = 12

Dans le cas où on a 12 sous-séries, la procédure de calcul serait trop lourde, le nombre de permutations devenant trop grand. Toutefois, pour $N = 2$, U_k ne peut prendre que deux valeurs : 0 ou 1. Donc :

$$U = \sum_{k=1}^{12} U_k$$

sera compris entre $12 \times 0 = 0$

$$12 \times 1 = 12$$

Par exemple, si $U = 4$, c'est qu'il y a quatre sous-séries qui ont la statistique 1 et huit qui ont la statistique 0.

Le nombre de permutations possibles de quatre chiffres dans 12 positions est donnée par A3.6 :

$$\frac{n!}{t! (n - t)!} \quad (\text{éq. A3.6})$$

n : nombre de positions;

t : nombre de chiffres à permuter.

Dans notre exemple on a $n = 12$ et $t = 4$. On a donc :

$$\frac{12!}{4! (12 - 4)!} = 482$$

En appliquant le même calcul pour les autres valeurs de U, on obtient les résultats de la 2^e colonne du tableau A3.6 reporté à la figure A3.3. La distribution théorique par la loi normale est calculée de la même façon que dans l'exemple précédent.

Vérification de l'ajustement

La qualité de l'ajustement peut être vérifiée par le khi-deux (χ^2) et par l'erreur relative maximale. Le chi-deux est donné par A3.7.

$$\chi^2 = \sum_U \frac{P(U) - G(U)}{G(U)} \quad (\text{éq. A3.7})$$

La valeur du chi-deux est utilisé pour vérifier l'hypothèse de l'égalité des distributions à un niveau de signification α . Cependant, ce test n'est pas très puissant et de peu d'intérêt dans les cas présentés ici; l'hypothèse d'égalité est toujours vérifiée. Les valeurs de chi-deux seront cependant utiles pour vérifier l'effet de N et de K sur l'ajustement.

L'erreur relative maximale est donnée par :

$$\Delta r = \frac{P(u) - G(u)}{G(u)} \times 100 \quad (\text{éq. A3.8})$$

au point où $P(u)$ et $G(u)$ sont le plus distant.

Les tableaux A3.7, A3.8 et A3.9 résument les résultats.

À l'examen des tableaux A3.7 et A3.8 on voit qu'une augmentation de N ou de K occasionne :

- une diminution du chi-deux calculé et une augmentation du seuil théorique de la chi-deux;
- une diminution de l'erreur relative;
- une diminution de l'erreur absolue.

On en conclut que l'ajustement s'améliore avec une augmentation de N et/ou de K .

Hirsch et al., 1982 sont satisfaits d'un ajustement où l'erreur relative maximale est à 14 %. À l'examen des tableaux A3.7, A3.8 et A3.9, on constate que l'on pourrait donc être satisfait de l'approximation à la loi normale lorsque :

- 3.8 -

$N = 4$ et $K = 4$

$N = 5$ et $K = 2$

$N = 2$ et $K = 12$

Évidemment, si N et K sont supérieurs, l'ajustement n'en sera qu'amélioré.

Tableau A3.1 : Permutations des rangs et statistique de Mann-Whitney pour une série de quatre observations

AVANT SAUT		APRÈS SAUT		U_k
0	1	2	3	0
0	1	3	2	0
1	0	2	3	0
1	0	3	2	0
0	2	1	3	1
0	2	3	1	1
0	3	1	2	2
0	3	2	1	2
1	2	0	3	2
1	2	3	0	2
1	3	0	2	3
1	3	2	0	3
2	0	1	3	1
2	0	3	1	1
2	1	0	3	2
2	1	3	0	2
2	3	1	2	4
2	3	2	1	4
3	0	1	2	2
3	0	2	1	2
3	1	0	2	3
3	1	2	0	3
3	2	0	1	4
3	2	1	0	4

Tableau A3.2 : Distribution de U pour une sous-série de 4 observations dont deux avant le saut.

