

République du Cameroun

Paix - Travail - Patrie

Université de Yaoundé I

Faculté des sciences

Département de Mathématiques

Master de statistique Appliquée



Republic of Cameroon

Peace - Work - Fatherland

The University of Yaoundé I

Faculty of science

Department of Mathematics

Master of Applied Statistics

**MODELE DE PREVISION DU PARC
DES COMPTES D'UNE BANQUE :
CAS AFRILAND FIRST BANK**

Mémoire présenté et soutenu en vue de l'obtention

Du

MASTER DE STATISTIQUE APPLIQUEE

Par

M. NANA YAKAM ANDRE

Matricule 04U602

Maître ès sciences en Mathématiques Appliquées

Sous la Direction de

DR. EUGENE PATRICE NDONG NGUEMA

Chargé de Cours, Ecole Nationale Supérieure Polytechnique

Sous la Supervision de

PR. HENRI GWÉT

Maître de Conférences, Ecole Nationale Supérieure Polytechnique

Année académique 2004 - 2005

SOMMAIRE :

Dédicace	3
Remerciements	4
Résumé/Abstract	5
Avant propos	6
Liste des Abréviations	7
Introduction	8
Chapitre I : Revue de la littérature et données utilisées pour l'Analyse	
I. Méthodes d'analyse des séries chronologiques	10
II. Exemple d'étude utilisant les trois types de modèles	11
III. Présentation de AFRILAND FIRST BANK	12
IV. Collecte et Traitement des données	12
Chapitre II : Analyse Descriptive des comptes	
I. Part de chaque produit dans les activités	16
II. Part de chaque agence dans le parc des comptes	17
III. Evolution des ouvertures des comptes	19
IV. Effets saisonniers	22
Chapitre III : Description de la démarche de BOX et JENKINS pour la prévision d'une série chronologique	
I. Identification a priori	26
II. Estimation des paramètres	35
III. Identification a posteriori	36
IV. Prévision	37
V. Algorithme de Box-Jenkins	38
Chapitre IV : Applications de la méthode de BOX-JENKINS à la prévision du nombre des comptes d'épargne de la First Bank	
I. Identification a priori	40
II. Estimation des coefficients et validation des modèles	46
III. Prévision	48
Conclusion	53
Annexe	54
Bibliographie	61

DEDICACE :

Je dédie ce travail à ma Maman, Madame
veuve YANKAM, née **MBAKOP**
Hélène dit « **MATCHOU** »

REMERCIEMENTS :

Mes remerciements vont d'abord à tous les enseignants du Master de statistique qui n'ont ménagé aucun effort pour nous donner une formation de qualité. Il s'agit de :

- Pr. Didier DACUNHA CASTELLE ;
- Pr. Jean COURSOL
- Pr. Bertrand AUVERT ;
- Pr. Danielle FLORENS ;
- Pr. Jean Christophe THALABARD ;
- Pr. Jean Louis GOLMARD ;
- Pr. Xavier NGUYON ;
- Pr. Elisabeth GASSIAT ;
- Pr. Henri GWET ;
- Dr. NDONG NGUEMA
- Dr. Michel NDOUMBE NKENG ;
- Dr. Maxime KIKI.

Je remercie tout le personnel de AFRILAND FIRST BANK pour l'accueil et l'encadrement qu'il m'a accordés pendant mon stage, un merci particulier à mon encadreur professionnel, Monsieur TIENTCHEU Janvier pour sa disponibilité tout au long de mon stage.

Je remercie Monsieur FOGAING KAMMOGNE Christian et Mlle NGUEYAP Annie Flore pour leur soutien avant et pendant mon stage à la FIRST BANK.

Je remercie toute ma famille, particulièrement mon papa, MBEU YAKEP né NANSHI Jean qui a toujours été à mes côtés quand j'en ai besoin.

Que toutes les personnes qui m'ont soutenu tout au long de cette formation trouvent ici ma profonde gratitude. Je pense notamment à :

- Personnel de la Maison de la Culture Française de Douala ;
- Mademoiselle WONGWE KAMANI Claire ;
- Monsieur NGOH Félix ;
- Monsieur KAMGHO TEZANOU ;
- Monsieur CHASSEM Narcisse ;
- Monsieur DONGMO Nicolas ;
- Monsieur NJOH NJOH Claude ;
- Monsieur PIEDJO K. Joachim ;
- Monsieur KEMGANG Joseph ;
- Monsieur TCHUISSEU NJOMO APPOLOS ;
- Monsieur TANKEU NGANSO Jean Michel ;
- Monsieur TCHOUAMANI Joseph Bertin ;
- Tous mes camarades du Master.

A tous ceux qui de près ou de loin ont contribué à la réussite de ce travail, je dis merci.

Loué soit le Seigneur tout-puissant qui m'a gardé en bonne santé et qui m'a permis de conduire ce travail jusqu'à son terme.

RESUME

De tout temps, la prévision a toujours été la préoccupation des Hommes, en ce qu'elle permet d'anticiper sur certains faits et événements. Elle prévient parfois des mauvaises surprises. Il est question dans ce mémoire de proposer un modèle de prévision du nombre des comptes pouvant être ouverts mensuellement dans une banque, avec l'application au cas AFRILAND FIRST BANK. Vu les besoins immédiats de cette Banque, il nous a semblé utile de consacrer une partie de notre travail à une description approfondie des comptes ouverts dans ses différentes agences, depuis le mois de juillet 1998 jusqu'à juin 2005. La prévision proprement dite dont il est question ici repose sur la théorie des séries chronologiques. La méthodologie que nous avons mise en œuvre dans ce mémoire pour l'analyse de ces séries chronologiques est celle de BOX et JENKINS.

ABSTRACT

Our aim in this memory is to build a **model for forecasting the number of accounts that can be opened monthly in a bank : case of AFRILAND FIRST BANK**. This is a microeconomic research which uses time-series models for forecasting.

There is a long tradition of building large microeconomic models with hundreds of variables. More recently, small models with only a handful of variables have become more common. In the years 1970, **George Box** and **Gwilym Jenkins** propose a methodology which views time-series as realizations of stochastic processes. This approach allows the model builder to consider that economic variables are integrated in the previous values of time series. Then, they can base their models only in those values. This is the method we use in this memory, for forecasting the series of accounts open per month in all AFRILAND FIRST BANK's agencies in Cameroon, between July 1998 and June 2005.

AVANT-PROPOS

Le Master de Statistique est une formation de 3^e cycle ouverte cette année 2004 à l'Université de Yaoundé I. Il s'agit d'une formation professionnalisante et d'initiation à la recherche, qui bénéficie de la collaboration et du soutien de plusieurs institutions Françaises et Africaines (Universités de Paris Orsay, Paris 5, Versailles, Paris Dauphine, l'INSERM, l'Université de Cocody et l'institut National Polytechnique de Côte d'Ivoire).

Son objectif général est de donner aux étudiants, cadres supérieurs d'entreprise et d'administration, et tout utilisateur de la statistique, une formation de haut niveau très pratique, classique quant aux techniques mathématiques utilisées, aussi moderne que possible, quant à l'informatique et aux logiciels spécialisés utilisés. Le Master apporte aux étudiants ayant les acquis fondamentaux en mathématiques et en statistique, une formation professionnelle complémentaire dans le domaine du traitement de l'information et de son exploitation.

L'admission au Master de statistique se fait sur étude de dossiers. Pour la première année, sont admissibles les personnes titulaires d'une licence de mathématiques ou d'un diplôme jugé équivalent, et pour la deuxième année, les personnes titulaires d'une maîtrise de mathématiques ou d'un diplôme jugé équivalent. Le recrutement est étendu aux médecins, agronomes et biologistes. Ces derniers suivent des cours de mise à niveau en mathématiques et en statistique.

Dans le souci de permettre aux étudiants de mettre en pratique les connaissances acquises tout au long de la formation, il leur est demandé d'effectuer un stage académique d'une durée de trois à six mois en entreprise. Ce stage consiste pour l'étudiant, à apporter une contribution à la résolution des problèmes statistiques auxquels l'entreprise est confrontée. C'est dans ce cadre que nous avons été reçu, du 13 juin au 12 septembre 2005, à la société **AFRILAND FIRST BANK**, une banque commerciale et d'investissement. Notre mission au sein de cette entreprise consistait, d'une part, à faire une étude descriptive du parc des comptes de la banque, et, d'autre part, à faire des prévisions sur l'évolution de ce parc des comptes.

Notre formation s'est bien déroulée, aussi bien dans le respect des programmes de formation que du calendrier académique. Cependant, vu son contenu dense, elle s'est déroulée dans un rythme très accéléré. Etant la première promotion du Master, nous n'avons pas bénéficié des infrastructures et de la logistique initialement prévues pour la formation, notamment la documentation. Nous remercions l'indulgence des lecteurs pour les insuffisances que pourrait présenter ce travail et nous les invitons à formuler les critiques en vue de son amélioration.

LISTE DES ABREVIATIONS

CE : Comptes d'épargne;

CDS : Comptes de Dépôt Spécial;

CCC : Comptes Courants Chèques;

MC² : Mutuelles Communautaires de Croissance;

CEL : Compte d'épargne sur livret;

CEI : Compte d'épargne Investissement;

i-Card : intelligent Card;

MCO : Moindre Carrés Ordinaires;

AR : Autoregressive;

MA : Mean Average;

ARMA : Autoregressive Mean Average;

ARIMA : Autoregressive Integrated Mean Average;

ARCH : Autoregressive Conditional Heteroscedastic;

GARCH : Generalise Autoregressive Conditionally Heteroscedastic;

AC : Autocorrelation;

PAC : Partial Autocorrelation;

ACF : Autocorrelation Function;

PACF : Partial Autocorrelation Function;

CVS : Corrigé des Variations Saisonnières;

TS : Trend Stationary;

DS : Difference Stationary;

INTRODUCTION :

Les données de toute entreprise constituent une richesse parce qu'elles contiennent des informations dont l'extraction judicieuse peut lui permettre d'avoir une interprétation claire de son passé afin de comprendre le présent et d'envisager ses actions futures. La statistique intègre des outils d'analyse des données qui s'imposent comme instrument de prévision et d'aide à la prise de décision. Aucune entreprise, encore moins une banque, ne saurait s'y dérober de nos jours.

La banque est un secteur d'activité très délicat dont l'une des priorités est de mettre ses clients en confiance, de par la qualité, mais aussi ses délais de service. La satisfaction des clients passe avant tout par la maîtrise de leurs effectifs, qui suppose non seulement leur nombre à un instant donné, mais surtout l'évolution de ce nombre au cours du temps. Ceci peut permettre de prendre les dispositions optimales pour les satisfaire. Les enjeux statistiques au niveau des banques vont bien au-delà de cette satisfaction de la clientèle. Ils peuvent s'étendre à d'autres agrégats, comme l'automatisation de l'acceptation, ou non d'un nouveau client, la mesure du risque avant tout investissement ou tout accord de crédit etc.

AFRILAND FIRST BANK est une banque commerciale et d'investissement créée en 1987 par deux groupes d'actionnaires, un groupe Africain conduit par le Docteur FOKAM KAMMOGNE Paul et un groupe financier Européen dont le chef est la FMO, un organisme Néerlandais pour le développement. Avec un capital social de 5.500.000.000 FCFA, la FIRST BANK compte aujourd'hui parmi les plus grandes banques installées au Cameroun.

La tâche qui nous a été confiée à la FIRST BANK pendant notre stage consistait à étudier d'une part l'évolution du parc des comptes, et à construire un modèle de prévision pour ce parc d'autre part. En d'autres termes, nous devons étudier l'évolution du nombre de comptes ouverts ou clôturés par les clients de cette institution financière lors des sept dernières années et puis, proposer un modèle statistique de prévision du nombre de comptes susceptibles d'être ouverts dans les douze mois à venir, c'est-à-dire de juillet 2005 à juin 2006. C'est ainsi qu'il nous a été proposé de réfléchir d'une manière générale sur le **Modèle de prévision du parc des comptes d'une banque** et de l'appliquer au **cas AFRILAND FIRST BANK**.

Le problème posé ici relève d'un phénomène économique qui évolue en fonction du temps. Il fait donc appel aux séries chronologiques. Généralement, plusieurs approches sont utilisées pour l'analyse d'une série chronologique. La méthodologie que nous adoptons dans ce mémoire est celle de BOX et JENKINS. Ils proposent une approche algorithmique dans la résolution des problèmes de modélisation et de prévision des séries chronologiques.

Dans le premier chapitre de ce mémoire, nous présentons une revue de la littérature, la société d'accueil et les données utilisées. Dans le second chapitre, nous faisons une description des données. Au troisième chapitre, nous présentons la méthodologie de Box et Jenkins pour l'analyse des séries chronologiques. Au dernier chapitre, nous appliquons cette approche pour la prévision du parc des comptes de la FIRST BANK. Pour ce faire, nous avons recueilli les données mensuelles d'ouverture et de clôture des comptes dans les différentes agences, que nous analysons après un regroupement des données de ces agences par ville.

CHAPITRE I :

**REVUE DE LA LITTERATURE ET
DONNEES UTILISEES POUR L'ANALYSE**

I. METHODES D'ANALYSE DES SERIES CHRONOLOGIQUES

« L'analyse statistique d'une série chronologique peut être simplement de nature descriptive. Il s'agit alors de dégager des éléments de synthèse, en général sous forme de nouvelles séries, qui résument au mieux la grandeur observée ou d'exhiber certaines caractéristiques. Cette phase descriptive utilise principalement des méthodes non paramétriques, qui permettent de travailler avec peu d'hypothèses préalables, mais peut faire appel à des techniques paramétriques importantes. Elle est souvent une étape préalable à une analyse plus complète dans laquelle on cherche à expliquer le mécanisme générateur de la série. A cette fin, on a recours à des modèles probabilistes qui, le plus souvent, sont de nature paramétrique. C'est en particulier le cadre idéal pour faire de la prévision, qui reste un des objectifs principaux de l'étude d'une série chronologique. » [4]

On distingue ainsi trois manières d'exprimer l'évolution d'une série [7] :

1. MODELES D'AJUSTEMENT

Parmi les méthodes les plus utilisées, la partie structurée de la grandeur étudiée est considérée comme étant une fonction déterministe du temps et représente la moyenne des variables aléatoires observées ou d'une transformation simple de celles-ci. La partie stochastique est alors une erreur additive et forme une séquence de variables aléatoires centrées, non corrélées et de même variance. L'hypothèse gaussienne est retenue pour la construction d'intervalles de confiance ou de tests. La régression linéaire par moindres carrés ordinaires, le lissage par moyennes mobiles, le lissage exponentiel simple ou double et les méthodes de HOLT et WINTERS constituent l'essentiel de cette approche.

2. MODELES AUTOPROJECTIFS

Cette méthodologie est dite « de BOX et JENKINS ». La partie prévisible de la série, à un instant donné, est exprimée comme une fonction mathématique des valeurs passées. La structure du phénomène se traduit donc par les corrélations entre les variables. Il consiste à supposer que les observations, après transformation éventuelle, sont issues d'une série stationnaire au second ordre. Les modèles autorégressifs moyennes mobiles (ARMA) constituent la base de cette seconde approche. Certaines méthodes permettent d'échapper aux hypothèses de stationnarité ou de linéarité. Les modèles autorégressifs moyennes mobiles intégrés (ARIMA), les modèles autorégressifs conditionnellement hétéroscédastiques (ARCH) et les modèles GARCH (ARCH généralisé) en sont les exemples principaux.

3. MODELES EXPLICATIFS

Le modèle explicatif général d'une série devrait pouvoir traduire l'influence du passé de la série sur son niveau présent, mais aussi d'autres variables exogènes sur la série.

Prenons l'exemple de l'ouverture des comptes dans une banque : l'évolution du nombre de comptes ouverts ou à ouvrir dépend de l'environnement socioéconomique (inflation, chômage, croissance,...), de la politique interne de la banque au cours du temps (marketing, nouveaux produits, création des nouvelles agences,...) et aussi du comportement des sociétés concurrentes (augmentation ou diminution du nombre, taux d'intérêt ou agios pratiqués, ...).

II. EXEMPLE D'ETUDE UTILISANT LES TROIS TYPES DE MODELES

En 2003, ALLAN RIDING et GEORGE HAINES FILS [5] du cabinet d'étude « EQUINOX MANAGEMENT CONSULTANTS LTD » du Canada ont fait une étude pour le compte « d'INDUSTRIE CANADA » et de « l'ADMINISTRATION DES PRETS AUX PETITES ENTREPRISES » au Canada. Cette étude portait sur la « **Prévision des coûts liés aux prêts accordés en vertu de la LPPE (Loi sur les prêts aux petites entreprises) et de la LFPEC (Loi sur le financement des petites entreprises du Canada)** »

En fait, de 1961 à 1999 au Canada, la *Loi sur les prêts aux petites entreprises* (LPPE) a facilité le prêt commercial aux petites entreprises canadiennes en prévoyant un mécanisme de partage des pertes sur prêts. Depuis 1999, cette loi a été remplacée par la *Loi sur le financement des petites entreprises du Canada* (LFPEC), qui poursuit le même objectif. Depuis avril 1995, le programme de partage des pertes sur prêts doit recouvrer les coûts liés au versement des indemnités portant sur les prêts enregistrés, afin de déterminer avec plus de précision si cette exigence sera satisfaite.

L'étude consistait à élaborer des modèles permettant de prévoir le volume des demandes d'indemnisation pour pertes sur prêts. Ainsi, l'ADMINISTRATION DES PRETS AUX PETITES ENTREPRISES sera en mesure de prévoir avec exactitude les coûts liés aux prêts consentis en vertu de la LPPE et de la LFPEC.

Quatre types de modèles ont été mis en œuvre dans ladite étude :

- Le **lissage exponentiel**, qui consiste à ajuster un modèle paramétrique de façon locale aux données, c'est-à-dire un modèle dont les paramètres évoluent au cours du temps ;
- Le **modèle ARIMA** (modèle autorégressif moyenne mobile intégré) ;
- La **modélisation économétrique**, qui établit des rapports quantitatifs entre la variable étudiée et la valeur des principaux paramètres économiques ;
- La **modélisation ARIMAX**, qui est la combinaison de la modélisation économétrique et de la méthode ARIMA.

Des comparaisons ont été ensuite faites entre les résultats des prévisions issues des différents modèles et les montants réels observés après la fin de la période de l'étude. Ce qui a permis de conclure que la modélisation économétrique et le modèle ARIMA ont des faibles taux d'erreurs par rapport aux deux autres modèles. Cette conclusion pourrait varier selon les données utilisées, c'est d'ailleurs la raison pour laquelle les auteurs de cet article conseillent, la moyenne des prévisions issues des différents modèles comme étant la meilleure prévision. Cette étude confirme aussi que le lissage exponentiel se prête mieux aux prévisions à court terme.

Dans le cadre de notre étude, dans l'impossibilité d'entrer en possession des informations pouvant nous permettre de déterminer les variables exogènes qui influencent significativement la variable étudiée, nous les considérerons comme intégrées aux données. Nous nous limiterons, dans la partie descriptive, à signaler leurs effets éventuels lorsqu'ils seraient fortement ressentis à des moments donnés dans l'évolution des données. Nous utiliserons donc les modèles autoprojectifs pour procéder à la prévision.

III. PRÉSENTATION DE AFRILAND FIRST BANK

Créée en 1987 au Cameroun par le Docteur FOKAM KAMMOGNE Paul, la CCEI (Caisse Commune d'Épargne et d'Investissement) devenue **AFRILAND FIRST BANK** en 2002 est une Société Anonyme (SA) au capital de 5.500.000.000 FCFA (cinq milliards cinq cent millions de francs CFA). La **FIRST BANK** (diminutif de **AFRILAND FIRST BANK**), est une banque privée qui a délibérément choisi d'être la banque du plus grand nombre.

