

Résumé

Jusqu'à 2020 plus que la moitié de la population Marocaine n'avait pas de couverture en matière de sécurité sociale, et plus précisément 58% de la population n'étaient jamais inscrites dans un tel système. Ceci a changé le mercredi 29 juillet 2020 par le discours de Sa Majesté le Roi Mohamed VI dont lequel, Sa Majesté a donné ces orientations pour **la généralisation de la protection Sociale**.

Notre mission est la réalisation d'une étude actuarielle sur le régime de retraite introduit par ce chantier Royal, pour s'assurer que le tarif défini précédemment est toujours valable en élaborant un modèle exprimant l'équivalence entre les engagements de chaque partie, dans le cadre des réglementations qui régissent ce régime. Le résultat de ce modèle est un tarif.

Nous avons étudié la sensibilité de ce tarif par rapport aux différents paramètres qui définissent sa valeur. Et Enfin, on a réalisé des projections pour étudier la pérennité du régime, en se basant sur un modèle de gestion actif-passif actualisé par une courbe des taux construits suivant un modèle stochastique.

Mots clés :

Tarifification, Assurance vie, TNS, Retraite, Courbe des taux, Projection, Valeur d'acquisition, Régime en points.

Dédicace

A mon cher père, qui nous a quittés il y a moins d'un an, nous demandons à Dieu de l'entrer au paradis.

A ma très chère maman, que Dieu tout puissant te remplisse de santé, de bonheur et te donne une longue vie.

A mes chers professeurs Durant mon parcours scolaire **M. Abdesslam AIT BAKHA, M. Zakaria BEHHAR, M. Abdellatif FIRAMI, M. Mohamed achraf DERRAZ**, leur encouragement, leur aide.

A mon cher binôme qui sans lui ce travail n'aurait pas vu le jour.

A mes chers amis, En souvenir des moments heureux passés ensemble, avec mes sincères vœux de réussite, de bonheur, de santé et de prospérité.

IBNJAA Said

Dédicace

To my most beautiful mother, my ever precious and loving father and my supporting brother, for their support, their patience and their unfailing encouragement.

To my uncle Si Mohamed, without whom we wouldn't have come this far and for his numerous pieces of advice.

To my best partner, without whom this work wouldn't have come to see the light.

To my friends whom with I shared the hardships, trials and the happy moments.

I hope they all fulfil their wishes and dreams and that this won't be the end but only a beginning of a new chapter.

ZAHRIOUI Mohamed

Remerciements

Nous n'oserions pas commencer ces remerciements sans exprimer toute la gratitude que nous ressentons envers celui qui nous a suivi, encadré et soutenu tout au long de cette étude :

Mme Asmaa ELANBARI et **Kawtar ELAFFANI** par leur disponibilité, leur professionnalisme et leur expertise technique.

Nous adressons notre profonde gratitude à notre encadrant interne et notre professeur **M. Pierre DEVOLDER** qui nous a honoré d'avoir accepté de nous encadrer sans oublier ses fructueux conseils et remarques tout au long de la réalisation du présent travail.

Nous exprimons aussi nous gratitude au **M. Khalil SAID** qui nous a honoré en acceptant de juger ce travail.

Nous sommes profondément reconnaissants à notre professeur de sciences actuarielles, **M. Fouad MARRI**, pour ses conseils et remarques tout au long de notre parcours académique à l'INSEA.

Nous adressons aussi nos plus sincères remerciements à toute l'équipe pédagogique de l'INSEA pour nous avoir transmis leur savoir tout au long de ces trois années de formations.

Enfin, nous tenons chacun à remercier nos chers parents pour leur soutien perpétuel.

Table des matières

1	Sécurité sociale	14
1.1	Introduction et Définition :	15
1.2	Régime de la sécurité sociale au Maroc	16
1.2.1	Prévoyance sociale au Maroc :	17
1.2.2	La Caisse Nationale de Sécurité Sociale :	17
2	Cadre d'Étude	21
2.1	Présentation de la problématique :	22
2.1.1	Régimes de Retraite :	22
2.1.2	Présentation de La Population Étudiée :	24
3	Tarification et automatisation sous R shiny	29
3.1	Tarification	30
3.1.1	Hypothèses :	30
3.1.2	Engagement des Assurés :	32
3.1.3	Engagement de l'assureur :	32

3.1.4	Automatisation :	38
3.1.5	Étude de Sensibilité :	44
4	Courbe des taux	47
4.0.1	Introduction :	48
4.0.2	Modèles stochastiques :	48
4.0.3	Réseaux de neurones Artificielles Introduction :	69
5	Projection	86
5.1	Introduction :	87
5.1.1	Outils :	87
5.1.2	Structure Financière :	90
5.2	Informations et Requises :	90
5.3	Projection Démographique :	95
5.4	Projection Financière :	98
5.4.1	Introduction :	98
5.4.2	Formulation :	98
5.4.3	Automatisation sous R shiny :	98

liste des figures

1.1	Organigramme CNSS	19
1.2	Pôles CNSS	20
1.3	Financement et plafonds	20
2.1	Distribution de la population par sexe	25
2.2	Âge moyen par catégorie pour les personnes moins de 65 ans	26
2.3	Âge moyen global par catégorie	27
2.4	Distrubution de la population par catégorie	27
2.5	Revenus par catégorie	28
3.1	Interface de la tarification sous R shiny	39
3.2	Interface d’affichage des résultats de la tarification	41
3.3	Taux d’Augmentation ed la VA et la VL au cours des années	42
3.4	Calcul de la rente - Interface	42
3.5	Taux d’Anticipation ou de Prorogation par âge	43
3.6	VA en fonction de l’âge	44
3.7	VA en fonction du Taux Technique	45

3.8	VA en fonction du taux d'augmentation	46
4.1	Evolution TMP entre 2016 et 2020	55
4.2	Courbe des taux CNSS 2022/01/01	56
4.3	Test Stationnarité Dickey-fuller	57
4.4	Graphique de Comparaison de l'Allure du modèle de Vasicek ajusté sans terme stochastique	58
4.5	Graphique de Comparaison de l'Allure du modèle de Vasicek ajusté avec le terme stochastique	59
4.6	Test normalité des résidus pour le modèle de Vasicek	59
4.7	Test d'homoscédasticité pour le modèle de Vasicek	60
4.8	Test d'Autocorrélation des Erreurs pour le modèle de Vasicek	60
4.9	Test normalité des résidus pour le modèle CIR	61
4.10	Test d'homoscédasticité pour le modèle CIR	62
4.11	Test d'Autocorrélation des Erreurs pour le modèle CIR	62
4.12	Comparaison entre le ZC CNSS et les courbes de taux estimées avec les deux modèles	63
4.13	Interface Rshiny pour la construction de la courbe des taux	64
4.14	Importation	65
4.15	Graphe TMP et résultats de l'estimation pour le modèle de Vasicek	66
4.16	Graphe de comparaison et table des taux pour le modèle de Vasicek	66
4.17	Graphe TMP et résultats de l'estimation pour le modèle CIR	67
4.18	Graphe de comparaison et table des taux pour le modèle CIR	67
4.19	Graphe de comparaison et table des taux pour les deux modèles	68

4.20	illustration de la possibilité de télécharger les données	68
4.21	Représentation de l'ACP sous une architecture ANN	71
4.22	Changement de la moyenne au long terme - TMP	73
4.23	Réseau de neurones récurrente	75
4.24	Réseau de neurones récurrent pour plus de deux entrées	76
4.25	Type des RNNs	76
4.26	Cellule LSTM	79
4.27	Architecture du réseau	80
4.28	Graphes d'Apprentissage	81
4.29	2 ^{me} Architecture.	82
4.30	Courbe d'évolution de la moyenne d'erreur quadratique.	83
4.31	Résultats de l'entraînement sur toute la base de données.	84
5.1	Évolution du taux d'entrée par année de projection	89
5.2	Illustration de la projection démographique	90
5.3	Taux de nuptialité par âge	91
5.4	Taux d'évolution globale et de naissance	92
5.5	Probabilité d'avoir un enfant normale de moins de 21 ans	93
5.6	Evolution du décalage d'âge pour les enfants normaux	93
5.7	Probabilité d'avoir un enfant handicapé	94
5.8	Lissage de la courbe de probabilité d'avoir enfant handicapé par moyenne mobile	95
5.9	Structure de projection démographique	96
5.10	Interface pour la projection des cotisations et des prestations	99

5.11 Taux d'entrée par année de projection	99
5.12 Projection pour L'Adoules	100
5.13 Prestations cumulées Vs Cotisations cumulées	101
5.14 projection des cotisations pour toutes les catégories	102
5.15 projection des prestations pour toutes les catégories	102

Introduction Générale

« Un Maroc des Compétences, où tous les citoyens disposent des capacités et jouissent du bien-être leur permettant de prendre en main leur projet de vie et de contribuer à la création de valeur : L'ambition ne peut être portée que par un capital humain aux capacités et aux compétences renforcées. Le capital humain est le moteur de la dynamique de développement, d'inclusion et de réactivation de l'ascenseur social. Il est également le principal déterminant de la capacité du pays à créer des richesses et à accélérer sa convergence vers les standards des nations avancées. » Rapport Générale du Nouveau Modèle de Développement.

Alors que ce capital Humain est encore en souffrance : « **des franges importantes de la population, notamment les femmes et les jeunes, souffrent de faible participation et de marginalisation faute d'accès à des opportunités d'autonomisation et d'accompagnement. Les inégalités sont également entretenues par la faiblesse des mécanismes de protection sociale et l'inefficience des filets sociaux.** » Rapport Général du nouveau Modèle de développement. C'est pour cette raison que Sa Majesté le Roi Mohammed VI a personnellement guidé et présidé l'établissement de ce chantier Royal de « **Généralisation de Protection Sociale** » pour assurer une meilleure subsistance à ses sujets ainsi qu'une forte stimulation de la croissance économique qui pourrait à elle

seule fournir les bases d'une vie meilleure pour la population. Le système de retraite pour les TNS sur lequel nous travaillons entrés dans le cadre ce chantier Royal et qui a pour objectif de Garantir une retraite pour les Travailleurs Non salariés 'TNS' dont les bénéficiaires peuvent compter pour 22 Millions de personnes.

Le système de retraite établie pour ces trois Millions d'assurés est un en point dont la méthode de calcul et de gestion sont régis par le Dahir N° 1-17-109 du 16 rabii I 1439 portant promulgation de la loi n° 99-15. Ce Dahir institut un régime de pensions pour les catégories professionnels, des travailleurs non-salariées exerçant une activité libérale.

Notre étude prend, en premier temps, sur : une Tarification de ce régime de pension, une étude de sensibilité et une projection pour prédire le déficit à titre d'intervention immédiate mais qui est régie par une loi.

Chapitre **1**

Sécurité sociale

1.1 Introduction et Définition :

Le terme sécurité sociale fait référence au système de compensation établi par les autorités compétentes pour le public, sujet de l'imprévisible de la vie (chômage, maladie, accident, vieillesse, décès d'un parent mineur, etc.). Il a pour objectif de base : l'assurance d'une source minimale de revenu et d'autres droits basique (éducation, eau, nourriture, santé) à tous les citoyens sauf en cas de catastrophe ou de guerre. Elle se base sur un système de solidarité entre les cotisants eux-même ou entre les générations, pourtant une cotisation minimale payée par chaque personne s'inscrit dans ce système.

Selon la convention N° 102 e l'OIT, on distingue 9 branches :

- Soins médicaux (AMO dans notre cas)
- Indemnisation de maladie
- Indemnisation de maternité
- prestation de chômage
- Pensions de Retraite ou assurance contre le vieillissement.
- Prestations des accidents de travail ou des maladies professionnelles
- Allocations familiales
- Prestations d'invalidité

En effet, la sécurité sociale était présente depuis des centaines d'années mais sous forme très réduite, concentrée principalement autour d'une solidarité entre les membres de la même famille fournie aux personnes moins chanceuses ou qui sont touchés par un des misérables événements de la vie. En d'autres termes, elle s'est constituée de la charité des autres.

La première forme de sécurité sociale, disant institutionnelle, était introduite au XIXe siècle par Bismarck à travers une assurance sociale au profit des travailleurs. Pour la fin de couvrir les risques émergents avec la naissance de l'industrie.

En 1942, W. Beveridge adressait au gouvernement Britannique à cette époque le rapport dit « The Beveridge Report, Social Insurance and Allied Services » visait à mettre en place un système complet d'assurance sociale "du berceau à la tombe". Il propose que tous les travailleurs versent une contribution hebdomadaire à l'État. En contrepartie, des prestations seraient versées aux chômeurs, aux malades, aux retraités et aux veufs.

Ce dernier était le premier pas envers le système de sécurité sociale dont nous connaissons les caractères mais la plupart des systèmes dont le terme « Moderne » peut-être visionné au vrai sens du terme n'était introduit qu'après la Deuxième Guerre mondiale.

1.2 Régime de la sécurité sociale au Maroc

Au cours des dernières décennies, les gouvernements ont porté une attention particulière à la protection sociale, résultant de la création d'un système pour la raison de maintenir le revenu des travailleurs et éradiquer la pauvreté. Ce système a réussi à coexister avec d'autres régimes obligatoires (régime de la sécurité sociale, couverture médicale de base...) et d'autres qui sont facultatives comme les régimes de retraite privés. Ce système souffre d'une défaillance résumée globalement dans les déficits qui arrivent plus souvent que prédit, comme le cas de la CNSS et cela peut être principalement dû à la couverture limitée à une seule catégorie de la population (42% de la population totale marocaine). Comme déjà mentionné le système de sécurité sociale comporte plusieurs organismes ayant chacun sa propre mission. Alors on ne s'intéressera qu'aux régimes de pension ou de la prévoyance sociale.

1.2.1 Prévoyance sociale au Maroc :

Les organismes de prévoyance sociale ou ceux responsable sur le système des pensions sont principalement :

- **la caisse nationale de Sécurité sociale (CNSS)** : établissement public ayant pour responsabilité la gestion des pensions pour les travailleurs salariés et non-salariés du secteur privé
- **la caisse Marocaine des Retraites (CMR)** : Ce système est un fait pour les employés du secteur publique, gérée principalement par le Ministère de la Finance et de la réforme Administrative.
- **le Régime collectif d'Allocation de Retraite (RCAR)** : une institution de prévoyance sociale gérée par la caisse des dépôts et de gestion (CDG), elle s'occupe des employés du secteur semi-publique non titulaire de l'état et des collectivités locales.
- Pour les salariés de Bank Al Maghrib et de l'Office National d'Electricité et de l'Eau Potable (ONEP) : **l'Autorité de Contrôle des Assurances et de la Prévoyance Sociale (ACAPS)**.

1.2.2 La Caisse Nationale de Sécurité Sociale :

La Caisse National de Sécurité Sociale est un organisme publique d'origine qui date de 1942 dont le nom était la caisse d'aide sociale. cette dernière était crée par le patronat pour servir des allocations familiales aux travailleurs des grandes entreprises et des allocations aux femmes en couche.

Le régime dont nous connaissons les propriétés n'a été instituer qu'après 1959 comme étant un pour les salariés de l'industrie. Il avait pour mission l'assurance contre l'impré-

vue de la vie et les événements qui peuvent toucher au revenu de ce salarié comme les maladies, l'invalidité et le vieillissement en addition des allocations familiales attribuer à leurs ayants-droits et une allocation de décès et une pension des survivants.

Le système est entré en vigueur le 1er avril 1961 et a depuis été étendu aux secteurs de l'agriculture et de l'artisanat. Il a également amélioré et élargi les services fournis et mis en place un réseau d'établissements de santé sous forme de polycliniques.

Conseil Administratif

La gestion de la caisse nationale de sécurité sociale est attribuée à un conseil administratif de 24 membres titulaires nommés par un décret pour une période de 3 ans qui constituent trois groupes représentatifs de trois parties : dont 8 représentent l'État, 8 représentants des employeurs et 8 d'autre en faveur des travailleurs.

Ce conseil est souvent présidé par le premier ministre ou par une autorité déléguée par les autorités

Organigramme :

La représentation schématique des liens fonctionnels, organisationnels et hiérarchiques de la CNSS est présentée sous l'organigramme suivant :

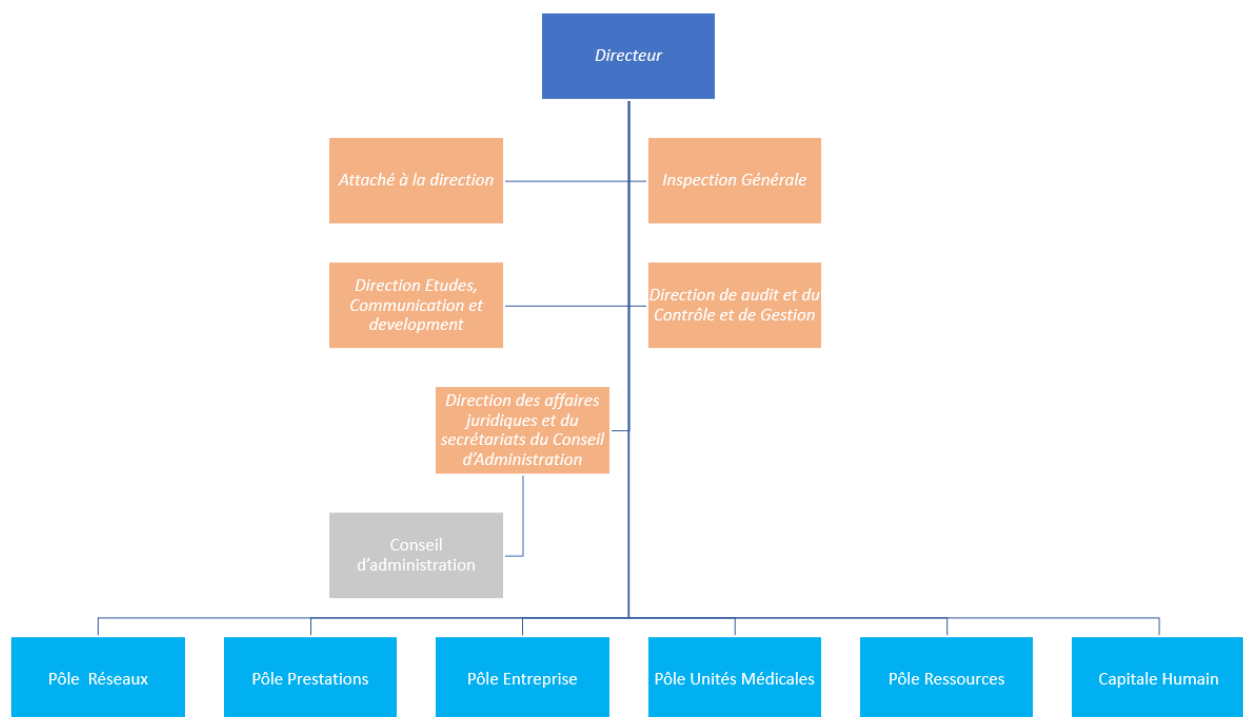


FIGURE 1.1 – Organigramme CNSS

Financement :

Le financement du régime est assuré par une contribution patronale et salariale assise sur les salaires. Les cotisations d'assurance maladie, maternité, décès et les cotisations vieillesse sont versés dans la limite d'un plafond fixé à 6.000 DH par mois. Les cotisations dues au titre des prestations familiales sont versées sur l'ensemble de la rémunération, sans limitation de plafond.