U_k	Fabs.	Frel.
0	4	1/6
1	4	1/6
2	8	2/6
3	4	1/6
4	4	1/6
	24	6/6

Fabs : fréquence absolue
Frel : fréquence relative

Tableau A3.3 : Calcul de la distribution exacte de la statistique totale U

U	1	2	P(u)
0	(0+0)	(1/6.1/6)	1/36 = 0,028
1	(0+1), (1+0)	(1/6.1/6) + (1/6.1/6)	2/36 = 0,056
2	(0+2), (2+0), (1+1)	(1/6.2/6) + (2/6.1/6) + (1/6.1/6)	5/36 = 0,139
3	(0+3), (3+0), (1+2), (2+1)	(1/6.1/6) + (1/6.1/6) + (1/6.2/6) + (2/6.1/6)	6/36 = 0,167
4	(0+4), (4+0), (1+3), (3+1), (2+2)	(1/6.1/6) + (1/6.1/6) + (1/6.1/6) + (1/6.1/6) + (2/6.2/6)	8/36 = 0,222
5	(1+4), (4+1), (2+3), (3+2)	(1/6.1/6) + (1/6.1/6) + (2/6.1/6) + (1/6.2/6)	6/36 = 0,167
6	(2+4), (4+2), (3+3)	(2/6.1/6) + (1/6.2/6) + (1/6.1/6) +	5/36 = 0,139
7	(3+4), (4+3)	(1/6.1/6) + (1/6.1/6) +	2/36 = 0,056
8	(4+4)	(1/6.1/6) +	1/36 = 0,028

Note 1 : permutations des U_k des sous-séries donnant le total U;
 2 : probabilités associées aux permutations en 1;
 P(u) : probabilité d'avoir la valeur U sous l'hypothèse nulle.

Tableau A3.4 : Distribution théorique (approximation normale) de U.

U	Z	F(Z)	G(U)
0	2,19	0,036	0,020
1	1,64	0,014	0,057
2	1,10	0,218	0,119
3	0,55	0,343	0,187
4	0	0,399	0,213

N.B. La distribution étant symétrique, les valeurs ne sont données que pour la moitié.

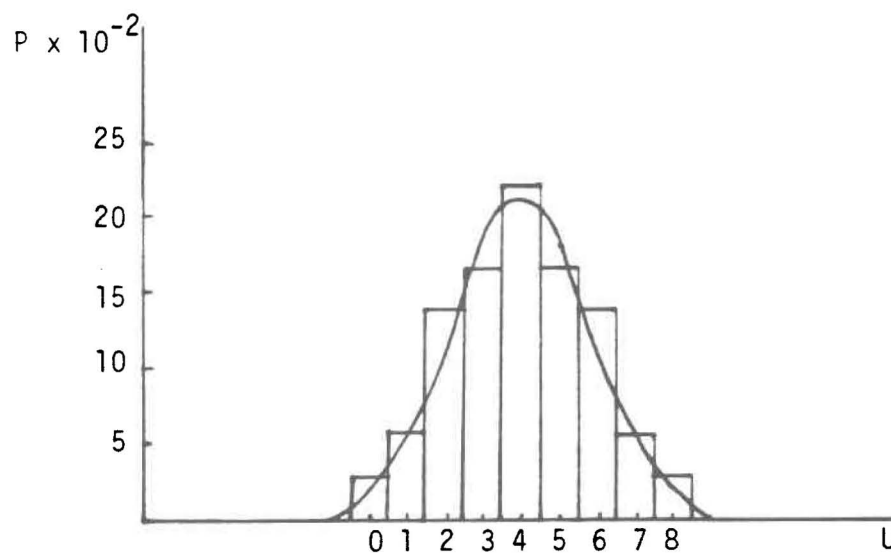


Figure A3.1 : Distributions exacte ($P(U)$: histogramme) et théorique ($G(U)$: courbe) pour $N = 2$ et $K = 4$

Tableau A3.5 : Distribution exacte ($P(U)$) et théorique ($G(U)$) pour $N = 4$ et $K = 4$

U	P(U)	Z	F(Z)	G(U)
0	0,001	- 3,10	0,004	0,002
1	0,003	- 2,71	0,010	0,004
2	0,011	- 2,32	0,028	0,011
3	0,025	- 1,94	0,061	0,024
4	0,050	- 1,55	0,120	0,046
5	0,080	- 1,16	0,204	0,080
6	0,118	- 0,77	0,297	0,115
7	0,139	- 0,39	0,370	0,143
8	0,151	0	0,399	0,155