La **FIRST BANK** offre à sa clientèle :

- **Les Comptes rémunérés** (Bon de caisse, Dépôt à terme et les Compte d'épargne)
- **Les Comptes non rémunérés** (Comptes chèque, Comptes de dépôt spécial, Compte de dépôt islamique)
- **FLASH CASH** (Une solution de sécurisation de l'épargne)
- **MoneyGram / Money First** (transfert d'argent);
- **MC² / MUFFA** (Mutuelles de lutte contre la pauvreté)
- **Fonds de capital risque** (Financement des micro-projets)
- **i-Card** (Monnaie électronique).

La **FIRST BANK** compte 13 agences installées dans sept grandes villes du territoire Camerounais (Yaoundé, Douala, Bafoussam, Nkongsamba, Garoua, Bamenda et Kousseri).

Elle a une filiale à SAO TOME et PRINCIPE, en GUINEE EQUATORIALE, en République Démocratique du CONGO, une représentation en CHINE et en France.

La **FIRST BANK** a un réseau de correspondants dans le monde entier (CITI BANK, DZ BANK, National BANK of Canada, CBIP...).

Son siège social est à Yaoundé, 1063 Place de l'Indépendance.

IV. COLLECTE ET TRAITEMENT DES DONNEES

1. COLLECTE

La base des données centrale de la FIRST BANK est gérée par la Direction des Systèmes d'Information Groupe (DSIG), vu la sécurité que requiert la gestion des données de toute entreprise, à fortiori celles d'une banque, les droits d'accès aux différents serveurs de la FIRST BANK sont limités pour chaque utilisateur du système informatique.

Lorsqu'un responsable ou un service de la banque a besoin des données qui ne lui sont pas directement accessibles, il s'adresse à la DSIG qui se charge de les rassembler pour les transmettre au demandeur, sur un support numérique ou par courrier électronique et au format sollicité par ce dernier.

Les données mises à notre disposition concernent le nombre de comptes ouverts ou clôturés par mois dans les différentes agences de la FIRST BANK du Cameroun (Au total 13 agences réparties dans sept villes du Cameroun). Il s'agit d'une étude rétrospective de la période allant du mois de juillet 1998 à juin 2005.

Les comptes bancaires à la FIRST BANK obéissent à une numérotation de 13 chiffres dont trois sont réservés au code du produit. Cette codification permet de distinguer les différents types de comptes à partir du numéro du compte. On compte plus d'une dizaine de types distincts (Comptes d'épargne, comptes chèques, ...).

Toutes les données dont nous avons besoin devaient être collectées auprès de la DSIG, tandis que notre stage s'effectuait à une autre direction, la Direction Commerciale et d'Exploitation (DCE). Ces données ont pris du temps à nous parvenir, car les données en question étaient stockées dans des serveurs différents et les responsables de la DSIG étaient occupés par leurs activités quotidiennes.

Les données mises à notre disposition étaient sur support numérique, compressées pour la plupart, aux formats ".txt " (**Bloc-notes**), ".doc" (**Microsoft Word**), ".xls" (**Excel**) ou ".unl" (**base des données de la banque**). Elles portaient sur le nombre de comptes ouverts ou clôturés classés par mois, par type, et par agence de la FIRST BANK. Ces fichiers classés par exercice budgétaire, de 1998 à 2005, comprenaient chacun trois colonnes : une pour les numéros des mois dans chaque exercice, l'autre pour les codes des produits et la troisième pour les nombres des comptes ouverts ou clôturés dans chaque agence.

2. TRAITEMENT DES DONNEES

Une fois les données reçues, la première opération à faire a consisté à les importer dans Excel, pour celles qui ne l'étaient pas encore, ce qui nous permettait d'avoir des tableaux en trois colonnes comme nous l'avons mentionné ci-dessus. Par la suite, il fallait se servir de la codification des produits, laquelle a connue une modification à partir de janvier 2004, pour reconnaître les différents types des comptes. Cette tâche d'identification des codes attribués à chaque compte a été facilitée par notre séjour dans les services d'ouverture des comptes lors de la première semaine de notre stage.

Après l'identification des produits, il fallait extraire les données pour chaque type de compte, agence par agence, les ranger dans des colonnes distinctes d'Excel pour enfin avoir les données de chaque agence dans un seul fichier Excel. Nous devions par la suite additionner les données des agences appartenant à une même ville pour avoir la situation dans chaque ville. Après, il ne nous restait plus qu'à regrouper les données des différentes villes, les sommer pour avoir le total sur l'ensemble du réseau national.

Au bout de cette procédure, nous nous sommes retrouvés avec huit fichiers Excel (un fichier pour chacune des sept villes citées ci-dessus et le 8^e pour l'ensemble du réseau national) à deux classeurs chacun (l'un pour l'ouverture et l'autre pour la clôture des comptes), contenant plus d'une dizaine de variables pour certains.

Après une séance de travail avec notre encadreur professionnel, il nous a été instruit de regrouper certains produits et d'orienter notre réflexion sur trois types de comptes. Les regroupements suivants ont ainsi été faits :

TYPES DE COMPTES	REGROUPEMENT
Compte d'épargne sur livret (CEL)	Comptes d'épargne
Compte d'épargne investissement (CEI)	
Comptes MC ²	
Comptes de dépôt spécial pour les personnes physiques	Comptes de dépôt spécial
Comptes de dépôt spécial pour les personnes morales	
Comptes courant pour les personnes physiques	Comptes courant chèque
Comptes courant pour les personnes morales	

C'est à ces trois types de comptes regroupés que nous devrions désormais nous intéresser. Nous étions appelés à faire les sommations nécessaires pour enfin obtenir, pour chaque ville, que nous appelons désormais agence, deux tableaux, un pour l'ouverture et l'autre pour la clôture des comptes, avec les trois variables issues du regroupement ci-dessus.

Ne pouvant pas appeler directement un tableau Excel dans **R**, le logiciel de statistique que nous avons utilisé pour l'analyse, nous étions obligés de les convertir au format texte pour en fin avoir nos **DATA FRAME**, structure de données de base dans **R**.

3. DONNEES MANQUANTES ET REMARQUES

Signalons qu'à Yaoundé où la FIRST BANK compte trois agences, nous n'avons pas pu obtenir les données de l'agence de MESSA au complet. Les données de Yaoundé n'intègrent celles de cette agence (MESSA) qu'à partir de Janvier 2003. En pareille situation, nous devons estimer les données manquantes. Mais dans le cas présent, presque toutes les données sont manquantes, car même les données de Yaoundé de 2004 et 2005 mises à notre disposition regroupe les données de ses trois agences et ne sont pas séparables agence par agence. Notre encadreur professionnel nous a rassuré que l'absence des données de MESSA ne devrait pas grandement influencer l'étude du fait du poids faible de cette agence à Yaoundé.

De même, les données de Nkongsamba concernant les comptes d'épargne n'intègrent que les données des comptes d'épargne sur livret (CEL), et sont par conséquent incomplètes. Compte tenu du regroupement opéré ci-dessus, l'étude des comptes d'épargne pourrait être faussée si les comptes MC² et CEI s'avéraient importants dans cette localité. Là encore notre requête pour l'avoir au complet n'a pas pu aboutir jusqu'au démarrage de l'étude, n'ayant aucune observation du tout, aucune technique ne nous permettait de les estimer. Cependant, ayant constaté que les comptes MC² et CEI n'avaient pas des chiffres mensuels élevés dans les autres localités, nous avons intégré les données de Nkongsamba comme telles dans l'étude.

Bafoussam présente une très grande variabilité dans ses données. Elles nous semblaient erronées au départ, mais après les avoir renvoyées trois fois et reçue la confirmation des services compétents qu'elles sont bonnes, nous les avons intégrées telles quelles dans notre étude.

CHAPITRE II :

ANALYSE DESCRIPTIVE DES COMPTES

I. PART DE CHAQUE PRODUIT DANS LES ACTIVITES

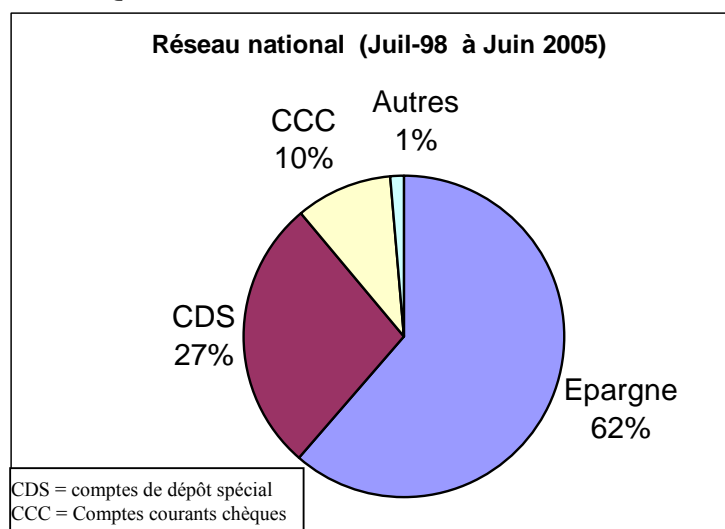


Figure 1 : Pourcentages des différents types de comptes ouverts de juillet 1998 à juin 2005.

Les activités de AFRILAND FIRST BANK sont fortement dominées par les comptes d'épargne, conséquence logique des objectifs qu'elle s'est assigné depuis sa création, qui consistent à ramener la banque à la portée du plus grand nombre de personnes et générer suffisamment de fonds pour financer les petites et moyennes entreprises. Cette politique se traduit par le montant minimum requis à l'ouverture des comptes d'épargne, 200.000 FCFA, qui reste inférieur à ce qui est exigé dans d'autres institutions concurrentes (500.000 FCFA en moyenne), sans négliger l'aspect culturel de banque africaine qui prévaut dans son fonctionnement. Or, ces comptes présentent un avantage en ce sens qu'ils constituent la ressource essentielle, source de prêts aux clients de la banque permettant à cette dernière d'en générer des intérêts. Mais, du fait que ce type de compte donne droit aux titulaires à un taux d'intérêt annuel de 5%, sa collecte effrénée peut réduire la rentabilité de la banque.

Les comptes de dépôt spécial (CDS) donnent droit à la quasi totalité des opérations bancaires, mais on n'exige qu'un minimum de 150.000 FCFA à l'ouverture et les frais de tenue de compte qui s'élèvent à 5965 FCFA TTC par trimestre. Ce type de compte occupe la deuxième place des produits de la FIRST BANK. Ils semblent voués à des lendemains meilleurs, grâce à la mesure prise par le Ministère des Finances en août 2005, obligeant les fonctionnaires dont le salaire est supérieur à 100.000 FCFA à ouvrir chacun un compte pour désormais recevoir leur salaire par virement bancaire.

Les comptes courants (CCC) viennent en 3^e position dans le parc des comptes de la FIRST BANK. L'ouverture d'un tel compte nécessite un minimum de 1.500.000 FCFA. Ils ne sont accessibles qu'aux clients dont le salaire mensuel est d'au moins 500.000 FCFA ou à toute personne physique ou morale pouvant justifier d'un chiffre d'affaires minimum de 5.000.000 FCFA par mois. Ce type de compte peut être débiteur : ses clients sont les principaux consommateurs de l'argent issu des comptes d'épargne. Pourtant, **le ratio comptes courants / comptes d'épargne ouverts à la FIRST BANK est passé d'un compte courant pour six comptes d'épargne en 1999 à un compte courant pour sept comptes d'épargne en 2004**. Cette évolution de ce ratio laisse voir la prédominance au cours du temps des comptes d'épargne sur les comptes courants. Ceci interpelle la FIRST BANK à prendre des dispositions adéquates pour y faire face dans l'avenir.

Cette description des proportions des produits sur le réseau national retrace la réalité dans toutes les agences où le classement est le même, sauf à Kousseri où les CCC rivalisent avec les CDS (voir détails annexe).

Les autres types de compte, du fait que la clientèle soit ciblée d'avance, fonctionnent très différemment d'une localité à l'autre. Ainsi, si les comptes de dépôt islamique (CDI) ont une part non négligeable dans les activités des agences situées dans le grand nord, les comptes Flash Cash sont beaucoup plus sollicités dans les grandes villes que sont Yaoundé et Douala. Ces autres types de comptes ne seront pas pris en compte dans notre étude.

II. PART DE CHAQUE AGENCE DANS LE PARC DES COMPTES

2.1 Comptes d'épargne

Yaoundé détient à lui seul plus de la moitié des comptes d'épargne, suivie de Douala dont la proportion est inférieure au quart du parc de ce type de compte, mais qui dépasse toutes les autres agences réunies.

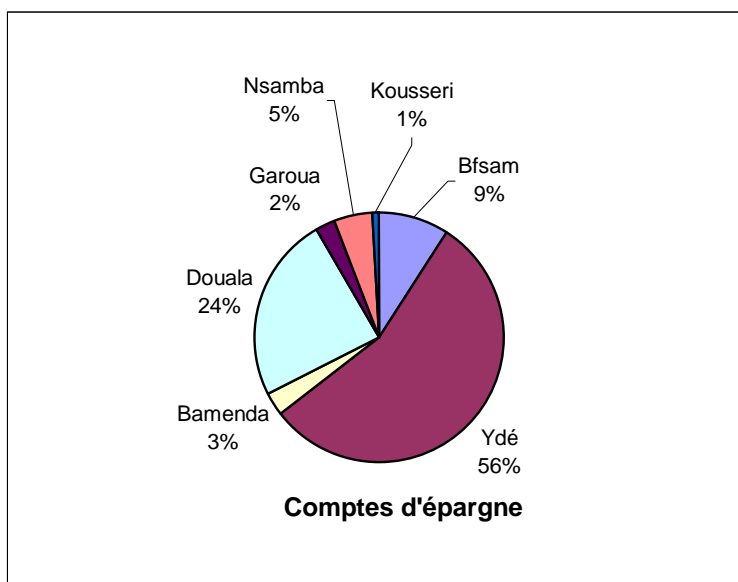


Figure 2 : Pourcentages des différentes agences dans le parc des comptes d'épargne.

2.2 Comptes CDS

Le parc des comptes CDS est également dominé par Yaoundé. Bien que les chiffres mensuels soient très inférieurs en valeur absolue aux comptes d'épargne, ce produit est très apprécié à Yaoundé, siège des institutions publiques et par ricochet, des marchés publics. Notons que toute entreprise désirant soumissionner aux marchés publics se doit d'avoir un compte bancaire qui donne droit au virement. Ainsi, les comptes de dépôt spécial constituent la solution idéale pour les petites et moyennes entreprises en quête des marchés publics. Car ce type de compte est intermédiaire entre les comptes d'épargne qui ne donnent pas droit au virement et les comptes courants très coûteux voire inaccessible pour elles.

Ce type de compte est moins sollicité à Douala du fait que les activités économiques de cette ville soient fortement dominées par le secteur informel dont les acteurs se contentent

bien des comptes d'épargne. Douala se partage ainsi moins de 30% du stock des comptes CDS avec les autres villes. Notons toutefois que c'est le type de compte qui a la durée de vie la moins élevée, car il est souvent sollicité par les clients pour résoudre des problèmes ponctuels, comme l'inscription dans les Universités étrangères et sont clôturés aussitôt leur problème résolu.

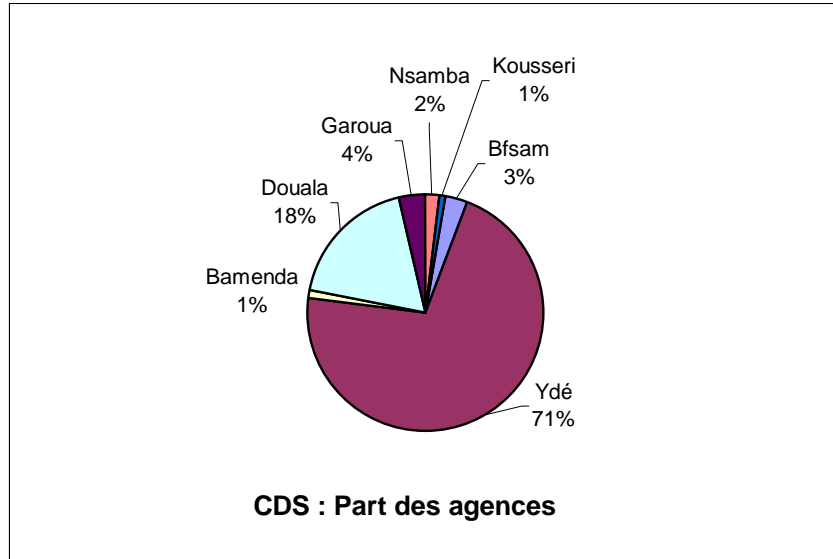


Figure 3 : Pourcentages des différentes agences dans le parc des comptes de dépôt spécial.

2.3 Comptes courants

Ce type de compte a l'avantage, entre autres, qu'il donne droit à un chéquier et qu'il peut être débiteur. Mais, du fait qu'il soit onéreux et sélectif, ils ne sont accessibles qu'aux personnes économiquement nanties et ils sont beaucoup plus sollicités par les grandes entreprises. Il n'est donc pas étonnant que Douala soit en tête du parc de ce compte, talonné par Yaoundé.

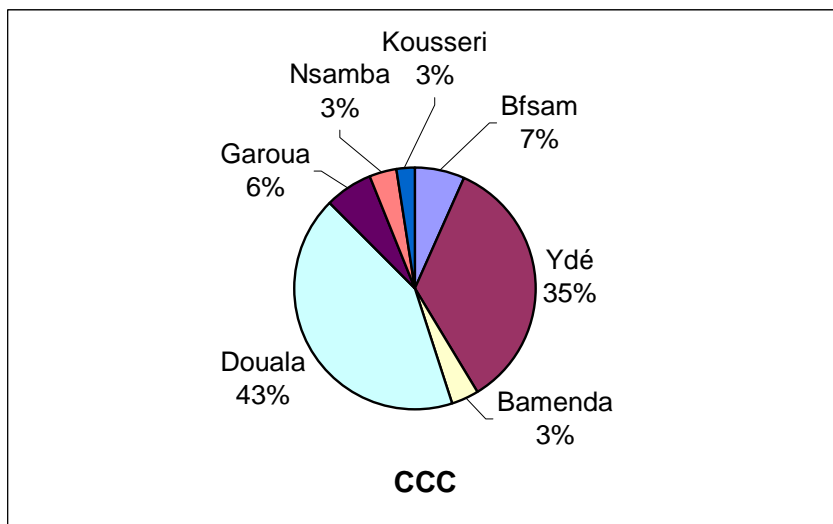


Figure 4 : Pourcentages des différentes agences dans le parc des comptes courants chèques.

III. EVOLUTION DES OUVERTURES DES COMPTES (juillet 1998 à juin 2005)

3.1 Evolution mensuelle

3.1.1 Comptes d'épargne

Agences	Minimum enregistré	1 ^{er} Quartile.	Moyenne mensuelle	3 ^e Quartile	Maximum enregistré
Bamenda	07	13.75	17.18	20	30
Garoua	05	11	13.86	15.25	27
Kousseri	01	05	7.232	09	15
Nkongsamba	08	18.25	27.80	34	66
Bafoussam	10	42.50	51.89	62.25	109
Douala	52	113.3	137.5	172.3	208
Yaoundé	100	194.3	312	372	668
Réseau nat.	230	412.8	565	661.5	1092

Le tableau précédent présente les moyennes mensuelles d'ouverture des comptes d'épargne dans les agences pendant notre période d'observation. Pour en avoir une idée nette, nous leur avons joint les minimums et les maximums déjà atteints dans chaque agence, ainsi que le 1^{er} et le 3^e quartiles, qui traduisent l'intervalle dans lequel se concentrent les 50% des chiffres mensuels d'ouverture des comptes dans ces agences. Le même travail fait sur les deux autres types de comptes est présenté ci-dessous : nous pouvons y remarquer qu'à Bafoussam, Nkongsamba, Kousseri et Bamenda, il se passe parfois des mois sans qu'aucun compte CDS ni CCC y soient ouverts.