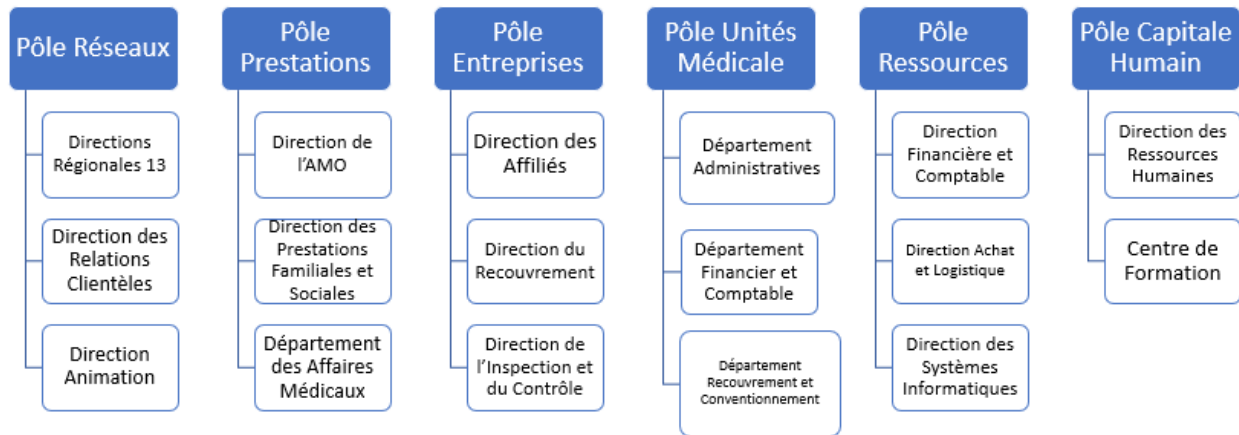


FIGURE 1.2 – Pôles CNSS

Risques	Part patronale	Part salariale	Plafond	Total
Maladie-maternité	0,67%	0,33%	6000 Dhs	1%
prestations décès				
Pension	7,93%	3,96%	6000Dhs	11,89%
Prestations familiales	6,4%	—	pas de plafond	6,4%
Solidarité AMO obligatoire	1,5%	—	pas de plafond	1,5%
AMO de base	2%	2%	pas de plafond	4%
Total	18,50%	6,29%	—	24,79%

FIGURE 1.3 – Financement et plafonds

La cotisation du régime des TNS compte pour 10% d'un revenu forfaitaire défini par la loi. Le système d'Assurance Maladie Obligatoire pour cette catégorie de la population prend 6.37% de ce revenu comme cotisation.

Chapitre **2**

Cadre d'Étude

2.1 Présentation de la problématique :

Notre étude concerne un régime de retraite en Points institué pour les Travailleurs non salariés. Cette étude se composera de trois parties principales, La retarification du régime (puisque le régime est déjà tarifié) en se basant sur des hypothèses qu'on procédera à présenter dans la partie concernée, une étude de sensibilité du tarif pré-mentionné en fonction de plusieurs paramètres et enfin une étude prédictive pour prévisionner l'état financier du régime. Premièrement, on commencera par une présentation des Régimes de pension que sont souvent utilisés ; les régimes en répartition, en capitalisation et celui sur lequel on travaillera.

2.1.1 Régimes de Retraite :

Systeme par répartition :

Un système de retraite par répartition est un système qui illustre parfaitement le sens de solidarité non pas seulement entre les générations mais parfois entre aussi ceux de la même génération, il se basent souvent sur une cotisation prédéfinie en pourcentage du salaire du cotisant en échange d'une prestation en cas de survie ou du survie des ayants droits après un âge de retraite, cette prestation peut ne pas être plafonnée comme elle peut l'être comme le cas pour le système de retraite géré par la CNSS pour les salariés du secteur privé. Ce type de changement est peu sensible aux aléas du marché financier mais il dépend fortement sur la démographie de la population concernée.

Le régime dont la CNSS établie pour les salariés du secteur privé est un peu différent d'un régime par répartition typique. il prend comme indicateur partiel de la pension de retraite le nombre de jours déclarés. Alors tant qu'un individu travaille tant que sa pension de retraite est élevée ce qui constitue une sorte d'équité entre les individus en garantissant la solidarité inhérente à ce type de régime.

Système par Capitalisation :

C'est un système d'équité parfaite dans lequel le cotisant ne cotise que pour lui même. les cotisations sont faites de façon régulière et contrairement au système par répartition, le taux de remplacement est indépendant de la démographie du pays, mais il dépend lourdement de l'état du marché.

Prenant en compte le caractère de singularité des cotisations, le taux de rendement peut être aussi différent d'une personne à une autre dépendant de leur cotisation.

Système en Points :

Dans un système de retraite par points, la pension de la retraite est calculé à partir du nombre de points acquis au cours d'une carrière, et non plus à partir du nombre de jours cotisés comme avec le système de retraite par répartition. Le nombre des points dépendent des revenus touchés par l'assuré. Un salarié peut partir à la retraite dès lors qu'il estime avoir obtenu assez de points, et s'il a atteint l'âge légal de départ à la retraite.

Le système en points est un qui est plus proche au système par capitalisation qu'à celui en répartition, les cotisations sont fait de façon individuelle et convertis en points sur le champ, la conversion se fait par devisions par une valeur appelée « Valeur d'Acquisition», les points sont cumulés jusqu'à l'âge de la retraite ou l'âge de décès de l'assuré, puis ils sont tous convertis en rente ou en capitale selon le cas. Ces conversion sont encadré par la loi n° 99-15 promulguée par le Dahir n° 1-17-109 du rabii I 1439.

À titre d'Exemple :

$$P_n = C_n / VA_n$$

Avec :

P_n : nombre de points acquis au titre de l'année n

C_n : Montant de la cotisation au titre de l'année n , nette de changement de gestion

VA_n : Valeur d'acquisition pour l'année n

La pension de vieillesse est le produit du totale de points inscrit au compte de l'individu et la valeur de liquidation pour l'année multipliée par un coefficient de réduction ou de prorogation selon l'âge de départ à la retraite qui est limité à l'intervalle 60-70.

$$PV = P * VL_n * z$$

VL_n : la valeur de liquidation de pour l'année n

P : Le nombre de points cumulés pendant la durée de cotisation

z : Le coefficient de prorogation ou d'Anticipation

2.1.2 Présentation de La Population Étudiée :

Note : Les statistiques présentées dans cette partie sont préliminaires; c'est à dire qu'ils peuvent toujours être changer car seulement la moitié de la population qui a adhéré au système jusqu'au moment d'acquis de la base de données. l'étude sera faite pour qu'elle soit adaptée au nombre de chaque catégorie ainsi que ses particularité-es.

Rappelons seulement que la base de données est pris de la base de données de l'Assurance de Maladie Obligatoire gérée par la CNSS.

La population étudiée est une composée de plus de 39 Catégories qu'on listera dans l'annexe.

La population compte pour un totale de 1.6 millions d'individus. La distribution de cette population ne peut être qualifiée que de très diverses sauf si on parle de la répartition

selon le sexe du cotisant, dans ce cas on a la distribution suivante :

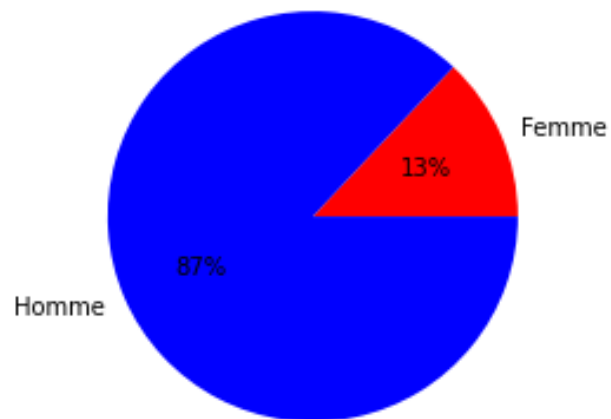


FIGURE 2.1 – Distribution de la population par sexe

Nos attentes à propos des âges moyens pour chaque catégorie étaient aux alentours de 40 ans, mais la réalité n'est que loin de nos estimations, la distribution était comme la suivante :

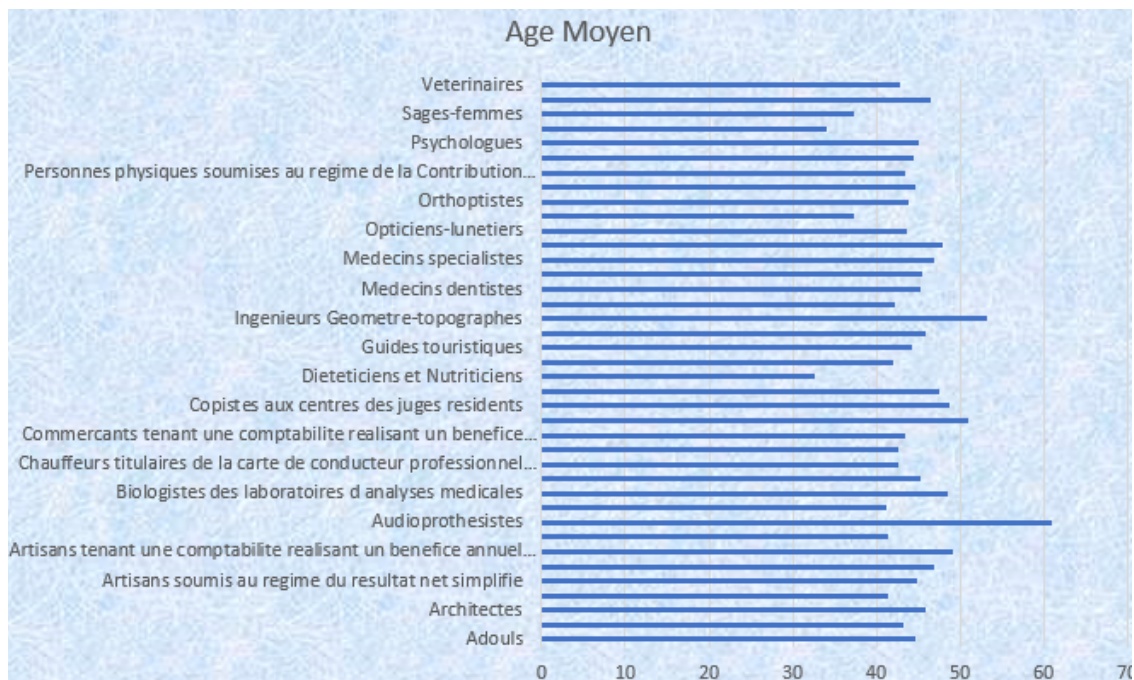


FIGURE 2.2 – Âge moyen par catégorie pour les personnes moins de 65 ans

La remarque qui induit une longue procédure de réflexion est la suivante : un âge moyen élevé est explicable pour les professions dont la formation peut durer une longue durée comme la médecine qui prend au moins 7 ans, mais pour d'autres, la seule raison qui peut expliquer d'un point de vue générale ce phénomène est l'intégration incomplète de tous les actifs de chaque catégorie.

Remarque :

les moyennes d'âge calculées sont toutes calculées pour les personnes moins de 65 ans, alors on doit garder en compte que l'âge moyen pour toute la population peut être plus grand qu'illustré.

Pour l'âge moyen globale, il est distribué de la façon suivante :

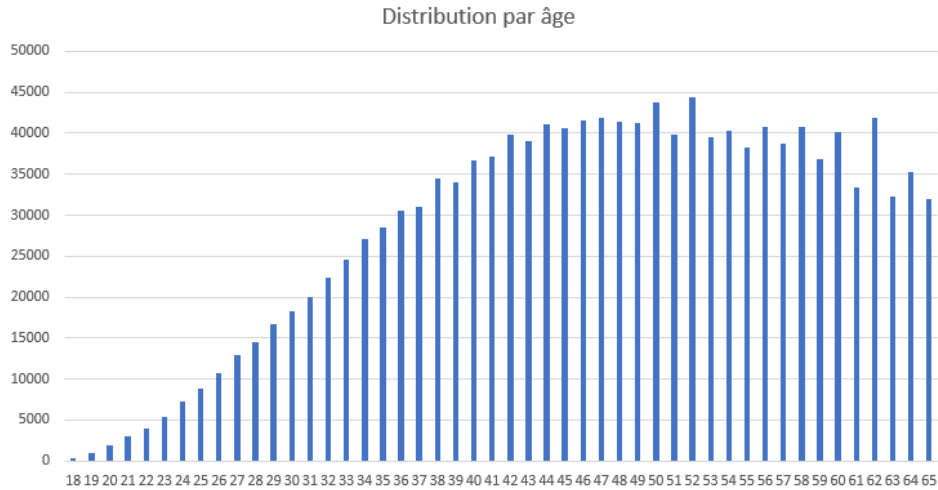


FIGURE 2.3 – Âge moyen global par catégorie

Alors le moyen sera concentrer autour de 47 ans.

Tandis que la distribution par âge apparaît moins concentrée autour d'une seule valeur, on ne peut pas dire la même chose pour les catégories.

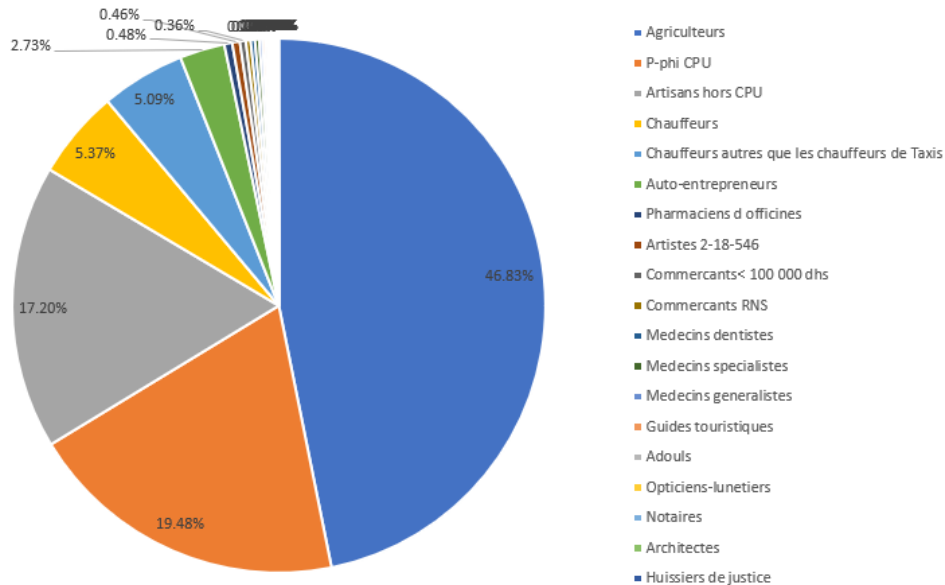


FIGURE 2.4 – Distrubution de la population par catégorie

Les agriculteurs constituent presque la moitié de toute la population sans même qu'elle soit complète. le nombre complet estimé pour la catégorie des agriculteurs est de 800

000 individus alors que le nombre estimé pour toute la population constitue presque 3.4 Millions de personnes. pour cette raison qu'on confirmera l'importance de cette catégorie pour le régime.

Enfin pour le facteur qui peut avoir un influence fort sur le système grâce à la possibilité d'acquérir plus de points que précis dans le tarif : Les Revenues Tarifaires :

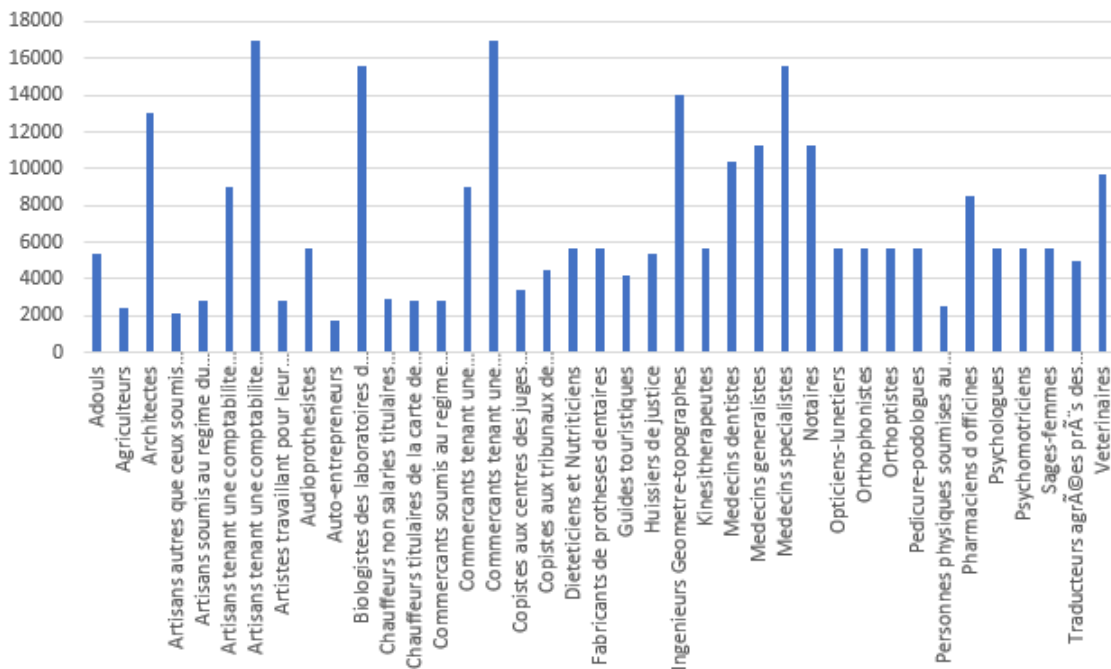


FIGURE 2.5 – Revenus par catégorie

La chose la plus remarquable est le revenu moyen de la catégorie la plus importante du système; les agriculteurs. Il est presque le plus faible parmi toutes les catégories. Alors, au cas où le tarif pour un point est élevé voyant que l'âge moyen est élevé (44 ans), le gouvernement doit intervenir en donnant une subvention à cette catégorie ou pour une solution plus permanente, elle doit investir de plus ou changer son approche envers ses investissements dans ce secteur pour augmenter le revenu de cette catégorie.

Chapitre 3

Tarification et automatisation sous R shiny

3.1 Tarification

3.1.1 Hypothèses :

Le premier pas de toute étude est de fixer les hypothèses. Pour notre étude, les hypothèses seront ceux qui auront un effet immédiat sur le calcul du tarif.

