Université du Québec
INRS-Eau, Terre et Environnement (ETE)

ÉTUDE COMPARATIVE DES TESTS DE STATIONNARITÉ

Par

Jean-Cléophas Ondo

(B.Sc., Mathématiques, Université Laval, 1994)
(M.Sc., Statistique, Université Laval, 1996)

Thèse présentée pour l'obtention du grade de
Philosophiae doctor (Ph.D.) en Sciences de l'Eau

Spécialité : Hydrologie Statistique

Jury d'évaluation

Examinateur externe	**Jean-Marie Dufour** Université de Montréal Département de sciences économiques
Examinateur externe	**Lynda Khalaf** Université Laval Département d'économique
Examinateur interne	**Marius Lachance** INRS-ETE
Directeur de recherche	**Bernard Bobée** INRS-ETE
Codirecteur de recherche	**Taha Ouarda** INRS-ETE

Avril 2002

© droits réservés de Jean-Cléophas Ondo, 2002

RÉSUMÉ

Une des hypothèses de base de l'analyse inférentielle en hydrométéorologie consiste à supposer que le phénomène observé est resté stationnaire. La stationnarité exprime que les propriétés statistiques du phénomène étudié sont indépendantes du temps. L'étude de la stationnarité de séries hydrométéorologiques revêt donc une grande importance en hydrologie et en climatologie.

La non-stationnarité des séries hydrométéorologiques a souvent été expliquée par l'existence de certaines dérives dans le temps, en particulier par une tendance déterministe ou stochastique. De nombreux tests statistiques ont été élaborés pour caractériser et détecter de telles dérives. Lubes et al. (1998) ont montré que la persistance et la présence d'une tendance dans les séries sont les deux caractéristiques qui pénalisent le plus les performances des tests habituellement utilisés en hydrométéorologie pour détecter une tendance déterministe. Or, il s'avère que ces deux caractéristiques sont généralement présentes dans les séries hydrométéorologiques. Il y a donc un réel besoin de méthodes statistiques nouvelles pour détecter une tendance déterministe dans de telles séries.

Dans cette thèse, on se préoccupe de la non-stationnarité de type rupture dans la pente déterministe. L'originalité de notre approche est double. D'une part, on présente une revue bibliographique des tests statistiques de stationnarité. D'autre part, on propose, pour la première fois dans la littérature hydrométéorologique, la technique des tests de Monte Carlo pour détecter une non-stationnarité de type rupture dans la pente déterministe. Dufour (1995) a démontré que ces tests, qui sont basés sur les méthodes randomisées, sont des tests exacts. En somme, la contribution de cette thèse peut être résumée ainsi :

1. *Sur la revue bibliographique* :
   a) on présente une revue de littérature extensive des tests de stationnarité ;
   b) on fait le parallèle entre trois littératures relativement disparates, à savoir, la littérature statistique, la littérature hydrométéorologique et la littérature économétrique.

2. *Sur ce qu'on apporte de nouveau par rapport à l'hydrologie*
   a) on introduit un test de Monte Carlo de détection d'une rupture ;
   b) on introduit un test de Monte Carlo qui formalise la méthode de segmentation binaire pour la recherche de plusieurs ruptures ;
   c) on généralise les tests qui sont introduits aux points a) et b) au cas de la dépendance sérielle (persistance) puis au cas régional (multivarié spatial) ;
   d) on conduit des expériences de simulations pour comparer la performance des tests proposés ;
   e) on réalise une application empirique des tests proposés.

REMERCIEMENTS

La réalisation de cette thèse fut un travail de longue haleine où le temps consacré à sa réalisation et les efforts pour le compléter ne se comptent que difficilement. Je peux dire, sans craindre de me tromper, que sans l'encadrement et le support de plusieurs personnes, il n'aurait pas été possible de le mener à terme. C'est pourquoi je tiens ici à remercier tous ces gens.

Je tiens à remercier en tout premier lieu M. Bernard Bobée et M. Taha Ouarda, mon directeur et mon codirecteur de thèse, pour m'avoir accueilli au sein de l'INRS-Eau, Terre et Environnement (ETE) et pour m'avoir proposé ce sujet de thèse qui m'a passionné durant ces quatre dernières années. Ils ont su me laisser une grande liberté pour la réalisation de ce travail.

Je suis infiniment reconnaissant envers M^{me} Lynda Khalaf pour sa disponibilité, ses conseils et tout l'aide quelle m'a apportée. C'est grâce à elle que je me suis initié à la technique des tests de Monte Carlo. Nos conversations m'ont convaincu de la pertinence et de la résonance de ma thèse dans des domaines autres que l'hydrologie statistique.

J'exprime toute ma reconnaissance envers tous les membres de mon jury, M. Marius Lachance, M. Jean-Marie Dufour, M^{me}. Lynda Khalaf, M. Bernard Bobée et M. Taha Ouarda, à qui je sais gré pour leur tolérance toute particulière.

Mes remerciements vont également à tous mes compatriotes et amis qui sont venus nombreux à ma soutenance. Un très grand merci va particulièrement à mon cher neveu Amvame-Bekale Auguste-Regis pour tout son soutient moral à la veille de cette soutenance.

Je n'oublierai pas le gouvernement de mon pays, le Gabon, pour son programme de bourse sans lequel cette thèse ainsi que ma formation universitaire n'auraient pas été possibles. Merci encore !

Finalement, je veux remercier mes parents pour qui je n'ai pas été aussi présent que j'aurais aimé l'être durant la réalisation de ce travail. Ils m'ont toujours soutenu, sans trop savoir jusqu'où tout cela pourrait me mener.

Merci à tous de m'avoir permis de vivre ça !

Obote Melac.

Je dédie cette thèse à la jeunesse africaine, et particulièrement celle du Gabon.

TABLE DES MATIÈRES

LISTE DES FIGURES

LISTE DES TABLEAUX

1 INTRODUCTION

1.1 Motivations

L'eau joue un rôle primordial dans la vie socio-économique des populations. Par exemple, Hydro-Québec, la plus grande entreprise hydroélectrique au monde et la septième plus grande compagnie d'électricité au monde, gère un réseau de systèmes hydriques de grande envergure, comprenant entre autres 40 grands réservoirs et 51 ouvrages hydroélectriques avec une puissance installée de 31512 MW (plus Churchil avec 5928 MW). L'exploitation d'un tel réseau, c'est-à-dire l'opération des réservoirs et des ouvrages est une activité importante pour Hydro-Québec. En effet, la gestion des réservoirs vise à optimiser la production tout en respectant certaines contraintes, telles que satisfaire la demande en électricité, garantir un débit minimal des rivières, et éviter des inondations (car pour atténuer l'impact d'une crue, il est souvent nécessaire d'évacuer un certain volume d'eau des réservoirs pour permettre à ceux-ci d'aborder de grands apports).

De manière générale, cette gestion ne se fait pas sans difficulté car la disponibilité de la ressource hydrique est incertaine. L'eau est de toute évidence une ressource aléatoire. Par conséquent, les apports aux réservoirs varient dans le temps et sont difficiles à prévoir dans le temps. Il s'agit là d'une question très pertinente que le gestionnaire d'Hydro-Québec doit trancher. Bien souvent, pour déterminer les risques de défaillance énergétique associés à différents scénarios d'exploitation du réseau et bien gérer la production électrique, ce dernier doit développer des modèles de planification (simulation et prévision) à court, moyen et long termes. Généralement ces modèles dépendent de paramètres inconnus (moyenne, écart-type, autocorrélation, etc.). L'estimation de ces paramètres est basée sur l'exploitation de séries de données recueillies pendant des périodes plus ou moins longues à partir des sites d'intérêt du réseau. De telles estimations constituent ce que l'on appelle habituellement des *intrants* des modèles de planification, en ce sens que se sont eux qui rendent opérationnels de tels modèles. Il en résulte que la qualité de ces estimations est très

importante. Les séries de données à partir desquelles on obtient ces estimations doivent bien représenter le phénomène hydrométéorologique sous étude si l'on veut que le modèle le représente adéquatement. Ainsi, il est indispensable, avant d'utiliser ces séries de données, de se préoccuper de leur qualité et de leur représentativité. En effet, il est possible que des modifications interviennent durant la période de mesure de ces données si bien que les séries de données résultantes soient dites non-stationnaires, c'est-à-dire qu'elles représentent deux ou plusieurs phénomènes hydrométéorologiques différents, lorsque c'est seulement un phénomène qui est à l'étude. Il y a également le problème de la qualité de ces données car une incertitude additionnelle peut être introduite par la mesure de la donnée elle-même. Les erreurs d'observation peuvent en effet être dues au comportement de l'observateur, de l'observé, des instruments utilisés dans l'observation, et/ou de l'environnement dans lequel l'observation est faite.

L'une des hypothèses de base des modèles de planification est que les données du phénomène hydrométéorologique étudié sont stationnaires. Cette hypothèse garantit, entre autres, la fiabilité sur les estimations obtenues à partir de telles données. Or, dans bien des cas, ces données mesurées à une échelle inférieure à l'année, voire même à l'année, sont fortement non-stationnaires (Mathier et al., 1992 ; Perreault et al., 1996). L'hypothèse de stationnarité est pour ainsi dire un précepte de base dans l'analyse des séries de données hydrométéorologiques car elle conditionne la qualité des intrants obtenus à partir de telles données.

1.2 Problématique de recherche

Les données de variables hydrométéorologiques, en particulier les informations sur la distribution spatio-temporelle des précipitations, constituent la base de notre connaissance du cycle hydrologique. Que ce soit dans le but de procéder à une étude du cycle de l'eau, d'impacts environnementaux ou pour procéder à la conception, la planification ou

l'optimisation de la gestion des systèmes hydriques, elles sont essentielles et constituent un préalable à toute analyse hydrologique.

La collecte de telles données consiste généralement à procéder, par le biais d'un instrument de mesure, à l'acquisition de l'information sur une variable hydrométéorologique donnée (par exemple : la hauteur d'eau d'une station limnimétrique, la température etc). Au vu de l'importance qualitative et quantitative des données recueillies, il importe de les contrôler en termes de fiabilité et de précision. Ce contrôle permet essentiellement de les valider avant leur mise en disposition à des fins opérationnelles. Pour l'hydrologue, l'élément ultime de cette chaîne opératoire est la mise en œuvre d'un modèle hydrologique pour faire la gestion et la planification des ressources en eau. En effet, certaines caractéristiques statistiques issues de telles données (moyenne, écart-type, autocorrélation) sont utilisées comme intrants à de tels modèles de planification. Ces intrants sont de ce fait liés à l'environnement hydrologique naturel de leur obtention. Il va sans dire qu'on peut établir un modèle hydrologique très sophistiqué, mais si les données d'entrée sont mauvaises, la sortie rendra caduque toute application.

En pratique, l'acquisition d'une série de données à partir d'une variable hydrométéorologique, constituant un échantillon au sens statistique du terme, est un processus long, parsemé d'embûches, et au cours duquel de nombreuses erreurs, de nature fort différente sont susceptibles d'être commises. En effet, des erreurs peuvent, d'une part, être perpétrées lors de l'acquisition des données (déplacement de la station de mesure, observateur unique, etc.) conduisant ainsi à des valeurs manquantes, imprécises ou erronées. D'autre part, des erreurs peuvent survenir lors de la mesure du phénomène. Ces dernières se traduisent principalement par une ou plusieurs ruptures de continuité dans l'échantillon de données enregistrées. C'est pourquoi il est important, avant d'utiliser de telles séries de données, de se préoccuper de leur qualité et de leur représentativité en utilisant diverses techniques mathématiques ou statistiques. Ces techniques peuvent être rassemblées dans les catégories suivantes :

- techniques de recherche des points singuliers ("outliers") ;
- techniques de détection de non-stationnarités ;
- techniques basées sur la modélisation.

Sans entrer dans le détail de ces techniques qui font pour chacune l'objet d'une littérature abondante, on abordera dans cette thèse les techniques de détection de non-stationnarité dans un échantillon de données : les tests de stationnarité.

Une série de données est dite stationnaire si elle provient d'un phénomène dont les caractéristiques sont indépendantes du temps. Inversement, une série de données est dite non-stationnaire si les caractéristiques statistiques du phénomène évoluent durant la période de mesure. La prise en compte de l'hypothèse de stationnarité est donc capitale pour garantir la fiabilité des modèles de planification des ressources en eau. On peut affirmer, sans exagération, que si cette hypothèse est négligée lors de la conception et du dimensionnement d'ouvrages hydrauliques cela peut entraîner des causes aux conséquences économiques et humaines importantes : risques de débordement du barrage, dommages matériels, inondations dévastatrices pour les populations riveraines, pertes en vie humaines, etc.

L'étude de la stationnarité des séries de données de variables hydrométéorologiques est pour ainsi dire une des tâches incontournables des hydrologues (Bernier, 1977). Et, depuis plusieurs années elle cherche aussi à répondre aux questions qui relèvent de la planification de réseaux hydrométriques pour le suivi des modifications climatiques et de ses effets sur les ressources en eau. Pour s'en rendre compte, il suffit de constater que le régime du cycle de l'eau est l'une des manifestations majeures du climat. Son suivi permet donc d'appréhender l'évolution de certains aspects climatiques. Cela est d'autant plus intéressant puisque depuis le début des années 1970, des scientifiques du monde entier ont commencé à soupçonner que certaines activités humaines, particulièrement l'utilisation des combustibles fossiles, ont un impact sur le climat. Selon toute vraisemblance, l'éventualité d'une modification climatique à l'échelle globale rendra plus complexe le problème de la prévision future des réserves hydrologiques et de leur exploitation. En effet, la modification climatique laisse craindre qu'il se produira au moins une évolution des

événements hydrologiques extrêmes qui touchent de manière directe les ressources en eau, l'écosystème entier et en particulier la population. Ainsi par exemple, à cause d'une augmentation des précipitations hivernales, les risques de crue (fréquence et intensité) de toutes les rivières seraient plus élevés qu'actuellement. Les implantations humaines dans les zones inondables seraient donc plus fortement menacées, requérant de ce fait le respect de normes plus strictes pour leur implantation. De plus, les infrastructures de contrôle des rivières comme les barrages d'écrêtement des crues construits sous les conditions actuelles pourraient ne plus être totalement adaptées. L'analyse et la caractérisation précise des manifestations de la variabilité du climat, et leur lien avec la variabilité des ressources en eau, constituent donc, aujourd'hui, un domaine de recherche d'actualité. L'étude de stationnarité qui résulte de la détection de variabilité du climat peut, dans un certain sens, aider à développer des modèles qui permettent de rendre compte de la relation pluie-débit, afin d'évaluer l'impact de différents scénarios sur la gestion de ressources en eau. La détection de non-stationnarité dans des séries hydrométéorologiques constitue donc un projet de recherche qui a toute sa mesure dans cette thèse.

1.3 Objectifs généraux

Pour les hydrométéorologues, les différents types de non-stationnarités considérés dans les séries de données sont :

- le changement de la moyenne (saut brusque de la moyenne, succession de changements de la moyenne, tendance continue de la moyenne, etc.) ;
- le changement de la variance (augmentation ou diminution).

Il existe un certain nombre de tests statistiques pour vérifier ces différents types de discontinuités ou dérives. Ce sont des tests du caractère aléatoire et simple d'une suite de variables indépendantes, qui impliquent l'hypothèse de non-organisation chronologique de l'échantillon observé. Par construction, ces tests varient généralement selon que l'on redoute tel ou tel défaut de non-stationnarité. Cependant, le choix entre les différentes

statistiques de test et leurs différents modes d'utilisation n'est pas aisé. Il dépend fortement d'hypothèses préalables faites sur les données (par exemple : les échantillons sont supposés être distribués selon une loi normale). De plus, ces tests ne tiennent pas souvent compte de l'autocorrélation temporelle qui peut caractériser les variables hydrométéorologiques de grande fréquence. Il aurait donc été souhaitable que leurs utilisateurs rendent compte des problèmes rencontrés afin d'améliorer les logiciels fournis tel que STATISTI (aujourd'hui KHRONOSTAT, logiciel élaboré par les équipes de l'IRD) par exemple.

L'ambition de cette thèse est de proposer de nouveaux tests de stationnarité dans la littérature hydrométéorologique. Ces tests sont construits à partir de la théorie des tests de Monte Carlo. Dufour (1995) a démontré que les tests de Monte Carlo sont des tests exacts. En exploitant méthodiquement cette importante propriété, on propose des tests exacts de détection de non-stationnarités. L'approche proposée est ambitieuse car elle vise d'une part à une prise en compte des caractéristiques que l'on retrouve dans les séries de données de variables hydrométéorologiques (séries de données à effectifs faibles qui ne sont pas nécessairement, indépendantes). D'autre part, elle se révèle utile dans la mesure où elle cherche à apporter des éléments de réponse nouveaux, complémentaires, pertinents et déterminants aux questions habituelles suivantes :

- à l'échelle locale y a-t-il eu un (ou plusieurs) changement de moyenne ou de variance ?
- à l'échelle régionale y a-t-il eu un changement (de moyenne ou de variance) ?

Afin de mettre en avant cette nouvelle méthode dans la littérature hydrométéorologique des tests de détection de non-stationnarités, notre initiative de recherche se préoccupe essentiellement de la détection de ruptures (ou encore changement de régime) qui est le type de non-stationnarité le plus courant dans les séries de données hydrométéorologiques (Hubert et al., 1989). Il en résulte aussi qu'une généralisation à d'autres types de non-stationnarité est tout aussi possible.

Le problème de rupture a fait et continue de faire l'objet d'une littérature abondante en hydrométéorologie. Les tests les plus utilisés et les mieux décrits se repartissent habituellement dans deux catégories. La première concerne le cas où la date de rupture est supposée connue. Les tests résultants comparent les moyennes avant et après la date de rupture. La deuxième catégorie concerne la situation la plus vraisemblable où la date de rupture et parfois même l'existence de la rupture est inconnue. Très peu de tests existent dans la littérature pour traiter ce cas de première importance en hydrométéorologie. Les principaux tests qui s'appliquent à la dernière catégorie et qui ont été largement appliqués aux analyses de séries de données d'une variable hydrométéorologique sont le test de Buishand (Buishand, 1982), le test de Pettitt (Pettitt, 1979) et le test de Worsley (Worsley, 1979). Ces tests sont exclusivement de type supremum (test de Pettitt et de Worsley) et ils sont construits sous l'hypothèse asymptotique. Lubes et al. (1998) ont évalué le comportement de ces tests dans des situations pratiques. Il semble, d'après leurs conclusions, que ces tests présentent plusieurs inconvénients dont les plus importants sont :

- l'absence de bonnes valeurs critiques ;
- le problème de contrôle de niveaux et de puissance pour des échantillons avec persistance (ou présentant une tendance);
- le problème du choix de la bonne date de rupture pour des échantillons avec persistance (ou présentant une tendance) ;

Ainsi, appliqués dans certaines conditions, ces tests usuels peuvent amener l'ingénieur hydrologue à prendre de mauvaises décisions dans sa tâche de planification des ressources en eau. Bien que la littérature hydrométéorologique recèle quelques tests pour traiter le problème de rupture à une date inconnue, les études qui se sont directement intéressées au problème relatif à plusieurs ruptures dans un échantillon sont peu nombreuses. La seule procédure connue à ce jour en hydrologie et qui est largement employée dans les applications pratiques est celle de Hubert et al. (1989). D'autre part, cette même littérature est dépourvue d'un test régional de détection de ruptures.

La technique des tests de Monte Carlo qu'on met en avant dans cette thèse se trouve donc être une alternative intéressante capable d'apporter des solutions novatrices dans le problème de détection de ruptures en hydrométéorologie. Parmi les principaux objectifs à atteindre, on cherche à étendre les travaux les plus usuels dans deux directions : on propose d'une part des tests exacts de Monte Carlo de type supremum pour détecter une (ou plusieurs) rupture dans un échantillon de données indépendantes ou autocorrélées. D'autre part, on propose une procédure régionale exacte de détection de ruptures. Dans ce contexte, les problèmes statistiques auxquels on apporte une solution sont :

- Y a-t-il une (ou plusieurs) rupture significative dans les données ?
- Si la rupture existe, quelle est la structure de cette rupture : brusque ou avec continuité ?
- Si la rupture existe localement, est-il raisonnable de l'admettre régionalement ?

À côté de ces objectifs généraux, on vise aussi des objectifs spécifiques. En tout premier lieu, il nous semble important de présenter un état de l'art des activités de recherche dans la littérature des tests statistiques de détection de non-stationnarité en hydrométéorologie et dans des domaines connexes comme la statistique et l'économétrie. Puis, il semble important de faire une étude comparative des tests qui sont proposés dans cette thèse avec ceux conventionnellement utilisés dans la littérature.

Bien que les travaux de cette thèse touchent principalement deux disciplines : le génie hydrologique et la statistique, les résultats exposés sont d'une portée très générale. Ils peuvent par conséquent être appliqués à tout autre domaine d'étude traitant du sujet des séries chronologiques. On peut citer des domaines tels que la statistique, l'économétrie, etc.

La suite de ce travail se divise en neuf chapitres :

Chapitre 2 : on présente quelques notions de base des termes statistiques qui reviennent le plus souvent dans le document. En particulier, on introduit les concepts fondamentaux de la notion de stationnarité.

Chapitre 3 : on développe un certain nombre de généralités sur la théorie des processus stochastiques afin de préciser et de définir la notion de stationnarité.

Chapitre 4 : on analyse dans ce chapitre la question de la non-stationnarité des processus hydrométéorologiques.

Chapitre 5 : on présente une revue bibliographique des principaux tests statistiques de stationnarité en hydrométéorologie.

Chapitre 6 : on s'intéresse d'abord à quelques généralités sur les tests statistiques. Puis, on présente la théorie fondamentale des tests Monte Carlo.

Chapitre 7 : on présente les tests de stationnarité qui sont proposés dans cette thèse.

Chapitre 8 : on effectue des études de simulations de type Monte Carlo pour évaluer les performances des tests qui sont proposés dans cette thèse.

Chapitre 9 : on présente une application des tests proposés aux données de suivi des changements climatiques de la province de Québec.

Chapitre 10 : ce dernier chapitre est consacré aux conclusions et aux perspectives de recherche.

2 FONDEMENTS STATISTIQUES DE LA NOTION DE STATIONNARITÉ

Une notion capitale pour aborder l'étude d'un phénomène stochastique est la notion de stationnarité. Il est en effet a priori intéressant de savoir si un processus qui représente un phénomène donné va stabiliser ou non sa distribution de probabilité, si cette distribution stable est atteinte au bout d'un temps fini ou non, si elle dépend ou non de l'état initial du processus. Avant donc d'aborder en détail les questions qui se rattachent à la notion de stationnarité, il semble indispensable de rappeler un certain nombre de généralités sur cette notion. C'est le but de ce chapitre qui a surtout un but pédagogique en hydrologie. Il est en effet bon de repartir de la base afin de mieux appréhender cette notion de stationnarité. C'est pourquoi ce chapitre est l'occasion de faire un rappel volontairement très bref et élémentaire sur la notion de stationnarité sans obstacles d'ordre théorique.

2.1 Généralités

Les phénomènes hydrométéorologiques sont généralement considérés comme des phénomènes réels. Le premier élément descriptif de tels phénomènes consiste à préciser ce qui peut se produire. Par exemple, les débits collectés à une station de mesure d'un cours d'eau constituent une réalisation d'un phénomène hydrologique. Ces réalisations décrivent avec précision le déroulement du phénomène, ici le débit. Signalons ici quelques questions techniques dont la solution fait intervenir l'étude du phénomène :

- Quelle valeur approchée peut-on proposer pour le débit moyen μ_t de cette rivière ?
- Est-ce que le débit moyen actuel de cette rivière est ou non supérieur à la valeur moyenne conventionnellement utilisée dans les études antérieures ?
- Quel serait le comportement futur du débit de cette rivière ?

Pour de telles questions concrètes, l'approche statistique se propose de construire des modèles mathématiques basés sur un certain nombre d'observations et d'hypothèses et cherchant à décrire le mieux possible le phénomène auquel on s'intéresse. L'objet de la statistique est donc de procéder à une analyse mathématique de phénomènes dans lesquels le hasard intervient. Ces phénomènes sont appelés des *phénomènes stochastiques*. Un phénomène est dit stochastique ou aléatoire si, reproduit maintes fois dans des conditions identiques, il se déroule chaque fois différemment de telle sorte que le résultat de l'expérience change d'une fois à l'autre de manière imprévisible (le vocabulaire de base en statistique est donné à l'appendice A). Bien que les comportements aléatoires sont a priori sujets à des variations imprévisibles, on est cependant capable de donner des renseignements sur ce type de phénomènes. L'idée majeure est que ces renseignements vont être donnés dans le cadre d'un modèle statistique. On peut donc, dans une certaine mesure, vérifier la cohérence de la représentation d'un tel modèle, en confrontant les conclusions auxquelles conduisent le traitement statistique avec les résultats observés dans le déroulement du phénomène. Toutefois, il ne faut pas oublier que cette vérification repose elle aussi sur une interprétation et ne représente pas ainsi un caractère absolu. En d'autres termes il serait erroné de croire que des méthodes statistiques permettent de démontrer «mathématiquement» que les événements associés au comportement réel du phénomène se déroulent de telle ou telle manière. Cependant, il convient de se poser la question de l'adéquation d'un modèle statistique donné ; en d'autres mots ce dernier offre-t-il une image suffisamment fidèle du phénomène qu'il veut représenter ou faut-il plutôt chercher à l'améliorer ? De telles questions sont du ressort de cette thèse.

2.2 Analyse fréquentielle et la notion de stationnarité

L'étude des phénomènes réels porte en général sur les résultats de l'épreuve (échantillon observé) et on souhaite habituellement utiliser ces informations dans un modèle. Par exemple, dans une étude du débit d'une rivière durant une période de 30 ans, on peut être intéressé à prévoir les crues de cette rivière pour dimensionner un barrage compte tenu des conséquences économiques. Si une telle approche doit être menée en utilisant les

techniques inférentielles, le problème évoqué précédemment implique une hypothèse supplémentaire sur la nature des observations que l'on analyse. En effet, celles-ci doivent obéir au postulat de base de l'analyse fréquentielle (voir appendice B). Autrement dit, la série de données observées doit être modélisée par une suite de variables aléatoires indépendantes et de même loi de probabilité.

2.2.1 Homogénéité de la série de données observées

Lorsqu'on étudie un phénomène relativement à un caractère mesurable X, si l'ordre dans lequel on fait les prélèvements dans la population du phénomène n'a pas d'importance, on choisira de considérer que les données observées $\{x_1,\ldots,x_n\}$, sont des réalisations de variables aléatoires indépendantes $\{X_1,\ldots,X_n\}$ et de même loi de probabilité. L'idée naturelle que l'on exploite pour étudier ces données en analyse inférentielle est de répéter l'épreuve un certain nombre de fois dans des conditions identiques pour obtenir plusieurs échantillons observés. Ainsi, disposant de plusieurs échantillons, on peut vérifier s'ils sont tirés dans la même population, c'est-à-dire de vérifier s'ils sont générés à partir de la loi de probabilité de la variable parente X. On peut ainsi chercher à obtenir certaines indications fiables sur diverses caractéristiques de la population parente. Par exemple, on pourrait effectuer des inférences sur les paramètres qui caractérisent l'essentiel de la loi de probabilité de la variable parente X, c'est-à-dire de les estimer à partir d'échantillons observés résultant de répétitions successives. Plus le nombre de répétitions seraient élevées, plus d'échantillons observés seraient à notre disposition et, par conséquent, plus on pourra juger de la stabilité de ces estimations. La stabilité des estimations fournit une idée approximative de ce qu'est la distribution de la population originelle. Elle représente ainsi un moyen efficace permettant de juger du degré de parenté des échantillons observés. En conséquence, elle permet de conclure qu'ils ont été tirés dans la même population.

La stabilité de telles estimations permet de conclure que les échantillons obtenus au cours de ces répétitions proviennent bien de la mesure d'un phénomène dont les caractéristiques restent stables. Ainsi, une série de données observées est dite homogène si elle provient de la mesure d'un phénomène aléatoire dont les caractéristiques n'ont pas évolué durant la période de mesure. À l'opposé, une série de données est dite hétérogène (non homogène)

lorsqu'elle provient de la mesure d'un phénomène aléatoire dont les caractéristiques évoluent durant la période de mesure. En reprenant l'exemple précédent, l'homogénéité garantit que toutes les mesures de débits observées appartiennent bien à la même famille, c'est-à-dire au même phénomène hydrologique. Par ailleurs, on dit d'une série de données qu'elle est hétérogène si elle reflète deux ou plusieurs phénomènes aléatoires différents. Par exemple, une série de mesures de débits de crue d'automne et de printemps à un endroit d'une rivière constitue un bon exemple de défaut d'homogénéité dû à l'influence de deux phénomènes différents.

2.2.2 Stationnarité de la série de données observées

En principe, il existe plusieurs façons de procéder au choix des prélèvements des éléments d'une population statistique d'un phénomène donné. Lorsque l'on est maître de ce choix, on considère, d'après le postulat de base de l'analyse fréquentielle, que l'échantillon observé peut être modélisé par une suite de variables aléatoires indépendantes et de même loi de probabilité : on peut alors utiliser les résultats des probabilités et ceux de l'analyse fréquentielle qui sont établis à partir de la répétition de l'épreuve dans les conditions identiques pour estimer les caractéristiques statistiques de la population. Mais quand considère-t-on qu'un échantillon peut être modélisé par une suite de n variables aléatoires et indépendantes de même loi de probabilité ? C'est précisément quand l'ordre dans lequel on prélève les données de l'échantillon observé n'a pas d'importance. Par contre, ce n'est pas le cas lorsque les données qui composent l'échantillon observé sont produites par la nature. C'est cette dernière situation qui prévaut dans de nombreux domaines comme l'hydrologie, la climatologie, l'économie,... où on enregistre des données au cours du temps pour constituer ce qu'on appelle une *série chronologique*. De telles données ne sont donc pas interchangeables et donc pas indépendantes.

De façon plus précise, une série chronologique (ou *série temporelle* ou encore *chronique*) est une suite d'observations correspondant à des mesures, repérées dans l'ordre croissant du temps, sur un même phénomène. Les débits enregistrés à chaque heure à une station de mesure constituent un bon exemple de série chronologique. Les échantillons observés produits par la nature diffèrent donc des premiers en ce sens que leurs éléments se

présentent nécessairement dans un ordre donné-celui de leur apparition dans le temps- et que c'est à cet ordre, au surplus, que se trouve attachée leur signification. Les propriétés habituelles de description et d'analyse inférentielle ne peuvent donc leur être étendues que sous certaines conditions. En effet, dans le cas d'une série chronologique, comme on ne peut appliquer le caractère reproductible des données d'après le postulat de base en statistique inférentielle, on ne connaît donc, pour un nombre fini de valeurs temporelles, qu'une seule réalisation du phénomène étudié. Sans une hypothèse supplémentaire, on ne peut utiliser les résultats des probabilités ni ceux de l'analyse fréquentielle qui sont établis à partir de la répétition de l'épreuve dans des conditions identiques. C'est d'autant plus préoccupant puisque pour estimer les paramètres caractéristiques d'un phénomène donné, on doit se contenter d'un échantillon observé (série chronologique) qui comprend les données d'un seul essai. D'autre part, on n'est pas à même de récréer un environnement expérimental identique à celui correspondant aux observations, dans l'espoir d'obtenir plus d'échantillons observés, comme cela est particulièrement le cas en théorie inférentielle classique (conséquence du postulat de base donné à l'appendice B). En conséquence, cela voudrait-il dire que, par l'impossibilité de répéter les essais, la stabilité et la fiabilité des estimations effectuées à partir du seul échantillon observé (la série chronologique) seraient faibles sinon totalement déficientes ? Fort heureusement, il n'en est pas nécessairement ainsi. L'hypothèse de stationnarité, que l'on définit dans le chapitre suivant permet de pallier cette difficulté. En effet, une série chronologique (échantillon observé du phénomène stochastique étudié) est stationnaire lorsque les propriétés de la loi de probabilité qui régit le phénomène stochastique (moyenne, variance ou moments d'ordre supérieur) sont invariables au cours du temps. L'hypothèse de stationnarité est donc un préalable dans l'analyse fréquentielle d'une série de données chronologiques. Elle est de ce fait une condition nécessaire qui donne lieu à une interprétation rigoureuse des propriétés habituelles de l'analyse inférentielle.

En somme, si les caractéristiques statistiques du phénomène stochastique étudié varient durant la période de mesure on dit de ce dernier qu'il est non-stationnaire. La stationnarité peut se résumer à l'homogénéité temporelle. On reviendra sur ce point au chapitre 4.

2.3 Conclusion

Dans ce chapitre, on a mis en avant les bases de la compréhension de la notion de stationnarité dans le cadre de l'inférence statistique. L'hypothèse de stationnarité est un préalable au traitement statistique d'une série de données temporelles. Si cette hypothèse est relâchée, et qu'en particulier on travaille sur des séries non-stationnaires, alors les méthodes de l'analyse inférentielle deviennent non fondées, et par conséquent les estimations des paramètres calculer à partir de la série temporelle ne sont plus fiables.

3 STATIONNARITÉ : DÉFINITION ET HYPOTHÈSES DE DÉTECTION

Le chapitre précédent a permis de prendre conscience de l'importance de l'hypothèse de stationnarité dans l'analyse fréquentielle d'une série chronologique. Dans ce chapitre, on va, dans un premier temps, définir de façon plus précise la notion de stationnarité. Puis, dans un second temps, on tentera de comprendre les causes de non-stationnarités d'une série chronologique. Les idées qu'on va développer ici seront indispensables pour entreprendre l'élaboration d'un test statistique de détection de non-stationnarité.

La plupart des notions que l'on trouve dans ce chapitre sont conformes à celles qu'on trouve dans les ouvrages de Doob (1953), de Bartlett (1956) et de Girault (1965).

3.1 Processus stochastiques et séries chronologiques

3.1.1 Processus stochastiques

La démarche conduisant à la notion de processus est un peu analogue à celle utilisée pour introduire les variables aléatoires à l'appendice A. On rappelle pour mémoire qu'une variable aléatoire réelle est une application particulière d'un espace probabilisé $(\Omega, \mathcal{A}, \mathrm{P})$ dans l'ensemble des nombres réels $\mathbb{R}$ et que la variable statistique est une suite de réalisations de cette variable aléatoire. En nous servant de cette définition et de celles données à l'appendice A, on peut définir rigoureusement un *processus aléatoire.*

Définition 3.1. *Soit* $(\Omega, \mathcal{A}, \mathrm{P})$ *un espace probabilisé et* $(\Omega', \mathcal{A}')$ *un espace probabilisable. Soit* $(\mathrm{T}, \Im)$ *un espace quelconque dans lequel* $\Im$ *est la tribu des événements de* T. *On appelle processus aléatoire, l'application X telle que :*

$$X : (\Omega, \mathcal{A}) \times (\mathrm{T}, \Im) \rightarrow (\Omega', \mathcal{A}')$$

Cette application associe au couple (ω,t) la quantité $X(\omega,t)$ notée $X_t(\omega)$. Pour tout t appartenant à T fixé, X_t est une variable aléatoire sur l'espace probabilisé $(\Omega,\mathcal{A}, \mathrm{P})$. En pratique, le couple $(\Omega',\mathcal{A}')$ appelé espace des états coïncide le plus souvent avec l'un ou l'autre des couples $(\mathbb{R},\mathcal{B})$ pour un *processus univarié* ou $(\mathbb{R}^n,\mathcal{B}^n)$ pour un *processus multivarié*. On rappelle ici que, $\mathbb{R}$ est l'ensemble des nombres réels et $\mathcal{B}$ la σ_algèbre des boréliens sur $\mathbb{R}$. À partir de ces constations, on peut maintenant définir un *processus stochastique*.

Définitions 3.2. *Un processus stochastique est une famille de variables aléatoires indicées par t noté* $\{X_t, \quad t\in \mathrm{T}\}$ *(ou encore* $\{X_t(\omega), \quad t\in \mathrm{T}\}$*).*

De façon formelle, un processus stochastique est une famille de variables aléatoires $\{X_t, \quad t\in \mathrm{T}\}$ définies sur un même espace de probabilité $(\Omega,\mathcal{A}, \mathrm{P})$ et à valeurs dans $\mathbb{R}^n$. On parle de série chronologique univariée (respectivement, multivariée) lorsque $n=1$ (respectivement, $n>1$). T est l'ensemble des valeurs possibles pour le paramètre t qui représente le temps dans cette thèse. Le *processus* est dit à *temps discret* si T est fini ou dénombrable et on dira qu'il est à *temps continu* dans le cas contraire. Dans le cadre de cette thèse, on supposera que T = $\mathbb{N}$.

D'un point de vue pratique, un processus $\{X_t, \quad t\in \mathrm{T}\}$ est défini comme une collection de variables aléatoires. Une distribution de probabilité peut donc être associée à chaque variable, indiquant les probabilités de réalisation de valeurs particulières. On caractérise la loi de probabilité du processus stochastique par une mesure dans $\mathbb{R}^n$. Cette mesure est représentée par sa fonction de répartition :

$$F_X(x_1,\ldots,x_k)=\mathrm{P}_X\left(X_{t_1}\le x_1,\ldots,X_{t_k}\le x_k\right)$$

Précisons que la loi de probabilité P_{X_i} de la variable aléatoire X_i est appelée *probabilité marginale* de P_X sur la $i^{\text{ième}}$ composante. Ces notions prolongent donc, de façon évidente, les notions identiques définies pour une variable aléatoire (appendice A). On indique,

d'ailleurs, que dans le cas des processus stochastiques la notion de fonction de répartition est notablement plus lourde que dans le cas d'une variable aléatoire, et est en conséquence assez peu utilisée. Comme on sait que le comportement d'une variable aléatoire est complètement décrit par sa distribution de probabilité, il en sera de même pour un processus stochastique $\{X_t, \quad t \in \mathrm{T}\}$ qui se trouvera complètement caractérisé par la distribution jointe des X_t : $\mathrm{P}_X\left(X_{t_1},\ldots X_{t_n}; t_1,\ldots,t_n\right)$. En pratique, la spécification des moments de la distribution jointe suffit à décrire complètement la loi de probabilité qui gouverne le processus. En particulier,

- la moyenne : $\mu_{X_t} = E(X_t)$
- la variance : $\sigma_{X_t} = Var(X_t)$
- l'autocovariance au retard θ : $\gamma_{X_t}(t, t+\theta) = Cov(X_t X_{t+\theta})$

En l'absence d'hypothèses on note que ces moments dépendent du temps t.

3.1.2 Séries chronologiques

D'après la définition 3.2, on peut assimiler un processus stochastique à un échantillon théorique $\{X_t, \quad t \in \mathrm{T}\}$. Ainsi, une réalisation de cet échantillon théorique, $\{x_t, \quad t \in \mathrm{T}\}$, est appelée série chronologique ou encore *trajectoire observée* du processus stochastique $\{X_t, \quad t \in \mathrm{T}\}$. De manière plus précise, on adoptera la définition suivante

Définition 3.3. *On appelle série chronologique, chronique ou encore série temporelle* $\{x_t, \quad t \in \mathrm{T}\}$ *une réalisation d'un processus stochastique* $\{X_t, \quad t \in \mathrm{T}\}$.

La série chronologique ainsi définie est dite trajectoire observée ou encore réalisation du processus stochastique. Ce dernier est appelé *processus générateur de la série chronologique*.

3.2 Stationnarité et ergodicité

3.2.1 Généralités

Une hypothèse simplificatrice qu'on fait souvent pour soumettre une série chronologique $\{x_t, \quad t \in \mathrm{T}\}$ à une analyse fréquentielle est de supposer que la série en question est une réalisation d'un processus stochastique $\{X_t, \quad t \in \mathrm{T}\}$. Le principe est que, d'après la définition 3.2, un processus stochastique peut être vu comme une population qui a la dimension «temps», c'est-à-dire que les éléments de cette population sont fonction du temps ou encore qu'il y a une population en chaque instant de temps *t*. De fait, X_1 est une variable aléatoire différente par exemple de X_2 ou X_3 etc. Une distribution de probabilité peut alors être associée à chaque variable aléatoire, indiquant les probabilités de réalisation de valeurs particulières. Ainsi, considérer une série chronologique comme un échantillon tiré de cette population, revient à admettre, d'après le postulat de base en analyse inférentielle, qu'on pourrait générer d'autres échantillons, c'est-à-dire d'autres séries chronologiques qui ne sont pas observées. De cette façon, on établit plus facilement le lien avec la théorie de l'analyse fréquentielle. Si maintenant on veut faire une analyse fréquentielle sur une série chronologique des débits journaliers d'une rivière à une station de mesure donnée, soit $\{x_1, x_2, \ldots, x_n\}$, on va considérer x_1 comme une observation unique sur la population de tous les débits possibles pour cette rivière au temps $t = 1$, x_2 comme une observation unique sur la population de tous les débits possibles au temps $t = 2$ et ainsi de suite. Il faut donc comprendre ici que le débit correspondant à chaque période est comme une observation générée par une population qui aurait pu en générer une infinité d'autres.

La connaissance d'une réalisation du processus stochastique ne permet pas cependant d'en déduire des propriétés intéressantes établies à partir du postulat de base de l'analyse fréquentielle : procéder à la répétition de l'épreuve dans des conditions identiques, pour estimer les caractéristiques de la population. L'hypothèse de stationnarité, qu'on définit ci-dessous, permet d'étendre les propriétés habituelles de l'analyse fréquentielle aux séries

chronologiques. Elle correspond à la modélisation suivante : chaque observation d'une série chronologique est considérée comme la somme d'une valeur réelle déterminée en fonction du temps et d'une valeur réelle résultant du hasard. Par conséquent, la suite d'observations constituant la réalisation du processus stochastique (la série chronologique) sont supposées être des réalisations d'une variable aléatoire X (la variable qui régit le phénomène sous-jacent) dont les paramètres ne dépendent pas du temps. Ce n'est que sous cette réserve que l'on peut appliquer les résultats de l'analyse fréquentielle aux séries chronologiques puisque l'on dispose alors d'un échantillon $\{X_t, \quad t \in \mathrm{T}\}$ de la variable aléatoire parente X.

En d'autres termes, un processus stochastique est stationnaire quand la loi de probabilité qui régit la valeur de chacune des variables aléatoires dont il est la succession, est indépendante de la date t. On voit donc apparaître l'importance de cette hypothèse pour estimer les moments de la loi de probabilité de la variable parente X à partir des moments empiriques d'une réalisation du processus stochastique sous-jacent. On peut donc dire que la stationnarité induit un équilibre statistique, en ce sens que les modèles de probabilités sous-jacents sont inchangés dans le temps.

3.2.2 Stationnarité : définitions

3.2.2.1 Stationnarité au sens strict

Une série chronologique de réalisations d'une grandeur aléatoire, à un pas de temps donné, est dite stationnaire si ses réalisations sont issues d'un même processus stochastique dont les paramètres (moyenne, variance, asymétrie, autocorrélation ...) restent constants au cours du temps. De façon plus mathématique, la stationnarité au sens strict sera mieux interprétée par la définition suivante.

Définition 3.4. *Un processus stochastique* $\{X_t, \quad t \in \mathrm{T}\}$ *est dit strictement ou fortement stationnaire si ses propriétés, c'est-à-dire probabilités, ne dépendent pas de l'instant t :*

$$(\forall h), (\forall n), \quad \mathrm{P}(X_1,\ldots,X_n : t_1,\ldots,t_n) = \mathrm{P}(X_1,\ldots,X_n : t_1+h,\ldots,t_n+h)$$

La définition précédente stipule qu'un processus sera dit stationnaire au sens strict quand les lois de probabilité pour chaque instant t sont identiques, et lorsque la loi de probabilité conjointe pour deux instants t_1 et t_2 est invariante pour toute translation du temps. Les caractéristiques (moments) d'un processus strictement stationnaire sont donc invariantes pour tout changement de l'origine du temps.

3.2.2.2 Stationnarité au sens faible

En principe, la stationnarité au sens strict est trop restrictive et on assouplit cette condition en définissant la stationnarité à l'ordre k. Dans ce cas précis, les probabilités peuvent évoluer mais on suppose que les moments d'ordre k ne dépendent pas du temps.

Définition 3.5. *Le processus* $\{X_t,\ t\in \mathrm{T}\}$ *est dit stationnaire à l'ordre k si, pour tout* $(t_1,\ldots,t_n)$ *avec* $t_i \in \mathrm{T}$, *pour* $i=1,\ldots,n$, *et si pour tout* $h\in \mathrm{T}$ *avec* $t_{i+h} \in \mathrm{T}$, *tous les moments joints d'ordre k de* $\{X_{t_1+h},\ldots,X_{t_n+h}\}$ *existent et sont identiques aux moments joints correspondants de* $\{X_{t_1},\ldots,X_{t_n}\}$ *:*

$$E\left\{\left(X_{t_1}\right)^{k_1},\ldots,\left(X_{t_n}\right)^{k_n}\right\}=E\left\{\left(X_{t_1+h}\right)^{k_1},\ldots,\left(X_{t_n+h}\right)^{k_n}\right\},$$

avec $k_1,\cdots,k_n \geq 0$ et $k_1+\ldots+k_n \leq k$

On se limite souvent au processus d'ordre 1 et au processus d'ordre 2. Ainsi,

Définition 3.6. *Le processus* $\{X_t,\ t\in \mathrm{T}\}$ *est dit stationnaire d'ordre 2 ou faiblement stationnaire ou bien stationnaire au sens large, ou encore à covariance stationnaire si :*

a) $E(X_t)=m=\mathrm{Cte},$ $\forall t\in \mathrm{T}$

b) $\mathrm{Var}(X_t)=\sigma^2=\gamma_x(0),$ $\forall t\in \mathrm{T}$

c) $\mathrm{Cov}(X_t,X_{t+\theta})=\gamma_x(\theta)$ $\forall t\in \mathrm{T},\ \forall \theta\in \mathrm{T}$

où, $\gamma_x(\theta)$ est la fonction d'autocovariance du processus. En langage de probabilité, les conditions a) et b) de la définition 3.6 signifient que la moyenne et la variance du processus sont constantes, ce qui traduit respectivement la stabilité de son comportement au cours du temps et le fait qu'il possède la propriété d'homoscédasticité (variance scalaire

notée $\gamma(0)$). La condition c) de la définition 3.6 traduit quant à elle le fait que la covariance entre deux périodes t et $t+\theta$ est uniquement fonction de la différence des temps θ. Ainsi, les processus stationnaires d'ordre 2 sont des processus générateurs de chroniques sans tendance en moyenne et sans tendance en variance. Toutefois, cela ne signifie pas que de telles chroniques ont une représentation graphique stable. Signalons en passant que le processus $\{X_t, \quad t \in \mathrm{T}\}$ est *stationnaire au premier ordre* si sa moyenne est indépendante du temps, c'est-à-dire si seule la première condition a) est vraie dans la définition 3.6.

Dorénavant, l'utilisation de la notion de stationnarité s'identifiera aux définitions données précédemment. Ainsi, lorsqu'on dira qu'un processus est stationnaire, cela reviendra donc à admettre que les moments du premier et second ordre sont finis et, par ailleurs, qu'ils ne sont pas affectés lorsque l'origine du temps est déplacée. La vérification de l'hypothèse de stationnarité s'avère donc comme une nécessité évidente avant de faire toute inférence statistique sur une série de données chronologiques. Plus le comportement du processus générateur de la série s'écarte des hypothèses de base, en l'occurrence celle de la stationnarité, moins sa capacité de décrire le phénomène sera grande, et par conséquent, quelle que soit l'optimalité de la méthode d'estimation correspondante, les résultats finaux seront fortement déformés. L'hypothèse de stationnarité conditionne donc la validité des résultats que l'on peut tirer des différentes procédures d'estimation sur une chronique.

3.2.3 Stationnarité et ergodicité

Dans l'analyse fréquentielle d'une série chronologique, il est aussi important que le processus générateur des données soit *ergodique*. Pour bien comprendre cette notion d'ergodicité que l'on utilise souvent de façon naturelle, il convient de revenir un peu en arrière. En théorie, un phénomène ne dépendant pas du temps peut, d'après le postulat de base (appendice B), être assimilé à une expérience ; on peut donc obtenir plusieurs échantillons résultant de ces expériences. La théorie de l'analyse fréquentielle garantit la stabilité des estimations de ces moments. Dans le cas des phénomènes aléatoires dépendant du temps, on doit se contenter des données d'une seule expérience : la série chronologique. L'hypothèse d'ergodicité garantit que les moments empiriques - calculés sur une série

chronologique - seront de bonnes estimations des moments décrivant la population statistique du phénomène. L'hypothèse d'ergodicité établit le lien entre les deux types de moments.

Définition 3.7. *Un processus stationnaire (condition nécessaire) est ergodique si la connaissance des moments peut être obtenue à partir de l'observation d'une seule réalisation choisie de façon quelconque (c'est-à-dire sans répéter l'expérience).*

Comme pour l'hypothèse de stationnarité, on se contente de l'ergodicité au second ordre c'est-à-dire que seules sont assurées les convergences quadratiques des deux premiers moments empiriques vers leur espérance mathématique correspondante. La stationnarité et l'ergodicité permettent donc la mise en œuvre d'une démarche statistique. Elles assurent qu'il est possible d'estimer les caractéristiques du processus générateur sous-jacent à partir d'une réalisation finie.

La discussion du théorème d'ergodicité est au-dessus des prétentions de cette thèse. Aussi, on n'oublie pas, cependant, qu'un processus n'a pas besoin d'être strictement stationnaire pour satisfaire la propriété d'ergodicité. C'est pourquoi une série chronologique stationnaire sera vue comme issue d'un processus stationnaire et ergodique du second ordre. Bendat et Piersol (1966) ont montré que toutes données représentant un phénomène hydrologique stationnaire sont généralement ergodiques.

3.3 Modèles de décomposition et d'évolution d'une série chronologique

Dans cette section, on veut simplement discuter intuitivement des différents types de changements (*non-stationnarités*) qui peuvent survenir dans une série chronologique. On pense qu'une telle approche est indispensable pour mieux comprendre les mécanismes statistiques générateurs de tels changements. D'autre part, elle permet aussi de mieux comprendre les problèmes de base associés et les procédures de tests sous-jacents pour détecter ces changements.

Pour figurer l'évolution dans le temps de certains phénomènes aléatoires, les statisticiens s'accordent pour décomposer une série chronologique $\{x_t, \quad t \in \mathrm{T}\}$ en une somme de deux composantes aux propriétés différentes, soit :

$$x_t = TD_t + S_t \qquad t = 1,\ldots,n \tag{3.1}$$

où,

- TD_t désigne la partie déterministe de la série (en générale, une fonction non-stationnaire);
- S_t la partie stochastique (en général, de moyenne nulle par hypothèse).

En pratique, de multiples spécifications sont possibles pour la composante déterministe TD_t. Celle qui se retrouve la plus couramment utilisée fait de TD_t une fonction polynomiale du temps, linéaire ou non linéaire. Dans les séries chronologiques de données hydrométéorologiques la spécification la plus utilisée fait de TD_t un polynôme de premier degré en t :

$$TD_t = \alpha + \beta t \qquad t = 1,\ldots,n \tag{3.2}$$

où,

- $(\alpha,\beta) \in \mathbb{R}^2$

En associant les variations accidentelles, ε_t, au modèle précédent on peut alors représenter cette série chronologique sous la forme suivante :

$$x_t = \alpha + \beta t + \varepsilon_t \qquad t = 1,\ldots,n \tag{3.3}$$

Dans ce modèle, ε_t désigne les irrégularités ou mouvement résiduel. Il regroupe ce qui n'a pas été pris en compte par la composante TD_t. On associe habituellement ε_t à un

processus stationnaire de type bruit blanc (gaussien ou non), c'est-à-dire à une suite de variables aléatoires $\{\varepsilon_t,\ \ t \in \mathrm{T}\}$ de moyenne nulle ($E(\varepsilon_t)=0$), de variance constante ($Var(\varepsilon_t)=\sigma_\varepsilon^2$) et non corrélées ($Cov(\varepsilon_t\, \varepsilon_{t+\theta})=0$, pour $\theta \neq 0$). Ainsi définie, la partie TD_t représente la fonction déterministe de la tendance de la série chronologique $\{x_t,\ \ t \in \mathrm{T}\}$. Cette fonction est le facteur représentant l'évolution à long terme de la grandeur x_t et traduit l'aspect général de la série. TD_t incorpore donc la moyenne de cette série chronologique. Afin de garder un cadre simple à cette exposition, on postule que la partie stochastique suit un processus autorégressif et de moyenne mobile (*ARMA*) de la forme :

$$A(L)S_t = B(L)\varepsilon_t \qquad t = 1,\ldots,n \tag{3.4}$$

où,

- L est l'opérateur de retard tel que $L\,x_t = x_{t-1}$;
- $A(L)$ est un polynôme en L de degré p ;
- $B(L)$ est un polynôme en L de degré q ;
- ε_t est un processus stationnaire de type bruit blanc (gaussien ou non) :

$$E(\varepsilon_t) = m = \mathrm{Cte}, \qquad \forall t \in \mathrm{T}$$

$$\mathrm{Var}(\varepsilon_t) = \sigma^2, \qquad \forall t \in \mathrm{T}$$

$$\mathrm{Cov}(\varepsilon_t, \varepsilon_{t+\theta}) = \gamma_x(\theta) = 0, \qquad \forall t \in \mathrm{T},\ \forall \theta \in \mathrm{T}$$

Pour bien comprendre les implications de l'hypothèse de non-stationnarité, il est nécessaire de décomposer S_t, la partie stochastique, comme suit :

$$S_t = TS_t + C_t \qquad t = 1,\ldots,n \tag{3.5}$$

où,

- TS_t est un processus caractérisant la composante stochastique de la fonction de tendance associée à la série chronologique $\{x_t, \quad t \in \mathrm{T}\}$;
- C_t est la composante cyclique et on suppose pour le moment qu'il s'agit d'un processus stationnaire de moyenne zéro.

Le plus souvent, TS_t est une fonction du passé de la série chronologique x_t :

$$x_t = f(x_{t-1}, x_{t-2}, x_{t-3}, \ldots) \qquad t = 1, \ldots, n \tag{3.6}$$

Cette dernière formulation suppose que la valeur de la grandeur x_t à la date t dépend des valeurs obtenues par celle-ci aux dates antérieures. Comme la «mémoire» du phénomène ayant généré la série chronologique $\{x_t, \quad t \in \mathrm{T}\}$ se trouve dans le passé le plus récent de la série, il n'est pas utile de remonter trop loin dans le temps. On exprime donc ces constatations par le fait que la composante TS_t représente la partie stochastique de la fonction de tendance associée à cette série chronologique.

Désignons à présent T_t la fonction de tendance et définissons $T_t = TD_t + TS_t$. En d'autres termes, cette fonction est la somme des tendances à la fois déterministe et stochastique. On se trouve donc avec la modélisation suivante d'une série chronologique :

$$x_t = TD_t + TS_t + C_t + \varepsilon_t \qquad t = 1, \ldots, n \tag{3.7}$$

Cette modélisation n'est en général pas unique puisque des conditions plus générales pourraient être spécifiées. En effet, hormis l'autocorrélation (phénomène de persistance) que l'on peut trouver entre les données de la série, chacune des composantes de x_t peut contenir des éléments saisonniers. Ainsi, on se retrouve avec la décomposition suivante d'une série chronologique :

$$x_t = TD_t + TS_t + C_t + Sa_t + R_t + \varepsilon_t \qquad t = 1, \ldots, n \tag{3.8}$$

où,

- TD_t, représente la tendance déterministe ;
- TS_t, représente la tendance stochastique ;
- C_t, représente la variation cyclique ;
- Sa_t, représente la variation saisonnière ;
- R_t, tient compte du phénomène de persistance ;
- ε_t, tient compte des variations résiduelles purement aléatoires.

La relation fonctionnelle exprimée à l'équation 3.8 est dite *forme linéaire additive*. Celle-ci est choisie lorsque la série chronologique présente un comportement stable et régulier. Si par contre on observe des amplitudes variables avec la tendance, on préférera une *forme multiplicative* :

$$x_t = TD_t \times TS_t \times C_t \times Sa_t \times R_t \times \varepsilon_t \qquad t = 1,\ldots,n \tag{3.9}$$

En adoptant la formulation exprimée à l'équation 3.8, les éléments T_t, C, Sa peuvent être représentés schématiquement tel qu'illustré à la figure 3.1. Malgré l'importance des éléments C, Sa et R, on les laisse d'abord de côté et on poursuit cet exposé en supposant qu'ils sont absents. Ceci simplifie l'analyse sans perte en généralité par rapport aux points qu'on veut illustrer. En effet, les principes qui seront énoncés demeurent valides dans leur grande majorité lorsque ces éléments sont pris en compte.

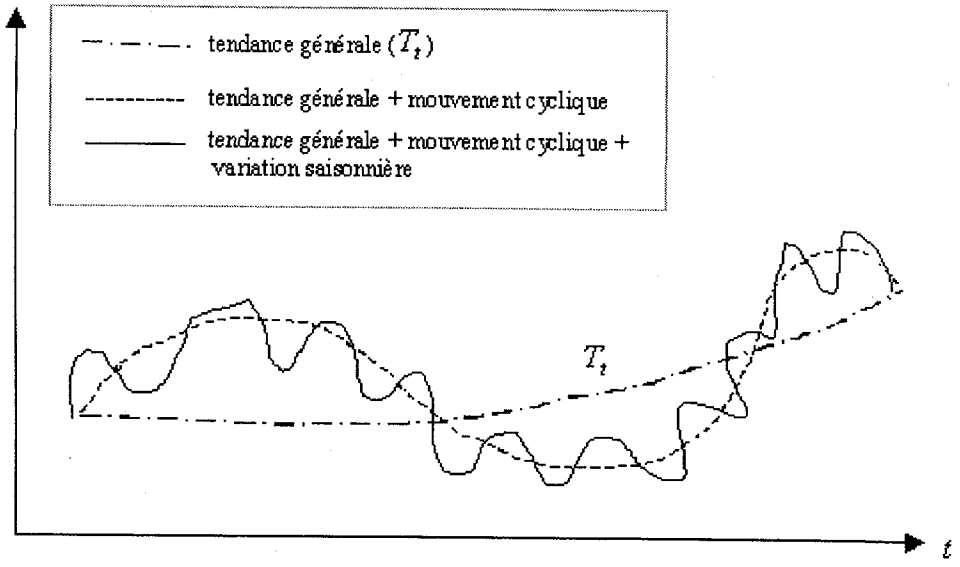


Figure 3.1 : Composition de la tendance générale (T_t) de la composante cyclique (C_t) et de la composante saisonnière (Sa_t)

3.4 Spécification des hypothèses de non-stationnarité

Pour rendre l'outil développé applicable à la problématique énoncée dans cette recherche, il est indispensable d'envisager une série d'hypothèses sur la structure d'une série chronologique, c'est-à-dire de distinguer différents niveaux de complexité associés à son modèle de génération. Il apparaît évident, à l'instar de la section précédente, que les structures TD_t et TS_t jouent un rôle important dans le traitement statistique d'une chronique. Choisir entre l'une ou l'autre des structures permet de mettre en évidence le caractère stationnaire ou non d'une série chronologique par la détermination d'une tendance déterministe ou stochastique ou encore par les deux. C'est dans cette optique que le problème de détection de tendances dans une série chronologique peut faire partie du problème de détection de non-stationnarités.

En reprenant les enseignements de la section précédente, on considère qu'une série chronologique, $\{x_t, \quad t \in \mathrm{T}\}$, est générée suivant le modèle ci-après :

$$x_t = TD_t + TS_t + \varepsilon_t \qquad t = 1,\ldots,n \tag{3.10}$$

où,

- TD_t, représente la tendance déterministe ;
- TS_t, représente la tendance stochastique ;
- ε_t, tient compte des variations résiduelles purement aléatoires,

l'étude de la stationnarité d'une série chronologique sera donc grandement simplifiée par la formulation de trois hypothèses de non-stationnarité :

- ***Hypothèse 1*** : la non-stationnarité est due à la présence dans la série d'une tendance déterministe, $T_t = TD_t$
- ***Hypothèse 2*** : la non-stationnarité est due à la présence dans la série d'une tendance stochastique, $T_t = TS_t$
- ***Hypothèse 3*** : *la non-stationnarité est due à la présence dans la série d'une tendance déterministe et d'une tendance stochastique,* $T_t = TD_t + TS_t$

En considérant ces trois hypothèses, on est alors conduit à définir de façon raisonnable les contours de cette étude. Pour l'instant, on s'attarde tout d'abord à caractériser la nature non-stationnaire de chacune de ces hypothèses.