3.1.2 Comptes CDS

Agences	Minimum enregistré	1 ^{er} Quartile.	Moyenne mensuelle	3 ^e Quartile	Maximum enregistré
Bamenda	00	01	02.76	03.25	11
Garoua	02	07	09.64	012	21
Kousseri	00	01	02.62	04	08
Nkongsamba	00	03	04.9	06	15
Bafoussam	00	04.75	07.96	10.25	28
Douala	28	38	45.51	51	83
Yaoundé	109	149.8	178.6	199.3	331
Réseau nat.	173	218.3	251.2	275.3	443

3.1.3 Comptes Courants Chèques

Agences	Minimum enregistré	1 ^{er} Quartile.	Moyenne mensuelle	3 ^e Quartile	Maximum enregistré
Bamenda	00	02	03.1	04	09
Garoua	01	04	05.81	08	14
Kousseri	00	01	03.46	04.25	19
Nkongsamba	00	02	03.36	04	14
Bafoussam	00	04	05.94	08	12
Douala	28	38	45.51	51	83
Yaoundé	09	22.75	31.32	39	65
Réseau nat.	38	70.75	89.8	106.3	150

3.2 Evolution annuelle

Nous présentons sur les histogrammes en annexe, l'évolution annuelle des ouvertures des comptes dans les agences lors de six dernières années ainsi que leurs taux de croissance. Au vu de ces graphiques, il se dégage les commentaires suivants :

3.2.1 Comptes d'épargne

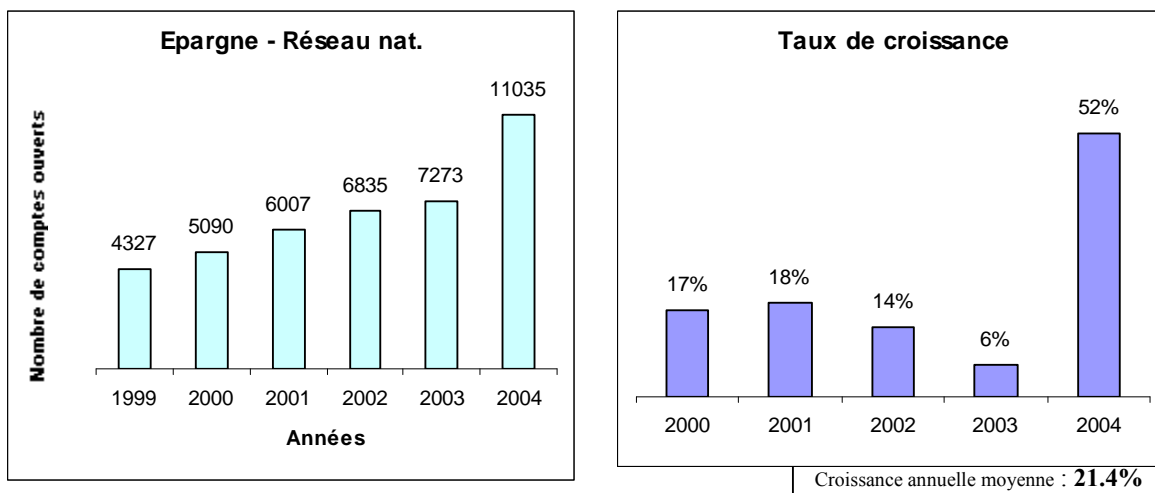


Figure 5 : Nombre des comptes d'épargne ouverts par an et taux de croissance.

Nous constatons une croissance continue sur l'ensemble du réseau national, de 1999 à 2004, qui ne traduit cependant pas la réalité dans les différentes agences. Car cette tendance continue à la hausse est largement influencée par Yaoundé, qui est la seule ville qui enregistre une croissance sans relâche depuis 1999.

Les autres agences, en dehors de Bafoussam, ont connu une baisse remarquable des activités au cours de l'année 2003 qui se traduit par une décroissance des ouvertures de comptes d'épargne, conséquence certaine d'un phénomène conjoncturel de l'économie camerounaise, peu ressenti cependant à Yaoundé et à Bafoussam. Cette baisse peut très bien être la conséquence de la diminution des parts de marché de la FIRST BANK par rapport aux entreprises concurrentes, nous pensons ainsi aux coopératives d'épargne qui ont connu un essor remarquable en 2003 et à l'expansion de la CBC (Commercial Bank of Cameroon), une banque concurrente qui vise le même type de clientèle que la FIRST BANK.

Ces effets ont très tôt disparu dans les activités de la banque au cours de l'année 2004 qui enregistre des chiffres records d'ouverture des comptes d'épargne dans l'ensemble des agences, avec un taux de croissance spectaculaire de 52% sur l'année précédente. Nous pouvons attribuer cette hausse brusque d'une part aux leçons tirées des mauvais résultats de l'année précédente, mais surtout aux problèmes internes qui ont ébranlé la Caisse d'épargne postale du Cameroun, et qui a certainement conduit ses anciens clients vers d'autres banques dont la FIRST BANK. Nous ne négligeons pas la perte de confiance aux coopératives d'épargne due au manque de sérieux de la part de leurs promoteurs, ayant même conduit à la fermeture de certaines d'entre elles ; Ceci aurait obliger les clients déçus à se diriger vers les autres banques.

Comptes CDS

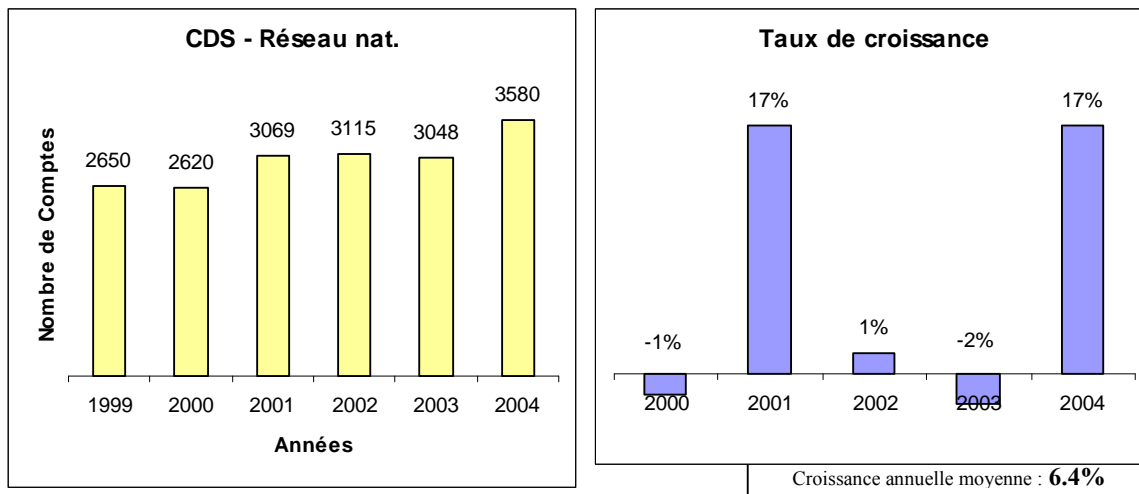


Figure 6 : Nombre des comptes CDS ouverts par an et taux de croissance.

Leur comportement varie très différemment d'une ville à l'autre. Ainsi, si Yaoundé, une fois de plus, enregistre une hausse continue depuis 1999, ce n'est pas le cas pour l'agence de Douala qui affiche un comportement stagnant au cours du temps (en six ans, il est resté entre 535 et 655 comptes par an). L'effet conjoncturel de 2003 se fait encore ressentir ici.

Concernant les autres agences (qui représentent moins de 12% de ce produit), quand les chiffres ne stagnent pas, l'on constate tout simplement une chute d'effectif, laissant croire que ce produit a déjà atteint sa maturité dans ces villes, on dirait qu'il est en voie de disparition dans ces villes. Ceci nécessite donc une révision des conditions pour y accéder ou bien une politique marketing efficace pour le booster dans les localités les plus concernées que sont Bamenda, Garoua et Kousseri.

Dans l'ensemble, malgré les années 2000 et 2003 qui enregistrent des taux de croissance négatifs, la croissance moyenne reste supérieure à 6%.

Comptes Courants

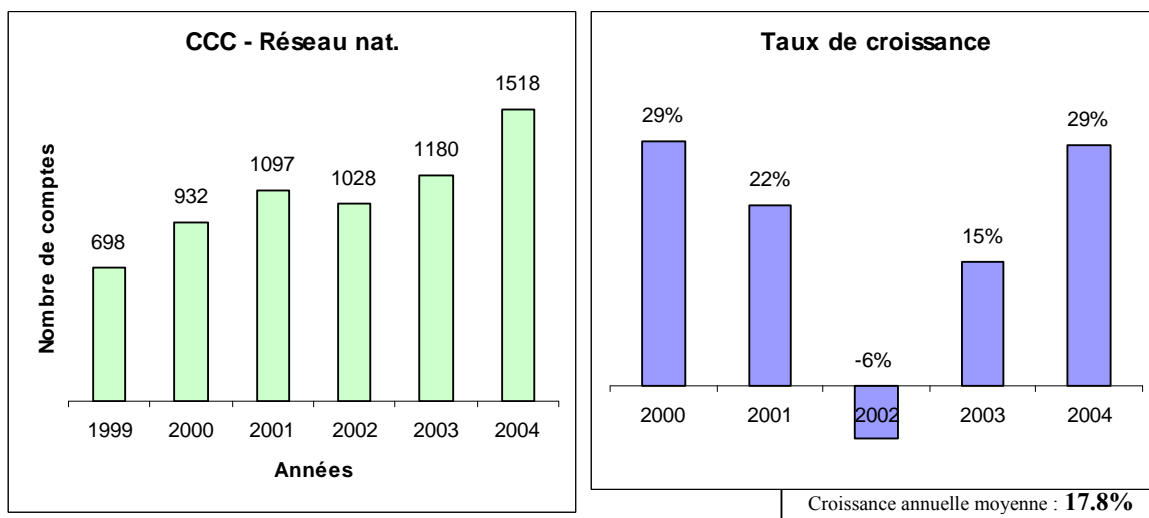


Figure 7 : Nombre des comptes CCC ouverts par an et taux de croissance.

Douala est la ville où ce produit se porte le mieux. Son évolution a connu une forte croissance entre 1999 et 2001, atteignant 63% de taux de croissance en 2000. Douala a cependant connu des moments difficiles entre 2002 et 2003, avec un taux de croissance négatif (-14%) en 2002. L'année 2004 quant à elle enregistre un bon comportement qui, sans aucun doute, est le résultat d'une politique interne à la FIRST BANK, notamment l'ouverture de l'agence de Mboppi, tout près du célèbre marché de ce quartier, qui lui aurait permis de conquérir d'avantage la clientèle. Sur le plan économique, même la relance annoncée ces dernières années ne suffit pas pour expliquer cette croissance de 50% constatée au niveau de l'ouverture des comptes courants à la FIRST BANK à Douala en 2004.

Yaoundé, bien que relégué en second rang dans le parc de ce produit, présente un comportement en hausse continue d'année en année. Le taux de croissance minimal enregistré ici est de 7% en 2001 et le maximal de 37% en 2003.

Pour ce qui concerne les autres agences, nous pouvons les classer en deux groupes. Le premier est constitué de Garoua et de Bafoussam dont le nombre d'ouverture de comptes annuel varie entre 56 et 90. Ces deux agences n'arrivent plus à atteindre leur niveau record respectif des années 2000 et 2001. Nkongsamba, Bamenda et Kousseri ont des chiffres annuels de moyenne inférieure à 45. Si Nkongsamba et Kousseri semblent avoir atteint leur maturité, Bamenda, avec son excellent résultat de 2004 (221% de taux de croissance) ne nous permet pas d'en tirer la même conclusion.

Dans l'ensemble, en dehors de l'année 2002 qui a un taux de croissance négatif, ce produit enregistre un comportement appréciable avec des taux de croissance régulièrement au dessus de 20%. Cependant, jusqu'à présent, son maximum annuel atteint sur le réseau national est de 1518 comptes ouverts, contre un minimum de 1749 comptes CDS à l'agence de Yaoundé. Les détails pour chaque agence sont présentés en annexe.

IV. EFFETS SAISONNIERS (évolution mensuelle moyenne des produits)

Il est important, voire indispensable, pour toute entreprise de connaître ses périodes annuelles d'intense activité ainsi que les saisons mortes. A cet effet, nous avons estimé les moyennes mensuelles d'ouverture des comptes pour avoir une idée de son évolution mensuelle moyenne d'année en année. Leurs représentations graphiques sont présentées et commentées ci-dessous. N'ayant pas suffisamment d'informations pour les expliquer de manière adéquate, nous nous limiterons juste à les décrire. Les détails agence par agence sont présentés en annexe.

4.1 Comptes d'épargne

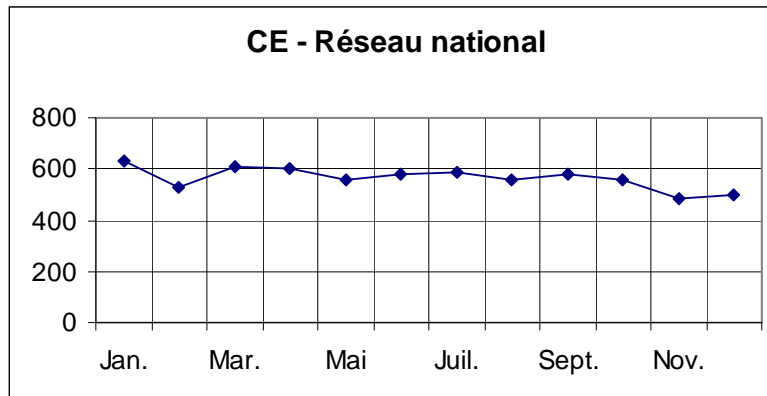


Figure 8 : Evolution mensuelle moyenne des comptes d'épargne.

Nous constatons que dans les grandes villes que sont Yaoundé et Douala, janvier est le mois où l'on épargne le plus, suivie d'une baisse relative en février, qui serait due au fait qu'il est le mois le plus court de l'année. Sinon, remarquons que l'ouverture des comptes d'épargne reste quasi constante de mars à septembre où commence une baisse continue jusqu'en décembre, le mois qui enregistre le niveau le plus bas. La baisse constatée en septembre est due à la rentrée scolaire et ses corollaires, tandis que celle de décembre peut être attribuée aux fêtes de fin d'année.

Dans les autres agences, les ouvertures des comptes sont grandement influencées par les activités économiques spécifiques à chaque ville. Le mois de juillet se distingue positivement des autres mois. C'est ainsi par exemple, qu'à l'opposée des villes du grand Sud (Bafoussam, Nkongsamba,...), les fin d'années sont les périodes d'ouverture maximale des comptes dans le grand Nord (Garoua et Kousseri) : Synonyme des retombés de la vente du bétail et autres produits dérivés pendant les fêtes de fin d'année.

4.2 Comptes CDS

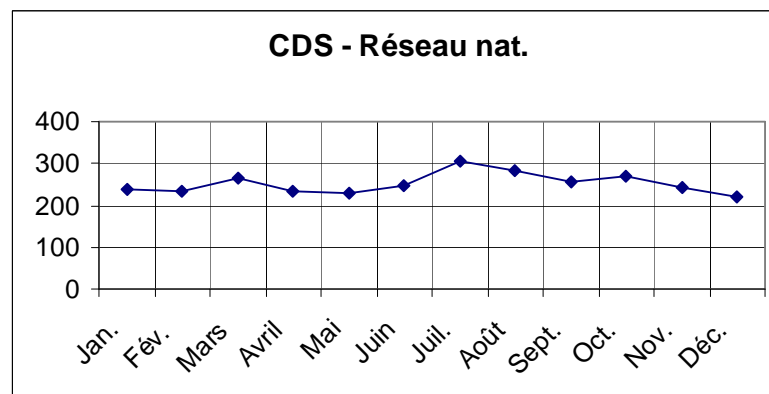


Figure 9 : Evolution mensuelle moyenne des CDS.

Visiblement, les comptes CDS sont sollicités à des fins très différentes d'une localité à l'autre. Ceci se traduit par une variabilité des effets saisonniers d'une ville à l'autre. Mais dans l'ensemble, les mois de juillet à octobre constituent la période de forte sollicitation, bien

que cette période commence un peu plus tôt ou un peu plus tardivement dans certaines villes. Cette période coïncide bien avec les inscriptions dans les universités ; ce qui confirme le fait que ce compte est très sollicité par les étudiants pour leur inscription dans les universités étrangères. C'est sans aucun doute l'une des raisons pour lesquelles Yaoundé, la principale ville estudiantine du Cameroun domine sur les autres villes.

Les mois d'avril et mai constituent dans l'ensemble, la période dite « morte » où ce produit est le moins sollicité.

4.3 Comptes Courants

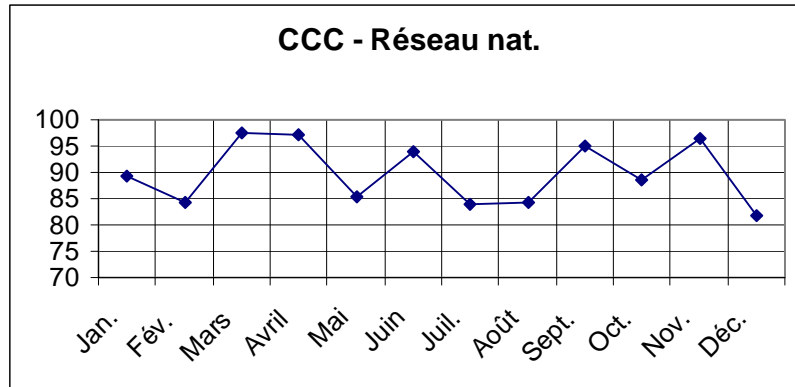


Figure 10 : Evolution mensuelle moyenne des comptes courants.

Ils se comportent de manière très similaire à Douala et à Yaoundé où les mois de juillet pour Yaoundé et août pour Douala sont les périodes où ce type de compte est le moins sollicité. Les mois de Mars à juin, puis de septembre à novembre semblent par contre favorables pour ce produit dans ces deux villes.

Les autres villes affichent des comportements différents, mais du fait des moyennes mensuelles encore faibles, il nous est difficile de ressortir leur saisonnalité, bien que leur graphique laisse voir des périodes de forte activité entrecoupées de saisons mortes.

CHAPITRE III :

**DESCRIPTION DE LA DEMARCHE DE BOX
ET JENKINS POUR LA PREVISION D'UNE
SERIE CHRONOLOGIQUE**

George BOX et **Gwilym JENKINS** sont deux statisticiens qui ont contribué, dans les années 1970, à populariser la théorie des séries temporelles univariées. Les procédures de modélisation sont présentées dans leur célèbre ouvrage intitulé « Time Series Analysis : Forecasting and control »

Ils ont proposé une démarche générale de prévision pour les séries chronologiques [1]. Cette démarche est fondée sur la notion de processus ARMA et elle comprend quatre (04) phases : l'identification a priori, l'estimation du modèle ARMA identifié, l'identification a posteriori et la prévision.

I. IDENTIFICATION A PRIORI

I.1 NOTIONS DE BASE SUR LES SERIES CHRONOLOGIQUES

I.1.1 GENERALITES SUR LES SERIES CHRONOLOGIQUES

a) Définition

On appelle *série chronologique* ou *chronique* toute suite d'observations $(X_t : t \in T)$ indexées par un ensemble ordonné T (le « temps »).