— Âge

Il est sûrement évident que l'âge peut différer d'un individu à un autre même s'ils appartiennent à la même catégorie. Mais pour garder le sens de solidarité entre la population, on définira un seul tarif pour les personnes ayant les mêmes caractéristiques, comme leurs espérances de vie, qui est directement ou indirectement touché par leurs activités, possiblement leur niveau d'éducation et enfin leurs revenus forfaitaires. La caractéristique qui peut englober ses qualifications est la catégorie socio-professionnelle. En choisissant cette manière d'identification, on peut aussi mieux cibler les individus concernés. Cette proposition est valable seulement si nous avons le choix de ne pas garder la solidarité entre toutes les parties de la population.

Pour cette raison, on utilisera l'âge moyen de toute la population pour définir le Tarif.

— Cotisation

Selon la réglementation, la cotisation est définie par un revenu forfaitaire prédéfini par la loi pour chaque catégorie. À titre d'exemple, pour les Aadouls, elle est définie comme 10% de 190% du Smig. Mais, afin de mieux illustrer la situation et tarifier la pension de façon plus précise, on utilisera la base de données précitée puisqu'elle fournit une idée sur la revenue Tarifaire de chaque catégorie. Pour renforcer l'éligibilité de notre proposition, on doit mentionner que pour certaines catégories comme les agriculteurs, on trouve des

différences énormes entre leurs revenus. Alors un revenu forfaitaire moyen pour chaque catégorie peut mieux représenter toute la population de cette catégorie qu'un revenu défini par la loi.

— **Taux Technique :**

Pour la raison de quasi-stabilité de cette valeur aux cours du temps et puisqu'elle est définie par un décret de façon régulière, on a décidé de la fixer au niveau de 3.5%.

— **Table de Mortalité :**

La table de mortalité est aussi fixer par la réglementation, c'est la Table TV 88-90.

— **Taux d'évolution de la Valeur d'Acquisition et de Liquidation :**

Rappelons que la valeur d'acquisition est la valeur par laquelle un cotisant peut acquérir un point. En bref, C'est la valeur qu'on vise à préciser.

Le taux d'évolution de cette valeur est défini pour qu'elle soit stable pendant les 5 premières années, augmentée de 1% chaque année pour les 5 années suivantes et enfin augmentée de 2% après. Cette méthode d'évolution est définie pour prendre en compte la stabilité du système et la durée dans laquelle il peut absorber toute la population concernée. En effet, le système ne sera stable qu'après un nombre d'années à cause du nombre faible d'adhésion et le manque partiel ou complet d'informations quant à certaines catégories. Ces empêchements rendront le tarif plus chère pour les catégories intégrant le régime à un temps ultérieur si le taux d'évolution commence de 2%, sauf que maintenant la différence ne sera pas énorme.

— **Taux d'évolution des revenus :**

Le Taux d'évolution des revenus est directement lié à la valeur des cotisations. Alors il est défini au niveau de 2% annuellement.

3.1.2 Engagement des Assurés :

L'assuré est engagé à payer une cotisation minimale dès la date de son adhésion jusqu'à la date de sa retraite. cette cotisation est convertie en points. Alors, la formule qui décrit cet engagement est la suivante :

$$\text{Engagement(assuré)} = \sum_{k=0}^{n-1} \text{cotisations}_k * (1 + i)^{-k} * {}_k p_x = \sum_{k=0}^{n-1} C_k * (1 + i)^{-k} * {}_k p_x$$

$$\text{Engagement(assuré)} = \sum_{k=0}^{n-1} C * (1 + g)^k * (1 + i)^{-k} * {}_k p_x$$

$$\text{on a } VA_k = VA_{k-1} * \frac{R_{k-1}}{R_{k-2}}$$

Avec :

R_k : le niveau du revenu moyen à l'année k, il définit le taux d'évolution de la valeur d'acquisition comme pré-mentionné.

g : Taux d'évolution des salaires.

i : Le Taux technique

${}_k p_x$: La probabilité qu'une personne âgée de x années survie pour les k prochaines années.

n : La durée entre l'âge x et l'âge de la retraite.

3.1.3 Engagement de l'assureur :

La pension garantis pour un assuré est définie comme le nombre de points cumulés jusqu'à l'âge de la retraite multiplié par la valeur de liquidation à cette date :

$$P = (\sum_{k=0}^{n-1} NP_k) * VL_n$$

Cet engagement est défini selon quatre cas principaux

1. L'assuré est encore vivant à l'âge de la retraite.
2. L'assuré décède après l'âge de la retraite laissant un conjoint/une conjointe et des enfants qui peuvent être en bonne santé et âgés de moins de 21 ans ou handicapé.
3. L'assuré décède avant l'âge de la retraite avec un conjoint/une conjointe et des enfants qui peuvent être en bonne santé et âgés de moins de 21 ans ou handicapé.
4. L'assuré décède avant l'âge de la retraite avec un conjoint/une conjointe et des enfants mais sans que son solde atteigne un nombre minimum de points.

Alors on aura un engagement envers l'assuré, son conjoint/sa conjointe et les enfants en cas de décès.

Pour les engagements reliés au décès de l'assuré avant la retraite, on procédera par une somme pondérée des engagements de l'assureur en cas de décès de l'assuré dans une année précise pondérés par la probabilité de décès de l'assuré à cette année.

Engagement de l'assureur envers l'assuré :

$$Eng_{assureur,assur} = \sum_{k=0}^{x_w - x_r - 1} P * \left(\frac{1}{1+i}\right)^{k+n} * {}_{k+n}p_x$$

$$Eng_{assureur,assur} = P * {}_{65-x}| \ddot{a}_x$$

Avec :

${}_{65-x}| \ddot{a}_x$ Une rente différée de 65-x années du début de période qui paye un dirham en début d'année tant et aussi longtemps que le rentier est vivant, mais pas durant les 65-x années.

Engagement de la pension du conjoint cas de décès après l'âge de la retraite :

la pension du conjoint constitue 50% de la pension de l'assurée :

$$Eng_{conjoint,APR}(y, x) = 50\% * \sum_{k=0}^{w-n-1} P * p_{m,x+k+n} * {}_{k+n}p_y * {}_{k+n}q_x * \left(\frac{1}{1+i}\right)^{k+n}$$

$$Eng_{conjoint,APR}(y, x) = 50\% * P * {}_{65-x}|\ddot{C}_{x|y}$$

Avec :

${}_{65-x}|\ddot{C}_{x|y}$ Une rente réversible avec différé 65-x année qui prévoit le versement de un dirhame en début d'année mais pas durant les premiers 65-x années au statut (y) à compter de la déchéance du statut (x) pondéré par la probabilité de mariage.

y : l'âge du conjoint, défini par une différence d'âge moyenne avec l'assurée (pris comme 5 ans).

$p_{m,x+k+n}$: Probabilité d'avoir un conjoint à l'âge x+k+n

${}_{k+n}q_x$: Probabilité d'un âgé de x de décéder dans les k+n année prochaine.

Engagement de la pension de l'enfant normal, cas de décès après l'âge de la retraite :

la pension d'un enfant normal est presque la même que celle du conjoint sauf qu'elle est fournie tant que l'enfant a moins de 21 ans et elle est réduite d'un pourcentage puisque 50% de la pension sont distribuées sur tous les enfants sans prendre en compte leur état de santé.

$$Eng_{EN,APR}(y, x) = \beta * 50\% * \sum_{k=0}^{21-y-n-1} P * p_{n,x+k+n} * {}_{k+n}p_y * {}_{k+n}q_x * \left(\frac{1}{1+i}\right)^{k+n}$$

$$Eng_{EN,APR}(y, x) = \beta * 50\% * P * {}_{65-x}|\ddot{N}_{x|y}$$

Avec :

${}_{65-x}|\ddot{N}_{x|y}$ Une rente réversible avec différé 65-x année qui prévoit le versement d'un dirham en début d'année mais pas durant les premiers 65-x années au statut (y) jusqu'à 21 ans compter de la déchéance du statut (x) pondéré par la probabilité d'avoir un enfant normal.

$p_{n,x+k+n}$: la Probabilité d'avoir un enfant normal à l'âge x+k+n.

$$\beta = \frac{47}{50}$$

Engagement de la pension e l'enfant handicapé, cas de décès après l'âge de la retraite :

La différence entre la pension d'un enfant normal et un enfant handicapé est que l'enfant handicapé reçoit la pension tant qu'il est vivant.

$$Eng_{EH,APR}(y, x) = (1 - \beta) * 50\% * \sum_{k=0}^{w-n-1} P * p_{h,x+k+n} * {}_{k+n}p_y * {}_{k+n}q_x * \left(\frac{1}{1+i}\right)^{k+n}$$

$$Eng_{EH,APR}(y, x) = (1 - \beta) * 50\% * P * {}_{65-x}|\ddot{H}_{x|y}$$

Avec :

${}_{65-x}|\ddot{H}_{x|y}$ Une rente réversible avec différé 65-x année qui prévoit le versement de un dirhame en début d'année mais pas durant les premiers 65-x années au statut (y) compter de la déchéance du statut (x) pondéré par la probabilité d'avoir un enfant handicapé.

$p_{h,x+k+n}$: la Probabilité d'avoir un enfant handicapé à l'âge x+k+n.

Engagement envers le conjoint en cas de décès avant l'âge de la retraite :

La pension du conjoint en cas de décès avant la retraite diffère de celle après la retraite à cause des cotisations non complètes. Alors on estimera l'engagement par une moyenne pondérée par la probabilité de décès de l'assuré à une année k. Sauf que le nombre de

points ne sera pas stable mais variable et en fonction de l'année de décès k. Alors :

$$Eng_{conjoint,AVR}(y, x) = 50\% * \sum_{k=0}^{n-1} P_k * p_{m,x+k} * {}_k p_x * q_x * \left(\sum_{j=k}^{x_w-y-k} {}_j p_y * \left(\frac{1}{1+i} \right)^j \right)$$

$$Eng_{conjoint,AVR}(y, x) = 50\% * \sum_{k=0}^{n-1} P_k * p_{m,x+k} * {}_k p_x * q_x * {}_{65-x-k} \ddot{a}_{x+k|y+k}$$

Engagement de la pension de l'enfant normal, cas de décès avant l'âge de la retraite :

$$Eng_{EN,AVR}(y, x) = \beta * 50\% * \sum_{k=0}^{n-1} P_k * p_{n,x+k} * {}_k p_x * q_x * \left(\sum_{j=k}^{21-y-k} {}_j p_y * \left(\frac{1}{1+i} \right)^j \right)$$

y ici illustre l'âge de l'enfant normal, attribué par une différence d'âge moyenne entre les personnes en même âge et leurs enfants ; Donc en connaissant l'âge de l'assuré, on peut déduire l'âge de l'enfant.

Engagement de la pension de l'enfant handicapé, cas de décès avant l'âge de la retraite :

$$Eng_{EH,AVR}(y, x) = (1 - \beta) * 50\% * \sum_{k=0}^{n-1} P_k * p_{h,x+k} * {}_k p_x * q_x * \left(\sum_{j=k}^{x_w-y-k} {}_j p_y * \left(\frac{1}{1+i} \right)^j \right)$$

$$Eng_{EH,AVR}(y, x) = (1 - \beta) * 50\% * \sum_{k=0}^{n-1} P_k * p_{h,x+k} * {}_k p_x * q_x * {}_{65-x-k} \ddot{a}_{x+k|y+k}$$

Donc pour résumé, notre objectif est de définir la valeur de la valeur d'acquisition depuis la formule d'équivalence actuarielle. Pour procéder à appeler à cette formule il faut en premier faire apparaître la valeur d'acquisition dans la formule donc notre point de départ sera 'P' : le nombre de points acquis.

on a :

$$P = \sum_{k=0}^{n-1} N P_k * V L_n \text{ et } V L_n = V L_0 * \frac{R_{n-1}}{R_0}$$

Et tant que

$$NP_k = \frac{C_k}{VA_k} = \frac{C*(1+g)^k}{VA_0*\frac{R_{k-1}}{R_0}}$$

Alors :

$$P = \frac{VL_0}{VA_0} * C * \sum_{k=0}^{n-1} (1+g)^k * \frac{R_{n-1}}{R_{k-1}} = \frac{VL_0}{VA_0} * C * CC_n$$

De même on trouve :

$$P_i = \frac{VL_0}{VA_0} * C * \sum_{k=0}^{i-1} (1+g)^k * \frac{R_{i-1}}{R_{k-1}} = \frac{VL_0}{VA_0} * C * CC_i$$

La quantité $\sum_{k=1}^{n-1} (1+g)^k * \frac{R_{n-1}}{R_{k-1}}$ est calculable puisque g est fixé à 2% et le rapport R_n/R_{n-1} est défini avant selon les années écoulées depuis le départ du régime. Donc ce qui nous reste est de définir le rapport $\frac{VL_0}{VA_0}$. Pour ce faire, on fixe au début la valeur de liquidation à 1 et on procède à intégrer le résultat au niveau de la formule d'équivalence actuarielle.

On notera la valeur $C * \sum_{k=1}^{n-1} (1+g)^k * \frac{R_{n-1}}{R_{k-1}}$ par la notation suivante CC_n et CC_i pour sa variable P_i . Enfin la formule d'équivalence actuarielle nous donne :

on applique la formule d'équivalence actuarielle :

$$VAP_{assur} = VAP_{assureur}$$

$$VAP_{assur} = \sum_{k=0}^{n-1} C * (1 + g)^k * (1 + i)^{-k} * {}_k p_x$$

$$\begin{aligned} VAP_{assureur} = & \frac{CC_n}{VA_0} * {}_{65-x}| \ddot{a}_x + 50\% * \frac{CC_n}{VA_0} * {}_{65-x}| \ddot{C}_{x|y} + \beta * 50\% * \frac{CC_n}{VA_0} * {}_{65-x}| \ddot{N}_{x|y} + (1 - \beta) * \\ & 50\% * \frac{CC_n}{VA_0} * {}_{65-x}| \ddot{H}_{x|y} + 50\% * \sum_{k=0}^{n-1} \frac{CC_k}{VA_0} * p_{m,x+k} * {}_k p_x * q_x * {}_{65-x-k}| \ddot{a}_{x+k|y+k} + \beta * 50\% * \\ & \sum_{k=0}^{n-1} \frac{CC_k}{VA_0} * p_{n,x+k} * {}_k p_x * q_x * \left(\sum_{j=k}^{21-y-k} {}_j p_y * \left(\frac{1}{1+i} \right)^j \right) + (1 - \beta) * 50\% * \sum_{k=0}^{n-1} \frac{CC_k}{VA_0} * p_{h,x+k} * \\ & {}_k p_x * q_x * {}_{65-x-k}| \ddot{a}_{x+k|y+k} \end{aligned}$$

Alors :

$$VA_0 = \frac{D}{N}$$

$$\begin{aligned} D = & \left[CC_n * {}_{65-x}| \ddot{a}_x + 50\% * CC_n * {}_{65-x}| \ddot{C}_{x|y} + \beta * 50\% * CC_n * {}_{65-x}| \ddot{N}_{x|y} + (1 - \beta) * 50\% * \right. \\ & CC_n * {}_{65-x}| \ddot{H}_{x|y} + 50\% * \sum_{k=0}^{n-1} CC_k * p_{m,x+k} * {}_k p_x * q_x * {}_{65-x-k}| \ddot{a}_{x+k|y+k} + \beta * 50\% * \\ & \left. \sum_{k=0}^{n-1} CC_k * p_{n,x+k} * {}_k p_x * q_x * \sum_{j=k}^{21-y-k} {}_j p_y * \left(\frac{1}{1+i} \right)^j + (1 - \beta) * 50\% * \sum_{k=0}^{n-1} CC_k * p_{h,x+k} * \right. \\ & \left. {}_k p_x * q_x * {}_{65-x-k}| \ddot{a}_{x+k|y+k} \right] \end{aligned}$$

$$N = \left[\sum_{k=0}^{n-1} C * (1 + g)^k * (1 + i)^{-k} * {}_k p_x \right]$$

3.1.4 Automatisation :

L'automatisation de cette partie de tarification est, comme pour toute automatisation sur cette étude, effectuée à l'aide de R shiny.

Nous avons la formule finale de la valeur d'acquisition, donc pour faire l'automatisation il suffit de créer des fonctions séparées sous R pour calculer chaque composant actuariel et introduire la table des probabilités, puis commencer l'automatisation.

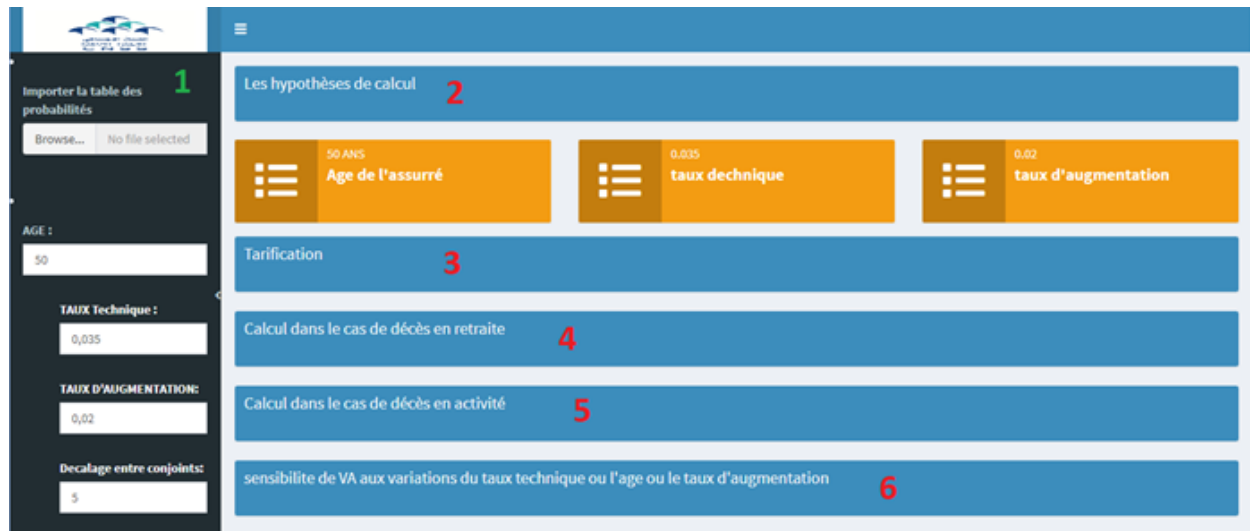


FIGURE 3.1 – Interface de la tarification sous R shiny

Cette automatisation est réalisée afin de répondre à plusieurs objectifs dont les importants sont :

- Tests de sensibilité de la valeur d'acquisition envers chaque paramètre
- La base de données à importer peut être dynamique, en effet, au cours des années, on aura des autres bases de données contenant d'autres informations, et le fait d'importer notre data dans l'interface répondre à ce problème.
- introduit une sorte de dynamité en terme de travail sur un serveur ou un serveur locale.
- visualisation simple et claire des résultats
- Cotisation annuelle initiale : cette case pour calculer la rente soit dans le cas de décès avant retraite ou après retraite

- Anticipation ou Prorogation : cette case pour insérer l'âge de retraite (entre 60 et 75) et dans ce cas la Pension de vieillesse sera multiplier par un facteur Z par âge.
- âge de décès avant retraite : cette partie pour calculer la pension des reversaires d'un assuré qui décéder avant retraite.