3.4.1 Hypothèse 1 : existence d'une tendance déterministe dans la série chronologique

On peut identifier deux types de tendances déterministes dans les données d'une série chronologique : la tendance en saut (ou rupture) et la tendance linéaire ou monotone. D'une façon générale, la tendance déterministe aura la forme :

$$x_t = \alpha + \beta t + \varepsilon_t \qquad t = 1,\ldots,n \tag{3.11}$$

où,

- $(\alpha,\beta) \in \mathbb{R}^2$;
- $t = \begin{cases} 0 & si \quad t \le t_0 \\ 1 & si \quad t \ge t_0 \end{cases}$, pour la tendance en saut ;
- t_0 est la date du saut éventuel et β l'amplitude ;
- $t = 1,2,\ldots,n$ pour la tendance linéaire ;
- $\varepsilon_t \overset{iid}{\sim} N(0,\sigma^2)$

L'hypothèse de non-stationnarité à vérifier est alors :

$$\text{H}: \quad \beta \neq 0 \tag{3.12}$$

D'autre part, les caractéristiques du modèle exprimé à l'équation 3.11 sont les suivantes :

$$\begin{cases} E(x_t) = \alpha + \beta t + E(\varepsilon_t) = \alpha + \beta t \\ Var(x_t) = E(x_t^2) - E(x_t)^2 = E(\varepsilon_t^2) - E(\varepsilon_t)^2 = \sigma^2 \\ Cov(x_t, x_s) = E(x_t - E(x_t))E(x_s - E(x_s)) = E(\varepsilon_t \varepsilon_s) = 0 \quad \text{pour } s \neq t \end{cases} \tag{3.13}$$

Il en résulte que le processus générateur de la série chronologique $\{x_t, \quad t \in \text{T}\}$ est non-stationnaire car $E(x_t)$ dépend du temps. De plus la variance est constante dans le temps et ses covariances sont nulles. Ainsi, vérifier qu'il existe une tendance déterministe dans une série de données, revient à vérifier l'existence d'une *non-stationnarité d'ordre un.*

3.4.2 Hypothèse 2 : existence d'une tendance stochastique dans la série chronologique

D'après les équations 3.4 et 3.5 on a spécifié que $S_t = TS_t + C_t$. De plus, on a imposé la structure $ARMA(p,q)$ pour S_t. Ainsi, l'hypothèse quant à la nullité de TS_t (ou encore hypothèse de stationnarité) peut se traduire par une hypothèse sur les coefficients du

polynôme autorégressif $A(L)$. Dans le cas où $TS_t = 0$ pour tout t, S_t est égal à C_t et est donc stationnaire. Cela signifie que les racines du polynôme $A(x) = 0$ sont toutes hors du cercle unitaire. Par contre, si ce même polynôme contient une racine sur le cercle unitaire alors S_t est non-stationnaire et TS_t est non nulle. Dans ce qui suit, et, sans perte en généralité, on considère le cas où S_t est un processus autorégressif d'ordre 1 ($AR(1)$) sans terme de moyenne mobile. On a alors : $(1-\rho L)S_t = \varepsilon_t$, où ε_t est un bruit blanc. L'hypothèse quant à la nullité de TS_t peut maintenant se traduire par une hypothèse sur les coefficients du polynôme autorégressif $A(L) = 1-\rho L$. Dans ce cas simplifié, une tendance stochastique est présente dans la série chronologique $\{x_t, \quad t \in \mathrm{T}\}$ si $\rho = 1$ (présence d'une racine unitaire), tandis que le processus est stationnaire (tendance stochastique absente) si $|\rho| < 1$. Bien que ce modèle simplifié ne considère pas le cas où une série chronologique peut contenir plus d'une racine unitaire, il permet une discussion plus aisée des questions se rapportant à l'hypothèse de l'existence d'une tendance stochastique. Ce faisant, la série chronologique $\{x_t, \quad t \in \mathrm{T}\}$ s'écrit alors sous la forme :

$$x_t = \beta_0 + \rho\, x_{t-1} + \varepsilon_t \qquad t = 1,\ldots,n \tag{3.14}$$

où,

- $\beta \in \mathbb{R}$
- $-1 \le \rho \le 1$
- $\varepsilon_t \overset{iid}{\sim} N(0,\sigma^2)$

L'hypothèse de non-stationnarité à vérifier est alors :

$$\mathrm{H}: \quad \rho = 1 \tag{3.15}$$

D'après le modèle exprimé à l'équation 3.14, on montre par récurrence que

$$x_t = \rho^{\tau} x_{t-1} + \beta_0 \sum_{j=0}^{\tau-1} \rho^j + \sum_{j=0}^{\tau-1} \rho^j \varepsilon_{t-j} \qquad t = 1,\ldots,n \tag{3.16}$$

Ainsi, si $|\rho|=1$ et $\tau=t \quad \Rightarrow \quad x_t = x_0 + t\beta_0 + \sum_{j=1}^{t}\varepsilon_t \quad$ avec $\quad \varepsilon_t \overset{iid}{\sim} .N(0,\sigma_\varepsilon^2)$, on obtient,

$$\begin{cases} E(x_t) = E(x_0 + t\beta_0) + \sum_{j=1}^{t} E(\varepsilon_j) \\ Var(x_t) = E(x_t^2) - \{E(x_t)\}^2 = t\sigma_\varepsilon^2 \\ Cov(x_t, x_s) = E\{x_t - E(x_t)\}E\{x_s - E(x_s)\} = E\left[\sum_{j=1}^{t}\varepsilon_j \sum_{j=1}^{s}\varepsilon_j\right] \\ \qquad\qquad = \min(t,s).\sigma_\varepsilon^2 \quad \text{pour} \quad s \neq t \end{cases} \tag{3.17}$$

À partir de l'équation 3.17, l'espérance (respectivement, la variance) de la série chronologique représentée à l'équation 3.14 suit une fonction de tendance linéaire. Autrement dit, elles sont dépendantes du temps. En conséquence, vérifier qu'il existe une tendance stochastique dans une série de données revient à vérifier l'existence d'une *non-stationnarité d'ordre deux.*

3.4.3 Hypothèse 3 : existence simultanée d'une tendance déterministe et d'une tendance stochastique

Dans la situation présente, la série chronologique $\{x_t, \quad t \in \mathrm{T}\}$ aura la forme suivante :

$$x_t = \alpha + \beta t + \rho x_{t-1} + \varepsilon_t \qquad t = 1,\ldots,n \tag{3.18}$$

Cette formulation permet une caractérisation simple de la question de la racine unitaire. En effet, faire l'hypothèse que la série $\{x_t, \quad t \in \mathrm{T}\}$ est caractérisée par la présence d'une racine unitaire revient à faire l'hypothèse que la composante stochastique non-stationnaire TS_t est non nulle $(\rho = 1)$. Si $TS_t = 0$ pour toute valeur de t, x_t est alors à tendance stationnaire dans le sens où x_t est caractérisée par des fluctuations transitoires C_t autour d'une fonction de tendance déterministe TD_t. Ainsi, en considérant les résultats obtenus aux équations 3.13 et 3.17 il appert que, vérifier qu'il existe simultanément une tendance déterministe et une tendance stochastique dans une série chronologique revient à vérifier l'existence d'une non-stationnarité d'ordre deux.

Les trois hypothèses présentées précédemment couvrent l'ensemble des simplifications que l'on peut envisager pour le modèle statistique de génération d'une série chronologique. On reviendra en détail sur chacune d'elles lorsqu'on aura plus d'éléments pour juger de leur opportunité. Le tableau 3.1 présente un résumé des trois hypothèses de non-stationnarité retenues.

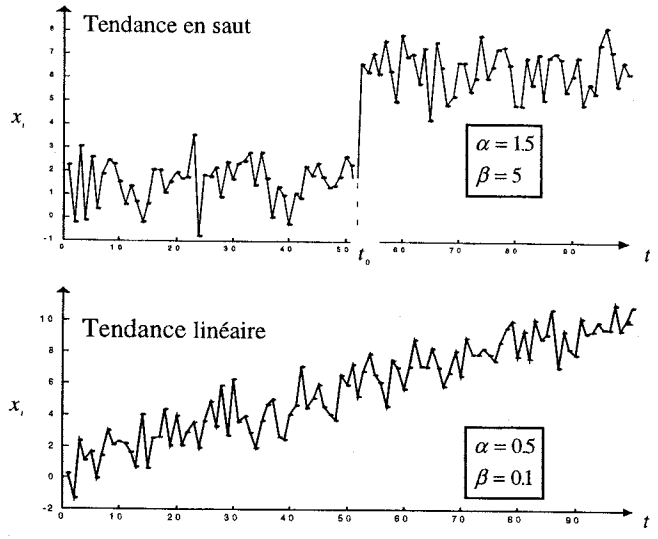
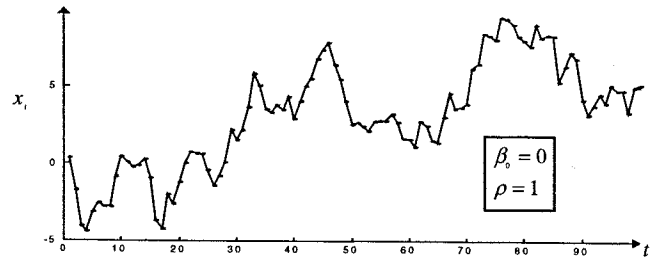
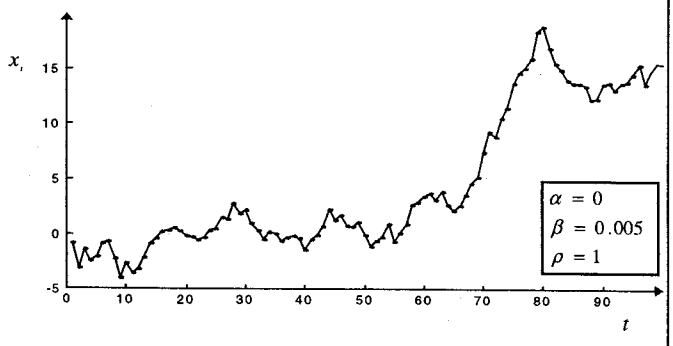
Type de tendance	**Figure illustrative**	**Hypothèse**
Déterministe : $T_t = TD_t$ $x_t = \alpha + \beta t + \varepsilon_t$		$H : \beta \neq 0$ (la non-stationnarité est de nature déterministe)
Stochastique : $T_t = TS_t$ $x_t = \beta_0 + \rho x_{t-1} + \varepsilon_t$		$H : \rho = 1$ (La non-stationnarité est de nature purement stochastique)
Déterministe et stochastique : $T_t = TD_t + TS_t$ $x_t = \alpha + \beta t + \rho x_{t-1} + \varepsilon_t$		$H : \rho = 1$ (La non-stationnarité est de nature déterministe et stochastique)

Tableau 3.1 : Illustration des différents types d'hypothèses de non-stationnarité

3.5 Existence de lacunes possibles au sein des séries

Certaines caractéristiques importantes des séries chronologiques de données hydrométéorologiques peuvent compliquer l'utilisation des différents types d'hypothèses de non-stationnarité présentées précédemment. Il s'agit de la persistance (ou encore de la présence d'une autocorrélation entre les données) et de la saisonnalité.

3.5.1 Persistance

La dépendance linéaire existant entre les données d'une série chronologique pourrait causer certains types de structures dans les données. Par exemple, dans certaines parties de la série chronologique, les petites valeurs (respectivement les grandes valeurs) seront fréquemment suivies par de petites valeurs (respectivement de grandes valeurs). Les hydrométéorologues réfèrent cette caractéristique comme étant la persistance et d'un point de vue statistique, cela signifie que les données sont probablement autocorrélées. Cette autocorrélation représente une dépendance à court terme des observations. L'existence de la persistance dans une série amène la violation d'au moins une hypothèse de base sur les modèles de représentation d'une série chronologique : soit l'indépendance des termes d'erreurs (ε_t) ou l'indépendance des observations x_t. D'autre part, comme la persistance traduit la nature des liens qui unissent la valeur présente aux valeurs passées de la série de données chronologiques, elle est la mémoire du processus hydrologique et est par le fait même très intimement liée à la loi qui le régit. Ainsi, sa présence dans la série peut conduire à des rejets fallacieux de l'hypothèse de stationnarité, c'est-à-dire qu'elle peut indiquer l'existence d'une tendance qui en réalité n'existe pas.

Dans les modèles de représentation des différents types de tendance retenues, pour extraire la dépendance à court terme, dans la série des erreurs ε_t induite par la présence de la persistance dans les données, on peut introduire, du côté des variables explicatives, les variables avec retard $x_{t-1}, x_{t-2}, \ldots, x_{t-p}$. Ces nouvelles variables explicatives permettent d'introduire, à l'intérieur du modèle, des corrélations entre les observations successives et,

donc de les exclure de la série des résidus. Les modèles de génération de données et les différentes hypothèses de non-stationnarité associées sont données au tableau 3.2.

Modèle de génération de données	Hypothèse
Tendance déterministe : $T_t = TD_t$ $x_t = \alpha + \beta t + \sum_{i=1}^{m} \varphi_i x_{t-i} + \varepsilon_t$	$H_0: \beta \neq 0$ (la non-stationnarité est de nature déterministe)
Tendance stochastique : $T_t = TS_t$ $x_t = \beta_0 + \rho x_{t-1} + \sum_{i=1}^{m} \varphi_i x_{t-i} + \varepsilon_t$	$H_0: \rho = 1$ (La non-stationnarité est de nature purement stochastique)
Tendances déterministe et stochastique : $T_t = TD_t + TS_t$ $x_t = \alpha + \beta t + \rho x_{t-1} + \sum_{i=1}^{m} \varphi_i x_{t-i} + \varepsilon_t$	$H_0: \rho = 1$ (La non-stationnarité est de nature déterministe et stochastique)

Tableau 3.2 : Illustration des différents types d'hypothèses de non-stationnarité en présence du phénomène de persistance

3.5.2 Saisonnalité

Habituellement la saisonnalité est connue d'avance. Pour les données issues des variables hydrométéorologiques, la saisonnalité est bien sûre causée par la rotation annuelle de la Terre autour du Soleil et d'ordinaire, seules les données annuelles ne sont pas saisonnières. Par exemple, dans une série de débits en rivière, on constatera que les observations prises en hiver s'éloignent plus de celles prises au printemps à l'été et à l'automne.

La présence de saisonnalités amène des changements de niveaux fréquents et de courte durée ; ces fluctuations à court terme doivent être prises en compte pour permettre une étude valable de l'évolution à long terme (tendance) d'une série chronologique. Les saisonnalités peuvent donc, de façon générale, masquer la présence de tendance à long terme ; elles peuvent également introduire des tendances fictives conduisant ainsi à un rejet erroné de l'hypothèse de non-stationnarité retenue jusqu'ici. Pour extraire la dépendance à

court terme induite par la saisonnalité, on peut utiliser des variables indicatrices dans le modèle de représentation des différents types de tendances retenus. Ainsi, en prenant l'exemple de quatre saisons : hiver, printemps, été et automne, les trois variables dichotomiques suivantes permettent d'extraire les composantes saisonnières dans les modèles retenus :

$$x_{\text{été}} = \begin{cases} 1 & \text{si l'observation est mesurée a l'été} \\ 0 & \text{sinon} \end{cases}$$

$$x_{aut} = \begin{cases} 1 & \text{si l'observation est mesurée a l'automne} \\ 0 & \text{sinon} \end{cases}$$

$$x_{pri} = \begin{cases} 1 & \text{si l'observation est mesurée au printemps} \\ 0 & \text{sinon} \end{cases}$$

Les modèles de génération de données et les différents types d'hypothèses de non-stationnarité sont présentés dans le tableau 3.3.

Modèle de génération de données	**Hypothèse**
Tendance déterministe : $T_t = TD_t$ $x_t = \alpha + \beta t + \delta_1 x_{ete} + \delta_2 x_{aut} + \delta_3 x_{pri} + \varepsilon_t$	$H_0: \beta \neq 0$ (la non-stationnarité est de nature déterministe)
Tendance stochastique : $T_t = TS_t$ $x_t = \beta_0 + \rho x_{t-1} + \delta_1 x_{ete} + \delta_2 x_{aut} + \delta_3 x_{pri} + \varepsilon_t$	$H_0: \rho = 1$ (La non-stationnarité est de nature purement stochastique)
Tendances déterministe et stochastique : $T_t = TD_t + TS_t$ $x_t = \alpha + \beta t + \rho x_{t-1} + \delta_1 x_{ete} + \delta_2 x_{aut} + \delta_3 x_{pri} + \varepsilon_t$	$H_0: \rho = 1$ (La non-stationnarité est de nature déterministe et stochastique)

Tableau 3.3 : Illustration des différents types d'hypothèses de non-stationnarité en présence de la saisonnalité

3.5.3 Persistance et saisonnalité

Si la fréquence d'échantillonnage d'une variable hydrologique est assez élevée, la série chronologique obtenue présentera à la fois les phénomènes de persistance et de saisonnalité. Dans une série de débits journaliers, en plus de la périodicité, on remarquera aussi une persistance. Les problèmes induits par la présence d'une persistance ou d'une saisonnalité ont maintenant un effet combiné dans ce cas-ci. Les différents types d'hypothèses de non-stationnarité associées sont illustrés au tableau 3.4.

Modèle de génération de données	Hypothèse
Tendance déterministe : $T_t = TD_t$ $x_t = \alpha + \beta t + \sum_{i=1}^{m} \varphi_i x_{t-i} + \delta_1 x_{ete} + \delta_2 x_{aut} + \delta_3 x_{pri} + \varepsilon_t$	$H_0: \beta \neq 0$ (la non-stationnarité est de nature déterministe)
Tendance stochastique : $T_t = TS_t$ $x_t = \beta_0 + \gamma x_{t-1} + \sum_{i=1}^{m} \varphi_i x_{t-i} + \delta_1 x_{ete} + \delta_2 x_{aut} + \delta_3 x_{pri} + \varepsilon_t$	$H_0: \rho = 1$ (La non-stationnarité est de nature purement stochastique)
Tendances déterministe et stochastique : $T_t = TD_t + TS_t$ $x_t = \alpha + \beta t + \rho x_{t-1} + \sum_{i=1}^{m} \varphi_i x_{t-i} + \delta_1 x_{ete} + \delta_2 x_{aut} + \delta_3 x_{pri} + \varepsilon_t$	$H_0: \gamma = 1$ (La non-stationnarité est de nature déterministe et stochastique)

Tableau 3.4 : Illustration des différents types d'hypothèses de non-stationnarité en présence du phénomène de persistance et de saisonnalité

3.6 Conclusion

Dans ce chapitre, on a défini de façon plus rigoureuse la notion de stationnarité. Un processus est stationnaire au second ordre si l'ensemble de ses moments d'ordre un et d'ordre deux sont indépendants du temps. Par opposition, un processus non-stationnaire est un processus qui ne satisfait pas l'une ou l'autre de ces deux conditions. Ainsi, l'origine de

la non-stationnarité peut provenir d'une dépendance du moment d'ordre un (l'espérance) par rapport au temps et/ou d'une dépendance de la variance ou des autocovariances par rapport au temps.

Une des causes évidentes de non-stationnarités d'une série chronologique est la présence d'une tendance déterministe ou stochastique ou encore une combinaison des deux. Il existe donc deux sortes de non-stationnarité : la non-stationnarité déterministe et la non-stationnarité stochastique. Le fait que la stationnarité puisse être de type déterministe ou stochastique nous a permis de définir trois hypothèses principales de non-stationnarité :

Hypothèse 1 : la non-stationnarité est due à la présence dans la série d'une tendance déterministe ;

Hypothèse 2 : la non-stationnarité est due à la présence dans la série d'une tendance stochastique ;

Hypothèse 3 : la non-stationnarité est due à la présence dans la série d'une tendance déterministe et d'une tendance stochastique.

Ces différentes hypothèses selon l'origine de la non-stationnarité sont essentielles pour développer des tests statistiques de stationnarité appropriés.

4 STATIONNARITÉ DES SÉRIES DE DONNÉES HYDROMÉTÉOROLOGIQUES

L'objectif de l'analyse hydrologique est de développer des représentations mathématiques de données qu'on appelle des modèles ou des hypothèses. Dans ce chapitre, on analyse la question de la non-stationnarité des processus hydrométéorologiques à la lumière des définitions données au chapitre précédent. On revoit d'abord quelques principes fondamentaux des processus stochastiques représentatifs de variables hydrométéorologiques ainsi que leurs réalisations. Puis, on souligne l'importance de se préoccuper de l'hypothèse de non-stationnarité en hydrologie. Finalement, on propose quelques pistes pouvant expliquer la non-stationnarité des processus hydrologiques.

4.1 Processus représentatifs d'une variable hydrométéorologique et leurs réalisations

4.1.1 Processus représentatifs de variables hydrométéorologiques

Pour les phénomènes hydrométéorologiques il est pratiquement impossible de prédire de façon certaine ce qui se produira dans le futur. Par exemple, les météorologues ne déclarent jamais qu'il y aura exactement 5 mm de pluie demain. Cependant, lorsqu'un événement comme «la chute de pluie de demain» s'est produit, c'est alors que cette valeur de la série temporelle de la précipitation est connue exactement. Toutefois, il continuera de pleuvoir dans le futur et la séquence de tous les enregistrements de la précipitation historique est tout simplement une réalisation de ce qui s'est produit. La précipitation est un exemple d'un phénomène aléatoire hydrométéorologique qui évolue dans le temps d'après les lois de probabilité.

De manière générale, les phénomènes hydrométéorologiques sont des phénomènes pour lesquels le déroulement ne peut être prédit avec certitude. Tout phénomène hydrométéorologique peut donc être décrit comme un phénomène qui change avec le temps conformément à la loi de probabilité aussi bien qu'avec la réalisation séquentielle entre ses occurrences (Salas et al., 1980). Par exemple, l'occurrence des débits d'un cours d'eau ne peut être considérée comme étant purement déterministe. Elle dépend, entre autres, des conditions météorologiques ou des facteurs humains qui ne peuvent être prévus de manière certaine. Cette incertitude est la caractéristique essentielle du phénomène aléatoire et l'oppose aux phénomènes déterministes. C'est pourquoi, on peut assimiler un phénomène hydrométéorologique à un processus de variable hydrométéorologique (Cavadias, 1966) ou encore processus stochastique $\{X_t, \quad t \in \mathrm{T}\}$. En principe, la forme des fonctions de distributions jointes (lois de probabilités jointes) des processus de variables hydrométéorologiques est rarement connue. Ces distributions, telles qu'exprimées au chapitre 3, dépendent du temps. Du point de vue pratique, la détermination des deux premiers moments de telles distributions suffit pour décrire le comportement des processus. En effet, nombre de propriétés importantes des processus de variables hydrométéorologiques peuvent être obtenues à partir de ces moments.

4.1.2 Réalisations de processus représentatifs de variables hydrométéorologiques

L'étude des processus hydrologiques à l'échelle locale et régionale passe par la mesure et l'analyse de différentes variables hydrométéorologiques (précipitations, débits, etc.). Cela nécessite des ensembles de réalisations possibles des populations définies respectivement par ces processus. Bien que mesurée par le biais d'un instrument de mesure, l'information disponible sur de tels processus est, dans la plupart des cas, systématiquement collectée et publiée sous forme d'enregistrements de données au cours du temps pendant des périodes plus ou moins longues pour constituer ce que l'on appelle une série chronologique (ou série temporelle ou encore chronique). L'échantillon de données ainsi constitué décrit une réalisation du processus. On le note $\{x_t, \quad t \in \mathrm{T}\}$. Des exemples simples sont fournis par

les températures enregistrées à chaque heure à une certaine station météorologique, une suite d'observations chronologiques du niveau d'eau d'un fleuve, des hauteurs de pluies recueillies à l'observatoire d'une station de mesure, les mesures de débits journaliers d'un cours d'eau, etc. En principe, l'acquisition des séries temporelles de variables hydrométéorologiques permet de compléter les séries de données provenant des services hydrologiques et météorologiques nationaux. De telles données sont stockées sur des systèmes de gestion de base de données. Elles forment ainsi un ensemble important de chiffres et de graphiques que l'on dépouille et classe.

4.2 Réalisations des processus hydrométéorologiques et leur usage en hydrologie

Pour l'ingénieur hydrologue, de nombreux problèmes économico hydrologiques ne font intervenir que certains aspects d'une réalisation d'un processus hydrométéorologique, notamment : les maximums annuels, les moyennes mensuelles, etc. Il est alors de pratique courante, d'extraire d'une série temporelle d'une variable hydrométéorologique les séquences de maxima ou de moyennes pour constituer de nouvelles séries temporelles. De telles séries de données calculées ou reconstituées constituent l'essentiel de l'information disponible pour un ensemble de stations qui permettent la détermination de la variabilité dans le temps et dans l'espace des caractéristiques des régimes hydrologiques d'un bassin hydrographique. Il va sans dire que cette information est nécessaire pour la compréhension des processus intervenant dans le cycle de l'eau ainsi que l'étude de leurs variations spatiales et temporelles. Ces séries de données constituent ainsi un préalable à toute analyse hydrologique que ce soit dans le but de réaliser des objectifs attendus très opérationnels liés à la gestion, à l'aménagement du territoire et des ressources en eau ou pour étudier l'impact d'éventuelles modifications climatiques sur l'hydrologie des cours d'eau (influence sur débits de pointe, débits d'étiage, etc.).

La connaissance des caractéristiques statistiques des séries temporelles de variables hydrométéorologiques est incontournable pour l'ingénieur hydrologue. Par exemple, en hydrologie, les observations effectuées sur une variable hydrologique montrent que cette grandeur est variable et ne peut être prévue avec exactitude. On peut donc associer une loi de probabilité régissant le comportement de cette variable. Le problème revient alors à juger si la réalisation de la variable hydrologique, c'est-à-dire la série temporelle, est compatible avec la loi de probabilité retenue. Si cette loi est complètement spécifiée, le choix des décisions que l'on doit prendre et dont les conséquences dépendent de la variable hydrométéorologique en question, pourra être fait en fonction de critères qui tiendront compte de cette loi de probabilité. De telles lois de probabilités connaissent diverses utilisations en hydrologie. Par exemple, pour la prédétermination des débits maxima de crues, on peut ajuster à la série temporelle des crues maxima annuelles une loi de probabilité théorique qui la représentera aussi fidèlement que possible. Par exemple, les hydrologues considèrent habituellement que la loi de probabilité de Gumbel ou Frechet permet le meilleur ajustement aux observations de débits maxima de crues. Si l'on considère que X est la variable aléatoire représentant la crue maximum annuelle et que $\mathrm{F}(X;\theta)=\mathrm{P}(X \le x)$ est sa fonction de probabilité, un usage important d'une telle modélisation est la détermination du quantile x_p tel que : $\mathrm{P}(X > x_p)=1-\mathrm{F}(x_p;\theta)=p$. Ainsi, une fois estimé la crue de probabilité p, les hydrologues l'utilisent alors pour introduire la notion de durée de retour ($1/p$ étant la durée de retour du quantile x_p). Par exemple, pour $p=1/10$, on définit la crue décennale, c'est-à-dire la valeur du débit dépassée en moyenne une fois tous les dix ans.

Les caractéristiques statistiques des séries temporelles de variables hydrologiques sont également utilisées dans des modèles de planification des ressources en eau. En effet, certaines caractéristiques statistiques issues de ces séries de données (moyenne, variance, autocorrelation, etc.) sont utilisées comme intrants dans de tels modèles. Ces derniers fournissent des informations nécessaires au dimensionnement d'ouvrages hydrauliques, de protection contre les crues ou pour la gestion hydraulique du bassin versant étudié. On

distingue principalement deux types d'utilisation des modèles de planification de ressources en eau :

- *La modélisation comme outil de prévision* : les prévisions météorologiques (les valeurs moyennes de la précipitation et de la température pour des dates particulières) peuvent être utilisées comme entrée d'un modèle pluie-débit afin de simuler le débit. Par exemple, une alternative d'un tel modèle est de générer une gamme de scénarios d'apports en considérant chaque année de météo observée comme données d'entrée du modèle pluie-débit. C'est le cas par exemple de la méthodologie appelée ESP ("*Extended Streamflow Prediction*"). Elle a été appliquée entre autre par Lettenmaier (1984) et Ouarda et al. (1998).
- *La modélisation comme outil de simulation* : la simulation consiste à utiliser un modèle en vue de générer un ensemble de scénarios synthétiques dont les propriétés statistiques reproduisent en moyenne celles de la série de données historiques. Par exemple le modèle MINERVE, qui est un modèle d'optimisation développé à l'IREQ, permet de déterminer l'aménagement optimal des installations hydroélectriques (taille et localisation des réservoirs et barrages, nombre et puissance des unités, etc.) sur une rivière donnée (Duquette, 1996). Dans ce type de modèle l'aménagement optimal est déterminé en utilisant les apports historiques. Pour tenir compte de la variabilité des apports, on peut introduire plusieurs scénarios d'apports (par exemple des apports mensuels, bimensuels ou trimestriels) qui sont générés à l'aide d'un modèle stochastique. Cette approche a été appliquée par Rasmussen et al (1998).

Une bonne connaissance des caractéristiques statistiques des réalisations de processus hydrologiques (séries temporelles hydrologiques) est donc nécessaire au bon fonctionnement des modèles de planification des ressources en eau. En effet, de cette connaissance dépend, bien évidemment, la fiabilité des estimations statistiques utilisées pour rendre ces modèles opérationnel.

4.3 Préalable à l'emploi des méthodes statistiques sur les réalisations de variables hydrométéorologiques

La plupart des études et travaux envisagés en hydrologie, à l'aide des méthodes statistiques, sont fondés sur des ensembles de données, résultats de mesures directes sur des processus représentatifs de variables hydrométéorologiques. Ces mesures ne sont en général pas ponctuelles et limitées dans le temps mais au contraire continues ou discontinues dans le temps et réparties dans l'espace au sein d'un réseau. Celui-ci est constitué principalement d'un ensemble de stations dont le but est la détermination de la variabilité dans le temps et dans l'espace des caractéristiques des régimes hydrologiques d'un bassin hydrographique. L'ensemble des données à une station constitue donc une information considérable qu'il est souhaitable de résumer à l'aide de caractéristiques statistiques bien choisies. On applique donc ainsi l'analyse fréquentielle, une méthode statistique, pour étudier les événements passés, caractéristiques d'un phénomène hydrologique donné, afin d'en définir des lois de probabilités d'apparition future. Par ailleurs, on estime aussi de la sorte les caractéristiques de tendance centrale (moyenne, médiane, etc.), de dispersion (variance, écart-type, etc.) et de manière plus générale les quantiles. Ces dernières sont des intrants de modèles de planification de ressources en eau.

Du point de vue pratique les réalisations des processus représentatifs de variables hydrométéorologiques ne sont pas des données indépendantes et identiquement distribuées au sens statistique du terme. Elles sont liées (le temps joue un rôle particulier) et présentent parfois des erreurs systématiques qui les rendent hétérogènes (absence de données, données défectueuses, données correspondant à deux ou plusieurs séries homogènes non défectueuses et groupées sous le nom d'une même station). Il est donc indispensable, avant de soumettre de telles séries de données à une analyse statistique, de se préoccuper de leur qualité et de leur représentativité. Il s'agit pour l'essentiel de procéder à leur validation avant leur mise en disposition à des fins opérationnelles.

L'hypothèse fondamentale de validation est qu'une série temporelle en tant que représentation d'un phénomène hydrologique est aléatoire et simple. Cette hypothèse assure que toutes les observations constituant la série temporelle sont issues de la même population (même phénomène) et qu'elles sont indépendantes entre elles (hypothèse iid donnée à l'appendice B). La non vérification (ou validation) du caractère aléatoire et simple d'une série d'observations peut avoir plusieurs causes, parfois simultanément. Ces causes se regroupent généralement en deux catégories :

1. autocorrélation : pour le caractère non aléatoire de la série ;
2. non-stationnarité : pour le caractère non simple de la série.

Le caractère aléatoire et simple d'une série d'observations est donc un préalable à l'emploi des méthodes conventionnelles de statistique sur les séries temporelles de variables hydrométéorologiques.

L'étude des processus hydrométéorologiques à l'échelle locale et régionale doit donc impérativement prendre en compte deux caractéristiques essentielles des variables hydrométéorologiques (précipitations, débits, variables météorologiques, etc.) : la non-stationnarité et une structure interne de dépendance. La prise en compte explicite de ces caractéristiques est capitale pour garantir la fiabilité des modèles statistiques utilisés en hydrologie pour la planification des ressources en eau. Seule la caractéristique non-stationnaire est au centre de cette thèse. Il est donc clair dans notre cas que l'estimation des intrants est conditionnée par l'hypothèse que les réalisations de variables hydrométéorologiques sont identiquement distribuées (stationnaire). Si cette hypothèse est relaxée, les codes existant pour la conception des digues, des barrages etc. doivent impérativement être révisés. L'étude de la stationnarité est pour ainsi dire, une tâche incontournable et récurrente des hydrologues (Bernier, 1977). Elle garantit entre autres que les estimations des intrants obtenus à partir des séries temporelles de variables hydrométéorologiques sont stables et fiables.

4.4 Hypothèses de stationnarité en hydrométéorologie

4.4.1 Stationnarité des séries temporelles hydrométéorologiques

L'emploi de méthodes d'inférence statistique sur n réalisations d'une variable hydrométéorologique X conduit à interpréter les observations portant sur X comme la réalisation d'un échantillon théorique $\{X_1,\ldots,X_n\}$. De fait, chaque variable aléatoire X_t a :

- une moyenne $E(X_t)=\mu_t$
- une variance $Var(X_t)=\sigma_t^2$

qui sont, a priori, des fonctions du temps t. De plus, pour tout couple de variables $(X_t, X_{t+\theta})$ on peut aussi définir une autocovariance avec un retard de θ périodes :

- $Cov(X_t, X_{t+\theta})=\gamma_t(t, t+\theta)=\gamma_t(\theta)$

qui dépend, a priori, des dates t et $t+\theta$ (et pas seulement du décalage temporel θ). Le processus $\{X_1,\ldots,X_n\}$ est stationnaire au sens strict si toutes ses caractéristiques statistiques (fonction de répartition et moments d'ordre quelconque) sont indépendantes du temps t. Il est stationnaire au sens faible si cette condition d'invariance par rapport à l'origine des temps n'est imposée qu'aux moments d'ordre 1 et 2 (stationnarité d'ordre 2), c'est-à-dire, d'après la section 3.2.2,

- $\mu_t=\mu \quad \forall t$
- $\sigma_t^2=\sigma^2 \quad \forall t$
- $\gamma_t(t, t+\theta)=\gamma(\theta) \quad \forall t$

En conséquence, une série temporelle de réalisations d'une variable hydrométéorologique X, à un pas de temps donné, est dite stationnaire si ses réalisations sont issues d'un même processus stochastique dont les caractéristiques statistiques (moyenne, variance, covariance) restent constants au cours du temps.

4.4.2 Causes de la non-stationnarité des processus hydrométéorologiques

La question de la stationnarité d'un processus représentatif de variables hydrométéorologiques n'est pas facile à résoudre car la stationnarité d'un tel processus peut être violée de plusieurs façons. Pour une compréhension correcte de cette question il importe de préciser le sens attribué à la non-stationnarité des processus représentatifs de variables hydrométéorologiques, surtout pour les processus représentatifs de variables hydrologiques.

Le régime hydrologique d'un cours d'eau résume l'ensemble de ses caractéristiques hydrologiques, définies statistiquement, et est habituellement décrit par les moyennes à long terme de ses débits. C'est donc un bon indicateur des variations des écoulements de la rivière. De fait, il peut soit se produire à l'identique d'une période à l'autre avec les mêmes traits caractéristiques, soit présenter une alternance de différents types de régimes. Généralement, le régime est considéré comme une caractéristique qui ne change pas au cours du temps, ce qui semble être une simplification abusive eu égard aux perpétuelles modifications du bassin versant et des conditions climatiques du milieu. Ainsi, du point de vue du climat (respectivement du bassin versant), il est certain que les conditions d'occurrence des débits mensuels caractérisant le régime hydrologique auront varié. Cette évolution des conditions dans lesquelles les débits mensuels se produisent met en cause la stationnarité du processus hydrologique sous-jacent et, par conséquent du régime hydrologique. Or, la stationnarité des échantillons de débits est, avec l'indépendance des observations, une des conditions essentielles d'applications des méthodes de la statistique classique sur de tels échantillons de débit.

Les régimes hydrologiques ne sont donc pas immuables. Ce sont des produits de la conjonction d'un climat et d'un bassin versant. Ce faisant, les séries temporelles de débits qui représentent de tels régimes sont susceptibles de refléter les modifications ou les transformations du climat ou du bassin versant. De telles modifications ou transformations caractérisent la non-stationnarité des processus hydrologiques sous-jacents. En somme, la non-stationnarité d'un processus représentatif de variable hydrométéorologique peut être due à :

- une transformation du bassin versant : causes anthropiques
- une modification climatique : causes climatiques

4.4.2.1 Non-stationnarité caractérisée par la transformation des bassins versants

On admet généralement que le caractère stationnaire de l'hydrologie du bassin versant et des séries temporelles de débits qui la représentent peut être perturbé à la longue par la transformation des bassins. En effet, le bassin versant est le siège de nombreuses activités humaines. Qu'elles soient le résultat d'une action volontaire ou involontaire, les conséquences de ces activités transforment profondément le fonctionnement du bassin versant. Par exemple, pour lutter contre les inondations l'homme a modifié, grâce à des aménagements et / ou à des ouvrages, le régime «naturel» d'écoulement des eaux (et par conséquent du bassin versant) dans un sens favorable à ses désirs et / ou à ses intérêts. Les barrages ou les digues sont par excellence les ouvrages qui permettent à l'homme de commander les débits à l'aval, c'est-à-dire de fabriquer un régime artificiel plus conforme à ses souhaits. En exploitant les ressources que recèlent les milieux aquatiques (l'eau pour l'alimentation domestique, l'industrie et l'agriculture, etc.), l'homme a également transformé les bassins versants, soit en modifiant la couverture végétale, essentielle pour retenir les eaux de ruissellement.

4.4.2.2 Non-stationnarité caractérisée par une modification climatique

La modification climatique agit sur les caractéristiques d'un régime hydrologique. Son effet peut contribuer à une modification du cycle hydrologique global en intensifiant certains flux comme l'évaporation. Ce qui conduit inexorablement à modifier les conditions d'occurrence des débits qui représentent le régime hydrologique : des crues plus accentuées et plus fréquentes en hiver, des étiages plus marqués en été.

La modification climatique est donc une cause potentielle de non-stationnarité des variables hydrométéorologiques. Elle est largement imputable à l'émission abusive des concentrations importantes de gaz à effet de serre dans l'atmosphère, notamment le dioxyde de carbone (CO_2), le méthane (CH_4) et le protoxyde d'azote (N_2O). Cette émission est due aux activités humaines, principalement à la combustion d'agents énergétiques fossiles, à des changements de l'affectation des sols, et à l'agriculture. Notons aussi qu'en plus des gaz à effet de serre, il faut prendre également en compte les aérosols, c'est-à-dire des infimes particules en suspension dans l'air, produites principalement par des processus de combustion.

4.5 Conclusion

Dans ce chapitre, on a vu que pour un régime hydrologique, le débit en un point de son cours d'eau à un instant t donné constitue une représentation du bassin versant (ainsi que des conditions climatiques du milieu) en évolution aléatoire dans le temps. La suite continue des débits, $\{x_1,\ldots,x_n\}$, mesurées durant un intervalle de temps $[t_1,t_n]$ se présente ainsi comme une réalisation d'un processus stochastique $\{X_1,\ldots,X_n\}$. Un cours d'eau possède donc une certaine "mémoire" qui fait que le débit à l'instant t est lié en partie par les états antérieurs du bassin versant et/ou des conditions climatiques du milieu. Il s'ensuit qu'au cours d'une période de temps donnée, un certain nombre de caractéristiques hydrométéorologiques d'un bassin versant (et/ou des conditions climatiques du milieu)

restent suffisamment stables ou évoluent suffisamment lentement pour qu'il en résulte dans la série temporelle des débits l'apparition d'une "tendance" autour de laquelle se répartissent des variations mensuelles par exemple. Cette tendance est la résultante de l'ensemble des facteurs explicatifs de l'évolution de la série temporelle des débits. La tendance à long terme est ainsi le facteur représentant l'évolution à long terme de la grandeur et traduit l'aspect général de la série temporelle tandis que la rupture est le facteur représentant l'évolution discontinue de la grandeur et est caractéristique d'un changement de régime dans la série temporelle. Ce sont principalement ces facteurs qui caractérisent la non-stationnarité des processus hydrologiques sous-jacents

5 TESTS STATISTIQUES DE STATIONNARITÉ EN HYDROMÉTÉOROLOGIE : ÉTAT DE L'ART

Après avoir établi dans le chapitre précédent qu'il existe différentes sources de non-stationnarité dans les processus représentatifs de variables hydrométéorologiques, on va maintenant s'intéresser aux tests statistiques qui permettent de vérifier si un processus donné est stationnaire ou non. Ce chapitre présente une revue bibliographique des tests statistiques de détection de non-stationnarité dans une série de données d'une variable hydrométéorologique. Le but poursuivi n'est pas de faire un inventaire détaillé et précis des tests de stationnarité. L'objet du chapitre est plutôt de donner un aperçu des différents tests et d'indiquer les tendances de la recherche actuelle, ainsi que certains liens avec d'autres disciplines comme l'économétrie et la statistique.

Le premier point développé réalise l'inventaire des tests largement employés dans les études de stationnarité en hydrométéorologie. Comme cette thèse s'intéresse à la non-stationnarité de type rupture, la partie suivante est une synthèse bibliographique dressant l'état de l'art des tests de rupture utilisés à ce jour dans la littérature statistique et économétrique. Il est en effet essentiel de faire le parallèle entre ces trois littératures qui sont relativement disparates.

5.1 Considérations d'ordre général

Lorsqu'on veut vérifier la stationnarité d'une série temporelle d'une variable hydrométéorologique, on veut montrer l'impact sur les observations d'un changement dans les conditions entourant l'occurrence de ces observations. On entend par conditions, les modifications climatiques du milieu et les transformations du bassin versant d'origine anthropique.

Souvent, dans la pratique, il arrive que l'on confonde la notion de stationnarité et la notion d'homogénéité en hydrométéorologie. Lorsque l'on vérifie l'homogénéité d'une série d'observations, on veut montrer l'impact sur les observations d'un changement dans les conditions entourant la saisie de ces observations. On entend par conditions, l'appareil de mesure et son environnement, l'observateur et la chaîne d'opérations exécutées par l'observateur avant, pendant et après la lecture de l'appareil de mesure. Ainsi, l'homogénéité garantit que la chaîne d'élaboration d'obtention de la mesure n'introduit pas, à chacune de ses étapes, des erreurs susceptibles de mettre en doute la valeur finale obtenue ou des incertitudes donnant finalement un intervalle de confiance trop grand pour permettre l'utilisation de la valeur obtenue. En revanche la stationnarité se définit quant à elle par rapport à une échelle d'espace et de temps. Ces échelles permettent en quelque sorte de définir le domaine de validité de la mesure qui est assuré par la stabilité du processus sous-jacent : la variable hydrométéorologique conserve la même définition sur toute la période objet de l'étude. Cela dit, la vérification de la stationnarité d'une série de données d'une variable hydrométéorologique n'a un sens que si on a vérifié au préalable qu'aucun aléa technique (l'appareil de mesure et son environnement, l'observateur et la chaîne d'opérations exécutées par l'observateur avant, pendant et après la lecture de l'appareil de mesure, observateur unique, conservation de l'échantillon de données) n'a perturbé la mesure durant la période d'observation. Autrement dit, on doit d'abord vérifier que la dite série est homogène.

Grâce aux efforts méthodologiques accomplis au cours des dernières années dans le domaine de validation des données hydrométéorologiques, il est aujourd'hui possible de spécifier correctement la non-stationnarité des séries temporelles hydrométéorologiques. Du point de vue statistique, celle-ci a souvent été expliquée par l'existence d'une singularité, en particulier par un changement de moyenne (ou de variance), provoquée par une modification d'origine naturelle et/ou anthropique du processus physique générateur. Cavadias (1992) a identifié différents types de modifications climatiques que l'on peut associer à des évolutions possibles des variables climatiques :

- un saut brusque de la moyenne de la série ;
- une succession de changements de la moyenne de la série (*shifting levels*) ;
- une tendance continue de la moyenne de la série ;
- une évolution cyclique de la série ;
- un changement de la variance de la série.

De manière plus générale, on peut considérer qu'il existe deux grands types de non-stationnarité à suivre pour les variables hydrométéorologiques (Ouarda et al., 1999) :

1. changement de la moyenne générale de la série temporelle ;
2. changement de la variance (diminution ou augmentation) de la série temporelle.

Ces différents types de changements dans une série temporelle peuvent correspondre soit

- à des modifications climatiques sur le régime hydrologique : un réchauffement du climat entraînera une certaine redistribution des régimes des écoulements, en particulier, en hiver, les crues pourraient être plus fréquentes alors qu'en été, elles auraient tendance à se raréfier ;
- à une influence humaine sur le régime hydrologique : conséquences des modifications de l'accessibilité à la ressource en eau sur l'aménagement du territoire et sur les pratiques agricoles par exemple ;
- aux variations du régime des débits dues à un changement spontané du cours d'eau (déforestation ou un reboisement), etc.

5.2 Synthèse bibliographique : outils de détection de non-stationnarités en hydrométéorologie

Diverses méthodes statistiques ont été proposées dans la littérature pour vérifier la stationnarité d'une série de données. Ces méthodes aboutissent principalement aux deux

catégories suivantes. La première concerne les méthodes graphiques qui sont généralement simples à mettre en œuvre et ne prétendent pas découvrir un type particulier de violation de stationnarité. Elles constituent une simplification des méthodes de détection de non-stationnarités dans une série de données. La seconde catégorie se rapporte aux méthodes d'inférence statistique. Ces dernières consistent le plus souvent à la mise en œuvre et à l'interprétation des tests statistiques d'homogénéité des séries. Ce sont des tests du caractère aléatoire d'une suite de variables indépendantes, qui impliquent l'hypothèse de stationnarité (jouant le rôle d'hypothèse "principale") et contre laquelle on teste diverses contre-hypothèses dont les plus importantes sont :

- l'existence d'une tendance déterministe dans la série (changement continu ou discontinu de la moyenne générale et/ou de la variance de la série) ;
- l'existence d'une tendance stochastique dans la série (changement de la moyenne de la série suivant une "marche aléatoire") ;
- l'existence d'une dépendance quelconque (existence d'une persistance ou d'une saisonnalité dans la série).

Chacune de ces "infractions" à la stationnarité est testée au moyen de procédures appropriées à chaque cas, de telle sorte que le rejet de l'hypothèse de stationnarité donne, en même temps une indication sur le type de non-stationnarité suspectée dans la série.

5.2.1 Procédures graphiques de détection de non-stationnarités

Les premiers outils auxquels on a eu recours pour détecter des non-stationnarités dans les séries de données de variables hydrométéorologiques étaient des procédures graphiques. Elles consistent à obtenir un certain aperçu de l'homogénéité temporelle des observations de la série temporelle. On connaît la popularité des techniques dites *des simples cumuls* et *des doubles cumuls* (appelées improprement *simple masse* et *double masse*). Ces techniques, qui furent proposées dans la littérature hydrométéorologique par Searcy et Hardison (1960), utilisent d'une part l'information chronologique par cumul des valeurs dans le temps (méthode des simples cumuls) et d'autre part une information extérieure en

la comparant au cumul de valeurs liées de façon statistique aux valeurs de la série temporelle à contrôler (méthode des doubles cumuls). Ainsi, on peut résumer le principe de détection de non-stationnarités comme suit (Brunet-Moret, 1971) :

- on suit le cumul des valeurs mesurées en un point en fonction du temps, à partir d'une date initiale. Des changements de pentes, indiquent des évolutions de valeurs moyennes, donc une série non-stationnaire ;
- en portant en ordonnée le cumul dans l'ordre chronologique des valeurs de la série de données à vérifier et en abscisse le cumul des valeurs concomitantes d'une série de contrôle, on repère les ruptures de pentes, et donc de non-stationnarités, en suivant le cumul des valeurs de la série en fonction du cumul des valeurs de la série de contrôle (voir l'illustration à la figure 5.1).

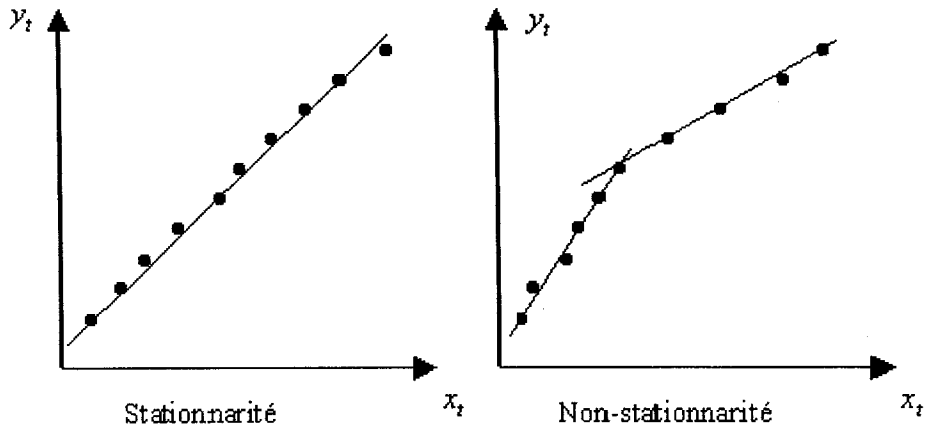


Figure 5.1 : Illustration de la méthode du cumul

Le *corrélogramme* est aussi un outil graphique destiné à mettre en évidence les composantes non-stationnaires d'une série temporelle telles que la variation saisonnière, la persistance et la tendance. Rappelons qu'on appelle coefficient d'autocorrélation d'ordre θ le coefficient de corrélation linéaire calculé entre la série temporelle et cette même série décalée de θ périodes de temps. Le corrélogramme est la représentation graphique de la

suite des autocorrélations en fonction du décalage θ. Il peut présenter diverses allures caractéristiques dans l'organisation d'une série temporelle. C'est-à-dire,

- si la série des autocorrélogrammes fluctue autour de zéro, on dira que la série est stationnaire (c'est la caractéristique d'un bruit blanc) ;
- si la série comporte une variation saisonnière, celle-ci sera clairement mise en évidence sur l'autocorrélogramme par une apparition de cyclicités ;
- si la série comporte une persistance, celle-ci sera mise en évidence par la décroissance des autocorrélations en fonction de leur ordre de décalage θ ;
- si la série des autocorrélogrammes décroît très lentement, la série n'est vraisemblablement pas stationnaire et il est important de mettre en évidence des évolutions de tendance. Ainsi lorsque les coefficients d'autocorrélation restent relativement élevés jusqu'à un rang important, la série peut être considérée comme monotone. Si par contre la tendance change de sens au cours de la période étudiée, certains coefficients d'autocorrélation prennent des valeurs négatives. Leurs variations sont alors caractéristiques des changements de régime (non-stationnaires).

En somme, l'utilisation du corrélogramme permet de mettre en évidence certains types de non-stationnarité. Mais il semble, d'après Bell et al. (1986), que son emploi doit rester limité à une simple indication de non-stationnarité. En effet, il ne permet pas de conclure avec beaucoup de confiance que la variable étudiée est non-stationnaire quand c'est effectivement le cas.

Les méthodes graphiques de détection qui viennent d'être présentées ont principalement l'avantage d'être simples à mettre en œuvre. Elles sont bien connues des praticiens et constituent une simplification des méthodes de détection de non-stationnarité dans un échantillon de données. Par contre, en dépit du fait qu'elles sont utiles pour détecter des changements dans une série, elles ne permettent pas de dissocier les changements réels des fluctuations aléatoires. De plus l'interprétation des graphiques n'est pas toujours aisée et, surtout ces méthodes ne proposent pas de tests statistiques pour quantifier en probabilité la

signification des changements apparents. C'est pourquoi la signification des résultats obtenus avec de telles procédures de détection doit être nuancée. Parallèlement, on note que Bois (1971, 1986) a proposé une extension de l'idée de la méthode du double cumul en y ajoutant un contenu statistique autorisant la pratique d'un véritable test de stationnarité. Cette méthode bien connue sous le nom d'ellipses de Bois complète le test de la statistique de Buishand (1982). Elle est un progrès décisif en soi, mais elle ne peut constituer un outil suffisant pour vérifier la stationnarité d'une série temporelle. D'où l'importance de se préoccuper des tests statistiques de détection de non-stationnarité dans une série temporelle.

5.2.2 Revue bibliographique des tests de stationnarité appliqués en hydrométéorologie

La littérature hydrométéorologique recèle un volume important de tests reliés au problème de détection de non-stationnarités dans une série temporelle. On rappelle les travaux pionniers de Kendall et Stuart (1943) et de l'Organisation Mondiale de la Météorologie (WMO, 1966). Ces premiers travaux préconisent la mise en œuvre de tests statistiques de stationnarité qui portent sur la constance de la moyenne (ou de la variance) de la série tout au long de sa période d'observation.

Depuis ces premiers travaux, les recherches se sont orientées vers une modélisation plus riche de la contre-hypothèse. De façon générale, l'hypothèse principale considère maintenant que la série est aléatoire simple tandis que les contre-hypothèses dépendent fortement de la mise en évidence d'une non-stationnarité basée sur le type d'infraction à la stationnarité. C'est-à-dire soit que la série démontre une tendance déterministe ou stochastique, soit que la série démontre une persistance, une saisonnalité ou encore une combinaison des deux. Dans ce qui suit, on fait un survol des tests de stationnarité les plus utilisés en hydrométéorologie, et de préférence déjà éprouvés dans les études pratiques.

5.2.2.1 Tests vérifiant une non-stationnarité de type tendance déterministe

Le problème des tests de détection de non-stationnarité de type tendance déterministe a reçu beaucoup d'attention dans la littérature des séries de données d'une variable hydrométéorologique. Les contributions les plus importantes considèrent deux catégories de tests :

1. les tests de détection d'une non-stationnarité de type tendance en saut(s) ;
2. les tests de détection d'une non-stationnarité de type tendance linéaire ;

5.2.2.1.1 Tests de détection d'une non-stationnarité de type tendance en saut(s) : détection de rupture(s)

Le problème de tendance en saut fait référence dans la littérature au problème de rupture dont le principe est le suivant. Lorsqu'on observe un phénomène hydrométéorologique, on peut se demander s'il est stable sur toute la période étudiée. Bien évidemment, il peut arriver que certaines circonstances provoquent une altération du modèle à partir d'une certaine date t_0. De façon plus concrète, soit $\{x_1,\ldots,x_n\}$ une réalisation du phénomène sur la période étudiée. On cherche laquelle des deux hypothèses suivantes est vraie :

$$\begin{aligned} &H_0 : \{x_1,\ldots,x_n\} \text{ suit le modèle } m_\theta \\ &H_1 : \text{il existe un instant } t_0 \text{ tel que} \\ &\quad \begin{cases} \{x_1,\ldots,x_{t_0}\} & \text{suit le modèle } m_\theta \\ \{x_{t_0+1},\ldots,x_n\} & \text{suit le modèle } m_{\theta'} \end{cases} \end{aligned}$$

où,

- m_θ représente une famille de modèles paramétrés par θ qui peut être un vecteur.

Si la contre-hypothèse H_1 est retenue la question subsidiaire est de déterminer la date de rupture t_0, l'intensité de la rupture et les paramètres θ et θ'. De manière générale, la façon dont l'altération se traduit dépend de l'objet étudié et de la modélisation adoptée. En

effet, il peut s'agir d'une rupture brusque (abrupte) ou avec continuité de la moyenne (ou variance) de la série $\{x_1,\ldots,x_n\}$. Par exemple, la figure 5.2 illustre le cas d'une rupture brusque survenue à la date t_0.

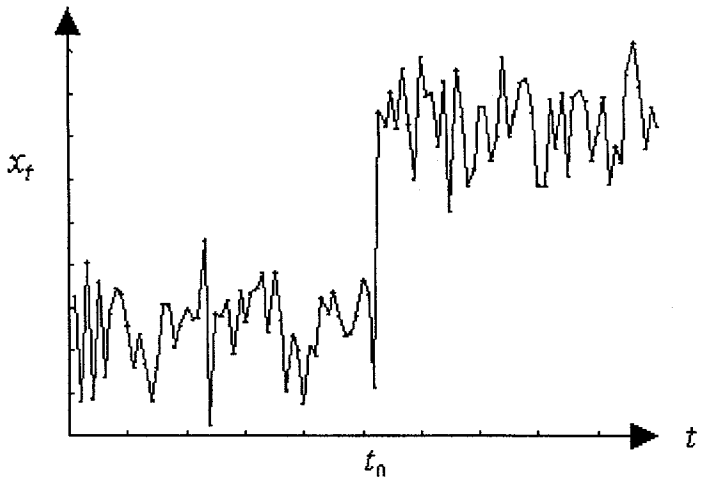


Figure 5.2 : Illustration d'une rupture brusque survenue à l'instant t_0

On appelle rupture tout changement dans la loi de probabilité du processus générateur de la série à un instant donné t_0. Ce changement survient instantanément ou très vite à l'échelle de la période d'échantillonnage. Il ne s'agit pas nécessairement d'un changement de forte amplitude. En réalité, un changement même de faible amplitude peut être révélateur d'une modification du processus étudié, et il peut être fondamental à plus d'un titre (pour la fiabilité des modèles de planification de la ressource en eau) de diagnostiquer cette modification. Le terme de rupture est donc étendu au sens de "date" à laquelle les propriétés statistiques du processus générateur de la série (moyenne, variance, etc.) ont subitement changé (non-stationnaire), étant entendu que ces propriétés restent stables pendant un certain temps avant et après la date de la rupture t_0. C'est cette idée qui est à la base de la formulation mathématique de la notion de rupture et c'est elle qui inspire les tests statistiques de détection qui sont présentés dans ce survol.

Il existe une littérature abondante sur les tests de détection de rupture en hydrométéorologie. Ces tests sont construits à partir du principe suivant. En régime stationnaire, la moyenne (ou la variance) d'une série $x_t = \{x_1, \ldots, x_n\}$ est constante par rapport au temps : $E(x_t) = \mu$ (respectivement $Var(x_t) = \sigma^2$). Partant de ce fait, on cherche à savoir si le processus physique générateur de la série x_t a subit une altération à partir d'une date t_0 supposée connue ou inconnue et comprise entre 1 et *n*, l'altération se traduisant par une rupture de la stationnarité de la moyenne μ (respectivement de la variance σ^2) du processus. Détecter une rupture dans la série x_t revient dans ces conditions à tester l'hypothèse principale H_0 d'absence de rupture, contre la contre-hypothèse H_1 de présence de rupture. De manière générale, la formulation de H_1 peut différer par

- le type d'altération redouté (rupture brusque ou avec continuité) ;
- la détermination d'une altération unique ou multiple (une date de rupture ou plusieurs dates de rupture) ;
- l'amplitude de l'altération (l'amplitude de la rupture) est supposée connue ou non.

Beaucoup d'auteurs se sont intéressés au cas de l'échantillon gaussien (ou éventuellement pour des distributions exponentielles). Le problème suppose que la rupture est brusque et sa détection est d'autant plus facile que son amplitude est plus grande. La procédure de détection procède ainsi : on considère un échantillon indépendant $\{x_1, \ldots, x_n\}$ et on veut tester les hypothèses

$$\begin{cases} H_0 : & x_1, \ldots, x_n & \sim & F(x_t/\theta) \\ H_1 : & x_1, \ldots, x_{t_0} & \sim & F(x_t/\theta) \\ & x_{t_0+1}, \ldots, x_n & \sim & F(x_t/\theta') \end{cases}$$

où,

- θ représente le vecteur des paramètres (en général de localisation d'échelle) ;
- t_0 désigne la date de rupture dans la série de données.

Lorsque la date de rupture t_0 est connue, il existe un certain nombre de tests statistiques que l'on utilise dans la littérature hydrométéorologique pour vérifier la stationnarité de la série. Ce sont des tests de conformité et des tests d'homogénéité. Les tests de conformité impliquent habituellement la comparaison de la moyenne (respectivement la variance) de l'échantillon avant ou après la date t_0 à la moyenne (respectivement la variance) de la loi théorique de la population dont il est issu. En revanche les tests d'homogénéité comparent des caractéristiques statistiques d'échantillons (moyenne ou variance) : étant donné l'échantillon avant et l'échantillon après la date de rupture t_0, on cherche principalement à vérifier s'ils ont été prélevés dans une même population indépendamment l'un de l'autre. Par ailleurs, on utilise aussi les tests d'ajustement pour vérifier si l'échantillon avant et après la date t_0 ont été tirés d'une même population parente. Faucher et al. (1997) donnent en référence une longue liste de ces tests. Parmi les tests non-paramétriques les plus usités pour détecter une rupture de la moyenne de la série, Berrymann et al. (1984) concluent que le test de Mann-Whitney (Mann et Whitney, 1947) est le plus efficace (son efficacité asymptotique est toujours supérieure à 0.864). Par ailleurs, il note aussi que d'autres tests comme ceux de Terry-Hoeffding (Terry, 1952 ; Hoeffding, 1950), de Van der Waerden (1952) et de Bell-Doksum (Conover, 1971) possèdent une efficacité asymptotique comparable à celle du test de Mann-Whitney. Pour les tests paramétriques, Berrymann (1988) note que le test de Student est le plus efficace (efficacité = 1).

Dans l'autre situation où la date de rupture t_0 est inconnue, la littérature hydrométéorologique est moins développée parce que la nature du problème est beaucoup plus complexe. Les quelques tests de détection que l'on retrouve dans cette littérature sont des tests de type supremum dont le principe consiste à rechercher la date la plus probable de la rupture en la faisant varier, et en supposant qu'il n'y a qu'une seule rupture dans l'échantillon. On calcule ainsi, pour chaque date probable de rupture, la valeur qu'une statistique de test d'homogénéité donnée prend sur la série étudiée. L'instant de rupture est

estimé à la date pour laquelle cette statistique de test atteint son maximum dans la série. Bien que ces tests soient la seule solution proposée dans la littérature hydrométéorologique, il reste cependant un certain nombre de points théoriques et méthodologiques récurrents à développer, notamment :

- l'absence de bons points critiques pour la prise de décision ;
- le problème de contrôle de niveau ;
- le choix de la bonne date de rupture.

Parmi les tests les plus utilisés dans cette littérature, on souligne le test du saut de la moyenne de Buishand (Buishand, 1982), le test du saut de la moyenne de Worsley (Worsley, 1979) et le test de Pettitt (Pettitt, 1979). Le test de Buishand est de nature bayésienne. Il fait référence au modèle simple qui suppose un changement dans la moyenne de la série :

$$x_t = \begin{cases} \mu + \varepsilon_t, & t = 1,\ldots,t_0 \\ \mu + \Delta + \varepsilon_t, & t = t_0 + 1,\ldots,n \end{cases} \qquad (5.1)$$

où,

- $\varepsilon_t \overset{iid}{\sim} N(0,\sigma^2)$;
- t_0 est la date de rupture ;
- $E(x_t) = \mu$ sous l'hypothèse principale d'absence de rupture ;
- Δ est l'amplitude de la rupture ;
- t_0 et les paramètres μ, σ^2 et Δ sont inconnus.

En supposant une distribution *a priori* uniforme pour la position de t_0, on définit la statistique de test U par :

$$U = [n(n+1)]^{-1} \sum_{k=1}^{n-1} \left(S_k^* / D_x\right)^2 \qquad (5.2)$$

où,

- $S_0^* = 0, \quad S_k^* = \sum_{t=1}^{k} (x_t - \bar{x})$ pour $k = 1, \ldots, n$
- $D_x = n^{-1} \sum_{t=1}^{n} (x_t - \bar{x})^2$

Les points critiques de cette statistique de test sont dérivés à partir de la méthode de simulation de type Monte Carlo (Buishand, 1982). L'estimation de la date de rupture est donnée par (Buishand, 1982) :

$$\hat{t}_0 = \arg\left(\max_{1 \le k \le n} \left\{ \left| S_k^* \right| \right\}\right) \qquad (5.3)$$

Le test de Worsley est un test paramétrique. Il fait aussi référence au modèle simple de l'équation 5.1. La statistique de test est

$$W = \max_{1 \le k \le n-1} \left\{ \left| T_k \right| \right\} \qquad (5.4)$$

où,

- T_k est la statistique de test de Student pour tester la différence de moyenne entre les k premières observations et les $(n-k)$ dernières observations.

En utilisant les sommes cumulées et le rapport de vraisemblance de Worsley, cette statistique de test s'exprime encore par :

$$W = \max_{1 \le k \le n-1} \left\{ \left| S_k^* \right| / D_x \sqrt{k(n-k)} \right\} \qquad (5.5)$$

Les points critiques exacts de cette statistique de test sont dérivés pour des tailles d'échantillons inférieures à 10 ($n \le 10$). Par contre, pour des échantillons excédant 10, les

points critiques correspondant sont obtenus à partir des méthodes de simulation (Worsley, 1979). L'estimation de la date de rupture est donnée par (Worsley, 1979):

$$\hat{t}_0 = \arg\left(\max_{1\le k\le n-1}\left\{\left|S_k^*\right| / D_x\sqrt{k(n-k)}\right\}\right) \qquad (5.6)$$

Finalement, le test de Pettitt (1979) est un test non-paramétrique. Il est dérivé à partir du test de Mann-Whitney (Mann et Whitney, 1947). Son principe est le suivant : on cherche la date la plus probable de la rupture t_0 en vérifiant, pour t_0 variant de 1 à n, si les deux séries $\{x_1,\ldots,x_{t_0}\}$ et $\{x_{t_0+1},\ldots,x_n\}$ appartiennent à la même population. La statistique de test est (Pettitt, 1979) :

$$K_n = \max_{1\le t\le n-1}\left\{\sum_{i=1}^{t}\sum_{j=t+1}^{n} D_{ij}\right\} \qquad (5.7)$$

où,

- $D_{ij} = \operatorname{sgn}(x_i - x_j) = \begin{cases} 1 & \text{si} \quad x_i > x_j \\ 0 & \text{si} \quad x_i = x_j \\ -1 & \text{si} \quad x_i < x_j \end{cases}$

Si k_n^{obs} désigne la valeur de la statistique de test K_n obtenue pour la série étudiée, sous l'hypothèse principale d'absence de rupture, la probabilité de dépassement de la valeur k_n^{obs} est donnée approximativement par (Pettitt, 1979) :

$$\mathrm{P}\left(K_n > k_n^{obs}\right) \approx 2\exp\left\{-6\left(k_n^{obs}\right)^2 / \left(n^3 + n^2\right)\right\} \qquad (5.8)$$

L'estimation de la date de rupture est donnée par (Pettitt, 1979) :

$$\hat{t}_0 = \arg\left(\max_{1\le t\le n-1}\left\{\sum_{i=1}^{t}\sum_{j=t+1}^{n} D_{ij}\right\}\right) \qquad (5.9)$$

En général, si une rupture brusque survient au milieu de la série, le test de Buishand et le test de Pettitt sont plus puissants pour la détecter. Par contre si la rupture survient aux extrémités de l'échantillon de la série de données, c'est le test de Worsley qui est le plus puissant pour la détecter (Buishand, 1982, 1984 ; Khodja et al., 1998). D'autre part, le test de Buishand et le test de Pettitt sont réputés pour leur robustesse, en ce sens qu'ils restent valides même pour des distributions de la variable étudiée qui s'écartent de la normalité (Lubès et al., 1998). Par ailleurs il apparaît que la persistance et la présence de tendance dans les séries sont les caractéristiques qui pénalisent le plus les performances de ces tests (Lubès et al., 1998). De plus, ces tests ne sont pas conçus pour détecter une rupture de moyenne avec continuité.