Les différents types des séries chronologiques sont [3] :

Série continue – Dans certains domaines, la variable X_t peut être observée de façon continue, *i.e.* l'indice t peut prendre toutes les valeurs dans un intervalle de nombres réels. Dans un tel cas, on parle de *série continue*. (De telles situations sont rares dans les données économiques)

Série discrète – Une série est *discrète* lorsque l'ensemble des valeurs possibles de t est un ensemble discret, *i.e.* T peut être considéré comme un sous-ensemble des nombres entiers.

Il y a deux types importants de série discrète, suivant que les observations correspondent à des **niveaux** : séries enregistrées instantanément (*e.g.*, prix, stocks), ou à des **flux** : séries cumulées sur un intervalle de temps (*e.g.*, revenu, consommation).

Lorsqu'on analyse une série de **flux**, il est important de tenir compte de l'intervalle de temps considéré.

b) Objectifs de l'analyse d'une chronique

Les objectifs de l'analyse d'une chronique sont :

- **Modélisation**

Elle consiste à :

- Développer des modèles permettant de décrire le comportement d'une ou plusieurs séries chronologiques.
- Mettre au point une méthodologie pour spécifier, estimer, valider (juger) un modèle approprié pour des données particulières.

- **Prévision**

Étant données des observations X_1, \dots, X_T , la prévision consiste à évaluer une valeur non observée, X_{T+h} . La prévision peut être ponctuelle, ou prendre la forme d'un intervalle de prévision.

I.1.2 SCHEMA DE DECOMPOSITION D'UNE SERIE CHRONOLOGIQUE

L'étude de toute série chronologique passe avant tout par sa décomposition, car l'on suppose qu'elle est constituée d'une partie prévisible (la tendance), d'une partie due aux variations saisonnières (effets saisonniers) et d'une partie non prévisible (le résidu).

Les problèmes de décomposition les plus fréquents sont les suivants :

- estimer et enlever la tendance (« *stationnarisation* »);
- estimer et enlever les variations saisonnières (« *déssaisonnalisation* »).

a) La Tendance

La composante fondamentale ou tendance (trend) traduit l'évolution à moyen terme du phénomène. On parle aussi de mouvement conjoncturel ou mouvement extra-saisonnier. La chronique correspondante, notée f_t , $t = 1 \dots T$, est une fonction à variation lente. Elle est le plus souvent estimée sous forme paramétrique (polynôme, exponentielle,...) ou comme le résultat d'une opération de lissage.

b) La composante saisonnière

La composante saisonnière ou mouvement saisonnier représente des effets périodiques de période connue p qui se reproduisent de façon plus ou moins identique d'une période sur l'autre. La chronique correspondante est notée S_t , ($t = 1..T$). Elle est généralement supposée rigoureusement périodique : $S_{t+p} = S_t$, et les valeurs $S_j = (S_{ij})$, $j = 1 \dots p$ d'une période sont appelées "coefficients saisonniers". Le bilan de l'effet saisonnier sur une période doit être nul car il est pris en compte dans la tendance. La composante saisonnière permet simplement de distinguer à l'intérieur d'une même période une répartition stable dans le temps d'effets positifs ou négatifs qui se compensent sur l'ensemble de la période.

c) La composante résiduelle

La composante résiduelle ou variations accidentelles est la partie non structurée du phénomène. Elle est modélisée par une suite de variables aléatoires ε_t , ($t = 1, \dots, T$) centrées, non corrélées et de même variance, on parle de bruit blanc.

Certains phénomènes économiques étudiés à très long terme présentent une composante cyclique (cycles d'activité) dont la période, de plusieurs années, est souvent mal définie. Cette composante est prise en compte dans la tendance sur les séries de taille moyenne.

d) Schémas de décomposition d'une chronique

Pour pouvoir séparer les trois composantes servant à décrire la série observée, il est nécessaire de préciser leur mode d'interaction. La plupart des séries chronologiques entrent dans l'un des schémas suivants :

- Schéma additif : $X_t = f_t + S_t + \varepsilon_t$,
- Schéma multiplicatif : $X_t = f_t \times S_t \times (1 + \varepsilon_t)$,
- Schéma mixte : $X_t = f_t \times S_t + \varepsilon_t$.

En utilisant $(I+\varepsilon_t)$ dans le cas multiplicatif, on conserve la même signification et les mêmes propriétés à chacune des trois composantes f_t , S_t et ε_t dans les trois schémas de composition. Cependant il est nécessaire de supposer que $(I+\varepsilon_t)$ reste positif dans le modèle multiplicatif car la composante résiduelle ne peut être responsable du signe de la grandeur observée.

e) Choix du schéma de décomposition

Dans les schémas multiplicatif et mixte, **les oscillations dues à l'effet saisonnier ont une amplitude proportionnelle à la valeur de la tendance**. C'est précisément l'argument utilisé pour faire le choix entre le schéma additif et les deux autres schémas au vu de la représentation graphique de la chronique [4].

La distinction entre le schéma multiplicatif et le schéma mixte peut également s'apprécier graphiquement selon le même principe. Elle peut aussi relever de considérations sur l'origine des erreurs : une erreur structurelle (de modélisation) a des chances d'être proportionnelle à la grandeur étudiée alors qu'une erreur de mesure pourrait ne pas en dépendre.

Notons que le schéma multiplicatif peut se ramener à un schéma additif en considérant le logarithme de la variable à étudier. Il peut être nécessaire d'appliquer plusieurs fois la transformation logarithmique. Ainsi, par la suite nous ne considérerons plus que le cas additif.

I.2 ANALYSE DE LA SAISONNALITE

L'analyse de la saisonnalité permet de détecter les effets saisonniers et leur périodicité pour pouvoir les enlever de la série : c'est la « *déssaisonnalisation* ». Il existe plusieurs méthodes de déssaisonnalisation.

I.2.1 METHODES DE DESSAISONALISATION [6]

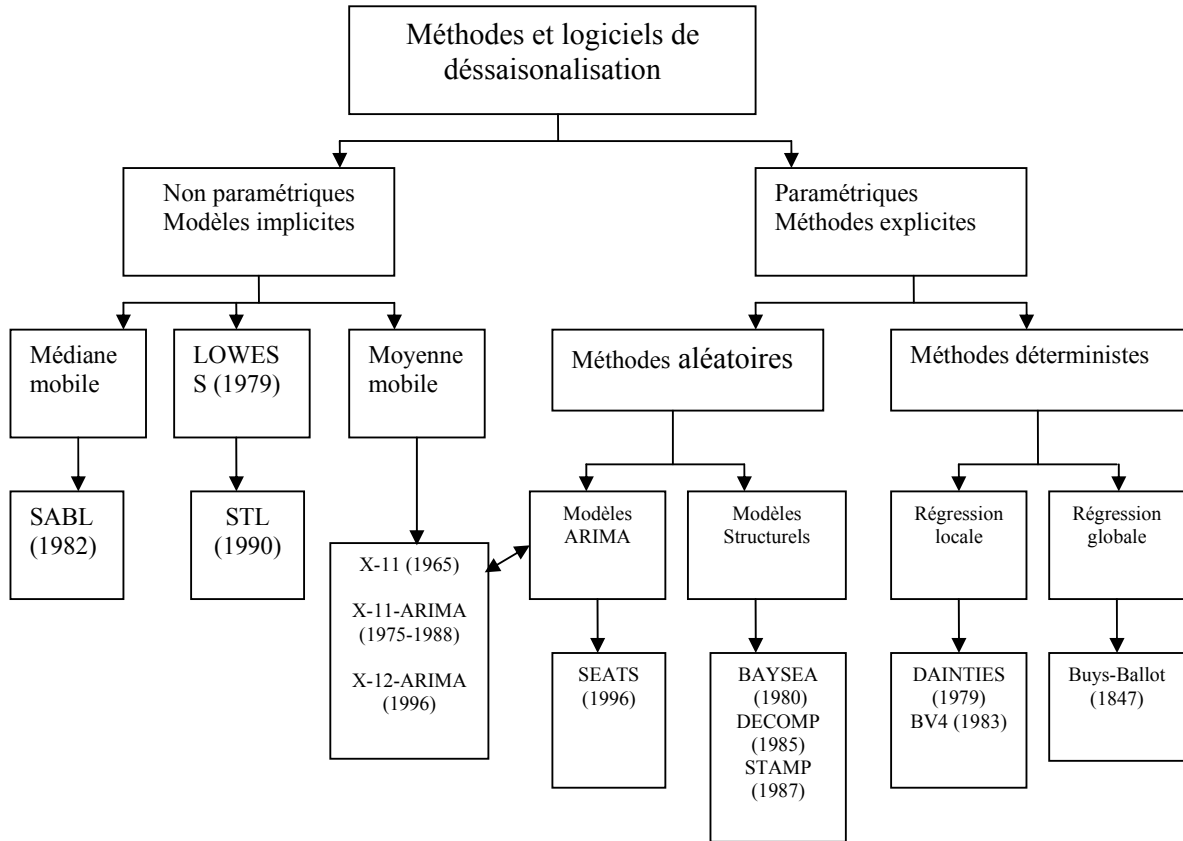


Figure 11 : Différents méthodes de déssaisonnalisation.

I.2.2 DESSAISONALISATION PAR REGRESSION

Les données sur lesquelles nous avons travaillé ont des tendances que nous pouvons considérer comme linéaires et exhibent une relative saisonnalité annuelle. Ainsi, dans cette étude, nous utiliserons une méthode de déssaisonnalisation par régression : celle de **Buys-Ballot** [4]. Elle fournit un schéma additif simple que l'on peut traiter très complètement par des méthodes élémentaires. La tendance est représentée par une droite, l'effet saisonnier est rigoureusement périodique de période p connue et la partie résiduelle est une suite de variables aléatoires indépendantes identiquement distribuées de loi normale centrée et de variance σ^2 :

$$X_t = \alpha t + \beta + S_t + \varepsilon_t, \quad t = 1, \dots, T; \quad S_t = S_{t+p}; \quad \varepsilon_t \sim i.i.d.N(0; \sigma^2).$$

La simplicité des calculs est obtenue en supposant que la série est observée pendant n "années" de p "mois" :

$$X_{ij} = \alpha[p(i-1) + j] + \beta_j + \varepsilon_{ij}, \quad \beta_j = \beta + S_j, \quad j = 1, \dots, p, \quad i = 1, \dots, n; \quad \sum_{j=1}^p S_j = 0. \quad (1)$$

Avec la convention, $t = (i, j) = (\text{année}, \text{mois})$.

Les coefficients saisonniers $S_j, j = 1, \dots, p$, caractérisent la composante périodique S_t dont l'effet annuel moyen est nul. Ainsi, la partie déterministe du modèle est décrite par $p+1$ paramètres linéairement indépendants : $\alpha, \beta_1, \dots, \beta_p$. La méthode des moindres carrés consiste à chercher, parmi les chroniques $x_{ij}(a, b_1, \dots, b_p) = a[p(i-1) + j] + b_j$, composées d'une tendance linéaire et d'un mouvement saisonnier périodique, celle qui est la plus proche de l'observation selon le critère :

$$\min_{a, b_1, \dots, b_p} \frac{1}{np} \sum_{i=1}^n \sum_{j=1}^p [X_{ij} - x_{ij}(a, b_1, \dots, b_p)]^2. \quad (2)$$

En d'autres termes, elle retient les valeurs des paramètres pour lesquels la moyenne des carrés des erreurs observées est minimum. Notons plus simplement $x_{ij} = x_{ij}(a, b_1, \dots, b_p)$ et introduisons les moyennes suivantes :

$$\bar{X}_{.j} = \frac{1}{n} \sum_{i=1}^n X_{ij}, \bar{x}_{.j} = \frac{1}{n} \sum_{i=1}^n x_{ij} = a[p(n-1)/2 + j] + b_j, \bar{X}_{..} = \frac{1}{np} \sum_{i,j} X_{ij}, i=1, \dots, n; j=1, \dots, p$$

La résolution du problème (2), sous les contraintes (1) permet d'obtenir les estimations suivantes :

$$\hat{\alpha} = \frac{12}{np(n^2-1)} \left[\sum_{i=1}^n i \bar{X}_{.i} - \frac{n(n+1)}{2} \bar{X}_{..} \right].$$

$$\hat{\beta}_j = \bar{X}_{.j} - \hat{\alpha}[p(n-1)/2 + j], j = 1, \dots, p,$$

$$\hat{\beta} = \frac{1}{p} \sum_{j=1}^p \hat{\beta}_j = \bar{X}_{..} - \hat{\alpha}_{..} = \bar{X}_{..} - \hat{\alpha}(np+1)/2$$

$$\hat{S}_j = \hat{\beta}_j - \hat{\beta} = \bar{X}_{.j} - \bar{X}_{..} - \hat{\alpha}[j - (p+1)/2], j = 1, \dots, p.$$

La désaisonnalisation de la série chronologique initiale s'obtient alors en retranchant à chaque X_{ij} , le coefficient saisonnier correspondant S_j . Ce qui donne la série corrigée des variations saisonnières (CVS).

I.3 STATIONNARITE DE LA SERIE

I.3.1 OBJECTIF DE LA STATIONNARISATION

Pour travailler avec des données temporelles, elles doivent conserver une distribution constante dans le temps. C'est le concept de *stationnarité*. Ainsi, si nos variables passées sont semblables à nos variables futures, on peut utiliser le passé pour tenter de prédire le futur. Si nos données ne sont pas stationnaires, on peut se retrouver avec: un **biais de prévision**, une **prévision inefficace** et par conséquent une **mauvaise inférence** [2].

Un concept de stationnarité généralement utilisé est celui de la stationnarité de second ordre. Nous dirons qu'une série chronologique X_t est stationnaire au second ordre si les trois conditions suivantes sont vérifiées :

$$\triangleright E[X_t] = \mu \text{ (l'espérance ne dépend pas de } t)$$

$$\triangleright \text{var}[X_t] = \sigma^2 \text{ (la variance ne dépend pas de } t)$$

$$\triangleright \text{cov}[X_t, X_s] = \gamma_k, k = t-s \text{ (la covariance ne dépend que de } t-s)$$

I.3.2 TEST DE DICKEY-FULLER

a) Principe du test

Il permet de tester la stationnarité en différence (hypothèse H0) contre la stationnarité en tendance (hypothèse H1) sur la série. Trois tests, sont effectués :

Tests	Hypothèse H0	Hypothèse H1
Test N°1	$X_t = X_{t-1} + \varepsilon_t$	$X_t = \alpha X_{t-1} + \varepsilon_t$ (où $ \alpha < 1$)
Test N°2	$X_t = X_{t-1} + \beta + \varepsilon_t$ (avec $\beta \neq 0$)	$X_t = \alpha X_{t-1} + \beta + \varepsilon_t$ (avec $ \alpha < 1$ et $\beta \neq 0$)
Test N°3	$X_t = X_{t-1} + \beta + \gamma.t + \varepsilon_t$ (avec $\gamma \neq 0$)	$X_t = \alpha X_{t-1} + \beta + \gamma.t + \varepsilon_t$ (avec $ \alpha < 1$ et $\gamma \neq 0$)

Les trois modèles alternatifs sont réécrits sous des formes algébriquement équivalentes :

- (1') $X_t - X_{t-1} = (\alpha - 1)X_{t-1} + \varepsilon_t$
- (2') $X_t - X_{t-1} = (\alpha - 1)X_{t-1} + \beta + \varepsilon_t$
- (3') $X_t - X_{t-1} = (\alpha - 1)X_{t-1} + \beta + \gamma.t + \varepsilon_t$

Soit encore :

- (1'') $\Delta X_t = \Phi X_{t-1} + \varepsilon_t$
- (2'') $\Delta X_t = \Phi X_{t-1} + \beta + \varepsilon_t$
- (3'') $\Delta X_t = \Phi X_{t-1} + \beta + \gamma.t + \varepsilon_t$

Avec $\Phi = (\alpha - 1)$ et Δ l'opérateur *différences-premières*. Il est clair qu'il est équivalent de tester l'hypothèse : $\alpha = 1$, dans les modèles sous la forme initiale, ou $\Phi = 0$ dans leur version transformée (hypothèse H0).

On estime par les MCO le modèle (3'') et on calcule la statistique de Student pour les valeurs estimées de γ , de β et de Φ .

Le test pour les coefficients γ et β est bilatéral, en ce qui concerne Φ , les choses diffèrent du test de Student traditionnel : ayant exclu le cas explosif ($\alpha > 1$ ou $\Phi > 0$) on procède au test unilatéral ($\alpha = 1$ ou $\Phi = 0$ contre $\Phi < 0$) ce qui donne une zone de rejet située du côté négatif, et les hypothèses des MCO n'étant pas satisfaites, notamment du fait que la variable considérée est endogène, on utilise une table particulière calculée par Dickey et Fuller (voir annexe).

En cas de rejet total ou partiel du test du modèle (3'') avec $\Phi = 0$, opéré comme indiqué, on teste de même le modèle (2'') et $\Phi = 0$ avec la table voulue, puis encore si nécessaire le modèle (1'') avec $\Phi = 0$.

b) Processus TS et DS

Un processus X_t est dit *difference-stationary* (DS) ou stationnaire en différence s'il peut s'écrire sous l'une des formes suivante :

- (1) $X_t = X_{t-1} + \varepsilon_t$
- (2) $X_t = X_{t-1} + \beta + \varepsilon_t$ (avec $\beta \neq 0$)
- (3) $X_t = X_{t-1} + \beta + \gamma.t + \varepsilon_t$ (avec $\gamma \neq 0$)

Il sera dit *Trend-stationary* (TS) ou stationnaire en tendance s'il peut se mettre sous l'une des formes suivantes :

- (4) $X_t = \alpha X_{t-1} + \beta + \varepsilon_t$ (avec $|\alpha| < 1$ et $\beta \neq 0$)
- (5) $X_t = \alpha X_{t-1} + \beta + \gamma.t + \varepsilon_t$ (avec $|\alpha| < 1$ et $\gamma \neq 0$)

La bonne manière de stationnariser une série TS consiste à estimer, en général par les moindres carrés ordinaires (MCO), l'expression de la tendance et à la retirer. Tandis que la stationnarisation des séries DS se fait par passage aux différences.

c) Estimation par Moindres carrés ordinaires (MCO)

Soit X_t , une série chronologique ayant une tendance linéaire, c'est-à-dire qu'elle peut se mettre sous la forme :

$$X_t = a + bt + \varepsilon_t \quad (t = 1, \dots, n)$$

L'estimation de la tendance par moindres carrés ordinaires consiste à déterminer les coefficients a et b qui minimisent la somme des carrés résiduels :

$$SCR(\Omega) = \sum_{i=1}^n (X_i - a - bt_i)^2$$

On montre que

$$\hat{b} = \frac{\text{cov}(X_t, t)}{\text{var}(X_t)} \quad \text{et} \quad \hat{a} = \bar{X}_t - \hat{b}\bar{t}$$

Généralisation [9] :

$$\text{Soit } Y = X\beta + \varepsilon, \quad Y \in R^n, \quad X \in M(n, p),$$

$$\beta \in R^p, \quad p \text{ paramètres inconnus}$$

$$\varepsilon = (\varepsilon_i), \quad i = 1, \dots, n, \quad (\varepsilon_i, \text{iid} \sim N(0, \sigma^2))$$

$$\text{alors } \hat{\beta} = ({}^tXX)^{-1} {}^tXY \sim N_p(\beta, \sigma^2 ({}^tXX)^{-1})$$

Remarque : L'estimateur des MCO est unique, sans biais et de moindre variance parmi les estimateurs linéaires.

d) Test de Student

$$(H0) : {}^t a \cdot \beta = b, \quad (a \in R^p \text{ et } b \in R \text{ connus})$$

$$\text{contre } (H1) : {}^t a \cdot \beta \neq b$$

$${}^t a \cdot \beta - b \underset{(H0)}{\sim} N(0, \sigma^2 \cdot {}^t a \cdot ({}^tXX)^{-1} a)$$

La statistique de test est donnée par :

$$T = \frac{{}^t a \cdot \beta - b}{S \sqrt{{}^t a \cdot ({}^tXX)^{-1} a}} \underset{(H0)}{\sim} T_{n-p}$$

$$\text{Avec } S^2 = \frac{1}{n-p} \|Y - X\hat{\beta}\|^2 \sim \frac{\sigma^2}{n-p} \chi_{n-p}^2$$

Règle de décision : on rejette $(H0)$ si $T > t(\alpha, n-p)$ (où $t(\alpha, n-p)$ est le point critique de Student au seuil α , à $n-p$ degrés de liberté)

I.4 DETERMINATION DES ORDRES p ET q DU PROCESSUS ARMA

I.4.1 LE CORRELOGRAMME

L'autocorrélation :

Les observations d'une série chronologique n'étant pas indépendantes, le corrélogramme permet de déceler les liens entre elles. Il s'agit de la représentation graphique de la fonction d'autocorrélation qui traduit la corrélation entre les observations à un instant donné, X_t , et les instants précédents X_{t-k} , (où X_t est la série étudiée et $k = 0 \dots n$, un entier naturel, n étant le nombre total d'observations).