Comme vous constatez, la figure ci-dessus a des zones numérotées. Cette numérotation serve comme indicateurs au descriptif suivant :

Zone 1 :

Dans cette partie on peut importer une table contenant la base de données des probabilités par âge, Probabilités de mariage, Probabilités d'avoir un enfant normal ou handicapé, les décalages d'âges....

on a aussi des cases pour saisir les hypothèses dont la modification est toujours possible comme :

- Âge : dans cette case on va saisir l'Âge moyenne de notre population, dans notre cas la valeur est : 47 ans
- Taux technique, nous avons déjà fixe ce taux dans la partie théorique a 3.5
- Taux d'augmentation des revenus : ce taux est fixe à 2
- Décalage entre conjointes : 5 ans
- âge de décès avant retraite : cette partie pour calculer la pension des reversaires d'un assuré qui décède avant retraite.

Zone 2 :

Cette partie est un simple affichage des hypothèses de calcul, à savoir l'âge et le taux technique ainsi le taux d'augmentation.

Zone 3 :

Dans cette partie on va afficher les résultats de notre tarification, à savoir la valeur d'acquisition et la valeur de liquidation.

Zone 4 :

Dans cette partie on va afficher quelques résultats du calcul dans le cas de décès en retraite comme le nombre de points acquis en retraite, la pension de vieillesse dans le cas normal (âge 65) et La pension de vieillesse dans d'anticipation ou Prorogation.

Zone 5 :

Dans cette partie on va afficher quelques résultats du Calcul dans le cas de décès avant retraite comme nombre de points acquis a l'âge de décès et La pension de vieillesse.

Zone 6 :

Cette partie pour étudier la sensibilité de la valeur d'acquisition aux variations du taux technique ou l'âge ou le taux d'augmentation

- Graphe 1 : la valeur d'acquisition en fonction de l'âge
- Graphe 2 : la valeur d'acquisition en fonction du taux technique
- Graphe 3 : la valeur d'acquisition en fonction du taux d'augmentation
- Graphe 4 : Nombre de points acquis en fonction de la cotisation annuelle initiale

Résultats de notre interface :

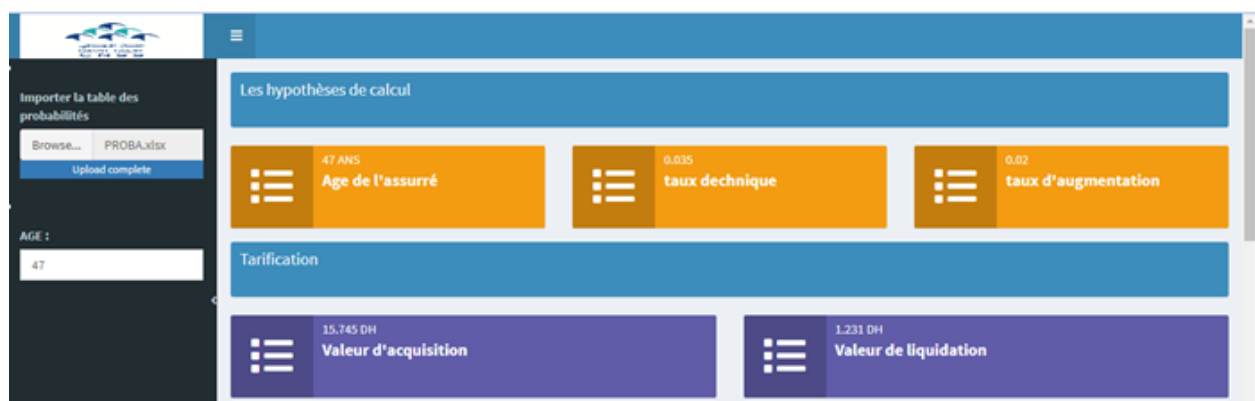


FIGURE 3.2 – Interface d'affichage des résultats de la tarification

On a une valeur d'acquisition égale a : 15.74 DH. C'est-à-dire la valeur d'achat d'un point lorsque le système démarre est égale à 15.74 DH.

On a une valeur de liquidation de 1.23 DH. c'est-à-dire le nombre de points acquis en retraite sera liquidé par 1.23DH par pointe.

Pour l'augmentation de VA et VL selon les années on a le tableau suivant :

Années	1--5	6--10	11+
augmenatation VA	0%	1%	2%
augmenatation VL	0%	1%	2%

FIGURE 3.3 – Taux d'Augmentation ed la VA et la VL au cours des années

Résultats concernant le calcul :



FIGURE 3.4 – Calcul de la rente - Interface

Supposons que pour une cotisation initiale annuelle de 5700DH pour un assuré d'âge 47 et un âge d'anticipation de 60 ans, la Pension de Vieilleses est 9428.66 DH

la Pension de Vieillesse en cas d'anticipation à l'âge de 60 ans

$$R_{A_{anticipation}} = R * Z_{60} = 8014.36 \text{ DH}$$

Le tableau et le graphe suivant résument les taux d'anticipation ou de prorogation selon la réglementation :

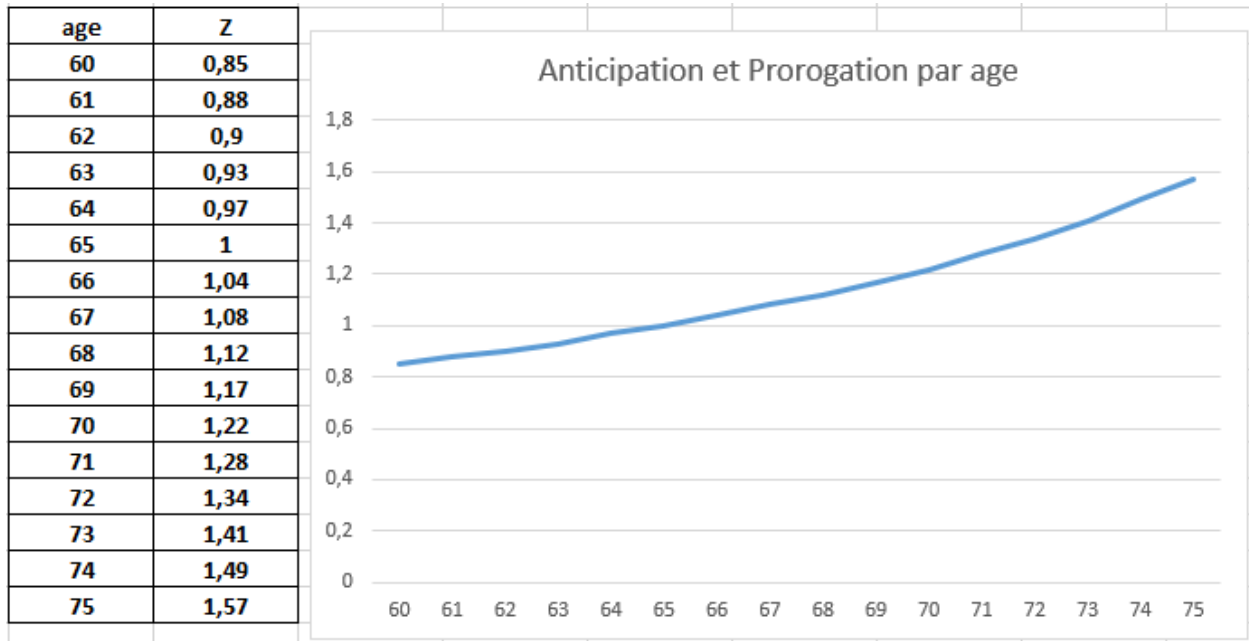


FIGURE 3.5 – Taux d'Anticipation ou de Prorogation par âge

La méthode de calcul de la pension en cas de décès en activité est presque la même sauf que le nombre de points acquis est naturellement réduit au cumule atteint à la date du décès. À titre d'exemple, pour un cotisant ayant :

- âge de 47 ans.
- Une cotisation initiale annuelle de 5700DH
- Un âge de décès 57 ans

La Pension de Vieillesse est :

$$R = \left(\sum_{i=0}^{10} NP_i \right) \times VL'$$

$$\text{avec } VL' = \prod_{j=0}^9 (1 + \gamma_j)$$

$$NP_i = \frac{\text{cot } i}{VA_i} = \frac{5700}{VA_0} \cdot \frac{(1+2\%)^{i-1}}{\prod_{j=0}^{i-1} (1+\gamma_j)}$$

$$\text{car : } VA_i = VA_0 \times \prod_{j=0}^{i-1} (1 + \gamma_j)$$

$$\text{avec } \gamma_j = \begin{cases} 0 & \text{si } 0 \leq j < 5 \\ 1\% & \text{si } 5 \leq j < 10 \\ 2\% & \text{si } 10 \leq j \end{cases}$$

$$\text{et } R = \frac{5700}{15,745} \times \sum_{i=0}^{10} \frac{(1+2\%)^{i-1}}{\prod_{j=0}^{i-1} (1+\gamma_j)} = 5388.72\text{DH}$$

3.1.5 Étude de Sensibilité :

Sensibilité envers la variable âge :

La valeur d'acquisition est par défaut dépendante de l'âge de l'assuré. Alors une étude de sensibilité est importante pour modéliser comment elle varie en fonction de l'âge de l'assuré et plus précisément si la structure de la formule de tarification est complexe dont la prestation est ainsi variable selon plusieurs cas. Voici un graphique illustrant ce type de relation.

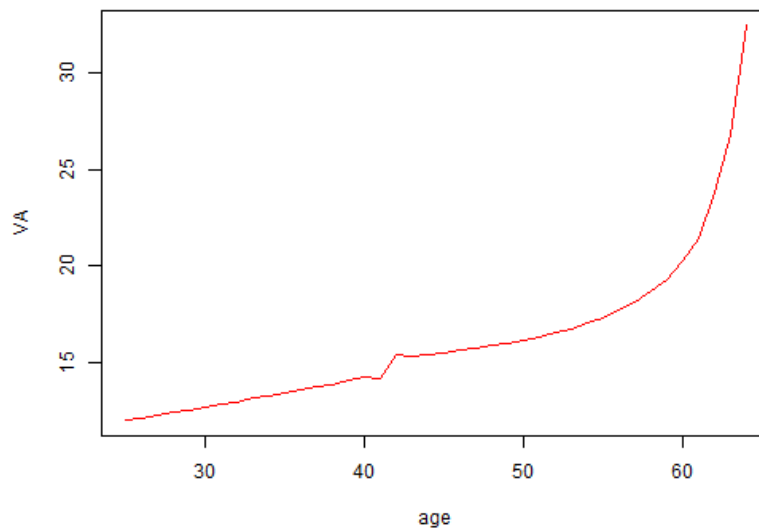


FIGURE 3.6 – VA en fonction de l'âge

Le tarif est bien évidemment croissant en fonction de l'âge, ce qui n'est que raisonnable. En effet, une différence d'une unité d'âge n'a pas de grand impact quant à la probabilité de survie d'un assuré à un âge ultérieur dépassant l'âge de la retraite. Mais il a un impact plus important sur les réserves qui garantissent le paiement des pensions, puisque plus de réserve sont construites avec une autre année. Ce qui explique cette nature croissante.

Sensibilité envers le Taux technique :

Comme indiqué précédemment, le taux technique est précisé par un décret de façon régulière, mais en vue de l'inflation que le monde a connue cette année, une étude de sensibilité envers cette variable apparaît plus importante puisqu'elle représente principalement le changement de valeur d'une même quantité de monnaie entre cette année et une année future.

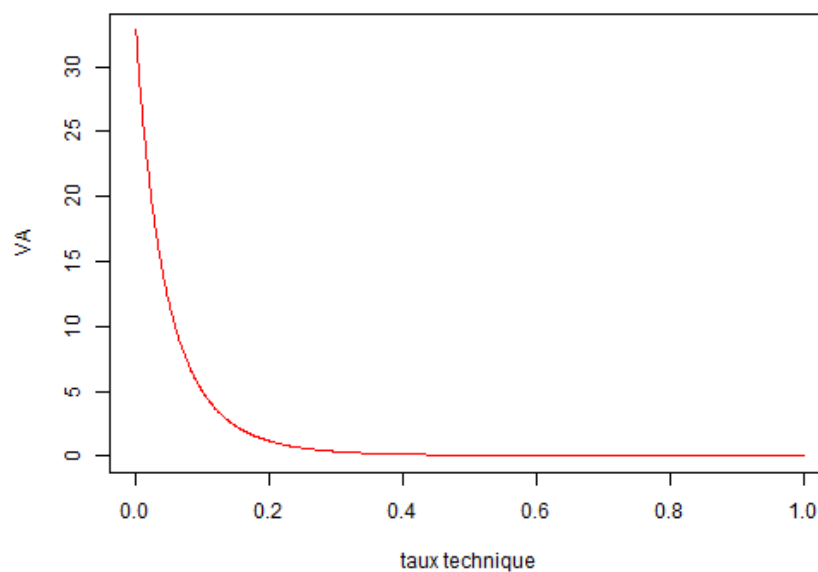


FIGURE 3.7 – VA en fonction du Taux Technique

La formulation trouvée dans la partie de la tarification montre que la valeur d'acquisition diminue avec le taux technique.

Sensibilité envers le Taux d'augmentation des salaires :

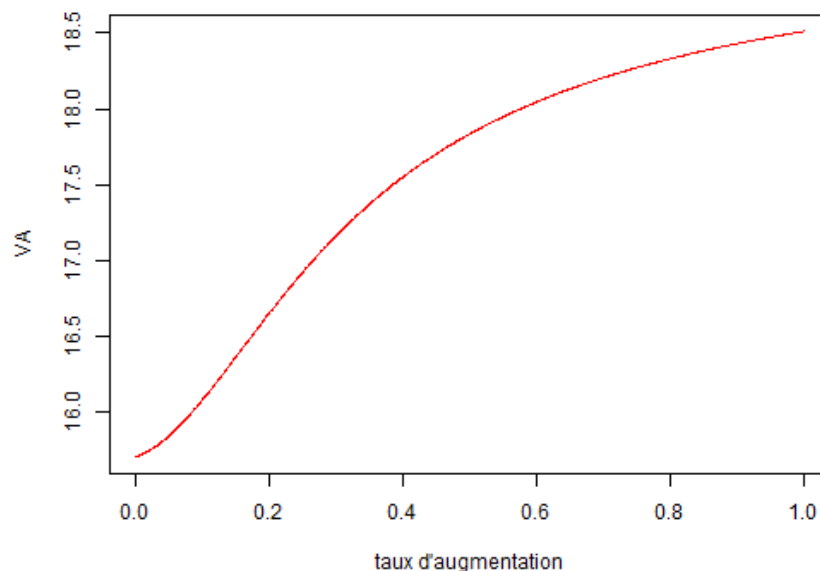


FIGURE 3.8 – VA en fonction du taux d'augmentation

Rappelons que la prestation est fortement liée à la valeur de la cotisation. Alors une augmentation plus forte implique une pension plus grande et par conséquent un tarif plus grand.

Chapitre **4**

Courbe des taux

4.0.1 Introduction :

La projection d'un régime de retraite doit être faite sur un laps de temps qui peut durer une dizaine. Mais par effet d'inflation, les valeurs futures ne seront jamais égales à ceux d'aujourd'hui. Pour résoudre ce problème, nous devons modéliser les taux d'intérêt long terme. Alors, pour cette partie, on se concentrera sur une modélisation d'une courbe des taux zéro-coupons à partir des taux courts.

Pour se faire, on introduira deux méthodes principales, une stochastique suivant les deux fameux modèles de Vasicek et le modèle de CIR. L'autre méthode sera une basée sur les réseaux de neurones artificiels, dont laquelle on comparera ses résultats avec les méthodes traditionnelles dans la modélisation des taux courts en introduisant aussi des résultats dont l'application peut bénéficier ce type de problèmes.

4.0.2 Modèles stochastiques :

Le Modèle de Vasicek :

Le modèle de Vasicek (1977) est fondé sur l'idée que le taux court est le seul paramètre qui régit les prix des titres obligataires, que celui-ci est censé suivre un processus d'Ornstein-Uhlenbeck dont l'évolution stochastique obéit à l'équation différentielle de type :

$$dr_t = a(b - r_t)dt + \sigma dW_t$$

où a, b et σ sont des constantes :

1. b : représente la moyenne à long terme du taux court.
2. a : représente la vitesse d'ajustement du taux court actuel vers sa moyenne de long terme b .
3. σ : représente l'Ecart-type du changement instantané de $r(t)$.

4. W_t : Un Mouvement brownien standard

On applique le lemme d'Ito sur le processus $X_t = r_t e^{-at}$, La solution de l'équation est donnée par :

$$r_t = r_0 e^{-at} + b(1 - e^{-at}) + \sigma e^{-at} \int_0^t e^{as} dW_s$$

Les variables aléatoires r_t suivent une loi normale de paramètres :

$$\mathbb{E}[r_t] = r_0 e^{-at} + b(1 - e^{-at}) \text{ et } \text{Var}[r_t] = \frac{\sigma^2 (1 - e^{-2at})}{2a}$$

Estimation paramètres :

Etape 1 : Discrétisation du modèle

$$r_s = r_t e^{-a(s-t)} + b(1 - e^{-a(s-t)}) + \sigma e^{-as} \int_t^s e^{au} dW_u$$

$$r_{t+1} = r_t e^{-a} + b(1 - e^{-a}) + \sigma \sqrt{\frac{1 - e^{-2a}}{2a}} \varepsilon$$

où ε est une variable aléatoire de loi normale centrée réduite.

Etape 2 : Régression

L'estimation des paramètres du modèle de Vasicek peut s'effectuer avec la méthode des moindres carrés ordinaires. Le point de départ de cette méthode consiste à régresser la série des taux courts à l'aide de la formule suivante

$$r_t = b(1 - e^{-a}) + e^{-a} r_{t-1} + \varepsilon_t$$

Avec $\epsilon_t = \mathcal{N}\left(0, \frac{\sigma^2}{2a}(1 - e^{-2a})\right)$ et correspond au résidu de la régression. Cette équation indique que les taux courts s'ajustent suivant un processus autorégressif d'ordre 1, hypothèse où il faut vérifier pour les données utilisées dans l'estimation. Donc : nous pouvons à présent, en utilisant les données discrètes du taux court, estimer les paramètres du modèle de Vasicek via l'équation simplifiée suivante :

$$r_t = \alpha + \beta r_{t-1} + \epsilon_t$$

Avec

$$\alpha = b(1 - e^{-a}), \beta = e^{-a} \text{ et } \epsilon_t \sim N\left(0, \frac{\sigma^2}{2a}(1 - e^{-2a})\right)$$

Lemme 1 : Supposons que la dynamique du taux d'intérêt instantané sous la probabilité risque-neutre est donnée par

$$dr_t = \mu(t, r_t) dt + \sigma(t, r_t) dW_t$$

Supposons de plus que μ et σ ont la forme suivante :

$$\begin{cases} \mu(t, r) = \alpha(t)r + \beta(t) \\ \sigma(t, r) = \sqrt{\gamma(t)r + \delta(t)} \end{cases}$$

Alors structure de $\{P(t, T); 0 \leq t \leq T\}$ est de la forme :

$$P(t, T) = \exp\{A(t, T) - B(t, T)r_t\}$$

On dit que le modèle possède une structure par terme affine (STA).