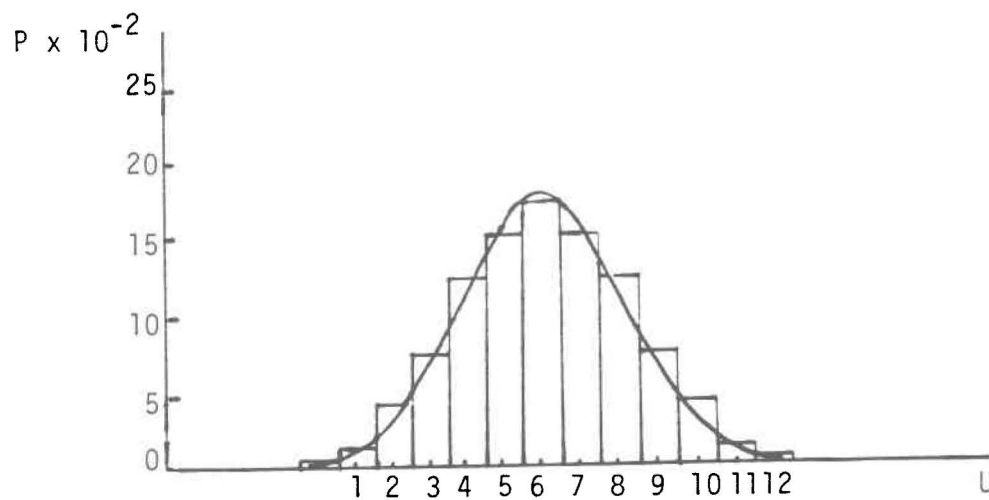


Figure A3.2 : Distributions exacte ($P(U)$: histogramme) et théorique ($G(U)$: courbe) pour $N = 4$ et $K = 4$

Tableau A3.6 : Distributions exacte et théorique pour $N = 2$ et $K = 12$

U	F(U)	P(U)	Z	F(Z)	G(U)
0	1	0,0002	- 3,46	0,005	0,0001
1	12	0,003	- 2,84	0,007	0,003
2	66	0,016	- 2,31	0,028	0,016
3	220	0,054	- 1,73	0,089	0,052
4	495	0,121	- 1,15	0,206	0,119
5	792	0,193	0,58	0,337	0,195
6	924	0,225	0	0,394	0,230

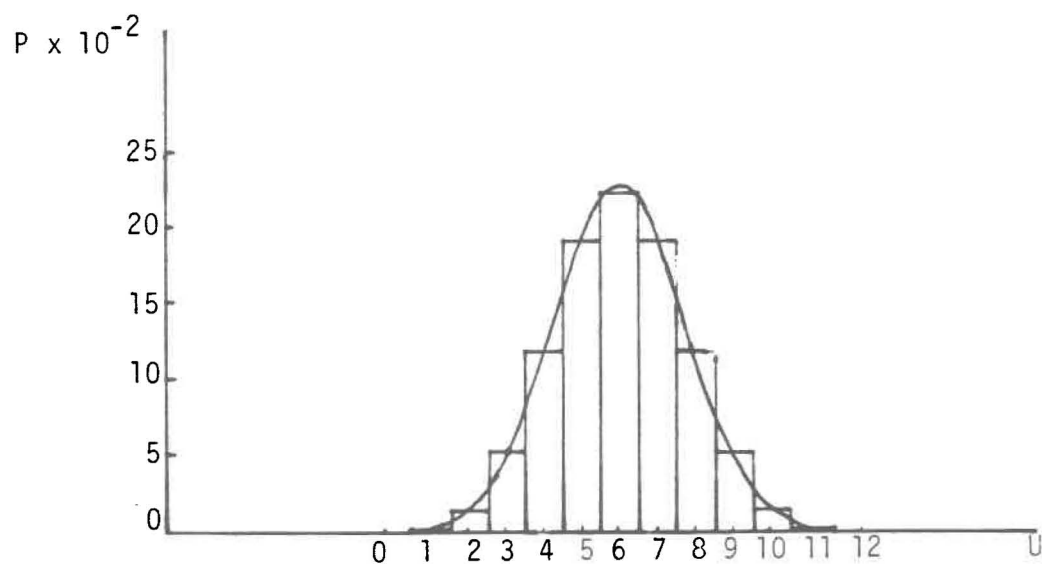


Figure A3.3 : Distributions exacte ($P(U)$: histogramme) et théorique ($G(U)$: courbe) pour $N = 2$ et $K = 12$