Les procédures bayésiennes ne sont pas en reste pour détecter une rupture lorsque la date de la rupture t_0 est inconnue. Dans la littérature hydrométéorologique, c'est la procédure de Lee et Heghinian (Lee et Heghinian, 1977) qui est largement employée dans les études pratiques. Elle fait référence au modèle simple de l'équation 5.1 et vise à confirmer ou à infirmer l'hypothèse d'un changement brusque de moyenne dans la série. Cette méthode est fondée sur les modes des distributions marginales a posteriori de la position de la rupture dans le temps (t_0) et de l'amplitude du changement sur la moyenne (Δ). À ces modes sont associés des probabilités dont l'exploitation permet de dégager des outils de diagnostic. Habituellement, on se limite à la distribution a posteriori de t_0. Ainsi, l'instant de la rupture est estimé par le mode de la distribution a posteriori de t_0. Bernier (1994) a généralisé la procédure de Lee et Heghinian pour différentes structures de changement de la moyenne de la série.

Un des inconvénients majeurs lié à l'utilisation de la procédure bayésienne proposée par Lee et Heghinian (1977) est qu'on travaille avec l'hypothèse qu'un changement de moyenne est effectivement présent dans la série. En effet, en présence d'une série de données stationnaires, cette procédure trouvera toujours un changement (non-stationnarité) à la date t_0 pour laquelle la probabilité a posteriori de réalisation du changement est maximale. Perreault (2000) a proposé une alternative à ce problème. Pour prendre en

compte "l'hypothèse d'absence de changement" et les différents types de changements, il assigne des probabilités a priori aux différentes alternatives et ils considèrent l'étude de la date de changement comme un problème bayésien de choix de modèles. Le but de cette approche est de construire des modèles de rupture traduisant la manifestation locale d'un changement sur une série de données hydrologiques et de développer les outils bayésiens d'estimation de la date de rupture. La perspective bayésienne et l'estimation font intervenir l'échantillonnage de Gibbs.

D'autres tests de rupture concernent l'utilisation des modèles de régression. En général, la variable explicative est le temps et l'on suppose que la rupture est brusque. Le modèle résultant et les hypothèses à tester sont donnés aux équations 3.11 et 3.12. Lorsque la date de rupture t_0 est connue, la procédure classique de la régression est appliquée : on détecte la rupture de pente, caractérisant le changement brusque de la moyenne, à l'aide du test de Student (Vincent, 1990 ; Laberge, 1995). Dans le cas où la date de rupture t_0 n'est pas connue, très peu de travaux existent dans la littérature hydrométéorologique. Soulignons cependant la contribution de Solow (1988). En s'inspirant des travaux de Hinkley (1969, 1971) Solow a développé une procédure bayésienne de détection à partir du modèle de régression suivant :

$$x_t = \begin{cases} \alpha_0 + \beta_0 t + \varepsilon_t, & t = 1, \ldots, t_0 \\ \alpha_1 + \beta_1 t + \varepsilon_t, & t = t_0 + 1, \ldots, n \end{cases} \qquad (5.10)$$

où,

- $\alpha_0, \alpha_1, \beta_0$ et β_1 sont des termes constants a priori inconnus.

En résumé, la plupart des tests de détection d'une rupture présentés dans ce survol procèdent en spécifiant une alternative à l'hypothèse principale d'absence de rupture. Bien évidemment, cette liste ne concerne pas les tests basés sur les écarts cumulés (test de

Worsley et test de Buishand) qui ne spécifient pas un type particulier de violation de l'hypothèse principale d'absence de rupture. Il va sans dire que ces tests ont des propriétés optimales dans le cas de changement brusque de la moyenne. Tous ces tests ne sont conçus que pour détecter une rupture brusque d'une caractéristique statistique (moyenne ou variance) dans une série de données indépendantes. En conséquence, ils vérifient une non-stationnarité d'ordre 1.

Bien que la littérature des tests de détection d'une rupture soit vaste, les études qui se sont directement intéressées au problème relatif à plusieurs ruptures sont peu nombreuses. À cet effet, on peut citer les tests non-paramétriques de Kruskal-Wallis (1952) et de Terpstra-Jonckheere (Terpstra, 1952 ; Jonckheere, 1954). Ces tests procèdent en découpant l'échantillon de données en plusieurs parties et en testant les moyennes (ou les variances). Pour les tests paramétriques, on note essentiellement le test d'analyse de variance pour la comparaison de plusieurs moyennes. Cette liste inclut également la procédure de segmentation des séries hydrométéorologiques proposée par Hubert et al. (1989). En effet, cette procédure peut être interprétée comme un test de stationnarité, l'hypothèse principale étant que la série étudiée est stationnaire. Son principe consiste à découper la série en segments ($m > 1$) de telle sorte que la moyenne calculée sur tout segment soit significativement différente de la moyenne du (ou des) segments(s) voisin(s). Une telle approche consiste alors à rechercher de multiples changements de moyenne. Ainsi, si la procédure ne produit pas de segmentation acceptable d'ordre supérieur ou égale à 2, l'hypothèse principale de stationnarité est acceptée. Il est cependant important de faire remarquer que cette procédure n'est pas en toute rigueur, un test statistique mais plutôt une méthode d'estimation du nombre de rupture(s).

5.2.2.1.2 Tests de détection d'une non-stationnarité de type tendance linéaire

Dans ce type de tendance, on cherche à vérifier si la distribution de probabilité du processus physique générateur de la série évolue avec le temps. L'illustration d'une tendance linéaire est donnée à la figure 5.2.

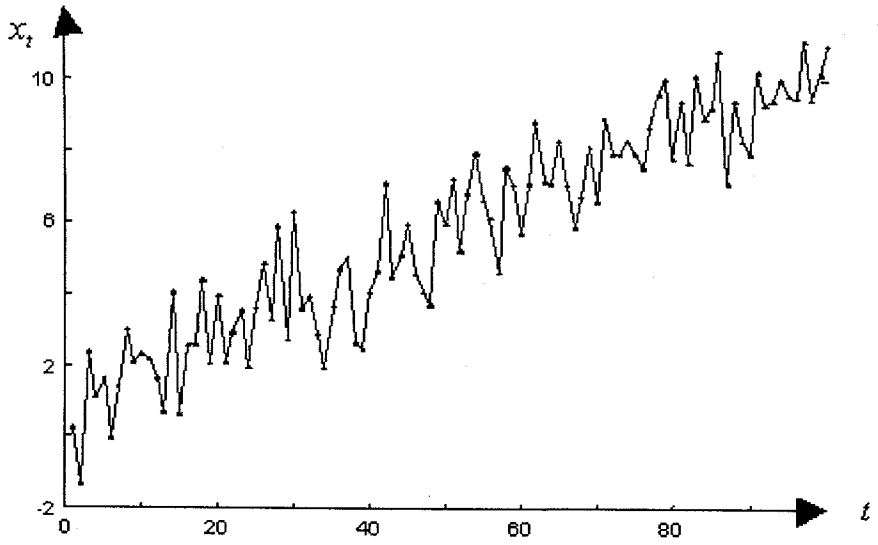


Figure 5.3 : Illustration d'une tendance linéaire

Un nombre important de tests paramétriques et non-paramétriques ont été proposés dans la littérature pour détecter une tendance linéaire dans les séries chronologiques de données hydrométéorologiques. Les tests paramétriques sont basés essentiellement sur le modèle de régression représenté à l'équation 3.11. Ils sont conçus pour tester l'hypothèse nulle de stationnarité H_0 : $\beta = 0$ (absence d'une tendance linéaire dans la série de données) contre la contre-hypothèse qui prétend le contraire. Le résultat de ce test est conditionné par la vérification des hypothèses standards d'un modèle de régression, notamment que les erreurs du modèle sont indépendantes et identiquement distribuées suivant la loi normale $N(0,\sigma^2)$. Très souvent, de telles hypothèses sont difficiles à vérifier avec les séries de données de variables hydrométéorologiques. C'est pourquoi on préfère utiliser les tests non-paramétriques pour détecter les non-stationnarités de type tendance linéaire. Parmi les tests les plus employés dans la littérature hydrométéorologique pour cette dernière catégorie, le test de Mann-Kendall (Mann, 1945; Kendall, 1975) est le plus performant. Il permet de vérifier l'hypothèse principale qu'il n'existe aucune tendance linéaire ni de structure particulière d'autocorrelation dans la série contre la contre-hypothèse qui affirme l'existence d'une tendance dans les données composant la série. Par ailleurs, il est capable de détecter une tendance linéaire de la moyenne (ou de la variance) de la série (Sneyers, 1990).

5.2.2.2 Tests de détection d'une non-stationnarité de type tendance stochastique

Lorsqu'une série chronologique est caractérisée par une tendance stochastique, on dit qu'elle provient d'un processus de «marche aléatoire». (on peut penser à un ivrogne qui prend de façon aléatoire un pas à gauche ou un pas à droite, et qui n'a aucune tendance à revenir à son point de départ). La figure 3.3 montre une illustration de ce type de tendance. À première vue, il est difficile de distinguer à l'œil entre tendances stochastiques et déterministes, d'où l'importance de développer des tests formels d'hypothèses.

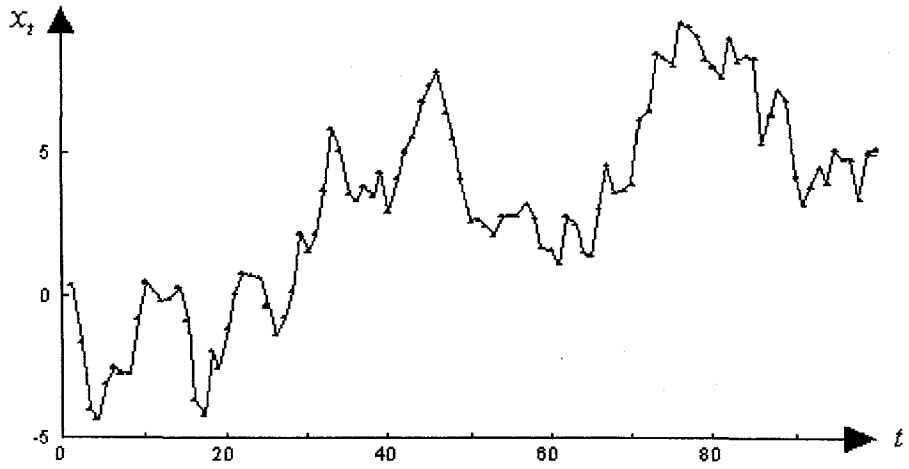


Figure 5.4 : Illustration d'une tendance stochastique

Dans la littérature hydrométéorologique, très peu de travaux ont été consacrés à la détection de tendance stochastique dans des séries de données. Il va de soi que ce type de tendance n'intéresse pas trop les hydrométéorologues car, bien souvent, l'autocorrélation que l'on calcule dans les séries hydrométéorologiques dépasse rarement la valeur 0.5. Malgré ce fait, il existe quand même quelques exceptions notables. Galbraith et Green (1990) et David et Robert (1997) ont utilisé les tests de Dickey et Fuller (1979) pour détecter ce type de non-stationnarité dans les séries de données de température. Dans la suite de ce document (section 5.3), on fournira une littérature plus appropriée pour étudier ce type de tendance qui a été proposé dans des disciplines comme l'économétrie.

5.2.3 Développements par rapport à certaines caractéristiques des séries de données hydrométéorologiques

Comme on l'a déjà mentionné à plusieurs reprises dans ce travail, les séries de variables hydrométéorologiques ne sont pas des suites de données indépendantes. Elles sont liées à cause de la présence, dans ces séries, de la persistance (la dépendance sérielle) et de la composante saisonnière. Ces deux caractéristiques peuvent sérieusement déformer les résultats des tests présentés ci-dessus.

5.2.3.1 Composante saisonnière

La variation saisonnière peut grandement compliquer la détection de tendance à long terme dans une série chronologique. Dans la littérature, trois procédures générales sont utilisées pour prendre en compte la présence de la composante saisonnière dans les tests de détection de tendance. La première utilise les fonctions périodiques pour décrire les variations saisonnières contenues dans la série. Cette approche est facile à appliquer en conjonction avec une analyse de régression pour tester la tendance linéaire ou la tendance en saut (ruptures). Le cas le plus simple fait référence au modèle de régression suivant :

$$x_t = f(\omega t, A) + \beta(\text{terme de tendance}) + \varepsilon_t \qquad (5.11)$$

avec, $f(\omega t, A) = A_0 + A_1 \sin(\omega t) + A_2 \cos(\omega t) + \ldots$

On ajoute autant de termes qu'il en faut pour caractériser la composante saisonnière présente dans la série et on teste la signification du terme de tendance à l'aide de la statistique de Fisher. Une référence complète est fournie par Hipel et McLeod (1994).

La deuxième approche traitant de la composante saisonnière utilise un test global. Pour évaluer la valeur observée de ce test global, on se sert d'un test classique de détection de tendance. Le principe est le suivant. On commence par diviser la série originale en sous-séries correspondant aux saisons. Puis, on calcule, pour chacune des sous-séries, la

statistique observée du test classique de détection de tendance. Par la suite, on fait la somme globale de ces statistiques observées. La somme ainsi calculée constitue la valeur observée de la statistique de test global. Dans la littérature, plusieurs tests ont été construits suivant cette approche. Il s'agit principalement du test de Hirsch et Gilroy (1985) pour la tendance en saut, du test saisonnier de Mann-Kendall (Hirsch et al. 1982), du test de Hirsch et Slack (1984), du test de Zetterqvist (1991)) et du test de Farell (1980) pour la tendance linéaire. Le test de Hirsch et Gilroy est une modification du test des rangs de Mann-Whitney ("Mann-whitney-rank-sum-test"). Farell (1980) s'est inspiré des travaux de Sen (1968) pour développer un test saisonnier de détection de tendance linéaire. Ce test est construit à partir des rangs des observations et il détecte la présence d'une tendance à l'intérieur d'une série désaisonnalisée. Van Bell et Hughes (1984) ont montré dans leurs travaux que le test de Farell est plus puissant dans la situation où les échantillons analysées ne possèdent pas de valeurs manquantes. Dans la situation contraire, Taylor et Loftis (1989) montrent que le test de Mann-Kendall est plus puissant. Il est important de signaler que les tests décrits ci-dessus ne parviennent pas à révéler les différences de comportement dans les différentes saisons considérées. En effet, il peut arriver qu'une saison donnée exhibe une longue tendance et que les autres n'en présentent aucune. Van Bell et Hughes (1984) ont proposé un test permettant de vérifier si toutes les saisons considérées se comportent de la même manière. La troisième et dernière approche est basée sur le modèle de régression de Wonnacott et Wonnacott (1980). Dans cette approche, on modélise les composantes saisonnières présentes dans la série à l'aide des variables dichotomiques.

5.2.3.2 Persistance

Les tests saisonniers présentés ci-dessus éliminent seulement l'effet de dépendance entre saisons, ils ne prennent pas en compte la corrélation dans la série à l'intérieur des saisons. Cela constitue une limitation, relativement aux tests présentés dans cette revue bibliographique. En effet, ces derniers supposent l'indépendance des observations. Dans la littérature, plusieurs méthodes traitent le problème des tests de tendance en présence de dépendance sérielle. Parmi ceux-ci, Hirsch et Slack (1984) ont proposé une modification du test de Mann-Kendall. Plus récemment, Khaled et al. (1998) ont suggéré une autre

version du test de Mann-Kendall pour tenir compte de l'autocorrélation des données dans la détection de tendance linéaire. Dans la famille des tendances en saut, Lettenmaier (1976) a proposé d'appliquer les principaux tests échantillonnés en remplaçant dans les formules appropriées la taille de l'échantillon de données (n) par le nombre d'observations indépendantes de la série. Pour les tests utilisant les modèles de régression, Box et Jenkins (1976) ont proposé les modèles ARIMA.

5.2.3.3 Composante saisonnière et persistance

Lorsqu'une série chronologique de variable hydrométéorologiques est caractérisée par la composante saisonnière et la persistance, Hirsch et Slack (1984) ont proposé un test non-paramétrique pour les deux types de tendances analysés.

5.3 Synthèse bibliographique des tests de détection de rupture appliqués en statistique et en économétrie

Une composante importante dans l'analyse des données temporelles concerne les tests de rupture. Cette question a depuis longtemps été reconnue dans la littérature statistique et économétrique. Dans cette section on ne prétend pas fournir un recueil exhaustif des divers tests de détection de non-stationnarité de type rupture qui ont fait l'objet de publications dans les littératures statistique et économétrique. Étant entendu que ce type de non-stationnarité est au centre de notre travail, on veut simplement exposer sommairement certaines avancées traitant de ce sujet dans ces deux littératures. Une telle approche peut s'avérer très utile. En effet, la détection de rupture est un problème important en économétrie et en statistique. On pourrait donc utiliser les travaux de ces deux domaines avec profit en hydrométéorologie.

5.3.1 Problème de rupture et tests de stationnarité en statistique et en économétrie

La théorie fondamentale de l'inférence statistique en économétrie et en statistique repose sur l'hypothèse que les données utilisées sont stationnaires. Bien souvent, bon nombre des séries chronologiques utilisées en statistique et en économétrie affichent une tendance à croître dans le temps. C'est le cas par exemple des séries de consommation, de stock monétaire, de PIB, etc. Ces séries ne satisfont donc pas la condition d'avoir une moyenne (ou variance) constante. Elles sont donc non-stationnaires. Cette constatation a donné naissance à un nombre important de travaux, tant théoriques qu'empiriques, sur les problèmes de non-stationnarité dans la représentation des séries chronologiques statistiques et économétriques (séries macroéconomiques).

Pour les statisticiens, il existe deux sortes de non-stationnarité : la non-stationnarité de type tendance déterministe et la non-stationnarité de type tendance stochastique. Il s'ensuit que les tests de rupture peuvent être vus comme des tests de détection de non-stationnarité de type tendance déterministe. Ces tests servent donc à discriminer entre séries stationnaires et séries non-stationnaires (de type tendance déterministe).

Du point de vue des économètres il y a une distinction fondamentale à faire entre une non-stationnarité de type tendance déterministe et une non-stationnarité de type tendance stochastique. Rappelons qu'une série chronologique non-stationnaire de type tendance déterministe est donnée à l'équation 3.11 tandis que une série chronologique non-stationnaire de type tendance stochastique est illustrée à l'équation 3.14. On peut voir que ces deux séries ont des comportements aléatoires radicalement différents. En effet, suite à un choc stochastique ε_t, donné au temps t_0, la série x_t dans l'équation 3.11 reviendra vers sa tendance à long terme $(\alpha + \beta t)$ d'après l'équation 3.13. C'est donc dire que lorsqu'un processus générateur de la série x_t (dans l'équation 3.11) est affecté par un choc stochastique, l'effet de ce choc tend à disparaître au fur et à mesure que le temps passe : c'est la propriété de non persistance des chocs. Dans le langage des économètres, cela signifie que la trajectoire de long terme de la série x_t est insensible aux aléas

conjoncturels. Par contre, un choc ε_t au temps t aura un effet permanent sur le niveau de la série x_t dans l'équation 3.14. C'est-à-dire, d'après l'équation 3.17, la non-stationnarité du processus générateur de la série x_t (dans l'équation 3.17) tient au fait que les chocs ε_t s'accumulent au cours du temps, ce qui accroît la variance de x_t au fur et à mesure que le temps passe. L'origine de la non-stationnarité provient donc ici de l'accumulation de chocs stochastiques ε_t. Ainsi, pour les processus générateurs de séries non-stationnaires de type tendance stochastique il existe une propriété de persistance des chocs qui n'existe pas dans les processus générateurs de séries non-stationnaires de type tendance déterministe. C'est cette constatation qui a amené Nelson et Plosser (1982), dans leur papier célèbre, à soutenir que la plupart des séries macroéconomiques contiennent des tendances stochastiques. En partie pour cette raison, les tests qui ont été développés en économétrie pour discriminer entre séries stationnaires et séries non-stationnaires prennent comme hypothèse principale la présence d'une racine unitaire (équations 3.14 et 3.15). Dans la version la plus générale de ces tests, on admet la possibilité de la présence simultanée d'une tendance déterministe et d'une tendance stochastique (équation 3.18). Fuller (1976), Dickey et Fuller (1979, 1981) sont considérés comme les pionniers de la littérature économétrique des tests de stationnarité. D'ailleurs, c'est à partir des tests de Dickey et Fuller (1979) que Nelson et Plosser (1982) ont démontré que la plupart des séries macroéconomiques étaient probablement caractérisées par une racine unitaire dans leur représentation autorégressive.

Intuitivement, la notion de rupture prend une autre dimension en économétrie. Pour les économètres, l'effet d'une rupture en moyenne (changement en moyenne) ressemble beaucoup à un choc qui a des effets permanents. Dans ce sens, la rupture de moyenne à l'intérieur de l'échantillon est perçue comme un phénomène exogène et non pas comme une réalisation du processus générateur sous-jacent. Ainsi, détecter une rupture dans la série ne revient plus à vérifier la présence d'une non-stationnarité dans celle-ci. On est plutôt intéressé à savoir s'il est possible qu'une série chronologique donnée soit stationnaire à part une rupture en moyenne. Si c'est effectivement le cas, la question subsidiaire est de savoir comment ceci devrait affecter la méthodologie statistique pour discriminer entre séries stationnaires et séries non-stationnaires ? Cela dit, dans la

littérature économétrique, le problème de rupture a reçu beaucoup d'attention en relation avec le débat portant sur la question de racine unitaire versus le changement dans la fonction de tendance pour une série chronologique univariée.

5.3.2 Revue bibliographique des tests de rupture en statistique

Le développement des techniques de détection de rupture en statistique a été très longtemps tributaire des travaux réalisés dans le domaine de la qualité du produit industriel. En effet, l'un des premiers problèmes de rupture étudié concernait le contrôle d'une chaîne de production, où il est en effet essentiel de détecter une baisse du régime de production pour pouvoir, par exemple, remplacer une machine défectueuse. La rupture est ici synonyme de baisse notable de la qualité du produit industriel concerné. Les travaux pionniers dans la littérature statistique sont ceux réalisés par Page (1954, 1955 et 1957) qui s'intéressait aux ruptures dans la qualité de production d'une machine industrielle. Pour détecter un changement de moyenne dans une série de données, Page a proposé deux statistiques de test. La première est basée sur les sommes cumulatives et sert à détecter un changement de moyenne dans un échantillon suivant une distribution binomiale (Page, 1954 ; 1955). La deuxième est construite à partir de la méthode du maximum de vraisemblance. Elle détecte un changement dans un échantillon gaussien (Page, 1957). Ces premiers travaux supposent que la date de rupture est connue a priori.

La littérature statistique qui s'est accumulée depuis les travaux de Page (1954, 1955 et 1957) a vu l'apparition de nouvelles techniques de détection de rupture. Dans le cas des échantillons gaussiens ou éventuellement pour des distributions exponentielles ainsi que pour des distributions binomiales, les premiers tests sont, par construction, utilisables lorsqu'on a pu déterminer la date de rupture. Pour l'échantillon gaussien, Chernoff et Zacks (1964), Gardner (1969), Smith (1975) et Hsu (1979) ont proposé des statistiques de test basées sur la moyenne initiale. Hinkley (1970), Siegmund (1986) et Ghorbanzadeh (1995) ont mis en avant des tests de type maximum de vraisemblance. Tandis que, Bhattacharya et Johnson (1968), Sen et Srisvastava (1975), Bhattacharya et Fierson (1981), Csörgö et Carlstein (1988) et Ferger et Stute (1992) ont proposé des tests non paramétriques basés sur

les rangs et la fonction de répartition empirique. Finalement, Worsley (1983, 1986) a proposé deux tests basés sur la moyenne initiale. Le premier concerne l'échantillon exponentiel et le second traite de l'échantillon binomial. Kim (1996) et Bhatti et Wang (1998) donnent en référence une bibliographie détaillée de ces travaux.

D'autres techniques de détection de rupture ont aussi été envisagées dans le cadre de la régression linéaire. Le plus souvent, ces techniques supposent que la date de rupture correspond à l'un des instants d'observation et tentent de répondre à la question : y a-t-il une rupture dans le modèle de régression considéré ou non ? Les travaux qui se sont intéressés à la régression linéaire simple ont traité le cas d'une rupture moins brutale. Le modèle de régression considéré par Quandt (1958), Chow (1960), Freeman (1986) et Kim et Siegmund suppose que la variable explicative est le temps. Par contre, dans le modèle de Maronna et Yohai (1978) la variable explicative n'est pas forcément le temps. En fin de compte, on retrouve aussi dans cette littérature quelques travaux qui se placent dans le cadre de la régression linéaire multiple : Brown, Durbin et Evans (1975). Bhattacharya (1994) présentent en référence un survol de ces différents travaux.

D'autres auteurs ont abordé le problème de rupture sous une forme qui diffère des deux premières. Leur démarche se place dans le cadre de la sélection de modèle. Le principe est le suivant. L'instabilité d'un processus peut revêtir plusieurs formes, le plus souvent on suppose que la série (ou encore le déroulement du processus) répond à différents modèles dans divers intervalles distincts. On dit alors qu'il y a un changement de structure du processus et le problème revient à identifier ces intervalles, ce qui équivaut encore à estimer les dates de rupture. Une fois supposé a priori que chaque observation de l'échantillon original peut appartenir à n'importe quel sous modèle d'une liste donnée, le problème de détection de ruptures revient alors à un problème statistique de choix de modèles. On peut donc recourir à l'ajustement de modèles, à des tests de la valeur des paramètres du modèle, etc. Parmi les travaux qui se sont intéressés à cette approche, soulignons ceux réalisés par Cox et Spjotvøll (1982). Ces auteurs ont proposé l'utilisation des méthodes telles que la *classification automatique*.

5.3.3 Revue bibliographique des tests de rupture en économétrie

Le problème des tests de rupture, encore appelé changement structurel, a reçu autant (ou même plus) d'attention dans la littérature économétrique que dans la littérature hydrométéorologique. Lorsque la date de changement ou sa forme est mal connue, les procédures de type prédictif s'avèrent à la fois faciles d'application et souvent très réalistes (Dufour, 1980 ; Dufour, 1981 ; Dufour, 1982c ; Dufour, 1982a ; Dufour, 1982b ; Dufour et al., 1994 e ; Dufour et Kiviet, 1996). D'autres procédures de type CUSUM ont également été proposées dans cette littérature (Dufour, 1982c ; Dufour, 1986 ; Dufour, 1989 ; Dufour et Kiviet, 1996 ; Hawkins and Olwell, 1998). Dans le cadre des modèles de régressions linéaires, les contributions récentes (et une revue de littérature) sont présentées par Dufour et Ghysels (1996), Perron (1997) et Bai et Perron (1998). Les principales difficultés que cette littérature a adressées concernent : (1) le cas où la(les) date(s) de rupture est (sont) endogène(s), et (2) les dates de rupture multiples. En effet, bien que la littérature économétrique soit vaste, les études qui se sont directement intéressées au problème relatif aux changements structurels multiples sont peu nombreuses. Notons, en particuliers, les travaux de Christiano (1992), Andrews (1993), Bai et Perron (1998) et de Banerjee et al. (1998). En effet, la majorité des tests couramment utilisés est fondée sur des approximations asymptotiques dont la fiabilité sur des échantillons finis n'est pas garantie (voir Diebold et Chen (1996)). Les problèmes qui en découlent dans les applications empiriques incluent : absence de bonnes valeurs critiques, problème de contrôle de niveaux et possibilité de détecter des dates de ruptures erronées. Ces inconvénients ont contribué à l'émergence d'une génération de tests basés sur des simulations de type Bootstrap par exemple (Christiano (1992), Diebold et Chen (1996) et Banerjee et al. (1998)). Pourtant, les méthodes proposées par ces derniers ne réalisent pas le contrôle de niveau.

5.4 Conclusion

La question de la présence ou de l'absence de la stationnarité des séries de données d'une variable hydrométéorologique continuera de préoccuper les ingénieurs hydrologues. Dans

ce chapitre on a fait un survol rapide des tests statistiques de détection de non-stationnarité dans une série de données. Ce survol devrait permettre de mieux comprendre la littérature hydrométéorologique sur le sujet et les conditions entourant l'applicabilité de ces tests. Bien que des progrès importants aient été faits dans la littérature hydrométéorologique, il n'en reste pas moins qu'il y a encore beaucoup de travail à faire pour mieux adapter ces tests aux réalités des séries hydrométéorologiques, tant sur le plan théorique que sur leur applicabilité.

Sur le plan de l'applicabilité, il apparaît important d'avoir les performances des tests présentés dans cette littérature. En effet, Lubès et al. (1998) ont montré que, pour les procédures largement employées dans les études de rupture en hydrométéorologie (le test de Pettitt, le test de Buishand, la procédure bayésienne de Lee et Heghinian, et la procédure de segmentation de Hubert et Carbonel), la persistance (l'autocorrélation) et la présence d'une tendance dans les séries sont les deux caractéristiques qui pénalisent le plus les performances des procédures. Des études de simulations seraient utiles à cet égard.

Sur le plan théorique, les points suivants méritent une investigation plus détaillée. Premièrement, il apparaît nécessaire d'augmenter la classe des tests considérés jusqu'à maintenant pour détecter une rupture dans une série de données. Dans la situation la plus vraisemblable où la date de rupture est inconnue, on devrait proposer de nouveaux tests qui prennent en compte les caractéristiques des séries de données d'une variable hydrométéorologique : la persistance et la présence d'une tendance. Deuxièmement, il apparaît important d'avoir des tests qui se préoccupent de la détection de plusieurs ruptures dans une série de données. Finalement, il serait intéressant d'avoir un test régional de détection de rupture (s). En effet, les tests statistiques de détection de non-stationnarité présentés dans ce survol concernent l'exploitation d'une série de données et une seule. Ils sont donc utiles pour une analyse ponctuelle (ou local) de la stationnarité d'un phénomène hydrométéorologique. Or, une non-stationnarité d'origine climatique, si elle existe, est un phénomène global qui se manifeste à une échelle au moins régionale. Il est donc naturel de souhaiter disposer d'outils nécessaires pour vérifier sa présence.

En somme, pour le type de non-stationnarité considéré dans cette thèse, on a présenté dans ce chapitre une courte revue bibliographique des tests de rupture en statistique et en économétrie. On a insisté sur ce sujet car compte tenu de l'importance de ce sujet en statistique et en économétrie, les travaux réalisés dans ces domaines pourraient être appliqués en hydrométéorologie avec profit.

6 THÉORIE DES TESTS DE MONTE CARLO

Il est généralement admis que le point de départ de la théorie des tests de Monte Carlo est le papier de Dwass (1957). Les idées présentées par Dwass ont été reprises quelques années plus tard dans les travaux de Barnard (1963) et Birnbaum (1974) pour constituer la forme actuelle des tests de Monte Carlo. Ce chapitre sert d'introduction aux tests de Monte Carlo qui seront utilisés ultérieurement dans l'originalité de cette thèse en matière de détection de non-stationnarité dans des séries de données d'une variable hydrométéorologique. Après quelques rappels de notions élémentaires sur les tests d'hypothèses, on introduit un certain nombre de généralités associées aux techniques des tests de Monte Carlo. Le présent chapitre constitue donc la pierre angulaire de ce document.

6.1 Rappels de quelques généralités sur les tests statistiques d'hypothèses

6.1.1 Problème posé

Dans les sciences de l'eau, l'ingénieur hydrologue est appelé à prendre des décisions concernant les populations de phénomènes hydrométéorologiques sur la base d'informations recueillies sur des échantillons (déroulements de tels phénomènes). De telles décisions sont appelées *décisions statistiques*. Il en est ainsi lorsqu'un ingénieur hydrologue, par exemple, souhaite décider sur la base de données recueillies sur le processus de débits en rivière, que les modifications climatiques ont eu un impact sur la disponibilité des ressources en eau ou encore que le régime des débits a varié.

Pour prendre des décisions de ce genre, l'ingénieur hydrologue va faire des hypothèses sur les populations concernées. De telles hypothèses sont en général des énoncés relatifs aux distributions de probabilité de ces populations. Ainsi, si sur la base de la supposition qu'une hypothèse donnée est vraie, l'ingénieur hydrologue trouve, à partir des réalisations du processus hydrométéorologique étudié, que les résultats observés diffèrent de ceux qu'il

pourrait espérer, il devra considérer ces différences comme *significatives*. Il devra donc être enclin à rejeter l'hypothèse proposée, ou tout au moins à ne pas l'accepter compte tenu des observations qu'il a obtenues dans la mesure de ce processus. De manière générale, les méthodes qui permettent de décider d'accepter ou de rejeter des hypothèses ou de déterminer si les échantillons observés diffèrent significativement des résultats attendus, sont appelées *tests d'hypothèses*, *tests de signification* ou *règles de décisions*.

Lorsqu'on effectue un test statistique, on se trouve devant une *hypothèse principale* H_0 qui représente une théorie qui a été proposée soit parce qu'on pense qu'elle est vraie, soit parce qu'elle sera utilisée comme base pour l'argument mais n'a pas été prouvée. Par exemple, si l'ingénieur hydrologue souhaite décider que le régime des débits d'une rivière donnée a varié, il va formuler l'hypothèse qu'aucune variation ne soit intervenue dans le régime des débits (hypothèse de stationnarité). Un test pour H_0 est alors une règle par laquelle on cherche à prendre une décision (admettre que H_0 est réalisée ou non). Pour montrer que H_0 n'est pas réalisée, on doit montrer que toutes les distributions compatibles avec H_0 sont inacceptables. Il existe plusieurs façons d'élaborer un test statistique relatif à une hypothèse principale H_0. En ayant une vision plus fréquentiste, on retiendra deux d'entre elles : le test statistique selon l'approche de Fisher et le test statistique selon la théorie de Neyman et Pearson. Ces deux approches sont résumées à l'appendice C. Dans la suite de ce document, on adopte l'approche de Neyman et Pearson qui est principalement employée dans les tests statistiques de détection de non-stationnarité en hydrométéorologie.

6.1.2 Notions essentielles de la théorie classique des tests de Neyman et Pearson

Pour Neyman et Pearson on doit introduire deux hypothèses dans l'élaboration d'un test, l'hypothèse principale H_0 et la contre-hypothèse H_1. On se trouve donc devant une alternative, c'est-à-dire deux hypothèses entre lesquelles on cherche à prendre une décision (admettre l'une des hypothèses comme réalisée). De façon générale, si l'on suppose le modèle statistique $(\chi, \mathcal{A}, \wp)$ où χ est l'espace d'échantillonnage, $\mathcal{A}$ est une tribu des

parties de χ et $\wp$ une famille de probabilités sur l'espace mesurable $(\chi, \mathcal{A})$ tel que : ($\exists\, p \in \mathbb{N}$) : $\wp = \{P_\theta : \theta \in \Theta \subset \mathbb{R}^p\}$, bâtir un test qui est défini par $H_0 \subset \Theta_0$ et $H_1 \subset \Theta_1 = \complement_{\Theta_0}$, avec $\Theta = \Theta_0 \cup \Theta_1$, consiste à définir une règle de décision au vu d'un échantillon $\{X_1, \ldots, X_n\}$. Une telle règle peut être représentée au moyen d'une fonction indicatrice $\phi(X_1, \ldots, X_n)$ qui prend la valeur 0 ou 1 :

$$\phi(X_1, \ldots, X_n) = \begin{cases} 1 & \text{si } H_0 \text{ est rejetée} \\ 0 & \text{si } H_0 \text{ est acceptée} \end{cases} \tag{6.1}$$

Habituellement, la fonction indicatrice $\phi(X_1, \ldots, X_n)$ est définie de la manière suivante :

$$\phi(X_1, \ldots, X_n) = \begin{cases} 1 & \text{si } \varphi(X_1, \ldots, X_n) > k_\alpha \\ 0 & \text{si } \varphi(X_1, \ldots, X_n) \leq k_\alpha \end{cases} \tag{6.2}$$

où, $T = \varphi(X_1, \ldots, X_n)$ est la statistique du test et k_α est le point critique.

Si la règle de décision que l'on vient de définir permet de prendre la décision au vu de l'échantillon, il est clair que cette décision peut correspondre ou non à la réalité. Et, comme on n'est pas en mesure de savoir dans quelle situation on se trouve, une fois la décision prise, on peut donc définir des risques d'erreur.

Définition 6.1. *Le risque de première espèce, noté α, est de rejeter à tort H_0 alors que H_0 est vraie. Sa probabilité s'écrit :*
$\alpha = P(\text{commettre une erreur de 1ère espèce}) = P(\text{rejeter } H_0 \text{ alors que } H_0 \text{ est vraie})$

Définition 6.2. *Le risque de deuxième espèce, noté β, est de décider à tort que H_0 est vraie. Sa probabilité s'écrit :*
$\beta = P(\text{commettre une erreur de 2ème espèce}) = P(\text{ne pas rejeter } H_0 \text{ alors que } H_0 \text{ est fausse})$

Pour qu'un test d'hypothèse soit convenable, il est nécessaire qu'il soit conçu de manière à minimiser les erreurs de décision. Ce n'est pas simple, dans la mesure où, pour un échantillon de taille donnée, toute initiative de nature à diminuer un type d'erreur

s'accompagne généralement d'une augmentation d'une erreur de l'autre type. Selon l'approche mise en avant par Neyman et Pearson, on doit traiter les deux risques d'erreur de façon non symétrique et limiter l'ensemble des tests possibles à la classe de tests ayant un risque de première espèce au plus égal à un niveau de signification α fixé au préalable. La procédure revient alors à rechercher un test optimal dans cette classe, c'est-à-dire un test dont le risque de seconde espèce, β, soit minimum (ou encore que la puissance $(1-\beta)$ soit maximale).

Étant donné le test défini précédemment, la fonction de risque d'erreur de première espèce est la fonction définie sur H_0 par :

$$\alpha(\theta) = P\{\varphi(X_1,\ldots,X_n) > k_\alpha \,/\, \theta \in \Theta_0\} \tag{6.3}$$

Ainsi, en suivant la logique de Neyman et Pearson, un *test* $\alpha\,\%$ implique

$$\alpha(\theta) \leq \alpha, \quad \forall \theta \in \Theta_0 \tag{6.4}$$

Par définition, un test $\phi(X_1,\ldots,X_n)$ est de *niveau* α si

$$\sup_{\theta\in\Theta_0}\{P(\varphi(X_1,\ldots,X_n) > k_\alpha)\} = \sup_{\theta\in\Theta_0}\{\alpha(\theta)\} \leq \alpha \tag{6.5}$$

Dans ce cas, on désigne par la *taille du test*, la grandeur $\sup_{\theta\in\Theta_0}\{\alpha(\theta)\}$ (que l'on note encore $\sup_{\theta\in\Theta_0}\{P(\varphi(X_1,\ldots,X_n) > k_\alpha)\}$). C'est la plus grande probabilité de rejeter H_0 avec raison. Ainsi, on dit d'un test $\phi(X_1,\ldots,X_n)$ qu'il a une taille α si et seulement si :

$$\sup_{\theta\in\Theta_0}\{P(\varphi(X_1,\ldots,X_n) > k_\alpha)\} = \sup_{\theta\in\Theta_0}\{\alpha(\theta)\} = \alpha \tag{6.6}$$

Ce qui caractérise l'approche de Neyman et Pearson, c'est que un test d'hypothèses n'est valide que si sa taille n'est pas plus grande que le niveau de signification α, et ce pour

toutes les valeurs possibles de α (Edgington, 1995). Dans ce cadre, il est raisonnable de penser que le concept d'un test plus puissant devient désuet. On est plutôt conduit à parler de tests $\alpha\%$ plus puissant. Dès lors, la recherche du test le plus puissant relatif à H_0 conduit tout d'abord à se placer dans la classe des tests ayant un niveau commun de probabilité d'erreur α. À partir de là, on dit alors qu'un test qui satisfait les conditions :

$$P_\theta\{\varphi(X_1,\ldots,X_n) > k_\alpha\} \leq \alpha, \qquad \forall \theta \in \Theta_0$$

$$P_\theta\{\varphi(X_1,\ldots,X_n) > k_\alpha\} \text{ est maximum pour tout } \theta \notin \Theta_0$$

est *uniformément plus puissant (UPP)* de niveau de signification α. En d'autres termes, sa fonction puissance est supérieure en tout point de H_1 à la fonction puissance de tout test de taille inférieure ou égale à α. Si par contre H_1 est une hypothèse simple, le test est dit le plus puissant de niveau de signification α. En pratique, un test uniformément plus puissant est difficile à obtenir. On se restreint habituellement à d'autres types de tests aux propriétés moins contraignantes mais qui sont d'un grand intérêt pratique :

- On dit d'un test qu'il est **sans biais de niveau** α si sa probabilité de rejeter H_0 à tort est majorée par α et celle de rejeter à juste titre est minorée par α :

  $$P_\theta\{\varphi(X_1,\ldots,X_n) \geq k_\alpha\} \leq \alpha, \qquad \forall \theta \in \Theta_0$$

  $$P_\theta\{\varphi(X_1,\ldots,X_n) \geq k_\alpha\} \geq \alpha, \qquad \forall \theta \in \Theta_1$$

- On dit d'un test qu'il est **convergent** si sa fonction de puissance tend vers 1 quand la taille de l'échantillon devient de plus en plus grande :

  $$P_\theta(\varphi(X_1,\ldots,X_n) > k_\alpha) \underset{\text{échantillon}\to\infty}{\to} 1 \qquad \forall \theta \in \Theta_1$$

- On dit d'un test qu'il est α**-semblable** si on a la propriété suivante :

  $$P_\theta(\varphi(X_1,\ldots,X_n) > k_\alpha) = \alpha, \qquad \forall \theta \in \Theta_0$$

En définitive, chaque procédé de test d'hypothèses, suivant l'approche de Neyman et Pearson, appelle plusieurs remarques. Premièrement, bien que la théorie classique des tests demande avant tout un contrôle de la taille du test (équation 6.6), dans de nombreuses situations pratiques, il est en fait impossible de déterminer une région critique de rejet de

telle sorte que la taille soit exactement α. On se contente donc de demander le contrôle du niveau du test (équation 6.5) néanmoins le contrôle de la taille (équation 6.6), si possible, est préférable. Deuxièmement, la majorité des tests reposent sur le respect d'un certain nombre de conditions (par exemple la référence aux lois de distribution de probabilité des données). Selon le degré de respect de ces conditions d'application, la validité des résultats se trouve plus ou moins affectée et elle l'est d'autant plus que le test est moins *robuste* (moins sensible au non-respect des conditions d'application). Finalement, la détermination du point critique k_α d'après l'équation 6.5 fait intervenir deux niveaux de complication. La première complication concerne la taille d'échantillon et les lois de probabilité pour le paramètre θ sous H_0 qui sont explicitement prises en considération dans l'expression de la taille du test. En effet, dans beaucoup de cas pratiques de telles lois de probabilité sont difficiles voire impossible à obtenir (par exemple elles pourraient dépendre de paramètres de nuisance). Le deuxième niveau de complication concerne le problème d'optimisation car, dans la pratique, la recherche d'un supremum est loin d'être un problème évident.

6.2 Principe des tests de Monte Carlo

6.2.1 Généralités

Lorsqu'on procède à un test statistique, la question fondamentale qu'on se pose toujours est la suivante : la valeur observée de la statistique de test, $\varphi(x_1,\ldots,x_n)$, est-elle "grande" ou "petite" ? On ne peut pas répondre à cette question dans l'absolu. En revanche, on ne peut y répondre que par comparaison avec les valeurs que la statistique de test, $\varphi(X_1,\ldots,X_n)$, pourrait prendre dans différentes autres expériences où l'hypothèse principale H_0 est vraie. Dès lors, on peut donc se demander : quelles valeurs pourraient prendre la statistique de test $\varphi(X_1,\ldots,X_n)$ si H_0 est vraie ?

Il existe deux façons de répondre à cette dernière question. La première qui est la méthode classique consiste à employer, sous certaines conditions, une loi de distribution théorique pour décrire la statistique de test sous H_0. La seconde est la méthode par permutations qui consiste à générer une loi de distribution empirique de la statistique de test sous H_0. Dans tous les cas, qu'on utilise l'une ou l'autre méthode, la décision statistique reste la même : on compare la valeur observée de la statistique de test à la distribution des valeurs que l'on pourrait obtenir sous H_0 ; ces valeurs sont fournies, dans le premier cas, par les tables de la loi théorique retenue et, dans le second cas, par la distribution d'échantillonnage obtenue sous H_0 en permutant aléatoirement les données.

L'inconvénient de l'approche classique est qu'elle est conditionnée par des conditions d'application particulières à chaque test. En effet, ce n'est que sous ces conditions que l'on peut admettre que la statistique de test obéit à une certaine loi de distribution particulière lorsque H_0 est vraie. Ces conditions ne sont donc pas inhérentes aux tests statistiques eux-mêmes. Pire encore, si les statistiques de test proposées contiennent des paramètres de nuisance, on peut s'attendre à ce que le test statistique ait des problèmes de niveau si l'on travaille sous l'hypothèse asymptotique. Dans pareille situation, les résultats que l'on obtient ne sont exacts qu'en présence d'échantillons de taille infinie, difficiles à obtenir dans notre univers fini. En fait, la théorie asymptotique n'est pas fiable dans des échantillons finis parce qu'elle est fondée sur des conditions de régularité que l'on ne peut tester.

La seconde approche, à la différence de la précédente, résolument théorique, est plutôt expérimentale. Elle se base sur une procédure qui permet de comparer la statistique de test observée avec une distribution qui est générée par permutations aléatoires des données. L'argument sous-jacent étant que si l'hypothèse principale H_0 est vraie, alors toutes les permutations possibles des données devraient pouvoir avoir lieu de manière égale. Une telle approche est donc intéressante lorsque l'échantillon est très restreint et qu'il n'est pas possible d'en obtenir un plus grand. On peut alors passer outre aux conditions qui concernent la distribution des données et les cas litigieux de paramètres de nuisance de

l'approche classique, en utilisant la méthode par permutation. Cette approche considérée jusqu'à récemment comme coûteuse en temps de calcul est désormais à la portée d'un grand nombre de personnes grâce à la puissance des ordinateurs actuels.

6.2.2 Principes de base de la technique des tests de Monte Carlo

Le tournant le plus important dans la construction des tests de Monte Carlo doit sans aucun doute se situer en 1957 lorsque Dwass a proposé une procédure de test randomisé exacte basée sur le principe des tests par permutations. Cette procédure a été généralisée par la suite par Barnard (1963) et Birnbaum (1974). On peut donc considérer à juste titre que ces auteurs sont les pionniers du développement de la technique des tests de Monte Carlo. Avant d'aborder la technique des tests de Monte Carlo proprement dite, on tient à dresser un bref aperçu historique des événements majeurs qui ont amené la construction de ces tests à leur état actuel. Le fil conducteur de cet historique sera basé sur la résolution du problème suivant.

> *On considère deux échantillons théoriques indépendants* $X_1,\ldots,X_n$ *et* $Y_1,\ldots,Y_m$, *où les* X_i *suivent une loi de distribution continue* $F(x)$ *tandis que les* Y_j *suivent une loi de distribution continue* $F(x-\delta)$; *en imposant aucune autre hypothèse distributionnelle sur les données, on souhaite tester l'hypothèse principale* $\mathrm{H}_0 : \delta = 0$. *Cela revient encore à tester que les deux échantillons théoriques viennent de la même population.*

6.2.2.1 Procédure d'un test par permutations

Pour tester l'hypothèse principale H_0 du problème précédent, la procédure classique des tests par permutations procède comme suit :

- on considère deux échantillons observés indépendants issus des deux échantillons théoriques, soit $x_1,\ldots,x_n$ et $y_1,\ldots,y_m$;
- on pose $z=(x_1,\ldots,x_n,y_1,\ldots,y_m)$ et on calcule la valeur observée du test, soit

$$s=\frac{1}{n}\sum_{i=1}^{n}x_i-\frac{1}{m}\sum_{i=1}^{m}y_i$$

- en supposant que les X_i et les Y_j sont *échangeables* (définition 6.4), on obtient toutes les $Q=(n+m)!$ permutations possibles de $z, z^{(1)},\ldots,z^{(Q)}$ et on calcule les valeurs analogues de la statistique de test S :

$$s^{(j)}=\frac{1}{n}\sum_{i=1}^{n} z_i^j - \frac{1}{m}\sum_{i=n+1}^{m+n} z_i^j, \qquad j=1,\ldots,Q$$

- soit r le nombre de $s^{(j)}$ pour lesquels $s \le s^{(j)}$. On rejette H_0 (par exemple contre $H_1 : \delta > 0$), si $r \le k$, où k est un entier prédéterminé.

L'avantage de ce test par permutations est qu'il est de taille k/Q. En effet, il est facile de voir que $P(r \le k) = k/Q$ sous H_0, car les X_i et les Y_j sont échangeables.

La méthode des tests par permutations avait été proposée au tout début par R. A. Fisher dès 1935, mais les statisticiens de l'époque ne disposaient pas d'ordinateurs pour réaliser les milliers de calculs nécessaires. En effet, dans l'exemple précédent, cette procédure intuitive demande de travailler avec $Q=(n+m)!$ permutations, ce qui peut être très fastidieux si l'on travaille avec des tailles d'échantillons relativement grandes. C'est le principal désavantage de cette méthode qui a conduit les statisticiens du début du siècle à développer les différentes distributions de probabilité qui servent plus couramment à réaliser les tests statistiques.

6.2.2.2 Procédure d'un test randomisé exact : Dwass (1957)

Pour contourner la difficulté de calcul d'un test par permutations, Dwass (1957) a proposé d'adapter cette procédure de test sur un échantillon de P permutations $\tilde{s}^{(1)},\ldots,\tilde{s}^{(P)}$ relativement bien choisies de sorte que la taille du test soit préservée. La procédure résultante est la suivante :

- fixer le nombre de permutations P à exécuter ;
- calculer $\tilde{s}^{(1)},\ldots,\tilde{s}^{(P)}$;

- soit $\tilde{r}$ le nombre de $\tilde{s}^{(j)}$ pour lesquels $s \leq \tilde{s}^{(j)}$, rejeter H_0 si $\tilde{r} \leq d$, où d est choisi tel que :

$$\frac{d+1}{P+1} = \frac{k}{Q} \tag{6.6}$$

Dwass (1957) montre ainsi qu'avec ce choix de d, la taille du test modifié reste exactement k/Q qui est la même que celle obtenue avec $Q = (n+m)!$ permutations. Cela veut dire que si l'on veut obtenir un test par permutations avec un niveau de signification de 5% ($\frac{k}{Q} = 0.05$) avec $P = 99$ permutations aléatoires, on doit poser $d = 4$ (d'après l'équation 6.6). Comme chaque $\tilde{s}^{(j)}$ est pondéré par la chance d'être choisie parmi toutes les $Q = (n+m)!$ permutations possibles, le test modifié de Dwass (1957) est donc un test randomisé exact. La particularité de ce résultat est que la forme exhaustive d'un test de randomisation est un test de permutation.

6.2.2.3 Procédure d'un test de Monte Carlo : Barnard (1963)

Le principe sous-jacent à la technique des tests de Monte Carlo est intimement lié à celui du test de Dwass (1957). L'idée est de remplacer la permutation par la simulation. On base ainsi le test de Monte Carlo sur une simulation de sorte à préserver le niveau du test. Cette idée qui a été préconisée par Barnard (1963) et Birnbaum (1974) est bien légitime. En effet, tout ce qu'un test de randomisation peut dire, c'est si une certaine structure dans les données peut ou non arriver par chance. Mais ceci est complètement spécifique aux données échantillonnées. C'est pourquoi il est opportun de préciser le processus générateur des données simulées à partir d'un modèle décrivant la population originale. De fait, un test de Monte Carlo est un test dans lequel une statistique de test observée calculée à partir de données échantillonnées est évaluée par comparaison avec la distribution obtenue en générant des ensembles alternatifs de données à partir d'un modèle. Ainsi, dans la procédure de Dwass (1957) qui a été présentée à la section 6.2.2.2, les valeurs $\tilde{s}^{(j)}$ sont générées à partir d'un modèle décrivant la population originale des données observées. Ces premiers travaux inspirent les remarques suivantes :

- On ne doit pas confondre la méthode de simulation dite de Monte Carlo et le test de Monte Carlo. La méthode de simulation de Monte Carlo s'applique généralement à des procédures où des grandeurs d'intérêt, normalement des intégrales, ou, en statistique, des espérances mathématiques, sont approchées par des moyennes générées par un grand nombre de simulations sur ordinateur. Par exemple, les statisticiens utilisent largement cette méthode pour vérifier les comportements de procédures statistiques d'une part (c'est le cas par exemple lorsqu'on étudie les performances d'un test par simulations) et d'autre part pour étudier les propriétés statistiques d'un processus générateur de données inconnu si l'on dispose d'une estimation du processus. À l'opposé, un test de Monte Carlo est une procédure inférentielle (tests d'hypothèses) dans laquelle on pose une hypothèse principale spécifique H_0 et on cherche à vérifier si les données observées sont consistantes avec la dite hypothèse. On peut donc se résumer en disant que les méthodes de simulation de Monte Carlo permettent une estimation de l'intervalle de confiance autour de la valeur observée, alors que les tests de Monte Carlo sont des procédures de tests.
- Selon la définition donnée précédemment, le *test bootstrap* (Efron, 1982) est un cas particulier de test de Monte Carlo. Le fait est que, un test bootstrap de signification consiste à comparer une statistique de test calculée à partir des données observées avec une distribution (de cette même statistique de test) obtenue en "rééchantillonnant" (c'est-à-dire on utilise l'échantillon observé pour modéliser la population en procédant à des tirages aléatoires avec remise). La seule différence, à présent, réside dans le fait que le test bootstrap est valide asymptotiquement alors que le test de Monte Carlo est valide pour un nombre fini de simulations. On reviendra sur ce point à la section 6.3.2.2.

6.2.3 Technique des tests de Monte Carlo

Le test de Monte Carlo tel qu'introduit originalement par Barnard (1963) est utilisé pour tester l'hypothèse que les données observées forment un échantillon aléatoire tiré d'une population spécifique. Cette méthode exige qu'un modèle décrivant la population spécifiée soit disponible. L'idée est de simuler plusieurs réalisations aléatoires de ce modèle sous

H_0, de calculer les valeurs de la statistique de test pour chacune de ces réalisations et d'estimer empiriquement la distribution de la statistique de test sous H_0 à partir de cet ensemble de valeurs.

On retrouve ici la même difficulté que celle évoquée dans la discussion de l'approche classique (section 6.1.2). En effet, si H_0 est une hypothèse multiple, on peut être confronté avec une statistique de test qui a des paramètres de nuisance. Dans pareille situation, le choix du processus générateur de données dépendrait normalement d'un ensemble de paramètres inconnus. L'idéal pour un test de Monte Carlo est alors de travailler avec une statistique de test pivotale (définition 6.3). La non pivotalité constitue une limitation dans l'élaboration d'un test de Monte Carlo. Dufour (1995) a proposé l'extension des tests de Monte Carlo aux cas de statistiques de test non pivotales. Dans cette section, on souligne les différentes étapes nécessaires à la construction d'un test de Monte Carlo.

6.2.3.1 Tests de Monte Carlo sans paramètres de nuisance

Il est souvent pratique de construire un test de Monte Carlo par l'intermédiaire des fonctions pivotales. Historiquement, les régions (ou intervalles) de confiance ont été rendues possibles grâce à de telles fonctions (Fisher, 1934). Ces dernières sont des fonctions des observations et des paramètres du modèle telles que la distribution de la fonction est connue (et donc ne dépend pas de paramètres de nuisance). De façon plus mathématique, on a la définition suivante :

Définition 6.3 *: On appelle fonction pivotale pour un paramètre θ une application mesurable $\pi : X^n \times \Theta \rightarrow$ ($\mathbb{R}$, $\Re$) dont la loi de probabilité Q est fixe, $\forall \theta$, c'est-à-dire est indépendante de θ.*

Cette définition signifie qu'une fonction pivotale n'est rien d'autre qu'une fonction des observations $x_1,\ldots,x_n$ et du paramètre θ dont la loi ne dépend d'aucun paramètre. Par exemple pour un échantillon théorique $X_1,\ldots,X_n \overset{iid}{\sim} N(\mu,\sigma)$, si σ est connu, la fonction $\pi(X_1,\ldots,X_n,\mu) = \sqrt{n}(\overline{X}-\mu)/\sigma$ est pivotale pour μ, de loi $N(0,1)$.

Pour illustrer la construction d'un test de Monte Carlo sans paramètres de nuisance on considère le problème qui consiste à tester une hypothèse principale H_0. Soit $\{T \geq k_\alpha\}$ la région critique du test, où $T_0 = \varphi(x_1,\ldots,x_n)$ est la valeur observée de la statistique de test $T = \varphi(X_1,\ldots,X_n)$ dont la distribution de probabilité sous H_0 est $F(x) = P(T \leq x)$. Le point critique k_α est déterminé à partir de l'équation :

$$P(T \geq k_\alpha) \leq \alpha \quad \Rightarrow \quad \{1 - F(k_\alpha) + P(T = k_\alpha)\} \leq \alpha \tag{6.7}$$

où α est le niveau de signification désiré $(0 < \alpha < 1)$. On note qu'on peut encore écrire la région de rejet du test $\{T \geq k_\alpha\}$ sous la forme alternative suivante :

$$G_N(T_0) \leq \alpha \tag{6.8}$$

avec $G_N(x) = P(T \geq x)$ la *fonction p-valeur* (appendice C), que l'on appelle encore la fonction critique du test. Elle est associée à la distribution de probabilité de la statistique de test sous H_0 :

$$G_N(x) = \begin{cases} P(T \geq x / H_0) & : \text{fonction critique d'un test unilatéral à droite} \\ P(|T| \geq x / H_0) & : \text{fonction critique d'un test bilatéral} \end{cases}$$

En général, la fonction G_N est inconnue. Pour l'estimer, on va supposer que la statistique de test T dont la distribution de probabilité $F(x) = P(T \leq x)$ est uniquement déterminée sous H_0, soit parce que H_0 est une hypothèse simple (c'est-à-dire, H_0 définit un processus unique de génération de données) ou bien parce que la distribution de T sous H_0 ne dépend pas de paramètres de nuisance (T est pivotale sous H_0). On peut donc simuler la statistique de test T sous H_0. Soit $T_1,\ldots,T_N$, N réalisations indépendantes ou à la rigueur échangeables. La propriété d'échangeabilité conditionne de façon suffisante la

majorité des résultats présentés ci-dessous. Elle se définit plus mathématiquement comme suit :

Définition 6.4 *On dit que les composantes d'un vecteur aléatoire* $(T_1,\ldots,T_N)'$ *sont échangeables si* $(T_{r_1},\ldots,T_{r_N})' \sim (T_1,\ldots,T_N)'$ *pour toute permutation* $(r_1,\ldots,r_N)'$ *d'entiers* $(1,\ldots,N)$.

D'après cette définition, les composantes d'un vecteur aléatoire sont échangeables si et seulement si la loi conjointe de ses composantes est invariante sous toutes les permutations. Une conséquence importante de cette définition est que les variables aléatoires échangeables sont forcément équidistribuées. Étant donné ce qui précède, la technique des tests de Monte Carlo fournit une méthode simple qui permet de remplacer la distribution de probabilité théorique $F(x)=\mathrm{P}(T\le x)=P\{(x-T)\ge 0\}$ par son analogue de l'échantillon calculé à partir des N réalisations précédentes :

$$\hat{F}_N(x)=\frac{1}{N}\sum_{i=1}^{N}\mathbf{1}_{[0,+\infty)}(x-T_i) \tag{6.9}$$

où,

$$\mathbf{1}_{(\mathrm{A})}(x)=\begin{cases}1, & \text{si} \quad x\in \mathrm{A}\\ 0, & \text{si} \quad x\notin \mathrm{A}\end{cases}$$

à partir de cette dernière expression, on peut déduire l'estimation de G_N :

$$\hat{G}_N(x)=\frac{1}{N}\sum_{i=1}^{N}\mathbf{1}_{[0,+\infty)}(T-x_i) \tag{6.10}$$

Pour la suite de l'inférence, la propriété fondamentale que l'on exploite ici est la suivante : pour obtenir la p-valeur du test, on va remplacer la distribution théorique de $F(x)$ sous H_0 par son estimé $\hat{F}_N(x)$ basé sur la simulation de façon à préserver la taille du test pour des tailles d'échantillons finis. Pour des distributions continues, cette propriété est exprimée par la proposition suivante (Dufour, 1995):

Proposition 6.1 *: Validité des tests de Monte Carlo pour les statistiques de distribution continue*

Soit $(T_0, T_1, \ldots, T_N)'$ *un vecteur de* $(N+1)$ *variables aléatoires réelles échangeables telles que*

$$\mathrm{P}(T_i = T_j) = 0 \quad pour \quad i \neq j, \quad i, j = 0,1,\ldots,N \tag{6.11}$$

En posant la p-valeur Monte Carlo $\hat{p}_N$ *comme :*

$$\hat{p}_N = \begin{cases} \dfrac{N\hat{G}_N(T_0)+1}{N+1}, & \text{pour un test unilatéral à droite} \\ 2\min\left[\dfrac{N\hat{G}_N(|T_0|)+1}{N+1}, 1-\dfrac{N\hat{G}_N(|T_0|)+2}{N+1}\right], & \text{pour un test bilatéral} \end{cases} \tag{6.12}$$

on a les résultats suivants :

$$\mathrm{P}\{\hat{G}_N(T_0) \leq \alpha_1\} = \mathrm{P}\{\hat{F}_N(T_0) \geq 1-\alpha_1\} = \frac{I[\alpha_1 N]+1}{N+1}, \quad pour \quad 0 \leq \alpha_1 \leq 1 \tag{6.13}$$

$$\mathrm{P}\{T_0 \geq \hat{F}_N^{-1}(1-\alpha_1)\} = \frac{I[\alpha_1 N]+1}{N+1}, \quad pour \quad 0 < \alpha_1 < 1 \tag{6.14}$$

$$\mathrm{P}\{\hat{p}_N(T_0) \leq \alpha\} = \frac{I[\alpha(N+1)]}{N+1}, \quad pour \quad 0 < \alpha < 1 \tag{6.15}$$

où $I[x]$ *est la partie entière de x.*

La conclusion que l'on peut tirer de cette proposition est la suivante : on peut choisir α_1 et N à partir de la relation suivante :

$$\alpha = \frac{I[\alpha_1 N]+1}{N+1} \tag{6.16}$$

Il en résulte que pour N relativement grand, α_1 sera très proche de α : en particulier, si $\alpha(N+1)$ est un entier, on peut prendre $\alpha_1 = \alpha - (1-\alpha)/N$. Il appert que les régions de rejet randomisées $\hat{p}_N(T_0) \leq \alpha$ et $\hat{G}_N(T_0) \leq \alpha$ ont le même niveau que la région de rejet non randomisée $G_N(T_0) \leq \alpha$.