L'autocovariance au décalage k ou d'ordre k , noté $\gamma(k)$, est définie par [10] :

$$\gamma(k) = \frac{1}{n} \sum_{j=1}^{n-k} (X_{j+k} - \bar{X}_{j+k})(X_j - \bar{X}_j) \text{ avec } \bar{X}_{t-k} = \frac{1}{n-k} \sum_{i=1}^{n-k} X_i$$

L'autocorrélation au décalage k , notée $\rho(k)$ ou ρ_k , est définie par :

$$\rho(k) = \frac{\gamma(k)}{\gamma(0)}$$

La fonction d'autocorrélation (AC) est la fonction qui à k associe $\rho(k)$. Le corrélogramme est le graphique de la fonction d'autocorrélation.

Autocorrélation Partielle :

L'autocorrélation partielle (PAC) entre X_t et X_{t-k} est la corrélation entre X_t et X_{t-k} , l'influence des variables décalées ($X_{t-1}, X_{t-2}, \dots, X_{t-k+1}$) ayant été retirée.

Ainsi, l'autocorrélation partielle d'ordre k d'une série X_t , notée $r(k)$ ou r_k est définie par le dernier coefficient de la projection linéaire de X_{t+1} sur ces k plus récentes valeurs [16]:

$$X_{t+1} = c_1 X_t + c_2 X_{t-1} + \dots + c_{k-1} X_{t-k+1} + r_k X_{t-k+1}$$

Ce qui est équivalent à

$$\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ r_k \end{pmatrix} = \begin{pmatrix} \gamma_0 & \gamma_1 & \dots & \gamma_{k-1} \\ \gamma_1 & \gamma_0 & \dots & \gamma_{k-2} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{k-1} & \gamma_{k-2} & \dots & \gamma_0 \end{pmatrix}^{-1} \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_k \end{pmatrix} = \begin{pmatrix} 1 & \rho_1 & \dots & \rho_{k-1} \\ \rho_1 & 1 & \dots & \rho_{k-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \dots & 1 \end{pmatrix}^{-1} \begin{pmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_k \end{pmatrix}$$

$$\text{d'où } r_k = \frac{|D_k|}{|D|}$$

$$\text{Avec } D = \begin{pmatrix} 1 & \rho_1 & \dots & \rho_{k-1} \\ \rho_1 & 1 & \dots & \rho_{k-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \dots & 1 \end{pmatrix}, D_k = \begin{pmatrix} 1 & \rho_1 & \dots & \rho_1 \\ \rho_1 & 1 & \dots & \rho_2 \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \dots & \rho_k \end{pmatrix}$$

La fonction d'autocorrélation partielle (PAC) est la fonction qui à k associe r_k . Le corrélogramme partiel est sa représentation graphique.

Au seuil 5%, nous dirons que l'autocorrélation au décalage k est significatif si :

$$|\rho(k)| > \frac{1.96}{\sqrt{\text{nombre des observations}}} \quad (\text{Il en est de même pour } r_k)$$

I.4.2 LES MODELES AUTOREGRESSIFS (AR)

Un processus autorégressif d'ordre p ($AR(p)$) est tout processus qui peut s'écrire :

$$Y_t = \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \dots + \alpha_p Y_{t-p} + \varepsilon_t ;$$

où ε_t est un bruit blanc et α_i , ($i=1; \dots; p$) une constante réelle (avec $|\alpha_i| < 1$)

Le corrélogramme (graphe de la fonction d'autocorrélation) d'un processus $AR(p)$ décroît de façon géométrique tandis que le corrélogramme partiel (graphe de la fonction d'autocorrélation partielle) a ses seuls p premiers termes significatifs (différents de 0).

I.4.3 LES MODELES MOYENNES MOBILES (MA)

Une série Y_t est dit processus Moyenne mobile d'ordre q ($MA(q)$) si chaque observation Y_t est générée par une moyenne pondérée d'aléas jusqu'à la q -ième période [10] :

$$Y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$$

Où $\theta_1, \theta_2, \dots, \theta_q$ sont des paramètres pouvant être négatifs.

Le corrélogramme d'un modèle $MA(q)$ a ses seuls q premiers retards différents de 0, le corrélogramme partiel décroît de manière géométrique.

I.4.4 LES MODELES AUTOREGRESSIFS MOYENNES MOBILES (ARMA)

On dit qu'un processus Y_t suit un $ARMA$ d'ordre (p,q) (noté $ARMA(p,q)$) si on peut écrire :

$$Y_t - \alpha_1 Y_{t-1} - \alpha_2 Y_{t-2} - \dots - \alpha_p Y_{t-p} = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}$$

On peut dire, en gros, que si on ne se retrouve pas d'une façon évidente en face d'un processus $AR(p)$ ou $MA(q)$, on a de fortes chances de se trouver en face d'un processus $ARMA(p,q)$.

NB : Les modèles AR , MA et $ARMA$ ne sont représentatifs que des chroniques stationnaires en tendance et corrigées des variations saisonnières.

I.4.5 CARACTERISTIQUES DES CORRELOGRAMMES DES PROCESSUS AR, MA ET ARMA

Modèle	Fonction d'autocorrélation (ACF)	Fonction d'autocorrélation partielle (PACF)
$AR(p)$	Décroissance exponentielle et/ou sinusoïdale	Pics significatifs seulement pour les p premiers retards
$MA(q)$	Pics significatifs seulement pour les q premiers retards	Décroissance exponentielle et/ou sinusoïdale
$ARMA(p,q)$	Décroissance exponentielle ou sinusoïdale tronquée après $(q-p)$ retards	Décroissance exponentielle ou sinusoïdale amortie tronquée après $(p-q)$ retards

I.4.6 LES MODELES ARIMA

Si la série étudiée est de type DS (stationnaire en différence), il convient de la stationnariser par passage aux différences selon l'ordre d'intégration $I = d$ (le nombre de fois qu'il faut différencier la série pour la rendre stationnaire) la série est alors étudiée selon les modèles BOX-JENKINS ($AR(p)$, $MA(q)$ ou $ARMA(p,q)$). On parle alors de modèle *Autoregressive Integrated Moving average* d'ordre (p,d,q) et on note $ARIMA(p,d,q)$.

Remarque : L'identification a priori consiste, à partir de la lecture du corrélogramme et du test de stationnarité, d'une part à détecter s'il s'agit d'un modèle $ARMA$ ou $ARIMA$ et d'autres part à déterminer les ordres (p,q) du modèle $ARMA$ ou (p,d,q) du modèle $ARIMA$.

II. ESTIMATION DES PARAMETRES

Les méthodes d'estimation diffèrent selon le type de processus diagnostiqué. Dans le cas d'un modèle $AR(p)$, nous appliquons la méthode des moindres carrés.

Pour l'estimation des paramètres d'un modèle $ARMA(p,q)$ BOX et JENKINS suggèrent une procédure itérative de type balayage décrit ainsi qu'il suit [1] :

Supposons le processus :

$$Y_t - \alpha_1 Y_{t-1} - \alpha_2 Y_{t-2} = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2}$$

Ce qui est équivalent à :

$$(1 - \alpha_1 D - \alpha_2 D^2) Y_t = (1 - \theta_1 D - \theta_2 D^2) \varepsilon_t,$$

$$\text{où } D \text{ est l'opérateur décalage défini par } DY_t = Y_{t-1} \text{ et } D^s Y_t = Y_{t-s}$$

Ceci revient à écrire

$$Y_t = \frac{(1 - \theta_1 D - \theta_2 D^2) \varepsilon_t}{1 - \alpha_1 D - \alpha_2 D^2}$$

$$\text{en posant } v_t = \frac{1}{1 - \alpha_1 D - \alpha_2 D^2} \varepsilon_t, \text{ il vient que, } v_t - \alpha_1 v_{t-1} - \alpha_2 v_{t-2} = \varepsilon_t \quad (3)$$

ce qui donne :

$$Y_t = v_t - \theta_1 v_{t-1} - \theta_2 v_{t-2} \Leftrightarrow v_t = Y_t + \theta_1 v_{t-1} + \theta_2 v_{t-2} \quad (4)$$

Initialisons la procédure de balayage en partant de deux valeurs plausibles pour $(\hat{\theta}_1; \hat{\theta}_2)$ et

d'un pas d'incrément. Puis, pour chaque couple de valeurs $(\hat{\theta}_1; \hat{\theta}_2)$ posons :

$\hat{v}_0 = 0$ et $\hat{v}_1 = 0$, calculons les valeurs estimées de \hat{v}_1 à partir de la relation (4) :

$$\hat{v}_2 = Y_2$$

$$\hat{v}_3 = Y_3 + \hat{\theta}_1 \hat{v}_2$$

$$\hat{v}_4 = Y_4 + \hat{\theta}_1 \hat{v}_3 + \hat{\theta}_2 \hat{v}_2$$

...

Après calcul de toutes les valeurs de v_t , nous estimons les paramètres α_1 et α_2 par la méthode des moindres carrés appliquée à l'équation (3) :

$$v_t = \alpha_1 v_{t-1} + \alpha_2 v_{t-2} + \varepsilon_t \quad (5)$$

Nous retenons les valeurs $\hat{\theta}_1, \hat{\theta}_2$ et $\hat{\alpha}_1, \hat{\alpha}_2$ qui rendent minimum la somme des carrés des résidus issue de la régression de l'équation (5).

III. IDENTIFICATION A POSTERIORI

Nous sommes partis des observations pour arriver à un modèle supposé représentatif de la série. Reste à valider si ce modèle est vraiment représentatif des observations. Cette validation du modèle passe par un certain nombre de tests.

III.1 TEST DE SIGNIFICATIVITE DES COEFFICIENTS

Les coefficients du modèle doivent être significativement différents de 0, le test de Student s'applique de manière classique. Si un coefficient n'est pas significativement différent de 0, on élimine le terme correspondant dans la formulation du modèle.

III.2 TEST DE SIGNIFICATIVITE GLOBALE DU MODELE

La significativité globale du modèle peut se mesurer par la qualité d'ajustement des données avec le modèle : le R^2

soit Y_t , la série des observations

\hat{Y}_t , la série des estimations issues du modèle

soit $\rho(Y_t, \hat{Y}_t)$, le coefficient de corrélation entre Y_t et \hat{Y}_t

$$\rho(Y_t, \hat{Y}_t) = \frac{\text{cov}(Y_t, \hat{Y}_t)}{\sqrt{\text{var}(Y_t) * \text{var}(\hat{Y}_t)}}$$

$R^2 = \rho^2(Y_t, \hat{Y}_t) \leq 1$, mesure la qualité d'ajustement.

Ainsi, R^2 sera d'autant plus proche de 1 que \hat{Y}_t expliquera bien Y_t .

III.3 TEST SUR LES RESIDUS

L'analyse des résidus s'effectue à partir de deux critères : ils doivent être un bruit blanc et leur distribution doit être Gaussienne. Deux tests s'imposent de ce fait : le test de bruit blanc et le test de normalité.

III.3.1 TEST DE BRUIT BLANC : Test de Box-Pierre et Ljung-Box

Les tests de Box-Pierre et Ljung-Box permettent d'identifier les processus de bruit blanc (suite de variables aléatoires de même distribution et indépendante entre elles). Il est question donc d'identifier si $\text{cov}(Y_t, Y_{t-k}) = 0$ ou encore $\rho_k = 0$.

Les deux hypothèses à tester sont les suivantes :

$$H_0 : \rho_1 = \rho_2 = \dots = \rho_h = 0$$

Contre H_1 : Il existe au moins un ρ_i significativement différent de 0

Pour effectuer ces tests, on recourt à la statistique Q de Box-Pierre ou Q' de Ljung-Box qui sont données par :

$$Q = \sum_{k=1}^h \hat{\rho}_k^2,$$

$$Q' = n(n+2) \sum_{k=1}^h \frac{\hat{\rho}_k^2}{n-k},$$

où h = nombre de retards, ρ_k = autocorrélation d'ordre k , n = nombre des observations

Les statistiques Q et Q' sont distribuées (sous l'hypothèse H_0) de manière asymptotique comme des χ^2 (chi-deux) à h degrés de liberté. Nous rejetons donc l'hypothèse de bruit blanc, au seuil α , si la statistique Q ou Q' est supérieur au χ^2 lu dans la table au seuil $(1-\alpha)$ et h degrés de liberté.

III.3.2 TEST DE NORMALITE : TEST de Shapiro-Wilk

Ce test non-paramétrique a pour principe de base le Q-Q plot. On cherche à mesurer la corrélation entre quantiles théoriques et quantiles empiriques. Les étapes du test sont les suivantes :

(a) On réordonne les valeurs de la série $\{\hat{u}_i\}_{i=1}^n$ par ordre croissant et l'on note $\{\hat{u}_{[i]}\}_{i=1}^n$ la série ordonnée.

(b) On calcule une somme pondérée des écarts entre les valeurs extrêmes de $\{\hat{u}_{[i]}\}_{i=1}^n$ et $\{\hat{u}_{[n-i+1]}\}_{i=1}^n$.

$$\delta = \sum_{i=1}^{n/2} a_i (\hat{u}_{[n-i+1]} - \hat{u}_{[i]})$$

où les poids a_i ont été tabulés par simulation de Monte-Carlo pour $i=1, \dots, n/2$ et pour différentes valeurs de n .

(c) On calcule la statistique de test SW dont la distribution a été tabulée par simulation de Monte-Carlo.

$$SW = \left[\frac{\delta}{\sqrt{n-1}\hat{\sigma}_{\hat{u}}} \right]^2$$

où $\hat{\sigma}_{\hat{u}}$ est l'écart type estimé de la série des résidus $\{\hat{u}_i\}_{i=1}^n$.

La région de rejet est définie pour des valeurs de SW inférieures à la valeur critique* sw_α au seuil α .

IV. PREVISION

Lorsque le modèle est validé, la prévision peut alors être calculée à un horizon de quelques périodes, limité car la variance de l'erreur de prévision croît très vite avec l'horizon. L'intervalle de prévision donne l'idée de la valeur que pourrait prendre la prévision si l'on tenait compte de ces erreurs.

* Voir table de Shapiro en annexe

V. ALGORITHME DE BOX-JENKINS

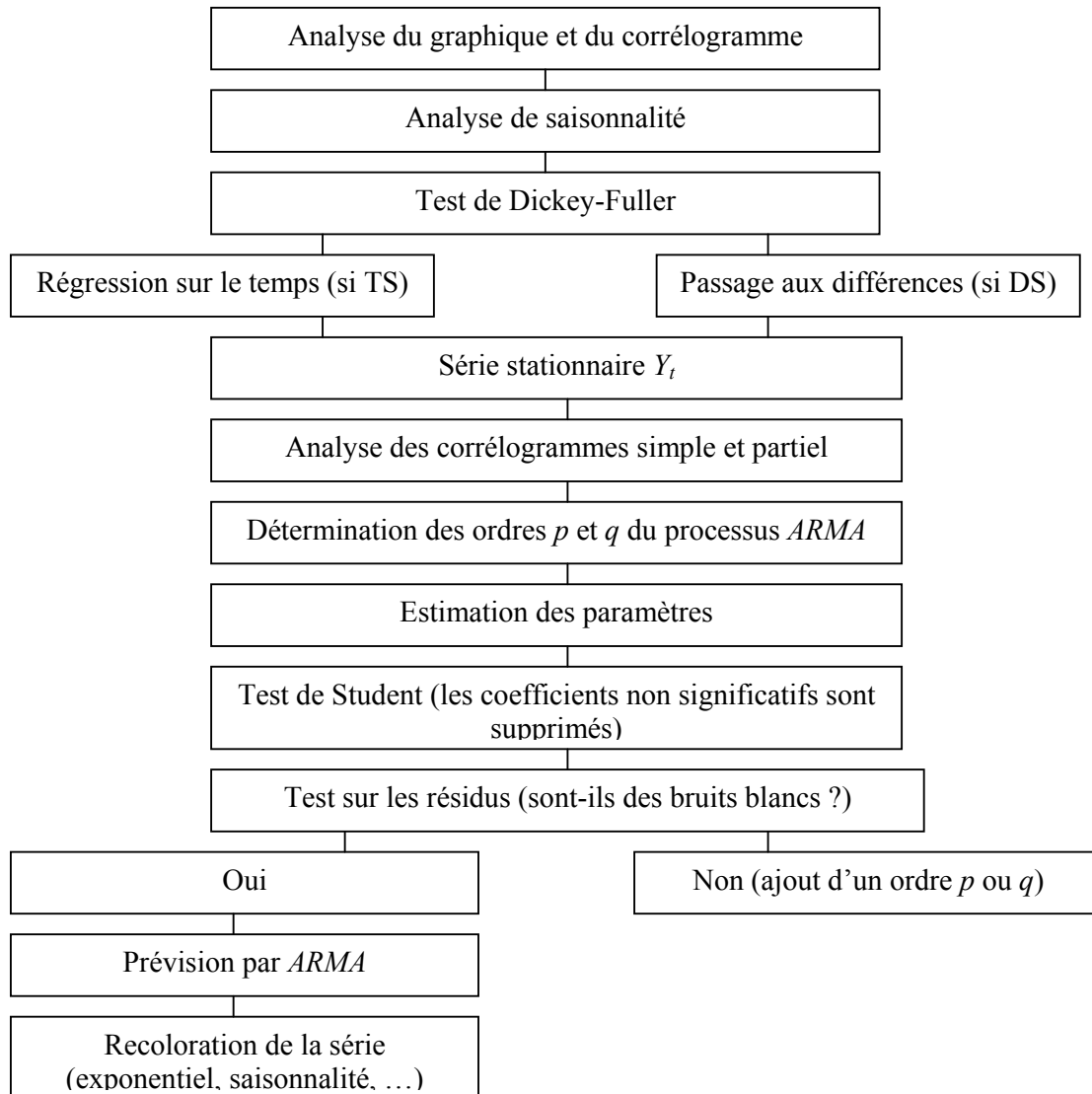


Figure 12 : Algorithme de BOX-JENKINS.