Tableau A3.7 : Effet de K sur l'ajustement N = 4

K	χ^2	χ^2_t	$\Delta \text{ ab.}$	$\Delta r.$
1	,0927	0,207	0,084	39,3
2	,0175	1,34	0,020	16,8
3	,0027	3,07	0,006	16,7
4	,0026	5,14	0,004	8,7

χ^2 : chi-deux calculé (ég. A3.7)
 χ^2_t : seuil théorique à $\alpha = 0,005$
 $\Delta \text{ ab}$: erreur absolue
 Δr : erreur relative (ég. A3.8)

Tableau A3.8 : Effet de N sur l'ajustement K = 2

N	χ^2	χ^2_t	$\Delta \text{ ab}$	Δr
4	0,0175	1,34	0,02	16,8
5	0,0052	3,07	0,01	6,67

Tableau A3.9 : Erreur maximale, N = 2, K = 12

U	$\Delta \text{ ab}$	Δr
6	,005	2,22 %
3	,002	3,85 %

ANNEXE 4

TEST D'ÉGALITÉ DES VARIANCES

Le test proposé ici est celui de Klotz (1962) tel que présenté par Marascuilo (1977). Ce test a la même puissance que son équivalent paramétrique (test du F) dans les conditions d'application de ce dernier. Dans d'autres cas, sa puissance est souvent supérieure. La statistique est donnée par :

$$K = \sum_{i=1}^n \theta_i^2 \quad (\text{éq. A4.1})$$

où :

θ_i = donnée par la table A.16 (Marascuilo, 1977). Dépend du rang i et de N ;

i = rangs des observations du plus petit échantillon dans l'ensemble des observations des deux échantillons;

n = nombre d'observations dans le plus petit échantillon;

N = nombre d'observations total.

On a ensuite :

$$E(K) = n_i \sum_{i=1}^N \theta_i^2 \quad (\text{éq. A4.2})$$

$$\text{Var} (K) = \frac{n_1 n_2}{N (N - 1)} \sum_{i=1}^N (\theta_i^4) - \frac{n_2 E(K)^2}{n_1 (N - 1)} \quad (\text{éq. A4.3})$$

$$Z = \frac{K - E(K)}{\sqrt{\text{Var} (K)}}, \text{ suit la loi } N(0, 1) \quad (\text{éq. A4.4})$$

Note : Ce test est sensible à une différence de moyennes des deux échantillons. Si ces moyennes sont différentes, il faut enlever ce biais en appliquant le test sur des résidus.

$$y'_i = y_i - \bar{y} \quad (\text{éq. A4.5})$$

$\bar{y}_i$ = moyenne du premier ou deuxième échantillon dépendamment de la provenance de y_i ;

y'_i = résidu.

ANNEXE 5

ESTIMÉ DE LA PENTE DE LA TENDANCE

La tendance détectée, il devient intéressant de la modéliser pour prédire les valeurs de la variable à différents horizons temporels.

En première approximation on peut être intéressé à une modèle linéaire, c'est-à-dire la pente moyenne de la tendance.

Pour les observations issues de populations normales, l'estimation paramétrique de la pente est:

$$B = \frac{n \sum x_i y_i - (\sum x_i) (\sum y_i)}{n \sum x_i^2 - (\sum x_i)^2} \quad (\text{éq. A5.1})$$

Si la distribution d'origine n'est pas normale, on utilise un estimateur non-paramétrique proposé par Theil (1950), repris par Sen (1968) et étendue aux cas des données subdivisées en sous-séries par Hirsch et al., 1982.

On doit calculer tous les d_{kij} définis par :

$$d_{kij} = (y_{kj} - y_{ki}) / (j - i) \quad (\text{éq. A5.2})$$

où :

k : indice de la sous-série;

i et j : indices des observations dans une sous-série;

n_k : nombre d'observations dans la sous-série K ;

$1 < i < j < n_k$;

L'estimé non-paramétrique de la pente est donné par la valeur médiane des d_{kij} .