Le fait que la distribution de probabilité de la statistique de test T soit continue sous H_0 joue un rôle important dans la démonstration de la proposition précédente. En effet, si tel n'est plus le cas, on peut rencontrer des égalités parmi $T_0, T_1, \ldots, T_N$. Dans pareille situation, le test randomisé n'est plus exact, en ce sens que l'égalité exprimée à l'équation 6.7 n'est plus toujours vraie. Pour éviter cet écueil, on utilise la technique de randomisation

de Dufour (1995) qui consiste à associer à chaque variable T_i, $i=1,\ldots,N$, une variable aléatoire uniforme U_i, $i=1,\ldots,N$ de sorte que $U_1,\ldots,U_N \overset{iid}{\sim} U(0,1)$ et que le vecteur $(T_1,\ldots,T_N)'$ soit indépendant du vecteur $(U_1,\ldots,U_N)'$. Ainsi, en considérant les paires $Z_i=(T_i,U_i)$, $i=1,\ldots,N$ qui sont ordonnées suivant l'ordre lexicographique :

$$(T_i,U_i)\geq(T_j,U_j) \;\Leftrightarrow\; \{T_i>T_j \quad ou \quad (T_i=T_j \quad et \quad U_i\geq U_j)\} \tag{6.17}$$

on a la proposition suivante (Dufour, 1995) :

Proposition 6.2 : *Validité des tests de Monte Carlo pour les statistiques de distribution générale*

Soit $(T_0,T_1,\ldots,T_N)'$ *un vecteur de* $(N+1)$ *variables aléatoires réelles échangeables. Soit* $(U_1,\ldots,U_N)'$ *un vecteur de N variables aléatoires* $\overset{iid}{\sim} U(0,1)$ *et indépendant du vecteur* $(T_0,T_1,\ldots,T_N)'$, *soient* $\hat{F}_N(x)$, $\hat{G}_N(x)$ *et* $\hat{p}_N(x)$ *comme précédemment et soient :*

$$\tilde{F}_N(x)=1-\hat{G}_N(x)+\frac{1}{N}\sum_{i\in\{j:T_j=x,1\leq j\leq N\}}^{N}\mathbf{1}_{[0,+\infty)}(U_0-T_i) \tag{6.18}$$

$$\tilde{G}_N(x)=1-\tilde{F}_N(x) \tag{6.19}$$

$$\tilde{p}_N=\begin{cases}\dfrac{N\tilde{G}_N(T_0)+1}{N+1}, & \text{pour un test unilatéral à droite}\\ 2\min\left[\dfrac{N\tilde{G}_N(|T_0|)+1}{N+1},1-\dfrac{N\tilde{G}_N(|T_0|)+2}{N+1}\right], & \text{pour un test bilatéral}\end{cases} \tag{6.20}$$

on alors les résultats suivants pour $0\leq\alpha_1\leq 1$:

$$\begin{aligned}\mathrm{P}\{\hat{G}_N(T_0)\leq\alpha_1\}&\leq\mathrm{P}\{\tilde{G}_N(T_0)\leq\alpha_1\}=\mathrm{P}\{\tilde{F}_N(T_0)\geq 1-\alpha_1\}\\ &=\frac{I[\alpha_1 N]+1}{N+1}\leq\mathrm{P}\{\hat{F}_N(T_0)\geq 1-\alpha_1\}\end{aligned} \tag{6.21}$$

avec $\mathrm{P}\{\hat{F}_N(T_0)\geq 1-\alpha_1\}=\mathrm{P}\{T_0\geq\hat{F}_N^{-1}(1-\alpha_1)\}$ *pour* $0\leq\alpha_1\leq 1$

$$\mathrm{P}\{\hat{p}_N(T_0)\leq\alpha\}\leq\mathrm{P}\{\tilde{p}_N(T_0)\leq\alpha\}=\frac{I[\alpha(N+1)]}{N+1}, \quad pour \;\; 0<\alpha<1 \tag{6.22}$$

En somme, les deux propositions précédentes disent que, pour effectuer un test de Monte Carlo, si la statistique de test T est pivotale et que $\alpha(N+1)$ est un entier, alors la région de rejet randomisée $\hat{p}_N(T_0) \leq \alpha$ est exact, en ce sens que

$$\mathrm{P}(\text{Rejeter } \mathrm{H}_0 \text{ à tort}) = \mathrm{P}\{\hat{p}_N(T_0) \leq \alpha / \mathrm{H}_0\} = \alpha \tag{6.23}$$

En d'autres termes, on a dans ces conditions un *test randomisé exact*. Une implication pratique de ces deux propositions est que, pour effectuer un test de l'hypothèse H_0 au niveau de signification 5%, on a seulement besoin d'une taille de simulation $N = 19, 39$, etc. ! C'est donc là un résultat très intéressant qui mérite d'être souligné. Les étapes nécessaires à l'application d'un test de Monte Carlo avec une statistique pivotale sont décrites dans le tableau 6.2.

Étape 1 *: définir l'hypothèse à tester* H_0.
Étape 2 *: choisir le niveau de signification* α *du test, puis définir le nombre N de simulations à effectuer de sorte que* $\alpha(N+1)$ *soit un entier.*
Étape 3 *: à partir des données observées, calculer la valeur observée de la statistique de test T, soit* T_0.
Étape 4 *: générer N réalisations indépendantes de la statistique T sous l'hypothèse* H_0, *soit* $T_1, \ldots, T_N$.
Étape 5 *: calculer la p-valeur Monte Carlo* $\hat{p}_N(T_0)$ *à partir de l'équation 6.12 (ou au besoin à partir de l'équation 6.20) et rejeter* H_0 *si* $\hat{p}_N(T_0) \leq \alpha$

Tableau 6.2 : Application d'un test de Monte Carlo avec une statistique pivotale

6.2.3.2 Tests de Monte Carlo avec paramètres de nuisance

Pour bien comprendre la pertinence des paramètres de nuisance dans un test de Monte Carlo, observons ce qui suit. Soit une expérience dont le résultat est représenté par un vecteur d'observation à valeurs réelles $X = \{X_1, \ldots, X_n\}$, et soit F_X sa fonction de

distribution de probabilité. Soit en outre, $\Im_n$ l'ensemble des fonctions de distributions de probabilité sur $\mathbb{R}^n$. Une hypothèse H_0 sur X est une assertion qui dit que :

$$H_0 : F_X \in \Im_n^{(0)} \tag{6.24}$$

où $\Im_n^{(0)}$ est un sous-ensemble de distributions possibles dans $\Im_n$. En particulier, dans le cas d'une loi normale, $\Im_n^{(0)}$ prend la forme suivante pour $x \in \mathbb{R}^n$:

$$\Im_n^{(0)} \equiv \left\{F(x): \quad F(x) = F_0(x/\theta_1, \theta_2) \quad et \quad \theta_1 = \theta_1^0\right\} \tag{6.25}$$

où F_0 est une loi normale de moyenne θ_1 et de variance θ_2 et $(\theta_1, \theta_2) \in \Theta_1 \times \Theta_2$. De façon plus concise, ont réécrit l'équation 6.24 comme suit :

$$H_0 : \theta_1 = \theta_1^0 \tag{6.26}$$

Dans ce cas, le paramètre θ_1 est qualifié de paramètre d'intérêt, car l'analyse s'intéresse spécialement à ce paramètre, tandis que θ_2 est ce qu'on appelle un paramètre de nuisance. Il est important de noter ici que dans le contexte d'un test de Monte Carlo la difficulté survient lorsque la distribution de probabilité de la statistique de test sous H_0 dépend de paramètres de nuisance. En effet, dans pareille situation, il n'est plus immédiatement clair de savoir, dans quelle famille de processus couverte par l'hypothèse H_0 on devrait simuler.

Dufour (1995) a proposé une extension des tests de Monte Carlo lorsque les paramètres de nuisance sont présents. L'idée est de générer la statistique de test, étant admise une valeur possible ou estimée du paramètre de nuisance, sous H_0. Cette approche est basée sur le fait que, pour une valeur donnée du paramètre de nuisance, la statistique de test résultante est pivotale. Toute la problématique se transpose alors sur la région critique à adopter de sorte

à préserver la taille du test. Les tests proposés par Dufour (1995) dans ce contexte peuvent être regroupés en deux catégories :

- le test local Monte Carlo (*"local Monte Carlo (LMC) test "*) ;
- le test maximisé Monte Carlo (*"maximised Monte Carlo (MMC) test"*).

Sans entrer dans le détail de ces deux types de test, on peut les résumer de la manière suivante. On considère une statistique de test T pour examiner une hypothèse principale H_0, et on suppose que la distribution de probabilité de T sous H_0 dépend d'un vecteur de paramètres inconnus θ. On désigne par T_0 la valeur observée de la statistique de test. Pour tester l'hypothèse principale H_0, le test local de Monte Carlo (LMC) procède avec un estimateur consistent $\hat{\theta}^0$ de θ. Ce faisant, on simule N valeurs de la statistique de test T, étant admis que $\theta = \hat{\theta}^0$, sous H_0 puis on calcule la p-valeur Monte Carlo étant donné $\theta = \hat{\theta}^0$, soit $\hat{p}_N\left(T_0/\hat{\theta}^0\right)$. Le test LMC peut être basé sur la région critique :

$$\hat{p}_N\left(T_0/\hat{\theta}^0\right) \leq \alpha \tag{6.27}$$

où, α est le niveau de signification du test. Dufour montre que le test LMC a un niveau, sous des conditions de régularité bien définies, correct asymptotiquement (c'est-à-dire lorsque la taille n d'échantillon devient très grande). Pour le test maximisé Monte Carlo (MMC), on examine H_0 en considérant un sous-ensemble M_0 de l'espace des paramètres compatible avec H_0 (c'est-à-dire l'espace des paramètres de nuisance). Puis, on simule, pour chaque valeur $\theta \in M_0$, N valeurs de la statistique de test T, et on calcule la p-valeur Monte Carlo $\hat{p}_N(T_0/\theta)$. Le test MMC est basé sur la région critique

$$\sup_{\theta \in M_0}\{\hat{p}_N(T_0/\theta)\} \leq \alpha \tag{6.28}$$

où, α est le niveau de signification du test. Dufour (1995) montre que le test MMC est exact (contrôle le niveau exactement). Les algorithmes illustrant les étapes d'application de ces deux tests sont rassemblés dans les tableaux 6.3 et 6.4.

Étape 1 : *définir l'hypothèse à tester* H_0.

Étape 2 : *choisir le niveau de signification* α *du test, puis définir le nombre N de simulations à effectuer de sorte que* $\alpha(N+1)$ *soit un entier.*

Étape 3 : *à partir des données observées, calculer la valeur observée de la statistique de test T, soit* T_0, *puis calculer l'estimateur contraint consistent* $\hat{\theta}^0$ *du paramètre inconnu* θ *de la distribution de la statistique de test T sous* H_0;.

Étape 4 : *générer, à partir de* $\hat{\theta}^0$, *N réalisations indépendantes de la statistique T sous l'hypothèse* H_0 : $T_1,\ldots,T_N$.

Étape 5 : *calculer la p-valeur Monte Carlo* $\hat{p}_N\left(T_0/\hat{\theta}^0\right)$ *à partir de l'équation 6.12 (ou au besoin à partir de l'équation 6.20) et rejeter* H_0 *si* $\hat{p}_N\left(T_0/\hat{\theta}^0\right) \leq \alpha$

Tableau 6.3 : Application d'un test de Monte Carlo avec paramètres de nuisance : test local

Étape 1 : *définir l'hypothèse à tester* H_0.

Étape 2 : *choisir le niveau de signification* α *du test, puis définir le nombre N de simulations à effectuer de sorte que* $\alpha(N+1)$ *soit un entier.*

Étape 3 : *à partir des données observées, calculer la valeur observée de la statistique de test T, soit* T_0.

Étape 4 : *maximiser la p-valeur Monte Carlo* $\hat{p}_N\left(T_0/\theta\right)$ *à travers* Θ_0 *(le sous-ensemble des paramètres inconnus compatibles avec l'hypothèse* H_0*).*

Étape 5 : *rejeter* H_0 *si* $\sup_{\theta\in\Theta_0}\{\hat{p}_N(T_0/\theta)\} \leq \alpha$

Tableau 6.4 : Application d'un test de Monte Carlo avec paramètres de nuisance : test maximisé

En définitive, le test local de Monte Carlo (LMC) exposé ci-dessus peut être interprété comme étant du bootstrap paramétrique, avec cependant une différence fondamentale. Alors que les tests bootstrap sont dans l'ensemble valides pour un nombre d'échantillons simulés qui tendent vers l'infini (asymptotique), le test LMC procède avec un petit nombre d'échantillons simulés explicitement pris en compte de façon à préserver la taille du test. Il en résulte que dans un test LMC, le fait que la procédure soit randomisée joue un rôle prépondérant dans la détermination de la taille du test. Par contre, dans un test de type bootstrap le nombre N d'échantillons simulés détermine la précision du test. En d'autres termes, on fait comme si le nombre de tirages était infini $(N \to \infty)$.

Tout compte fait, un test de Monte Carlo et un test bootstrap sont asymptotiquement équivalents à un test par permutations (Good, 1994). Des informations complémentaires sur les tests de Monte Carlo et leur relation avec les tests bootstrap sont données par Barnard (1963), Dufour (1995), Dufour et Kiviet (1998), et Dufour et Khalaf (2001).

6.3 Conclusion

Ce chapitre a été l'occasion de rappeler d'une part quelques notions sur les tests d'hypothèses, et, d'autre part, d'introduire certains notions de base de la théorie des tests de Monte Carlo.

Lorsqu'on cherche à développer un test statistique, il est important de ne pas perdre de vue trois conditions essentielles :

1. pour qu'un test d'hypothèses soit valide il faut que sa taille ne soit pas supérieure au niveau de signification α (équation 6.6) ;
2. pour qu'une statistique de test soit interprétable, il faut que sa distribution soit bornable sous l'hypothèse principale H_0 (équation 6.2) ; autrement seul $k_\alpha = \infty$ permet de satisfaire la contrainte de niveau de l'équation 6.5 ;

3. les régions de rejet du test doivent être basées sur des statistiques pivotales (et donc ne dépendent pas de paramètres de nuisance).

Bien souvent, ces conditions ne satisfont pas certains tests de détection de non stationnarité de type rupture qui sont habituellement employés dans les études de séries de données d'une variable hydrométéorologique (Lubès et al., 1998). En effet, on a vu au chapitre 5 que ces tests sont généralement construits sous l'hypothèse asymptotique. Or, bien souvent, pour les échantillons de taille finie, la théorie asymptotique ne donne que des résultats approximatifs. On doit toujours se méfier des tests de détection qui sont construits sous l'hypothèse asymptotique. Le fait est que, en travaillant sous l'hypothèse asymptotique on peut toujours définir un point critique asymptotique comme étant une approximation du point critique k_α. Un tel choix va produire un test de taille approximativement α qui pourrait être très différent de la valeur critique exacte. En procédant ainsi, on doit savoir trois choses :

1. une loi asymptotique avec des paramètres de nuisance est inutilisable ;
2. l'approximation asymptotique retenue peut être très mauvaise, car parfois on oublie les hypothèses sous lesquelles ces lois asymptotiques ont été établies ;
3. la théorie asymptotique est fondée sur des conditions de régularité faibles qui ne sont pas testables.

L'hypothèse asymptotique pourrait conduire à des décisions fallacieuses.

La technique des tests Monte Carlo présentée dans ce chapitre nous offre une alternative intéressante pour développer des tests de détection de rupture dans un échantillon fini. Cette technique, simple dans ses fondements, a la grande particularité de fournir des tests exacts (randomisés) basés sur toute statistique de test dont la distribution sous l'hypothèse principale est inconnue mais pouvant être simulée (à condition qu'elle ne dépend pas de paramètre de nuisance). Dans ce sens, les trois conditions énumérées précédemment satisfont ces tests.

7 TESTS DE MONTE CARLO DE DÉTECTION DE NON-STATIONNARITÉS

Une des hypothèses fondamentales de l'analyse fréquentielle en hydrométéorologie consiste à supposer que le phénomène observé est resté stationnaire. La caractérisation d'éventuelles non-stationnarités ou dérives dans le temps peut être menée par l'utilisation de tests statistiques sur des réalisations de tels phénomènes. Dans ce chapitre, on développe des tests exacts Monte Carlo relativement simples qui permettent de détecter certains types de non-stationnarités dans les séries de données de variables hydrométéorologiques. L'approche proposée se préoccupe essentiellement de la détection de ruptures (ou encore changement de régime) qui est le type de non-stationnarité le plus courant dans des séries hydrométéorologiques (Hubert et al., 1989). Une des conséquences les plus importantes des tests proposés est de répondre aux questions habituelles suivantes :

- Y a-t-il une (ou plusieurs) rupture significative dans les données ?
- Si la rupture existe, quelle est la structure de cette rupture : brusque ou avec continuité ?
- Si la rupture existe localement, est-il raisonnable de l'admettre régionalement ?

Puisque la démarche proposée vise à introduire les potentialités offertes par la technique des tests de Monte Carlo, on est d'avis que l'approche développée ici peut tout aussi bien se généraliser pour traiter les autres types de non-stationnarité qui ont été évoqués au chapitre 5.

7.1 Tests ponctuels de Monte Carlo de détection de non-stationnarités hydrométéorologiques

Les tests de Monte Carlo qui sont proposés ici concernent l'exploitation d'une série de données et d'une seule à des fins de détection d'une non-stationnarité possible. L'analyse résultante est donc qualifiée de ponctuelle, "par site" ou encore locale. On se doit donc de

préciser qu'elle est une étape préliminaire qui s'impose avant de procéder à des interprétations prenant en compte la dimension spatiale des dites séries ponctuelles.

On a vu qu'une rupture est généralement définie comme étant un changement dans la loi de probabilité de la série chronologique à un instant donné (chapitre 5). Le changement dont il est question ici est celui de la moyenne de la série. Il est expliqué par l'existence d'un changement brusque ou continu de moyenne (figure 7.1), provoquée par une modification brusque ou continue du processus physique générateur (Hubert et al.,1989). Dans la littérature hydrométéorologique, les tests statistiques les plus utilisés traitent habituellement la détection d'une rupture brusque (Bernier, 1994 ; Lubès et al., 1998). Or, les scientifiques argumentent que dans le phénomène de rupture il est tout aussi vraisemblable de s'attendre à un changement continu de la moyenne (Jaruskova, 1997).

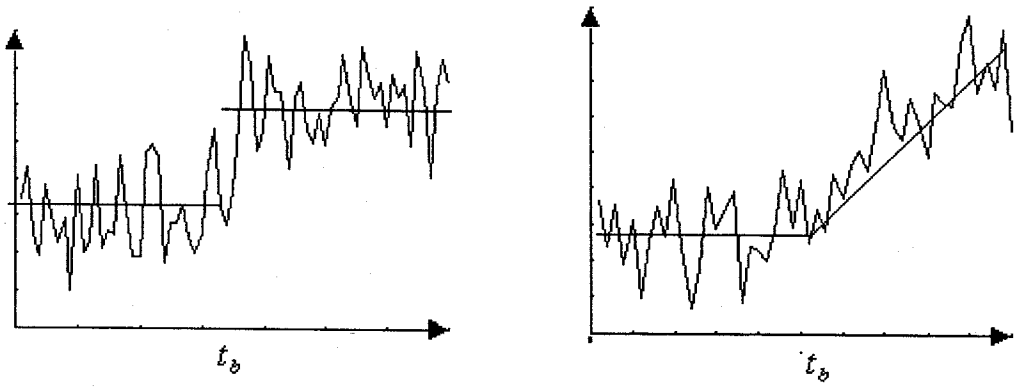


Figure7.1:Illustration d'une rupture brusque et d'une rupture avec continuité à la date t_b

Le problème de rupture a été abondamment traité par de nombreux auteurs dans la littérature hydrométéorologique. Lorsque la date de rupture t_b est supposée connue, les tests les plus répandus comparent les moyennes avant et après la date d'intérêt t_b. Dans la situation la plus vraisemblable où la date de rupture est inconnue, très peu de tests existent dans la littérature. Bien souvent, les conditions généralement admises pour ces tests sont que les observations sont indépendantes, une situation qui est moins réaliste pour les séries hydrométéorologiques. On peut aussi rencontrer dans des applications pratiques le

problème de l'existence de plusieurs ruptures dans une série de données. Là encore, la littérature hydrométéorologique est dépourvue d'outils statistiques adéquats pour résoudre ce problème. On tente ici une approche basée sur la technique des tests de Monte Carlo afin d'apporter des éléments de solution à ces différents problèmes. Le but recherché consiste à obtenir des tests exacts de détection de non-stationnarité pour les différents problèmes évoqués. En premier lieu, il est capital de proposer un test exact de détection d'une rupture à la fois dans un échantillon indépendant et dans un échantillon de données autocorrélées. De même il est également important de proposer un test exact de détection de plusieurs ruptures dans ces mêmes types d'échantillons.

7.1.1 Tests de Monte Carlo de détection d'une rupture

7.1.1.1 Position du problème

Les tests de Monte Carlo de détection d'une rupture qui sont proposés dans le présent contexte sont des tests de type *supremum*. L'idée d'un test supremum est la suivante. On suppose qu'on a une série de données temporelles de taille n pour laquelle on veut vérifier l'existence d'une rupture. L'hypothèse principale H_0 stipule qu'il n'existe aucune rupture dans la série de données, tandis que la contre-hypothèse H_1 prétend qu'il existe une rupture à la date $t_b \in \{1,\ldots,n\}$. Tout ce qu'on sait assurément est que si une rupture existe dans la série, alors elle s'est produite à l'instant t_b. Soit T_{t_b} la statistique de test qui sert à tester H_0 pour t_b connu. Le test du supremum est construit de la façon suivante : pour chaque instant d'observation susceptible de correspondre avec une date de rupture, t_b ($t_b = 1, \ldots, n$), on calcule la statistique T_{t_b} observée correspondante et, la date de rupture est estimée à la date t_b pour laquelle la statistique T_{t_b} atteint sa valeur maximale. La statistique du supremum s'exprime alors comme suit :

$$T_{\sup} = \sup_{1 \le t_b \le n} \left\{ T_{t_b} \right\} \tag{7.1}$$

Soient $T_{\sup}^{obs}$ la réalisation de la statistique de test $T_{\sup}$ calculée à partir de l'échantillon empirique et k_α le point critique déterminé pour un niveau de signification α. Si $T_{\sup}^{obs} > k_\alpha$, on rejette H_0 (il y a une rupture). Si par contre $T_{\sup}^{obs} \leq k_\alpha$, on accepte H_0 (il n'y a pas de rupture).

En pratique, l'utilisation de la statistique de test $T_{\sup}$ en vue de prendre une décision fait appel à deux problèmes cruciaux. Le premier est posé par les effets de bords, c'est-à-dire les ruptures intervenant aux bords de l'intervalle d'observations. En effet, pour les instants t_b proches de 1 ou n il peut être judicieux de bien construire le test de type supremum. Par exemple, pour t_b proche de 1 on compare un estimé calculé à partir d'un grand nombre d'observations – les $(n - t_b)$ observations après la rupture – avec un estimé calculé à partir d'un petit nombre d'observations – les t_b observations avant la rupture. On voit donc qu'il y a un problème d'information aux bords de l'intervalle d'observations. Pour pénaliser ces effets de bords, plusieurs approches ont été proposées dans la littérature. James et al.(1987) ont proposé de travailler avec une certaine proportion d'instants temporels dans l'intervalle d'observations tandis que Deshayes et Picard (1986) ont suggéré de pondérer la statistique de test correspondant à t_b dans l'intervalle d'observations. Le deuxième problème est celui de la distribution de la statistique de test $T_{\sup}$ sous H_0. En effet, le test du supremum essaie de répondre à la question pertinente suivante : «la valeur observée par la statistique $T_{\sup}$ est-elle significative à un seuil de signification donné ? ». Dès lors, la question pratique est de savoir si la date t_b correspondant au maximum des valeurs de la statistique est une rupture ou non. La connaissance de la distribution exacte de la statistique $T_{\sup}$ est donc essentielle à la prise de décision, puisqu'elle permet de dériver les valeurs critiques associées au test. Comme on peut le remarquer malheureusement, cette loi est difficile à dériver sous H_0. Plusieurs méthodes différentes ont été proposées dans la littérature hydrométéorologique pour dériver des valeurs critiques approximatives (Jaruskova, 1997). Il s'agit principalement de la méthode de Bonferroni et ses améliorations, de la méthode des distributions asymptotiques et de la méthode par simulation.

La technique des tests de Monte Carlo qu'on utilise ici concerne le deuxième problème soulevé précédemment. Il s'agit de proposer une solution qui consiste à résoudre le problème litigieux de la distribution de la statistique de test $T_{\sup}$ sous H_0, de sorte à obtenir des points critiques exacts pour le test de type supremum. L'approche proposée comporte deux étapes principales. La première consiste à spécifier un modèle, qui décrit la population, ou encore le phénomène de rupture que l'on étudie. Ce modèle est ensuite mis en œuvre pour calculer la valeur observée ($T_{\sup}^{obs}$) de la statistique de test $T_{\sup}$. Dans la deuxième étape, on cherche à savoir si la valeur observée ($T_{\sup}^{obs}$) de la statistique de test $T_{\sup}$ est la réalisation d'un processus stationnaire (sans rupture). Pour ce faire, en supposant que la distribution de la statistique de $T_{\sup}^{obs}$ est pivotale sous H_0 (et donc simulable sous H_0), on va générer N valeurs de la statistique de test $T_{\sup}$ sous H_0 (d'après la section 6.2.3.1 la valeur de N est choisie de sorte que $\alpha(N+1)$ soit un entier). Ce qui revient à répéter les opérations suivantes N fois : (1) on génère un échantillon à partir du modèle, en supposant que H_0 est vraie ; (2) on calcule, une réalisation de la statistique de test $T_{\sup}$ à partir de l'échantillon généré en (1), soit $T_{\sup}^{i}$ pour l'étape i.

Les N réalisations ($T_{\sup}^{1}, T_{\sup}^{2}, \ldots, T_{\sup}^{N}$) obtenues par ce procédé représentent les valeurs que la statistique de test $T_{\sup}$ pourrait prendre dans différentes autres expériences où H_0 est vraie. Elles représentent donc, toutes ensemble, une estimation de la distribution d'échantillonnage de la statistique de test $T_{\sup}$ sous H_0. En conséquence, la décision statistique est prise en comparant la valeur $T_{\sup}^{obs}$ (valeur observée de $T_{\sup}$) à cette distribution d'échantillonnage. Une telle décision se prend plus facilement au moyen de la p-valeur Monte Carlo (équation 6.12). C'est le principe de base des tests de Monte Carlo qui va nous permettre de construire des tests originaux de détection de non-stationnarité.

7.1.1.2 Test de Monte Carlo de détection d'une rupture dans un échantillon sans persistance

7.1.1.2.1 Modèle

Le modèle le plus connu en hydrométéorologie pour représenter un phénomène de rupture est le modèle de régression linéaire présenté à l'équation 3.11. Ce modèle a été abondamment étudié dans la littérature : voir par exemple Refsgaard et al. (1989), Laberge (1995), et Jaruskova (1996) pour une synthèse et une bibliographie complète. On peut re-écrire ce modèle sous la forme synthétique suivante :

$$x_t = \beta_0 + \beta_1 ID(t_b) + \varepsilon_t \qquad t = 1,\ldots,n \tag{7.2}$$

où,

- la rupture (changement de moyenne ou de régime de la série x_t) se fait au moyen de la variable indicatrice $ID(t_b)$ qui répond à l'une ou l'autre des définitions suivantes. Pour une rupture brusque à la date $t = t_b$:

$$ID(t_b) = \begin{cases} 0 & \text{si } t \le t_b \\ 1 & \text{autrement} \end{cases} \tag{7.3}$$

  par contre dans le cas d'une rupture avec continuité à la date $t = t_b$:

$$ID(t_b) = \begin{cases} 0 & \text{si } t \le t_b \\ t - t_b & \text{autrement} \end{cases} \tag{7.4}$$

- β_0 est l'ordonnée à l'origine dans le modèle de régression et β_1 est l'amplitude (ou la pente) de la rupture à la date $t = t_b$.
- Les erreurs $\varepsilon_1,\ldots,\varepsilon_n$ sont mutuellement indépendantes et homoscédastiques. Par ailleurs, on suppose que leur distribution est connue à un paramètre d'échelle près ($E(\varepsilon_t) = 0$ et $Var(\varepsilon_t) = \sigma_\varepsilon^2$).
- La date de rupture $t = t_b$ est inconnue.

Le modèle de régression qu'on vient de définir correspond à une évolution de la série x_t avec un changement de pente à la date $t = t_b$. On se place donc ici dans le cadre d'un

problème de détection de rupture dans la pente déterministe. Détecter la rupture dans la série x_t, revient à tester l'hypothèse principale de stationnarité $H_0 : \beta_1 = 0$ («il n'y a pas de rupture dans la série x_t ») contre la contre-hypothèse $H_1 : \beta_1 \neq 0$ («il y a une rupture dans la série x_t à l'instant $t = t_b$ »). L'instant de rupture $t = t_b$ est supposé inconnu.

7.1.1.2.2 Statistique de test

En s'inspirant de l'équation 7.1, la statistique du test de Monte Carlo que nous proposons pour tester H_0 est définie par :

$$F_{\max} = \max_{3 \le t_b \le n-1} \{F(t_b)\} \tag{7.5}$$

où,

- on suppose que t_b est l'instant d'observation susceptible de correspondre avec une date de rupture, vérifiant $3 \le t_b \le n-1$;
- pour chaque instant d'observation $t = t_b$, $F(t_b)$ est la statistique de test utilisée pour tester $H_0 : \beta_1 = 0$ dans le modèle de régression de l'équation 7.2.

On note que, dans le cas du modèle de régression de l'équation 7.2, la statistique de test, $F(t_b)$, pour tester $H_0 : \beta_1 = 0$ est donnée par :

$$F(t_b) = \frac{\dfrac{\sum_{t=1}^{n} (\hat{x}_t - \bar{x})^2}{\hat{\sigma}^2}}{\dfrac{\sum_{t=1}^{n} (x_t - \hat{x}_t)^2}{\hat{\sigma}^2}} \times \frac{n-2}{1} \tag{7.6}$$

où,

- $\hat{x}_t = \hat{\beta}_0 + \hat{\beta}_1 ID(t_b)$

- $\bar{x} = \dfrac{\sum_{t=1}^{n} x_t}{n}$
- $\hat{\sigma}^2 = \dfrac{\sum_{t=1}^{n} (x_t - \hat{x}_t)^2}{n-2}$
- avec $\hat{\beta}_0$ et $\hat{\beta}_1$ les estimateurs du maximum de vraisemblance (c'est-à-dire des moindres carrés ordinaires) de β_0 et β_1 sous H_0.

Sous H_0, la statistique de test $F(t_b)$ est distribuée d'après la loi de Fisher avec 1 et $(n-2)$ degrés de liberté.

Dans notre règle de détection d'une éventuelle rupture, si $F_{\max}^{obs}$ représente la valeur observée de la statistique de test $F_{\max}$ et si α est le niveau de signification choisi à l'avance, l'hypothèse principale d'absence de rupture (H_0) est rejetée si $F_{\max}^{obs} > k_\alpha$ (ici, k_α est le point critique). D'autre part, l'instant $\hat{t}_b$ appartenant à $[3, n-1]$ tel que : $\hat{t}_b = \arg \max_{3 \le t_b \le n-1} \{F(t_b)\}$ correspond à la date de rupture quand H_0 est rejetée.

Pour déterminer le point critique k_α en fonction du niveau de signification α, une démarche analytique est extrêmement difficile. En effet, d'après l'équation 7.5, la statistique de test $F_{\max}$ représente le maximum de la statistique de test $F(t_b)$ qui sert à tester $H_0 : \beta_1 = 0$ pour chaque instant d'observation t_b appartenant à $[3, n-1]$. Pour contourner ce problème, on va dériver une estimation de la distribution d'échantillonnage de $F_{\max}$ sous H_0 à l'aide de la technique des tests de Monte Carlo. Pour y arriver, on doit d'abord montrer que la statistique de test $F_{\max}$ est simulable sous H_0, ce qui revient encore démontrer que la statistique $F_{\max}$ est pivotale.

Pour démontrer que la statistique de test $F_{\max}$ est pivotale, il suffit de montrer que la statistique $F(t_b)$ est pivotale, car d'après la définition 6.3, le supremum d'une statistique

pivotale est pivotale. Le principe de la démonstration est le suivant. Tout d'abord, on réécrit, sans perte de généralité, le modèle de régression de l'équation 7.2 sous la forme matricielle suivante :

$$Y = X\theta + \varepsilon \tag{7.7}$$

où,

- $X = [1 \quad ID(t_b)]$, est la matrice des variables explicatives de dimension $(n \times k)$ de rang k;
- $Y = (x_1, \ldots, x_n)$, est la variable dépendante de dimension $(n \times 1)$;
- $\theta = (\beta_0, \beta_1)$, est le vecteur des paramètres de régression de dimension $(k \times 1)$;
- $\varepsilon \sim N(0, \sigma^2 I)$, est le vecteur des erreurs de régression de dimension $(n \times 1)$

L'hypothèse d'absence de rupture relative à ce modèle de régression multiple a la forme matricielle suivante : H_0 : $R\theta = r$, où $R = (0,1)'$ est une matrice des contraintes sous H_0 de dimension $(q \times k)$ et de rang q. Il en résulte que l'estimateur non contraint de θ est

$$\hat{\theta} = (X'X)^{-1} X'Y \tag{7.8}$$

tandis que son estimateur contraint est donné par

$$\hat{\theta}_c = \hat{\theta} + (X'X)^{-1} R'\left[R\,(X'X)^{-1} R'\right]^{-1}\left(r - R\hat{\theta}\right) \tag{7.9}$$

Dans ce cas, les résidus non contraints, $\hat{\varepsilon}$, et les résidus contraints, $\hat{\varepsilon}_c$, de ce modèle de régression sont définis par

$$\hat{\varepsilon} = Y - X\hat{\theta} = \left[I - X\,(X'X)^{-1} X'\right]Y = M_{(t_b)} Y = M_{(t_b)}\varepsilon \tag{7.10}$$

$$\hat{\varepsilon}_c = Y - X\hat{\theta}_c = \left\{\left[I - X(X'X)^{-1} X'\right] + X(X'X)^{-1} R'\left[R(X'X)^{-1} R'\right]^{-1} R(X'X)^{-1} X'\right\}\varepsilon$$

$$= M_c \varepsilon \tag{7.11}$$

Il vient que M_c et $M_{(t_b)}$ sont les matrices de projections usuelles sur l'espace contraint et non contraint engendré par les colonnes des régresseurs pertinents.

Étant donné les expressions précédentes, on peut réécrire, d'après l'équation 7.6, la statistique $F(t_b)$ comme suit :

$$F(t_b) = \frac{\hat{\varepsilon}_c' \hat{\varepsilon}_c - \hat{\varepsilon}' \hat{\varepsilon}}{\hat{\varepsilon}' \hat{\varepsilon}} \times \frac{n-k}{q} \tag{7.12}$$

En utilisant les expressions obtenues aux équations 7.10 et 7.11 dans l'équation 7.13, on obtient

$$F(t_b) = \frac{\varepsilon' M_c \varepsilon - \varepsilon' M_{(t_b)} \varepsilon}{\varepsilon' M_{(t_b)} \varepsilon} \times \frac{n-k}{q} \tag{7.13}$$

Si maintenant on divise le numérateur et le dénominateur de l'équation 7.13 par σ, on a :

$$F(t_b) = \frac{\left(\frac{\varepsilon}{\sigma}\right)' M_c \left(\frac{\varepsilon}{\sigma}\right) - \left(\frac{\varepsilon}{\sigma}\right)' M_{(t_b)} \left(\frac{\varepsilon}{\sigma}\right)}{\left(\frac{\varepsilon}{\sigma}\right)' M_{(t_b)} \left(\frac{\varepsilon}{\sigma}\right)} \times \frac{n-k}{q} \tag{7.14}$$

d'où,

$$F(t_b) = \frac{w' M_c w - w' M_{(t_b)} w}{w' M_{(t_b)} w} \times \frac{n-k}{q} \tag{7.15}$$

avec $w = \frac{\varepsilon}{\sigma}$.

Comme w est distribuée d'après une loi normale centrée réduite ($w \sim N(0,1)$), il en résulte que la statistique $F(t_b)$ est pivotale. En conséquence, d'après l'équation 7.6, la statistique de test $F_{\max}$ est simulable sous H_0. Elle est donc pivotale. C.Q.F.D.

Il est important de noter que d'après l'équation 7.15, comme w est distribuée d'après une loi normale centrée réduite, il n'est pas nécessaire d'avoir une loi spécifique pour les erreurs ε_t. Il suffit juste d'avoir une loi connue. Autrement dit, l'hypothèse de normalité des erreurs que l'on fait habituellement dans un modèle de régression classique n'est pas nécessaire ici. La pivotalité ne requière donc pas la normalité.

7.1.1.2.3 Région critique du test

Pour la construction du test, on prend comme région de rejet de H_0, la région randomisée. Ainsi, puisque $F_{\max}^{obs}$ est la valeur observée de la statistique du test ($F_{\max}$), la région critique est définie par :

$$\hat{p}_N\left(F_{\max}^{obs}\right) \leq \alpha \tag{7.16}$$

où,

- $\hat{p}_N\left(F_{\max}^{obs}\right)$ est la p-valeur de Monte Carlo (équation 6.12) et α est le niveau de signification du test fixé à l'avance.

Nous décrivons maintenant l'algorithme du test Monte Carlo que nous proposons, en nous limitant au cadre qui nous intéresse ici, à savoir la détection d'une rupture dans un échantillon sans persistance. On dispose d'une série de données $\{x_1,\ldots,x_n\}$ d'une variable hydrométéorologique et un modèle décrivant le phénomène de rupture. La décomposition des étapes du test de Monte Carlo pour détecter une rupture brusque (le modèle de rupture est donné aux équations 7.2 et 7.3) et une rupture avec continuité (le modèle de rupture est donné aux équations 7.2 et 7.4) est donnée respectivement aux tableaux 7.1 et 7.2.

*% **Le modèle considéré pour représenter le phénomène de rupture avec continuité est***

*% **celui exprimé à l'équation 7.2 avec** $ID(t_b)$ **telle qu'exprimée à l'équation 7.3***

Étape 1 : *Calcul de la valeur observée de la statistique de test,* $F_{\max}^{obs}$, *et de la date éventuelle de rupture,* $\hat{t}_b$, *à partir des données*

Pour $t_b = 3,\ldots,n-1$

  Calculer $F_{obs}(t_b)$, la valeur observée de la statistique $F(t_b)$, dans l'équation 7.6

t_b suivant

$\rightarrow \quad F_{\max}^{obs} = \max_{3\le t_b \le n-1} \{F_{obs}(t_b)\}$

$\rightarrow \quad \hat{t}_b = \arg \max_{3\le t_b \le n-1} \{F_{obs}(t_b)\}$

Étape 2 : *Génération de* N *réalisations de la statistique* $F_{\max}$ *sous* H_0

Pour $i = 1,\ldots,N$

  Pour $t_b = 3,\ldots,n-1$

    Générer $w \sim N(0,1)$

    Calculer $F_{res}(t_b)$, la réalisation de la statistique $F(t_b)$, dans l'équation 7.15

  t_b suivant

  Calculer $F_{\max}^{i} = \max_{3\le t_b \le n-1} \{F_{res}(t_b)\}$

i suivant

$\rightarrow \quad \{F_{\max}^{1}, F_{\max}^{2}, \ldots, F_{\max}^{N}\}$

Étape 3 : *Calcul de la p-valeur à partir de l'échantillon* $\{F_{\max}^{obs}, \; F_{\max}^{1}, F_{\max}^{2}, \ldots, F_{\max}^{N}\}$

Calculer la p-valeur Monte Carlo $\hat{p}_N$ dans l'équation 6.12

Étape 4 : *Décision*

Rejeter H_0 si $\hat{p}_N$ est plus petite ou égale à un niveau de signification α

Tableau 7.1 : Test de Monte Carlo de détection d'une rupture brusque dans un échantillon sans persistance

*% **Le modèle considéré pour représenter le phénomène de rupture avec continuité est***

*% **celui exprimé à l'équation 7.2 avec** $ID(t_b)$ **telle qu'exprimée à l'équation 7.4***

Étape 1 : *Calcul de la valeur observée de la statistique de test,* $F_{\max}^{obs}$, *et de la date éventuelle de rupture,* $\hat{t}_b$, *à partir des données*

Pour $t_b = 3,\ldots,n-1$

  Calculer $F_{obs}(t_b)$, la valeur observée de la statistique $F(t_b)$, dans l'équation 7.6

t_b suivant

$\rightarrow \quad F_{\max}^{obs} = \max_{3\le t_b \le n-1} \{F_{obs}(t_b)\}$

$\rightarrow \quad \hat{t}_b = \arg \max_{3\le t_b \le n-1} \{F_{obs}(t_b)\}$

Étape 2 : *Génération de* N *réalisations de la statistique* $F_{\max}$ *sous* H_0

Pour $i = 1,\ldots,N$

  Pour $t_b = 3,\ldots,n-1$

    Générer $w \sim N(0,1)$

    Calculer $F_{res}(t_b)$, la réalisation de la statistique $F(t_b)$, dans l'équation 7.15

  t_b suivant

  Calculer $F_{\max}^{i} = \max_{3\le t_b \le n-1} \{F_{res}(t_b)\}$

i suivant

$\rightarrow \quad \{F_{\max}^{1}, F_{\max}^{2}, \ldots, F_{\max}^{N}\}$

Étape 3 : *Calcul de la p-valeur à partir de l'échantillon* $\{F_{\max}^{obs}, \quad F_{\max}^{1}, F_{\max}^{2}, \ldots, F_{\max}^{N}\}$

Calculer la p-valeur Monte Carlo $\hat{p}_N$ dans l'équation 6.12

Étape 4 : *Décision*

Rejeter H_0 si $\hat{p}_N$ est plus petite ou égale à un niveau de signification α

Tableau 7.2 : Test de Monte Carlo de détection d'une rupture avec continuité dans un échantillon sans persistance

7.1.1.3 Test de Monte Carlo de détection d'une rupture dans un échantillon avec persistance

Les tests de Monte Carlo de détection de rupture présentés précédemment ont été développés sous la condition d'indépendance des observations. Cela signifie que chaque valeur observée de l'échantillon à examiner doit être non influencée par les autres valeurs observées du même échantillon. Comme on l'a maintes fois soulignée dans cette thèse, cette condition est rarement remplie pour les séries de données d'une variable hydrométéorologique. En effet, à la section 3.5.1, on a vu que le phénomène de persistance est très présent dans de telles séries de données, d'où l'importance de proposer des tests de Monte Carlo de détection de rupture qui tiennent compte de ce phénomène.

7.1.1.3.1 Modèle

Le modèle le plus connu en hydrométéorologie pour représenter un phénomène de rupture est le modèle de régression linéaire présenté à l'équation 3.11. Sa formulation la plus simple est donnée par :

$$x_t = \beta_0 + \rho\, x_{t-1} + \beta_1\, ID(t_b) + \varepsilon_t \qquad t = 1,\ldots,n \tag{7.17}$$

où,

- la rupture (changement de moyenne ou de régime de la série x_t) se fait au moyen de la variable indicatrice $ID(t_b)$ qui répond à l'une ou l'autre des définitions données aux équations 7.3 et 7.4.
- les paramètres β_0 et β_1 sont identiques à l'équation 7.2 tandis que ρ est le paramètre d'autocorrélation tel que $|\rho| < 1$.
- Les erreurs $\varepsilon_1,\ldots,\varepsilon_n$ sont mutuellement indépendantes et homoscédastiques. Par ailleurs, on suppose que leur distribution est connue à un paramètre d'échelle près ($E(\varepsilon_t) = 0$ et $Var(\varepsilon_t) = \sigma_\varepsilon^2$).
- La valeur x_0 est indépendante des erreurs $\varepsilon_1,\ldots,\varepsilon_n$. Elle est soit fixée ou aléatoire.
- La date de rupture $t = t_b$ est inconnue.

Ce modèle de régression a déjà été utilisé dans la littérature hydrométéorologique (Laberge, 1995). On rappelle que l'introduction de la variable retardée x_{t-1} permet de modéliser la dynamique de la série x_t. Cette variable retardée permet d'introduire, à l'intérieur du modèle, les corrélations entre les observations successives et, donc, de les exclure de la série des résidus. Ainsi, le modèle de régression de l'équation 7.17 décrit une évolution de la série autocorrélée x_t par morceaux avec une rupture dans la pente déterministe à la date $t = t_b$, pourvu que la corrélation entre les données soit de nature stationnaire et donc, ici, que $|\rho| < 1$. On fait donc face au problème de détection d'une rupture dans la pente déterministe de la série autocorrélée x_t. La date de rupture $t = t_b$ est supposée inconnue.

Détecter la rupture dans la série autocorrélée x_t, revient à tester l'hypothèse principale de stationnarité $H_0 : \beta_1 = 0$ («il n'y a pas de rupture dans la série x_t ») contre la contre-hypothèse $H_1 : \beta_1 \neq 0$ («il y a une rupture dans la série x_t à la date $t = t_b$ »). On suppose ici que le paramètre ρ est stable sous les hypothèses H_0 et H_1. Ainsi, sous H_0, la série autocorrélée x_t est stationnaire.

7.1.1.3.2 Statistique de test

En s'inspirant toujours de l'équation 7.1, la statistique du test de Monte Carlo que nous proposons dans ce cas-ci pour tester H_0 est définie par :

$$F_{\max_\rho} = \max_{3 \le t_b \le n-1} \{F(t_b)\} \tag{7.18}$$

où,

- on suppose encore que t_b est l'instant d'observation susceptible de correspondre avec une date de rupture, vérifiant $3 \le t_b \le n-1$;

- pour chaque instant d'observation $t = t_b$, $F(t_b)$ est la statistique de Fisher (équation 7.6) pour tester $H_0 : \beta_1 = 0$ dans le modèle de régression de l'équation 7.17.

La règle de détection d'une éventuelle rupture est la suivante. Soit $F_{\max_\rho}^{obs}$ la valeur observée de la statistique de test $F_{\max_\rho}$, l'hypothèse principale d'absence de rupture (H_0) est rejetée si $F_{\max_\rho}^{obs} > k_\alpha^*$ (ici, k_α^* est le point critique et α est le niveau de signification choisi à l'avance). L'instant $\hat{t}_b$ appartenant à $[3, n-1]$ tel que : $\hat{t}_b = \arg \max_{3 \le t_b \le n-1} \{F(t_b)\}$ correspond à la date de rupture quand H_0 est rejetée.

On remarquera aussi que, comme pour la statistique de test $F_{\max}$ (équation 7.5), une démarche analytique est extrêmement difficile pour déterminer le point critique k_α^* en fonction du niveau de signification α. Pour contourner ce problème, on va utiliser la technique des tests de Monte Carlo.

Un raisonnement analogue à celui de la propriété de pivotalité de la statistique de test $F_{\max}$ (équation 7.5) montre que, sous H_0, le paramètre ρ du modèle de régression de l'équation 7.17 est inconnu. Il en résulte que la loi de distribution de la statistique de test $F_{\max_\rho}$, exprimée à l'équation 7.18, dépend de ce paramètre inconnu. Le paramètre ρ est donc un paramètre de nuisance pour la loi de distribution $F_{\max_\rho}$. Ainsi, sur la base de cette observation et d'après les équations 7.18 et 7.15, la statistique de test $F_{\max_\rho}$ n'est pas pivotale. Elle n'est donc pas simulable sous H_0.

Il est facile de voir que si l'on fixe le paramètre ρ (et donc connue), la loi de distribution de $F_{\max_\rho}$ est pivotale sous H_0. En effet, pour ρ fixé, on peut réécrire le modèle de régression de l'équation 7.17 sous la forme $x_t - \rho\, x_{t-1} = \beta_0 + \beta_1\, ID(t_b) + \varepsilon_t \quad t = 1, \ldots, n$.

Si dans cette dernière formulation on pose $x_t^* = x_t - \rho x_{t-1}$, le modèle de régression résultant est l'analogue du modèle de régression de l'équation 7.2. Il en résulte que la statistique $F_{\max_\rho}$ est l'équivalent de $F_{\max}$ et donc pivotale. Il en résulte que, pour un ρ fixé, on peut obtenir une p-valeur de Monte Carlo ($\hat{p}_N\left(F_{\max_\rho}^{obs}/\rho\right)$) à partir de l'équation 6.12. Sur la base de ce résultat et compte tenu de ce qui précède, on va construire le test maximisé de Monte Carlo pour détecter une rupture dans un échantillon avec persistance.

7.1.1.3.3 Région critique du test

Pour la construction du test Monte Carlo (test MMC) de détection de rupture dans le modèle de régression de l'équation 7.17, on prend comme région de rejet de H_0, la région randomisée :

$$\sup_{\rho \in M_0} \left\{\hat{p}_N\left(F_{\max_\rho}^{obs}/\rho\right)\right\} \leq \alpha \qquad (7.19)$$

où,

- $F_{\max}^{obs}$ est la valeur observée de la statistique du test ($F_{\max_\rho}$) ;
- $M_0 \subseteq]-1;1[$ est un sous-ensemble des valeurs de ρ ;
- $\hat{p}_N\left(F_{\max_\rho}^{obs}/\rho\right)$ est la p-valeur de Monte Carlo pour un ρ fixé (équation 6.12) et α est le niveau de signification du test choisi à l'avance.

La décomposition des étapes du test est donnée au tableau 7.3. Ce test peut être appliqué au problème de détection de rupture brusque (les équations 7.17 et 7.3) de même que pour le problème de détection de rupture continue (équations 7.17 et 7.4).

% ***Le modèle considéré pour représenter le phénomène de rupture est celui***
% ***exprimé à l'équation 7.17***

Étape 1 : *Calcul de la valeur observée de la statistique de test,* $F_{\max_\rho}^{obs}$, *et de la date éventuelle de rupture,* $\hat{t}_b$, *à partir des données*

Pour $t_b = 3,\ldots,n-1$
  Calculer $F_{obs}(t_b)$, la valeur observée de la statistique $F(t_b)$, dans l'équation 7.6
t_b suivant
→ $F_{\max_\rho}^{obs} = \max_{3\le t_b \le n-1} \{F_{obs}(t_b)\}$

→ $\hat{t}_b = \arg \max_{3\le t_b \le n-1} \{F_{obs}(t_b)\}$

Étape 2 : Calcul des p-valeurs $\hat{p}_N\left(F_{\max_\rho}^{obs}/\rho\right)$ pour $\rho \in M_0$

Pour $\rho \in \mathrm{M}_0$
  Pour $i = 1,\ldots,N$ (*Génération de* N *statistique* $F_{\max_\rho}$ *sous* H_0)
    Pour $t_b = 3,\ldots,n-1$
      Générer $w \sim N(0,1)$
      Calculer $F_{res}(t_b)$, la réalisation de $F(t_b)$, dans les équations 7.17 et 7.15
    t_b suivant
      → $F_{\max_\rho}^{i} = \max_{3\le t_b \le n-1} \{F_{res}(t_b)\}$
  i suivant
  Calculer $\hat{p}_N\left(F_{\max_\rho}^{obs}/\rho\right)$ dans l'échantillon $\left\{F_{\max_\rho}^{obs}, F_{\max_\rho}^{1}, F_{\max_\rho}^{2}, \ldots, F_{\max_\rho}^{N}\right\}$
ρ suivant
→ $\left\{\rho \in \mathrm{M}_0 : \hat{p}_N\left(F_{\max_\rho}^{obs}/\rho\right)\right\}$

→ $\sup_{\rho\in M_0}\left\{\hat{p}_N\left(F_{\max_\rho}^{obs}/\rho\right)\right\}$

Étape 3 : *Décision*

Rejeter H_0 si $\sup_{\rho\in M_0}\left\{\hat{p}_N\left(F_{\max_\rho}^{obs}/\rho\right)\right\}$ est plus petit ou égale à α

Tableau 7.3 : Test de Monte Carlo de détection d'une rupture dans un échantillon avec persistance

7.1.2 Tests de Monte Carlo de détection de plusieurs ruptures

Un des problèmes inhérents aux tests de rupture est celui de détecter non pas une mais plusieurs ruptures susceptibles d'être présentes dans un échantillon de données. Ce problème est apparent en hydrométéorologie. Par exemple, si une rupture a déjà été détectée dans une réalisation d'un phénomène hydrométéorologique, on peut vouloir vérifier que les sous-échantillons avant et après la date de rupture sont stationnaires, ou encore examiner si une autre rupture existe dans les données.

7.1.2.1 Formulation du problème et approches proposées

La détection de plusieurs ruptures dans un échantillon de données est une généralisation du problème de détection d'une rupture présenté dans les sections précédentes. Il peut être représenté de la façon suivante : on observe dans le temps une certaine série de données x_t et on suppose que les t_1 premières observations de cette série sont régies par un modèle M_{θ_1}, les $t_2 - t_1$ observations suivantes par un modèle M_{θ_2}, etc, les instants d'observation $t_1, t_2, \ldots$ s'appellent dates de ruptures. Le problème revient à détecter ces dates de rupture. Pour résoudre ce problème, on propose une procédure séquentielle et un test de Monte Carlo de détection de ruptures.

7.1.2.2 Procédure séquentielle de détection de ruptures

7.1.2.2.1 Méthodologie

Pour détecter des ruptures dans une série de données, Vostrikova (1981) a proposé une méthode, bien connue sous le nom de *procédure de segmentation binaire*, et prouvé son efficacité. Cette méthode a le mérite de détecter le nombre de ruptures et leurs positions simultanées dans la série et de faire des économies en temps de calcul. Elle peut être décrite comme suit. On commence, à la première étape, par appliquer un test de détection de rupture dans la série originale. Si ce test ne détecte pas une rupture, la procédure de segmentation binaire s'arrête et l'hypothèse principale d'absence de rupture est acceptée.

Si par contre le test détecte une date de rupture, alors cette date divise la série de données en deux sous-échantillons. Pour chaque sous-échantillon ainsi constitué, on applique le test de détection d'une rupture, comme dans la première étape, et on continue le processus jusqu'à ce que le test ne détecte plus de rupture dans tous les sous-échantillons formés. Le critère d'arrêt est donc : d'une itération à la suivante, aucune rupture n'a été détectée par le test de détection dans les sous-échantillons formés, c'est-à-dire les sous-échantillons ainsi composés sont stationnaires.

Vostrikova (1981) a montré que le nombre de ruptures détectées avec cette méthode converge en probabilité vers le nombre de ruptures réelles de la série de données. Cette procédure est particulièrement efficace pour détecter plusieurs ruptures brusques dans une série de données (Vostrikova, 1981). En effet, si l'on veut approximer une fonction par étapes (avec plusieurs sauts) au sens de la norme L^2 par une fonction ayant deux valeurs (une fonction indicatrice par exemple), c'est-à-dire ayant seulement un saut, alors le point du saut de la meilleure approximation coïncide avec l'un des points de saut de la fonction à approximer. En règle générale, si l'on veut détecter plusieurs ruptures brusques de la moyenne dans un échantillon de données, la procédure de segmentation binaire est convergente. En revanche, dans le cas de la détection de ruptures avec continuité, le raisonnement précédent laisse présager que la procédure de Vostrikova (1981) sera moins performante.

La mise en œuvre de la méthode de Vostrikova (1981) comporte deux étapes. La première consiste à obtenir un test statistique qui permet de vérifier l'hypothèse d'absence de rupture dans une série de données. La seconde étape consiste à intégrer ce test dans l'algorithme de segmentation binaire. Ainsi, ayant un test de détection de rupture, il ne reste plus qu'à répéter le processus de segmentation pour chaque sous-échantillon jusqu'à ce que l'hypothèse principale d'absence de rupture soit acceptée. La procédure séquentielle de détection de rupture qu'on propose dans cette thèse est de coupler un test de Monte Carlo de détection d'une rupture (tableaux 7.1, 7.2 ou 7.3) avec l'algorithme de segmentation binaire de Vostrikova (1981).

7.1.2.2.2 Algorithme de la procédure séquentielle de détection de ruptures

Soit une série temporelle de n valeurs numériques : x_t, $t = 1,2,\ldots,n$. On considère, sans perte de généralité, le problème de détection de plusieurs ruptures brusques de la moyenne de la série. Ce problème revient à examiner l'hypothèse H_0 : «il n'y a pas de rupture brusque de moyenne dans la série x_t». Le principe de la méthode de Vostrikova (1981) est de "découper" la série originale x_t en k segments (sous-échantillons stationnaires) de sorte qu'on rejette H_0 dans tout segment. Autrement dit, la segmentation est effectuée de telle sorte que la moyenne calculée sur tout segment soit significativement différente de la moyenne du (ou des) segment(s) voisin(s).

La procédure séquentielle de détection qu'on propose pour examiner H_0 consiste à introduire le test Monte Carlo de détection d'une rupture (dans un échantillon avec ou sans persistance) dans l'algorithme de segmentation binaire de Vostrikova (1981). Ce test permet ainsi de vérifier la validité des segmentations concernant la série temporelle étudiée complète $x_t, t = 1,\ldots,n$. Cela dit, la segmentation retenue dans la procédure séquentielle qu'on propose doit être telle que pour un ordre l de segmentation donné, le test de Monte Carlo de détection d'une rupture (tableaux 7.1 ou 7.3) rejette H_0 au niveau de signification α dans toute série x_i, $i = i_k, i_{k-1}$ avec $i_{k-1} \geq 1$ et $i_k \leq n$ où $(i_{k-1} < i_k)$ constituant un segment de la série initiale $x_t, t = 1,\ldots,n$.

La solution fournie par l'algorithme de segmentation binaire est optimale si les segments sont relativement grands (Vostrikova, 1981). Il est donc important d'imposer un critère supplémentaire dans cet algorithme. Dans le cadre de notre procédure séquentielle, au cours de l'exploration de l'arborescence de segmentation de la série, on applique le test Monte Carlo seulement dans les segments dont la longueur est plus grande qu'une valeur h, fixée à l'avance. Autrement dit, si la longueur du $k$ième segment est telle que $n_k = i_k - i_{k-1}$ avec $i_0 = 1 < i_1 < \ldots < i_k < \ldots < i_l = n$, on applique le test Monte Carlo dans toute série x_i, $i = i_k, i_{k-1}$ avec $i_{k-1} \geq 1$ et $i_k \leq n$ où $(i_{k-1} < i_k)$ constituant un segment de la série initiale telle que $n_k \geq h$.

La décomposition des étapes de la procédure séquentielle de détection de ruptures qu'on propose dans cette thèse est donnée dans le tableau 7.4.

% Procédure séquentielle de détection de ruptures brusques dans une série temporelle % dont la longueur est supérieure à h
Étape 1 : Tester H_0 dans la série à l'aide du test de Monte Carlo (tableau 7.1 ou 7.3). Si le test ne rejette pas H_0, arrêt de la procédure. Il n'y a pas de rupture dans la série initiale. Si par contre le test rejette H_0, il y a une rupture à la date t_{b_1} et aller à l'étape 2.
Étape 2 : Appliquer le test de Monte Carlo (tableau 7.1 ou 7.3) dans les sous-échantillons avant et après la date de rupture détectée pourvu que leur longueur soit au moins égale à h .
Étape 3 : Répéter l'étape 2 jusqu'à ce que le test de Monte Carlo (tableau 7.1 ou 7.3) ne rejette plus H_0 dans les sous échantillons formés.
Étape 4 : Rejeter H_0 si le nombre de ruptures détectées aux étapes 1 à 3 est supérieur ou égal à 1. Les dates détectées sont données par $\{\hat{t}_{b_1},\ldots,\hat{t}_{b_q}\}$

Tableau 7.4 : Procédure séquentielle de détection de ruptures

Le critère d'arrêt de cette procédure séquentielle est : d'une itération à la suivante, aucune rupture n'a été détectée par le test de détection dans les sous-échantillons formés. Ainsi, pour l'exécution de la procédure, on n'a pas besoin de spécifier un nombre de ruptures a priori.

La procédure séquentielle de détection qu'on propose ici est appropriée à la recherche de multiples ruptures de moyenne dans une série. L'intérêt principal de cette procédure réside dans sa simplicité de mise en œuvre et dans l'existence de résultats de convergence de la

méthode de segmentation binaire de Vostrikova (1981). Cette procédure peut être vue comme un test de stationnarité où l'hypothèse principale H_0 affirme : «il n'y a pas de rupture brusque de moyenne dans la série x_t» tandis que la contre-hypothèse de non-stationnarité H_1 affirme : «il y a q ruptures brusques de moyenne localisées aux instants $t_{b_1},\ldots,t_{b_q}$». Si cette procédure ne produit pas un nombre de ruptures significatives supérieures ou égal à 1, l'hypothèse H_0 est acceptée. Toutefois, aucun niveau de signification n'est attribué à ce test.

7.1.2.3 Test de Monte Carlo de détection de plusieurs ruptures

Un problème important dans la procédure de détection de rupture présentée précédemment (section 7.1.2.2) est que le niveau de ce "test" est contrôlé seulement à chaque étape de la procédure. Il est donc difficile d'interpréter les résultats produits, c'est-à-dire d'attribuer un niveau de signification global aux résultats obtenus. Le test de Monte Carlo de détection de ruptures qui est proposé ici permet d'apporter une solution novatrice à ce problème.

7.1.2.3.1 Position du problème

En particularisant l'hypothèse principale d'absence de rupture H_0, la procédure séquentielle telle que décrite au tableau 7.4 peut être interprétée comme une procédure de test d'hypothèses multiples. En effet, d'après le tableau 7.4, on désire tester successivement plusieurs hypothèses : à l'étape 1, on commence par tester l'hypothèse principale d'absence de rupture H_{01} dans la série originale ; si H_{01} est refusé, on arrête la procédure ; sinon on teste les hypothèses d'absence de rupture H_{02} et H_{03} dans les sous-échantillons avant et après la date de rupture détectée et ainsi de suite, jusqu'à ce que le test ne rejette plus les hypothèses d'absence de rupture $H_{0k}, k = 2,\ldots,l$ avec $l \geq 3$.

Selon la théorie classique des tests d'hypothèses (appendice C), si la valeur de la statistique de test calculée sur l'échantillon dépasse un certain seuil, l'hypothèse testée (absence de rupture dans le cas présent) est rejetée et par suite l'hypothèse principale est acceptée. Ce

seuil de rejet dépend du risque de première espèce choisi par l'utilisateur du test : celui-ci " prend le risque " de rejeter à tort l'hypothèse testée avec une probabilité α (le niveau de signification du test). Or, lors de l'exécution de la procédure séquentielle (tableau 7.4), on applique plusieurs fois le de test Monte Carlo (tableau 7.1 ou 7.3), au cours de la même expérience de détection de rupture, dans différentes sections d'une même série de données. Du fait de ces tests multiples, le risque d'un rejet fallacieux de H_0 doit donc être raisonné, non pas pour chaque test individuel, mais au niveau global de la procédure séquentielle. Toutefois, même en tenant compte du nombre de fois qu'on applique le test Monte Carlo (tableau 7.1 ou 7.3) sur les différents segments pratiqués sur la série de données, le calcul du seuil de rejet global de H_0 n'est pas simple. En effet, les différents tests Monte Carlo réalisés au niveau des segments de l'arborescence de l'algorithme de segmentation ne sont pas indépendants entre eux car les données issues de ces segments sont liées. Ces différentes difficultés nous font dire que le niveau global de la procédure séquentielle de détection de ruptures (tableau 7.4) n'est pas contrôlé. Autrement dit, il n'est pas possible d'établir le seuil global de rejet de H_0 et être sûr du risque α réellement pris au cours de l'exécution de cette procédure.

7.1.2.3.2 Approche proposée

Pour contrôler le niveau global de la procédure séquentielle de détection de ruptures (tableau 7.4), on propose de combiner les statistiques de test de Monte Carlo obtenues au cours de l'exploration de l'arborescence des segmentations de la série de données complète à analyser. De manière équivalente cela revient aussi à combiner les différentes p-valeurs Monte Carlo obtenues avec ces statistiques. En combinant ces p-valeurs Monte Carlo, il est facile d'obtenir un test de H_0 : il suffit de rejeter H_0 lorsque au moins une des p-valeurs est inférieure au niveau de signification α fixé à l'avance. On appelle une telle procédure un *test combiné*.

Le problème des tests combinés peut être posé en ces termes : l statistiques de test T_k, $k = 1,\ldots,l$ sont disponibles pour éprouver les hypothèses principales H_{0k}, $k = 1,\ldots,l$, qui pourraient être de même nature. La question qui se pose est de savoir comment on peut

combiner ces différentes statistiques de tests pour arriver à un test de l'hypothèse globale H_0: "toutes les hypothèses H_{0k}, $k=1,\ldots,l$, sont vraies simultanément", tout en contrôlant le niveau de signification global α. Deux étapes principales sont nécessaires à la construction d'un test combiné :

1. le choix du critère de combinaison des tests ;
2. le contrôle du niveau global de signification.