CHAPITRE IV :

**APPLICATIONS DE LA
METHODE DE BOX-JENKINS A
LA PREVISION DU NOMBRE DES
COMPTES D'EPARGNE DE LA
FIRST BANK**

I. IDENTIFICATION A PRIORI

I.1 ANALYSE DU GRAPHIQUE

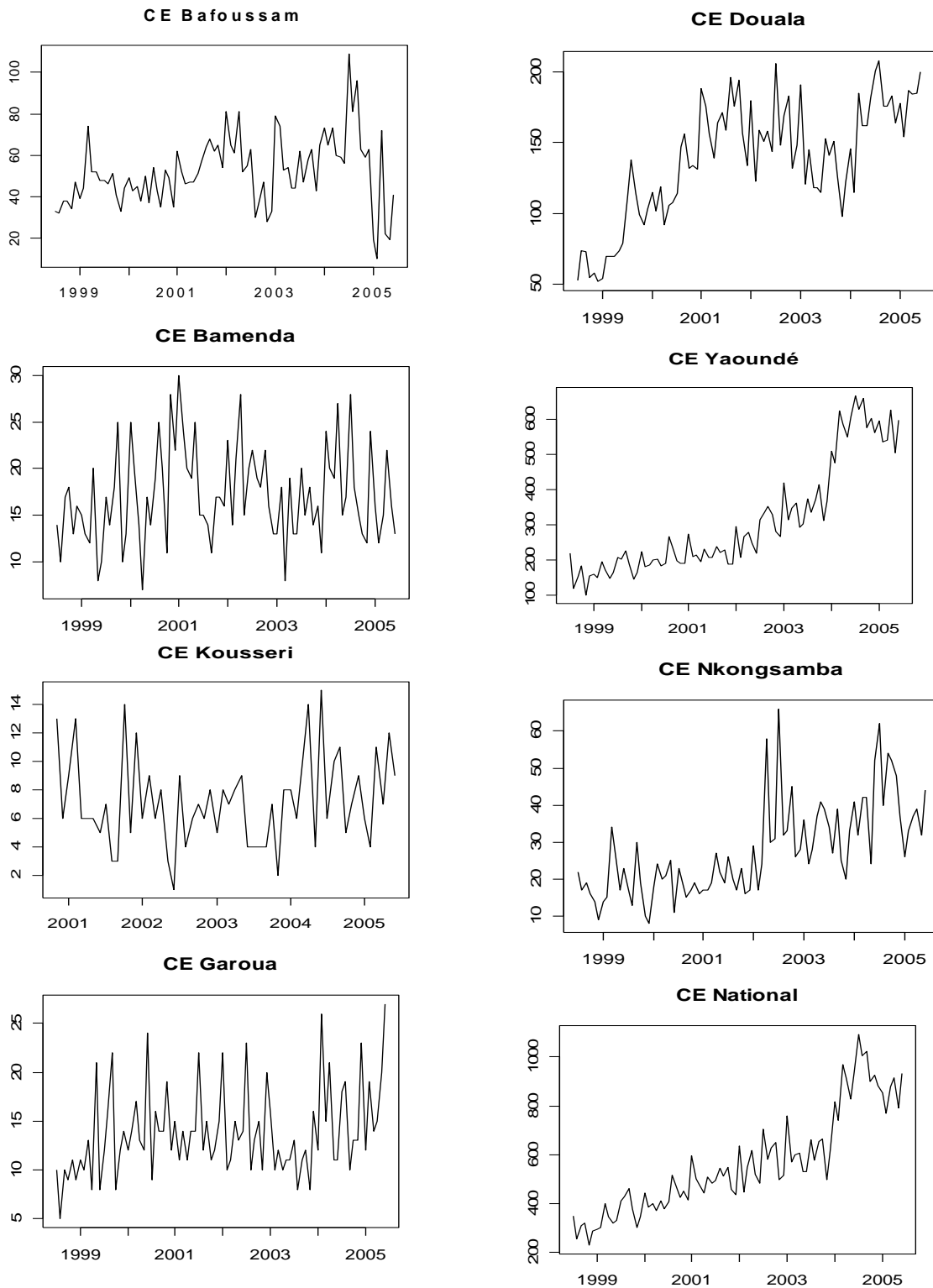


Figure 13 : Graphiques des séries brutes

L'allure du graphique du **Réseau National** (cf. graphique « CE National », page précédente) présente un processus à tendance linéaire fortement croissante, avec une baisse dans l'ensemble des agences en 2003, suivie d'une brusque élévation en 2004, synonyme des phénomènes conjoncturels internes ou externes à la First Bank, qui se seraient produits au cours de ces deux années. Ce genre d'événements dont nous avons donné les causes probables à la partie descriptive de ce travail est difficilement prévisible.

Nous pouvons aussi constater que Bamenda, Garoua et Kousseri avec leurs autocorrélations (AC) et autocorrélations partielles (PAC) presque toutes nulles (cf. corrélogramme des séries brutes en annexe), semblent stationnaires. Tandis que les autres agences affichent des tendances linéaires croissantes, malgré la perturbation générale constatée en 2003. Elles sont donc non stationnaires et doivent être stationnarisées avant toute étude.

I.2 ANALYSE DE LA SAISONNALITE :

I.2.1 Choix du schéma de décomposition et dessaisonnalisation

Un test de bande peut nous permettre de décider si une série peut être décomposée selon un schéma additif, multiplicatif ou bien mixte. Il s'agit de constater à partir de la représentation graphique de la série, si les oscillations de la série ont une amplitude proportionnelle à la valeur de la tendance. Si tel est le cas, la série peut être décomposée selon un schéma multiplicatif ou mixte.

Or nos séries ont toutes des tendances soit constantes soit croissantes (voir graphiques à la page précédente) qui ne sont pas proportionnelles aux oscillations des séries. C'est dire qu'elles ne peuvent être décomposées ni selon le schéma multiplicatif, ni un schéma mixte. Nous les décomposerons donc toutes selon le schéma additif.

La dessaisonnalisation consistera donc à soustraire l'effet saisonnier des valeurs observées de la série après l'avoir estimé. Ceci nous permettra d'obtenir des séries corrigées des variations saisonnières (CVS) que nous étudierons par la suite (voir graphiques des séries dessaisonnalisées en annexe).

I.2.2 Calcul des coefficients saisonniers

La saisonnalité que nous examinerons ici est annuelle (série périodique de période $P=12$). Pour le calcul des coefficients saisonniers, nous nous sommes servis du modèle de BUYS-BALLOT que nous avons décrit au chapitre précédent. Une régression par moindres carrés ordinaire nous a permis d'estimer les pentes pour chaque série et un simple calcul dans Excel en nous servant de la formule de Buys-Ballot que nous rappelons ci-dessous, nous a permis d'obtenir les coefficients saisonniers. Les résultats sont consignés dans le tableau qui suit, pour chaque série :

Rappel : $\hat{S}_j = \bar{X}_j - \bar{X}_{..} - \hat{\alpha}[j - (p+1)/2], j = 1, \dots, p, p = 12.$

où \hat{S}_j est le coefficient Saisonnier, $\hat{\alpha}$ la pente, \bar{X}_j , la moyennemensuelle et $\bar{X}_{..}$ la moyenne de la série

Coefficients saisonniers													
Agences	Jan.	Fév.	Mars	Avril	Mai	Juin	Juil.	Août	Sept.	Oct.	Nov.	Déc.	Pente
Yaoundé	74,2	10,9	48,0	48,0	8,6	17,3	-3,3	-16,3	-8,7	-28,7	-77,6	-72,6	5,72
Douala	19,5	-9,0	12,6	-3,6	2,5	5,8	4,3	10,9	4,4	-4,0	-21,1	-22,3	1,23
Nkongsamba	0,0	-3,0	2,6	9,0	0,1	3,7	7,8	-4,3	0,9	-0,8	-7,5	-8,5	0,35
Réseau Nat	105,5	0,5	72,4	56,9	5,1	21,7	20,0	-17,2	-3,6	-36,6	-118,3	-106,4	7,68
Bafousam	6,5	-0,7	9,3	-0,9	-5,5	-4,4	8,9	-3,2	2,4	-0,1	-8,3	-4,2	0,18
garoua	0,7	0,7	0,0	-0,4	0,8	1,7	1,4	-1,6	-0,4	-2,3	-1,9	1,4	0,05
bamenda	3,7	0,1	-1,6	3,1	-1,6	-2,6	2,1	-0,8	-0,5	0,0	-1,2	-0,8	0
Kousseri	0,0	2,0	0,2	2,0	-1,5	-0,8	-0,5	-1,8	-1,0	1,2	-2,0	2,2	0

I.2.3 Courbes des effets saisonniers

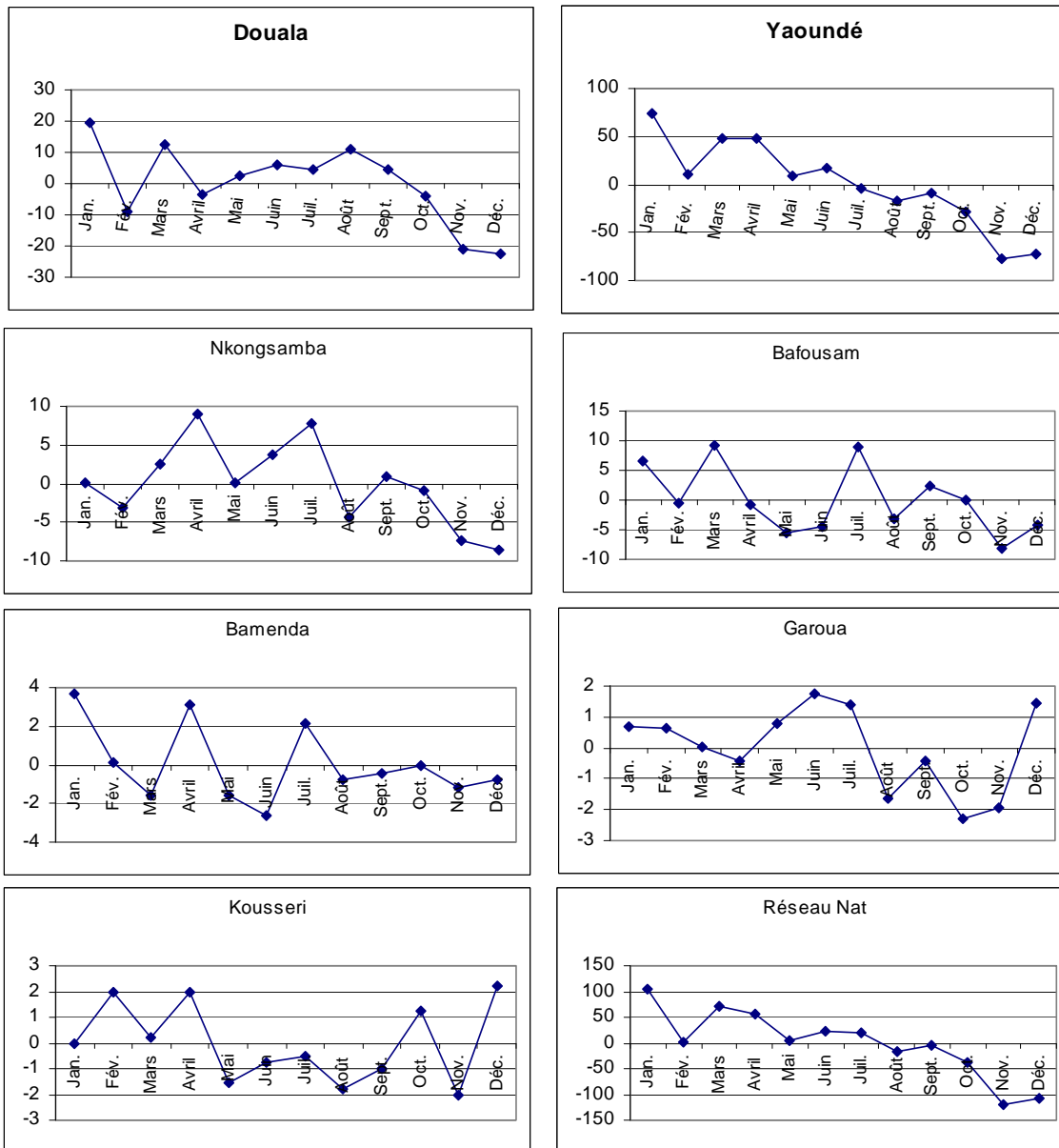


Figure 14 : Graphiques des effets saisonniers, selon le modèle de Buys-Ballot

I.3 ANALYSE DE LA STATIONNARITE

I.3.1 Test de Stationnarité

Le test de stationnarité (Test de Dickey-Fuller) appliqué aux différentes séries, qui teste de la **stationnarité en différence** contre la **stationnarité en tendance**, retourne les valeurs de la statistique de Dickey-Fuller consignées dans le tableau suivant. Comme il s'agit d'un test unilatéral, l'hypothèse nulle est rejetée lorsque la statistique obtenue est inférieure à la valeur critique contenue dans la table de Dickey-Fuller^{**}, au seuil correspondant.

Agences	Bfsam	Nsamba	Douala	Garoua	Kousseri	Bamenda	Yaoundé	Réseau
Statistiques	-6.0861	-6.7534	-3.9254	-9.5025	-8.3949	-8.0308	-3.2227	-4.2411

Ce test rejette l'hypothèse nulle pour toutes les séries au seuil de 5 %. En d'autres termes, toutes les séries sont stationnaires en tendance. Leur stationnarisation consistera donc à estimer et à prélever la tendance pour chacune d'entre elles.

I.3.2 Stationnarisation

La régression sur le temps par moindres carrés ordinaires, suivie du test de significativité de Student sur les coefficients nous permet d'obtenir la tendance de chaque série, que nous prélevons des séries CVS pour les stationnariser (cf. chapitre précédent).

Les séries stationnalisées s'obtiennent ainsi qu'il suit :

$$\text{Bafoussam : } Y_t = X_t - 0.19*t - 43.85$$

$$\text{Nkongsamba : } Y_t = X_t - 0.34*t - 13.09$$

$$\text{Douala : } Y_t = X_t - 1.23*t - 85.25$$

$$\text{Yaoundé : } Y_t = X_t - 5.61*t - 73.77$$

$$\text{Bamenda : } Y_t = X_t - 16.26$$

$$\text{Garoua : } Y_t = X_t - 0.05*t - 11.6$$

$$\text{Kousseri : } Y_t = X_t - 6.8$$

$$\text{Réseau National : } Y_t = X_t - 7.56*t - 243.83$$

Où X_t représente la série CVS, Y_t la série stationnarisée et " t ", le temps (l'ordre du mois dans la série).

Remarque : Comme processus stationnaires en tendance, toutes les séries seront modélisées par des processus $AR(p)$, $MA(q)$ ou bien $ARMA(p,q)$.

^{**} Voir table de Dickey-Fuller en annexe

I.3.3 choix des ordres p et q des modèles

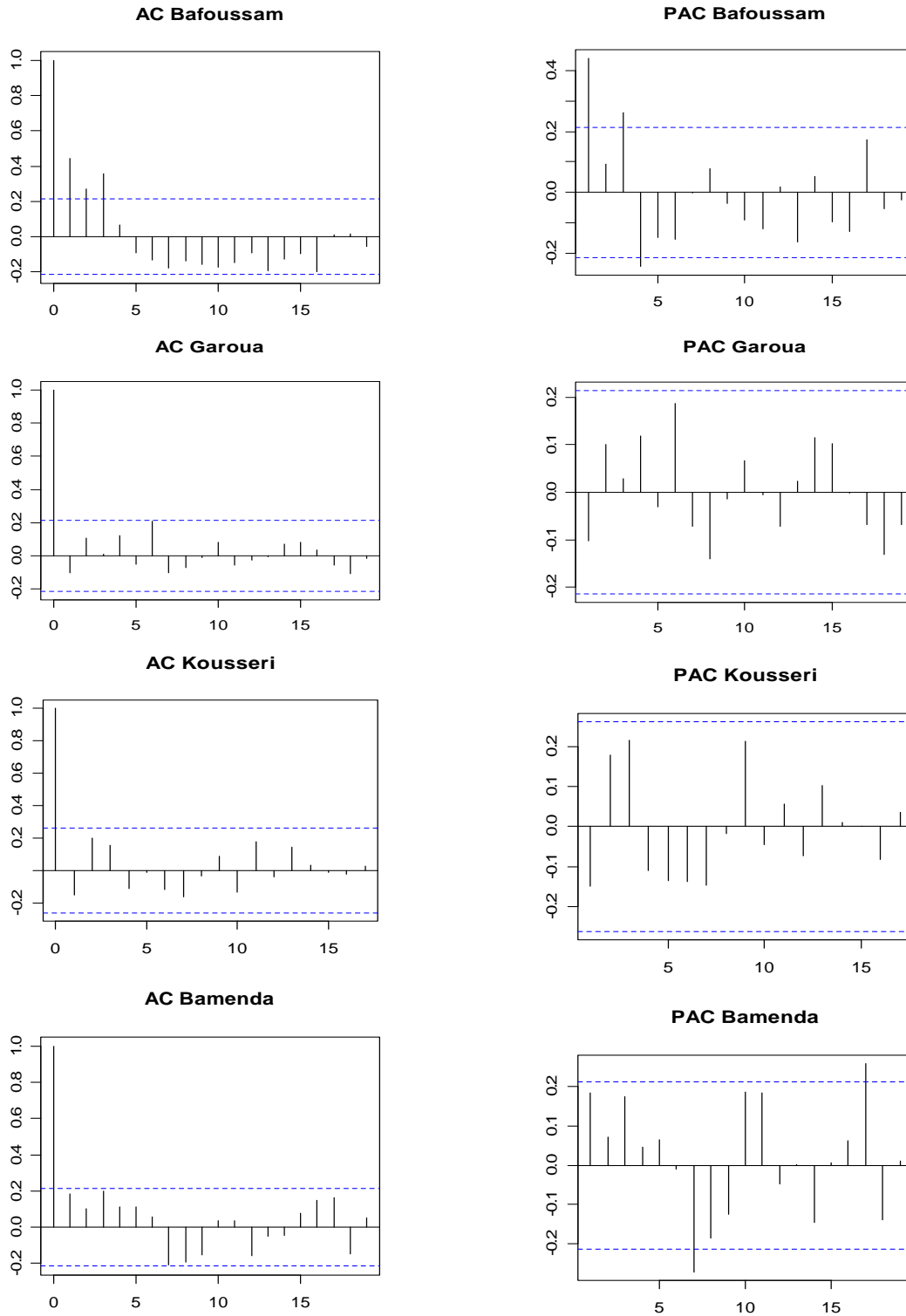


Figure 15 : Corrélogrammes des séries stationnarisées

AC = Auto Corrélation
PAC = Auto Corrélation Partiel

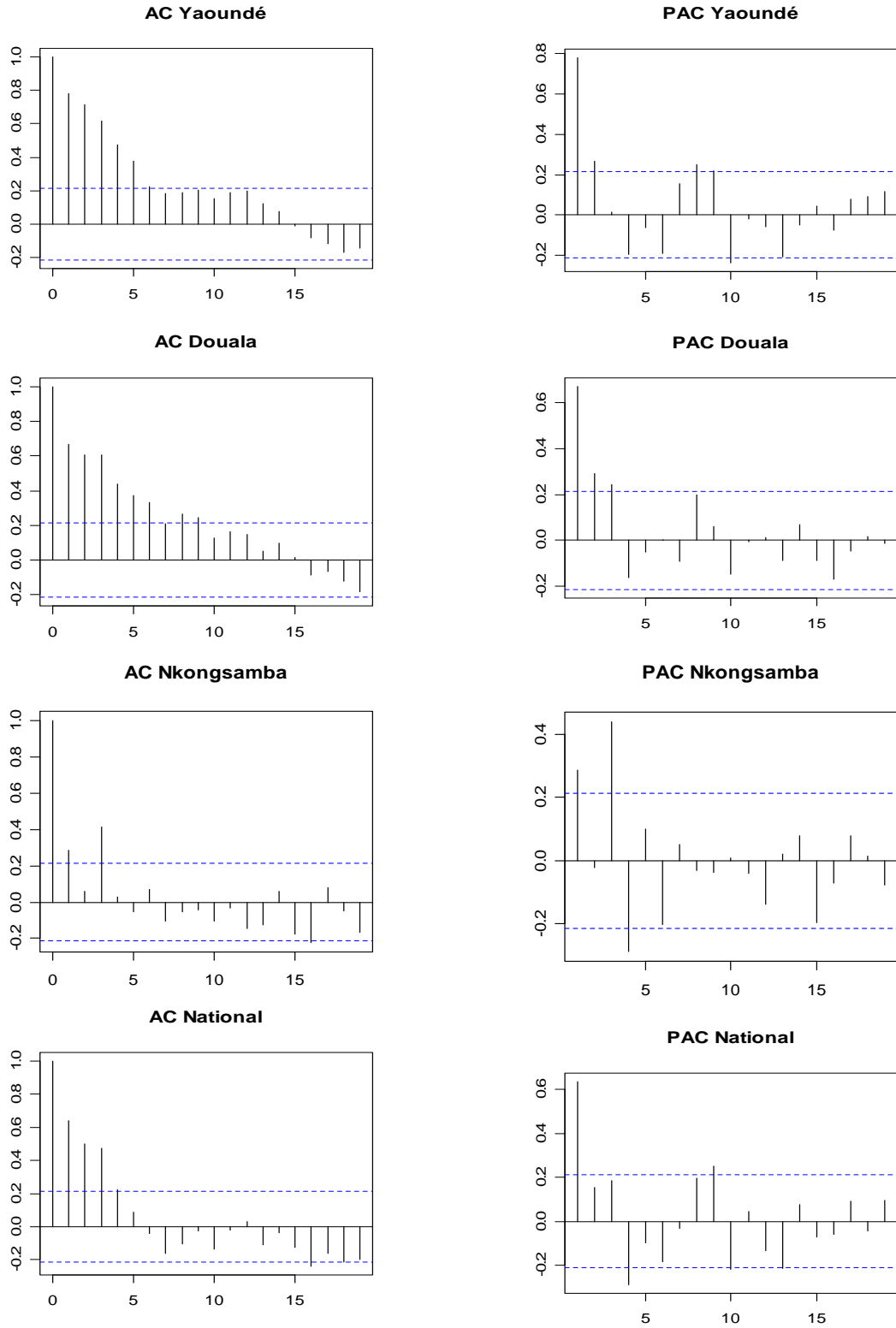


Figure 15 (suite): Corrélogrammes des séries stationnarisées

Yaoundé : Les autocorrélations partielles (PAC) sont non significatives, sauf les deux premières, tandis que les autocorrélations (AC) décroissent. Il s'agit certainement d'un $AR(2)$.