Etape 3 : Construction de la courbe des taux

la dynamique du modèle de Vasicek vérifie les conditions du **Lemme 1**, donc le modèle

possède une structure par terme affine (STA) on écrit :

$$P(t, T) = A(t, T) \exp(-B(t, T)r_t)$$

Avec :

$$A(t, T) = \exp \left[\frac{(B(t, T) - T + t) \left(a^2 b^* - \frac{\sigma^2}{2} \right)}{a^2} - \frac{\sigma^2 B(t, T)^2}{4a} \right]$$

$$\text{Et : } b^* = b - q \frac{\sigma}{a}$$

Etape 4 : Estimation de q

Soit n le nombre des maturités \underline{q} est la constante qui minimise l'erreur entre la courbe à coupon zéro de la CNSS et notre $P(t_0, T_i)$ pour toutes les maturités

$$q = \min_x \left\{ \sum_{i=1}^n (R(t_0, T_i, x) - ZC(t_0, T_i))^2 \right\}$$

Avec :

$$R(t, T) = -\frac{\ln P(t, T)}{T-t}$$

Modèle de Cox-Ingersoll-Ross (CIR) :

Le modèle de Cox, Ingersoll et Ross occupe depuis les années 80, une position privilégiée au sein du champ théorique de l'évaluation des actifs financiers. La raison essentielle du succès des travaux de Cox, Ingersoll et Ross tient au fait qu'ils parviennent à nous fournir un modèle parfaitement compatible avec les hypothèses de la théorie standard.

Afin d'éviter l'apparition des taux d'intérêt négatif comme dans le modèle de Vasicek (car il est gaussien), Cox, Ingersoll et Ross (CIR) ont eu recours au processus "racine

carrée".

Dans ce modèle le taux d'intérêt instantané r_t suit l'EDS :

$$dr_t = a(b - r_t)dt + \sigma\sqrt{r_t}dW_t$$

où a, b, σ sont des constantes positives Il est possible de montrer que si $2ab > \sigma^2$, alors, la solution n'atteint pas 0, dans ce cas, le drift de l'équation est suffisamment important pour contraindre la diffusion d'atteindre 0. Nous supposons cette hypothèse qui assure la non-négativité du taux court.

Estimation paramètres :

Etape 1 : Discrétisation du modèle

Le procédé de discrétisation d'Euler consiste, dans le modèle de Cox, Ingersoll et Ross, en l'approximation du processus continu r_t par le processus discret r_t suivant :

$$\tilde{r}_{t+1} = \tilde{r}_t + a^* (b^* - \tilde{r}_t) + \sigma\sqrt{\tilde{r}_t}\varepsilon$$

$$\text{où : } a^* = \alpha\sigma + a \text{ et } b^* = \frac{ab}{a^*}$$

Etape 2 : régression

L'estimation des paramètres du modèle de CIR peut s'effectuer avec la méthode des moindres carrés ordinaires. Le point de départ de cette méthode consiste à régresser la série des taux courts à l'aide de la formule suivante :

$$Y_t = a^* \times b^* X_t + (1 - a^*) Z_t + \sigma\varepsilon_t$$

$$\text{avec } Y_t = \frac{r_t}{\sqrt{r_{t-1}}} \text{ et } X_t = \frac{1}{\sqrt{r_{t-1}}} \text{ et } Z_t = \sqrt{r_{t-1}}$$

Etape 3 : Construction de la courbe des taux

la dynamique du modèle de CIR vérifie les conditions du Lemme 1 , donc le modèle possède une structure par terme affine (STA) on écrit :

$$P(t, T) = A(t, T) \exp(-B(t, T)r_t)$$

Avec :

$$A(t, T) = \left[\frac{2\gamma \exp\left(\frac{(a+\gamma)(T-t)}{2}\right)}{(a+\gamma)(\exp(\gamma(T-t)) - 1) + 2\gamma} \right]^{\frac{2ab}{\sigma^2}}$$

$$B(t, T) = \left[\frac{2 \exp(\gamma(T-t)) - 1}{(a+\gamma)(\exp(\gamma(T-t)) - 1) + 2\gamma} \right]$$

et

$$\gamma = \sqrt{a^2 + 2\sigma^2}$$

Etape 4 : Estimation de a

On a : $a^* = \alpha\sigma + a$ et $b^* = \frac{ab}{a^*}$ et n le nombre des maturités, pour estimer a il faut estimer

α qui minimise l'erreur entre la courbe à coupon zéro de la CNSS et notre $P(t_0, T_i)$ pour toutes les maturités

$$\alpha = \min_x \left\{ \sum_{i=1}^n (R(t_0, T_i, x) - ZC(t_0, T_i))^2 \right\}$$

Avec :

$$R(t, T) = -\frac{\ln P(t, T)}{T-t}$$

Partie pratique :

Présentation des données :

Pour estimer les paramètres souhaités, nous avons choisi comme taux court, le taux moyen pondéré au jour le jour, qui est le taux monétaire au jour le jour du marché interbancaire. Il s'agit d'un taux moyen pondéré par les montants des transactions déclarées, pour toutes les institutions admises sur le marché interbancaire.

Les données journalières du 1er janvier 2016 au 01 janvier 2020, sont au total 1044 observations. Ces données sont représentées sur le graphe suivant :

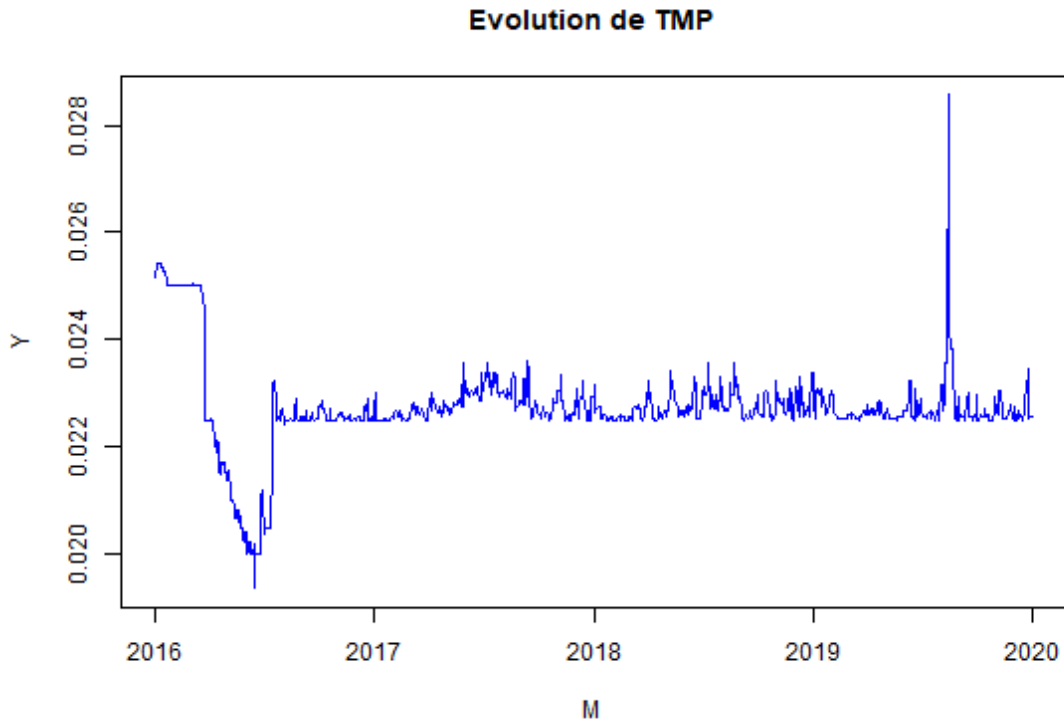


FIGURE 4.1 – Evolution TMP entre 2016 et 2020

Quoi qu'on remarque qu'il y a une sorte de "jump" au début, les taux restent stationnaires. on peut ignorer ces sauts, puisque leur fréquence n'est pas élevée ainsi que la différence introduite en cas de saut n'est pas énorme et n'influe pas notre étude puisqu'elle vise le long terme.

Afin de calibrer les autres paramètres, nous utiliserons la courbe des taux fournis par la CNSS représentée sous le graphique ci-dessous :

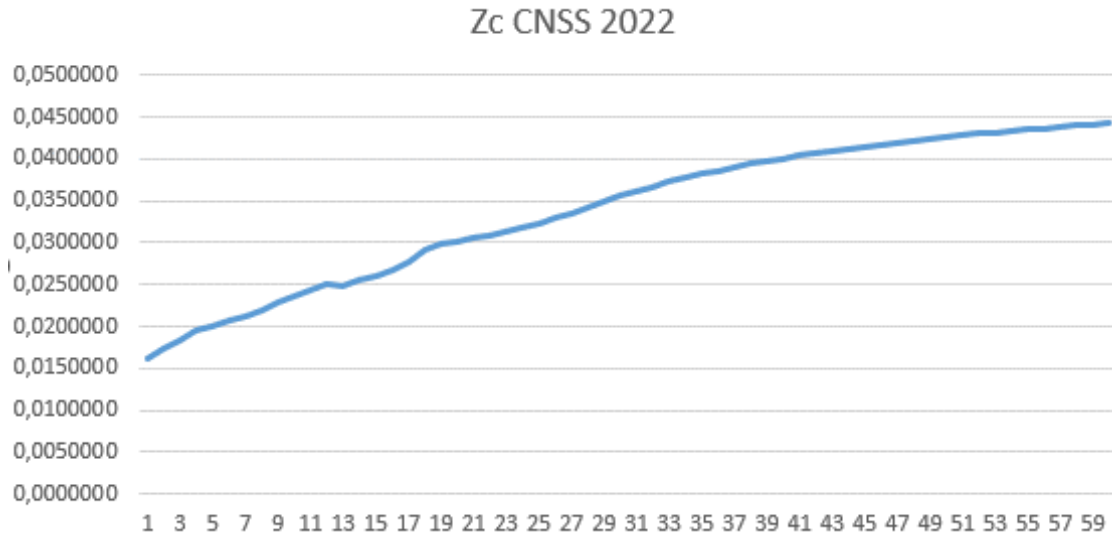


FIGURE 4.2 – Courbe des taux CNSS 2022/01/01

la courbe contient les taux pour 60 maturités ce qui va être suffisant pour estimer les paramètres pour que le modèle soit le plus proche possible à $R(t,T)$.

Avant de Commencer, nous introduisons la base des prérequis d'une analyse des données chronologiques : la stationnarité :

Une stationnarité au sens large signifie tout simplement l'indépendance temporaire de l'espérance de la série, sa variance et sa covariance avec les valeurs précédentes. Les tests souvent utilisés sont les tests de Dickey-Fuller avancé ou le test de Kwiatkowski-Phillips-Schmidt-Shin (KPSS). L'importance de cette propriété réside dans le fait que sans elle l'estimation des paramètres sera totalement biaisée. Alors commençons par le test de Dickey-Fuller.

Teste de Dickey-Fuller :

Notre test consiste à tester l'hypothèse nulle : $\rho = 1$ qui signifie l'existence d'une racine unitaire contre l'hypothèse alternative de $\rho < 1$.

Le test est fait sur R et les résultats sont les suivantes :

```

> T=read_excel('DATA_ESTIMATION.xlsx')
> adf.test(T$TMP)

Augmented Dickey-Fuller Test

data: T$TMP
Dickey-Fuller = -3.9319, Lag order = 10, p-value = 0.0124
alternative hypothesis: stationary

```

FIGURE 4.3 – Test Stationnarité Dickey-fuller

La *p-value* = 0.01 donc on rejette l'hypothèse nulle du test (H_0 : Il y'a non stationnarité).
D'où la stationnarité de la série du taux moyens pondérés.

Estimation : Pour le modèle de **Vasiček** :

Les paramètres estimés sont les suivants :

<i>Paramtres</i>	<i>Valeur</i>
α	0.0619914114
β	0.0226855032
σ	0.0002786217
λ	-0.0019184826

Pour la raison de s'assurer de l'ajustement du modèle, on estimera tout le rang des valeurs qu'on a sur la base des données d'un point de départ (prix comme le premier point de la base de données) en utilisant chaque prédiction \hat{r}_t pour prédire la prochaine valeur \hat{r}_{t+1} . le résultat est présenté ci-dessous.

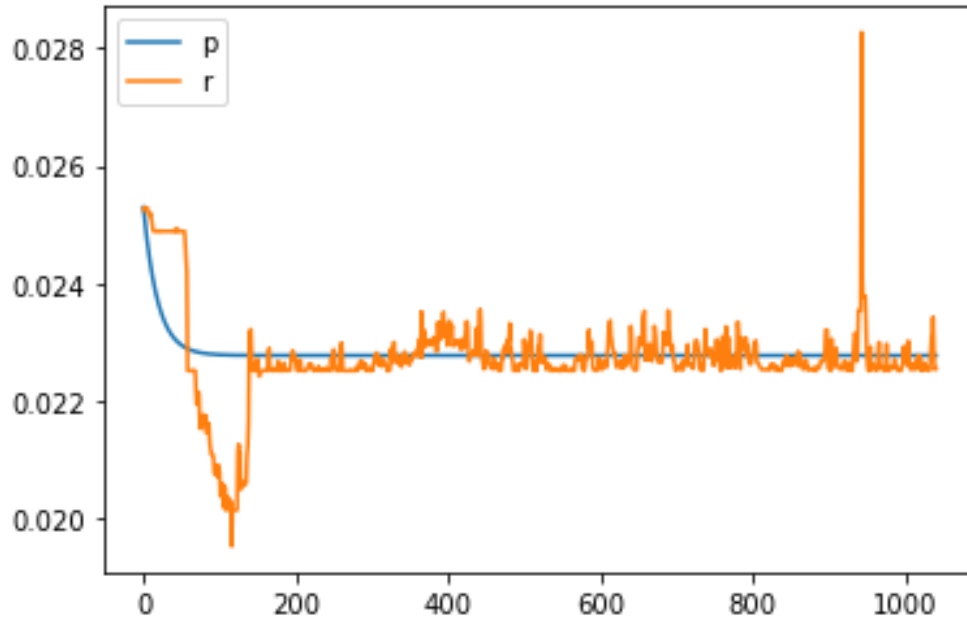


FIGURE 4.4 – Graphique de Comparaison de l'Allure du modèle de Vasicek ajusté sans terme stochastique

Donc les prédictions s'ajustent totalement aux valeurs réelles. En ajoutant le terme stochastique on aura le graphique suivant :

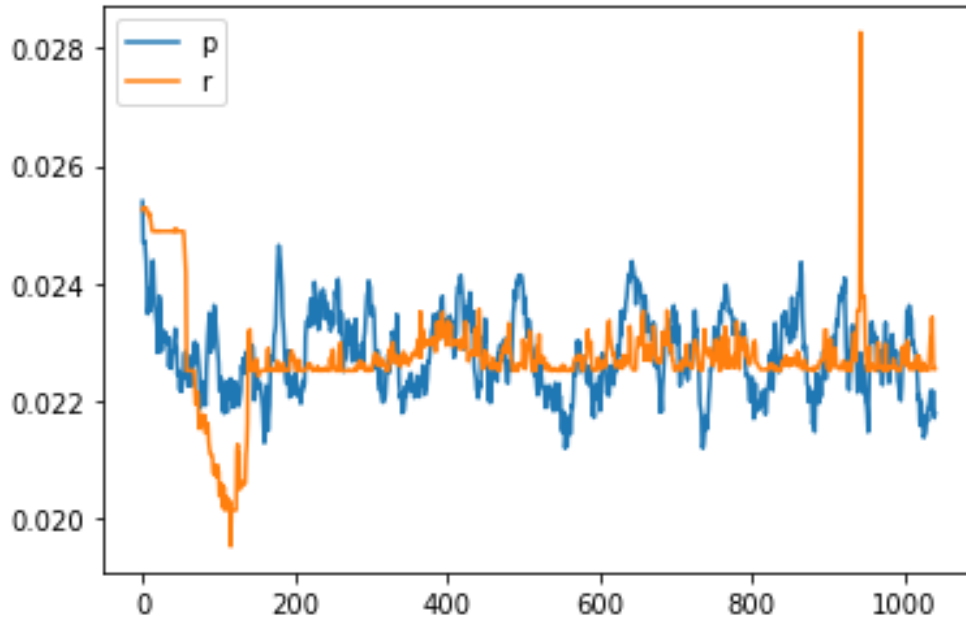


FIGURE 4.5 – Graphique de Comparaison de l’Allure du modèle de Vasicek ajusté avec le terme stochastique

Après chaque estimation des paramètres, les tests qui suivent sont toujours les tests sur les erreurs. en premier Temps on a :

Test de normalité des erreurs :

Pour ce test, on présente les résultats du fameux test de Shapiro-wilk.

```
> shapiro.test(regg$residuals)
      shapiro-wilk normality test
data:  regg$residuals
w = 0.51617, p-value < 2.2e-16
```

FIGURE 4.6 – Test normalité des résidus pour le modèle de Vasicek

La **p-value** = $2e - 16 < 5\%$. Donc on rejette l’hypothèse nulle du test (H_0 : les erreurs suivent une loi normale), d’où la non-normalité des erreurs.

Le deuxième test est celui d’**homoscédasticité** (ARCH) :

Homoscédasticité

La détection de l'hétéroscédasticité par le processus ARCH se fait avec l'hypothèse nulle l'existence d'homoscédasticité. Alors :

```
> ArchTest(regg$residuals)
      ARCH LM-test; Null hypothesis: no ARCH effects
data:  regg$residuals
Chi-squared = 265.51, df = 12, p-value < 2.2e-16
```

FIGURE 4.7 – Test d'homoscédasticité pour le modèle de Vasicek

La $p\text{-value} < 5\%$ donc on rejette l'hypothèse nulle du test (H_0 : Il n'y a pas d'effet ARCH), d'où l'hétéroscédasticité des erreurs.

Enfin le test d'**autocorrélation** des erreurs.

Autocorrélation des Erreurs

Le test qu'on a utilisé est celui de Durbin-Watson, il cherche à vérifier la significativité du coefficient ρ dans la formule :

$$\epsilon_t = \rho\epsilon_{t-1} + u_t$$

Où le ϵ_t est le résidu estimé du modèle et u_t un bruit blanc. L'hypothèse nulle (H_0) stipule qu'il y a non autocorrélation donc $\rho = 0$. Les résultats sont les suivants :

```
> dwtest(regg)
      Durbin-watson test
data:  regg
DW = 2.54, p-value = 1
alternative hypothesis: true autocorrelation is greater than 0
```

FIGURE 4.8 – Test d'Autocorrélation des Erreurs pour le modèle de Vasicek

La **p-value** > 5% alors on accepte l'hypothèse nulle du test (H_0 :), d'où l'absence d'auto-corrélation des erreurs.