Il existe une littérature abondante sur les tests combinés (Birnbaum, 1954 ; Holm, 1979 ; Folks, 1984). L'approche généralement admise consiste à utiliser les niveaux de signification (p-valeurs) des l tests comme une statistique de test combiné. Cette approche a été proposée originalement par Tippett (1931), Fisher (1932) et Wilkinson (1951), dans le cas de statistiques de test indépendantes. Sans perte de généralité, la p-valeur individuelle de chaque test servant à éprouver les hypothèses principales H_{0k}, $k=1,\ldots,l$, est donnée par :

$$p_k = \mathrm{P}\left(T_k \geq t_0^k / \mathrm{H}_{0k}\right) \tag{7.20}$$

où,

- t_0^k est la valeur observée de la statistique de test T_k

Si on suppose que les statistiques de test T_k, $k=1,\ldots,l$ sont indépendantes et continues, les p_k sont des variables uniformes, $U(0,1)$, indépendantes. Dans ce cas, le test proposé par Fisher (1932) conduit au niveau de signification combiné p_α :

$$p_a = \mathrm{P}\left(\chi_{2l}^2 \geq -2\sum_{k=1}^{l}\ln(p_k)\right) \tag{7.21}$$

Le test combiné de Fisher a déjà été utilisé dans la littérature hydrométéorologique (Bahadur, 1967 ; Littell et Folks, 1971 ; Hipel et McLeod, 1994). Une autre variante de la

statistique de test de Fisher, qui a été originalement proposée par Fisher (1932) et Pearson (1933), est de considérer le produit des p-valeurs. La statistique de test combiné est alors :

$$T_{\times} = \prod_{i=1}^{k} p_i \tag{7.22}$$

Ces deux procédures de test telles que proposées originellement par leurs auteurs ne sont valides que lorsque les tests combinés sont indépendants. Or, d'après ce qui a été dit à la section (7.1.2.3.1), cette condition d'indépendance est difficile à obtenir avec des données hydrométéorologiques. Pour contourner ce problème, on propose d'appliquer la technique de tests de Monte Carlo afin de résoudre le problème de combinaison des p-valeurs. Dans notre cas, cette approche consiste à appliquer la technique des tests Monte Carlo sur la statistique combinée de l'équation 7.22 de sorte à obtenir un test exact.

7.1.2.3.3 Test de Monte Carlo de détection de ruptures : statistique de test

La procédure séquentielle de détection de ruptures qui a été présentée au tableau (7.4) sert comme fondement au test de Monte Carlo que nous proposons dans cette thèse. Le principe est le suivant. En premier lieu, on applique la procédure séquentielle (tableau 7.4) basée sur la statistique de test Monte Carlo $F_{\max}$ (tableau 7.1 ou 7.3). Soient $F_{\max}^1,\ldots,F_{\max}^m$ les statistiques significatives retenues dans l'exécution de cette procédure séquentielle, et $\hat{p}_N\left(F_{\max}^k\right),\ k=1,\ldots,m$, leurs p-valeurs Monte Carlo respectives. À partir de ces dernières, on peut construire la statistique combinée

$$T_{mul} = 1 - \prod_{k=1}^{m} \hat{p}_N\left(F_{\max}^k\right) \tag{7.23}$$

Pour tester l'hypothèse principale d'absence de rupture H_0, on propose comme statistique de test, la statistique combinée T_{mul} (ainsi écrite pour avoir un test à droite). Il est important de mentionner que l'obtention des p-valeurs $\hat{p}_N\left(F_{\max}^k\right),\ k=1,\ldots,m$, comme

détaillé ci-haut, consiste à ce qu'on appelle un contrôle de niveau à la première étape. Il s'ensuit que pour réaliser un contrôle du niveau global de la procédure séquentielle, on applique la technique des tests Monte Carlo pour la statistique combinée T_{mul}, où pour chaque tirage d'échantillon simulé, les p-valeurs qui interviennent dans le calcul de la statistique simulée sont re-simulées. Bien évidemment, pour calculer les p-valeurs $\hat{p}_N\left(F_{\max}^k\right)$, $k=1,\ldots,m$, le modèle qui sert à générer les échantillons simulés est adapté à chaque cas, mais ceci n'affecte pas le caractère pivotal des statistiques $F_{\max}$ impliquées.

7.1.2.3.4 Test de Monte Carlo de détection de ruptures : région critique du test

La statistique de test T_{mul} de l'équation 7.23 est libre de paramètres de nuisance sous H_0, ce qui justifie l'application des tests de Monte Carlo. Ainsi, si T_{mul}^{obs} est la valeur observée de la statistique de test T_{mul}, on prend comme région de rejet de H_0, la région randomisée :

$$\hat{p}_N\left(T_{mul}^{obs}\right) \leq \alpha \qquad (7.24)$$

où,

- $\hat{p}_N\left(T_{mul}^{obs}\right)$ est la p-valeur de Monte Carlo (équation 6.12) et α est le niveau de signification du test fixé à l'avance.

Le déroulement des étapes d'élaboration du test Monte Carlo de détection de ruptures, dans le cas de ruptures brusques, est donné au tableau 7.5.

% Test Monte Carlo de détection de multiples ruptures brusques dans une série
% temporelle de n valeurs

Étape 1 : *Calcul de la valeur observée de la statistique de test,* T_{mul}^{obs}, *et des dates éventuelles de rupture,* $\{\hat{t}_{b_1},\ldots,\hat{t}_{b_m}\}$, *à partir des données observées*

Appliquer la procédure séquentielle (tableau 7.4) à la série temporelle à analyser. Soient $F_{\max}^1,\ldots,F_{\max}^m$ les statistiques significatives retenues dans l'exécution de cette procédure, et $\hat{p}_N\left(F_{\max}^k\right)$, $k=1,\ldots,m$, leurs p-valeurs Monte Carlo respectives

→ $T_{mul}^{obs}=1-\prod_{k=1}^{m}\hat{p}_N\left(F_{\max}^k\right)$

→ les dates de rupture sont données par : $\{\hat{t}_{b_1},\ldots,\hat{t}_{b_m}\}$

Étape 2 : *Génération de* N *statistiques* T_{mul} *sous* H_0

Pour $i=1,\ldots,N$

Générer, sous H_0 (à partir du modèle 7.2), une série de n valeurs.

Appliquer la procédure séquentielle (tableau 7.4) à cette série. Soient $F_{\max}^1,\ldots,F_{\max}^m$ les statistiques retenues dans l'exécution de cette procédure séquentielle, et $\hat{p}_N\left(F_{\max}^k\right)$, $k=1,\ldots,m$, leurs p-valeurs Monte Carlo respectives.

$T_{mul}^i=1-\prod_{k=1}^{m}\hat{p}_N\left(F_{\max}^k\right)$

i suivant

→ $\{T_{mul}^1,T_{mul}^2,\ldots,T_{mul}^N\}$

Étape 3 : *Calcul de la p-valeur à partir de l'échantillon* $\{T_{mul}^{obs},\ T_{mul}^1,T_{mul}^2,\ldots,T_{mul}^N\}$

Calculer la p-valeur Monte Carlo $\hat{p}_N$ dans l'équation 6.12

Étape 4 : *Décision*

Rejeter H_0 si $\hat{p}_N$ est plus petite ou égale à un niveau de signification α.

Tableau 7.5 : Test de Monte Carlo de détection de multiples ruptures

7.2 Test régional de Monte Carlo de détection de non-stationnarités hydrométéorologiques

7.2.1 Position du problème

À la section 7.1, on a proposé des tests de Monte Carlo pour détecter une(des) rupture(s) à un(des) instant(s) inconnu(s) dans une série de données. Ces tests concernent l'exploitation d'une série de données et une seule. L'utilisation de ces tests dans l'analyse d'une série de données d'une variable hydrométéorologique aura donc pour but d'améliorer la compréhension des mécanismes statistiques générateurs de cette série d'observations. Une telle analyse est qualifiée de ponctuelle, locale ou "par site".

Bien souvent, la compréhension des mécanismes statistiques générateurs de séries ne peut être atteinte en considérant une seule série de données puisque toute série de données d'une variable hydrométéorologique peut n'être qu'une représentation partielle d'un phénomène complexe générant un nombre substantiel de séries différentes. Par exemple, une modification climatique est un phénomène global qui se manifeste à une échelle au moins régional. Il est donc nécessaire de chercher à apporter une dimension régionale à la connaissance hydrométéorologique, en consolidant les observations faites en un site de l'espace par des observations faites sur un secteur géographique. Cette prise en compte spatiale des différentes observations donne une plus grande solidité aux conclusions hydrométéorologiques que l'on peut établir. En conséquence, elle améliore la compréhension des phénomènes générateurs de dites séries ponctuelles. Développer un test régional de stationnarité constitue donc un enjeu majeur dans la compréhension des phénomènes hydrométéorologiques.

7.2.2 Directions de recherche envisagées en hydrométéorologie

Lorsqu'on aborde la question de la non-stationnarité d'un phénomène hydrométéorologique, il est primordial d'envisager cette question dans un contexte régional. Il est en effet important d'exploiter la dimension spatiale du phénomène, c'est-à-

dire d'analyser cette question dans plusieurs séries d'observations à des sites soumis à ce même phénomène. Le problème de la cohérence régionale de non-stationnarité éventuelle a donc une incidence sur la prise de décision. Mais, comment transférer des informations à caractère plus ou moins disparate depuis les échelles ponctuelles ou "par site" vers l'échelle régionale où les facteurs du milieu sont généralement modifiés de façon non uniforme dans l'espace ? Construire un test régional conduit nécessairement à répondre à cette question.

Pour répondre à la question précédente, l'hydrologie statistique a été peu développée dans la perspective classique. Les méthodes régionales de détection de non-stationnarité qui existent dans la littérature reposent sur une analyse indépendante de séries, correspondant chacune à un site déterminé, en vue de procéder à des interprétations prenant en compte la dimension spatiale des phénomènes générateurs des dites séries ponctuelles. Par exemple, pour répondre à la question précédente, Westmacott et Burn (1997) et Gan (1998) tirent des conclusions globales en analysant indépendamment la non-stationnarité de chacune des séries. Ce type d'approche est contestable dans la mesure où la conclusion globale n'est pas basée sur une mesure qui combine l'information obtenue à l'échelle locale. En outre, la validité de la régionalisation est entachée à de multiples erreurs liées d'une part à la non formalisation des relations structurelles entre les caractéristiques des non-stationnarités à chaque site. D'autre part, les corrélations spatiales existant entre les observations d'un site à l'autre ne sont pas prises en compte. Dans la perspective bayésienne, la question précédente a longtemps été abordée en adoptant le point de vue de la statistique classique (Slivitzky et Mathier (1994) et Perreault et al. (1999)). Ce n'est que très récemment que Perreault (2000) a proposé une procédure formelle de détection d'une non-stationnarité régionale. Cette procédure est basée sur la construction des modèles multisite de rupture traduisant les manifestations régionales d'une rupture sur des séries d'observations correspondant chacune à un site déterminé. Grâce à cette procédure, il est désormais possible d'intégrer dans un modèle probabiliste les particularités du phénomène d'intérêt.

De notre point de vue une réponse à la question précédente pourrait être recherchée dans un transfert d'information d'un site à un autre par la construction d'une statistique

régionale qui combinerait l'information obtenue à l'échelle locale. Les enjeux de la construction de cette statistique sont donc importantes car, comme il a été exposé plus haut, un test statistique construit avec cette statistique doit guider un choix décisionnel sur des bases qu'on espère les plus objectives. C'est la raison pour laquelle on propose l'approche des tests combinés comme base de la méthodologie du test régional. L'idée centrale induite par notre approche privilégie l'exploitation "par site" des séries comme un préalable à toute interprétation prenant en compte la dimension spatiale des phénomènes responsables des dites séries de données. Par ce formalisme, on cherche à étendre l'approche classique, mais en cherchant à construire un test formel de détection de manifestations régionales d'une non-stationnarité sur un ensemble de séries.

7.2.3 Construction du test régional de stationnarité

7.2.3.1 Approche de modélisation préconisée

Le type de non-stationnarité envisagé ici est le problème de rupture. On suppose que l'étude de ce problème porte sur une région géographique qui concerne plusieurs sites de mesure. Soit $I = \{1,\ldots,k\}$ la collection de ces sites. Pour simplifier la présentation, on suppose qu'en chaque site i on dispose, au temps t, d'une mesure hydrométéorologique (débit par exemple) notée x_{it} qui est générée par une variable X_i dont la distribution de probabilité est $F(X_i/\theta_i)$ avec θ_i le paramètre caractérisant localement la loi de l'occurrence du phénomène hydrométéorologique. Pour présumer qu'il existe une certaine cohérence régionale entre les caractéristiques de rupture présente dans chaque série, on va supposer que les sites d'intérêt forment un ensemble hydrométéorologique homogène. Autrement dit, on admet que les durées d'observations en ces divers sites sont égales. Dans ce contexte, le modèle de rupture qu'on considère ici est l'extension naturelle du modèle 7.2 (pour une rupture brusque) au cas multidimensionnel :

$$x_{it} = \beta_{0i} + \beta_{1i}\, ID(t_{ib}) + \varepsilon_{it} \qquad \begin{matrix} t = 1,\ldots,n \\ i = 1,\ldots,k \end{matrix} \tag{7.25}$$

Ce modèle caractérise un ensemble de k "régressions empilées". Puisque, $X_i = X_j,\ i,j = 1,\ldots,k$, c'est un modèle linéaire multivarié (Anderson, 1984). Le système d'équations de l'équation 7.25 peut encore s'écrire sous la forme matricielle suivante :

$$\mathrm{X} = \mathrm{Y}\beta + \varepsilon \tag{7.26}$$

où,

- $\mathrm{X} = \begin{bmatrix} n\times 1\{Y_1 \\ n\times 1\{Y_2 \\ \vdots \\ \underbrace{n\times 1\{Y_k}_{kn\times 1} \end{bmatrix}$, $\varepsilon = \begin{bmatrix} n\times 1\{\varepsilon_1 \\ n\times 1\{\varepsilon_2 \\ \vdots \\ \underbrace{n\times 1\{\varepsilon_k}_{kn\times 1} \end{bmatrix}$, $\beta = \begin{bmatrix} 2\times 1\{(\beta_{01},\beta_{11})' \\ 2\times 1\{(\beta_{02},\beta_{12})' \\ \vdots \\ \underbrace{2\times 1\{(\beta_{0k},\beta_{1k})'}_{2k\times 1} \end{bmatrix}$

- $\mathrm{Y} = \begin{bmatrix} \underbrace{\begin{matrix} n\times 2\{[\mathbf{1} & ID(t_{1b})] \\ 0 & \end{matrix}}_{} & 0 & \cdots & 0 \\ 0 & \underbrace{\begin{matrix} n\times 2\{[\mathbf{1} & ID(t_{2b})] \end{matrix}}_{} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \underbrace{0}_{} & \underbrace{0}_{} & \cdots & \underbrace{n\times 2\{[\mathbf{1} \quad ID(t_{kb})]}_{} \end{bmatrix}_{kn\times 2k}$

Pour éviter de disperser nos efforts, on décide d'approcher le problème posé en ne prenant pas en compte les corrélations entre les équations (pour $i \neq j \quad E(\varepsilon_i \varepsilon_j') = 0.\mathbf{1}_n$). On considère également que les erreurs d'une même équation sont homoscédastiques ($E(\varepsilon_i \varepsilon_i') = \sigma^2.\mathbf{1}_n,\ i = 1,\ldots,k$). Ce choix est éclairé par le fait que le formalisme mathématique que nous développons dans la construction du test régional s'adapte facilement en plus d'être compatible avec ce qui a été fait à la section 7.1.

D'après le modèle posé, pour chaque équation du modèle exprimé à l'équation 7.25, l'hypothèse d'absence de rupture est donnée par :

$$H_{0i} : \beta_{1i} = 0 \qquad \text{pour} \quad i = 1,\ldots,k \tag{7.27}$$

Construire un test régional de stationnarité revient à construire un test qui vérifie l'hypothèse principale suivante : H_0 : "toutes les hypothèses H_{0i}, $i = 1,\ldots,k$ sont vraies simultanément".

L'hypothèse H_0 formalise qu'il existe une cohérence régionale entre les caractéristiques de rupture dans une zone homogène. On laisse donc ici la possibilité à la date de rupture de varier d'une série à l'autre tout en conservant la cohérence régionale.

7.2.3.2 Statistique du test régional

La question qui se pose dans la construction du test de l'hypothèse H_0 est de savoir comment on peut combiner les *k* statistiques de tests obtenues pour tester H_{0i} au niveau $\alpha_i = \alpha$, $i = 1,\ldots,k$ tout en contrôlant le niveau de signification global α. Tout en particularisant l'hypothèse principale H_0, on formule ce problème comme un problème de test combiné dont le principe a été présenté à la section 7.1.2.3.2. En s'inspirant de l'équation 7.22, si pour $i = 1,\ldots,k$, $F^i_{\max}$ est la statistique du test Monte Carlo (voir tableau 7.1) pour tester H_{0i} au niveau α et $\hat{p}_N\left(F^i_{\max}\right)$ est sa p-valeur Monte Carlo alors, la statistique de test pour tester H_0 est donnée par :

$$T_{reg} = 1 - \prod_{i=1}^{k} \hat{p}_N\left(F^i_{\max}\right) \tag{7.28}$$

D'après les résultats obtenus à la section 7.1.1.2, il est facile de voir que les statistiques de test $F^i_{\max}$ $i = 1,\ldots,k$ sont conjointement pivotales (car pour $i \neq j$ $E\left(\varepsilon_i \varepsilon'_j\right) = 0.\mathbf{1}_n$ de même, $E\left(\varepsilon_i \varepsilon'_i\right) = \sigma^2.\mathbf{1}_n$, $i = 1,\ldots,k$). Il en résulte qu'il est possible d'appliquer la technique des tests de Monte Carlo pour cette statistique combinée, où pour chaque tirage

d'échantillon simulé, les k p-valeurs qui interviennent dans le calcul de la statistique simulée sont re-simulées.

7.2.3.3 Région critique du test régional

D'après ce qui précède, la statistique de test T_{reg} de l'équation 7.28 est libre de paramètres de nuisance sous H_0, ce qui justifie l'application des tests de Monte Carlo. Si T_{reg}^{obs} est la valeur observée de la statistique de test T_{reg}, on prend comme région de rejet de H_0, la région randomisée :

$$\hat{p}_N\left(T_{reg}^{obs}\right) \leq \alpha \tag{7.29}$$

où,

- $\hat{p}_N\left(T_{reg}^{obs}\right)$ est la p-valeur de Monte Carlo (équation 6.12) et α est le niveau de signification du test fixé à l'avance.

L'algorithme décrivant les étapes relatives à l'application du test régional est donné au tableau 7.6.

% Test régional de Monte Carlo de détection de rupture brusque dans une série
% temporelle de n valeurs

Étape 1 : *Calcul de la valeur observée de la statistique de test,* T_{reg}^{obs} *, éventuelles de rupture, à partir des données observées*

Appliquer indépendamment le test Monte Carlo de détection d'une rupture brusque (tableau 7.1) aux séries correspondant chacune à un site déterminé.
Soient $F_{\max}^1,\ldots,F_{\max}^k$ les statistiques de tests obtenues, et $\hat{p}_N\left(F_{\max}^i\right)$, $i=1,\ldots,k$, leurs p-valeurs Monte Carlo respectives

$\rightarrow \quad T_{reg}^{obs} = 1-\prod_{i=1}^{k}\hat{p}_N\left(F_{\max}^i\right)$

Étape 2 : *Génération de* N *statistiques* T_{reg} *sous* H_0

Pour $j=1,\ldots,N$
  Pour $i=1,\ldots,k$
    Générer, sous H_0 (à partir du modèle 7.2), une série de n valeurs.
    Appliquer le test Monte Carlo (tableau 7.1) à cette série.
    Soit $F_{\max}^i$ la statistique de test obtenue et $\hat{p}_N\left(F_{\max}^i\right)$ sa p-valeur Monte Carlo
  i suivant
    $\rightarrow T_{reg}^i = 1-\prod_{i=1}^{k}\hat{p}_N\left(F_{\max}^i\right)$
j suivant
  $\rightarrow \quad \left\{T_{reg}^1, T_{reg}^2, \ldots, T_{reg}^N\right\}$

Étape 3 : *Calcul de la p-valeur à partir de l'échantillon* $\left\{T_{reg}^{obs},\ T_{reg}^1, T_{reg}^2,\ldots,T_{reg}^N\right\}$

Calculer la p-valeur Monte Carlo $\hat{p}_N$ dans l'équation 6.12

Étape 4 : *Décision*

Rejeter H_0 si $\hat{p}_N$ est plus petite ou égale à un niveau de signification α.

Tableau 7.6 : Test régional de Monte Carlo de détection d'une rupture

7.3 Conclusion

Dans ce chapitre, on a défini le cadre général dans lequel l'originalité de cette thèse s'inscrit. On veut en conclusion réitérer certains points importants. Premièrement, on a insisté sur la nécessité d'augmenter la classe de tests considérée jusqu'à maintenant pour détecter une (des) rupture (s) dans une série de données. On pensait plus particulièrement à une classe de tests qui permettrait de considérer les caractéristiques que l'on trouve habituellement dans les séries de données hydrométéorologiques. Par exemple, Lubès et al. (1998) ont montré que la persistance et la présence de tendance dans les séries sont les deux caractéristiques qui affectent le plus les performances des tests de détection d'une rupture qui sont habituellement employés en hydrométéorologie. Ce qu'on apporte par rapport à ce point est :

1. introduction d'un test de Monte Carlo de détection d'une rupture (brusque ou avec continuité) dans un échantillon sans persistance ;
2. introduction d'un test de Monte Carlo de détection d'une rupture (brusque ou avec continuité) dans un échantillon avec persistance ;
3. introduction d'une procédure séquentielle de détection de plusieurs ruptures brusque dans un échantillon avec persistance ;
4. introduction d'une procédure séquentielle de détection de plusieurs ruptures brusques dans un échantillon sans persistance ;
5. introduction d'un test de Monte Carlo de détection de plusieurs ruptures brusques dans un échantillon avec persistance ;
6. introduction d'un test de Monte Carlo de détection de plusieurs ruptures brusques dans un échantillon sans persistance ;

On a aussi abordé le problème de la quasi-absence d'un test régional de détection de rupture en hydrométéorologie. En effet, un test régional peut s'avérer utile dans la compréhension des phénomènes hydrométéorologiques. Par exemple, l'étude et la détection des modifications climatiques, phénomènes qui se produisent en général à grande échelle, doivent se faire sur une base régionale rigoureuse plutôt que sur la base de séries de données provenant de différents sites étudiés indépendamment les unes des autres

(Ouarda et al., 1999). Notre contribution par rapport à ce point est l'introduction dans la littérature hydrométéorologique d'un test régional de détection de rupture. Ce test ne tient pas compte des corrélations spatiales pouvant exister entre les observations des sites d'intérêt.

En fin de compte, pour les tests qui sont proposés dans cette thèse, l'hypothèse de normalité des erreurs que l'on fait habituellement dans un modèle de régression classique n'est pas nécessaire (conséquence de l'équation 7.15). Il suffit juste de spécifier une loi connue sur les erreurs.

8 ÉTUDE DE PERFORMANCE PAR SIMULATION

Les résultats théoriques exposés sur les tests de Monte Carlo supposent que ces derniers sont exacts. Afin d'évaluer dans quelle mesure les tests de Monte Carlo de stationnarité dépendent de ce postulat, on effectue dans ce chapitre des expériences de simulations de type Monte Carlo pour explorer le comportement de ces tests. Le but d'une telle approche permet aussi d'évaluer, dans des situations contrôlées, la performance de ces tests. La présente étude s'attache donc à donner des estimations du niveau, de la puissance et de la robustesse des principaux tests de Monte Carlo qui sont proposés dans cette thèse. L'intérêt de cette étude est de faire une comparaison rigoureuse des résultats en terme de puissance et de robustesse du test de Monte Carlo de détection d'une rupture avec d'autres tests largement employés dans les études de rupture en hydrométéorologie : le test de Pettitt, le test de Buishand et le test de Worsley.

8.1 Méthodologie de simulation

L'étude du comportement des tests proposés sera menée au moyen d'expériences Monte Carlo. Pour effectuer une expérience Monte Carlo, on doit préciser le processus générateur des données simulées. Ensuite, ce processus est mis en œuvre sur ordinateur par l'emploi d'un générateur de nombres pseudo-aléatoires. Ce générateur permet de faire un très grand nombre de simulations, qui auront la propriété essentielle d'avoir les mêmes caractéristiques que le même nombre de tirages indépendants d'un processus réellement aléatoire. Il est donc possible d'étudier ce qui se passerait en «moyenne» si l'on disposait de données générées par le processus choisi par l'expérimentateur.

8.1.1 Spécifications pour l'étude du niveau et de la puissance

Pour étudier le niveau des principaux tests proposés dans cette thèse, on va générer sous l'hypothèse principale de stationnarité H_0 (absence de rupture) des séries artificielles. Ces séries seront, dans certains cas, transformées afin de contenir des caractéristiques

restrictives comme la persistance (autocorrélation). Chaque série est ensuite soumise aux tests avec un risque de première espèce α (risque de rejeter H_0 à tort). En théorie, si le générateur respecte le caractère aléatoire des séries qu'il a produit, la proportion de rejets de H_0 de chacun des tests devrait correspondre à l'erreur de première espèce. Autrement dit, le niveau empirique, $\hat{\alpha}$, d'un test est correctement estimé par la proportion de rejets de H_0 :

$$\hat{\alpha} = \frac{\text{nombre de rejets de } H_0}{\text{nombre de séries simulées}} \tag{8.1}$$

Dans l'étude de la puissance des tests, on va répéter le même exercice que précédemment, mais cette fois-ci en générant des séries sous la contre-hypothèse H_1 de présence d'une ou plusieurs dates de ruptures (séries non-stationnaires). Chaque série est ensuite soumise aux tests avec un risque de première espèce α (risque de rejeter H_0 à tort) et un risque de deuxième espèce β (risque de rejeter H_1 à tort). La puissance d'un test, $(1-\beta)$, correspond à sa capacité de rejeter H_0. La puissance empirique d'un test donné, $(1-\hat{\beta})$, sera estimée par la proportion de rejets de H_0 :

$$\begin{cases} \hat{\beta} = \dfrac{\text{nombre d'acceptations de } H_0}{\text{nombre de séries simulées}} \\ (1-\hat{\beta}) = \dfrac{(\text{nombre de séries simulées}) - (\text{nombre d'acceptations de } H_0)}{\text{nombre de séries simulées}} \\ \qquad = \dfrac{\text{nombre de rejets de } H_0}{\text{nombre de séries simulées}} \end{cases} \tag{8.2}$$

Bien souvent, on est plus intéressé à examiner la proportion de rejets de H_0 associés à une estimation correcte de la date de rupture. Pour ce faire, on va réitérer l'exercice précédent. La proportion de rejets de H_0 associés à l'estimation correcte de la date de rupture, $(1-\hat{\beta})_{\text{det}}$, sera correctement estimée par le pourcentage de séries non-stationnaires

simulées pour lesquelles ce test rejette H_0 et estime correctement la date (ou les dates) de rupture simulée (s).

Afin de porter un jugement convenable sur les estimations du niveau empirique (respectivement puissance empirique) des tests, il est judicieux de connaître la précision de ces estimations. Pour ce faire, on considère que sur chaque lot de M séries générés par simulations de Monte Carlo, la proportion de rejets de H_0 de chacun des tests devrait correspondre au niveau théorique du test (ou encore le risque de première espèce du test). En pratique, pour tenir compte des fluctuations d'échantillonnage, chaque série générée est considéré comme une épreuve indépendante au cours de laquelle la statistique du test donné a une probabilité p égale au niveau de signification du test, c'est-à-dire au risque de première espèce du test d'être hors de l'intervalle d'acceptation de H_0. Pour chaque lot de M séries générées et pour un test particulier, il en résulte que la probabilité que k séries générées rejettent H_0, $P(k)$, est représentée par une distribution binomiale $B(M,k,p)$: $P(k)=C_M^k p^k (1-p)^{M-k}$. L'estimation de la variance du niveau empirique de ce test, $Var(\hat{\alpha})$, est équivalente à l'estimation de la variance d'une loi binomiale :

$$Var(\hat{\alpha})=\frac{\hat{\alpha}(1-\hat{\alpha})}{M} \tag{8.3}$$

En adoptant un raisonnement similaire, il vient que l'estimation de la variance de la puissance empirique de ce test, $Var\{(1-\hat{\beta})\}$, est

$$Var\{(1-\hat{\beta})\}=\frac{\hat{\beta}(1-\hat{\beta})}{M} \tag{8.4}$$

8.1.2 Spécifications générales

Chaque générateur de nombres aléatoires est utilisé pour produire $M=1000$ échantillons, sous une hypothèse spécifique (H_0 pour l'étude de la taille et H_1 pour l'étude de puissance), pour des effectifs couramment observés sur des séries hydrologiques annuelles

$n = 30, 50$ ou 100 . Chaque génération d'échantillon est initialisée avec un état initial différent obtenu à partir de l'horloge interne de l'ordinateur. Pour réaliser ces différentes opérations, on utilise le logiciel MatLab (Version 5.2.0.3084 on PCWIN).

Étant donné qu'on explore la performance des tests en fonction de plusieurs conditions : la taille n, le niveau de signification théorique α, l'autocorrélation ρ ou encore la position des dates de rupture t_b, il est nécessaire d'utiliser un plan d'échantillonnage. Par exemple, dans la situation où on étudie le niveau des tests en fonction de la taille n et du niveau de signification théorique α, ce plan d'échantillonnage consiste à répéter les étapes suivantes : (1) on choisit de façon aléatoire une valeur de n et une valeur de α ; (2) on soumet les tests avec un risque de première espèce α à chacune des M séries d'effectif n générée par simulation Monte Carlo. Cette attribution aléatoire a pour but d'empêcher que les résultats ne soient influencés de façon systématiques par des facteurs inconnus ou non parfaitement contrôlés. De cette façon, on pense pouvoir établir des critères de comparaison et de qualité adéquats.

En fin de compte, pour réaliser ces expériences de simulations, on s'est inspiré de certains travaux qui sont susceptibles d'intéresser les hydrologues confrontés à des problèmes d'évaluations d'outils statistiques. Il s'agit entre autres des travaux de El-Shaarawi et Damsleth (1988), Laberge (1995) et Lubes et al. (1998).

8.1.3 Spécifications sur les tests

Pour le test de Monte Carlo de détection d'une rupture (tableaux 7.1, 7.2 ou 7.3), on utilise un échantillon de $N = 99$ simulations de la statistique de test. Dans le cas particulier du test de Monte Carlo de détection d'une rupture dans un échantillon avec persistance (tableau 7.3), on pose $M_0 = \{0.1, 0.2, \ldots, 0.9\}$. Pour le test de Monte Carlo de détection de plusieurs ruptures (tableau 7.5), on utilise un échantillon de $N = 19$ simulations de la statistique de test.

8.2 Étude de la performance du test de Monte Carlo dans un échantillon sans persistance

Dans cette section, on explore la performance du test de Monte Carlo de détection d'une rupture qui a été proposé dans cette thèse. Trois cas nous intéressent. Premièrement, celui où l'on cherche à comparer son niveau à celui des tests habituellement employés dans la littérature hydrométéorologique. Deuxièmement, celui où l'on cherche à comparer sa puissance à celui de ces mêmes tests. Troisièmement celui où l'on cherche à connaître sa robustesse, c'est-à-dire s'il est performant pour des échantillons contenant une persistance (situation habituellement rencontrée en hydrométéorologie). Pour des raisons de commodité, on va adopter la notation suivante :

MC_1rup : test de Monte Carlo détection d'une rupture ;

Buishand : test de Buishand ;

Worsley : test de Worsley ;

Pettitt : test de Pettitt

Ces quatre tests sont détaillés à la section 5.2.2.1.1

8.2.1 Étude comparative de niveau : Test de Monte Carlo versus test de Pettitt, test de Buishand et Test de Worsley

On procède ici à une étude comparative de niveau du test de Monte Carlo de détection d'une rupture et des tests largement employés dans des études de rupture en hydrométéorologie (section 5.2.2.1.1) : le test de Pettitt, le test de Buishand et le test de Worsley. On rappelle que le test de Monte Carlo dont il s'agit ici est décrit au tableau 7.1.

8.2.1.1 Modèle de simulation de séries stationnaires et indépendantes

La présente étude nécessite de générer des séries artificielles sous l'hypothèse principale d'absence de rupture H_0. Le modèle de génération s'écrit :

$$x_t = \mu_0 + \varepsilon_t \tag{8.5}$$

où, $\mu_0 = 181$ (moyenne régionale des séries réelles du tableau 9.3) et les erreurs $\varepsilon_t \sim \overset{iid}{N}(0,1)$. Les méthodes de simulation Monte Carlo sont utilisées pour générer $M = 1000$ échantillons de $n = 30, 50$ ou 100 valeurs. Chaque série est soumise aux tests MC_1rup , Buishand, Worsley et Pettitt pour tester H_0 au niveau de signification théorique $\alpha = 1\%, 5\%$ ou 10%. Ce choix est bien évidemment arbitraire. On se réfère ici à l'emploi qui est fait habituellement dans la littérature $\alpha = 0.10$ (résultat significatif), $\alpha = 0.05$ (résultat hautement significatif) ou $\alpha = 0.01$ (résultat très hautement significatif).

8.2.1.2 Analyse des résultats

Les résultats relatifs à l'étude comparative de niveau sont rassemblés dans le tableau A1 de l'annexe A et dans la figure 8.1. Pour les trois tailles d'échantillons retenues, ces résultats déclarent performants les tests de Pettitt, de Buishand et de Monte Carlo (MC_1rup). En effet, le niveau empirique de chacun de ces tests est, aux fluctuations d'échantillonnage près, conforme aux niveaux de signification théoriques $\alpha = 1\%, 5\%$ et 10%. Ces tests ne présentent donc aucune distorsion de niveau. Par contre le test de Worsley présente un léger problème de distorsion de niveau pour les trois niveaux théoriques $\alpha = 1\%, 5\%$ ou 10%. Les résultats du tableau A1 démontrent aussi que les niveaux de signification estimés diminuent à mesure que la taille d'échantillon augmente. On serait donc en droit de penser que la performance des tests étudiés est meilleure pour des échantillons de grande taille. En somme, les conditions sont réunies pour effectuer une étude de puissance des tests.

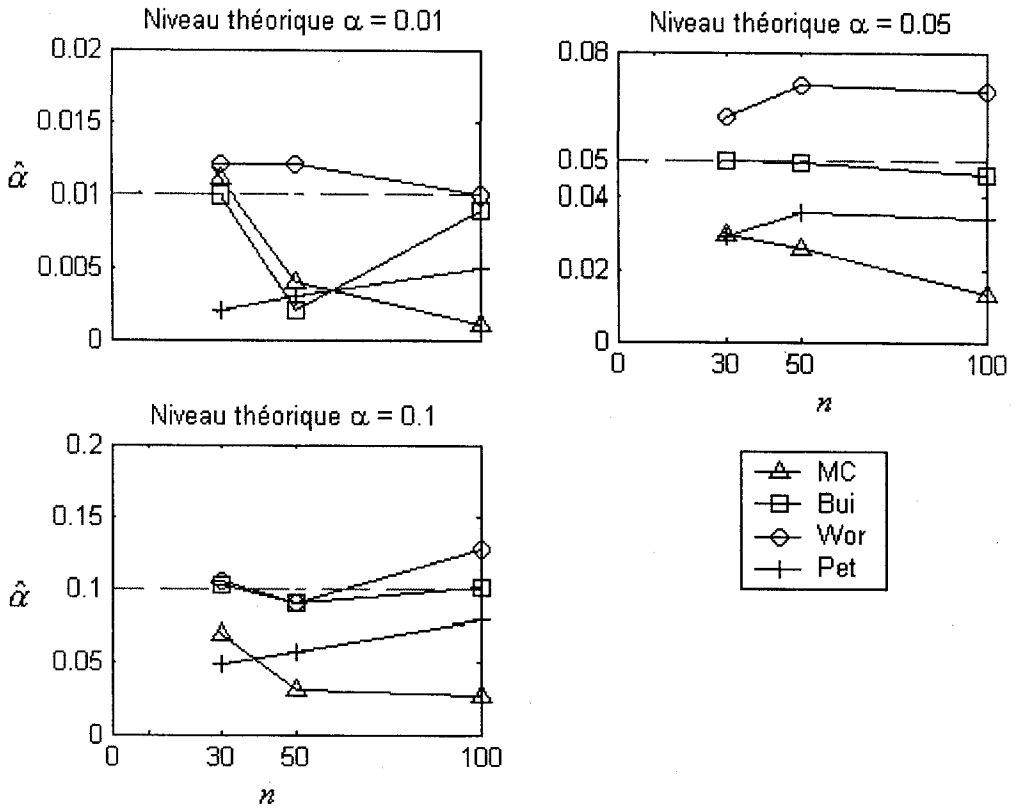


Figure 8.1 : Étude comparative du niveau dans une série sans persistance

8.2.2 Étude comparative de puissance : Test de Monte Carlo versus test de Pettitt, test de Buishand et Test de Worsley

La comparaison entre les quatre tests est menée cette fois-ci en considérant leur puissance à détecter une rupture dans un échantillon de données. Cette étude est réalisée à deux niveaux. Dans une première partie, la puissance de détection d'une rupture brusque de chaque test est évaluée et comparée. Dans une deuxième partie, c'est la puissance de détection d'une rupture avec continuité de chaque test qui est évaluée et comparée.

L'objectif de cette étude est d'évaluer la performance des tests à détecter une rupture située au milieu ou aux extrémités de l'échantillon de données. La question d'intérêt est de savoir si les tests sont performants pour estimer une date de rupture survenant au milieu ou aux bords de l'échantillon, et ce pour différentes amplitudes (respectivement, différentes

pentes) de la rupture. Pour estimer la puissance d'un test, l'étude de simulation nécessite de générer des séries artificielles sous la contre-hypothèse H_1 de la présence de rupture à la date t_b. Compte tenu de ce qui précède, on considère trois dates de rupture : $t_b = [n/3]$ (pour une rupture survenue au début de l'échantillon), $t_b = [n/2]$ (pour une rupture survenue au milieu de l'échantillon) et $t_b = [2n/3]$ (pour une rupture survenue à la fin de l'échantillon

8.2.2.1 Comportement des tests dans la détection d'une rupture brusque de la moyenne

Les études de simulation sont menées ici pour différentes amplitudes de rupture Δ. Les valeurs de Δ représentent respectivement un taux de rupture de 0.5, 1, 1.5, 2 et 2.5% de la moyenne $\mu_0 = 181$ (moyenne régionale des séries réelles du tableau 9.3).

8.2.2.1.1 Modèle de simulation de séries non-stationnaires et indépendantes : cas de la rupture brusque

Les séries artificielles sont générées sous la contre hypothèse H_1 de présence de rupture d'amplitude Δ à la date t_b par le processus

$$x_t = \begin{cases} \mu_0 + \varepsilon_t & t = 1,\ldots,t_b \\ (\mu_0 + \Delta) + \varepsilon_t & t = t_b + 1,\ldots,n \end{cases} \qquad (8.6)$$

où, $\mu_0 = 181$ et les erreurs $\varepsilon_t \sim^{iid} N(0,1)$. Pour les valeurs de Δ et de t_b, on génère $M = 1000$ échantillons de $n = 30, 50$ ou 100 valeurs. Chaque série est soumise aux tests MC_1rup, Buishand, Worsley et Pettitt pour tester H_0 au niveau de signification théorique $\alpha = 1\%, 5\%$ ou 10%.

8.2.2.1.1 Analyse des résultats

Les tableaux A2 à A19 de l'annexe A présentent les résultats obtenus pour l'analyse de la puissance des tests à détecter une rupture brusque de la moyenne : A2 à A9 pour l'estimation de la puissance $(1-\hat{\beta})$ et A10 à A19 pour l'estimation de la proportion de rejets de H_0 associés à une estimation correcte de la date de rupture, $(1-\hat{\beta})_{det}$.

Dans l'ensemble, les estimations obtenues sont de bonne qualité. En présence d'une rupture brusque de la moyenne, il apparaît que la puissance des tests augmente avec l'amplitude de la rupture et avec la taille de l'échantillon. Cette performance est d'autant meilleure lorsque la rupture survient au milieu de la série. D'autre part, la date de rupture est estimée dans une faible proportion de cas si l'amplitude de la rupture est faible même si elle est proche du milieu de l'échantillon. Par contre, les quatre tests sont performants pour des amplitudes suffisamment fortes même si la rupture a lieu au début ou vers la fin de la série. Toutefois, dans ce dernier cas, le test de Monte Carlo s'avère plus performant que les tests habituellement utilisés dans la littérature hydrométéorologique.

En fin de compte, le test de Monte Carlo (MC_1rup) paraît, d'après les figures 8.2 à 8.4 et les tableaux A10 à A19, plus puissant que les tests de Pettitt, de Buishand et de Worsley lorsque la rupture a lieu au début ou vers la fin de l'échantillon. Par exemple, pour une rupture d'amplitude $\Delta = 2.5\%$ simulée au début d'une série de taille $n = 100$ ($t_b \approx [n/3]$), le test de Monte Carlo (MC_1rup) estime la vraie date de rupture dans 97% des fois et les tests de Buishand et Pettitt près de 93% et 69% (tableau A11 de l'annexe A).

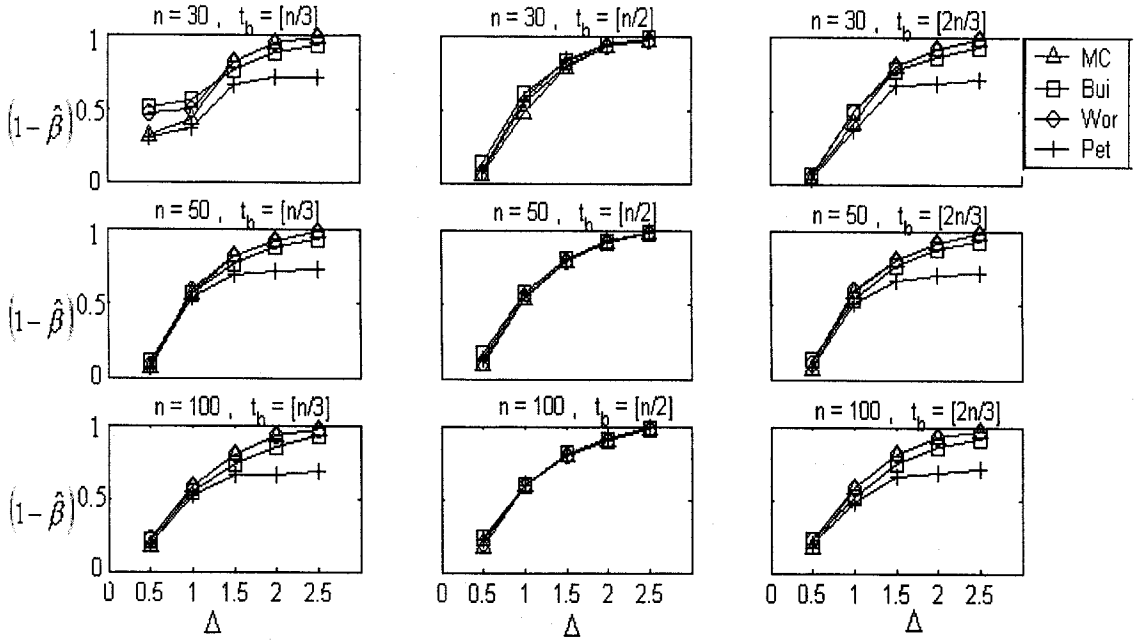


Figure 8.2 : Étude comparative de puissance de détection dans une série sans persistance ($\alpha = 0.01$)

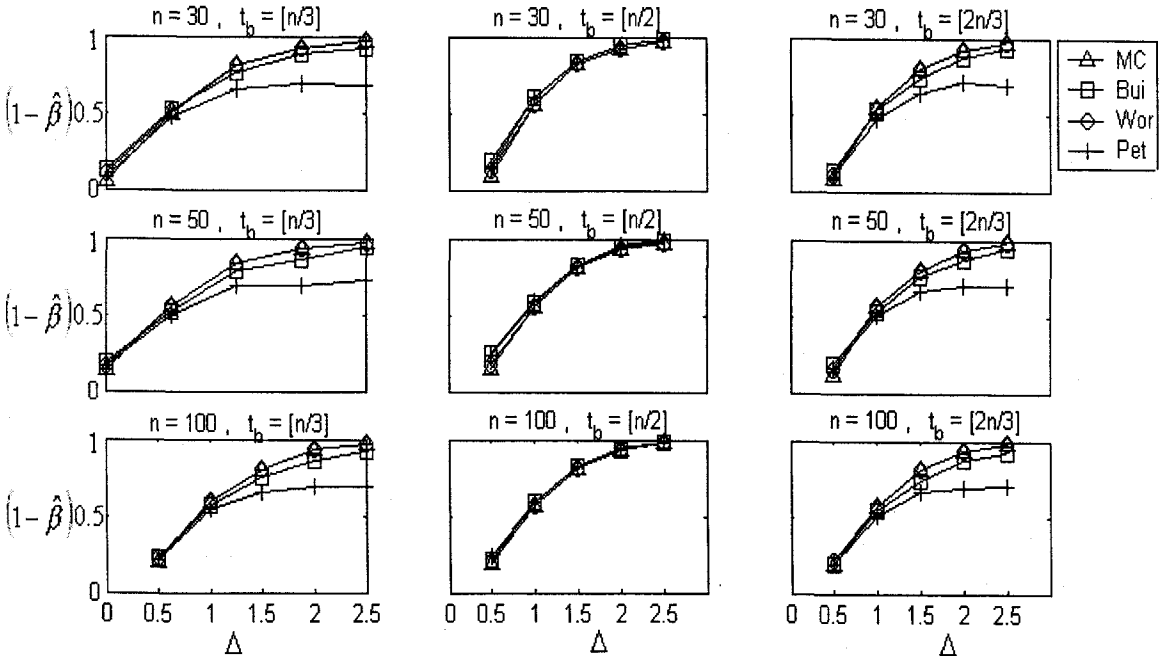


Figure 8.3 : Étude comparative de puissance de détection dans une série sans persistance ($\alpha = 0.05$)

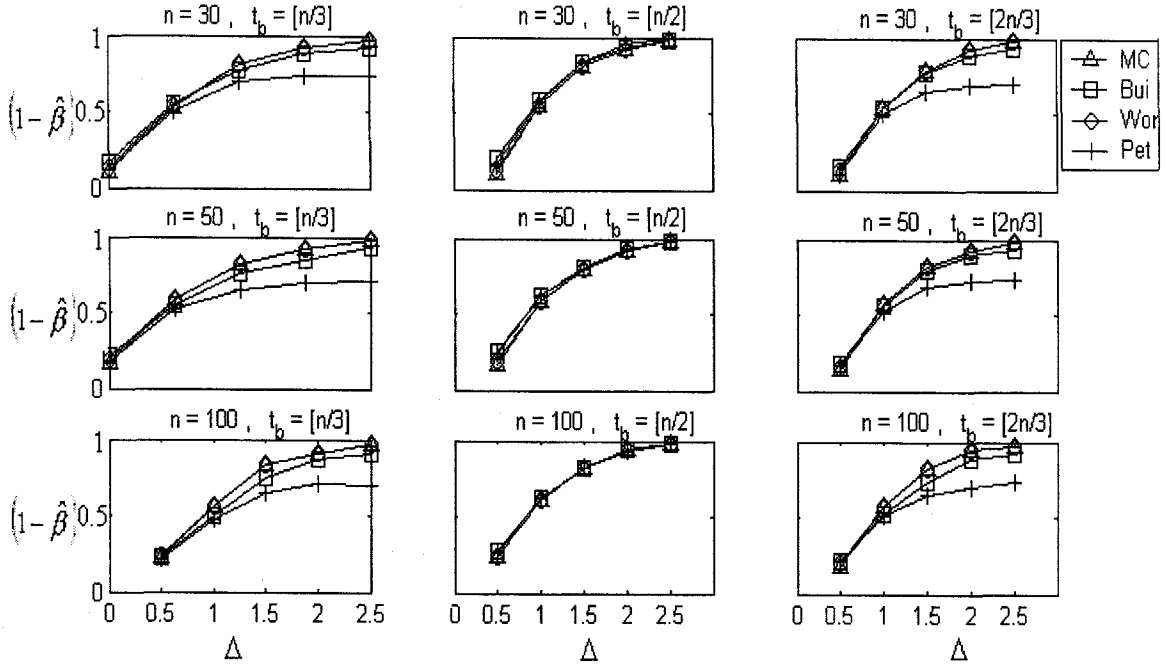


Figure 8.4 : Étude comparative de puissance de détection dans une série sans persistance ($\alpha = 0.10$)

8.2.2.2 Comportement des tests dans la détection d'une rupture avec continuité de la moyenne

Les études de simulations sont menées ici pour différentes pentes δ de la rupture avec continuité. Les valeurs de δ considérées sont de 0.1, 0.2, 0.3 et 0.4.

8.2.2.2.1 Modèle de simulation de séries non-stationnaires et indépendantes : cas de la rupture avec continuité

Les séries artificielles sont générées sous la contre-hypothèse de présence de rupture avec continuité de pente δ à la date t_b par le processus

$$x_t = \mu_0 + \delta\, ID(t_b) + \varepsilon_t \tag{8.7}$$

où,

$$ID(t_b) = \begin{cases} 0 & t = 1,\ldots,t_b \\ (t - t_b) & t = t_b + 1,\ldots,n \end{cases} \tag{8.8}$$

avec $\mu_0 = 181$ (moyenne régionale des séries réelles du tableau 9.3) et les erreurs $\varepsilon_t \sim \overset{iid}{N}(0,1)$. Pour les valeurs de δ = 0.1, 0.2, 0.3 ou 0.4 et les valeurs de t_b les méthodes de simulation Monte Carlo sont utilisées pour générer $M = 1000$ échantillons de $n = 30, 50$ ou 100 valeurs. Chaque série est soumise aux tests MC_1rup (voir tableau 7.2 et les spécifications données à la section 8.1.3), Buishand, Worsley et Pettitt pour tester H_0 au niveau de signification théorique $\alpha = 5\%$.

8.2.2.2.2 Analyse des résultats

Les résultats obtenus sont présentés dans les tableaux B1 à B6 de l'annexe B : B1 à B3 pour l'estimation de la puissance $(1 - \hat{\beta})$ et B4 à B6 pour l'estimation de la proportion de rejets de H_0 associés à une estimation correcte de la date de rupture, $(1 - \hat{\beta})_{det}$.

Pour le niveau de signification $\alpha = 5\%$ retenu, la qualité des estimations obtenues est satisfaisante. En faisant une lecture attentive des tableaux, on remarque que les tests ont une bonne puissance même si la pente de la rupture est faible ou modérée. Cette puissance s'accroît avec la taille de l'échantillon. Les choses changent cependant lorsqu'on examine l'estimation de la proportion de rejets de H_0 associés à une estimation correcte de la date de rupture. En effet, d'après les tableaux B4, B5 et B6 les tests ont de la difficulté à estimer la date de rupture (même si celle-ci a lieu au milieu de l'échantillon) et cette difficulté est d'autant plus sévère pour les tests habituellement employés en hydrométéorologie. Par exemple, lorsqu'une rupture a lieu au début ou au milieu de l'échantillon, pour une taille d'échantillon n = 100, le test de Monte Carlo estime la date de rupture près de 49% des fois et les tests classiques près de 0% de fois (excepté le test de Pettitt qui atteint 1% de fois en milieu de l'échantillon). Les choses ne semblent guère s'améliorer lorsque la rupture a lieu à la fin de l'échantillon. Toutefois, on observe une nette amélioration du test de Buishand et du test de Pettitt.

Plusieurs enseignements peuvent être tirés de ces premiers résultats. Il apparaît tout d'abord que sur l'ensemble des cas étudiés, le test de Monte Carlo, le test de Pettitt et le test de Buishand ne présentent pas de problème de distorsion de niveau. Par contre, malgré le fait que les résultats obtenus doivent être interprétés avec prudence compte tenu d'artefacts dus à l'échantillonnage, il apparaît clair que le test de Worsley présente de légers problèmes de distorsion de niveau. Ce test, lorsque appliqué à une série de données hydrométéorologiques, peut concourir à sur-rejeter l'hypothèse principale d'absence de rupture.

Cette étude a ensuite abordé le problème de la puissance des quatre tests étudiés. Au regard des résultats obtenus, on peut retenir que les tests étudiés sont puissants pour détecter une rupture brusque survenue au milieu de l'échantillon. Cette conclusion est conforme aux résultats obtenus par Lubès et al. (1998) à propos des tests de Buishand et de Pettitt.

En fin de compte, le test de Monte Carlo s'avère séduisant à maints égards : grâce à ces études de simulations, on a vérifié que c'est un test exact. On sait qu'il est plus puissant

pour détecter une rupture brusque de la moyenne survenue au début ou à la fin de l'échantillon. On sait également qu'il est plus puissant pour détecter une rupture avec continuité de la moyenne. Par ailleurs, on a aussi observé que les proportions de rejets de l'hypothèse principale associés à une estimation correcte de la date de rupture brusque (respectivement avec continuité) sont élevées (respectivement relativement élevée). Le revers de la médaille réside certainement dans le temps de calcul nécessaire pour le calcul de la statistique de test. Quoi qu'il en soit, étant donné la puissance des ordinateurs actuels, si on devait choisir une méthode, on aurait tendance à recommander le test de Monte Carlo en raison de ses performances.

8.2.3 Étude de robustesse : Test de Monte Carlo versus test de Pettitt, test de Buishand et Test de Worsley

Par définition, un test est dit robuste s'il n'est pas sensible à des petits écarts sur les conditions pour lesquelles la statistique de test obéit à une certaine loi de distribution, lorsque l'hypothèse principale H_0 est vraie. Mais, quelles sont ces conditions ? Tout dépend bien évidemment des tests. Une condition qui se trouve dans les quatre tests examinés dans cette étude est la condition d'indépendance des observations de l'échantillon. Cette condition est rarement remplie lorsqu'on considère en pratique les séries de données d'une variable hydrométéorologique. En effet, le phénomène de persistance (phénomène d'autocorrélation temporelle) est très présent dans de telles séries de données. L'étude de robustesse qui est faite ici cherche à évaluer le comportement d'un test vis à vis de séries normales autocorrélées.

8.2.3.1 Modèle de simulation de séries stationnaires autocorrélées

Dans cette étude, le modèle de génération de données est simplement un modèle autorégressif d'ordre 1, $AR(1)$. Il s'écrit comme suit :

$$x_t = \rho\, x_{t-1} + \varepsilon_t \tag{8.9}$$

avec les erreurs $\varepsilon_t \sim \overset{iid}{N}(0,1)$ et $x_0 = 0$. Le paramètre autorégressif ρ prend les valeurs $\rho = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8$ et 0.9. Finalement, les méthodes de simulation de Monte Carlo sont utilisées pour générer $M = 1000$ séries de $n = 30, 50$ ou 100 valeurs. Comme dans les expériences précédentes, chaque série simulée est soumise aux tests MC_1rup, Buishand, Worsley et Pettitt pour tester H_0 au niveau de signification théorique $\alpha = 1\%, 5\%$ ou 10%.

8.2.3.2 Analyse des résultats

Les résultats obtenus dans le cadre de l'étude de robustesse sont rapportés dans les tableaux C1, C2 et C3 de l'annexe C ainsi qu'à la figure 8.5. À l'examen de ces tableaux, on observe que la performance des tests est sérieusement compromise en présence d'autocorrélation. C'est lorsque la mesure d'autocorrélation excède la valeur de 0.4 que les tests ont de sérieux problèmes de distorsion de niveau. Ce type de comportement est d'ailleurs en tout point semblable à celui que Lubes-Niel et al. (1998) ont observé au cours d'une étude de robustesse sur les tests de Buishand et de Pettitt.

En résumé, on peut retenir que les tests habituellement employés dans les études de rupture en hydrométéorologie sont fortement influencés par la présence de persistance dans l'échantillon de données. Il est donc clair que, lorsqu'ils sont appliqués à des séries de données hydrométéorologiques, ces tests peuvent conduire à des rejets fallacieux de l'hypothèse principale d'absence de rupture.

Cette conclusion doit nous inciter à la prudence, car comme on l'a maintes fois souligné dans cette thèse, la condition d'indépendance des observations est rarement remplie lorsqu'on étudie des séries de données hydrométéorologiques. Ceci met donc en évidence l'importance de développer un test prenant en compte cette réalité des phénomènes hydrométéorologiques.

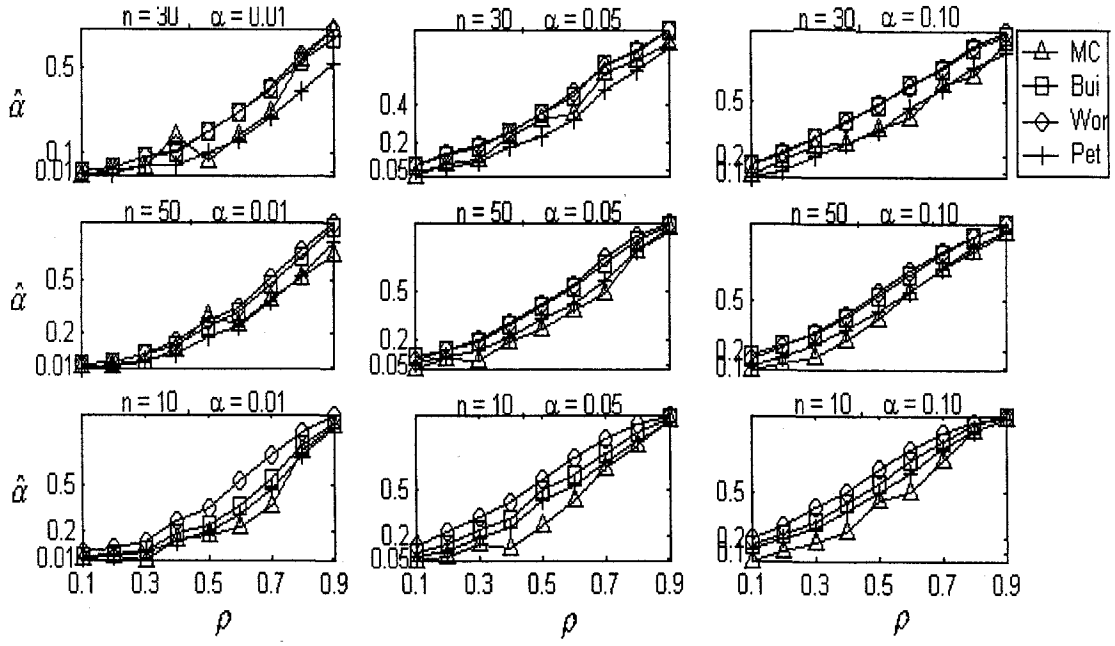


Figure 8.5 : MC_1rup, une rupture en absence de persistance : étude comparative du niveau (échantillon autocorrélé)

8.3 Étude de la performance du test de Monte Carlo dans un échantillon avec persistance

On se propose à présent d'étudier le comportement du test de Monte Carlo de détection d'une rupture dans un échantillon avec persistance (voir tableau 7.3 et les spécifications de la section 8.1.3). Les expériences de simulations qui suivent vont permettre :

- de vérifier que le test de Monte Carlo proposé garantit bien le niveau de signification théorique α, c'est-à-dire qu'il ne connaît pas de problème de distorsion de niveau,
- d'évaluer la puissance de ce test, c'est-à-dire sa capacité de détecter une rupture dans un échantillon de données.

Le test de Monte Carlo de détection d'une rupture dont il est question ici sera noté MC_1rup, c'est le test maximisée Monte Carlo (MMC). D'après le tableau 7.3 et les spécifications de la section 8.1.3, ce test est mis en œuvre en considérant l'ensemble $M_0 = \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$.

8.3.1 Étude de niveau : test de Monte Carlo de détection d'une rupture dans un échantillon avec persistance

On cherche à vérifier si le test de Monte Carlo de détection d'une rupture dans un échantillon avec persistance contrôle bien le niveau dans un échantillon de données autocorrélées.

8.3.1.1 Modèle de simulation de séries stationnaires autocorrélées

La génération de données autocorrélées sous H_0 s'effectue dans les mêmes conditions que précédemment : on utilise le modèle autorégressif d'ordre 1 de l'équation 8.9. Puis, pour différentes valeurs de $\rho = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8$ ou 0.9, les méthodes de simulation de Monte Carlo sont utilisées pour générer $M = 1000$ échantillons de

$n = 30, 50$ ou 100 valeurs. Chaque échantillon est alors soumis au test MC_1rup pour vérifier H_0 au niveau de signification théorique $\alpha = 0.05$.

8.3.1.2 Analyse des résultats

Les résultats obtenus pour un niveau $\alpha = 5\%$ sont rassemblés dans le tableau C4 de l'annexe C. Ces résultats sont aussi repris et illustrés à la figure 8.6.

À la lecture de ces résultats, il apparaît clairement que le test de Monte Carlo de détection d'une rupture dans un échantillon avec persistance contrôle parfaitement le niveau. En effet, le niveau empirique estimé est, dans toutes les situations d'étude considérées inférieur au niveau théorique $\alpha = 5\%$, sans regard à la taille d'échantillon ou aux valeurs du paramètre autorégressif ρ. Ce qui traduit la robustesse du test par la même occasion. C'est un résultat encourageant qui mérite d'être souligné en comparaison des piètres résultats obtenus, dans les mêmes conditions d'expérience, avec les tests de Pettitt, de Buishand et de Worsley (tableaux C1 à C3 de l'annexe C).

Il est important de noter ici que les résultats obtenus ici ne sont pas incompatibles avec ceux obtenus avec le test de Monte Carlo de détection d'une rupture dans un échantillon sans persistance lorsque le paramètre autorégréssif est proche de 0. En effet, pour $\rho = 1$ (et par conséquent proche de 0), au niveau théorique $\alpha = 5\%$, les résultats de la figure 8.1 sont comparables à ceux de la figure 8.6.

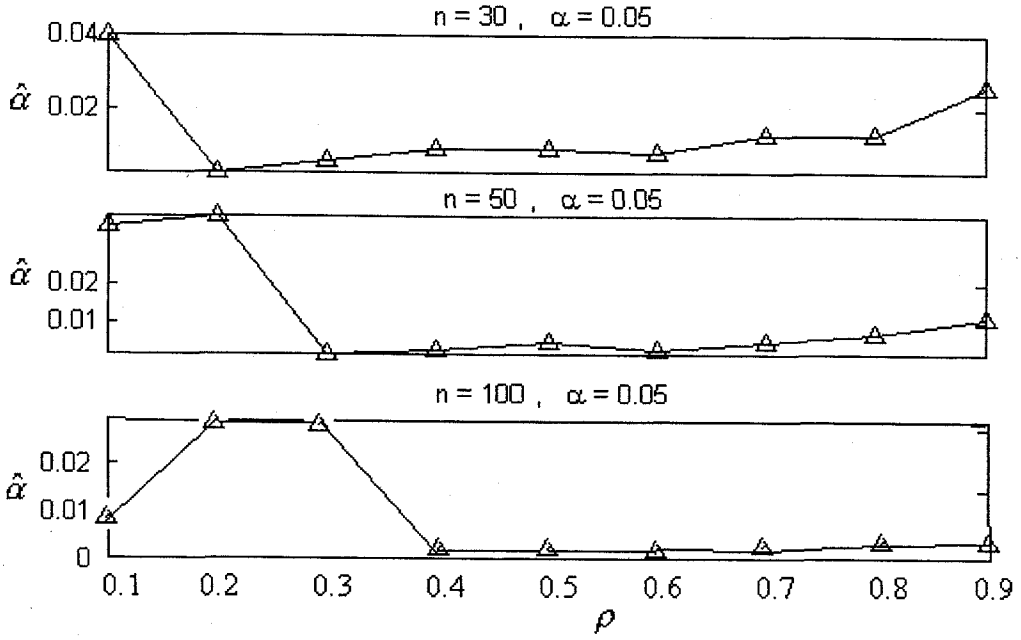


Figure 8.6 : MC_1rup, une rupture en présence de persistance : étude comparative du niveau (échantillon autocorrélé)

8.3.2 Étude de puissance : test de Monte Carlo de détection d'une rupture dans un échantillon avec persistance

On s'intéresse maintenant à évaluer la puissance du test de Monte Carlo de détection d'une rupture dans un échantillon avec persistance. Pour ce faire, on va examiner dans un premier temps sa puissance de détection d'une rupture brusque de la moyenne dans un échantillon autocorrélé. Dans un second temps, on explore cette fois-ci sa puissance de détection d'une rupture avec continuité de la moyenne.