Douala : Les AC décroissent, tandis que les PAC sont nulles sauf les trois premiers. Il s'agirait d'un $AR(3)$.

Réseau National : Les AC sont significatives pour les trois premiers retards tandis que les PAC sont non significatives sauf pour le premier retard. Nous suspectons un processus $ARMA(1,3)$.

Bamenda, Garoua et Kousseri ont tous leurs AC non significatives : il se pourrait qu'ils soient des Bruits blancs.

Nkongsamba et Bafoussam : Leurs AC et PAC ne présentent pas de forme identifiable à première vue. Il s'agit d'un processus $ARMA(p,q)$ dont les ordres p et q seront choisis dans l'intervalle $[0 ;3] \times [0 ;3]$ et arrêtés après la validation du modèle.

II. ESTIMATION DES COEFFICIENTS ET VALIDATION DES MODELES

Modèles pressentis :

Agences	Modèles
Bafoussam	$ARMA(1,3)$
Nkongsamba	$ARMA(3,1)$
Douala	$AR(3)$
Garoua	Bruit Blanc
Kousseri	Bruit Blanc
Bamenda	Bruit Blanc
Yaoundé	$AR(2)$
Réseau National	$ARMA(1,3)$

II.1 Estimation des coefficients

Bafoussam :

Coefficients	ar1	ma1	ma2	ma3	intercept
Estimation	0.3281	0.0460	0.1632	0.5308	-1.0291
Standard error	0.1544	0.1264	0.1034	0.1161	3.3870
Test de Student	Significatif	Non significatif	Non significatif	Significatif	Non significatif
Modèle retenu	$Y_t = 0.33*Y_{t-1} + 0.53*\varepsilon_{t-3} + \varepsilon_t$				

Nkongsamba :

Coefficients	ar1	ar2	ar3	ma1	intercept
Estimation	-0.1717	-0.0061	0.4266	0.6554	-0.0262
Standard error	0.1338	0.1052	0.1049	0.1289	1.5298
Test de Student	Non significatif	Non significatif	Significatif	Significatif	Non significatif
Modèle retenu	$Y_t = 0.43*Y_{t-3} + 0.66*\varepsilon_{t-1} + \varepsilon_t$				

Douala :

Coefficients	ar1	ar2	ar3	intercept
Estimation	0.4027	0.1794	0.2451	-2.4716
Standard error	0.1053	0.1124	0.1049	10.6411
Test de Student	Significatif	Non signif.	Significatif	Non signif
Modèle retenu	$Y_t = 0.4*Y_{t-1} + 0.25*Y_{t-3} + \varepsilon_t$			

Yaoundé :

Coefficients	ar1	ar2	intercept
Estimation	0.5994	0.2535	10.6609
Standard error	0.1091	0.1094	30.8055
Test de Student	Significatif	Significatif	Non significatif
Modèle retenu	$Y_t = 0.6*Y_{t-1} + 0.25*Y_{t-2} + \varepsilon_t$		

Réseau National :

Coefficients	ar1	ma1	ma2	ma3	intercept
Estimation	0.6009	-0.1360	0.0314	0.6302	2.7634
Standard error	0.1120	0.0953	0.0932	0.1063	24.8501
Test de Student	Significatif	Non significatif	Non significatif	Significatif	Non significatif
Modèle retenu	$Y_t = 0.6*Y_{t-1} - 0.63*\varepsilon_{t-3} + \varepsilon_t$				

Bamenda, Garoua et Kousseri : $Y_t = \varepsilon_t$ (bruits blancs)

Tableau récapitulatif des modèles

AGENCES	σ^2	R^2	AIC	MODELES
Bafoussam	148.4	0.950063	671.65	$Y_t = 0.33*Y_{t-1} + 0.53*\varepsilon_{t-3} + \varepsilon_t$
Nsamba	42.25	0.9542837	566.08	$Y_t = 0.43*Y_{t-3} + 0.66*\varepsilon_{t-1} + \varepsilon_t$
Douala	338.6	0.9835828	738.64	$Y_t = 0.4*Y_{t-1} + 0.25*Y_{t-3} + \varepsilon_t$
Yaoundé	1987	0.9837072	885.47	$Y_t = 0.6*Y_{t-1} + 0.25*Y_{t-2} + \varepsilon_t$
Bamenda	21.6	0.9325789	699.76	$Y_t = \varepsilon_t$ (bruits blancs)
Garoua	17.46	0.9176877	699.76	$Y_t = \varepsilon_t$ (bruits blancs)
Kousseri	9.73	0.8439897	699.76	$Y_t = \varepsilon_t$ (bruits blancs)
Réseau Nat.	3768	0.989597	944.27	$Y_t = 0.6*Y_{t-1} - 0.63*\varepsilon_{t-3} + \varepsilon_t$

II.2 VALIDATION DES MODELES

Les tests de normalité (Test de Shapiro : shapiro.test) et de bruit blanc (Test de Box-Pierre : Box.test) sur les résidus donnent les résultats suivants:

Agences	Test de normalité		Test de bruit blanc
	p-value	décision	décision
Bafoussam	0.1050	OK	OK
Nkongsamba	0.1122	OK	OK
Douala	0.931	OK	OK
Garoua	0.007044	-	OK
Kousseri	0.06252	OK	OK
Bamenda	0.4855	OK	OK
Yaoundé	0.4442	OK	OK
Réseau National	0.9663	OK	OK

Les résultats de ces tests nous permettent de conclure que tous les modèles présentés précédemment peuvent être considérés comme valides, à l'exception de Garoua qui est un bruit blanc non gaussien.

III. PREVISION

Les observations de nos séries sont mensuelles et s'étalent sur une période de sept ans. Ainsi, vu le nombre d'observations, nous avons cru bon de limiter la prévision sur un an.

Ces prévisions ne tiennent pas compte des variables extérieures. Elles seront d'autant plus réalisables qu'aucun phénomène conjoncturel, pouvant bouleverser considérablement l'ordre actuel d'évolution des activités de la banque ne se produisait.

Nous avons procédé à la prévision ville par ville, chacune avec ses réalités locales. Nous avons ensuite sommé les résultats des prévisions issues de toutes les villes que nous comparons dans un tableau récapitulatif à la prévision issue des données consolidées du réseau national.

Les colonnes des tableaux de prévision se décrivent de la manière suivante :

Temps : il s'agit de l'ordre des observations (en mois) à partir de juillet 1998 (par exemple, juillet 2005 correspond au 85^e)

Modèle : prévisions obtenues sur les séries stationnalisées à partir des modèles retenus précédemment.

Tendance : valeurs de la tendance que nous avons ôtée pour stationnariser les séries.

Coefficients saisonniers : coefficients soustraits des séries pour les dessaisonnaliser.

Prévision : prévisions finales obtenues après avoir ajouté la tendance et le coefficient saisonnier à la prévision obtenue des différents modèles.

$$(\text{Prévision} = \text{Modèle} + \text{Tendance} + \text{Coefficient Saisonnier}).$$

Intervalle de prévision : intervalles dans lesquels les valeurs attendues devraient tombées avec 95% de chance. Ils sont obtenus à partir du calcul des erreurs sur les résidus.

Taux d'erreur : estimation de l'erreur de prévision.

$$(\text{Taux d'erreur} = \text{demi longueur de l'intervalle de prévision} / \text{prévision})$$

1) Douala :

Mois	temps	Modèle	Tendance	Coefficient saisonnier	Prévision	intervalle de prévision	Taux d'erreur
juil-05	85	1	189	20	210	[174 ; 246]	17%
août-05	86	0	191	-9	182	[143 ; 221]	21%
sept-05	87	1	192	13	206	[165 ; 247]	20%
oct-05	88	0	193	-4	190	[146 ; 234]	23%
nov-05	89	0	194	2	197	[151 ; 243]	23%
déc-05	90	0	196	6	201	[154 ; 248]	23%
janv-06	91	0	197	4	201	[152 ; 249]	24%
févr-06	92	-1	198	11	208	[159 ; 257]	24%
mars-06	93	-1	199	4	203	[154 ; 253]	24%
avr-06	94	-1	201	-4	196	[146 ; 246]	26%
mai-06	95	-1	202	-21	180	[129 ; 231]	28%
juin-06	96	-1	203	-22	180	[129 ; 231]	28%

2) Yaoundé :

Mois	temps	Modèle	Tendance	Coefficient saisonnier	Prévision	intervalle de prévision	Taux d'erreur
juil-05	85	13	551	74	638	[550 ; 725]	14%
août-05	86	18	556	11	585	[483 ; 687]	17%
sept-05	87	16	562	48	626	[511 ; 741]	18%
oct-05	88	16	567	48	631	[507 ; 754]	20%
nov-05	89	15	573	9	597	[467 ; 727]	22%
déc-05	90	14	579	17	610	[475 ; 746]	22%
janv-06	91	14	584	-3	595	[456 ; 734]	23%
févr-06	92	14	590	-16	588	[446 ; 729]	24%
mars-06	93	13	596	-9	600	[456 ; 744]	24%
avr-06	94	13	601	-29	585	[439 ; 731]	25%
mai-06	95	13	607	-78	542	[395 ; 689]	27%
juin-06	96	13	612	-73	553	[405 ; 700]	27%

3) Nkongsamba :

Mois	temps	Modèle	Tendance	Coefficient saisonnier	Prévision	intervalle de prévision	Taux d'erreur
juil-05	85	-1	42	0	41	[28 ; 54]	32%
août-05	86	-4	42	-3	35	[21 ; 49]	40%
sept-05	87	0	43	3	45	[31 ; 59]	31%
oct-05	88	0	43	9	52	[36 ; 67]	30%
nov-05	89	-2	43	0	41	[26 ; 57]	37%
déc-05	90	0	44	4	47	[32 ; 63]	33%
janv-06	91	0	44	8	52	[36 ; 67]	30%
févr-06	92	-1	44	-4	39	[24 ; 55]	40%
mars-06	93	0	45	1	46	[31 ; 62]	34%
avr-06	94	0	45	-1	44	[28 ; 59]	35%
mai-06	95	0	45	-7	38	[22 ; 53]	41%
juin-06	96	0	46	-9	37	[21 ; 53]	43%

4) Bafoussam :

Mois	temps	Modèle	Coefficient saisonnier	Tendance	Prévision	intervalle de prevision	Taux d'erreur
juil-05	85	-16	7	60	51	[27 ; 75]	48%
août-05	86	-13	-1	60	47	[21 ; 72]	55%
sept-05	87	-12	9	60	58	[31 ; 84]	46%
oct-05	88	-5	-1	61	55	[25 ; 86]	56%
nov-05	89	-2	-5	61	53	[22 ; 83]	57%
déc-05	90	-1	-4	61	56	[25 ; 86]	55%
janv-06	91	-1	9	61	69	[38 ; 100]	45%
févr-06	92	-1	-3	61	57	[26 ; 88]	54%
mars-06	93	-1	2	62	63	[32 ; 96]	51%
avr-06	94	-1	0	62	61	[30 ; 92]	51%
mai-06	95	-1	-8	62	53	[22 ; 84]	59%
juin-06	96	-1	-4	62	57	[26 ; 88]	54%

5) Garoua :

Mois	temps	Tendance	Coefficient saisonnier	Prévision	intervalle de prevision	Taux d'erreur
juil-05	85	16	1	17	[9 ; 26]	49%
août-05	86	16	-2	14	[6 ; 23]	60%
sept-05	87	16	0	16	[8 ; 25]	55%
oct-05	88	16	-2	14	[6 ; 23]	62%
nov-05	89	16	-2	14	[6 ; 23]	60%
déc-05	90	16	1	18	[10 ; 27]	48%
janv-06	91	16	1	17	[9 ; 26]	50%
févr-06	92	16	1	17	[9 ; 26]	50%
mars-06	93	16	0	16	[8 ; 25]	52%
avr-06	94	16	0	16	[8 ; 25]	54%
mai-06	95	16	1	17	[9 ; 26]	50%
juin-06	96	16	2	18	[10 ; 27]	47%

6) Bamenda :

Mois	temps	Tendance	Coefficient saisonnier	Prévision	intervalle de prevision	Taux d'erreur
juil-05	85	16	4	20	[11 ; 29]	45%
août-05	86	16	0	16	[7 ; 25]	55%
sept-05	87	16	-2	15	[6 ; 24]	61%
oct-05	88	16	3	19	[10 ; 28]	46%
nov-05	89	16	-2	15	[6 ; 24]	61%
déc-05	90	16	-3	14	[5 ; 23]	66%
janv-06	91	16	2	18	[9 ; 27]	49%
févr-06	92	16	-1	16	[7 ; 25]	58%
mars-06	93	16	0	16	[7 ; 25]	57%
avr-06	94	16	0	16	[7 ; 25]	55%
mai-06	95	16	-1	15	[6 ; 24]	60%
juin-06	96	16	-1	16	[7 ; 25]	58%

7) Kousseri :

Mois	temps	Tendance	Coefficient saisonnier	Prévision	intervalle de prevision	Taux d'erreur
juil-05	85	7	0	7	[1 ; 13]	89%
août-05	86	7	2	9	[3 ; 15]	68%
sept-05	87	7	0	7	[1 ; 13]	85%
oct-05	88	7	2	9	[3 ; 15]	68%
nov-05	89	7	-2	5	[0 ; 11]	114%
déc-05	90	7	-1	6	[0 ; 12]	100%
janv-06	91	7	-1	6	[0 ; 12]	96%
févr-06	92	7	-2	5	[0 ; 11]	119%
mars-06	93	7	-1	6	[0 ; 12]	104%
avr-06	94	7	1	8	[2 ; 14]	75%
mai-06	95	7	-2	5	[0 ; 11]	126%
juin-06	96	7	2	9	[3 ; 15]	66%

8) Réseau National :

Mois	temps	Modèle	Tendance	Coefficient saisonnier	Prévision	intervalle de prevision	Taux d'erreur
juil-05	85	72	886	106	1064	[943 ; 1 184]	11%
août-05	86	33	894	1	928	[795 ; 1 060]	14%
sept-05	87	68	902	72	1042	[905 ; 1 180]	13%
oct-05	88	42	909	57	1008	[839 ; 1 178]	17%
nov-05	89	26	917	5	948	[769 ; 1 128]	19%
déc-05	90	17	924	22	963	[780 ; 1 146]	19%
janv-06	91	11	932	20	963	[779 ; 1 147]	19%
févr-06	92	8	939	-17	930	[745 ; 1 114]	20%
mars-06	93	6	947	-4	949	[764 ; 1 133]	19%
avr-06	94	5	954	-37	923	[738 ; 1 107]	20%
mai-06	95	4	962	-118	848	[663 ; 1 033]	22%
juin-06	96	3	970	-106	866	[682 ; 1 051]	21%

9) Tableau récapitulatif

	Ydé	Dla	Bfsam	Nsamba	bmda	groua	Kseri	Total des Agences	Réseau National
juil-05	638	210	51	41	20	17	7	983	1064
août-05	585	182	47	35	16	14	9	888	928
sept-05	626	206	58	45	15	16	7	972	1042
oct-05	631	190	55	52	19	14	9	970	1008
nov-05	597	197	53	41	15	14	5	922	948
déc-05	610	201	56	47	14	18	6	952	963
janv-06	595	201	69	52	18	17	6	958	963
févr-06	588	208	57	39	16	17	5	929	930
mars-06	600	203	63	46	16	16	6	949	949
avr-06	585	196	61	44	16	16	8	926	923
mai-06	542	180	53	38	15	17	5	849	848
juin-06	553	180	57	37	16	18	9	869	866

Une comparaison de la somme des prévisions des agences et de la prévision du réseau national laisse voir une différence significative pour les cinq premiers mois, et un rapprochement à partir du sixième mois. Ceci rassure quant aux résultats obtenus. Toutefois, vu les caractéristiques des modèles et des taux d'erreur de prévision, les résultats du réseau national (dernière colonne du tableau précédent) sont beaucoup plus vraisemblables.