Estimation des paramètres du Modèle de CIR

Les paramètres estimée sont les suivants :

<i>Paramtres</i>	<i>Valeur</i>
<i>a</i>	0.0019421733
<i>b</i>	0.6255803036
<i>v</i>	0.0031408385
σ	-0.0017453981
α	29.5793564400

Test d'Hypothèses :

Normalité des résidus :

```
> shapiro.test(regg2$residuals)
      shapiro-wilk normality test
data:  regg2$residuals
w = 0.53146, p-value < 2.2e-16
```

FIGURE 4.9 – Test normalité des résidus pour le modèle CIR

Les résidus ne sont pas normaux puisque la **p-value** est inférieure à 5%.

Homoscédasticité. :

```

> ArchTest(regg2$residuals)

      ARCH LM-test; Null hypothesis: no ARCH effects

data:  regg2$residuals
Chi-squared = 205.83, df = 12, p-value < 2.2e-16

```

FIGURE 4.10 – Test d’homoscédasticité pour le modèle CIR

La **p-value** < 5% donc on rejette l’hypothèse nulle du test (H_0 : Il n’y a pas d’effet ARCH), d’où l’hétéroscédasticité des erreurs

Autocorrélation :

```

> m=length(T$TMP)
> Y=T$TMP[-1]/sqrt(T$TMP[-m])
> X=1/sqrt(T$TMP[-m])
> Z=sqrt(T$TMP[-m])
> regg2=lm(Y~0+X+Z)
> dwtest(regg2)

      Durbin-watson test

data:  regg2
DW = 2.5255, p-value = 1
alternative hypothesis: true autocorrelation is greater than 0

```

FIGURE 4.11 – Test d’Autocorrélation des Erreurs pour le modèle CIR

Le fonctionnement du test de Durbin Watson est déjà introduit de façon détaillée dans la même partie du modèle de Vasicek. Les résultats sont présentés dans la figure ci-dessus où la **p-value** est égale à 1, d’où l’absence d’autocorrélation des erreurs.

Remarque : Malgré le rejet de la plupart des hypothèses pour les deux modèles, la discrétisation de l’équation différentielle stochastique du modèle Vasicek et CIR, nous impose la normalité des erreurs et l’homoscédasticité, qui provient de l’utilisation du mouvement brownien dans les deux modèles. Donc, pour nous, ces hypothèses que nous avons mises à l’essai sont plus des contraintes au modèle que des hypothèses. Pour estimer les

paramètres des deux modèles, nous nous appuyerons sur l'hypothèse selon laquelle les hypothèses de base sont vérifiées.

Ajustement au courbe de Taux :

L'ajustement des modèles se fait par minimisation de la distance entre les valeurs des modèles estimés et les valeurs réelles de la courbe fournis par la CNSS. Les résultats sont les suivants :

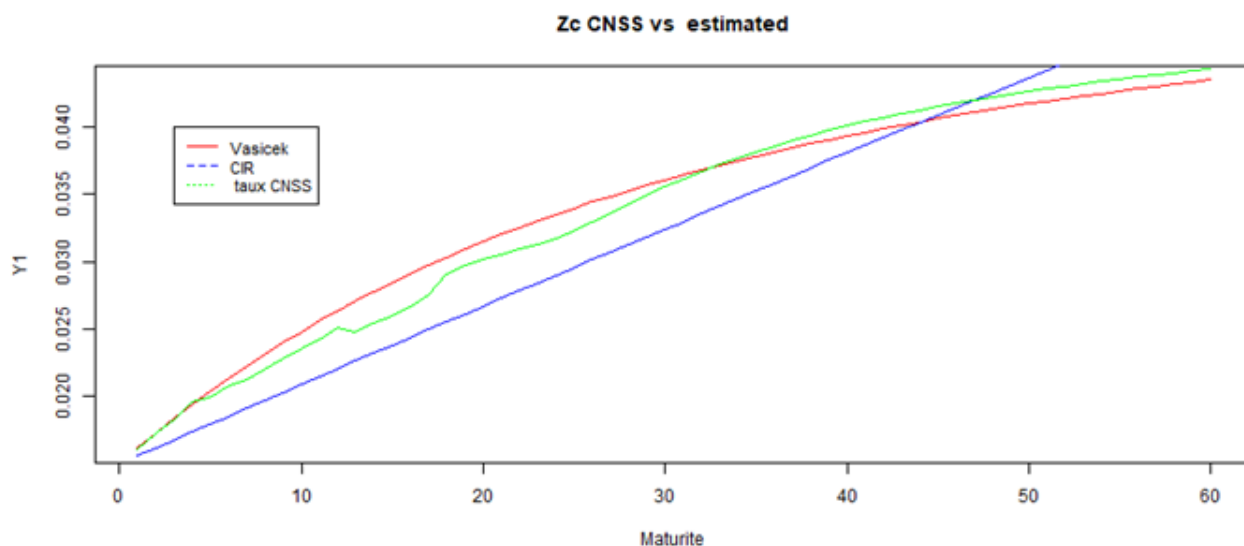


FIGURE 4.12 – Comparaison entre le ZC CNSS et les courbes de taux estimées avec les deux modèles

Le graphe indique que le modèle le plus proche est celui du Vasicek.

Création d'une application pour la construction de courbe des taux

Tant que la théorie ainsi que la procédure de modélisation restent les mêmes, Les données peuvent changer. Alors L'automatisation de la modélisation est faisable et importante . Puisqu'une installation d'un fichier exécutable n'est pas possible, notre automatisation sera sous forme d'une interface web construite par **Rshiny**. La forme sera la suivante :

Construction Courbe des taux



FIGURE 4.13 – Interface Rshiny pour la construction de la courbe des taux

L'interface, comme indiqué par les rectangles rouges, se compose de 7 zones principales, les zones de 1 à 3 sont des zones pour la fournissions des données et les autres pour l'affichage des résultats.

- **zone 1** : une zone faite pour la fournissions des données des taux courts afin de les modéliser et estimer les paramètres. Les données peuvent être sous forme de Taux moyen pondéré ou sous d'autre formes. Elle requis seulement que les deux colonnes soient nommées sous la même nomenclature utilisée et que les données soient triées temporairement en croissance.
- **zone 2 : Zéro-coupons** : une zone pour les données relatives à la courbe des taux utilisée pour la calibration des paramètres. Les contraintes sont les mêmes que celles de la zone précédente.
- **zones 3 : Valeur d'initialisation** : La valeur d'initialisation des projections des taux courts. La valeur que nous avons utilisée est $r_0 = 0.015$.
- La zone 4 est faite pour permettre à l'utilisateur de télécharger les résultats de l'estimation.
- Les zones 5 et 6 :Affichage des résultats

- Paramètres Estimés
 - Graphe représentatif de la différence entre la courbe à coupon zéro et la structure par terme estimé.
 - Tableau récapitulatif des résultats des maturités jusqu'à 60 ans.
- La dernière zone est réservée à la comparaison entre les courbes modélisées est la réelle et pour l'affichage des erreurs de de chaque modèle.

Résultats de notre application

importation data pour l'estimation

Construction Courbe des taux

DATA journalier pour l'estimation

Browse... DATA_ESTIMATION.xlsx

Upload complete

Zero coupon reel

Browse... Zerro coupon CNSS.xlsx

Upload complete

format Excel

Taux le jour d'evaluation

0,015

Download Data Vasicek Download Data CIR

FIGURE 4.14 – Importation

Résultats pour le modèle de Vasicek

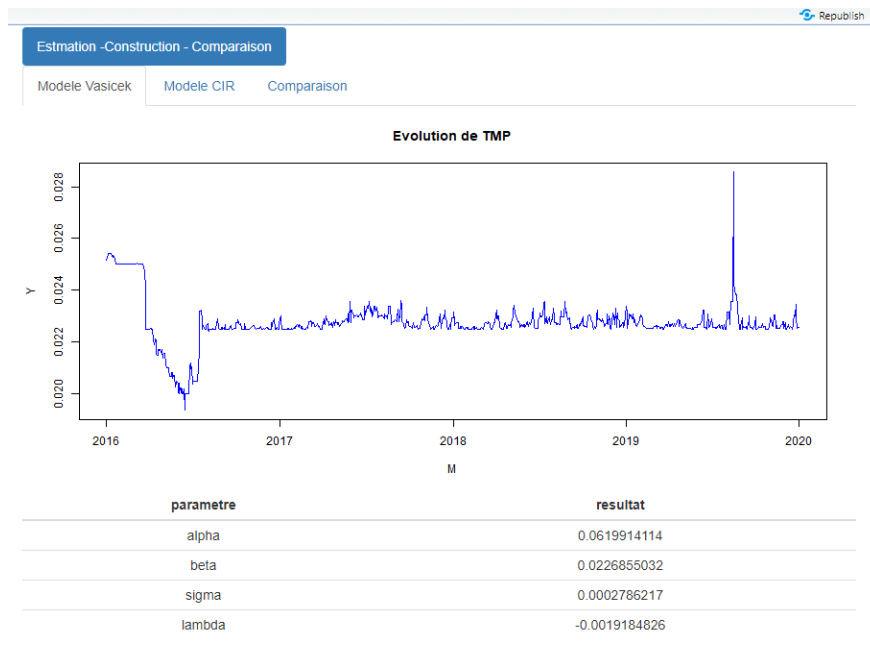


FIGURE 4.15 – Graphe TMP et resultats de l'estimation pour le modèle de Vasicek

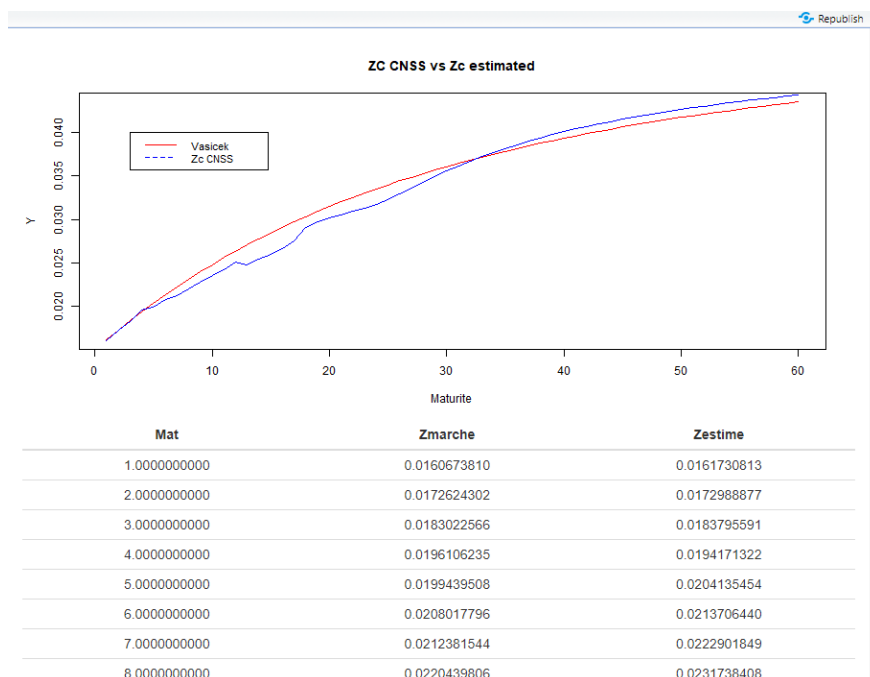


FIGURE 4.16 – Graphe de comparaison et table des taux pour le modèle de Vasicek

Résultats pour le modèle CIR

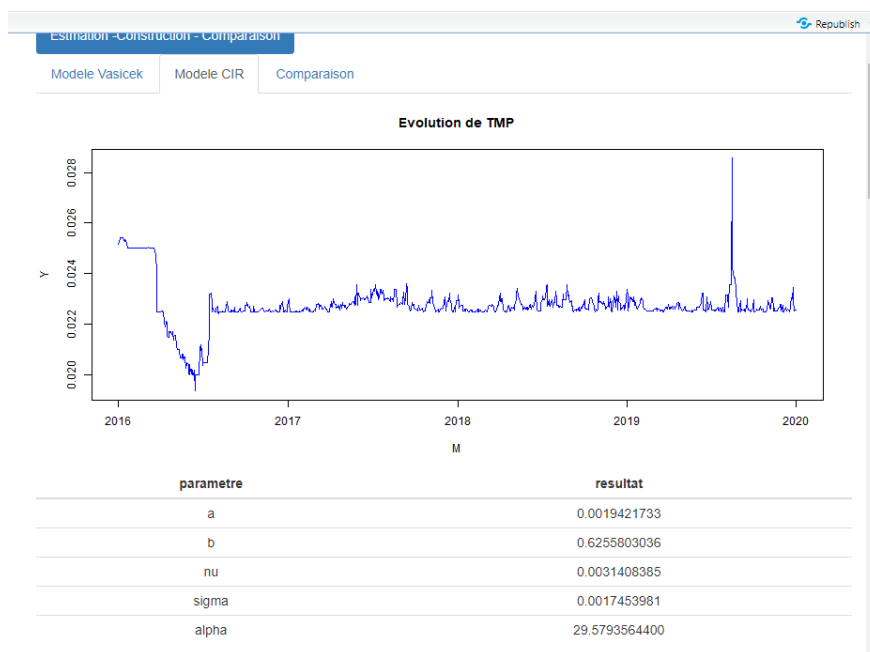


FIGURE 4.17 – Graphe TMP et resultats de l'estimation pour le modèle CIR

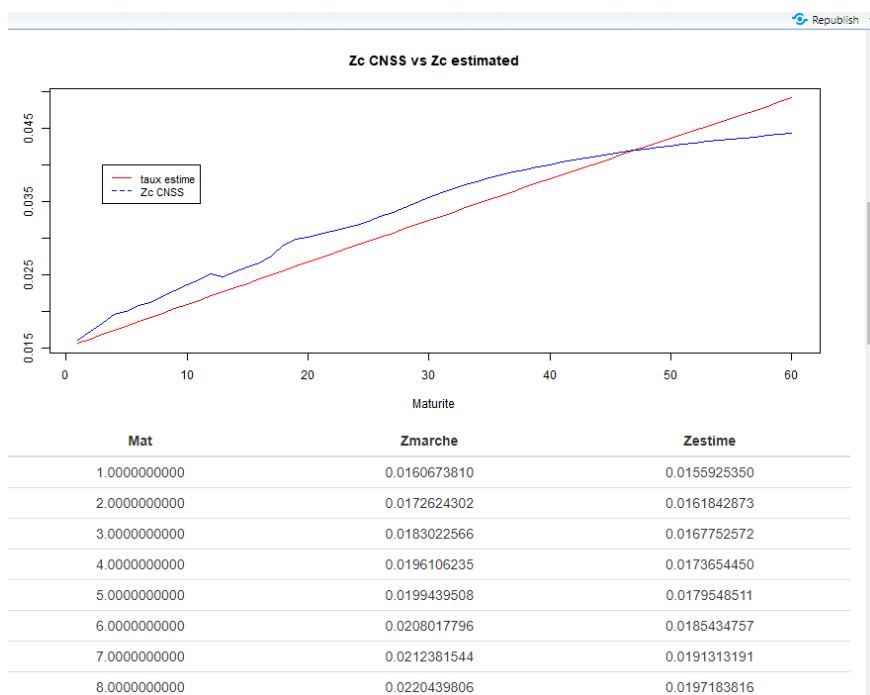


FIGURE 4.18 – Graphe de comparaison et table des taux pour le modèle CIR

Comparaison et téléchargement



FIGURE 4.19 – Graphe de comparaison et table des taux pour les deux modèles

Construction Courbe des taux

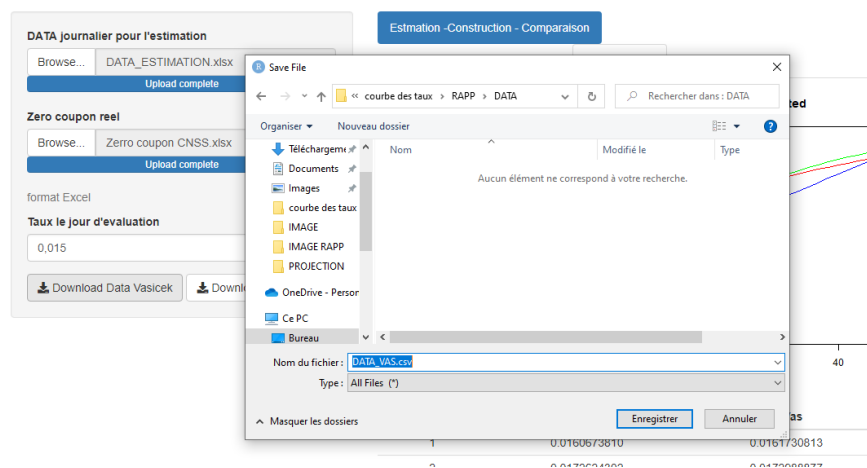


FIGURE 4.20 – illustration de la possibilité de télécharger les données

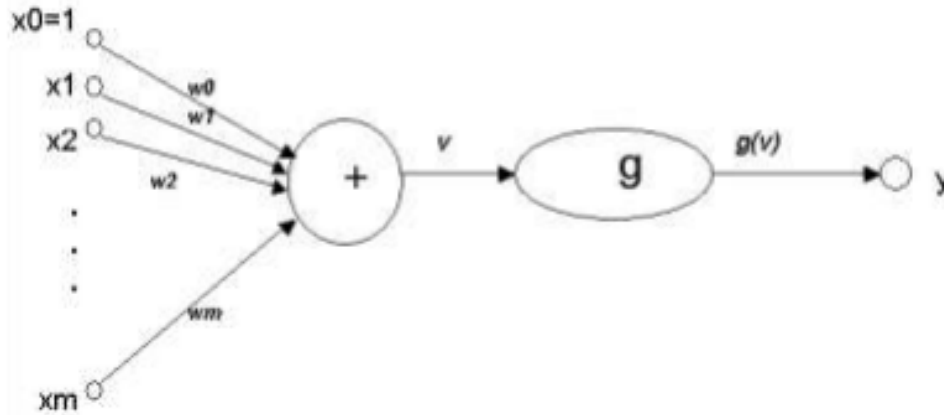
4.0.3 Réseaux de neurones Artificielles Introduction :

les ANN existent depuis un certain temps : ils ont été introduits pour la première fois en 1943 par le neurophysiologiste Warren McCulloch et le mathématicien Walter Pitts. Dans leur article historique intitulé "A Logical Calculus of Ideas Immanent in Nervous Activity", McCulloch et Pitts ont présenté un modèle de calcul simplifié de la façon dont les neurones biologiques pouvaient travailler ensemble dans le cerveau des animaux pour effectuer des calculs complexes en utilisant la logique propositionnelle.. Il s'agissait de la première architecture de réseau neuronal artificiel. Depuis, de nombreuses autres architectures ont été inventées, comme nous allons le voir.