8.3.2.1 Comportement du test étudié dans la détection d'une rupture brusque

8.3.2.1.1 Modèle de simulation de séries non-stationnaires autocorrélées : cas de la rupture brusque

Les séries artificielles autocorrélées présentant une rupture brusque de la moyenne à la date t_b d'amplitude Δ sont générées par le processus

$$x_t = \begin{cases} \mu_0 + \rho\, x_{t-1} + \varepsilon_t & t = 1,\ldots,t_b \\ (\mu_0 + \Delta) + \rho\, x_{t-1} + \varepsilon_t & t = t_b + 1,\ldots,n \end{cases} \qquad (8.10)$$

avec les erreurs $\varepsilon_t \overset{iid}{\sim} N(0,1)$ et $x_0 = 0$. Pour un coefficient d'autocorrélation ρ prenant les valeurs $\rho = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8$ ou 0.9, les méthodes de simulation Monte Carlo sont utilisées pour générer $M = 1000$ échantillons de $n = 30, 50$ ou 100 valeurs comportant un taux de rupture Δ sur la moyenne de 0.5%, 1%, 1.5%, 2% et 2.5%. Chaque échantillon est ensuite soumis au test MC_1rup pour vérifier H_0 au niveau théorique $\alpha = 0.05$.

8.3.2.1.2 Analyse des résultats

Les résultats obtenus sont donnés dans les tableaux C5 à C10 de l'annexe C. Ils sont aussi reproduits sous la forme de figures (figure 8.7 et 8.8). La première observation est que les estimations obtenues sont d'une précision plus que satisfaisante.

Sur les tableaux C5, C6 et C7, ainsi que sur la figure 8.7, on constate que la puissance du test de Monte Carlo de détection d'une rupture dans un échantillon avec persistance varie en fonction du taux de rupture appliqué (Δ), et de la valeur du coefficient d'autocorrélation ρ. Pour des tailles d'échantillons n supérieures ou égales à 30, la relation entre la puissance du test et le coefficient d'autocorrélation est croissante, sans égard au taux de rupture appliqué (Δ) même si la rupture a lieu au début, au milieu ou vers la fin de l'échantillon (figure 8.7). On note aussi qu'une augmentation de l'autocorrélation pour des tailles d'échantillons $n = 100$ n'a pas d'incidence sur la puissance du test (égale à 1) pour des taux de rupture strictement supérieurs à 1.5% ($\Delta \geq 1.5\%$).

Dans tous les cas, la puissance du test est bonne pour des taux de rupture strictement supérieurs à 1.5% (figure 8.7). En revanche, les proportions de rejets de H_0 associés à une estimation correcte de la date de rupture sont relativement faibles (tableaux C8, C9 et C10, ainsi que la figure 8.8). En effet, quand le taux de rupture de la moyenne est de 2.5%, les proportions passent de 5% à 50% (respectivement de 50% à 5%), et ce, lorsque le coefficient d'autocorrélation varie de 0.1 à 0.5 (respectivement de 0.5 à 0.7). Par ailleurs, sur des échantillons présentant un coefficient d'autocorrélation de l'ordre de 0.8 ou 0.9, la proportion de rejets de H_0 associés à une estimation correcte de la date de rupture est pratiquement nulle quel que soit le taux de rupture de la moyenne considéré.

Finalement, ces résultats confirment que, pour ρ = 1 (c'est-à-dire proche de 0), la puissance du test de Monte Carlo de détection d'une rupture dans un échantillon avec persistance est comparable à la puissance du test Monte Carlo de détection d'une rupture dans un échantillon sans persistance. Cette observation doit cependant être relativisé dans le cas de la proportion de rejets de H_0 associés à une estimation correcte de la date de rupture. En effet, on remarque que, pour $\rho = 1$ (c'est-à-dire proche de 0), cette proportion est faible avec le test de Monte Carlo de détection d'une rupture dans un échantillon avec persistance (figure 8.8)

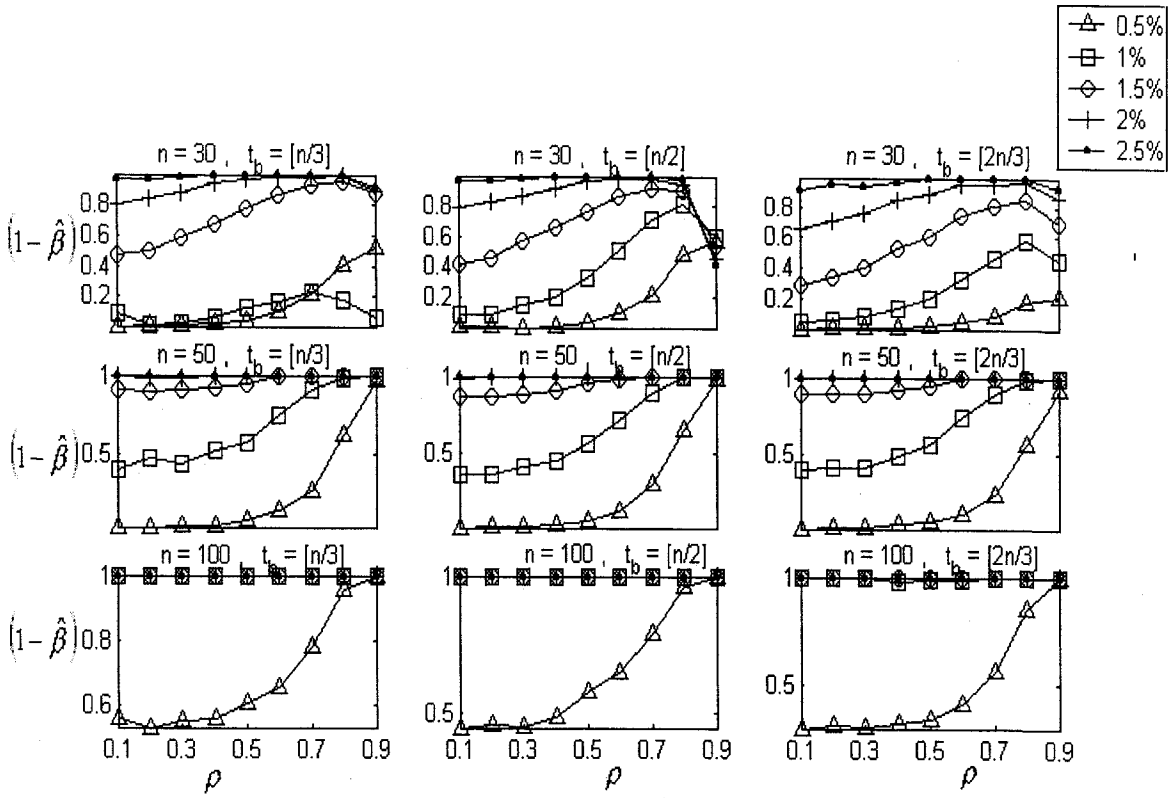


Figure 8.7 : MC_1rup, une rupture en présence de persistance : puissance du test (échantillon autocorrélé, $\alpha = 0.05$)

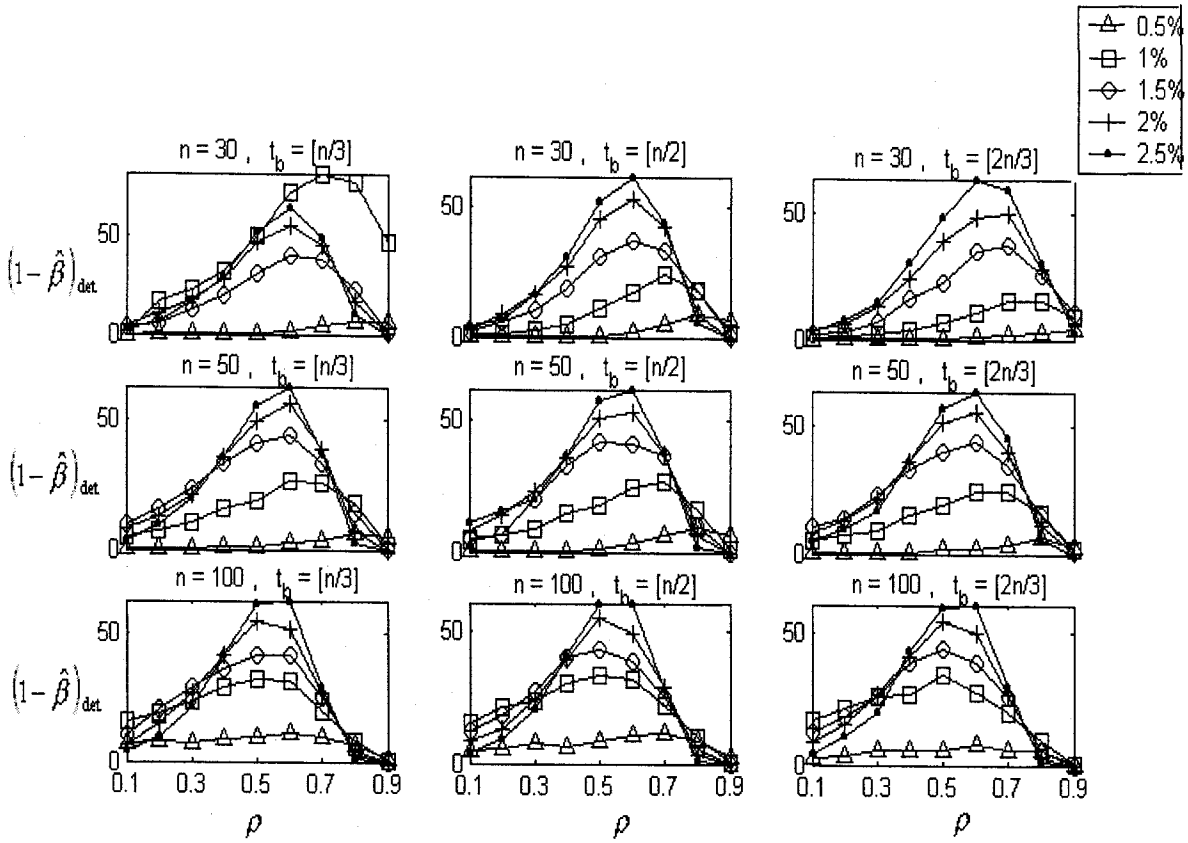


Figure 8.8 : MC_1rup, une rupture en présence de persistance : pourcentage d'estimation de la date de rupture (échantillon autocorrélé, $\alpha = 0.05$)

8.3.2.2 Comportement du test étudié dans la détection d'une rupture avec continuité

8.3.2.2.1 Modèle de simulation de séries non-stationnaires autocorrélées : cas de la rupture avec continuité

Les séries artificielles autocorrélées présentant une rupture avec continuité de la moyenne à la date t_b de pente δ sont générées par le processus

$$x_t = \mu_0 + \rho\, x_{t-1} + \delta\, ID(t_b) + \varepsilon_t \qquad (8.11)$$

où ε_t, ρ, μ_0 et δ sont définis comme précédemment. La variable $ID(t_b)$ est de la même manière qu'à l'équation 8.8. pour un coefficient d'autocorrélation $\rho = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8$ ou 0.9, les méthodes de simulation Monte Carlo sont utilisées pour générer $M = 1000$ échantillons de $n = 30, 50$ ou 100 valeurs comportant une rupture de pente $\delta = 0.1, 0.2, 0.3$ ou 0.4. Chaque échantillon est soumis au test MC_1rup pour vérifier H_0 au niveau théorique $\alpha = 0.05$.

8.3.2.2.2 Analyse des résultats

Les résultats obtenus dans ce dernier cas sont portés dans les tableaux D1 à D6 de l'annexe D : D1 à D3 pour l'estimation de la puissance $(1-\hat{\beta})$ et D4 à D6 pour l'estimation de la proportion de rejets de H_0 associés à une estimation correcte de la date de rupture, $(1-\hat{\beta})_{det}$.

On observe que le test étudié est performant pour détecter une rupture avec continuité de la moyenne dans un échantillon autocorrélé (Tableaux D1, D2 et D3). Par ailleurs, on constate que la relation entre la puissance du test et le coefficient d'autocorrélation est croissante, indépendamment de la pente de rupture considérée. Par ailleurs, les résultats des tableaux (D4, D5 et D6) montrent que les proportions de rejets de l'hypothèse principale associés à une estimation correcte de la date de rupture sont relativement faibles, et ce, même lorsque la rupture survient au milieu de l'échantillon. On obtient un

pourcentage de 48% pour une rupture survenue au début de l'échantillon, 46% pour une rupture survenue au milieu de l'échantillon et 42% pour une rupture survenue vers la fin de l'échantillon. À l'évidence, il semble que la proportion de rejets de l'hypothèse principale associés à une estimation correcte de la date de rupture est relativement meilleure pour une rupture survenue au début de l'échantillon. Outre cela, ces mêmes résultats mettent aussi en évidence la détérioration graduelle de cette proportion à mesure que le coefficient d'autocorrélation croit (figure 8.9).

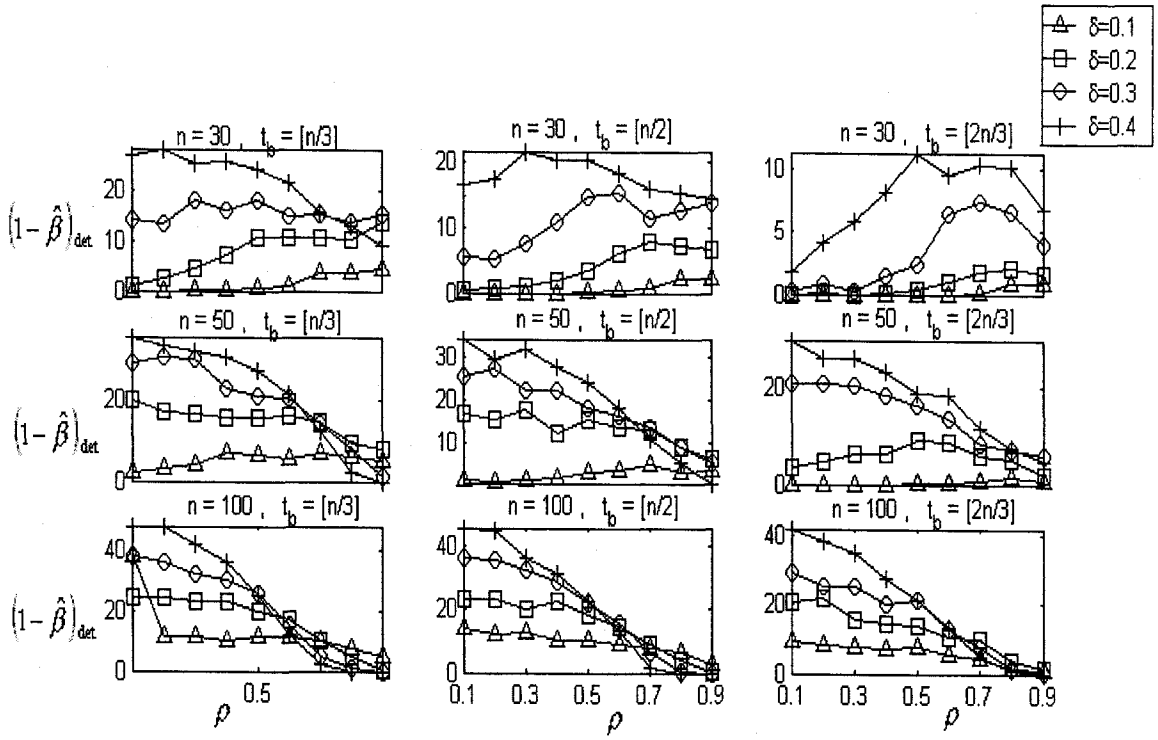


Figure 8.9 : MC_1rup, rupture avec continuité en présence de persistance : pourcentage d'estimation de la date de rupture (échantillon autocorrélé, $\alpha = 0.05 , t_b = [n/2]$)

8.4 Étude de la performance du test de Monte Carlo de détection de plusieurs ruptures

On illustre dans cette partie le comportement du test de Monte Carlo de détection de plusieurs ruptures brusques de la moyenne qui est proposé dans cette thèse. Des simulations de type Monte Carlo sont mises en œuvre afin d'évaluer le niveau et la puissance de ce test. En particulier, on veut comparer les résultats des simulations avec la théorie exposée sur les tests de Monte Carlo. Cela dit, on s'intéresse au test décrit au tableau 7.5. Sans perte de généralité, on notera ce test MC_krup.

8.4.1 Comportement du test de Monte Carlo de détection de ruptures dans un échantillon sans persistance

Les étapes de déroulement du de Monte Carlo (test MC_krup) qu'on explore ici sont données dans le tableau 7.5 (c'est-à-dire, une combinaison des tableaux 7.4 et 7.1). Les autres spécifications sont données à la section 8.1.3.

8.4.1.1 Étude de niveau du test de Monte Carlo de détection de ruptures brusque : séries indépendantes

On veut explorer le niveau du test de Monte Carlo de détection de ruptures dans un échantillon de données indépendantes. Pour ce faire, les séries artificielles sont générées sous H_0, à partir du modèle décrit à l'équation 8.5. Les simulations de Monte Carlo sont alors utilisées pour produire $M = 1000$ échantillons de $n = 30, 50$ ou 100 valeurs. Chaque série générée est ensuite soumise au test MC_krup pour vérifier H_0 au niveau de signification $\alpha = 5\%$. La longueur minimale de segmentation admise dans la procédure séquentielle est $h = 10$ (voir tableau 7.4).

Le tableau E1 de l'annexe E donne les résultats obtenus. Il appert que pour les différentes tailles d'échantillons considérées, le test de Monte Carlo de détection de ruptures contrôle parfaitement le niveau. En effet, on observe que dans tous les cas examinés, le niveau empirique estimé est inférieur au niveau théorique $\alpha = 5\%$. Il apparaît donc approprié de

dire que ce test ne connaît pas de problème de distorsion de niveau. D'autre part, les estimations obtenues sont d'une qualité satisfaisante.

8.4.1.2 Étude de puissance du test de Monte Carlo de détection de ruptures : séries indépendantes

La présente étude concerne deux situations d'intérêt. On veut évaluer d'une part, la puissance du test relativement à sa capacité de détecter une rupture dans un échantillon de données indépendantes lorsque c'est vraiment le cas. D'autre part, on veut réitérer le même exercice pour deux ruptures.

8.4.1.2.1 Étude de la performance du test dans la détection d'une rupture brusque de la moyenne : séries indépendantes

Pour générer des séries artificielles indépendantes présentant une rupture brusque de la moyenne d'amplitude Δ et située au milieu de l'échantillon ($t_b = [n/2]$), on considère le modèle décrit à l'équation 8.6. Les méthodes de simulation de type Monte Carlo sont alors utilisées pour générer $M = 1000$ échantillons de $n = 30, 50$ ou 100 valeurs. Chaque échantillon généré est soumis au test MC_krup pour tester H_0 au niveau de signification $\alpha = 5\%$. La longueur minimale de segmentation admise est $h = 10$.

Les résultats issus de ces simulations sont présentés dans les tableaux E2 et E5 de l'annexe E. On observe que le test détecte bien la rupture même si le taux de rupture (ou encore amplitude Δ) est faible (tableau E2). En revanche, les résultats montrent aussi que la proportion de rejets de l'hypothèse principale associés à une estimation correcte de la date de rupture est faible dans les petits échantillons pour un taux de rupture Δ = 0.5%. Par exemple, pour n = 30 cette proportion est de 9% (tableau E5). Quoi qu'il en soit, ces résultats révèlent de manière globale que la proportion de rejets de H_0 associés à une estimation correcte de la date de rupture est très bonne. Elle atteint même 91.4% quand la taille d'échantillon est égale à 100 ! Ce résultat est vraiment impressionnant.

8.4.1.2.2 Étude de la performance du test dans la détection de deux ruptures brusques de la moyenne : séries indépendantes

Les séries artificielles indépendantes présentant une rupture brusque au début ($t_{b_1} = [n/3]$) et à la fin de l'échantillon ($t_{b_2} = [2n/3]$) sont générées par le processus

$$x_t = \begin{cases} \mu_0 + \varepsilon_t & t = 1, \ldots, t_{b_1} \\ (\mu_0 + \Delta) + \varepsilon_t & t = t_{b_1} + 1, \ldots, t_{b_2} \\ \{(\mu_0 + \Delta) + \Delta\} + \varepsilon_t & t = t_{b_2} + 1, \ldots, n \end{cases} \qquad (8.12)$$

les paramètres μ_0, Δ et les erreurs ε_t sont analogues à ceux de l'équation 8.6. À partir des méthodes de simulation Monte Carlo, on génère $M = 1000$ échantillons de $n = 30, 50$ ou 100 valeurs. Chaque échantillon est ensuite soumis au test MC_krup pour vérifier H_0 au niveau de signification $\alpha = 5\%$. On explore ici deux longueurs minimales de segmentation : $h = 5$ *et* 10.

Les tableaux E3, E4, E6 et E7 de l'annexe E rassemblent les résultats obtenus. Si on commence par examiner la puissance du test, on voit que cette dernière croît avec le taux de rupture. Le test arrive à détecter les deux ruptures simulées dans des proportions très élevées. Celles-ci atteignent la valeur de 0.80 pour des taux de rupture de 1.5%, 2% et 2.5%, et ce sans égard à la taille d'échantillon et à la longueur minimale de segmentation *h*. D'autre part, le test estime correctement les deux dates de rupture simulées avec un pourcentage plus qu'acceptable : près de 50% des cas pour un taux de rupture de 1.5% ! En revanche, la puissance du test et sa proportion de rejets de H_0 associés à une estimation correcte de la date de rupture sont moins grandes pour un taux de rupture de 0.5%.

Dans tous les cas, les résultats obtenus sont vraiment encourageants. Il n'est pas inutile de mentionner que, en regard des tableaux E3, E4, E6 et E7 de l'annexe E, la qualité des estimations est satisfaisante.

8.4.2 Comportement du test de Monte Carlo de détection de plusieurs ruptures dans un échantillon avec persistance

Les étapes de déroulement du test de Monte Carlo (MC_krup) qu'on étudie ici sont consignées dans le tableau 7.5 (c'est-à-dire, une combinaisons des tableaux 7.4 et 7.3).

8.4.2.1 Étude de niveau du test de Monte Carlo de détection de ruptures : séries autocorrélées

Le modèle de génération des données est le même que celui exprimé à l'équation 8.9. Pour $\rho = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8$ et 0.9, on génère, à l'aide de simulations de Monte Carlo, $M = 1000$ séries de $n = 30, 50$ ou 100 valeurs. Chacune de ces séries est soumise au test MC_krup pour vérifier H_0 au niveau de signification théorique $\alpha = 5\%$. La longueur minimale d'un segment dans la procédure séquentielle est $h = 10$.

Le tableau E9 de l'annexe E nous donne un aperçu des résultats obtenus avec les différentes simulations. À première vu, le test MC_krup contrôle le niveau dans des échantillons autocorrélés. En effet, le tableau E9 montre que le niveau de signification empirique du test est strictement inférieur au niveau de signification théorique $\alpha = 0.05$ pour toute valeur du coefficient d'autocorrélation ρ. Cela confirme donc que le test de Monte Carlo de détection de multiples ruptures dans un échantillon avec persistance est un test exact.

8.4.1.2 Étude de puissance du test de Monte Carlo de détection de ruptures : séries autocorrélées

La présente étude est menée de façon analogue aux études de simulations qui ont été effectuées sur des séries indépendantes.

8.4.1.2.1 Étude de performance du test dans la détection d'une rupture brusque : séries autocorrélées

Le modèle de génération de séries artificielles autocorrélées présentant une rupture brusque de la moyenne d'amplitude Δ et située au milieu de l'échantillon ($t_b = \lfloor n/2 \rfloor$) est le même

que celui qui est exprimé à l'équation 8.10. À partir des méthodes de simulation Monte Carlo, on génère $M = 1000$ échantillons de $n = 30, 50$ ou 100 valeurs. Chaque échantillon est ensuite soumis au test MC_krup pour vérifier H_0 au niveau de signification $\alpha = 5\%$. La longueur minimale de segmentation est $h = 10$.

Les tableaux E10 et E11 de l'annexe E présentent les résultats obtenus. Pour interpréter ces résultats, remarquons que le test a de la difficulté à détecter une rupture dans un échantillon avec persistance. En effet, pour toute valeur du paramètre autorégressif ρ, la puissance du test est faible, même pour un taux de rupture de 2.5% (tableau E10). Pire encore, les proportions de rejets de H_0 associés à une estimation correcte de la date de rupture ne sont pas plus grand que 2% (tableau E11). Ces résultats tendent donc à montrer que, en présence d'échantillons avec persistance, le test de Monte Carlo de détection de multiples ruptures a une puissance beaucoup moins bonne.

Quelles raisons pourrait-on invoquer pour expliquer l'affaiblissement de la puissance du test de Monte Carlo de détection de multiples ruptures en présence d'échantillons avec persistance ? De notre point de vue, ceci peut avoir deux origines. Soit que ce test n'est pas capable de discerner le phénomène de rupture et celui de persistance, soit que ce test est performant pour des taux de rupture strictement supérieurs à 2.5%. Pour cette dernière hypothèse, on remarque effectivement que la puissance s'accroît légèrement avec le taux de rupture (tableaux E10 et E11).

8.5 Conclusion

Les simulations qui viennent d'être d'effectuées dans ce chapitre ont permis d'une part, de confirmer le bien-fondé théorique des tests de Monte Carlo et d'autre part de montrer l'intérêt des tests de Monte Carlo qui sont proposés dans cette thèse pour la détection de non-stationnarité.

En premier lieu, on a pu constater le fait que dans l'étude comparative, le test de Monte Carlo de détection d'une rupture apporte cinq perfectionnements par rapport aux tests habituellement employés dans les études de rupture en hydrométéorologie :

1. Il ne connaît pas de problème de distorsion de niveau, à l'opposé du test de Worsley (figure 8.1).
2. Sa puissance dans le pire des cas est égale à celle du test de Pettitt, de Buishand ou de Worsley pour détecter une rupture brusque de la moyenne au milieu de l'échantillon (figures 8.2, 8.3 et 8.4).
3. Il est plus puissant que les tests habituellement employés en hydrométéorologie pour détecter (respectivement estimer) une rupture (respectivement date de rupture) brusque de la moyenne au début ou à la fin de l'échantillon (tableaux A11 à A19 de l'annexe A).
4. Il est plus puissant pour détecter (respectivement estimer) une rupture (respectivement une date de rupture) avec continuité de la moyenne dans un échantillon (tableaux de l'annexe C).
5. Il est tout simplement plus robuste, en ce sens que, comparé aux autres tests qui sont habituellement employés en hydrométéorologie, il donne des meilleurs résultats dans un échantillon dépendant (figures 8.5 et 8.6).

Les résultats obtenus confirment aussi que le test de Monte Carlo de détection de plusieurs ruptures brusques de la moyenne est très performant dans le cas d'échantillons de données indépendantes (tableaux E1 à E8 de l'annexe E). En revanche, cette performance est beaucoup moins bonne dans le cas d'échantillons de données autocorrélées. Par ailleurs, outre le fait que les résultats obtenus paraissent très satisfaisants, il n'en demeure pas moins que ces tests de Monte Carlo, aussi bien que les autres tests classiques, peuvent avoir une puissance moindre qui augmente de façon substantielle lorsque l'on augmente la taille de l'échantillon. Donc apparemment, ces tests donneraient de meilleurs résultats dans des échantillons de grande taille.

Finalement, après avoir analysé les résultats exposés dans ce chapitre, on peut se demander s'ils peuvent se généraliser. La réponse est oui à cause des faits suivants. D'abord, les résultats obtenus sont comparables à ceux obtenus dans d'autres études qui se sont consacrées aux tests de Pettitt et Buishand (Lubes-Niel, 1998). D'autre part, on a réalisé également d'autres simulations en considérant $\mu_0 = 0$ et en retenant les taux de rupture utilisés dans le cadre de la présente étude. Les résultats obtenus étaient en tout point semblables à ceux présentés ici. En somme, il resterait à vérifier comment les tests de Monte Carlo qui sont proposés dans cette thèse se comportent dans des séries comportant d'autres types de non-stationnarité. On pense en particulier au problème de rupture sur la variance etc.

9 EXEMPLE D'APPLICATION SUR DES DONNÉES RÉELLES

Pour mettre en évidence les tests de Monte Carlo de détection de rupture (s) qui ont été proposés dans cette thèse, on va appliquer certains d'entre eux sur des données réelles. Dans ce chapitre, on propose une application des tests de Monte Carlo sur un ensemble de séries de données hydrologiques. Ces séries ont déjà été utilisées dans le cadre du réseau de suivi des changements climatiques de la province de Québec (Ouarda et al., 1999). Le but poursuivi est de confronter nos conclusions avec celles obtenues par Ouarda et al. (1999) qui ont utilisé la procédure bayésienne de Lee et Heghinian (1977) pour détecter les changements en moyenne dans ces séries de données.

9.1 Données de l'étude

9.1.1 Cadre de l'étude

La mise en place des réseaux de mesures spécialement conçus pour l'étude des modifications climatiques et leurs impacts sur les ressources en eau est fort utile pour mieux déceler les effets d'un réchauffement à l'échelle planétaire. Le but d'une telle opération est de délimiter les bassins hydrographiques pour lesquels il existe suffisamment de données sur le débit d'eau (et au besoin sur les variables climatologiques) pour le suivi des modifications climatiques. C'est dans cette optique que Ouarda et al. (1999) ont conduit une étude dont l'objectif était la conception d'un réseau hydrométrique pour le suivi des modifications climatiques dans la province de Québec. Les données utilisées dans le cadre de cette étude sont les débits moyens annuels d'un ensemble de 19 stations préalablement choisies, et ce, à partir de critères spécifiques. La figure 9.1 présente une répartition de ces stations sur la carte du Québec. Les caractéristiques des 19 stations sont présentées dans le tableau 9.1.

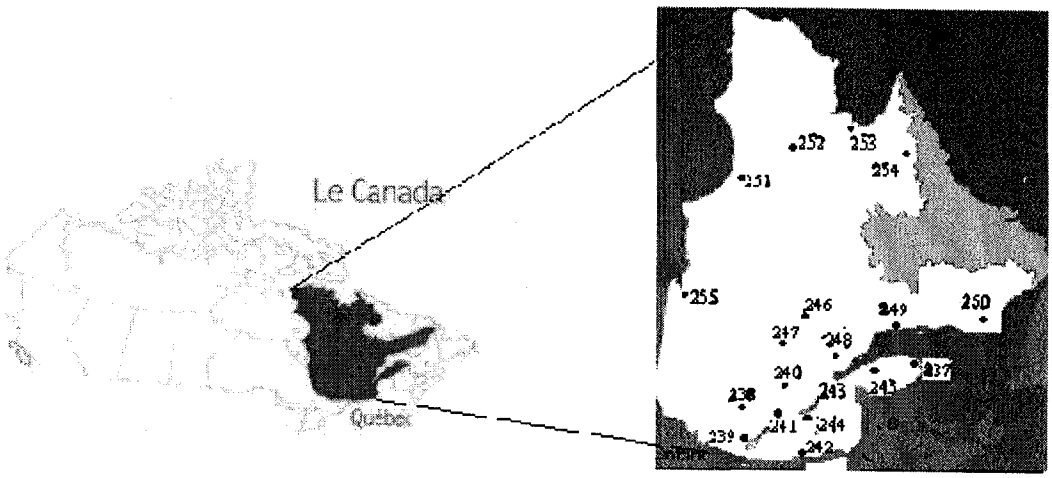


Figure 9.1 : Répartition des stations retenues dans le cadre du réseau de suivi des changements climatiques de la province de Québec

N° d'identification	Nom de la station	Aire (km^2)	Début d'enregistrement
237	Dartmouth	645	1970
238	Gatineau	6 840	1974
239	Picanoc	1 290	1926
240	Croche	1 570	1965
241	Matawin	1 390	1931
242	Eaton	642	1953
243	Sainte-Anne	642	1965
244	Beaurivage	704	1925
245	Rimouski	1 610	1962
246	Mistassibi	9 320	1953
247	Chamouchouane	15 300	1915
248	Metabetchouane	2 280	1964
249	Moisie	19 000	1965
250	Romaine	13 000	1956
251	Lac des Loups Marins	8 390	1974
252	Aux Mélèzes	42 700	1965
253	À la Maleine	29 800	1956
254	George	24 200	1975
255	Harricana	3 680	1933

Tableau 9.1 : Liste de stations retenues dans le cadre du réseau de suivi des changements climatiques de la province de Québec (Ouarda et al., 1999)

Le but poursuivi par Ouarda et al. (1999) était de détecter des changements brusques de la moyenne dans les séries de données obtenues à chacune de ces 19 stations. La méthode de détection qu'ils ont utilisée est la procédure bayésienne de Lee et Heghinian (1977). Au vu des résultats qu'ils ont obtenus, il semble que, pour 5 des 19 stations retenues, la date la plus probable d'occurrence du changement se situe autour des années 1983-1985. Leurs résultats sont repris dans le tableau 9.2.

No d'identification	Date probable de changement
237	1983
249	1984
250	1985
253	1985
254	1985

Tableau 9.2 : Principaux résultats obtenus par Ouarda et al. (1999)

Dans le cadre de notre étude, nous reprenons la base de données ayant servi à Ouarda et al. (1999) pour détecter des changements brusques de la moyenne des séries (tableau 9.1). Nous nous attachons donc à analyser, à l'aide des tests de Monte Carlo, les changements brusques de la moyenne dans les 19 séries de débits moyens annuels de bassins versants situés dans la région de Québec (figure 9.1). Le but recherché dans notre approche est de confronter nos conclusions avec celles obtenues par Ouarda et al. (1999).

9.1.2 Présentation de la base de données

Avant toute chose, il importe de se faire une idée globale de la base de données à notre disposition. À cet effet, la figure 9.1 représente la répartition des 19 stations et le tableau 9.1 donne quelques caractéristiques sur ces stations. Les séries de données obtenues à ces différentes stations sont des débits moyens annuels.

Une des caractéristiques essentielles de ces séries est qu'elles ne sont pas influencées par l'activité humaine (Ouarda et al., 1999). En effet, d'après l'étude menée par Ouarda et al.

(1999), ces stations sont surtout caractérisées par la qualité de leur courbe de tarage, l'absence de projets futurs dans les différentes régions qui délimitent leur bassin versant (pas de nouvel aménagement), la taille relativement grande de leur bassin de drainage, un emplacement géographique adéquat et des séries relativement longues. Cette base de données est donc adéquate pour le suivi de modification climatique.

9.2 Applications

On s'intéresse ici à détecter une ou plusieurs ruptures brusques de la moyenne dans les séries de données à notre disposition.

9.2.1 Caractéristiques statistiques des séries étudiées

Après avoir nettoyé les données, on a rassemblé les statistiques relatives aux 19 stations dans le tableau 9.3. En jetant un regard sur ce tableau, on note que la plus grande taille d'échantillon est égale à 69 (la station 244) tandis que la plus petite taille enregistre une valeur de 17 (la station 248). C'est donc dire que l'on est en présence de données relativement longues. La moyenne régionale calculée sur cet ensemble de données est $\mu_0 = 181$ m^3/s.

Le tableau 9.3 montre que certaines stations ont un coefficient d'aucorrélation relativement élevé. La série enregistrée à la station 238 a le plus grand coefficient d'autocorrélation, elle paraît donc autocorrélée. Au vu de la valeur de son coefficient d'autocorrélation ($|r_1| = 0.64$), on serait en droit de penser que l'hypothèse d'indépendance des observations peut sérieusement être mise en défaut dans cette série. On pourrait également être tenté d'étendre ces observations aux séries de données enregistrées aux stations 254 ($|r_1| = 0.49$), 249 ($|r_1| = 0.48$) et 239 ($|r_1| = 0.34$).

No d'identification	Taille (n)	Q moyen (m^3/s)	$\lvert r_1 \rvert^*$
237	25	14.6	0.24
238	18	127.9	0.64
239	23	18.7	0.34
240	29	29.5	0.02
241	61	24.1	0.18
242	41	13.2	0.04
243	29	19.1	0.06
244	69	14.2	0.19
245	32	30.7	0.20
246	34	198.3	0.16
247	32	300.0	0.03
248	17	47.2	0.38
249	29	427.2	0.48
250	38	297.0	0.26
251	18	195.7	0.26
252	28	615.7	0.03
253	32	511.0	0.03
254	19	495.7	0.49
255	62	59.4	0.12

* valeur absolue du coefficient d'autocorrélation d'ordre 1

Tableau 9.3 : Caractéristiques statistiques des stations retenues dans le cadre du réseau de suivi des changements climatiques de la province de Québec

9.2.2 Caractéristiques des tests de Monte Carlo

On rappelle que les objectifs de la présente étude se situent dans la ligne des études menées par Ouarda et al. (1999) sur ces mêmes séries de données : détection des changements brusques de la moyenne. Il convient donc de travailler avec l'hypothèse d'indépendance des données (hypothèse admise par la procédure bayésienne de Lee et Heghinian (1977)).

On se limite donc dans le cadre de ce travail aux procédures de Monte Carlo de détection d'une(des) rupture(s) brusque(s) de la moyenne dans un échantillon sans persistance. Ainsi, pour la détection d'une rupture brusque de la moyenne dans les séries, on applique le test de Monte Carlo de détection d'une rupture brusque dans un échantillon sans persistance (tableau 7.1). Par contre, pour la détection de plusieurs ruptures brusques de la moyenne, c'est le test de Monte Carlo de détection de ruptures brusques dans un échantillon sans persistance (tableau 7.5, c'est-à-dire une combinaison des tableaux 7.1 et 7.4) qui est appliqué. Dans ce dernier cas, on prend comme intervalle minimal de

segmentation $h = 5$. Le niveau de signification qu'on considère dans cette étude est $\alpha = 5\%$. Par ailleurs, pour les différents tests de Monte Carlo qui sont utilisés dans le cadre de l'étude, on génère la statistique de test 99 fois. On pose donc que $N = 99$ (voir tableaux 7.1 et 7.5). Ainsi, avec ces choix, on montre que $\alpha(N+1) = 0.05(99+1) = 5$ est un entier. Ce qui est conforme à la condition d'application des tests de Monte Carlo (proposition 6.2). En somme, ces tests sont appliqués aux séries pour vérifier l'hypothèse principale d'absence de rupture, H_0, au niveau de signification $\alpha = 5\%$.

9.2.3 Application du test de Monte Carlo de détection d'une rupture brusque de la moyenne

Le test de Monte Carlo de détection d'une rupture brusque de la moyenne dans un échantillon sans persistance (tableau 7.1) a été appliqué aux séries correspondant chacune à l'une des 19 stations qui composent la base de données à notre disposition. Afin de mieux voir l'étendue des résultats obtenus, on les a rassemblés dans le tableau 9.4.

No d'identification	Date de changement	p-valeur
237	1982	0.03 *
238	1987	1.00
239	1929	0.02
240	1976	0.21
241	1949	0.96
242	1983	0.95
243	1976	0.99
244	1930	0.20
245	1990	1.00
246	1963	0.01 *
247	1979	0.99
248	1980	0.78
249	1983	0.01 *
250	1984	0.02 *
251	1991	0.98
252	1973	0.17
253	1984	0.09
254	1984	0.01 *
255	1941	0.96

* p-valeur significative au niveau de signification $\alpha = 5\%$

Tableau 9.4 : Résultats obtenus avec le test de Monte Carlo de détection d'une rupture brusque de la moyenne dans un échantillon indépendant

Il appert que, au niveau de signification $\alpha = 5\%$, pour quatre des stations retenues, le changement brusque en moyenne se situe autour des années 1982-1984 : 1982 pour la station 237, 1983 pour la station 249, 1984 pour les stations 250 et 254.

À l'examen des tableaux 9.2 et 9.4, on constate qu'on obtient des résultats similaires à ceux de Ouarda et al. (1999). Il y a en effet la possibilité qu'une rupture de la moyenne, survenue autour des années 1982-1984, soit présente dans les séries de données enregistrées aux stations 237, 249, 250 et 254.

9.2.4 Détection de plusieurs ruptures brusques de la moyenne

On se propose à présent de détecter plusieurs ruptures brusques de la moyenne dans les 19 séries qui forment notre base de données. Le but poursuivit est d'aller un peu plus loin que les travaux de Ouarda et al. (1999). En effet, si on arrive à montrer qu'il existe un changement brusque de la moyenne dans une série de données, il ne faut pas non plus abandonner tout espoir de détecter d'autres changements brusques de la moyenne dans cette même série.

En bref, on ne sait pas si, dans la première partie de cette étude, on se trouve en présence d'un seul changement brusque de la moyenne dans les séries enregistrées aux stations 237, 239, 246, 249, 250 et 254 ou de plusieurs changements brusques de la moyenne. L'analyse doit donc obligatoirement passer par d'autres tests pour fournir un résultat fiable. C'est le but de cette deuxième partie de l'étude.

En adoptant toujours un niveau de signification $\alpha = 5\%$, on a appliqué le test Monte Carlo de détection de plusieurs ruptures brusques dans un échantillon sans persistance (tableau 7.5, c'est-à-dire une combinaison des tableaux 7.4 et 7.1) aux séries de données enregistrées dans les 19 stations. On rappelle que pour chacune de ces séries, ce test est appliqué pour vérifier l'hypothèse principale d'absence de rupture, H_0, au niveau de signification $\alpha = 5\%$. Le tableau 9.5 présente les résultats obtenus pour l'ensemble des 19 séries.

No d'identification	Date de changement	p-valeur
237	1982	0.11
238	1987	0.99
239	1929	0.02 *
240	1976	0.07
241	1949	0.50
242	1983	0.43
243	1976	0.76
244	1930	0.10
245	1990	1.00
246	1963	0.01 *
247	1979	1.00
248	1980	0.56
249	1983	0.01 *
250	1984	0.01 *
251	1991	1.00
252	1973	0.06
253	1984	0.04 *
254	1984	0.03 *
255	1941	0.33

* p-valeur significative au niveau de signification $\alpha = 5\%$

Tableau 9.5 : Résultats obtenus avec le test de Monte Carlo de détection de ruptures brusques de la moyenne dans un échantillon indépendant

La conclusion ne fait aucun doute : au niveau de signification $\alpha = 5\%$, aucune des 19 séries étudiées ne présente plus d'une rupture en son sein. Il est cependant intéressant de constater que les résultats obtenus ici ne sont pas incompatibles avec ceux obtenus à la section 9.2.3. En effet, à l'exception de la station 237, ces résultats viennent confirmer la possibilité qu'une rupture de la moyenne, survenue autour des années 1982-1984, soit présente dans les séries de données enregistrées aux stations 249, 250 et 254.

9.3 Conclusion

Les exemples d'application qui viennent d'être présentés dans ce chapitre montrent l'intérêt et le bien fondé des tests de détection de rupture qui sont mis en avant dans cette thèse. En effet, en utilisant la même base de données que Ouarda et al. (1999), on arrive à des conclusions similaires. On constate en effet la possibilité qu'un changement de la moyenne, survenue autour des années 1982-1984, soit présente dans les séries de données enregistrées aux stations 237, 249, 250 et 254. Il est cependant difficile d'expliquer les causes du phénomène observé aux stations 237, 249, 250 et 254. En effet, il se peut notamment que le changement observé à ces stations soit dû à un changement climatique ou au déplacement de la station de jaugeage par exemple. Il importe donc de faire très attention à l'interprétation finale des résultats obtenus.

En fin de compte, on n'a pu appliquer le test régional de Monte Carlo qui a été développé dans cette thèse (tableau 7.6). Le fait est que, d'après le tableau 9.1, les durées d'observations aux diverses stations de mesure ne sont pas égales. Il y a donc là une cause d'hétérogénéité. Dans ce cas précis, l'application du test régional n'est pas appropriée. En effet, si les stations d'intérêt ne forment pas un ensemble hydrologique homogène, il est difficile de présumer qu'il existe une certaine cohérence régionale entre les caractéristiques de rupture de moyenne présentes dans chaque série, notamment entre les dates où surviennent ces ruptures. C'est pour cette raison que cette étude s'est limitée à l'analyse indépendante de séries correspondant chacune à l'une des 19 stations de la base de données à notre disposition.

10 CONCLUSIONS GÉNÉRALES ET PERSPECTIVES

10.1 Résumé de la problématique de recherche

L'étude de la stationnarité de séries de données d'une variable hydrométéorologique est une tâche incontournable pour l'ingénieur hydrologue. Cette tâche est en particulier un problème central dans l'analyse fréquentielle de telles séries de données. En effet, une des hypothèses de base de l'analyse fréquentielle en hydrométéorologie consiste à supposer que le phénomène observé est resté stationnaire et que ses propriétés statistiques n'évolueront pas dans un proche futur.

Il existe beaucoup de questions sur l'aspect aléatoire d'un processus hydrométéorologique sur lesquelles un statisticien peut s'attaquer pour solutionner le problème de la détection de non-stationnarités dans les séries de données hydrométéorologiques. En effet, si par non-stationnarités on signifie des modifications naturelles et/ou artificielles des caractéristiques statistiques des processus hydrométéorologiques générateurs de ces séries qui peuvent être stables ou non, et si on ne se préoccupe pas de l'origine de ces non-stationnarités, mais que l'on considère que les dates d'échantillonnage de données parlent par elles-mêmes, alors le problème de détection de non-stationnarités dans des séries de données hydrométéorologiques peut être vu comme un problème de test de détection de changements dans une série de données. C'est dans cet esprit que s'est développée la littérature hydrométéorologique des tests de détection de non-stationnarités.

Au cours de ce travail de thèse, on a abordé principalement la détection de rupture en moyenne (encore appelée changement de régime) qui est le type de non-stationnarité le plus courant dans des séries de données d'une variable hydrométéorologique (Hubert et al., 1989). Il existe un volume important d'études reliées au problème de rupture dans la littérature hydrométéorologique. Les contributions les plus intéressantes de cette littérature concernent les tests statistiques qui sont conçus pour caractériser d'éventuelles discontinuités ou dérives dans le temps. Nombre de ces tests ont fait l'objet d'applications

pratiques sur des séries de données de variables hydrométéorologiques. Il s'agit principalement des tests de Buishand (1982), de Pettitt (1979) et de Worsley (1979). Ce sont des tests de type supremum. Lubès et al. (1998) ont montré que l'autocorrélation et la présence de tendance dans les séries sont les deux caractéristiques qui pénalisent le plus les performances de ces tests. Or, ce sont essentiellement ces caractéristiques que l'on retrouve dans des séries de données hydrométéorologiques.

Bien que des progrès aient été faits dans la littérature hydrométéorologique pour développer des tests de détection d'une rupture dans un échantillon de données, il n'en reste pas moins qu'il y a encore beaucoup de travail à faire. Premièrement, il apparaît nécessaire d'élaborer des tests de détection de rupture qui prennent en compte les caractéristiques que l'on retrouve couramment dans les séries de données hydrométéorologiques (la persistance et la présence de tendance dans les données). Deuxièmement, il paraît aussi important d'étendre ces tests au problème de détection de plusieurs ruptures dans une série de données hydrométéorologiques. Finalement, il serait également intéressant de généraliser ces tests au cas multidimensionnel. En effet, l'aspect régional, dans le problème de détection de rupture, est central en hydrométéorologie.

10.2 Contribution de la thèse

Au terme de cette thèse, nous ne pouvons nous empêcher de résumer le travail que nous avons accompli sur le problème de la détection de non-stationnarité en hydrométéorologie. Nous avons d'abord commencé par analyser la question de la non-stationnarité des données aux chapitres 2 3 et 4. Puis, nous avons passé en revue les principales approches de détection de non-stationnarité tant dans le domaine de l'hydrométéorologie que dans des domaines connexes comme la statistique et l'économétrie. Ensuite, au chapitre 6, nous avons exposé une technique que l'on peut mettre à profit pour la détection de non-stationnarité dans des séries de données en hydrométéorologie : la technique des tests de Monte Carlo. Finalement, au chapitre 7, nous avons développé, à partir de la théorie des tests de Monte Carlo, des tests originaux de détection de non-stationnarité dans des séries

de données hydrométéorologiques. Les performances de ces tests ont été évaluées à l'aide d'expériences de Monte Carlo au chapitre 8. En fin de compte, nous avons appliqué ces tests sur des données réelles (chapitre 9).

Sans revenir dans les détails des conclusions apportées dans chaque chapitre, on pense avoir atteint les objectifs spécifiques qu'on s'est assignés dans cette thèse de doctorat en hydrologie statistique. En somme, la contribution principale de cette thèse peut être résumée ainsi :

1. présentation d'une revue de littérature extensive des tests de stationnarité en hydrométéorologie et dans les domaines connexes comme la statistique et l'économétrique ;
2. proposition d'un test de Monte Carlo de détection d'une rupture (brusque ou avec continuité) dans un échantillon sans persistance ;
3. proposition d'un test de Monte Carlo de détection d'une rupture (brusque ou avec continuité) dans un échantillon avec persistance ;
4. proposition d'une procédure séquentielle de détection de plusieurs ruptures brusques dans un échantillon avec persistance ;
5. proposition d'une procédure séquentielle de détection de plusieurs ruptures brusques dans un échantillon sans persistance ;
6. proposition d'un test de Monte Carlo de détection de plusieurs ruptures brusques dans un échantillon avec persistance ;
7. proposition d'un test de Monte Carlo de détection de plusieurs ruptures brusques dans un échantillon sans persistance ;
8. proposition d'un test régional de Monte Carlo de détection de rupture(s) ;
9. réalisation d'expériences de Monte Carlo pour évaluer les performances des tests proposés ;
10. réalisation d'une application empirique des tests proposés.

Cette thèse n'offrait un réel intérêt que si elle était concrétisée par la proposition de nouveaux tests de détection de non-stationnarité qui prennent en compte les

caractéristiques les plus couramment présentes dans les séries de données hydrométéorologiques. Les simulations qui ont été effectuées au chapitre 8 ont permis d'une part, de confirmer le bien-fondé théorique des tests de Monte Carlo et d'autre part de montrer l'intérêt des tests de Monte Carlo qui sont proposés dans cette thèse. Dans bien des cas, ces tests sont plus performants que les tests habituellement employés en hydrométéorologie (voir la conclusion à la section 8.5). D'autre part, on y voit également un avantage d'utiliser ces tests dans la mesure où on peut les adapter pour détecter les autres types de non-stationnarités qui ont été évoqués au chapitre 5.

On fait donc le double vœu en terminant que d'une part les tests proposés dans cette thèse facilitent par leur simplicité de mise en œuvre et leur richesse théorique, une redécouverte du problème de détection de non-stationnarité en hydrométéorologie, et que d'autre part, ils donnent à l'ingénieur hydrologue la pleine assurance dans l'interprétation de ses données, lorsqu'il est confronté à ce problème.

10.3 Perspectives de recherche

Cette thèse constitue une base de travail à partir de laquelle de nouvelles activités de recherche peuvent être lancées. Les perspectives que nous proposons ont pour objet principal de donner lieu à divers prolongements et applications des tests de Monte Carlo en hydrométéorologie. que ce soit sur le plan des performances, de l'adaptabilité ou de la généralisation. Ainsi, les travaux qui restent à accomplir peuvent s'orienter sur les grandes directions suivantes :

Sur le plan des performances

- Améliorer le test de Monte Carlo de détection de plusieurs ruptures dans un échantillon avec persistance. L'intérêt de continuer ce travail est d'améliorer la puissance de ce test. En effet, lors des études de simulations (section 8.4.1.2.1) on a constaté que sa puissance est beaucoup moins bonne pour détecter une(des) rupture(s) dans le cas d'échantillons de données autocorrélées.

- Améliorer le test régional de détection de rupture(s) (tableau 7.6). On a proposé dans cette thèse un test régional de détection de non-stationnarité. Les principes d'élaboration de ce test sont encourageants mais insuffisants car ce test ne tient pas compte des corrélations spatiales existant entre les séries de données formant un ensemble hydrométéorologiquement homogène. Il serait donc important d'approfondir l'étude de ce test régional car c'est central en hydrométéorologie.

Sur le plan de l'adaptabilité

- Une adaptation des tests de Monte Carlo à d'autres types de non-stationnarité hydrométéorologique reste à accomplir. On pense principalement à développer un test de Monte Carlo de détection d'une(des) rupture(s) dans la variance d'une série de données hydrométéorologiques.

Sur le plan de la généralisation

- Un autre axe de travail identifié consiste à s'interroger sur une généralisation des tests de Buishand et de Pettitt en une version naturelle des tests Monte Carlo. Étant donné que le test de Buishand est construit à partir des sommes cumulées et que celui de Pettitt est un test non-paramétrique, il pourrait être intéressant de voir dans quelle mesure on peut faire de ces tests des versions de tests exactes (tests de Monte Carlo).

REVUE BIBLIOGRAPHIQUE

Anderson, T. W. (1984). An Introduction to Multivariate Statistical Analysis (Second ed.). New York : John Wiley & Sons.

Andrews, D. K. (1993). Tests for Parameter Instability and Structural Change with Unknown Change Point. *Econometrica, 61* pp. 821-856.

Arbuthnot, J. (1710). An argument for Divine Providence, taken from the constant regularity observed in the births of both sexes.Philosophical Transactions of the Royal Society, 27, 186-190.

Bahadur, R. R. (1967). An Optimal Property of the Likelihood Ratio Statistic. In Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, I : 13-26.

Bai, J. et Perron P. (1998). Estimating and Testing Linear Models with Multiple Structural Changes. *Econometrica, Vol., 66*, No 1, pp. 71-78.

Banerjee, A., Lazarova S. et Urga G. (1998). Bootstrapping Sequential Tests for Multiple Structural Breaks, *Discussion Paper No 17-98*, Institut of Economics and Statistics (Oxford).

Barnard, G. A. (1963). Comment on "The Spectral Analysis of Point Processes" by M. S.

Bartlett, M. S. (1956). An Introduction to stochastic processes, Cambridge University Press.

Bell, W. R., D. A. Dickey, R. B. Miller (1986). Unit Roots in Time Series Models : Tests and Implications. *The American Statistician*, 40, 1, pp 12-26.

Bendat, J. S. et A. G. Piersol (1966). Measurement and Analysis of Random Data. John Wiley & Sons, Inc., New York – London – Sydney, 390 pp.

Bernier, J. (1977). Étude de la Stationnarité des Séries Hydrométéorologiques. *La Houille Blanche*. N^0 4.313-319.

Bernier, J. (1994). Statistical Detection of Change in Geophysical Series. In : Natural Resouces Management under Changes of Physical or Climatic Environment. Proc. NATO Advanced Study Institute on "Engineering Risk and Reliability in a Changing Physical Environment" (France, May-June, 1993).

Berrymann, D. (1984). La Détection des Tendances dans les Séries Temporelles de Paramètres de la Qualité de l'eau à l'aide de Tests Non-parametriques. Thèse de Maîtrise, Institut National de la Recherche Scientifique (INRS-Eau), 130pp.

Berryman, D., B. Bobée, D. Cluis, J. Haemmerli (1988). Nonparametric Tests for Trend Detection in Water Quality Time Series. *Water Resources Bulletin*, 24 (3) :pp. 545-556.

Bhattacharya P. K. (1994). Some Aspects of Change-Point Analysis, dans E. Carlstein, F-G. Muller et D. Siegmund (eds), change-Point Problems, *IMS lecture notes- monograph series, vol. 23* : pp.28-56.

Bhattacharya, P. K. et Johnson, R. A. (1968). Nonparametric Tests for Shifts at an Unknown Time Point. *Ann. Math. Stat.* 39 pp. 1731-1743.

Bhattacharya P. K. et Fierson D, JR (1981). A Nonparametric Control Chart for Detecting Disorder. *Ann. of Stat.* 9. pp. 544-554.

Bhatti, M. I. et Wang, J. (1998). Power Comparaison of Some Tests for Testing Change Point. *Statistica, Anno LVIII*, N 1, pp.127-137.

Birnbaum, Z. W. (1974). Computers and Unconventional Test-Statistics, in Reliability and Biometry, ed. by F. Proschan, and R. J. Serfling, pp. 441-458. SIAM, Philadelphia, PA.

Birnbaum, A. (1954). Combining Independent Tests of Significance. *Journal of the American Statistical Association* 49, pp.559-574.

Bois, Ph. (1971). Une Méthode de Contrôle de Séries Chronologiques Utilisées en Climatologie et en Hydrologie. Laboratoires de Mécanique des Fluides. Université de Grenoble. "Section Hydrologie". 49 pages.

Bois, Ph. (1986). Contrôle des Séries Chronologiques Corrélées par Étude du Cumul des Résidus. Deuxièmes Journées Hydrologiques de l'Orstom. Montpellier, pp. 89-100.

Box., G. E. P. et G. C. Jenkins (1976). Time Series Analysis : Forecasting and Control. *2nd ed.* Holden-Day, San Francisco.

Brown, R. L., Durbin J., et Evans J. M. (1975). Techniques for Testing the Constancy of Regression Relationships over Time. *Journal of the Royal Statistical Society, Series B*, 37 : pp. 149-192.

Brunet-Moret, Y. (1971). Étude de l'Homogénéité de Séries Chronologiques de Précipitations Annuelles par la Méthode des Doubles Masses. *Cahiers O.R.S.T.O.M., Série Hydrologie*, Vol. VIII, n^0 4, pp. 3-31.

Buishand, T. A. (1982). Some Methods for Testing the Homogeneity of Rainfall Records. *Journal of Hydrology*, Vol. 58, pp. 11-27.

Buishand, T. A. (1984). Tests for Detecting a Shift in the Mean of Hydrological Time Series. *Journal of Hydrology*, Vol. 58, pp. 51-69.

Carlstein, E. (1988). Nonparametric Change-Point Estimation. *Ann, statisti*, 16, pp. 188-197.

Cavadias, G. S. (1992). A survey of Current Approaches to Modelling of Hydrological Time-Series with respect to Climate Variability and Change. Report prepared for the World Climate Programme-Water-Project A2, WMO/TD-N° 534, 50pp.

Chernoff, H. et Zacks S. (1964). Estimating the current mean of a normal distribution which is subjected to change in time. *Ann. Math. Stat.* 35 pp. 999-1018.

Chow, G. C. (1960). Tests of Equality between Set of Coefficients in two Linear Regressions, *Econometrica 28*, pp. 591-603.

Christiano, L. J. (1992). Searching for a Break in GNP. *Journal of Business and Economic Statistics*, 10, pp. 237-249.

Conover, W. J. (1971). Practical Non-Parametric Statistics, *John Wiley Pub.*, 462 pp.

Cox, D. R. et Spjotvøll E. (1982). On Partitioning Means into Groups, Scand. *J. Statist.*, 9, pp. 147-152.

Csörgö, M. et Horvàth (1988). Nonparametric Methods for Changepoint Problems. *Handbook of statistic*. Vol. 7 pp. 403-425.

David, I. S. et Robert K. K. (1997). Time Series Properties of Global Climate Variables : Detection and Attribution of Climate Change. *Working Papers* in Ecological Economics N^0 9702, 37 pages.

Deshayes, J. et Picard D. (1986). Off-line statistical Analysis of Change-Point Models. Detection of abrupt changes in Signals and dynamic Systems, *Lecture Notes* in Control and Information Sciences, 77, pp. 103-168.

Dickey, D. et W. Fuller (1979). Distribution of the Estimatiors for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association*. V. 74, pp. 427-431.

Dickey, D. et W. Fuller (1981). Likelihood Ration Statistics for Autoregressive Time Series with a Unit Root. *Econometrica*. V. 49, pp. 1057-1072.

Diebold, F. X. et Chen C. (1996). Testing Structural Stability with Endogenous Breakpoint : A size Comparaison of Analytic and Bootstrap Procedures, *Journal of Econometrics, 70*, pp.
221-241.

Doob, J. L. (1953). Stochastic Processes, Wiley

Dufour, J.-M. (1995). Monte Carlo Tests in the Presence of Nuisance Parameters with Econometric Applications, Working Paper, Centre de Recherche et Développement en Économique (CRDE) and Département de Sciences Économiques, Université de Montréal.

Dufour, J-M. (1980). Dummy Variables and Predictive Tests for Structural Change, *Economics Letters*, 6, pp. 241-247.

Dufour, J-M. (1981). Variables Binaires et Tests Prédictifs contre les Changements Structurels : une Application à l'Équation de St-Louis, *L'Actualité Économique*, 57, pp.376-385.

Dufour, J-M. (1982a). Generalized Chow Tests for Structural Change : A Coordinate-Free Approach, *International Economic Review*, 23, pp. 565-575.

Dufour, J-M. (1982b). Predictive Tests for Structural Change and St-Louis Equation, in Proceedings of the Business and Economic Statistics Section of the American Statistical Association, pp. 323-327, Washington (D.C.).

Dufour, J-M. (1982c). Recursive Stability Analysis of Linear Regression Relationships : An Exploratory Methodology, Journal of Econometrics, 19, pp. 31-76.

Dufour, J-M. (1986). Recursive Stability Analysis : The Demand for Money during the German Hyperinflation, in Model Reliability, ed. by D. A. Belsley, and E. Kuh, pp. 18-61. M.I.T. Press, Boston.

Dufour, J-M. (1989). Investment, Taxation and Econometric Policy Evaluation : Some Evidence on the Lucas Critique, in Statistical Analysis and Forecasting of Economic Structural Change, ed. by P. Hackl, pp. 441-473. Springer-Verglag, Berlin.

Dufour, J-M. et J. F. Kiviet (1994). Exact Inference Methods in First-Order Autoregressive Distribued Lag Models. *Discussion Paper, C.R.D.E*, Université de Montréal, and Faculty of Economics and Econometrics, University of Amsterdam.

Dufour, J-M. et J. F. Kiviet (1996). Exact Tests for Structural Change in First-Order Dynamic Models, Journal of Econometrics, 70, pp. 39-68.

Dufour, J. M. et Ghysels, E. (1996). Recent Developments in the Econometrics of Structural Change. *Journal of Econometrics*, Vol. 70, 317 pages.

Dufour, J.-M. et Khalaf L. (2001). Monte Carlo Test Methods in Econometrics. In : B. Baltagi, ed., Companion to Theoretical Econometrics, Blackwell Companions to Contemporary Economics, Chapter 23, pp.494-519. Oxford, U.K. : Basil Blackwell.

Dufour, J. M., E. Ghysels and A. Hall (1994). Generalized Predictive Tests and Structural Change Analysis in Econometrics, International Economic Review, 35, pp. 199-229.

Duquette, R. (1996). Description de la Nouvelle Modélisation retenue pour le problème lié à l'Aménagement Optimal d'un Nouveau Bassin. Rapport IREQ-96-351, Institut de Recherche d'Hydro-Québec, Varenne, Québec.

Dwass, M. (1957). Modified Randomization Tests for Non-parametric hypotheses, *Annals of Mathematical Statistics* 28, pp.181-187.

Edgington, E. S. (1995). Randomization tests. 3rd edition. Marcel Dekker Inc., New York.

Efron, B. (1982). The Jacknife, the Bootstrap and Other Resampling Plans, CBS-NSF *Regional Conference Series in Applied Mathematics*, Monograph N0. 38,
Philadelphia, PA : Society for Industrial and Applied Mathematics

El-Shaarawi, A. H. and Damsleth (1988). Parametric and Nonparametric tests for Dependent Data. *Water Resources Bulletin, American Water Resources Association*, Vol. 24 No. 3, pp. 513-519.

Farrell, R. (1980). Methods for Classifying Changes in Environmental Conditions. Technical Report VRF-EP AF.4-FR80-1, *Vector Research Inc.*, Ann. Arbor., Michigan.