10) Graphiques des prévisions

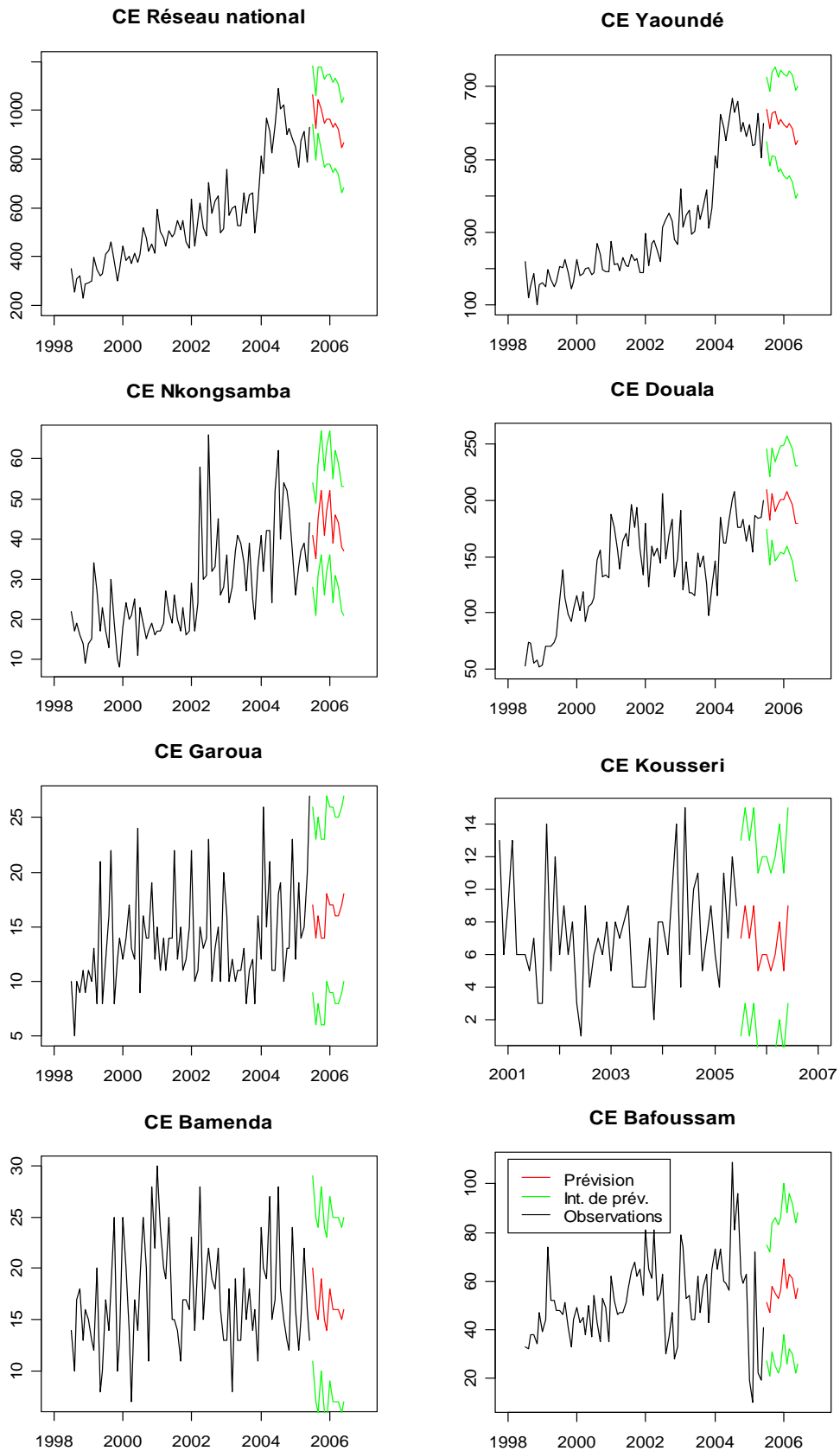


Figure 16 : Graphiques des prévisions

CE = Comptes d'épargne

CONCLUSION

A l'heure de la globalisation et de son corollaire d'ouverture des marchés nationaux, les entreprises africaines ont plus que jamais besoin de connaître les leviers qu'elles peuvent actionner pour ne pas rester en marge de la dynamique économique mondiale. La question est alors de savoir ce qui explique le positionnement actuel des entreprises africaines au sein de ce vaste marché mondial et quelles sont leurs forces et leurs faiblesses. Evoluant dans un contexte hautement concurrentiel, la maîtrise de ses parts de marché et son évolution au cours du temps devraient être inscrites dans les priorités de toute entreprise soucieuse de son devenir. Ce travail est une tentative de contribution à cette réflexion. Notre étude devrait permettre à la FIRST BANK de voir le chemin qu'elle a parcouru de juillet 1998 jusqu'en juin 2005 en terme du nombre de comptes ouverts à ses guichets. Notre travail consistait aussi à proposer un modèle de prévision et à prévoir l'évolution de son parc de comptes les douze prochains mois.

Notre analyse descriptive nous a permis de constater que la FIRST BANK est une entreprise en pleine croissance. Les comptes les plus sollicités ici sont les comptes d'épargne qui représentent 62% du parc des comptes dont 56% pour la seule ville de Yaoundé. Ce type de compte enregistrent un taux de croissance moyen de 21.4% lors des sept dernières années. Il s'est particulièrement bien comporté en 2004 tandis l'année 2003 ne lui a pas beaucoup souri. Les comptes de dépôt spécial (CDS) viennent en deuxième position dans le parc des comptes. Ils enregistrent un taux de croissance moyen de 6.4%. Ce type de compte est fortement prisé à Yaoundé qui détient 71% du parc ; cependant, il est de courte durée de vie et presque en voie de disparition dans certaines agences. Les comptes courants (CCC) sont beaucoup plus sollicités à Douala. Malgré leur taux de croissance moyenne de 17.8%, ils restent très inférieurs en valeur absolue aux deux autres types de comptes.

L'étude prévisionnelle nous a amené à construire un modèle de prévision pour chacune des agences de la FIRST BANK, puis, un autre modèle pour les données du réseau national consolidées. Loin d'être redondant, cette approche nous a permis non seulement d'avoir une idée de ce qui se passe localement dans les agences, elle nous permet aussi de faire un rapprochement entre la somme des prévisions issues des agences et la prévision faite sur l'ensemble des données consolidées. Ce rapprochement est satisfaisant à partir du 5^e mois de prévision. Les modèles obtenus, après être soumis aux différents tests de validation, ont été choisis selon les critères statistiques (parmi les modèles valides, ils sont de moindre variance, ils maximisent la vraisemblance et ils sont de moindre AIC). Les prévisions qui en sont issues nous font savoir que la bonne tenue au niveau de l'ouverture des comptes d'épargne en 2004 devrait continuer jusqu'en novembre 2005, période à partir de la quelle ils pourront commencer à enregistrer une baisse dans l'ensemble du réseau national.

Compte tenu du temps alloué pour cette étude, nous n'avons pas pu effectué la prévision sur les deux autres types de comptes. Nous invitons la FIRST BANK à nous permettre de poursuivre l'étude. Les résultats obtenus auraient été plus intéressants si l'étude intégrait l'évolution de la clôture des comptes, ce qui n'a pas été possible faute des données. Ce travail ouvre des portes aux autres études qui, menés sur le parc des comptes de la FIRST BANK pourraient apporter plus d'informations aux décideurs de cette banque. Ainsi, cette réflexion devrait être approfondie en intégrant d'autres paramètres comme les caractéristiques de la clientèle des différents types de comptes, les montants mensuels d'encaissement et de consommation des comptes d'épargne en terme de prêts à la clientèle. Ceci pourra permettre d'établir des indicateurs permettant d'apprécier la rentabilité d'un produit ou d'une agence. Pour la pérennité de ce travail, il est souhaitable que la FIRST BANK prenne des dispositions pour que les prévisions sur le parc des comptes et sur d'autres agrégats soient renouvelées d'année en année. Pour l'instant, il est irréaliste de faire des prévisions de plus d'un an à partir des données actuellement disponibles.

ANNEXE



Appel et formatage des données

```
Bfsam=read.table("Bfsam.txt",h=T)
Bda=read.table("bamenda.txt",h=T)
Kous=read.table("kousseri.txt",h=T)
Gra=read.table("garoua.txt",h=T)
Nkong=read.table("nsamba.txt",h=T)
Dla=read.table("douala.txt",h=T)
Yde=read.table("Yde.txt",h=T)
Cons=read.table("consolidation.txt",h=T)
```

```
cebaf=Bfsam[,2]
cebda=Bda[,2]
cekous=Kous[,2]
cegra=Gra[,2]
cedla=Dla[,2]
ceyde=Yde[,2]
cenkg=Nkong[,2]
cenat=Cons[,2]
```

```
bfs=ts(cebaf,f=12,start=c(1998,7))
bmd=ts(cebda,f=12,start=c(1998,7))
ksr=ts(cekous,f=12,start=c(2000,11))
gar=ts(cegra,f=12,start=c(1998,7))
dou=ts(cedla,f=12,start=c(1998,7))
ynd=ts(ceyde,f=12,start=c(1998,7))
nkg=ts(cenkg,f=12,start=c(1998,7))
nat=ts(cenat,f=12,start=c(1998,7))
```

Représentation Graphique

```
plot(bfs,main="CE Bafoussam",xlab="",ylab="")
plot(bmd,main="CE Bamenda",xlab="",ylab="")
plot(ksr,main="CE Kousseri",xlab="",ylab="")
plot(gar,main="CE Garoua",xlab="",ylab="")
plot(dou,main="CE Douala",xlab="",ylab="")
plot(ynd,main="CE Yaoundé",xlab="",ylab="")
plot(nkg,main="CE Nkongsamba",xlab="",ylab="")
plot(nat,main="CE National",xlab="",ylab="")
```

Corrélogramme simple et partiel

```
acf(bfs,main="AC Bafoussam",xlab="",ylab="")
acf(bmd,main="AC Bamenda",xlab="",ylab="")
acf(ksr,main="AC Kousseri",xlab="",ylab="")
acf(gar,main="AC Garoua",xlab="",ylab="")
acf(dou,main="AC Douala",xlab="",ylab="")
acf(ynd,main="AC Yaoundé",xlab="",ylab="")
acf(nkg,main="AC Nkongsamba",xlab="",ylab="")
acf(nat,main="AC National",xlab="",ylab="")
```

```
pacf(bfs,main="PAC Bafoussam",xlab="",ylab="")
pacf(bmd,main="PAC Bamenda",xlab="",ylab="")
pacf(ksr,main="PAC Kousseri",xlab="",ylab="")
pacf(gar,main="PAC Garoua",xlab="",ylab="")
pacf(dou,main="PAC Douala",xlab="",ylab="")
pacf(ynd,main="PAC Yaoundé",xlab="",ylab="")
pacf(nkg,main="PAC Nkongsamba",xlab="",ylab="")
pacf(nat,main="PAC National",xlab="",ylab="")
```

Dessaisonalisation

```
dla=read.table("dla.txt",h=T)
yde=read.table("yde.txt",h=T)
nkong=read.table("nkong.txt",h=T)
nat=read.table("nat.txt",h=T)
```

```
gar=read.table("gra.txt",h=T)
bda=read.table("bda.txt",h=T)
bfs=read.table("bfs.txt",h=T)
kous=read.table("kous.txt",h=T)
```

```
dlacvs=dla[,4]
ydecvs=yde[,4]
natevs=nat[,4]
nkcvvs=nkong[,4]
```

Stationnarisation

```
y.stat=function(x){
n=length(x)
t=1:n
reg=lm(x~t)
ynouv= x-predict(reg)
}
```

```
dla=y.stat(dlacvs)
yde=y.stat(ydecvs)
nat=y.stat(natevs)
nkg=y.stat(nkcvvs)
```

```
bfs=bfscvs-predict(bf)
bda=bdacvs-predict(bd)
kous=koucvs-predict(k)
gar=garcvs-predict(g)
```

Estimation des coef. et validation des modèles

```
y=arima(yde,c(2,0,0))
d=arima(dla,c(3,0,0))
n=arima(nat,c(1,0,3))
```

```
nk=arima(nkg,c(3,0,1))
```

```
baf=arima(bfs,c(1,0,3))  
bam=arima(bfs,c(0,0,0))  
ksr=arima(bfs,c(0,0,0))  
gra=arima(bfs,c(0,0,0))
```

```
tsdiag(b)  
tsdiag(n)  
tsdiag(d)  
tsdiag(y)  
tsdiag(r)
```

```
shapiro.test(b$res)  
shapiro.test(n$res)  
shapiro.test(d$res)  
shapiro.test(y$res)  
shapiro.test(r$res)
```

```
##### Prévision #####
```

```
f=function(x,mod){  
  p=predict(arima(x,order=mod),n.ahead=12)  
  s=1.96*p$se  
  min=p$pred-s  
  max=p$pred+s  
  int=paste(round(min),round(max))  
  prediction=round(p$pred)  
  list(prediction,int)  
}
```

```
f(dla,c(3,0,0))  
f(nkg,c(3,0,1))  
f(yde,c(2,0,0))  
f(nat,c(1,0,3))
```

```
f(bfs,c(1,0,3))  
f(bda,c(0,0,0))  
f(kous,c(0,0,0))  
f(gar,c(0,0,0))
```

```
Bfs=read.table("bfs.txt",h=T)  
Bda=read.table("bda.txt",h=T)  
Kous=read.table("kous.txt",h=T)  
Gra=read.table("gar.txt",h=T)
```

```
pbfs=Bfs[,2]  
pbda=Bda[,2]  
pkous=Kous[,2]  
pgra=Gra[,2]
```

```
pdla=Dla[,2]
pyde=Yde[,2]
pnkg=Nkg[,2]
pnat=Nat[,2]
```

```
mbfs=Bfs[,3]
mbda=Bda[,3]
mkous=Kous[,3]
mgra=Gra[,3]
```

```
Mbfs=Bfs[,4]
Mbda=Bda[,4]
Mkous=Kous[,4]
Mgra=Gra[,4]
```

```
mbs=ts(mbfs,f=12,start=c(2005,7))
mbd=ts(mbda,f=12,start=c(2005,7))
mks=ts(mkous,f=12,start=c(2005,7))
mgr=ts(mgra,f=12,start=c(2005,7))
```

```
Mbs=ts(Mbfs,f=12,start=c(2005,7))
Mbd=ts(Mbda,f=12,start=c(2005,7))
Mks=ts(Mkous,f=12,start=c(2005,7))
Mgr=ts(Mgra,f=12,start=c(2005,7))
```

```
bs=ts(pbfs,f=12,start=c(2005,7))
bd=ts(pbda,f=12,start=c(2005,7))
ks=ts(pkous,f=12,start=c(2005,7))
gr=ts(pgra,f=12,start=c(2005,7))
```

Graphique des prévisions

```
plot(bfs,main="CE Bafoussam",xlab="",ylab="",xlim=c(1998,2007))
lines(bs,col="red")
lines(mbs,col="green")
lines(Mbs,col="green")
legend(1998, 110, c(paste(c("Prévision","Int. de prév.,"Observations"))), lty = 1, col =
c("red", "green", "black"))
```

```
plot(bmd,main="CE Bamenda",xlab="",ylab="",xlim=c(1998,2007))
lines(bd,col="red")
lines(mbd,col="green")
lines(Mbd,col="green")
```

```
plot(ksr,main="CE Kousseri",xlab="",ylab="",xlim=c(2001,2007))
lines(ks,col="red")
lines(mks,col="green")
lines(Mks,col="green")
```

Table de Dickey-Fuller (DF)

Critical values for the Dickey-Fuller t statistics

Sample size	Level of significance		
	0.01	0.05	0.10
Testing $\Phi = 0$ in $\Delta x_t = \Phi \cdot x_{t-1} + \eta_t$ "t-test" statistic			
25	-2.66	-1.95	-1.60
50	-2.62	-1.95	-1.61
100	-2.60	-1.95	-1.61
250	-2.58	-1.95	-1.62
Infinite	-2.58	-1.95	-1.62
Testing $\Phi = 0$ in $\Delta x_t = \Phi \cdot x_{t-1} + \beta + \eta_t$ "t-test" statistic			
25	-3.75	-3.00	-2.62
50	-3.58	-2.93	-2.60
100	-3.51	-2.89	-2.58
250	-3.46	-2.88	-2.57
Infinite	-3.44	-2.87	-2.57
Testing $\Phi = 0$ in $\Delta x_t = \Phi \cdot x_{t-1} + \beta + \gamma \cdot t + \eta_t$ "t-test" statistic			
25	-4.38	-3.60	-3.24
50	-4.15	-3.50	-3.18
100	-4.04	-3.45	-3.15
250	-3.99	-3.43	-3.13
Infinite	-3.98	-3.42	-3.13
Standard t-distribution (Student)			
Infinite	-2.33	-1.65	-1.28
(Source: W.A. Fuller (1976), Table 8.5.2)			

Extrait du site « **Séries chronologiques et problèmes de stationnarité** »
www.chez.com/deaecodemo/stationnarite.doc

Table de Shapiro-Wilk

Nombre d'observations	Risque 5%	Risque 1%
10	0.842	0.781
11	0.850	0.792
12	0.859	0.805
13	0.856	0.814
14	0.874	0.825
15	0.881	0.835
16	0.837	0.844
17	0.892	0.851
18	0.897	0.858
19	0.901	0.863
20	0.905	0.868
21	0.908	0.873
22	0.911	0.878
23	0.914	0.881
24	0.916	0.884
25	0.918	0.888
26	0.920	0.891
27	0.923	0.894
28	0.924	0.896
29	0.926	0.898
30	0.927	0.900
31	0.929	0.902
32	0.930	0.904
33	0.931	0.906
34	0.933	0.908
35	0.934	0.910
36	0.935	0.912
37	0.936	0.914
38	0.938	0.916
39	0.939	0.917
40	0.940	0.919
41	0.941	0.920
42	0.942	0.922
43	0.943	0.923
44	0.944	0.924
45	0.945	0.926
46	0.945	0.927
47	0.946	0.928
48	0.947	0.929
49	0.947	0.929
50	0.947	0.930

Extrait du site « **Ressources nationales de chimie** »

<http://www.educnet.education.fr/rnchimie/math/benichou/tests/normalite/normalite.htm>

BIBLIOGRAPHIE

- [1] **Régis BOURBONNAIS**, Université Paris-IX Dauphine « Manuel et exercices corrigés, Econométrie », DUNOD, 4^e édition, 2002.
- [2] **Simon LEBLOND-Isabelle Belley-Ferris**, « Guide d'économétrie appliquée », Département de sciences économiques, Université de Montréal, Document de travail, 2004.
- [3] **Jean-Marie DUFOUR**, Département de sciences économiques Université de Montréal. « Techniques de séries chronologiques », Document de travail, 1998.
- [4] **Serge DEGERINE**, UNIVERSITE JOSEPH FOURIER, « Cours de séries chronologiques », Document de travail, 2004.
- [5] **ALLAN RIDING et GEORGE HAINES FILS**, EQUINOX MANAGEMENT CONSULTANTS LTD, Canada, « Préviation des coûts liés aux prêts accordés en vertu de la LPPE et de la LFPEC », Document de travail, 2003.
- [6] **Dominique LADIRAY et Benoît QUENNEVILLE**, « Désaisonnaliser avec la méthode X-11 », Document de travail, 2000.
- [7] **Xavier BRY**, Ecole Nationale de statistique et d'économie Appliquée (ENEA), « Analyse et prévision élémentaires des séries temporelles », X. BRY – ENEA-1998.
- [8] **Danielle FLORENS**, Université Paris-IX Dauphine, cours de processus de ruine, Master de Statistique, Université de Yaoundé I, 2005.
- [9] **Xavier GUYON**, Université Paris I, cours de modèle linéaire et ses extensions, Master de Statistique, Université de Yaoundé I, 2005
- [10] **Elisabeth GASSIAT**, Université Paris-Sud, cours de séries chronologiques, Master de Statistique, Université de Yaoundé I, 2005
- [11] **Henri GWET**, Université de Yaoundé I, cours de Modèles Probabilistes, Master de Statistique, Université de Yaoundé I, 2005
- [12] **NDONG NGUEMA**, Université de Yaoundé I, cours de Statistique Mathématique, Master de Statistique, Université de Yaoundé I, 2005
- [13] **Bernard RAPACCHI**, centre interuniversitaire de calcul de Grenoble, « Dur, dur!! Les séries chronologiques !! », Document de travail, 1993.
- [14] **Walter ZUCCHINI, Oleg NENADIE**, « Time series analysis with R – Part I », Document de cours.
- [15] **R Development Core Team** (2005). **R** : A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
- [16] **Christophe HURLIN**, U.F.R, « Econométrie Appliquée Séries Temporelles », Document de cours.