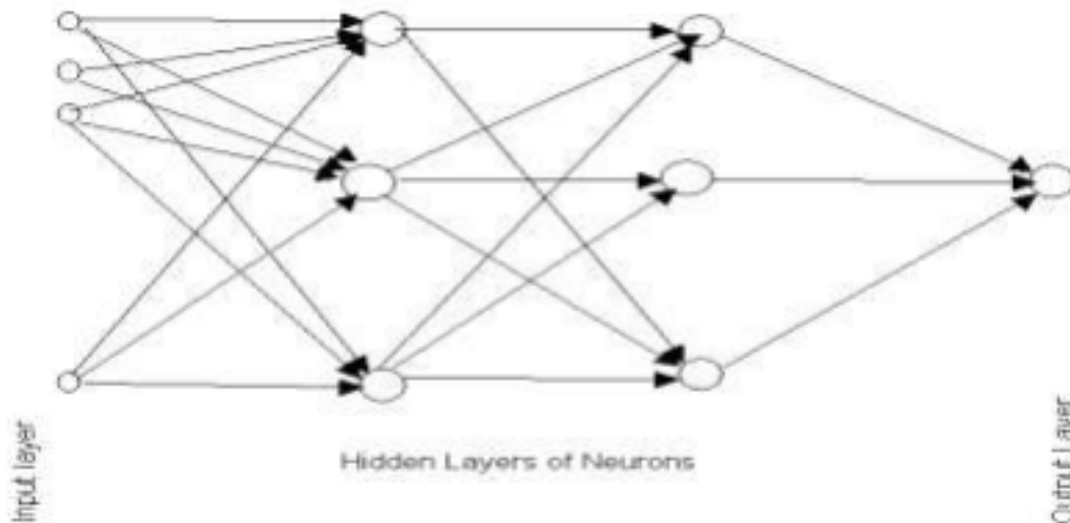
Les réseaux de neurones artificiels ont été inspirés par découvertes biologiques relatives au comportement du cerveau en tant que réseau d'unités appelées neurones. On estime que le cerveau humain compte environ 10 milliards de neurones connectés chacun en moyenne à 10 000 autres neurones. Chaque neurone reçoit des signaux via des synapses qui contrôlent les effets du signal sur le neurone. On pense que ces connexions synaptiques jouent un rôle clé dans le comportement du cerveau. Le bloc de construction fondamentale dans un réseau de neurones artificiels est le modèle mathématique d'un neurone, comme le montre la figure présentée ci-dessous. Les trois composants de base du neurone (artificiel) sont :

1. Les synapses ou liens de connexion qui fournissent des poids, w_j , aux valeurs d'entrée, x_j pour $j = 1, \dots, m$.
2. Un additionneur qui additionne les valeurs d'entrée pondérées pour calculer les entrées à la fonction d'activation $v = w_0 + \sum_{j=0}^m w_j x_j$ où w_0 est appelé le biais (à ne pas confondre avec le biais statistique dans la prédiction ou l'estimation) est une valeur numérique associée au neurone. Il est pratique de considérer le biais comme le poids d'une entrée x_0 dont la valeur est toujours égale à 1.
3. Une fonction d'activation g (également appelée fonction d'écrasement) qui mappe v

à $g(v)$ la valeur de sortie du neurone. Cette fonction est une fonction monotone. Elle peut différer selon le type de problème ainsi que sa compatibilité avec l'optimisateur utilisé.



un groupe d'unités neuronales artificielles formerait une couche, chaque couche alimente la couche suivante et reçoit en même temps de la couche précédente. Cette architecture présente la forme la plus connue et la plus discrète de l'apprentissage profond, à savoir un réseau neuronal standard.



Chaque neurone est alimenté par tous les neurones de la couche précédente jusqu'au neurone de sortie où la comparaison avec les résultats réelles s'effectue. Cette comparai-

son nous introduit une forme d'erreur dont le méthode de calcul peut différer selon le type de problème (Régression ou Classification) où tout simplement selon l'objectif de l'analyste, puisque la performance de l'algorithme ou de l'approche peut différer selon la méthode.

Après le calcul de l'erreur, l'optimisateur vient pour réduire cet erreur en étudiant l'effet de la variation des différents variable entraînable comme les poids et les biais de chaque couche, la plupart des optimisateurs utilisés sont des optimisateurs locaux basés principalement sur l'algorithme du gradient descendant. Pour le moment l'optimisateur le plus utilisée et recommandé est ADAM ou estimateur **AD**aptif du **M**oment.

L'architecture du réseau de neurone est aussi importante puisque chacune peut exprimer un autre forme de fonctionnement. Par exemple, on peut modéliser les algorithmes de réduction de dimensions ou l'analyse principale des composants sous forme d'une architecture dont les couches contiennent moins de neurone en avançant vers le centre ou la fin, d'où la concentration de l'information. Dans la figure suivante nous présentons des architectures introduisant cette fonctionnalité :

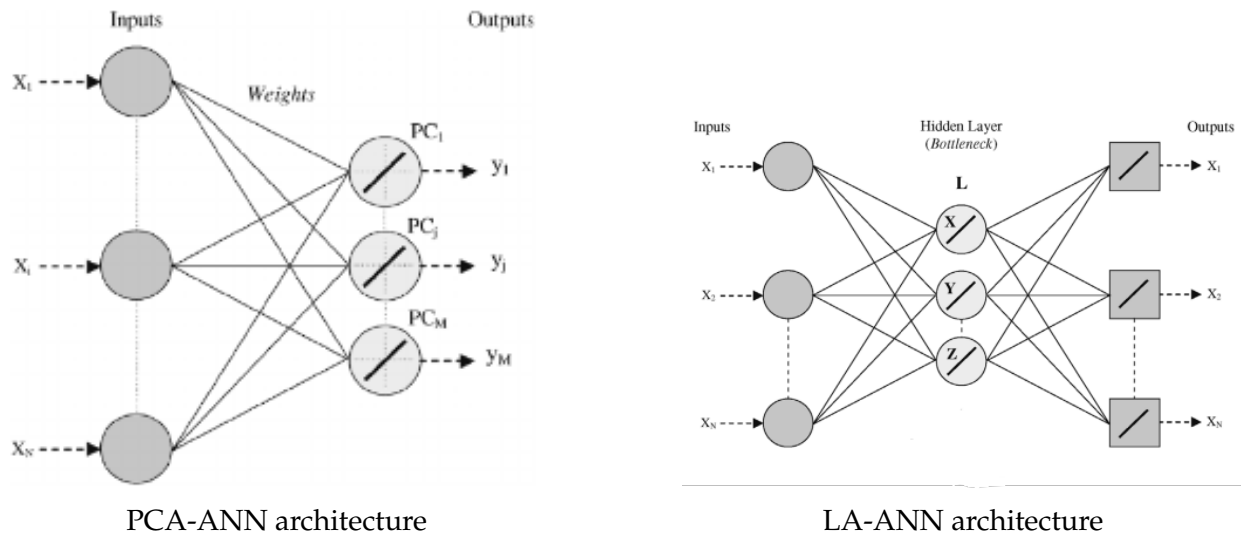


FIGURE 4.21 – Représentation de l'ACP sous une architecture ANN

L'intuition est que l'information se concentre pendant son passage d'une couche à une suivante avec un nombre de neurones inférieur au nombre d'entrées (un goulot d'étranglement) car, pendant la rétro-propagation, l'optimisateur donne plus de poids aux variables les plus représentatives. Le nombre de neurones inclus dans cette couche représente le nombre de dimension désiré.

Il existe plusieurs types d'architecture de réseau de neurones, citons parmi eux :

- RNN ou réseau de neurone récurrente : ils sont principalement utilisés pour la modélisation des problèmes de nature séquentielle ou chronologique ; c'est à dire les problèmes dont l'ordre des données est important, comme les problèmes de traduction, où l'ordre des mots est important, ou les problèmes relatifs aux données auditives comme la génération de l'audio.
- CNN ou Réseau neuronal convolutif : ils sont souvent utilisés pour les problèmes en relation avec le traitement d'image. Mais ils peuvent performer mieux que les RNNs dans certaines problèmes séquentiels.

Objectif d'utilisation :

Les données que nous utilisons des taux moyens pondérés présentent des changements au niveau de la moyenne au cours du temps. Ces changements se caractérisent par une soudaineté ou une rapidité de survenance comme présentée dans la figure ci-dessous.

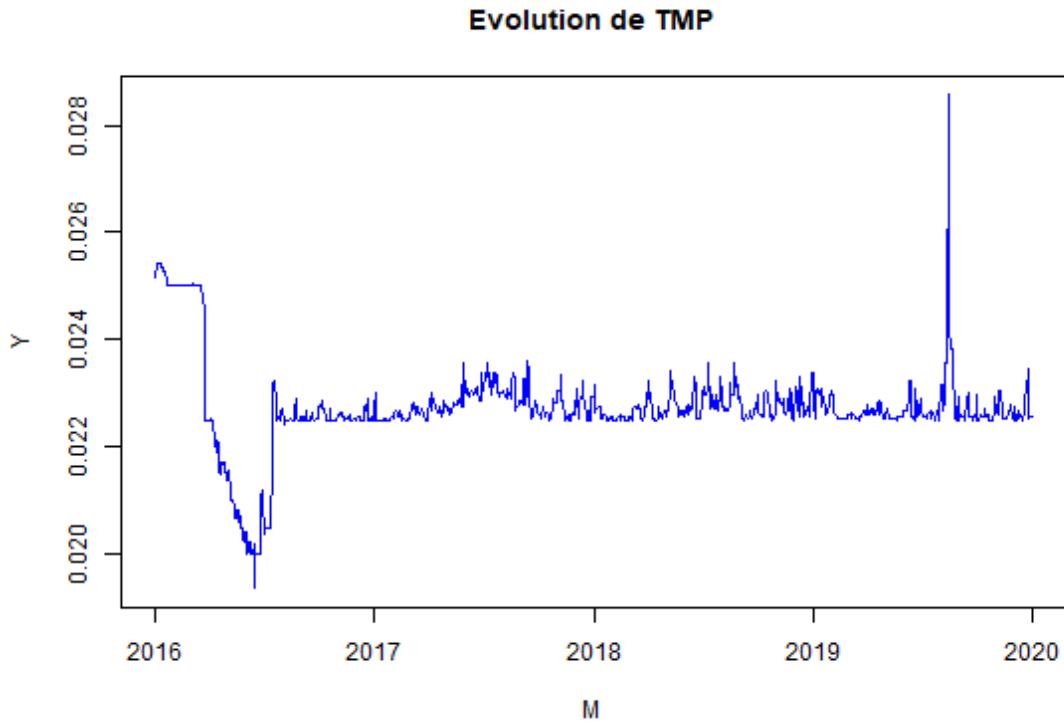


FIGURE 4.22 – Changement de la moyenne au long terme - TMP

Notre étude sera autour de la modélisation des taux courts au long terme et plus précisément sur leur performances par rapport aux modèles stochastiques quant à leur capacité à capter le changement soudain de la moyenne ou les "jumps" dans les séries financières, à savoir que les modèles stochastiques de Vasicek et de CIR n'introduisent pas cette propriété.

Notre étude est inspirée de l'étude de Ymir Mäkinen, Juho Kannianen, Moncef Gabbouja, Alexandros Iosifidis : *Forecasting of Jump Arrivals in Stock Prices : New Attention-based Network Architecture using Limit Order Book Data* , où les auteurs proposent une nouvelle architecture composée de plusieurs types de réseaux neuronaux : les réseaux convolutifs, les réseaux de mémoire à long et court terme et les couches d'attention, conçue pour prédire les sauts du rendement sur les marchés d'actions avec des données de carnet d'ordres à haute fréquence. Les résultats indiquent que l'architecture proposée est la

plus performante des architectures testées suivie d'une architecture LSTM pure (qu'on va introduire par la suite).

Alors selon le type de données qu'on dispose qui est de type séquentiel, le type le plus convenant est le réseau récurrent, mais notre problème est une modélisation long terme. Aors que Une architecture basique est connue pour sa performance faible dans les pré-diction long terme, pour cette raison nous introduisant les réseaux LSTM. Dans notre cas on se concentrera sur les réseau LSTM pures. Alors C'est quoi un réseau LSTM?

LSTM

introduction des RNNs :

Un réseau de neurones récurrents ressemble beaucoup à un réseau neuronal simple, sauf qu'il possède également des connexions orientées vers l'arrière. Alors pour le RNN le plus simple possible, composé d'un neurone qui reçoit des entrées, produit une sortie et se renvoie cette sortie à lui-même, comme le montre la figure ci-dessous. À chaque pas de temps t (également appelé trame), ce neurone récurrent reçoit les entrées $x(t)$ ainsi que sa propre sortie du pas de temps précédent, $y(t-1)$. Comme il n'y a pas de sortie précédente au premier pas de temps, elle est généralement fixée à 0. Nous pouvons représenter ce minuscule réseau sur l'axe du temps, comme le montre la figure ci-dessous. Cela s'appelle un déroulement du réseau dans le temps (il s'agit du même neurone récurrent représenté une fois par pas de temps). Un réseau dans le temps (il s'agit du même neurone récurrent représenté une fois par pas de temps).

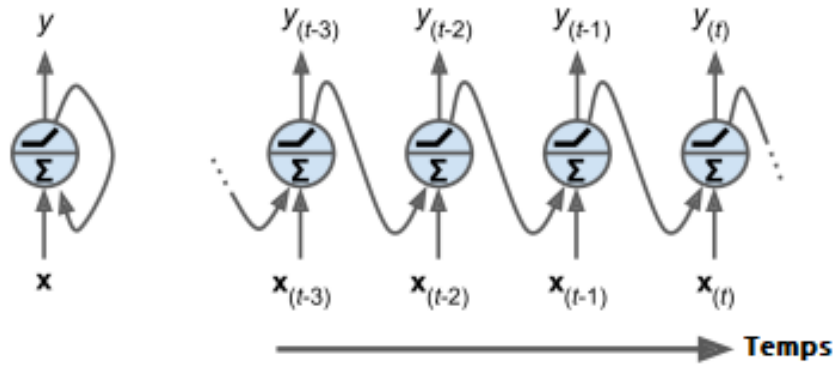


FIGURE 4.23 – Réseau de neurones récurrente

À chaque pas de temps t , chaque neurone reçoit à la fois le vecteur d'entrée $\mathbf{x}(t)$ et le vecteur de sortie du pas de temps précédent $\mathbf{y}(t-1)$. Notez que les entrées et les sorties sont des vecteurs (lorsqu'il n'y avait qu'un seul neurone, la sortie était un scalaire).

Pour introduire le cas général, nous considérons avoir un vecteur des entrées temporel $\mathbf{x}(t)$.

Chaque neurone récurrent possède deux ensembles de poids : un pour les entrées $\mathbf{x}(t)$ et l'autre pour les sorties du pas de temps précédent, $\mathbf{y}(t-1)$. Appelons ces vecteurs de poids w_x et w_y . Si nous considérons la couche récurrente entière au lieu d'un seul neurone récurrent, nous pouvons placer tous les vecteurs de poids dans deux matrices de poids, w_x et w_y . Le vecteur de sortie de l'ensemble de la couche récurrente peut alors être calculé à peu près comme on pourrait s'y attendre :

$$\mathbf{y}(t) = g \left(\mathbf{W}_x^\top \mathbf{x}(t) + \mathbf{W}_y^\top \mathbf{y}(t-1) + \mathbf{b} \right)$$

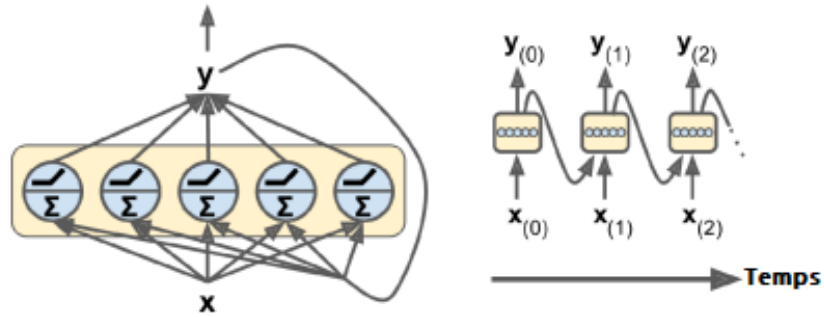


FIGURE 4.24 – Réseau de neurones récurrent pour plus de deux entrées

Avec :

- g : la fonction d'activation définie précédemment.
- $x_{(t)}$: Vecteur des entrées
- $y_{(t-1)}$: Vecteur des sorties du neurone précédent.
- W_x : Vecteur de Poids relatif aux entrées x_t .
- W_y : Vecteur de Poids relatif aux sorties du neurone précédent y_{t-1} .
- b : Biais

Il y a plusieurs type de RNN :

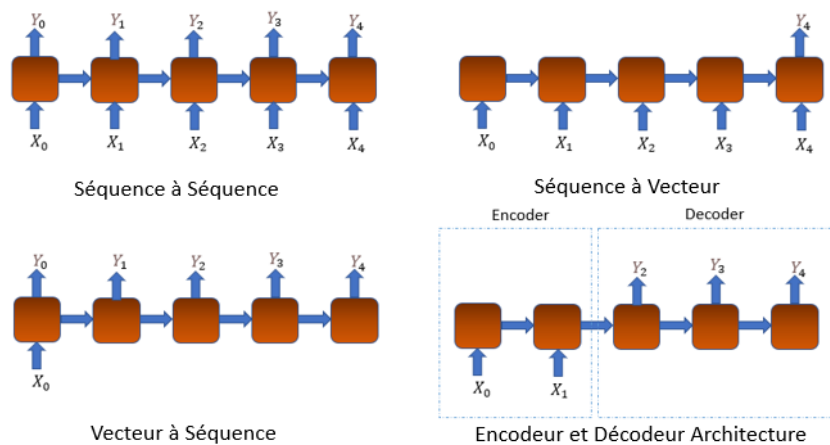


FIGURE 4.25 – Type des RNNs

Un RNN peut simultanément prendre une séquence d'entrée et produire une séquence de sortie (voir le réseau en haut à gauche). Ce type de réseau séquence à séquence est utile pour prédire des séries chronologiques telles que les cours boursières en lui fournissant les cours des N derniers jours et il doit produire les cours décalés d'un jour dans le futur (c'est-à-dire d'il y a N - 1 jours à demain).

On peut aussi alimenter le réseau avec une séquence d'entrée et ignorer toutes les sorties sauf la dernière (voir le réseau en haut à droite). En d'autres termes, il s'agit d'un réseau de séquence à vecteur. Par exemple, nous pourrions alimenter le réseau avec une séquence de mots correspondant à une critique du film, et le réseau produirait un score de sentiment dont -1 représente la haine et +1 représente l'amour.

À l'inverse, nous pouvons alimenter le réseau avec le même vecteur d'entrée à chaque pas de temps et le laisser produire une séquence (voir le réseau en bas à gauche). Il s'agit d'un réseau vecteur-séquence. Par exemple, l'entrée pourrait être une image animée (ou la sortie d'un CNN), et la sortie pourrait être une légende pour cette image.