Faucher, D., T. B. M. J. Ouarda et B. Bobée (1997). Revue Bibliographique des Tests de Stationnarité. Québec, *I.N.R.S.-Eau*, 66 pp

Ferger, D. et Stute, W. (1992). Convergence of changepoint estimators. *Stochastic Processes and their Applications*. North-Holland. 42, pp.345-351.

Fisher, R, A. (1932). Statistical Methods for Research Workers, Oliver and Boyd, Edinburgh 4th Ed.

Fisher, R. (1934). Two New Properties of Mathematical Likelihood, Proceedings of the Royal Society of London A 144, pp. 285-307.

Fisher, R. (1935). The Design of Experiments (8 th ed., 1966). Edinburgh: Oliver & Boyd.

Folks, J. L. (1984). Combination of Independent Tests. In Handbook of Statistics, Volume 4, Non Parametric Methods, Ed.. bq P. R. Krishnaiah & P.K. Sem. Amsterdam : North Holland, pp. 113-121

Freeman, J. M. (1986). An Unknown Change-Point and Goddness of Fit, *the Statistician*, 35, pp. 335-344.

Fuller, W. A. (1976). Introduction to Statistical Time Series. New York, *John Wiley*.

Galbraith, J. W. et C. Green (1990). Trends and Stationarity in Global Temperature Data. Centre for Climate and Global Change Reserarch (C^2GCR), McGill University, Report No. 12pp.

Gan, T. Y. (1998). Hydroclimatic Trends and Possible Climatic Warming in the Canadian Prairies. *Water Resources Research*, Vol. 34, N^0 11, pp.3009-3015.

Gardner, L. A. Jr. (1969). On detecting changes in the mean of normal variates. *Ann. Math. Stat.* 40. pp. 114-115.

Ghorbanzadeh, D. (1995). Un Test de Détection de Rupture de la Moyenne. *Revue de Statistique Appliquée, XLIII* (2), pp. 68-76.

Girault, M. (1965). Processus Aléatoires, Dunod

Good, P. (1994). Permutation Tests : A Pratical Guide to Resampling Methods for Testing Hypotheses. New York : Springer-Verlag New York.

Gouriéroux, C. and A. Monfort (1994). Simulation Based Econometric Methods. *Technical Report*, INSEE, Paris.

Gosset, W. S. (1908). The probable error of the mean., *Biometrika*, 6, pp. 1-25.

Hawkins, D. M., and D. H. Olwell (1998). Cumulative Sum Charts and Charting for Quality Improvement. Springer-Verlag, New York.

Hinkley, D. V. (1969). Inference about the Intersection in Two-Phase Regression, *Biometrika*, 56, pp. 495-504.

Hinkley, D. V. (1970). Inference about the change-point in a sequence of binomial random variable. *Biometrika Vol. 57*. pp. 477-488.

Hinkley, D. V. (1971). Inference in Two-Phase Regression. *Journal of the American Statistical Association*, 66, pp. 736-743.

Hipel, K. W. et McLeod A. I. (1994). Time Series Modelling of Water Resources and Environmental Systems, Elsevier. Developpements in *Water Science, 45*, 1013 pp.

Hirsch, R. M et E. S. Gilroy (1985). Detectability of Step Trends in the Rate of Atmospheric Deposition of Sulphate. *Water Resources Bulletin*, 21 (5) : pp. 773-784

Hirsch, R. M. et J. R. Slack , R. A. (1984). Non-parametric Trend Test for Ssaisonal Data with Serial Dependance. *Water Resources Research*, 20 (6), pp. 727-732.

Hirsch, R. M., J. R. Slack , R. A. Smith (1982). Techniques of Trend Analysis for Monthly Water Quality Data. *Water Resources Research*, 18 (1), pp. 107-121.

Hoeffding, W. (1950). Optimum Non-parametric Tests. Proc. *11nd Berkeley Symposium*, pp.83-92.

Hogg, R. V., and Craig, A. T. (1978). Introduction to Mathematical Statistics, Fourth Edition, Macmillan Publishing Co., Inc. New York.

Holm, G. (1979). A Simple Sequentially Rejective Multiple Test Procedure. Scand. *J. Statist.* 6, pp. 65-70

Hsu, D. A. (1979). Detecting, shifts of parameter in gamma sequences with applications to stock price and air traffic flow analysis. *Journal of the American Statistical Association*, 74, pp. 31-40.

Hubert, P., J. P. Carbonnel J. P., A. Chaouche (1989). Segmentation des Séries Hydro-météorologiques . Application à des séries de précipitation et de débits de l'Afrique de l'Ouest. *Journal of Hydrology*, Vol. 110, pp. 349-367.

Huberty, C. (1993). Historical origins of statistical testing practices: The treatment of Fisher versus Neyman-Pearson views in textbooks. Journal of Experimental Education, 61, pp. 317-333.

Jaffard, P. (1973). Initiation aux Méthodes de la Statistique et du Calcul des Probabilités, Masson et C^{ie}.

James, B. James K. L. et Siegmund D. (1987). Test for a change-point, *Biometrika*, 74, pp. 71-83.

Jaruskova, D. (1996). Change-Point Detection in Meteorological Measurement. *Monthly Weather Review*, 124, pp. 1535-1543.

Jaruskova, D. (1997). Somme Problems With applications of Change-Point Detection Methods to Environmental Data. *Environmetrics*, Vol. 8. Pp-469-483.

Jonckheere, A. R. (1954). A Distribution-Free k-Sample Test against Ordered Alternatives. *Biometrika*, 41 : pp.133-145.

Kendall, M. G. (1975). Rank Correlation Methods. *4th ed.* Charles Griffin, London.

Kendall, M. S. and Stuart A. (1943). The Advanced Theory of Statistics. Charles Griffin Londres. 2ème volume, 690 pp.

Khaled, H. A. Hamed et R. Rao (1998). A modified Mann-Kendall Trend Test for Autocorrelated Data. *Journal of Hydrology*, 204, pp. 182-196.

Khodja, H., Lubes-Niel H., Sabatier R., Masson J. M., Servat E. et Paturel J. E. (1998). Analyse Spatio-Temporelle de Données Pluviométriques en Afrique de l'Ouest. Recherche d'une rupture en moyenne. Une Alternative Intéressante : les Tests de Permutations. *Revue de Statistique Appliquée*. XLVI (1), pp. 95-110.

Kim, H. J. (1996). Change-Point Detection for Correlated Observations. *Statistica Sinica*, 6, pp. 275-287.

Kim, H. et Siegmund D. (1989). The Likelihood Ratio Test for a Change-Point in Sample Linear Regression, *Biometrika, 76*, 3, pp. 409-23.

Kruskall, W. H et W. A. Wallis (1952). Use of ranks in one-criterion Variance Analysis. *Journal of American Statistical Association*, 48 : pp.583-612.

Laberge, C. (1995). Utilisation de la Régression Linéaire pour la Détection de Tendance dans des Séries Chronologiques Environnementales. Thèse de doctorat présentée à l'INRS-Eau, Université du Québec, 214pp.

Laplace, P. (1814). Essai Philosophique sur les Probabilités. Paris : Courcier.

Lee, A. F. S. et S. M. Heghinian (1977). A Shift of the Mean Level in a Sequence of Independent Normal Random Variables. A Bayesian Approach. *Technometrics* 19 (4), pp. 503-506.

Lehmann, E. L. (1959). Testing Statistiqcal Hypotheses. John Wiley & Sons, Inc., New York.

Lettenmaier, D. P. (1976). Detecting Trends in Water Quality Data from Records with Dependent Observations, *Water Resources Research*, 12 (5), pp. 1037-1046.

Lettenmaier, D. P. (1984). Limitations on Seasonal Snowmelt Forecast Accuracy. J. Water. Resour. Plann. Manage., ASCE, 110 (3) : pp. 255-269.

Littell, R. C. and Folks, J. L. (1971). Asymptotic Optimality of Fisher's Method for Combining Independent Tests. *Journal of the American Statistical Association*, 66 : 802-806.

Lubès, N. Masson J. M., Paturel J. E., Servat E. (1998). Variabilité climatique et statistiques. Étude par simulation de la puissance et de la robustesse de quelques tests utilisés pour vérifier l'homogénéité de chroniques. *Revue des sciences de l'Eau*, 11 (3), pp. 384-408.

Mann, H. B. (1945). Non-parametric Tests against Trend. *Econometrica*, 13 : pp. 245-259.

Mann, H. B. et D. R. Whitney (1947). On the test of Whether one of two Random Variables is stochastically Larger than the other. *Ann. Math. Statist.*, 18 : pp. 50-60.

Maronna, R. et V. J. Yohai (1978). A Bivariate Test for the Detection of a Symetric Change in Mean. *Journal of American. Statistical. Association.*, 73 : pp 640-645.

Mathier, L., Fagherazzi, L., Rassam, J. C. et Bobée B. (1992). Great Lakes net Basin Supply Simulation by a Stochastic Approach. Rapport de Recherche No R-362, Université du Québec, Québec Canada..

Nelson, C. R. et C. Plosser (1982). Trends and Random Walks in Macroeconomic Time Series : Somme Evidence and Implications. *Journal of Monetary Economics*, 10, pp. 139-162.

Ouarda, T.B.M.J., Haché M., Rasmussen, P. F. et Bobée B. (1998). Risque et Fiabilité en Hydrologie. Rapport de Recherche C6, INRS-Eau, Université du Québec, Québec Canada.

Ouarda, T.B.M.J., Rasmussen P. F., Cantin J.-F., Bobée B., Laurence R., Hoang V. D. et Barabé G. (1999). Identification d'un Réseau hydrométrique pour le suivi des Modifications Climatiques dans la Province de Québec. *Revue des Sciences de l'Eau*, 12/2 pp. 425-448.

Page. E. S. (1954). Continuous inspection schemes. *Biometrika, Vol. 41*, pp. 100-155.

Page. E. S. (1955). A test for a change in a parameter occurring at an unknown point. *Biometrika Vol 42* pp. 523-526.

Page. E. S. (1957). On problems in which a change of parameters occurs at an unknown point. *Biometrika Vol 44* pp.248-252.

Pearson, K. (1900). On a criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can reasonably be supposed to have arisen in random sampling. Philosophical Magazine, 50, pp. 157-175.

Pearson, K. (1933). On a Method of Determining Whether a Sample of Side n Supposed to have been drawn from a Parent Population having a known Probability Integral has Probably been drawn at Random. *Biometrika 25*, pp. 379-410.

Perreault, L. (2000). Analyse Bayésienne Rétrospective d'une Rupture dans les Séquences de Variables Aléatoires Hydrologiques. Thèse présentée à l'INRS-Eau, Université du Québec,

Perreault, L., Haché M. Bobée B. (1996). Analyse Statistique des Séries d'apports Énergétiques. Rapport de Recherche No R-474, INRS-Eau, Université du Québec, Québec Canada.

Perreault, L., Haché M., Slivitzky M., Bobée B. (1999). Detection of Change in Precipitation and Runoff over Eastern Canada and U.S. using a Bayesian Approach. *Stochastic Environmental Research and Risk Assessment*, 13 (3), pp. 201-216.

Perron, P (1997). L'Estimation de Modèles avec Changements Structurels Multiples. *L'Actualité Économique, Revue d'Analyse Économique*, Vol., 73, No 1-2-3.

Pettitt, A. N. (1979). A Non-Parametric Approach to the Change-Point Problem. *Applied Statistics*, 28 N^0 2, pp. 126-135.

Quandt, R. (1958). Tests of Hypothesis that a Linear Regression obeys two Separate Regimes. *Journal of the American Statistical Association, 55*, pp. 324-330.

Rasmussen, P. F., Haché M. et Bobée B. (1998). Modélisation Stochastique des Apports Naturels. Rapport de Recherche C4, No R-533, INRS-Eau, Université du Québec, Québec Canada

Refsgaard, J. C., Alley W. M., Vuglinsky V. S. (1989). Methodology for Distinguishing between Man's Influence and Climatic Effets on the Hydrological Cycle. Technical Documents in Hydrology. International Hydrological Programme United Nations Educational Scientific and Cultural Organization. Unesco Paris,

Salas, J. D., J. W. Delleur, V. Yevjevich et W. L. Lane (1980). Applied Modeling of Hydrologic Time Series, *Water Resources Publications*, 484 pages.

Sen, P. K. (1968). Estimates of the Regression Coefficient based on Kendall's tau. *Journal of American Statistical Association*, 63 : pp. 1379-1389.

Sen, A. et Srivastava, M. S. (1975). On tests for detecting change in mean. *Ann. of Stat.* 3 pp. 90-108.

Searcy, J. K. and Hardison C. H. (1960). Double-Mass Curves. In : Manual of Hydrology, Part I. General Surface Water Techniques. U.S. Geol. Surv., *Water-Supply Pap.*, 1541-B : pp. 31-59.

Siegmund, D. (1986). Boundary Crossing probabilities and statistical applications. *Ann. Statisti.* 14. Pp. 361-404.

Slivitzky, M. and Mathier L. (1994). Climatic Changes during the 20th century on the Laurentian Great Lakes and their Impacts on Hydrologic Regime in Engineering Risk in Natural Resources Management, L. Duckstein and E. Parent (eds.), NATO ASI Series E, Vol. 275, Kluwer : pp. 235-251.

Smith. A. F. M. (1975). A Bayesian approach to inference about change-point in a sequence of random variables. *Biometrika Vol. 62*. Pp. 407-416.

Sneyers, R. (1990). Sur l'analyse statistique des séries d'observations. OMM, Genève, *Note Technique* N° 143, 192 pp.

Solow, A. R. (1988). A Bayesian Approach to Statistical Inference about Climatic Change. *Journal of Climate*, Vol.1 : pp 512-521.

Taylor, C. et J. C. Loftis (1989). Testing for Trend in Lake and Groung Water Quality Time Series. *Water Resources Bulletin*, 25 (4) : pp. 715-726.

Terpstra, J. J. (1952). The Asymptotic Normality and Consistency of Kendall's Test agains Trend when Ties are Present in one ranking. *Indag. Math.*, 14 : pp. 327-333.

Terry, M. E. (1952). Some Rank Order Tests Which are most Powerful against Specific Parametric Alternatives. *Ann. Math. Statist.* 23 : pp. 346-366.

Tippett, L. H. C. (1931). The Methods of Statistics, Williams and Norgate, London, 1st. Ed.

Van Belle, G. et Hughes J. P. (1984). Nonparametric Tests for Trend in Water Quality. *Water Resources Research, 20* (1) : pp. 127-136.

Van der Waerden, B. L. (1952). Order Tests for the Two-Sample Problem and their Power I, II, III Indigationes math. 14 (Proc. Kon. Nederl. Akad. Wet. 55), pp. 453-458 ; *Indigationes math.* 15 (Proc. 56), pp. 303-310, pp. 311-316; Correction : *Indigationes math.* 15 (Proc. 56), 80.

Vincent L. (1990). Time Series Analysis : Testing the Homogeneity of Monthly Temperature Series. *Survey Paper* 90-105, York University, 50 pages.

Vostrikova, L. Ju. (1981). Detecting Disorder in Multidimensional random process. *Soviet Math. Dokl.* 24, pp. 55-59.

Westmacott, J. R. and Burn D. H. (1997). Climate Change Effects on the Hydrologic Regime within the Churchill-Nelson River Basin. *Journal of Hydrology* 202, pp. 263-279.

Wilkinson, B. (1951). A Statistical Consideration in Psychological Research. *Psychology Bulletin 48*, pp. 156-158.

Wonnacot, T. H., R. J. Wonnacott (1980). Regression a Second Course in Statistics. *Wiley*, New York, 556 pp.

Woodward, W. A. and Gray H. L. (1993). Global Warming and the Problem of Testing for Trend in Time Series Data. Journal of Climate, V. 6, pp. 953-962.

World Meteorological Organization (WMO, 1966). Climatic Change, by a working Group of the Commission for Climatology. World Meteorological Organization, WMO 195, TP 100, *Technical Note N0 79* : 78pp.

Worsley, K. J. (1979). On the Likelihood Ratio Test for a Shift in Location of Normal Populations. *Journal of the American Statistical Association,* 74 pp. 365-367.

Worsley, K. J. (1983). The Power of Likelihood Ratios and Cumulative Sum Tests for a Change in Binomial Probability. *Biometrika Vol. 70* pp. 91-104.

Worsley, K. J. (1986). Confidence Regions and Tests for a Change-Point in a Sequence of Exponential Family Random Variables. *Biometrika Vol. 73*, pp. 455-464.

Zetterqvist, L. (1991). Statistical Estimation and Interpretation of Trends in Water Quality Time Series. *Water Resources Research*, 27 (7), pp. 1637-1648.

Appendice A

Vocabulaire de base en statistique

Notions d'expérience stochastique ou aléatoire, d'épreuve, d'ensemble fondamental, d'événement, de tribu et de probabilité

Une *expérience stochastique* ou *aléatoire* est une expérience où on ne peut prévoir de façon certaine (avant l'expérience) le résultat de l'expérience, mais ce résultat est clairement identifiable. On appelle *épreuve* une expérience qui conduit à un résultat et un seul, qui en général est imprévisible mais qui appartient à un ensemble de résultats possibles. On notera cet ensemble Ω tandis que l'épreuve sera notée ξ. L'ensemble Ω est aussi mieux connu sous le nom d'ensemble *fondamental* ou *ensemble des épreuves* ou *univers* ou encore *population statistique*. Un élément, ω, de cet ensemble est qualifié de *résultat* ou *individu*. Selon la nature de l'expérience, l'ensemble fondamental Ω peut être *fini* (donc de la forme $\Omega = \{\omega_1, \ldots, \omega_N\}$), *infini dénombrable* (c'est-à-dire pouvant être mis en bijection avec tout ou partie de l'ensemble $\mathbb{N}$ des entiers naturels) ou *infini non dénombrable* (exemple un intervalle). Étant donné une épreuve ξ d'univers Ω, un *événement* associé à ξ est une proposition dont on peut dire, pour chaque résultat de l'épreuve, si elle est vérifiée, ou pas. L'événement dont on a la certitude de la réalisation est désigné par Ω (*événement certain*), tandis que l'événement dont on a la certitude qu'il ne se produira pas est désigné par $\varnothing$ (*événement impossible*). Parmi les classes d'événements que l'on est amené à considérer, les classes qui sont telles que si on effectue les opérations élémentaires $(\cap, \cup, \boldsymbol{C})$ on obtienne encore un élément de cette classe, jouent un rôle important.

Définition 1. *Une classe* $\boldsymbol{A}$ *de parties de* Ω *est appelée tribu ou* σ_algèbre *si elle vérifie les trois axiomes suivants :*

a) $\boldsymbol{A}$ *contient les parties* $\varnothing$ *et* Ω *:* $\varnothing \in \boldsymbol{A}$ et $\Omega \in \boldsymbol{A}$

b) $\boldsymbol{A}$ *est stable par complémentation :* $(\mathrm{A} \in \boldsymbol{A}) \Rightarrow (\boldsymbol{C}_{\mathrm{A}} \in \boldsymbol{A})$

c) $\boldsymbol{A}$ *est stable par réunion dénombrable :* $(\mathrm{A}_n \in \boldsymbol{A}),\ (\forall n), (n \in \mathrm{N}); \left(\bigcup_{n \in \mathrm{N}} \mathrm{A}_n \in \boldsymbol{A} \right)$

En somme, le premier élément descriptif d'un phénomène aléatoire consiste à préciser ce qui peut se produire. Il importe donc de définir rigoureusement l'ensemble Ω de telle sorte que toutes les réalisations possibles du phénomène (caractéristiques observables du phénomène) puissent être décrites comme les éléments de cet ensemble. Le deuxième élément de ce modèle est la tribu $\boldsymbol{A}$ qui est la famille des événements associés à Ω. Ainsi,

une fois défini ces deux éléments, on dit que le couple $(\Omega, \mathcal{A})$ est un *espace probabilisable*, c'est-à-dire l'espace sur lequel on peut définir des *probabilités*.

Définition 2. *Étant donné un phénomène aléatoire auquel est associé un espace probabilisable* $(\Omega, \mathcal{A})$, Ω *ensemble des réalisations,* $\mathcal{A}$ *une tribu des parties de* Ω, *on appelle probabilité sur* $(\Omega, \mathcal{A})$, *l'application* $P: \mathcal{A} \to [0,1]$ *telle que :*

- $P(\Omega) = 1$;
- *Pour des événements* $\{A_i \in \mathcal{A}\}$ *incompatibles* $(i \neq j \Rightarrow A_i \cap A_j = \varnothing)$:

$$P\left(\bigcup_i A_n\right) = \bigcup_i P(A_n), \quad i \in \mathbb{N}$$

Notion de variable aléatoire

Une *variable aléatoire* est une grandeur numérique attachée au résultat d'une épreuve. Chacune de ses valeurs est associée à une probabilité d'apparition :

Définition 3. *Soit* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé. On appelle variable aléatoire* X *toute application mesurable, telle que :*
$X : (\Omega, \mathcal{A}) \to (\Omega', \mathcal{A}')$ *possédant les propriétés suivantes :*

- $(\forall A' \in \mathcal{A}')$; $X^{-1}(A') \in \mathcal{A}$ (propriété de mesurabilité)
- $(\forall A' \in \mathcal{A}')$; $P_X(A') = P\left(X^{-1}(A')\right)$

Définition 4. *Soit* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé. On appelle variable aléatoire réelle X une application mesurable, telle que :* $X : (\Omega, \mathcal{A}) \to (\mathbb{R}, \mathcal{B})$
où : $\mathbb{R}$ *est l'ensemble des nombres réels* et $\mathcal{B}$ *est la tribu des boréliens sur* $\mathbb{R}$

Une fois que l'on a défini cette transformation de X et Ω dans $\mathbb{R}$, il devient alors intéressant de définir une fonction qui à chaque valeur x prise par la variable aléatoire X va faire correspondre la probabilité que X prenne la valeur x. Ainsi, à l'aide de X, on introduit donc une loi de probabilité sur l'ensemble des valeurs de la variable aléatoire X. On l'appelle loi de la variable aléatoire X.

Définition 5. *Soit un espace probabilisé* $(\Omega, \mathcal{A}, P)$ et $\mathfrak{R} \in \mathcal{B}$ ($\mathcal{B}$ tribu borélienne sur $\mathbb{R}$). *On appelle loi de la variable aléatoire X la loi de probabilité* P_X *sur* $\mathbb{R}$, *définie pour tout borélien* $\mathfrak{R}$ *de* $\mathbb{R}$ *par :* $P_X(\mathfrak{R}) = P(X \in \mathfrak{R}) = P(X^{-1}\{\mathfrak{R}\})$
où : $X^{-1}\{\mathfrak{R}\} \in \mathcal{A}$

Notions d'échantillon et de statistique

La majorité des phénomènes réels, qu'ils soient provoqués ou simplement observés, entrent dans un schéma général que l'on a désigné sous le nom d'épreuve. Ainsi, si l'on répète une expérience un nombre fini de fois, soit n ; on disposera d'un ensemble $\{x_1,\ldots,x_n\}$ appelé *échantillon observé* de valeurs de la variable aléatoire X. un échantillon observé de taille n apparaît donc comme l'ensemble de n résultats d'une même épreuve. Le rôle d'un échantillon est de fournir des renseignements sur la population du phénomène et plus particulièrement sur certains paramètres caractérisant cette population. Plus l'échantillon est de grande taille, plus on doit espérer que les renseignements seront précis. Le mot échantillon prend généralement en statistique deux sens différents, selon que l'on parle des données observées ou du modèle probabiliste. En effet, l'hypothèse de modélisation d'un phénomène aléatoire consiste à voir l'échantillon observé comme une réalisation d'un échantillon théorique $\{X_1,\ldots,X_n\}$.

Soient $\{X_1,\ldots,X_n\}$ l'échantillon théorique de la variable aléatoire X , on appelle *statistique* sur un échantillon toute fonction mesurable $\varphi(X_1,\ldots,X_n)$ de l'échantillon $\{X_1,\ldots,X_n\}$. Une statistique est donc une variable aléatoire dont on peut chercher à déterminer les caractéristiques (loi de probabilité, fonction de répartition, moments, ...).

Les définitions qui ont été données dans cet appendice sont conforment à celles que l'on trouve dans les ouvrages de référence de Lehmann (1959), de Jaffard (1973) et de Hogg et Craig (1978).

Appendice B

Postulat de base en statistique fréquentielle

Généralités

Les méthodes statistiques dites *classiques* ont pour objectif principal de faire, au vu d'observations de l'échantillon, une inférence au sujet de la loi générant ces observations. Dans l'inférence statistique traditionnelle, les seules probabilités considérées sont les probabilités d'échantillonnage, conditionnelles au paramètre (ou caractéristique de la population). Ces probabilités peuvent s'interpréter comme des fréquences. C'est pourquoi un grand nombre de statisticiens en faveur de cette approche dite classique l'appelle *inférence fréquentiste*. On peut regrouper les méthodes employées en analyse fréquentielle en deux catégories : les *estimations* et les *tests d'hypothèses*. Les différentes étapes d'une analyse fréquentielle peuvent être représentées très succinctement selon le diagramme de la figure B1.1. Sur cette figure, ce sont essentiellement les étapes 3 et 5 qui concernent le problème de stationnarité. C'est donc elles qui sont au centre de cette thèse.

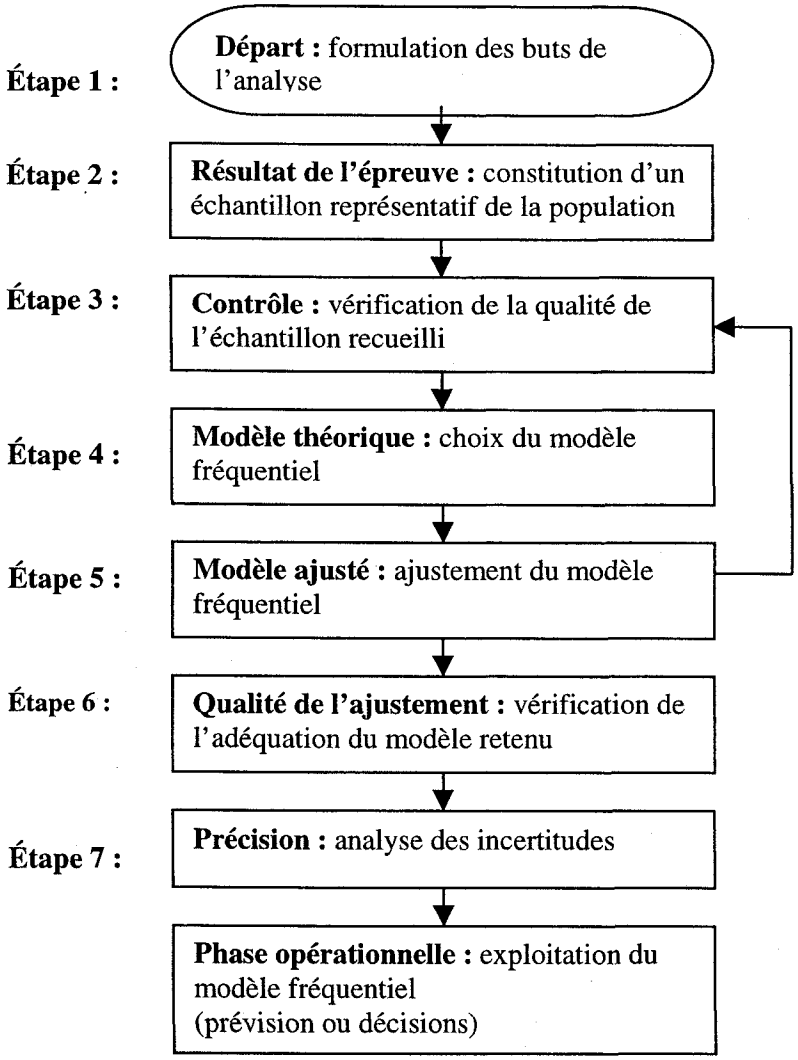


Figure B1.1 : Principales étapes de l'analyse fréquentielle

Principe de base en analyse fréquentielle

Lorsqu'on définit la nature de la population statistique et le mode d'échantillonnage (voir appendice A), on établit en fait un modèle statistique (c'est-à-dire une formulation mathématique des hypothèses faites sur les observations). L'emploi des méthodes d'analyse fréquentielle est donc subordonné à de strictes conditions d'application et l'interprétation correcte des résultats dépend de la vérification du respect de ces conditions. Le postulat de modélisation sur lequel toute étude d'analyse fréquentielle est basée est :

> ***les données observées sont des réalisations de variables aléatoires indépendantes et de même loi de probabilité (Hypothèse iid** : indépendantes et identiquement distribuées).*

Ce postulat permet, dans une certaine mesure de substituer à la notion purement numérique d'échantillon numérique (échantillon observé) le concept probabiliste d'échantillon aléatoire (échantillon théorique). En somme, l'intuition derrière ce postulat est que, le mot échantillon prend en statistique deux sens différents, selon que l'on parle de données observées ou du modèle probabiliste. L'hypothèse de modélisation consiste donc à voir l'échantillon observé comme une réalisation d'un échantillon théorique d'une certaine distribution de probabilité F. En d'autres termes, on considère que les données qui composent l'échantillon observé auraient pu être produites en simulant de façon répétée la distribution de probabilité F comme cela est particulièrement illustré au tableau B1.1.

Échantillons observés	**Échantillon théorique**				
	X_1	X_2	$\cdots$	X_i	X_n
échantillon 1	x_{11}	x_{21}	$\cdots$	x_{i1}	x_{n1}
échantillon 2	x_{12}	x_{22}	$\cdots$	x_{i2}	x_{n2}
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
échantillon k	x_{1k}	x_{2k}	$\cdots$	x_{ik}	x_{nk}
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$

Tableau B1.1 : Illustration d'une génération d'échantillons observés à partir d'un échantillon théorique

La conséquence directe du postulat de base en analyse fréquentielle est qu'il permet de fonder l'induction, en donnant le moyen, à partir de l'information nécessairement incomplète fournie par des échantillons observés, de formuler certains jugements concernant les populations d'où ils sont prélevés. En effet, si l'on considère la population infinie d'échantillons observés de taille n (comme le montre le tableau B1.1) qu'on peut extraire de la population statistique étudiée, les valeurs prises par une certaine statistique (par exemple la moyenne arithmétique) dans ces différents échantillons observés ne sont évidemment pas identiques, elles se répartissent suivant une loi de probabilité ou distribution d'échantillonnage de la statistique. La forme analytique de cette loi et les paramètres dont elle dépend sont fonction de celle et de ceux de la loi de probabilité de la population dont proviennent les différents échantillons observés. Cette fluctuation échantillonnale fournit un bon moyen de juger de la stabilité de la loi de probabilité de la population de base. De plus, elle permet de juger de la stabilité des estimations ainsi que de l'écart maximal entre une estimation ponctuelle et la vraie valeur du ou des paramètres inconnus de la population.

En somme, le principe de base en analyse fréquentielle stipule que les séries de données observées sont des réalisations de variables aléatoires indépendantes et de même loi de probabilité. La théorie des probabilités fournit des outils, comme la loi des grands nombres, permettant d'extraire des données ce qui est reproductible et qui pourra donc fonder une prédiction ou une décision. Cela dit, quand les hypothèses de modélisation conduisent à supposer que les données observées sont des réalisations de variables aléatoires indépendantes et de même loi de probabilité, la loi des grands nombres justifie que l'on considère cette loi comme proche de la distribution empirique. Toutes les caractéristiques usuelles de la distribution empirique seront proches des caractéristiques analogues de la loi théorique.

Appendice C

Théorie des tests : approche de Fisher et de Neyman et Pearson

La paternité des tests statistiques a toujours été attribuée à tord à Ronald Fisher, Jerzy Neyman, Egon Pearson ou même au père de ce dernier, Karl Pearson. En effet, le premier document qui mentionne l'utilisation d'un test statistique est celui de Arbuthnot (1710). On trouve également des tests statistiques dans les travaux de Laplace en 1814 (Laplace, 1814). C'est en 1900, que Karl Pearson a proposé le test de Chi-deux (χ^2) qui est un test d'adéquation d'une distribution théorique à une distribution observée (Pearson, 1900). Gosset (dont le pseudonyme est Student) fera de même peu de temps après à l'aide de la statistique *t* afin de comparer les moyennes de deux populations (Gosset, 1908). Profitant principalement des travaux de Laplace (1814), Fisher développera une méthode générale de détermination des intervalles de confiance pour un paramètre (Fisher, 1935). Puis, c'est finalement Neyman qui développera la théorie générale des intervalles de confiance et le lien avec la théorie des tests. Bien que Fisher, Neyman, Pearson et fils aient présenté leurs travaux plus tardivement, cela ne diminue en rien leur contribution au développement formel des tests statistiques au $20^{ème}$ siècle (Huberty, 1993). Pour comprendre la logique d'élaboration des tests qui est en vigueur depuis près d'un siècle, on présente dans cet appendice quelques fondamentales de Fisher et de Neyman et Pearson.

Généralités

Bien souvent, nous conduisons une recherche de façon à déterminer l'acceptabilité d'hypothèses découlant de nos connaissances théoriques. Au vu des informations directes que nous récoltons sur des données empiriques, notre décision concernant la signification des données nous conduit alors soit à retenir, soit à réviser ou soit à rejeter de telles hypothèses et la théorie qui en est la source. Un test est un procédé permettant de décider si une hypothèse donnée, appelée hypothèse principale et notée généralement H_0, peut être considérée comme vraie ou fausse. Ce procédé suit généralement les 7 étapes suivantes : (1) établir l'hypothèse principale H_0 ; (2) choisir la statistique du test (*T*) pour tester H_0 ; (3) spécifier un niveau de signification (α) et la taille de l'échantillon (*n*) ; (4) trouver la distribution d'échantillonnage de la statistique du test sous H_0 ; (5) sur la base de (2), (3) et (4), définir la région de rejet du test ; (6) calculer la valeur observée de la statistique du test à partir des données de l'échantillon (T_0) ; (7) décider (rejeter ou accepter H_0).

La prise de décision (étape 7) dans un test statistique est toujours basée sur une information partielle (échantillon). Il en résulte forcément un risque d'erreur inévitable due au processus d'échantillonnage. Du point de vue de la statistique fréquentiste, il existe deux grandes approches découlant de l'interprétation du risque d'erreur encourue dans l'élaboration d'un test : l'approche de Fisher et l'approche de Neyman et Pearson.

Approches de Fisher et de Neyman et Pearson dans l'élaboration d'un test statistique

Pour bien comprendre la pertinence de l'approche de Fisher et de celle de Neyman et Pearson, on va tout d'abord définir la notion de *p-valeur* (en anglais, "*p-value*"). La p-valeur d'un test est la probabilité, sous H_0, que la statistique de test T prenne une valeur aussi extrême ou plus extrême que la valeur observée T_0. Sans perte de généralité, la p-valeur à droite se définit comme : $p-\text{valeur} = P(T \geq T_0 / H_0)$

Dans la procédure de prise de décision (étape 7), la $p-\text{valeur}$ est perçue comme une mesure du poids de l'évidence d'un test statistique. Elle donne une mesure précise du risque que l'on est prêt à tolérer pour rejeter l'hypothèse H_0 quand elle est vraie. Ainsi, si la $p-\text{valeur}$ est plus petite que la probabilité d'erreur que l'on est prêt à tolérer (α), on rejette H_0. Si par contre, la $p-\text{valeur}$ est plus grande que α, la conclusion qui s'impose diffère selon que l'on opte pour la logique de Fisher ou la logique de Neyman et Pearson. C'est le point de divergence des deux approches.

Pour Fisher, on doit accorder plus de poids au test de l'hypothèse principale H_0. C'est pourquoi la logique de Fisher est communément rapportée des *tests de signification*. Dans un test de signification il n'y a que l'hypothèse principale H_0 qui est explicitement posée. Le niveau exact de signification est exclusivement une propriété des données, c'est-à-dire une relation entre un ensemble de données et une hypothèse principale, ici H_0. En d'autres termes, le niveau de signification informe sur la nature vraie ou fausse de H_0 dans une expérience particulière. Cette interprétation rapproche Fisher d'une position quasi-Bayésienne dans la mesure où, en théorie bayésienne, le théorème de Bayes permet de

calculer la probabilité d'une hypothèse en fonction des données récoltées. Ainsi, pour Fisher, le cas de figure où la $p-$valeur est plus grande que la probabilité d'erreur tolérée n'amène pas de décision conclusive. On doit simplement prendre un jugement de réserve en précisant que l'hypothèse principale H_0 doit être tolérée jusqu'à preuve du contraire, en l'occurrence de nouvelles expériences examinant H_0. En somme, l'approche de Fisher conçoit le test statistique comme une procédure d'épreuve de propositions scientifiques dans un contexte d'information limitée (échantillon). Cette approche favorise alors la suspension du jugement : autrement dit si on est amené à rejeter H_0, la théorie est considérée comme vraie jusqu'à nouvel ordre, et si l'on accepte (on ne peut rejeter H_0) on rejette la théorie.

Pour Neyman et Pearson on doit introduire deux hypothèses dans l'élaboration d'un test, l'hypothèse principale, H_0 et la contre-hypothèse H_1. C'est pour cette raison que la logique de Neyman et Pearson se rapporte à ce que l'on nomme généralement *tests d'hypothèses*. Dans les tests d'hypothèses, lorsque l'hypothèse principale est vraie, elle sera quand même rejetée avec une fréquence relative ou probabilité α qui dépend de l'emplacement des limites des régions d'acceptation et de rejet. Ce rejet erroné de H_0 est appelé une *erreur de première espèce* (ou de *type I*). Si par contre l'hypothèse principale H_0 est fausse, elle sera quand même acceptée avec une probabilité β qui dépend aussi de l'emplacement des limites des régions d'acceptation et de rejet. Cette acceptation erronée de H_0 est appelée une *erreur de seconde espèce* (ou de *type II*). Le risque de première espèce α est une propriété du test en ce sens qu'il doit ordinairement être spécifié avant même que l'évidence statistique soit obtenue. On peut illustrer les cas possibles d'une décision avec l'approche de Neyman et Pearson dans le tableau C1.1. Les deux cases vides de ce tableau correspondent aux probabilités complémentaires à 1 de α et de β, mais ne traduisent pas des risques puisque dans les deux cas il n'y a pas d'erreur de décision. Dans celle de la première ligne s'inscrirait la probabilité de retenir l'hypothèse H_0 quand celle-ci est vraie; cette probabilité doit être normalement élevée. En revanche, dans la case vide de la deuxième ligne se trouverait l'expression $(1-\beta)$; c'est-à-dire, la probabilité de

rejeter l'hypothèse H_0 quand celle-ci est fausse. Cette dernière probabilité, appelée *puissance d'un test*, est tout naturellement retenue comme caractéristique désirable, car on souhaite détecter la contre-hypothèse H_1 lorsque c'est cette dernière qui est vraie.

États de H_0	Décisions	
	Ne pas rejeter H_0	*Rejeter* H_0
H_0 *est vraie*		Mauvaise décision Erreur de type I ou erreur de 1ère espèce
H_0 *est fausse*	Mauvaise décision Erreur de type II ou erreur de 2ème espèce	

Tableau C1.1 : Risques d'erreur d'un test d'hypothèses

Ainsi, pour Neyman et Pearson, une fois que α est fixé dans la procédure de test, on rejetterait H_0 si la p-valeur est plus petite que α. Le rejet de H_0 suggère alors que H_1 pourrait être vraie. Par contre si la p-valeur est plus grande que α, la décision serait conclusive, c'est-à-dire, "ne pas rejeter H_0". Toutefois, cela ne signifie pas nécessairement que H_0 est vraie mais suggère seulement qu'il n'y a pas de preuve suffisante contre H_0 en faveur de H_1.

Cela dit, dans le schéma Fishérien , il n'y a que l'hypothèse principale H_0 qui est explicitement posée tandis que, chez Neyman et Pearson, il existe deux hypothèses clairement définies, à savoir l'hypothèse principale H_0 et la contre-hypothèse H_1. Le niveau de signification et les risques d'erreur de première et de deuxième espèce sont deux concepts différents. Dans le schéma Fishérien, le niveau exact de signification est une propriété des données tandis que pour Neyman et Pearson, les risques d'erreur de première et de deuxième espèce sont des propriétés du test. Les différentes étapes de mise en œuvre de ces deux approches sont représentées dans le tableau C1.2.

Fisher : **approche "probabilité variable"** (*Test de signification*)	Neyman-Pearson : **approche "erreur α fixée"** (*Test d'hypothèses*)
Étape 1 : Formuler l'hypothèse principale H_0	**Étape 1** : Formuler l'hypothèse principale H_0 et la contre-hypothèse H_1
Étape 2 : Spécifier la statistique du test à utiliser (*T*) et sa distribution	**Étape 2** : Spécifier la statistique du test à utiliser (*T*) et sa distribution
Étape 3 : Collecter les données et calculer la valeur observée de *T*	**Étape 3** : Spécifier l'erreur de première espèce α et déterminer la région de rejet du test
Étape 4 : Calculer la p-valeur (probabilité de dépassement)	**Étape 4** :Collecter les données et calculer la valeur observée de *T*
Étape 5 : Rejeter l'hypothèse H_0 si la p-valeur est petite	**Étape 5** : Rejeter H_0 en faveur de H_1 si la valeur observée de *T* se trouve dans la région de rejet sinon tolérer H_0

Tableau C1.2 : Étapes de mise en œuvre d'un test statistique

En somme, la logique d'un test statistique prend donc des éléments propres à chacune des deux approches présentées précédemment. Le fait de ne considérer explicitement qu'une seule hypothèse a du mal à tenir dans les tests de détection de non-stationnarités en hydrométéorologie. En effet, comme on l'a déjà mentionnée au chapitre 5, on voudrait tester l'hypothèse de stationnarité (H_0) contre l'une ou l'autre des infractions à la stationnarité (H_1) de telle sorte que le rejet de H_0 donne en même temps une indication sur le type de non-stationnarité suspecte dans les données. De plus, on voudrait aussi évaluer, en terme de puissance (et donc erreur de deuxième espèce), le comportement des tests de stationnarité largement employés en hydrométéorologie. C'est pour cette raison que la théorie des tests de stationnarité en hydrométéorologie se fait quasi exclusivement à l'aide de la technique de Neyman et Pearson (tests d'hypothèses).

Annexe A

Résultats de simulations pour une rupture : échantillon sans persistance

Tableau A1 : MC_1rup, une rupture : niveau (échantillon non autocorrélée)

Données générées par $x_t = \mu_0 + \varepsilon_t$

$\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $N = 99$, $M = 1000$ répétitions

n	Test	Niveaux critiques α		
		1%	**5%**	**10%**
30	MC_1rup	0.011	0.030	0.069
	Buishand	0.010	0.050	0.103
	Worsley	0.012	0.062	0.105
	Pettitt	0.002	0.029	0.049
50	MC_1rup	0.004	0.026	0.031
	Buishand	0.002	0.049	0.090
	Worsley	0.012	0.071	0.090
	Pettitt	0.003	0.036	0.057
100	MC_1rup	0.001	0.013	0.027
	Buishand	0.009	0.046	0.102
	Worsley	0.010	0.069	0.127
	Pettitt	0.005	0.034	0.079

Tableau A2 : MC_1rup, une rupture brusque en absence de persistance : puissance (échantillon indépendant , $\alpha = 0.01$, $t_b = [n/3]$)

Données générées par $x_t = \begin{cases} \mu_0 + \varepsilon_t & t = 1,\ldots,t_b \\ (\mu_0 + \Delta) + \varepsilon_t & t = t_b + 1,\ldots,n \end{cases}$

$\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $t_b = [n/3]$, $N = 99$, $M = 1000$ répétitions

Niveau critique $\alpha = 1\%$ $(t_b = [n/3])$						
n		$\Delta = 0.5\%$	$\Delta = 1\%$	$\Delta = 1.5\%$	$\Delta = 2\%$	$\Delta = 2.5\%$
30	MC_1rup	0.10	0.72	0.99	1.00	1.00
	Buishand	0.23	0.90	1.00	1.00	1.00
	Worsley	0.18	0.84	1.00	1.00	1.00
	Pettitt	0.11	0.69	0.99	1.00	1.00
50	MC_1rup	0.26	0.97	1.00	1.00	1.00
	Buishand	0.45	1.00	1.00	1.00	1.00
	Worsley	0.36	1.00	1.00	1.00	1.00
	Pettitt	0.31	0.97	1.00	1.00	1.00
100	MC_1rup	0.64	1.00	1.00	1.00	1.00
	Buishand	0.87	1.00	1.00	1.00	1.00
	Worsley	0.87	1.00	1.00	1.00	1.00
	Pettitt	0.80	1.00	1.00	1.00	1.00

Tableau A3 : MC_1rup, une rupture brusque en absence de persistance : puissance (échantillon indépendant , $\alpha = 0.01$, $t_b = [n/2]$)

Données générées par $x_t = \begin{cases} \mu_0 + \varepsilon_t & t = 1,\ldots,t_b \\ (\mu_0 + \Delta) + \varepsilon_t & t = t_b + 1,\ldots,n \end{cases}$

$\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $t_b = [n/2]$, $N = 99$, $M = 1000$ répétitions

Niveau critique $\alpha = 1\%$ ($t_b = [n/2]$)

n		$\Delta = 0.5\%$	$\Delta = 1\%$	$\Delta = 1.5\%$	$\Delta = 2\%$	$\Delta = 2.5\%$
30	MC_1rup	0.17	0.80	1.00	1.00	1.00
	Buishand	0.35	0.96	1.00	1.00	1.00
	Worsley	0.24	0.90	1.00	1.00	1.00
	Pettitt	0.20	0.89	1.00	1.00	1.00
50	MC_1rup	0.32	0.98	1.00	1.00	1.00
	Buishand	0.60	1.00	1.00	1.00	1.00
	Worsley	0.46	1.00	1.00	1.00	1.00
	Pettitt	0.47	1.00	1.00	1.00	1.00
100	MC_1rup	0.73	1.00	1.00	1.00	1.00
	Buishand	0.93	1.00	1.00	1.00	1.00
	Worsley	0.90	1.00	1.00	1.00	1.00
	Pettitt	0.89	1.00	1.00	1.00	1.00

Tableau A4 : MC_1rup, une rupture brusque en absence de persistance : puissance (échantillon indépendant , $\alpha = 0.01$, $t_b = [2n/3]$)

Données générées par $x_t = \begin{cases} \mu_0 + \varepsilon_t & t = 1,\ldots,t_b \\ (\mu_0 + \Delta) + \varepsilon_t & t = t_b + 1,\ldots,n \end{cases}$

$\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $t_b = [2n/3]$, $N = 99$, $M = 1000$ répétitions

Niveau critique $\alpha = 1\%$ ($t_b = [2n/3]$)

n		$\Delta = 0.5\%$	$\Delta = 1\%$	$\Delta = 1.5\%$	$\Delta = 2\%$	$\Delta = 2.5\%$
30	MC_1rup	0.15	0.73	1.00	1.00	1.00
	Buishand	0.24	0.90	1.00	1.00	1.00
	Worsley	0.21	0.85	1.00	1.00	1.00
	Pettitt	0.12	0.73	0.99	1.00	1.00
50	MC_1rup	0.27	0.96	1.00	1.00	1.00
	Buishand	0.47	0.99	1.00	1.00	1.00
	Worsley	0.39	0.99	1.00	1.00	1.00
	Pettitt	0.34	0.98	1.00	1.00	1.00
100	MC_1rup	0.65	1.00	1.00	1.00	1.00
	Buishand	0.87	1.00	1.00	1.00	1.00
	Worsley	0.87	1.00	1.00	1.00	1.00
	Pettitt	0.80	1.00	1.00	1.00	1.00

Tableau A5 : MC_1rup, une rupture brusque en absence de persistance : puissance (échantillon indépendant , $\alpha = 0.05$, $t_b = [n/3]$)

Données générées par $x_t = \begin{cases} \mu_0 + \varepsilon_t & t = 1,\ldots,t_b \\ (\mu_0 + \Delta) + \varepsilon_t & t = t_b + 1,\ldots,n \end{cases}$

$\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $t_b = [n/3]$, $N = 99$, $M = 1000$ répétitions

Niveau critique α = 5% ($t_b = [n/3]$)						
n		$\Delta = 0.5\%$	$\Delta = 1\%$	$\Delta = 1.5\%$	$\Delta = 2\%$	$\Delta = 2.5\%$
30	MC_1rup	0.31	0.91	1.00	1.00	1.00
	Buishand	0.49	0.97	1.00	1.00	1.00
	Worsley	0.42	0.95	1.00	1.00	1.00
	Pettitt	0.34	0.91	1.00	1.00	1.00
50	MC_1rup	0.50	1.00	1.00	1.00	1.00
	Buishand	0.70	1.00	1.00	1.00	1.00
	Worsley	0.64	1.00	1.00	1.00	1.00
	Pettitt	0.62	1.00	1.00	1.00	1.00
100	MC_1rup	0.84	1.00	1.00	1.00	1.00
	Buishand	0.97	1.00	1.00	1.00	1.00
	Worsley	0.95	1.00	1.00	1.00	1.00
	Pettitt	0.95	1.00	1.00	1.00	1.00

Tableau A6 : MC_1rup, une rupture brusque en absence de persistance : puissance (échantillon indépendant , $\alpha = 0.05$ $t_b = [n/2]$)

Données générées par $x_t = \begin{cases} \mu_0 + \varepsilon_t & t = 1,\ldots,t_b \\ (\mu_0 + \Delta) + \varepsilon_t & t = t_b + 1,\ldots,n \end{cases}$

$\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $t_b = [n/2]$, $N = 99$, $M = 1000$ répétitions

Niveau critique α = 5% ($t_b = [n/2]$)						
n		$\Delta = 0.5\%$	$\Delta = 1\%$	$\Delta = 1.5\%$	$\Delta = 2\%$	$\Delta = 2.5\%$
30	MC_1rup	0.35	0.96	1.00	1.00	1.00
	Buishand	0.57	0.99	1.00	1.00	1.00
	Worsley	0.45	0.98	1.00	1.00	1.00
	Pettitt	0.42	0.98	1.00	1.00	1.00
50	MC_1rup	0.57	1.00	1.00	1.00	1.00
	Buishand	0.81	1.00	1.00	1.00	1.00
	Worsley	0.70	1.00	1.00	1.00	1.00
	Pettitt	0.73	1.00	1.00	1.00	1.00
100	MC_1rup	0.90	1.00	1.00	1.00	1.00
	Buishand	0.99	1.00	1.00	1.00	1.00
	Worsley	0.98	1.00	1.00	1.00	1.00
	Pettitt	0.99	1.00	1.00	1.00	1.00

Tableau A7 : MC_1rup, une rupture brusque en absence de persistance : puissance (échantillon indépendant , $\alpha = 0.05$, $t_b = [2n/3]$)

Données générées par $x_t = \begin{cases} \mu_0 + \varepsilon_t & t = 1,\ldots,t_b \\ (\mu_0 + \Delta) + \varepsilon_t & t = t_b + 1,\ldots,n \end{cases}$

$\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $t_b = [2n/3]$, $N = 99$, $M = 1000$ répétitions

Niveau critique α = 5% ($t_b = [2n/3]$)						
n		$\Delta = 0.5\%$	$\Delta = 1\%$	$\Delta = 1.5\%$	$\Delta = 2\%$	$\Delta = 2.5\%$
30	MC_1rup	0.31	0.93	1.00	1.00	1.00
	Buishand	0.46	0.98	1.00	1.00	1.00
	Worsley	0.40	0.96	1.00	1.00	1.00
	Pettitt	0.33	0.94	1.00	1.00	1.00
50	MC_1rup	0.47	1.00	1.00	1.00	1.00
	Buishand	0.71	1.00	1.00	1.00	1.00
	Worsley	0.63	1.00	1.00	1.00	1.00
	Pettitt	0.59	1.00	1.00	1.00	1.00
100	MC_1rup	0.83	1.00	1.00	1.00	1.00
	Buishand	0.97	1.00	1.00	1.00	1.00
	Worsley	0.95	1.00	1.00	1.00	1.00
	Pettitt	0.94	1.00	1.00	1.00	1.00

Tableau A8 : MC_1rup, une rupture brusque en absence de persistance : puissance (échantillon indépendant , $\alpha = 0.10$, $t_b = [n/3]$)

Données générées par $x_t = \begin{cases} \mu_0 + \varepsilon_t & t = 1,\ldots,t_b \\ (\mu_0 + \Delta) + \varepsilon_t & t = t_b + 1,\ldots,n \end{cases}$

$\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $t_b = [n/3]$, $N = 99$, $M = 1000$ répétitions

Niveau critique α = 10% ($t_b = [n/3]$)						
n		$\Delta = 0.5\%$	$\Delta = 1\%$	$\Delta = 1.5\%$	$\Delta = 2\%$	$\Delta = 2.5\%$
30	MC_1rup	0.41	0.96	1.00	1.00	1.00
	Buishand	0.61	0.99	1.00	1.00	1.00
	Worsley	0.54	0.99	1.00	1.00	1.00
	Pettitt	0.47	0.97	1.00	1.00	1.00
50	MC_1rup	0.61	1.00	1.00	1.00	1.00
	Buishand	0.82	1.00	1.00	1.00	1.00
	Worsley	0.74	1.00	1.00	1.00	1.00
	Pettitt	0.74	1.00	1.00	1.00	1.00
100	MC_1rup	0.88	1.00	1.00	1.00	1.00
	Buishand	0.98	1.00	1.00	1.00	1.00
	Worsley	0.97	1.00	1.00	1.00	1.00
	Pettitt	0.96	1.00	1.00	1.00	1.00

Tableau A9 : MC_1rup, une rupture brusque en absence de persistance : puissance (échantillon indépendant , $\alpha = 0.10$, $t_b = [n/2]$)

Données générées par $x_t = \begin{cases} \mu_0 + \varepsilon_t & t = 1,\ldots,t_b \\ (\mu_0 + \Delta) + \varepsilon_t & t = t_b + 1,\ldots,n \end{cases}$

$\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $t_b = [n/2]$, $N = 99$, $M = 1000$ répétitions

Niveau critique α = 10% ($t_b = [n/2]$)						
n		$\Delta = 0.5\%$	$\Delta = 1\%$	$\Delta = 1.5\%$	$\Delta = 2\%$	$\Delta = 2.5\%$
30	MC_1rup	0.43	0.98	1.00	1.00	1.00
	Buishand	0.69	1.00	1.00	1.00	1.00
	Worsley	0.56	0.99	1.00	1.00	1.00
	Pettitt	0.56	0.99	1.00	1.00	1.00
50	MC_1rup	0.68	1.00	1.00	1.00	1.00
	Buishand	0.90	1.00	1.00	1.00	1.00
	Worsley	0.82	1.00	1.00	1.00	1.00
	Pettitt	0.84	1.00	1.00	1.00	1.00
100	MC_1rup	0.93	1.00	1.00	1.00	1.00
	Buishand	1.00	1.00	1.00	1.00	1.00
	Worsley	0.99	1.00	1.00	1.00	1.00
	Pettitt	0.99	1.00	1.00	1.00	1.00

Tableau A10 : MC_1rup, une rupture brusque en absence de persistance : puissance (échantillon indépendant , $\alpha = 0.10$, $t_b = [2n/3]$)

Données générées par $x_t = \begin{cases} \mu_0 + \varepsilon_t & t = 1,\ldots,t_b \\ (\mu_0 + \Delta) + \varepsilon_t & t = t_b + 1,\ldots,n \end{cases}$

$\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $t_b = [2n/3]$, $N = 99$, $M = 1000$ répétitions

Niveau critique α = 10% ($t_b = [2n/3]$)						
n		$\Delta = 0.5\%$	$\Delta = 1\%$	$\Delta = 1.5\%$	$\Delta = 2\%$	$\Delta = 2.5\%$
30	MC_1rup	0.39	0.96	1.00	1.00	1.00
	Buishand	0.61	0.99	1.00	1.00	1.00
	Worsley	0.51	0.98	1.00	1.00	1.00
	Pettitt	0.43	0.97	1.00	1.00	1.00
50	MC_1rup	0.59	1.00	1.00	1.00	1.00
	Buishand	0.82	1.00	1.00	1.00	1.00
	Worsley	0.72	1.00	1.00	1.00	1.00
	Pettitt	0.73	1.00	1.00	1.00	1.00
100	MC_1rup	1.00	1.00	1.00	1.00	1.00
	Buishand	1.00	1.00	1.00	1.00	1.00
	Worsley	1.00	1.00	1.00	1.00	1.00
	Pettitt	1.00	1.00	1.00	1.00	1.00

Tableau A11 : MC_1rup, une rupture brusque en absence de persistance : pourcentage d'estimation de la date $t_b = [n/3]$ (échantillon indépendant, $\alpha = 0.01$)

Données générées par $x_t = \begin{cases} \mu_0 + \varepsilon_t & t = 1,\ldots,t_b \\ (\mu_0 + \Delta) + \varepsilon_t & t = t_b + 1,\ldots,n \end{cases}$

$\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $t_b = [n/3]$, $N = 99$, $M = 1000$ répétitions

Niveau critique α = 1% $(t_b = [n/3])$

n		$\Delta = 0.5\%$	$\Delta = 1\%$	$\Delta = 1.5\%$	$\Delta = 2\%$	$\Delta = 2.5\%$
30	MC_1rup	31.0	43.3	81.0	94.4	98.2
	Buishand	51.2	56.0	76.2	87.3	93.5
	Worsley	47.0	49.8	81.7	94.4	98.2
	Pettitt	30.0	36.6	66.6	70.6	71.3
50	MC_1rup	6.50	56.9	81.3	92.3	97.7
	Buishand	11.3	56.6	76.3	87.5	93.5
	Worsley	8.20	58.8	81.3	92.3	97.7
	Pettitt	7.40	53.2	68.9	71.8	72.4
100	MC_1rup	16.7	58.0	79.8	93.5	97.0
	Buishand	20.6	54.1	73.3	85.5	92.7
	Worsley	21.9	58.0	79.8	93.5	97.0
	Pettitt	18.7	52.4	65.5	66.3	69.2

Tableau A12 : MC_1rup, une rupture brusque en absence de persistance : pourcentage d'estimation de la date $t_b = [n/2]$ (échantillon indépendant, $\alpha = 0.01$)

Données générées par $x_t = \begin{cases} \mu_0 + \varepsilon_t & t = 1,\ldots,t_b \\ (\mu_0 + \Delta) + \varepsilon_t & t = t_b + 1,\ldots,n \end{cases}$

$\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $t_b = [n/2]$, $N = 99$, $M = 1000$ répétitions

Niveau critique α = 1% $(t_b = [n/2])$

n		$\Delta = 0.5\%$	$\Delta = 1\%$	$\Delta = 1.5\%$	$\Delta = 2\%$	$\Delta = 2.5\%$
30	MC_1rup	5.50	48.2	79.0	92.9	97.1
	Buishand	13.6	61.1	82.8	93.7	97.8
	Worsley	7.20	54.0	79.5	92.9	97.1
	Pettitt	8.40	56.5	83.5	94.2	98.5
50	MC_1rup	9.40	53.9	78.9	91.3	97.5
	Buishand	17.3	58.1	79.8	92.4	97.7
	Worsley	11.3	54.5	78.9	91.3	97.5
	Pettitt	13.1	57.2	81.0	93.1	98.6
100	MC_1rup	17.3	59.0	80.2	89.8	98.2
	Buishand	23.8	60.2	80.8	90.8	98.4
	Worsley	20.5	59.0	80.2	89.8	98.2
	Pettitt	22.7	59.9	81.1	91.2	98.9

Tableau A13 : MC_1rup, une rupture brusque en absence de persistance : pourcentage d'estimation de la date $t_b = [2n/3]$ (échantillon indépendant, $\alpha = 0.01$)

Données générées par $x_t = \begin{cases} \mu_0 + \varepsilon_t & t = 1,\ldots,t_b \\ (\mu_0 + \Delta) + \varepsilon_t & t = t_b + 1,\ldots,n \end{cases}$

$\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $t_b = [2n/3]$, $N = 99$, $M = 1000$ répétitions

Niveau critique α = 1% ($t_b = [2n/3]$)

n		$\Delta = 0.5\%$	$\Delta = 1\%$	$\Delta = 1.5\%$	$\Delta = 2\%$	$\Delta = 2.5\%$
30	MC_1rup	3.40	42.3	79.7	92.2	98.5
	Buishand	6.30	49.6	78.0	86.7	93.4
	Worsley	5.00	49.2	80.1	92.2	98.5
	Pettitt	2.70	36.8	66.8	69.0	71.5
50	MC_1rup	7.10	58.4	81.0	91.9	97.6
	Buishand	12.7	53.5	76.5	87.9	93.5
	Worsley	10.3	59.5	81.0	91.9	97.6
	Pettitt	8.40	51.7	66.4	70.5	71.7
100	MC_1rup	17.3	58.2	81.0	93.0	97.3
	Buishand	21.8	51.5	75.0	87.0	92.0
	Worsley	22.1	58.2	81.0	93.0	97.3
	Pettitt	19.5	48.7	66.0	68.8	71.0

Tableau A14 : MC_1rup, une rupture brusque en absence de persistance : pourcentage d'estimation de la date $t_b = [n/3]$ (échantillon indépendant, $\alpha = 0.05$)

Données générées par $x_t = \begin{cases} \mu_0 + \varepsilon_t & t = 1,\ldots,t_b \\ (\mu_0 + \Delta) + \varepsilon_t & t = t_b + 1,\ldots,n \end{cases}$

$\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $t_b = [n/3]$, $N = 99$, $M = 1000$ répétitions

Niveau critique α = 5% ($t_b = [n/3]$)

n		$\Delta = 0.5\%$	$\Delta = 1\%$	$\Delta = 1.5\%$	$\Delta = 2\%$	$\Delta = 2.5\%$
30	MC_1rup	5.00	49.2	80.9	92.2	96.6
	Buishand	13.3	51.5	76.3	88.0	92.4
	Worsley	10.3	51.0	80.9	92.2	96.6
	Pettitt	8.50	46.9	65.4	69.5	67.4
50	MC_1rup	14.2	55.3	84.0	93.7	96.9
	Buishand	18.9	52.1	78.3	86.3	94.0
	Worsley	17.0	55.4	84.0	93.7	96.9
	Pettitt	16.4	49.3	69.0	69.1	72.0
100	MC_1rup	20.0	59.5	80.0	93.2	96.8
	Buishand	22.7	56.5	74.5	85.9	92.5
	Worsley	22.2	59.5	80.0	93.2	96.8
	Pettitt	21.6	54.0	65.7	69.0	69.3

Tableau A15 : MC_1rup, une rupture brusque en absence de persistance : pourcentage d'estimation de la date $t_b = [n/2]$ (échantillon indépendant, $\alpha = 0.05$)

Données générées par $x_t = \begin{cases} \mu_0 + \varepsilon_t & t = 1,\ldots,t_b \\ (\mu_0 + \Delta) + \varepsilon_t & t = t_b + 1,\ldots,n \end{cases}$

$\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $t_b = [n/2]$, $N = 99$, $M = 1000$ répétitions

Niveau critique α = 5% ($t_b = [n/2]$)						
n		$\Delta = 0.5\%$	$\Delta = 1\%$	$\Delta = 1.5\%$	$\Delta = 2\%$	$\Delta = 2.5\%$
30	MC_1rup	9.70	54.9	81.8	92.6	97.2
	Buishand	18.7	60.5	83.4	94.3	97.7
	Worsley	12.6	55.3	81.8	92.6	97.2
	Pettitt	15.1	59.9	83.8	94.9	98.3
50	MC_1rup	14.8	55.4	81.7	93.8	97.3
	Buishand	24.6	58.2	82.6	94.2	97.6
	Worsley	18.0	55.4	81.7	93.8	97.3
	Pettitt	23.2	58.9	82.0	95.2	98.9
100	MC_1rup	18.4	56.8	81.0	92.8	98.0
	Buishand	21.0	58.8	81.8	93.3	98.0
	Worsley	20.0	56.8	81.0	92.8	98.0
	Pettitt	23.3	58.9	82.5	94.0	98.0

Tableau A16 : MC_1rup, une rupture brusque en absence de persistance : pourcentage d'estimation de la date $t_b = [2n/3]$ (échantillon indépendant, $\alpha = 0.05$)

Données générées par $x_t = \begin{cases} \mu_0 + \varepsilon_t & t = 1,\ldots,t_b \\ (\mu_0 + \Delta) + \varepsilon_t & t = t_b + 1,\ldots,n \end{cases}$

$\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $t_b = [2n/3]$, $N = 99$, $M = 1000$ répétitions

Niveau critique α = 5% ($t_b = [2n/3]$)						
n		$\Delta = 0.5\%$	$\Delta = 1\%$	$\Delta = 1.5\%$	$\Delta = 2\%$	$\Delta = 2.5\%$
30	MC_1rup	8.20	53.7	80.0	91.9	96.4
	Buishand	12.7	52.7	73.6	87.1	92.9
	Worsley	9.90	54.7	80.1	91.9	96.4
	Pettitt	8.90	47.0	63.6	71.0	68.7
50	MC_1rup	11.2	56.5	80.1	93.2	97.9
	Buishand	18.5	53.3	76.2	86.8	94.3
	Worsley	15.5	56.5	80.1	93.2	97.9
	Pettitt	16.0	50.3	66.0	70.1	70.5
100	MC_1rup	19.1	56.9	81.0	93.0	97.9
	Buishand	19.3	54.3	74.0	87.0	92.1
	Worsley	21.2	56.9	81.0	93.0	97.9
	Pettitt	19.0	50.6	66.0	68.7	70.0

Tableau A17 : MC_1rup, une rupture brusque en absence de persistance : pourcentage d'estimation de la date $t_b = [n/3]$ (échantillon indépendant, $\alpha = 0.10$)

Données générées par $x_t = \begin{cases} \mu_0 + \varepsilon_t & t = 1,\ldots,t_b \\ (\mu_0 + \Delta) + \varepsilon_t & t = t_b + 1,\ldots,n \end{cases}$

$\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $t_b = [n/3]$, $N = 99$, $M = 1000$ répétitions

Niveau critique α = 10% $(t_b = [n/3])$

n		$\Delta = 0.5\%$	$\Delta = 1\%$	$\Delta = 1.5\%$	$\Delta = 2\%$	$\Delta = 2.5\%$
30	MC_1rup	11.0	52.2	80.6	92.3	96.8
	Buishand	16.3	55.2	76.7	88.8	91.6
	Worsley	12.8	53.4	80.6	92.3	96.8
	Pettitt	11.6	50.3	69.5	72.9	73.2
50	MC_1rup	16.2	58.1	81.5	91.6	97.9
	Buishand	21.0	55.3	76.4	85.0	92.8
	Worsley	18.9	58.2	81.5	91.6	97.9
	Pettitt	18.0	52.1	65.1	70.3	70.7
100	MC_1rup	22.6	56.4	83.3	91.3	97.0
	Buishand	22.6	50.6	74.1	86.5	91.1
	Worsley	24.2	56.4	83.3	91.3	97.0
	Pettitt	21.6	47.1	65.3	71.5	69.7

Tableau A18 : MC_1rup, une rupture brusque en absence de persistance : pourcentage d'estimation de la date $t_b = [n/2]$ (échantillon indépendant, $\alpha = 0.10$)

Données générées par $x_t = \begin{cases} \mu_0 + \varepsilon_t & t = 1,\ldots,t_b \\ (\mu_0 + \Delta) + \varepsilon_t & t = t_b + 1,\ldots,n \end{cases}$

$\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $t_b = [n/2]$, $N = 99$, $M = 1000$ répétitions

Niveau critique α = 10% $(t_b = [n/2])$

n		$\Delta = 0.5\%$	$\Delta = 1\%$	$\Delta = 1.5\%$	$\Delta = 2\%$	$\Delta = 2.5\%$
30	MC_1rup	10.5	54.5	80.4	92.0	98.1
	Buishand	20.9	59.1	83.7	93.6	98.7
	Worsley	13.1	54.8	80.4	92.0	98.1
	Pettitt	16.5	57.9	83.4	95.3	99.2
50	MC_1rup	16.5	59.2	79.7	92.1	97.8
	Buishand	25.5	62.0	80.6	93.2	98.3
	Worsley	19.6	59.2	79.7	92.1	97.8
	Pettitt	23.8	62.0	80.1	93.4	98.5
100	MC_1rup	23.6	61.4	81.6	94.0	97.6
	Buishand	27.7	62.8	82.2	94.2	97.8
	Worsley	24.3	61.4	81.6	94.0	97.6
	Pettitt	26.5	61.9	82.5	93.6	98.6

Tableau A19 : MC_1rup, une rupture brusque en absence de persistance : pourcentage d'estimation de la date $t_b = [2n/3]$ (échantillon indépendant, $\alpha = 0.10$)

Données générées par $x_t = \begin{cases} \mu_0 + \varepsilon_t & t = 1,\ldots,t_b \\ (\mu_0 + \Delta) + \varepsilon_t & t = t_b + 1,\ldots,n \end{cases}$

$\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $t_b = [2n/3]$, $N = 99$, $M = 1000$ répétitions

Niveau critique α = 10% ($t_b = [2n/3]$)

n		$\Delta = 0.5\%$	$\Delta = 1\%$	$\Delta = 1.5\%$	$\Delta = 2\%$	$\Delta = 2.5\%$
30	MC_1rup	10.3	53.3	79.0	91.6	97.8
	Buishand	15.4	53.7	76.7	87.8	93.8
	Worsley	12.8	54.2	79.0	91.6	97.8
	Pettitt	11.1	50.5	64.6	68.4	69.8
50	MC_1rup	13.6	57.9	81.6	92.1	97.8
	Buishand	17.4	55.7	79.4	89.2	93.3
	Worsley	15.8	57.9	81.8	92.1	97.8
	Pettitt	16.5	52.6	68.3	71.8	73.5
100	MC_1rup	18.0	57.0	82.0	94.0	96.9
	Buishand	21.0	53.0	74.0	88.0	92.1
	Worsley	20.0	57.0	82.0	94.0	96.9
	Pettitt	22.0	51.0	65.0	70.0	74.0

Annexe B

Résultats de simulations pour une rupture : échantillon sans persistance

(Rupture avec continuité)

Tableau B1 : MC_1rup, rupture avec continuité en absence de persistance : puissance (échantillon indépendant, $\alpha = 0.05$, $t_b = [n/3]$)

Données générées par $x_t = \mu_0 + \delta\, ID(t_b) + \varepsilon_t$, où $ID(t_b) = \begin{cases} 0 & t = 1,\ldots,t_b \\ (t - t_b) & t = t_b + 1,\ldots,n \end{cases}$

$\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $t_b = [n/3]$, $N = 99$, $M = 1000$ répétitions

Niveau critique α = 5% ($t_b = [n/3]$)

n		$\delta = 0.1$	$\delta = 0.2$	$\delta = 0.3$	$\delta = 0.4$
30	MC_1rup	0.681	1.000	1.000	1.000
	Buishand	0.891	1.000	1.000	1.000
	Worsley	0.844	1.000	1.000	1.000
	Pettitt	0.797	1.000	1.000	1.000
50	MC_1rup	0.681	1.000	1.000	1.000
	Buishand	0.891	1.000	1.000	1.000
	Worsley	0.844	1.000	1.000	1.000
	Pettitt	0.797	1.000	1.000	1.000
100	MC_1rup	1.000	1.000	1.000	1.000
	Buishand	1.000	1.000	1.000	1.000
	Worsley	1.000	1.000	1.000	1.000
	Pettitt	1.000	1.000	1.000	1.000

Tableau B2 : MC_1rup, rupture avec continuité en absence de persistance : puissance (échantillon indépendant, $\alpha = 0.05$, $t_b = [n/2]$)

Données générées par $x_t = \mu_0 + \delta\, ID(t_b) + \varepsilon_t$, où $ID(t_b) = \begin{cases} 0 & t = 1,\ldots,t_b \\ (t - t_b) & t = t_b + 1,\ldots,n \end{cases}$

$\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $t_b = [n/2]$, $N = 99$, $M = 1000$ répétitions

Niveau critique α = 5% ($t_b = [n/2]$)

n		$\delta = 0.1$	$\delta = 0.2$	$\delta = 0.3$	$\delta = 0.4$
30	MC_1rup	0.368	0.983	1.000	1.000
	Buishand	0.596	0.995	1.000	1.000
	Worsley	0.568	0.995	1.000	1.000
	Pettitt	0.439	0.967	1.000	1.000
50	MC_1rup	0.994	1.000	1.000	1.000
	Buishand	0.997	1.000	1.000	1.000
	Worsley	0.996	1.000	1.000	1.000
	Pettitt	0.995	1.000	1.000	1.000
100	MC_1rup	1.000	1.000	1.000	1.000
	Buishand	1.000	1.000	1.000	1.000
	Worsley	1.000	1.000	1.000	1.000
	Pettitt	1.000	1.000	1.000	1.000

Tableau B3 : MC_1rup, rupture avec continuité en absence de persistance : puissance (échantillon indépendant, $\alpha = 0.05$, $t_b = [2n/3]$)

Données générées par $x_t = \mu_0 + \delta\, ID(t_b) + \varepsilon_t$, où $ID(t_b) = \begin{cases} 0 & t = 1,\ldots,t_b \\ (t - t_b) & t = t_b + 1,\ldots,n \end{cases}$

$\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $t_b = [2n/3]$, $N = 99$, $M = 1000$ répétitions

Niveau critique 5% $(t_b = [2n/3])$

n		$\delta = 0.1$	$\delta = 0.2$	$\delta = 0.3$	$\delta = 0.4$
30	MC_1rup	0.105	0.612	0.964	0.999
	Buishand	0.202	0.703	0.973	0.999
	Worsley	0.234	0.752	0.982	1.000
	Pettitt	0.108	0.494	0.849	0.984
50	MC_1rup	0.718	1.000	1.000	1.000
	Buishand	0.797	1.000	1.000	1.000
	Worsley	0.827	1.000	1.000	1.000
	Pettitt	0.675	0.999	1.000	1.000
100	MC_1rup	1.000	1.000	1.000	1.000
	Buishand	1.000	1.000	1.000	1.000
	Worsley	1.000	1.000	1.000	1.000
	Pettitt	1.000	1.000	1.000	1.000

Tableau B4 : MC_1rup, rupture avec continuité en absence de persistance : pourcentage d'estimation de la date $t_b = [n/3]$ (échantillon indépendant, $\alpha = 0.05$)

Données générées par $x_t = \mu_0 + \delta\, ID(t_b) + \varepsilon_t$, où $ID(t_b) = \begin{cases} 0 & t = 1,\ldots,t_b \\ (t - t_b) & t = t_b + 1,\ldots,n \end{cases}$

$\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $t_b = [n/3]$, $N = 99$, $M = 1000$ répétitions

Niveau critique α = **5%** ($t_b = [n/3]$)					
n		$\delta = 0.1$	$\delta = 0.2$	$\delta = 0.3$	$\delta = 0.4$
30	MC_1rup	5.90	13.5	20.7	29.8
	Buishand	1.20	0.20	0.00	0.00
	Worsley	1.50	0.60	0.10	0.00
	Pettitt	1.30)	0.30	0.00	0.00
50	MC_1rup	9.30	19.4	27.6	37.3
	Buishand	0.10	0.00	0.00	0.00
	Worsley	0.60	0.00	0.00	0.00
	Pettitt	0.10	0.00	0.00	0.00
100	MC_1rup	11.7	25.3	35.5	49.6
	Buishand	0.00	0.00	0.00	0.00
	Worsley	0.00	0.00	0.00	0.00
	Pettitt	0.00	0.00	0.00	0.00

Tableau B5 : MC_1rup, rupture avec continuité en absence de persistance : pourcentage d'estimation de la date $t_b = [n/2]$ (échantillon indépendant, $\alpha = 0.05$)

Données générées par $x_t = \mu_0 + \delta\, ID(t_b) + \varepsilon_t$, où $ID(t_b) = \begin{cases} 0 & t = 1,\ldots,t_b \\ (t - t_b) & t = t_b + 1,\ldots,n \end{cases}$

$\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $t_b = [n/2]$, $N = 99$, $M = 1000$ répétitions

Niveau critique α = **5%** ($t_b = [n/2]$)					
n		$\delta = 0.1$	$\delta = 0.2$	$\delta = 0.3$	$\delta = 0.4$
30	MC_1rup	2.20	12.2	23.6	29.1
	Buishand	4.50	4.70	2.50	1.70
	Worsley	1.70	2.50	0.60	0.30
	Pettitt	3.50	6.90	6.90	8.40
50	MC_1rup	8.30	17.9	25.9	33.0
	Buishand	2.60	0.40	0.00	0.00
	Worsley	1.30	0.10	0.00	0.00
	Pettitt	3.10	2.20	3.00	4.00
100	MC_1rup	9.00	25.4	36.4	49.3
	Buishand	0.00	0.00	0.00	0.00
	Worsley	0.00	0.00	0.00	0.00
	Pettitt	0.20	0.00	0.50	1.20

Tableau B6 : MC_1rup, rupture avec continuité en absence de persistance : pourcentage d'estimation de la date $t_b = [2n/3]$ (échantillon indépendant, $\alpha = 0.05$)

Données générées par $x_t = \mu_0 + \delta\, ID(t_b) + \varepsilon_t$, où $ID(t_b) = \begin{cases} 0 & t = 1,\ldots,t_b \\ (t - t_b) & t = t_b + 1,\ldots,n \end{cases}$

$\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $t_b = [2n/3]$, $N = 99$, $M = 1000$ répétitions

Niveau critique α = 5% ($t_b = [2n/3]$)

n		$\delta = 0.1$	$\delta = 0.2$	$\delta = 0.3$	$\delta = 0.4$
30	MC_1rup	0.90	8.00	15.9	24.4
	Buishand	2.50	9.40	14.7	17.6
	Worsley	1.90	4.20	5.30	4.10
	Pettitt	1.50	7.90	14.6	21.4
50	MC_1rup	5.60	13.2	23.1	29.4
	Buishand	6.20	7.90	9.20	6.10
	Worsley	2.60	1.30	0.70	0.10
	Pettitt	5.00	11.0	16.0	20.0
100	MC_1rup	11.2	21.6	31.0	42.4
	Buishand	3.00	1.40	0.50	0.00
	Worsley	0.10	0.00	0.00	0.00
	Pettitt	5.30	10.6	13.4	18.4

Annexe C

Résultats de simulations pour une rupture : échantillon avec persistance

Tableau C1 : MC_1rup, une rupture en absence de persistance: niveau (échantillon autocorrélé, $\alpha = 0.01$)

Données générées par $x_t = \rho\, x_{t-1} + \varepsilon_t$

$\varepsilon \overset{iid}{\sim} N(0,1)$, $N = 99$, $M = 1000$ répétitions

Niveau critique α = 1%					
n	ρ	MC_1rup	Buishand	Worsley	Pettitt
30	0.1	0.000	0.016	0.015	0.003
	0.2	0.027	0.033	0.033	0.008
	0.3	0.038	0.086	0.084	0.041
	0.4	0.187	0.101	0.111	0.045
	0.5	0.069	0.204	0.201	0.104
	0.6	0.182	0.290	0.290	0.160
	0.7	0.289	0.401	0.413	0.275
	0.8	0.542	0.531	0.564	0.394
	0.9	0.687	0.645	0.684	0.521
50	0.1	0.014	0.020	0.018	0.005
	0.2	0.015	0.032	0.040	0.020
	0.3	0.040	0.081	0.082	0.047
	0.4	0.110	0.129	0.149	0.085
	0.5	0.298	0.241	0.277	0.173
	0.6	0.243	0.317	0.347	0.237
	0.7	0.395	0.473	0.509	0.381
	0.8	0.505	0.625	0.662	0.519
	0.9	0.640	0.793	0.824	0.714
100	0.1	0.010	0.031	0.053	0.019
	0.2	0.023	0.046	0.083	0.034
	0.3	0.009	0.072	0.131	0.061
	0.4	0.138	0.179	0.257	0.123
	0.5	0.173	0.238	0.344	0.191
	0.6	0.222	0.365	0.520	0.306
	0.7	0.360	0.543	0.682	0.474
	0.8	0.711	0.755	0.840	0.685
	0.9	0.879	0.897	0.954	0.867

Tableau C2 : MC_1rup, une rupture en absence de persistance: niveau (échantillon autocorrélé, $\alpha = 0.05$)

Données générées par $x_t = \rho\, x_{t-1} + \varepsilon_t$

$\varepsilon \overset{iid}{\sim} N(0,1)$, $N = 99$, $M = 1000$ répétitions

Niveau critique α = 5%					
n	ρ	MC_1rup	Buishand	Worsley	Pettitt
30	0.1	0.027	0.080	0.081	0.050
	0.2	0.085	0.128	0.145	0.071
	0.3	0.116	0.181	0.184	0.102
	0.4	0.232	0.259	0.258	0.172
	0.5	0.323	0.356	0.360	0.238
	0.6	0.356	0.441	0.466	0.323
	0.7	0.566	0.608	0.598	0.480
	0.8	0.636	0.677	0.691	0.579
	0.9	0.721	0.792	0.796	0.700
50	0.1	0.029	0.098	0.072	0.058
	0.2	0.091	0.140	0.135	0.097
	0.3	0.079	0.196	0.207	0.137
	0.4	0.189	0.288	0.311	0.215
	0.5	0.267	0.406	0.423	0.327
	0.6	0.388	0.520	0.541	0.420
	0.7	0.490	0.662	0.709	0.568
	0.8	0.749	0.810	0.829	0.745
	0.9	0.883	0.900	0.918	0.862
100	0.1	0.046	0.098	0.134	0.073
	0.2	0.061	0.142	0.225	0.107
	0.3	0.146	0.241	0.329	0.190
	0.4	0.119	0.311	0.423	0.256
	0.5	0.274	0.500	0.565	0.430
	0.6	0.430	0.590	0.707	0.527
	0.7	0.636	0.733	0.825	0.675
	0.8	0.784	0.872	0.925	0.811
	0.9	0.955	0.966	0.984	0.940

Tableau C3 : MC_1rup, une rupture en absence de persistance: niveau (échantillon autocorrélé, $\alpha = 0.10$)

Données générées par $x_t = \rho\, x_{t-1} + \varepsilon_t$

$\varepsilon \overset{iid}{\sim} N(0,1)$, $N = 99$, $M = 1000$ répétitions

Niveau critique α = 10%					
n	ρ	MC_1rup	Buishand	Worsley	Pettitt
30	0.1	0.093	0.154	0.145	0.086
	0.2	0.176	0.203	0.219	0.118
	0.3	0.252	0.305	0.304	0.197
	0.4	0.270	0.394	0.389	0.273
	0.5	0.354	0.472	0.495	0.343
	0.6	0.408	0.600	0.584	0.464
	0.7	0.592	0.677	0.691	0.568
	0.8	0.641	0.796	0.811	0.690
	0.9	0.830	0.861	0.893	0.770
50	0.1	0.099	0.174	0.160	0.123
	0.2	0.134	0.243	0.225	0.164
	0.3	0.163	0.298	0.313	0.236
	0.4	0.269	0.401	0.417	0.325
	0.5	0.388	0.525	0.545	0.440
	0.6	0.544	0.646	0.663	0.543
	0.7	0.684	0.770	0.779	0.684
	0.8	0.776	0.867	0.865	0.804
	0.9	0.884	0.937	0.944	0.904
100	0.1	0.058	0.166	0.207	0.130
	0.2	0.113	0.230	0.276	0.195
	0.3	0.177	0.310	0.410	0.265
	0.4	0.240	0.436	0.496	0.370
	0.5	0.441	0.551	0.650	0.497
	0.6	0.496	0.703	0.765	0.624
	0.7	0.708	0.830	0.874	0.780
	0.8	0.889	0.928	0.950	0.897
	0.9	0.973	0.985	0.993	0.977

Tableau C4 : MC_1rup, une rupture en présence de persistance: niveau (échantillon autocorrélé, $\alpha = 0.05$)

Données générées par $x_t = \rho\, x_{t-1} + \varepsilon_t$

$\varepsilon \overset{iid}{\sim} N(0,1)$, $N = 99$, $M = 1000$ répétitions

Niveau critique α = 5%		
n	ρ	MC_1rup
30	0.0	0.003
	0.1	0.006
	0.2	0.009
	0.3	0.009
	0.4	0.008
	0.5	0.013
	0.6	0.013
	0.7	0.026
	0.8	0.035
	0.9	0.038
50	0.0	0.002
	0.1	0.003
	0.2	0.005
	0.3	0.003
	0.4	0.005
	0.5	0.007
	0.6	0.011
	0.7	0.009
	0.8	0.029
	0.9	0.028
100	0.0	0.001
	0.1	0.001
	0.2	0.001
	0.3	0.001
	0.4	0.002
	0.5	0.002
	0.6	0.002
	0.7	0.003
	0.8	0.004
	0.9	0.004

Tableau C5 : MC_1rup, une rupture en présence de persistance: puissance (échantillon autocorrélé, $\alpha = 0.05$, $t_b = [n/3]$)

Données générées par $x_t = \begin{cases} \mu_0 + \rho x_{t-1} + \varepsilon_t & t = 1,\ldots,t_b \\ (\mu_0 + \Delta) + \rho x_{t-1} + \varepsilon_t & t = t_b + 1,\ldots,n \end{cases}$

$\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $t_b = [n/3]$, $N = 99$, $M = 1000$ répétitions

Niveau critique α = 5% ($t_b = [n/3]$)						
n	ρ	$\Delta = 0.5\%$	$\Delta = 1\%$	$\Delta = 1.5\%$	$\Delta = 2\%$	$\Delta = 2.5\%$
30	0.1	0.007	0.095	0.465	0.797	0.969
	0.2	0.010	0.016	0.501	0.839	0.972
	0.3	0.013	0.030	0.590	0.878	0.985
	0.4	0.020	0.064	0.671	0.939	0.995
	0.5	0.037	0.123	0.774	0.964	0.994
	0.6	0.098	0.167	0.869	0.979	0.997
	0.7	0.223	0.224	0.935	0.985	0.999
	0.8	0.413	0.184	0.950	0.993	0.993
	0.9	0.523	0.063	0.882	0.895	0.920
50	0.1	0.030	0.398	0.912	0.990	0.999
	0.2	0.030	0.467	0.898	0.986	0.999
	0.3	0.031	0.437	0.908	0.983	0.999
	0.4	0.041	0.520	0.923	0.994	1.000
	0.5	0.071	0.574	0.950	1.000	1.000
	0.6	0.136	0.740	0.988	1.000	1.000
	0.7	0.259	0.902	0.998	1.000	1.000
	0.8	0.616	0.982	1.000	1.000	1.000
	0.9	0.967	1.000	1.000	1.000	1.000
100	0.1	0.561	0.999	1.000	1.000	1.000
	0.2	0.531	0.999	1.000	1.000	1.000
	0.3	0.555	1.000	1.000	1.000	1.000
	0.4	0.560	1.000	1.000	1.000	1.000
	0.5	0.607	1.000	1.000	1.000	1.000
	0.6	0.655	1.000	1.000	1.000	1.000
	0.7	0.783	1.000	1.000	1.000	1.000
	0.8	0.954	1.000	1.000	1.000	1.000
	0.9	1.000	1.000	1.000	1.000	1.000

Tableau C6 : MC_1rup, une rupture en présence de persistance: puissance (échantillon autocorrélé, $\alpha = 0.05$, $t_b = [n/2]$)

Données générées par $x_t = \begin{cases} \mu_0 + \rho x_{t-1} + \varepsilon_t & t = 1,\ldots,t_b \\ (\mu_0 + \Delta) + \rho x_{t-1} + \varepsilon_t & t = t_b + 1,\ldots,n \end{cases}$

$\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $t_b = [n/2]$, $N = 99$, $M = 1000$ répétitions

Niveau critique α = 5% ($t_b = [n/2]$)						
n	ρ	$\Delta = 0.5\%$	$\Delta = 1\%$	$\Delta = 1.5\%$	$\Delta = 2\%$	$\Delta = 2.5\%$
30	0.1	0.009	0.088	0.417	0.785	0.964
	0.2	0.007	0.088	0.455	0.828	0.963
	0.3	0.006	0.147	0.570	0.865	0.978
	0.4	0.016	0.203	0.661	0.911	0.989
	0.5	0.034	0.333	0.763	0.964	0.992
	0.6	0.104	0.505	0.861	0.975	0.997
	0.7	0.220	0.717	0.912	0.977	0.995
	0.8	0.477	0.812	0.902	0.946	0.972
	0.9	0.575	0.596	0.541	0.453	0.423
50	0.1	0.020	0.370	0.862	0.983	0.999
	0.2	0.028	0.370	0.862	0.990	1.000
	0.3	0.021	0.411	0.884	0.990	1.000
	0.4	0.036	0.457	0.901	0.989	1.000
	0.5	0.070	0.571	0.961	0.998	1.000
	0.6	0.127	0.720	0.976	0.999	1.000
	0.7	0.307	0.898	0.997	1.000	1.000
	0.8	0.654	0.991	1.000	1.000	1.000
	0.9	0.980	1.000	1.000	1.000	1.000
100	0.1	0.443	0.999	1.000	1.000	1.000
	0.2	0.457	0.999	1.000	1.000	1.000
	0.3	0.448	0.994	1.000	1.000	1.000
	0.4	0.480	0.995	1.000	1.000	1.000
	0.5	0.572	0.996	1.000	1.000	1.000
	0.6	0.647	0.999	1.000	1.000	1.000
	0.7	0.791	0.999	1.000	1.000	1.000
	0.8	0.961	1.000	1.000	1.000	1.000
	0.9	1.000	1.000	1.000	1.000	1.000

Tableau C7 : MC_1rup, une rupture en présence de persistance: puissance (échantillon autocorrélé, $\alpha = 0.05$, $t_b = [2n/3]$)

Données générées par $x_t = \begin{cases} \mu_0 + \rho x_{t-1} + \varepsilon_t & t = 1,\ldots,t_b \\ (\mu_0 + \Delta) + \rho x_{t-1} + \varepsilon_t & t = t_b + 1,\ldots,n \end{cases}$

$\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $t_b = [2n/3]$, $N = 99$, $M = 1000$ répétitions

Niveau critique α = 5% $(t_b = [2n/3])$						
n	ρ	$\Delta = 0.5\%$	$\Delta = 1\%$	$\Delta = 1.5\%$	$\Delta = 2\%$	$\Delta = 2.5\%$
30	0.1	0.002	0.047	0.291	0.656	0.910
	0.2	0.005	0.063	0.334	0.709	0.947
	0.3	0.007	0.088	0.403	0.763	0.94
	0.4	0.008	0.131	0.528	0.843	0.963
	0.5	0.018	0.200	0.599	0.880	0.985
	0.6	0.052	0.327	0.742	0.947	0.992
	0.7	0.089	0.466	0.811	0.952	0.989
	0.8	0.179	0.583	0.849	0.958	0.984
	0.9	0.216	0.458	0.689	0.854	0.916
50	0.1	0.019	0.405	0.888	0.997	1.000
	0.2	0.023	0.409	0.887	0.991	1.000
	0.3	0.026	0.413	0.891	0.993	1.000
	0.4	0.051	0.494	0.920	0.994	1.000
	0.5	0.059	0.569	0.940	1.000	1.000
	0.6	0.113	0.740	0.989	1.000	1.000
	0.7	0.235	0.888	0.994	1.000	1.000
	0.8	0.571	0.981	1.000	1.000	1.000
	0.9	0.920	1.000	1.000	1.000	1.000
100	0.1	0.295	0.996	0.999	1.000	1.000
	0.2	0.310	0.995	1.000	1.000	1.000
	0.3	0.297	0.991	1.000	1.000	1.000
	0.4	0.318	0.981	1.000	1.000	1.000
	0.5	0.337	0.989	1.000	1.000	1.000
	0.6	0.408	0.987	1.000	1.000	1.000
	0.7	0.563	0.991	1.000	1.000	1.000
	0.8	0.855	0.999	1.000	1.000	1.000
	0.9	1.000	1.000	1.000	1.000	1.000

Tableau C8 : MC_1rup, une rupture en présence de persistance: pourcentage d'estimation de la date $t_b = \lfloor n/3 \rfloor$ (échantillon autocorrélé, $\alpha = 0.05$)

Données générées par $x_t = \begin{cases} \mu_0 + \rho x_{t-1} + \varepsilon_t & t = 1,\ldots,t_b \\ (\mu_0 + \Delta) + \rho x_{t-1} + \varepsilon_t & t = t_b+1,\ldots,n \end{cases}$

$\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $t_b = \lfloor n/3 \rfloor$, $N = 99$, $M = 1000$ répétitions

Niveau critique α = 5% ($t_b = \lfloor n/3 \rfloor$)						
n	ρ	$\Delta = 0.5\%$	$\Delta = 1\%$	$\Delta = 1.5\%$	$\Delta = 2\%$	$\Delta = 2.5\%$
30	0.1	0.00	0.50	3.60	4.80	2.80
	0.2	0.10	17.0	5.30	10.5	6.30
	0.3	0.30	22.3	12.2	17.3	16.9
	0.4	0.60	32.0	19.9	28.1	28.4
	0.5	0.50	49.3	31.0	46.2	51.5
	0.6	2.00	71.5	39.9	54.5	64.1
	0.7	4.40	81.3	38.0	44.8	48.2
	0.8	6.40	76.3	23.0	16.8	11.0
	0.9	6.30	46.5	1.80	0.20	0.00
50	0.1	0.40	5.20	9.40	8.20	3.10
	0.2	0.40	7.50	15.7	12.4	9.20
	0.3	0.40	10.3	23.5	21.1	19.8
	0.4	1.00	15.3	33.4	35.7	35.5
	0.5	1.10	18.8	41.1	49.3	55.3
	0.6	2.80	26.0	44.1	56.4	62.6
	0.7	3.60	25.4	33.5	38.7	37.6
	0.8	6.50	18.1	13.9	7.20	2.70
	0.9	5.10	2.00	0.20	0.00	0.00
100	0.1	7.30	15.6	10.7	6.70	3.90
	0.2	7.80	19.3	20.6	15.3	9.80
	0.3	7.60	23.2	29.1	26.9	21.4
	0.4	9.10	29.3	36.1	41.4	41.4
	0.5	9.70	32.1	41.7	54.4	61.3
	0.6	11.2	31.4	41.0	51.3	62.6
	0.7	9.40	19.9	25.1	26.3	28.9
	0.8	6.20	8.30	4.00	2.80	1.00
	0.9	1.00	0.40	0.00	0.00	0.00

Tableau C9 : MC_1rup, une rupture en présence de persistance: pourcentage d'estimation de la date $t_b = [n/2]$ (échantillon autocorrélé, $\alpha = 0.05$)

Données générées par $x_t = \begin{cases} \mu_0 + \rho\, x_{t-1} + \varepsilon_t & t = 1,\ldots,t_b \\ (\mu_0 + \Delta) + \rho\, x_{t-1} + \varepsilon_t & t = t_b + 1,\ldots,n \end{cases}$

$\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $t_b = [n/2]$, $N = 99$, $M = 1000$ répétitions

Niveau critique α = 5% $(t_b = [n/2])$						
n	ρ	$\Delta = 0.5\%$	$\Delta = 1\%$	$\Delta = 1.5\%$	$\Delta = 2\%$	$\Delta = 2.5\%$
30	0.1	0.20	0.80	2.20	2.90	3.90
	0.2	0.20	1.20	5.80	8.50	6.20
	0.3	0.10	2.60	10.1	16.3	16.4
	0.4	0.30	4.70	18.5	27.2	31.2
	0.5	0.60	11.0	30.9	45.5	52.9
	0.6	1.90	17.0	37.6	53.4	62.2
	0.7	4.80	24.5	33.7	42.9	43.9
	0.8	9.00	17.7	18.4	9.00	5.40
	0.9	6.30	1.60	0.10	0.00	0.00
50	0.1	0.20	4.40	4.60	8.00	11.1
	0.2	0.50	6.30	6.60	13.9	15.3
	0.3	0.40	8.90	20.0	23.1	19.5
	0.4	0.50	14.6	32.6	35.1	36.3
	0.5	1.30	17.6	41.3	50.7	57.3
	0.6	3.20	24.1	40.6	52.8	61.2
	0.7	6.30	26.3	36.1	37.3	37.8
	0.8	9.60	15.7	10.4	6.30	2.10
	0.9	6.60	0.80	0.00	0.00	0.00
100	0.1	4.70	15.6	12.6	8.30	4.30
	0.2	5.80	21.2	18.3	13.0	9.10
	0.3	8.00	23.9	27.4	24.3	20.9
	0.4	6.70	30.2	40.0	39.3	41.0
	0.5	8.60	33.6	43.4	55.2	60.9
	0.6	10.6	31.7	38.6	49.4	60.4
	0.7	11.7	22.3	25.5	29.1	28.1
	0.8	8.70	9.90	5.40	3.20	1.10
	0.9	2.50	0.30	0.00	0.00	0.00

Tableau C10 : MC_1rup, une rupture en présence de persistance: puissance de détection de la date $t_b = [2n/3]$ (échantillon autocorrélé, $\alpha = 0.05$)

Données générées par $x_t = \begin{cases} \mu_0 + \rho x_{t-1} + \varepsilon_t & t = 1,\ldots,t_b \\ (\mu_0 + \Delta) + \rho x_{t-1} + \varepsilon_t & t = t_b + 1,\ldots,n \end{cases}$

$\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $t_b = [2n/3]$, $N = 99$, $M = 1000$ répétitions

Niveau critique α = 5% ($t_b = [2n/3]$)						
n	ρ	$\Delta = 0.5\%$	$\Delta = 1\%$	$\Delta = 1.5\%$	$\Delta = 2\%$	$\Delta = 2.5\%$
	0.1	0.00	0.30	1.60	3.50	3.00
	0.2	0.10	1.00	2.90	6.00	7.50
	0.3	0.10	1.80	6.60	12.9	14.9
	0.4	0.30	3.40	16.4	24.2	30.4
30	0.5	0.30	6.30	22.5	39.2	48.1
	0.6	1.10	10.4	35.2	48.4	63.6
	0.7	1.90	15.6	37.2	50.2	59.3
	0.8	2.40	15.6	25.9	27.9	29.9
	0.9	4.00	8.80	10.7	6.60	2.50
	0.1	0.30	5.00	10.5	8.50	5.20
	0.2	0.40	7.30	13.9	13.0	9.40
	0.3	0.20	9.00	22.7	21.2	16.7
	0.4	0.40	15.5	32.8	35.9	36.1
50	0.5	1.70	19.3	39.8	51.8	56.9
	0.6	1.70	24.5	44.0	55.1	63.6
	0.7	3.70	24.3	34.5	39.7	45.0
	0.8	6.40	15.8	15.3	11.3	5.50
	0.9	2.80	2.10	1.10	0.20	0.00
	0.1	3.00	16.8	12.6	8.40	4.30
	0.2	3.70	21.7	18.2	15.6	10.9
	0.3	5.80	26.3	26.2	26.6	19.8
	0.4	5.70	26.5	39.0	40.9	43.1
100	0.5	5.60	34.5	44.3	54.5	60.3
	0.6	7.90	27.4	38.8	50.2	61.1
	0.7	5.80	20.0	25.9	26.5	29.6
	0.8	5.70	9.40	5.90	2.90	1.30
	0.9	0.50	0.70	0.00	0.00	0.00

Annexe D

Résultats de simulations pour une rupture : échantillon avec persistance

(Rupture avec continuité)

Tableau D1 : MC_1rup, rupture avec continuité en présence de persistance : puissance (échantillon autocorrélé, $\alpha = 0.05$, $t_b = [n/3]$)

Données générées par $x_t = \mu_0 + \rho\, x_{t-1} + \delta\, ID(t_b) + \varepsilon_t$

où, $ID(t_b) = \begin{cases} 0 & t = 1,\ldots,t_b \\ (t - t_b) & t = t_b + 1,\ldots,n \end{cases}$

$\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $t_b = [n/3]$, $N = 99$, $M = 1000$ répétitions

Niveau critique α = 5% $(t_b = [n/3])$					
n	ρ	$\delta = 0.1$	$\delta = 0.2$	$\delta = 0.3$	$\delta = 0.4$
30	0.1	0.001	0.113	0.653	0.943
	0.2	0.004	0.214	0.777	0.982
	0.3	0.006	0.358	0.915	0.999
	0.4	0.017	0.565	0.974	0.999
	0.5	0.079	0.782	0.998	1.000
	0.6	0.190	0.950	1.000	1.000
	0.7	0.513	0.999	1.000	1.000
	0.8	0.885	1.000	1.000	1.000
	0.9	0.997	1.000	1.000	1.000
50	0.1	0.328	0.991	1.000	1.000
	0.2	0.407	0.994	1.000	1.000
	0.3	0.470	1.000	1.000	1.000
	0.4	0.603	1.000	1.000	1.000
	0.5	0.814	1.000	1.000	1.000
	0.6	0.950	1.000	1.000	1.000
	0.7	0.999	1.000	1.000	1.000
	0.8	1.000	1.000	1.000	1.000
	0.9	1.000	1.000	1.000	1.000
100	0.1	1.000	1.000	1.000	1.000
	0.2	1.000	1.000	1.000	1.000
	0.3	1.000	1.000	1.000	1.000
	0.4	1.000	1.000	1.000	1.000
	0.5	1.000	1.000	1.000	1.000
	0.6	1.000	1.000	1.000	1.000
	0.7	1.000	1.000	1.000	1.000
	0.8	1.000	1.000	1.000	1.000
	0.9	1.000	1.000	1.000	1.000

Tableau D2 : MC_1rup, rupture avec continuité en présence de persistance : puissance (échantillon autocorrélé, $\alpha = 0.05$, $t_b = [n/2]$)

Données générées par $x_t = \mu_0 + \rho\, x_{t-1} + \delta\, ID(t_b) + \varepsilon_t$

où, $ID(t_b) = \begin{cases} 0 & t = 1,\ldots,t_b \\ (t - t_b) & t = t_b + 1,\ldots,n \end{cases}$

$\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $t_b = [n/2]$, $N = 99$, $M = 1000$ répétitions

Niveau critique α = 5% ($t_b = [n/2]$)					
n	ρ	$\delta = 0.1$	$\delta = 0.2$	$\delta = 0.3$	$\delta = 0.4$
30	0.1	0.000	0.023	0.222	0.608
	0.2	0.001	0.039	0.335	0.743
	0.3	0.002	0.082	0.498	0.917
	0.4	0.005	0.147	0.675	0.964
	0.5	0.013	0.309	0.873	1.000
	0.6	0.059	0.622	0.980	0.999
	0.7	0.175	0.882	1.000	1.000
	0.8	0.547	0.984	1.000	1.000
	0.9	0.948	1.000	1.000	1.000
50	0.1	0.082	0.890	0.997	1.000
	0.2	0.080	0.917	1.000	1.000
	0.3	0.136	0.968	1.000	1.000
	0.4	0.233	0.989	1.000	1.000
	0.5	0.399	0.997	1.000	1.000
	0.6	0.675	1.000	1.000	1.000
	0.7	0.907	1.000	1.000	1.000
	0.8	0.997	1.000	1.000	1.000
	0.9	1.000	1.000	1.000	1.000
100	0.1	0.999	1.000	1.000	1.000
	0.2	0.999	1.000	1.000	1.000
	0.3	1.000	1.000	1.000	1.000
	0.4	1.000	1.000	1.000	1.000
	0.5	1.000	1.000	1.000	1.000
	0.6	1.000	1.000	1.000	1.000
	0.7	1.000	1.000	1.000	1.000
	0.8	1.000	1.000	1.000	1.000
	0.9	1.000	1.000	1.000	1.000

Tableau D3 : MC_1rup, rupture avec continuité en présence de persistance : puissance (échantillon autocorrélé, $\alpha = 0.05$, $t_b = [2n/3]$)

Données générées par $x_t = \mu_0 + \rho\, x_{t-1} + \delta\, ID(t_b) + \varepsilon_t$

où, $ID(t_b) = \begin{cases} 0 & t = 1,\ldots,t_b \\ (t - t_b) & t = t_b + 1,\ldots,n \end{cases}$

$\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $t_b = [2n/3]$, $N = 99$, $M = 1000$ répétitions

Niveau critique α = 5% $(t_b = [2n/3])$					
n	ρ	$\delta = 0.1$	$\delta = 0.2$	$\delta = 0.3$	$\delta = 0.4$
30	0.1	0.000	0.000	0.029	0.102
	0.2	0.000	0.003	0.027	0.149
	0.3	0.000	0.003	0.050	0.287
	0.4	0.001	0.012	0.125	0.420
	0.5	0.002	0.039	0.264	0.670
	0.6	0.011	0.111	0.462	0.838
	0.7	0.026	0.260	0.710	0.961
	0.8	0.161	0.570	0.904	0.994
	0.9	0.701	0.917	0.988	1.000
50	0.1	0.004	0.288	0.870	0.994
	0.2	0.009	0.319	0.902	0.999
	0.3	0.018	0.414	0.939	1.000
	0.4	0.024	0.604	0.986	1.000
	0.5	0.050	0.779	0.996	1.000
	0.6	0.148	0.937	1.000	1.000
	0.7	0.439	0.996	1.000	1.000
	0.8	0.832	1.000	1.000	1.000
	0.9	1.000	1.000	1.000	1.000
100	0.1	0.845	1.000	1.000	1.000
	0.2	0.814	1.000	1.000	1.000
	0.3	0.823	1.000	1.000	1.000
	0.4	0.866	1.000	1.000	1.000
	0.5	0.916	1.000	1.000	1.000
	0.6	0.978	1.000	1.000	1.000
	0.7	1.000	1.000	1.000	1.000
	0.8	1.000	1.000	1.000	1.000
	0.9	1.000	1.000	1.000	1.000

Tableau D4 : MC_1rup, rupture avec continuité en présence de persistance : pourcentage d'estimation de la date $t_b = [n/3]$ (échantillon autocorrélé, $\alpha = 0.05$)

Données générées par $x_t = \mu_0 + \rho\, x_{t-1} + \delta\, ID(t_b) + \varepsilon_t$

où, $ID(t_b) = \begin{cases} 0 & t = 1,\ldots,t_b \\ (t - t_b) & t = t_b + 1,\ldots,n \end{cases}$

$\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $t_b = [n/3]$, $N = 99$, $M = 1000$ répétitions

Niveau critique α = 5% $(t_b = [n/3])$					
n	ρ	$\delta = 0.1$	$\delta = 0.2$	$\delta = 0.3$	$\delta = 0.4$
	0.1	0.00	1.10	14.3	27.2
	0.2	0.00	2.40	13.5	28.5
	0.3	0.30	4.40	18.1	25.7
	0.4	0.30	7.40	16.3	26.0
30	0.5	0.60	10.7	18.1	24.5
	0.6	0.90	10.6	15.2	21.8
	0.7	3.80	10.8	15.6	15.7
	0.8	3.60	10.2	14.0	12.9
	0.9	4.40	13.9	15.4	9.10
	0.1	2.30	20.2	29.1	35.6
	0.2	3.20	17.2	30.7	33.3
	0.3	4.30	16.7	29.9	31.9
	0.4	6.80	15.4	22.7	30.7
50	0.5	6.70	15.7	21.0	27.3
	0.6	5.70	16.3	20.3	21.5
	0.7	7.10	14.5	14.4	12.5
	0.8	6.20	9.60	8.00	2.40
	0.9	5.20	7.80	1.40	0.00
	0.1	38.0	24.3	38.0	48.1
	0.2	11.7	24.5	36.4	47.9
	0.3	11.7	23.1	32.1	41.9
	0.4	10.2	23.1	30.4	36.2
100	0.5	11.3	19.6	25.9	25.0
	0.6	11.2	17.5	15.0	12.3
	0.7	10.2	10.0	5.10	2.60
	0.8	7.60	5.10	0.70	0.00
	0.9	4.70	0.10	0.00	0.00

Tableau D5 : MC_1rup, rupture avec continuité en présence de persistance : pourcentage d'estimation de la date $t_b = [n/2]$ (échantillon autocorrélé, $\alpha = 0.05$)

Données générées par $x_t = \mu_0 + \rho\, x_{t-1} + \delta\, ID(t_b) + \varepsilon_t$

où, $ID(t_b) = \begin{cases} 0 & t = 1,\ldots,t_b \\ (t - t_b) & t = t_b + 1,\ldots,n \end{cases}$

$\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $t_b = [n/2]$, $N = 99$, $M = 1000$ répétitions

Niveau critique α = 5% $(t_b = [n/2])$					
n	ρ	$\delta = 0.1$	$\delta = 0.2$	$\delta = 0.3$	$\delta = 0.4$
30	0.1	0.00	0.30	5.40	16.4
	0.2	0.00	0.60	5.20	17.5
	0.3	0.00	0.90	7.60	21.7
	0.4	0.00	1.90	10.8	20.4
	0.5	0.10	3.30	14.6	20.4
	0.6	0.40	6.20	15.3	18.2
	0.7	0.80	8.00	11.4	15.9
	0.8	2.10	7.20	12.7	15.3
	0.9	2.30	6.60	13.9	14.5
50	0.1	1.00	16.8	25.2	34.1
	0.2	0.90	15.4	27.0	29.5
	0.3	1.40	17.4	22.0	31.7
	0.4	1.70	12.3	22.1	27.3
	0.5	3.00	15.4	18.0	23.8
	0.6	3.50	13.4	16.2	18.1
	0.7	4.90	12.9	13.4	10.8
	0.8	3.20	8.80	9.00	5.20
	0.9	3.50	6.10	5.30	0.50
100	0.1	14.0	23.4	36.1	45.9
	0.2	12.1	23.1	35.6	45.0
	0.3	12.6	19.9	32.3	36.0
	0.4	10.5	22.5	28.7	31.5
	0.5	10.4	18.3	21.9	22.7
	0.6	9.30	14.3	15.6	13.4
	0.7	8.00	8.80	6.70	1.50
	0.8	6.60	4.00	0.60	0.00
	0.9	2.60	0.50	0.00	0.00

Tableau D6 : MC_1rup, rupture avec continuité en présence de persistance : pourcentage d'estimation de la date $t_b = [2n/3]$ (échantillon autocorrélé, $\alpha = 0.05$)

Données générées par $x_t = \mu_0 + \rho\, x_{t-1} + \delta\, ID(t_b) + \varepsilon_t$

où, $ID(t_b) = \begin{cases} 0 & t = 1,\ldots,t_b \\ (t - t_b) & t = t_b + 1,\ldots,n \end{cases}$

$\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $t_b = [2n/3]$, $N = 99$, $M = 1000$ répétitions

Niveau critique α = 5% $(t_b = [2n/3])$					
n	ρ	$\delta = 0.1$	$\delta = 0.2$	$\delta = 0.3$	$\delta = 0.4$
30	0.1	0.00	0.00	0.30	1.80
	0.2	0.00	0.10	0.90	4.10
	0.3	0.00	0.00	0.20	5.80
	0.4	0.00	0.10	1.40	8.20
	0.5	0.00	0.40	2.40	11.3
	0.6	0.00	1.00	6.50	9.50
	0.7	0.10	1.80	7.30	10.3
	0.8	0.80	2.10	6.60	10.1
	0.9	0.80	1.70	4.00	6.70
50	0.1	0.00	3.60	21.0	30.1
	0.2	0.00	4.70	20.8	26.2
	0.3	0.00	6.40	20.6	26.2
	0.4	0.00	6.20	18.5	23.4
	0.5	0.30	9.10	16.4	18.9
	0.6	0.40	8.60	13.5	18.7
	0.7	0.60	6.10	8.50	11.4
	0.8	1.30	5.10	7.00	7.60
	0.9	0.50	2.00	6.00	4.60
100	0.1	9.40	20.9	29.5	42.1
	0.2	8.40	22.0	25.3	38.6
	0.3	7.90	15.9	25.6	35.0
	0.4	6.90	14.6	20.0	27.4
	0.5	7.70	13.9	21.1	21.3
	0.6	5.40	10.4	12.9	13.6
	0.7	4.00	10.0	7.90	5.00
	0.8	1.50	3.70	1.60	0.80
	0.9	0.70	1.40	0.10	0.00

Annexe E

Résultats de simulations pour plusieurs ruptures

Tableau E1 : MC_krup, plusieurs ruptures en absence de persistance: niveau (échantillon indépendant, $\alpha = 0.05$)

Données générées par $x_t = \mu_0 + \varepsilon_t$, avec $\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $N = 19$, $M = 1000$

n	**Niveau critique α = 5%**
30	0.043
50	0.028
100	0.016

Tableau E2 : MC_krup, plusieurs ruptures brusques en absence de persistance : puissance (échantillon indépendant , $\alpha = 0.05$ $t_b = [n/2]$)

Données générées par $x_t = \begin{cases} \mu_0 + \varepsilon_t & t = 1,\ldots,t_b \\ (\mu_0 + \Delta) + \varepsilon_t & t = t_b + 1,\ldots,n \end{cases}$

$\mu_0 = 181$, $t_b = [n/2]$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $N = 19$, $M = 1000$, $\boldsymbol{k}$: nombre de ruptures détectées

Niveau critique α = 5% (h = 10)						
n	k^1	$\Delta = 0.5\%$	$\Delta = 1\%$	$\Delta = 1.5\%$	$\Delta = 2\%$	$\Delta = 2.5\%$
30	0	0.687	0.077	0.000	0.000	0.000
	1	0.313	0.923	1.000	1.000	1.000
50	0	0.290	0.000	0.000	0.000	0.000
	1	0.710	1.000	1.000	1.000	1.000
100	0	0.042	0.000	0.000	0.000	0.000
	1	0.958	1.000	1.000	1.000	1.000

Tableau E3 : MC_krup, plusieurs ruptures brusques en absence de persistance : puissance (échantillon indépendant, $\alpha = 0.05$, $h = 5$, $t_b = [n/3]$, $t_b = [2n/3]$)

Données générées par $x_t = \begin{cases} \mu_0 + \varepsilon_t & t = 1,\ldots,t_{b_1} \\ (\mu_0 + \Delta) + \varepsilon_t & t = t_{b_1} + 1,\ldots,t_{b_2} \\ \{(\mu_0 + \Delta) + \Delta\} + \varepsilon_t & t = t_{b_2} + 1,\ldots,n \end{cases}$

$\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $N = 19$, $M = 1000$, $\boldsymbol{k}$: nombre de ruptures détectées

Niveau critique α = 5% (h = 5)						
n	k^1	$\Delta = 0.5\%$	$\Delta = 1\%$	$\Delta = 1.5\%$	$\Delta = 2\%$	$\Delta = 2.5\%$
30	0	0.908	0.000	0.000	0.000	0.000
	1	0.037	0.444	0.063	0.006	0.000
	2	0.052	0.502	0.808	0.824	0.868
	>2	0.003	0.054	0.129	0.170	0.132
50	0	0.005	0.000	0.002	0.000	0.000
	1	0.904	0.144	0.000	0.000	0.000
	2	0.088	0.763	0.796	0.809	0.831
	>2	0.003	0.093	0.202	0.191	0.169
100	0	0.000	0.000	0.000	0.000	0.000
	1	0.676	0.000	0.000	0.000	0.000
	2	0.317	0.787	0.793	0.801	0.876
	>2	0.007	0.213	0.207	0.199	0.124

Tableau E4 : MC_krup, plusieurs ruptures brusques en absence de persistance : puissance (échantillon indépendant, $\alpha = 0.05$, $h = 10$, $t_b = [n/3]$, $t_b = [2n/3]$)

Données générées par $x_t = \begin{cases} \mu_0 + \varepsilon_t & t = 1,\ldots,t_{b_1} \\ (\mu_0 + \Delta) + \varepsilon_t & t = t_{b_1}+1,\ldots,t_{b_2} \\ \{(\mu_0 + \Delta) + \Delta\} + \varepsilon_t & t = t_{b_2}+1,\ldots,n \end{cases}$

$\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $N = 19$, $M = 1000$, k : nombre de ruptures détectées

		Niveau critique α = 5% (h = 10)				
n	k^1	$\Delta = 0.5\%$	$\Delta = 1\%$	$\Delta = 1.5\%$	$\Delta = 2\%$	$\Delta = 2.5\%$
30	0	0.217	0.254	0.000	0.004	0.000
	1	0.729	0.202	0.046	0.000	0.001
	2	0.054	0.497	0.824	0.825	0.862
	>2	0.000	0.047	0.130	0.171	0.137
50	0	0.007	0.000	0.000	0.000	0.000
	1	0.894	0.121	0.000	0.000	0.000
	2	0.098	0.799	0.813	0.814	0.823
	>2	0.001	0.080	0.187	0.186	0.177
100	0	0.000	0.000	0.000	0.000	0.000
	1	0.661	0.000	0.000	0.000	0.000
	2	0.326	0.754	0.793	0.812	0.843
	>2	0.013	0.246	0.207	0.188	0.157

Tableau E5 : MC_krup, plusieurs ruptures brusques en absence de persistance : puissance de détection de la date $t_b = [n/2]$ (échantillon indépendant , $\alpha = 0.05$)

Données générées par $x_t = \begin{cases} \mu_0 + \varepsilon_t & t = 1,\ldots,t_b \\ (\mu_0 + \Delta) + \varepsilon_t & t = t_b+1,\ldots,n \end{cases}$

$\mu_0 = 181$, $t_b = [n/2]$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $N = 19$, $M = 1000$, k : nombre de ruptures détectées

		Niveau critique α = 5% (h = 10)				
n	k^1	$\Delta = 0.5\%$	$\Delta = 1\%$	$\Delta = 1.5\%$	$\Delta = 2\%$	$\Delta = 2.5\%$
30	0	91.9	50.2	25.4	16.1	11.7
	1	9.10	49.8	74.6	83.9	88.3
50	0	84.6	45.1	22.2	13.4	10.8
	1	15.4	54.9	77.8	86.6	89.2
100	0	76.9	44.4	22.5	10.9	8.60
	1	23.1	55.6	77.5	89.1	91.4

Tableau E6 : MC_krup, plusieurs ruptures brusques en absence de persistance : pourcentage d'estimation des dates $t_b = [n/3]$ et $t_b = [2n/3]$ (échantillon indépendant, $\alpha = 0.05$, $h = 5$)

Données générées par $x_t = \begin{cases} \mu_0 + \varepsilon_t & t = 1,\ldots,t_{b_1} \\ (\mu_0 + \Delta) + \varepsilon_t & t = t_{b_1}+1,\ldots,t_{b_2} \\ \{(\mu_0 + \Delta) + \Delta\} + \varepsilon_t & t = t_{b_2}+1,\ldots,n \end{cases}$

$\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $N = 19$, $M = 1000$, k : nombre de ruptures détectées

Niveau critique α = 5% (h = 5)						
n	k^1	$\Delta = 0.5\%$	$\Delta = 1\%$	$\Delta = 1.5\%$	$\Delta = 2\%$	$\Delta = 2.5\%$
30	0	90.8	0.00	0.00	0.00	0.00
	1	0.60	19.7	3.50	0.60	0.00
	2	0.10	12.3	45.2	62.8	74.9
	>2	0.30	5.40	12.9	17.0	13.2
50	0	0.50	0.00	0.20	0.00	0.00
	1	21.7	5.10	0.00	0.00	0.00
	2	0.40	19.7	45.9	61.4	73.2
	>2	0.30	9.30	20.2	19.1	16.9
100	0	0.00	0.00	0.00	0.00	0.00
	1	14.5	0.00	0.00	0.00	0.00
	2	2.10	26.0	48.0	63.4	70.1
	>2	0.70	12.4	19.9	20.7	21.3

Tableau E7 : MC_krup, plusieurs ruptures brusques en absence de persistance : pourcentage d'estimation des dates $t_b = [n/3]$ et $t_b = [2n/3]$ (échantillon indépendant, $\alpha = 0.05$, $h = 10$)

Données générées par $x_t = \begin{cases} \mu_0 + \varepsilon_t & t = 1,\ldots,t_{b_1} \\ (\mu_0 + \Delta) + \varepsilon_t & t = t_{b_1}+1,\ldots,t_{b_2} \\ \{(\mu_0 + \Delta) + \Delta\} + \varepsilon_t & t = t_{b_2}+1,\ldots,n \end{cases}$

$\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $N = 19$, $M = 1000$, k : nombre de ruptures détectées

Niveau critique α = 5% (h = 10)						
n	k^1	$\Delta = 0.5\%$	$\Delta = 1\%$	$\Delta = 1.5\%$	$\Delta = 2\%$	$\Delta = 2.5\%$
30	0	21.7	25.4	0.00	0.40	0.00
	1	21.2	9.80	2.90	0.00	0.10
	2	0.20	13.6	43.5	63.3	75.9
	>2	0.00	4.70	13.0	17.1	13.7
50	0	0.70	0.00	0.00	0.00	0.00
	1	23.1	4.70	0.00	0.00	0.00
	2	0.30	22.3	45.6	61.9	72.9
	>2	0.10	8.00	18.7	18.6	17.7
100	0	0.00	0.00	0.00	0.00	0.00
	1	13.2	0.00	0.00	0.00	0.00
	2	1.60	23.6	48.1	63.4	68.6
	>2	1.30	15.7	18.8	20.7	24.6

Tableau E8 : MC_krup, plusieurs ruptures en absence de persistance niveau (échantillon autocorrélé, $\alpha = 0.05$)

Données générées par $x_t = \rho x_{t-1} + \varepsilon_t$

$\varepsilon \overset{iid}{\sim} N(0,1)$, $N = 19$, $M = 1000$ répétitions

n	ρ	**Niveau critique α = 5%**
30	0.1	0.131
	0.2	0.162
	0.3	0.231
	0.4	0.312
	0.5	0.399
	0.6	0.548
	0.7	0.670
	0.8	0.739
	0.9	0.830
50	0.1	0.036
	0.2	0.078
	0.3	0.106
	0.4	0.201
	0.5	0.428
	0.6	0.560
	0.7	0.685
	0.8	0.822
	0.9	0.843
100	0.1	0.062
	0.2	0.096
	0.3	0.171
	0.4	0.283
	0.5	0.390
	0.6	0.550
	0.7	0.680
	0.8	0.815
	0.9	0.954

Tableau E9 : MC_krup, plusieurs ruptures en présence de persistance: niveau (échantillon autocorrélé, $\alpha = 0.05$)

Données générées par $x_t = \rho\, x_{t-1} + \varepsilon_t$

$\varepsilon \overset{iid}{\sim} N(0,1)$, $N = 19$, $M = 1000$ répétitions

n	ρ	**Niveau critique α = 5%**
30	0.1	0.001
	0.2	0.001
	0.3	0.001
	0.4	0.002
	0.5	0.002
	0.6	0.002
	0.7	0.002
	0.8	0.003
	0.9	0.003
50	0.1	0.001
	0.2	0.001
	0.3	0.001
	0.4	0.001
	0.5	0.002
	0.6	0.002
	0.7	0.002
	0.8	0.002
	0.9	0.002
100	0.1	0.001
	0.2	0.001
	0.3	0.001
	0.4	0.001
	0.5	0.001
	0.6	0.001
	0.7	0.001
	0.8	0.001
	0.9	0.001

Tableau E10 : MC_krup, plusieurs ruptures brusques en présence de persistance : puissance (échantillon indépendant , $\alpha = 0.05$ $t_b = [n/2]$)

Données générées par $x_t = \begin{cases} \mu_0 + \rho x_{t-1} + \varepsilon_t & t = 1,\ldots,t_b \\ (\mu_0 + \Delta) + \rho x_{t-1} + \varepsilon_t & t = t_b + 1,\ldots,n \end{cases}$

$\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $N = 19$, $M = 1000$, $\boldsymbol{k}$: nombre de ruptures détectées

Niveau critique α = 5% (h = 10)							
n	ρ	k^1	$\Delta = 0.5\%$	$\Delta = 1\%$	$\Delta = 1.5\%$	$\Delta = 2\%$	$\Delta = 2.5\%$
30	0.1	1	0.009	0.020	0.053	0.045	0.039
	0.2	1	0.005	0.023	0.032	0.049	0.034
	0.3	1	0.004	0.018	0.031	0.048	0.043
	0.4	1	0.003	0.019	0.024	0.031	0.035
	0.5	1	0.002	0.018	0.025	0.035	0.034
	0.6	1	0.001	0.009	0.029	0.030	0.040
	0.7	1	0.004	0.013	0.020	0.033	0.027
	0.8	1	0.003	0.003	0.020	0.037	0.041
	0.9	1	0.002	0.010	0.019	0.034	0.035
50	0.1	1	0.013	0.026	0.036	0.040	0.037
	0.2	1	0.019	0.032	0.050	0.040	0.037
	0.3	1	0.013	0.039	0.027	0.032	0.039
	0.4	1	0.009	0.022	0.038	0.040	0.045
	0.5	1	0.007	0.025	0.043	0.047	0.047
	0.6	1	0.002	0.019	0.046	0.041	0.038
	0.7	1	0.007	0.021	0.031	0.043	0.043
	0.8	1	0.001	0.022	0.040	0.035	0.035
	0.9	1	0.010	0.044	0.068	0.053	0.048
100	0.1	1	0.052	0.083	0.104	0.093	0.081
	0.2	1	0.066	0.078	0.095	0.075	0.080
	0.3	1	0.055	0.090	0.093	0.087	0.079
	0.4	1	0.049	0.069	0.072	0.073	0.072
	0.5	1	0.043	0.079	0.082	0.077	0.070
	0.6	1	0.029	0.073	0.081	0.078	0.069
	0.7	1	0.034	0.070	0.079	0.076	0.068
	0.8	1	0.031	0.067	0.070	0.073	0.059
	0.9	1	0.042	0.065	0.069	0.070	0.053

Tableau E11 : MC_krup, plusieurs ruptures brusques en présence de persistance : pourcentage d'estimation de la date $t_b = [n/2]$ (échantillon indépendant , $\alpha = 0.05$)

Données générées par $x_t = \begin{cases} \mu_0 + \rho x_{t-1} + \varepsilon_t & t = 1,\ldots,t_b \\ (\mu_0 + \Delta) + \rho x_{t-1} + \varepsilon_t & t = t_b + 1,\ldots,n \end{cases}$

$\mu_0 = 181$, $\varepsilon \overset{iid}{\sim} N(0,1)$, $N = 19$, $M = 1000$, $\boldsymbol{k}$: nombre de ruptures détectées

Niveau critique α = 5% (h = 10)

n	ρ	k^1	$\Delta = 0.5\%$	$\Delta = 1\%$	$\Delta = 1.5\%$	$\Delta = 2\%$	$\Delta = 2.5\%$
	0.1	1	0.00	0.00	0.00	0.00	0.00
	0.2	1	0.00	0.00	0.00	0.00	0.00
	0.3	1	0.00	0.00	0.00	0.00	0.00
	0.4	1	0.00	0.00	0.00	0.00	0.00
30	0.5	1	0.00	0.00	0.00	0.00	0.00
	0.6	1	0.00	0.00	0.00	0.00	0.00
	0.7	1	0.00	0.00	0.00	0.00	0.00
	0.8	1	0.00	0.00	0.00	0.00	0.00
	0.9	1	0.00	0.00	0.00	0.00	0.00
	0.1	1	0.10	0.20	0.20	0.01	0.01
	0.2	1	0.10	0.10	0.00	0.00	0.00
	0.3	1	0.00	0.20	0.00	0.00	0.00
	0.4	1	0.00	0.00	0.10	0.00	0.00
50	0.5	1	0.00	0.10	0.10	0.00	0.00
	0.6	1	0.00	0.00	0.00	0.00	0.00
	0.7	1	0.00	0.10	0.00	0.00	0.00
	0.8	1	0.00	0.20	0.00	0.00	0.00
	0.9	1	0.00	0.10	0.30	0.00	0.00
	0.1	1	0.50	0.80	1.50	1.00	0.80
	0.2	1	0.70	0.90	1.10	0.90	0.81
	0.3	1	0.40	0.70	1.20	1.00	0.78
	0.4	1	0.40	0.70	1.10	0.80	0.74
100	0.5	1	0.20	0.70	1.10	0.90	0.75
	0.6	1	0.10	0.50	0.90	0.70	0.64
	0.7	1	0.50	0.40	0.80	0.50	0.47
	0.8	1	0.10	0.30	0.60	0.40	0.37
	0.9	1	0.20	0.30	0.50	0.30	0.25