Enfin, nous pourrions avoir un réseau séquence-vecteur, appelé encodeur, suivi d'un réseau vecteur-séquence, appelé décodeur (voir le réseau en bas à droite). Cette structure est largement utilisée en traduction et dans les travaux où l'entrée et la sortie ont une longueur différente.

Pour **L'entraînement** de ce type de réseau de neurones, il s'effectue d'une manière normale comme la rétropropagation de l'optimisateur sur le réseau de neurone durant laquelle il calcule la valeur d'ajustement de chaque élément entraînable, comme les poids et les biais de chaque couche. La nature temporelle qu'un réseau récurrent présente rend la rétropropagation, d'une manière, un déroulement dans le temps. Mais cette notion reste juste un nom raffiné, car elle reste toujours une rétropropagation normale de l'optimisateur sur les couches du réseau.

Le problème que ce type de réseaux, après l'entraînement, est le sur-apprentissage, mais des solutions ont été déjà introduites à cet égard, comme le '*Dropout*' ou de façon générale

les moyens de régularisation.

LSTM

Les Réseaux LSTM sont des réseaux de neurones de type récurrente. Mais dont les cellules de neurone effectue des opération plus complexe qu'une cellule récurrente simple. La cellule reçoit les entrées de l'instant t , le état cachée de la cellule précédente, dont le résultat est une transformation des entrées de cette dernière (pour un réseau neurale récurrent simple elle représente les sorties de cette cellule elles même), et un terme représentatif de l'état à long terme des séquences.

L'idée clé est que le réseau peut apprendre ce qu'il faut stocker dans l'état à long terme, ce qu'il faut jeter et ce qu'il faut lire. Lorsque l'état à long terme $c(t-1)$ traverse le réseau de gauche à droite, vous pouvez voir qu'il passe d'abord par une porte d'oubli, en éliminant certains composants, puis qu'il ajoute de nouveaux composants via l'opération d'addition (qui ajoute les composants qui ont été sélectionnés par une porte d'entrée). Le résultat $c(t)$ est envoyé directement à l'extérieur, sans autre transformation. Ainsi, à chaque pas de temps, des éléments précédents sont abandonnés et des éléments sont ajoutés. De plus, après l'opération d'addition, l'état à long terme est copié et passé par la fonction \tanh , puis le résultat est filtré par la porte de sortie. Cela produit l'état à court terme $h(t)$ (qui est égal à la sortie de la cellule pour ce pas de temps, $y(t)$).

Le calcul se fait de la façon suivante :

$$\begin{aligned} \mathbf{i}_{(t)} &= \sigma \left(\mathbf{W}_{xi}^T \mathbf{x}_{(t)} + \mathbf{W}_{hi}^T \mathbf{h}_{(t-1)} + \mathbf{b}_i \right) \\ \mathbf{f}_{(t)} &= \sigma \left(\mathbf{W}_{xf}^T \mathbf{x}_{(t)} + \mathbf{W}_{hf}^T \mathbf{h}_{(t-1)} + \mathbf{b}_f \right) \\ \mathbf{o}_{(t)} &= \sigma \left(\mathbf{W}_{xo}^T \mathbf{x}_{(t)} + \mathbf{W}_{ho}^T \mathbf{h}_{(t-1)} + \mathbf{b}_o \right) \\ \mathbf{g}_{(t)} &= \tanh \left(\mathbf{W}_{xg} \mathbf{x}_{(t)} + \mathbf{W}_{hg} \mathbf{h}_{(t-1)} + \mathbf{b}_g \right) \\ \mathbf{c}_{(t)} &= \mathbf{f}_{(t)} \otimes \mathbf{c}_{(t-1)} + \mathbf{i}_{(t)} \otimes \mathbf{g}_{(t)} \\ \mathbf{y}_{(t)} = \mathbf{h}_{(t)} &= \mathbf{o}_{(t)} \otimes \tanh \left(\mathbf{c}_{(t)} \right) \end{aligned}$$

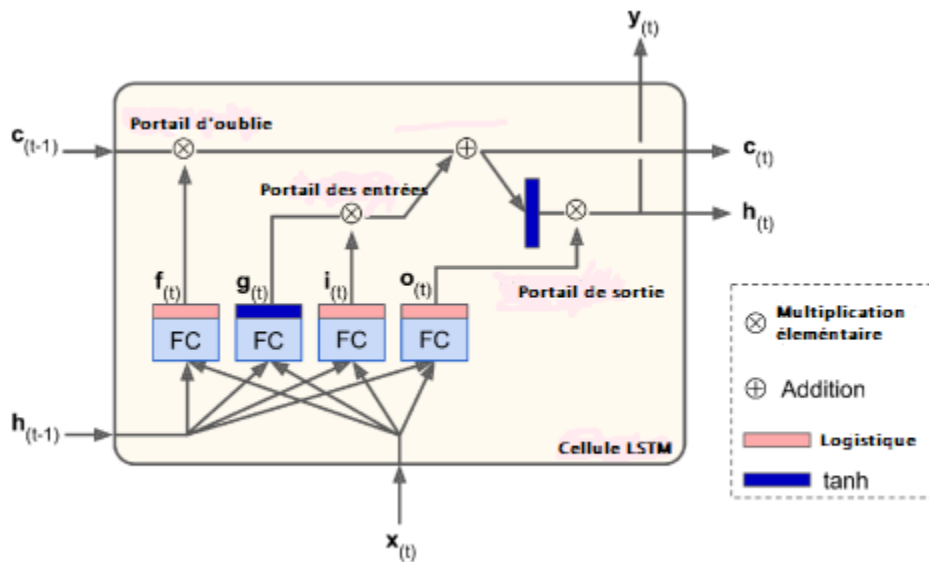


FIGURE 4.26 – Cellule LSTM

La couche principale est celle qui sort $g(t)$. Elle a le rôle habituel d'analyser les entrées actuelles $x(t)$ et l'état précédent (à court terme) $h(t-1)$. Dans une cellule de base, il n'y a rien d'autre que cette couche, et sa sortie va directement vers $y(t)$ et $h(t)$. Par contre, dans une cellule LSTM, la sortie de cette couche ne sort pas directement, mais ses parties les plus importantes sont stockées dans l'état à long terme (et le reste est abandonné).

Les trois autres couches sont des contrôleurs de porte. Comme elles utilisent la fonction d'activation logistique, leurs sorties vont de 0 à 1. Comme vous pouvez le voir, leurs sorties sont alimentées par des opérations de multiplication par éléments, donc si elles produisent des 0, elles ferment la porte, et si elles produisent des 1, elles l'ouvrent. Plus précisément :

1. La porte d'oublie (contrôlée par $f_t v$) contrôle quelles éléments de l'état long terme à écraser.
2. La porte des entrées (contrôlée par $i_t v$) contrôle quelles éléments de g_t à ajouter à l'état long terme.

3. Enfin, la porte de sortie (contrôlée par $o(t)$) contrôle les éléments de l'état à long terme qui doivent être lus et émis à ce pas de temps, à la fois vers $h(t)$ et vers $y(t)$.

Ce qui nous a poussé à utiliser les réseaux LSTM est l'état long terme de ce type de cellule qui peut capter le changement du moyen long terme des taux courts.

Application :

L'architecture choisie est un réseau de neurones LSTM multicouches stockés à 3 couches cachées dont la fonction d'activation est la fonction d'activation typique d'identité linéaire rectifiée 'relu'. L'optimisateur utilisé est le fameux optimisateur ADAM pour sa propriété de moment qui ajuste de la vitesse d'apprentissage se fait graduellement. Les entrées sont une séquence de 30 pas des valeurs du taux courts antérieurs et la sortie de la couche de sortie $y(t)$ sont la valeur qui suit directement ces séquences. La structure est décrite sur la figure la suivante :

Les résultats sont illustrés sur le graphe d'apprentissage suivant :

```
Model: "sequential_1"
-----
Layer (type)                Output Shape              Param #
-----
lstm (LSTM)                  (None, 7, 32)            4352
lstm_1 (LSTM)                (None, 7, 32)            8320
lstm_2 (LSTM)                (None, 7, 32)            8320
dense (Dense)                (None, 7, 1)              33
-----
Total params: 21,025
Trainable params: 21,025
Non-trainable params: 0
-----
```

FIGURE 4.27 – Architecture du réseau

Le graphique d'apprentissage illustre le développement de la moyenne absolue des erreurs de la base de données de validation et d'entraînement pour chaque époque.

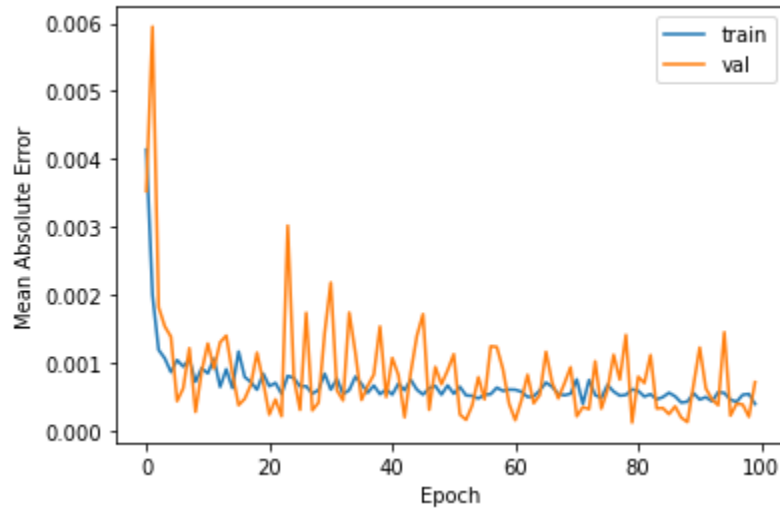


FIGURE 4.28 – Graphes d'Apprentissage

Quoique la vitesse de convergence est grande, la non-stabilité du modèle pendant le traitement de la base de données de validation rend le modèle moins performant d'un modèle stable à erreur absolue moyenne un peu supérieure. ce qui indique que le modèle n'est pas bon.

Modèle introduisant un element de diffusion de saut.

Pour résoudre ce problème, nous inspirons des modèles de diffusion des sauts, dont le modèle principal, qui est dans notre cas les modèles de Vasicek, comprend un élément additionnel représentant les sauts. Cet élément est composé de la probabilité de ce saut à cette instance-là multipliée par la différence que ce saut a créée.

Alors pour modèle comme celui de Vasicek, il serait ;

$$dr_t = a(b - r_t)dt + \sigma dW_t + Jump$$

où :

$$Jump = jump_t * N_t$$

avec :

- $jump_t = 0$ si on a pas de saut et 1 si oui
- N_t : est la valeur du saut ; c'est à dire la différence entre la valeur à t et celle à t-1.

Cette quantité est souvent modélisée par un processus de poisson. Alors pour représenter les variables requis pour estimer cette quantité, on introduira deux nouvelles variable ; un représentant l'arrivée d'un saut $jump_t$ et l'autre présentant le nombre d'instant écoulés entre après le dernier saut Nt .

La classification d'un arrivé de saut ($jump_t = 1$) est fait par comparaison de la différence entre la valeur à une instance et la valeur qu'elle suit, si cette différence dépasse un seuil prédéfini alors on a un saut et par consequent $jump_t = 1$, sinon $jump_t = 0$. Les valeurs de la variable de comptage des instants écoulés entre les sauts sont affectés de la même façon comme la variable est décrite.

L'architecture est la suivante :

Les résultats sont les suivants :

```

Model: "sequential_7"
-----
Layer (type)                Output Shape              Param #
-----
lstm_18 (LSTM)              (None, 7, 31)            4340
lstm_19 (LSTM)              (None, 7, 15)            2820
lstm_20 (LSTM)              (None, 24)                3840
dense_6 (Dense)             (None, 1)                  25
-----
Total params: 11,025
Trainable params: 11,025
Non-trainable params: 0

```

FIGURE 4.29 – 2^{me} Architecture.

Notons que la vitesse d'apprentissage était réduite à $1e - 6$ pour permettre à l'optimiseur de mieux estimer le sens d'optimisation et la quantité de variation requise pour la

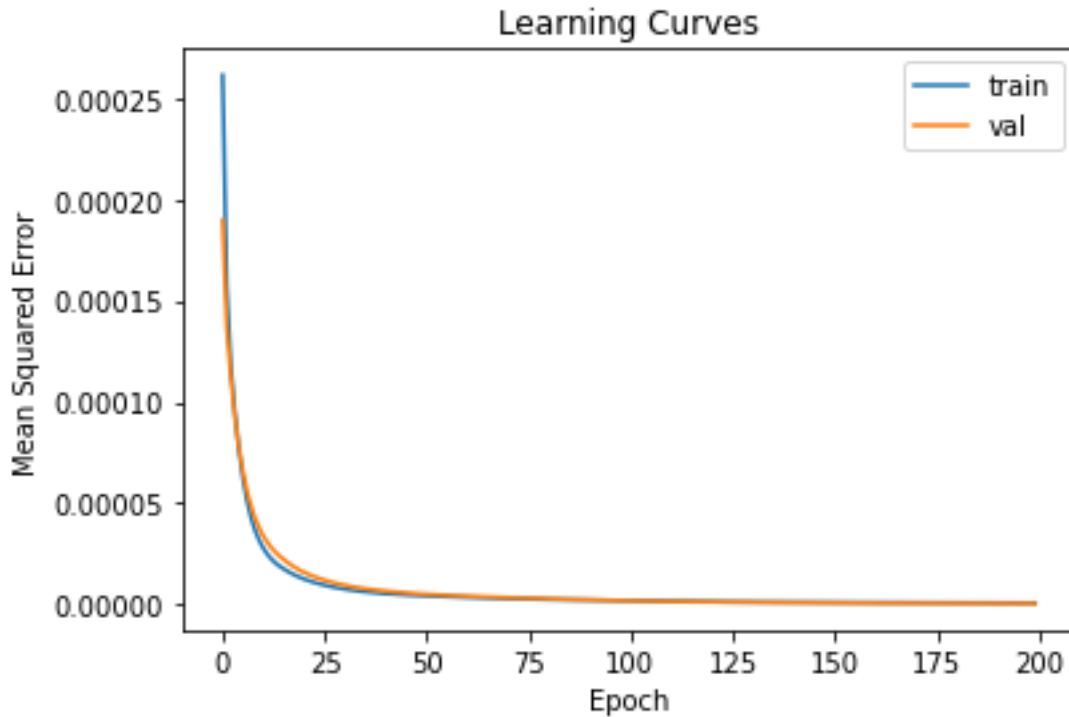


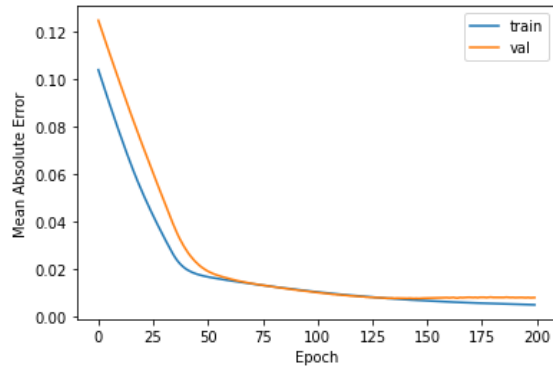
FIGURE 4.30 – Courbe d'évolution de la moyenne d'erreur quadratique.

convergence et que la base de données utilisées est celle sur laquelle on a travaillé pendant la modélisation de la courbe des taux. Par la suite on travaillera sur une base de données avec un historique plus long (de 01/01/2013 à 31/12/2021) pour tester la performance de l'architecture en comparaison avec un modèle de AR(1).

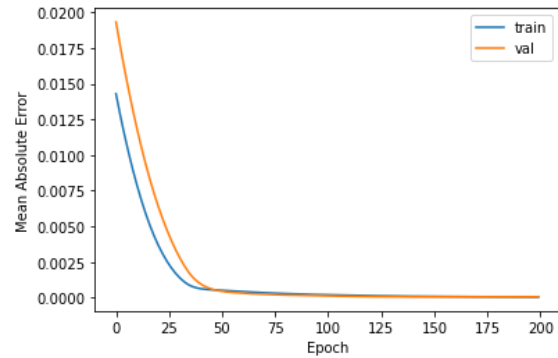
Quoique la vitesse d'apprentissage était réduite, la vitesse de convergence est rapide ainsi que la vitesse de traitement des données. A valeur finale pour l'erreur quadratique est $5.0354e - 07$ et $4.3488e - 04$ l'erreur moyenne absolue de la base de données de validation. Pour la base de données de traînement, on a $7.6805e - 07$ et $5.6910e - 04$ respectivement.

Pour la base de donnée entière :

Voici les graphes d'apprentissage ;



Moyenne absolu des erreurs



Moyenne quadratique des erreurs

FIGURE 4.31 – Résultats de l’entraînement sur toute la base de donn e.

Un mod le AR(1) simple a donn e un r sultat de $7.11e - 4$ pour l’erreur moyen absolu et $6.09e - 7$ pour l’erreur quadratique. Quant   notre r seau de neurones, l’erreur quadratique final  tait $7.32e - 5$ et 0.008 pour l’erreur absolu.

	AR(1)	LSTM
mse	$6.09e-7$	$7.32e-5$
mae	$7.11e-4$	0.008

L’utilisation des r seaux s quence   s quence n’a pas donn e de mieux r sultat malgr  l’ largissement de la longueur de s quence utilis e pour pr dire le prochain pas.

Le r sultat final indique que les r seaux LSTM, avec plus d’ajustement, peuvent donner des r sultats mieux que les mod les traditionnels aux courts termes. Mais finalement, ils n’ont pas pu capter les changements du moyen long terme des taux courts, peut- tre   cause de la fr quence tr s faible de ces sauts qui r sulte d’un manque d’information ce qui implique que le r seau ignore ces sauts. Donc, il faut plus de d veloppement et de recherche   ce niveau ou on peut toujours retourner   l’utilisation de d’autres type de r seau comme les r seaux convolutifs 1D ou les ‘Dilated Wavenets’ Utilis e par Google dans ses travaux en relation avec la g n ration des s quences auditives, puisqu’ils peuvent garder en m moire $2^{\text{Nombre de couche}}$  l ments de l’historique, ou des r seaux de neurones dont

l'intégration de règles est faisable comme les DeepCTRL.

Chapitre **5**

Projection

5.1 Introduction :

La projection est un élément indispensable de toute étude, elle permet la prévision du déficit du régime ainsi que l'étude des chocs que le régime peut rencontrer pendant sa durée de vie.

Cette procédure se fait par prévision de la possibilité de pérennité du régime par rapport à la situation actuelle. Pour cela nous nous permettons de réaliser une projection selon des hypothèses en se basant sur un modèle de gestion actif-passif.

Pour commencer, on introduit les bases et la méthode de calcul des engagements futurs. Les hypothèses sont principalement ceux utilisés pour la tarification du régime, le tarif employé pour déterminer le nombre de points cumulé est celui précis précédemment.

5.1.1 Outils :

Les outils dont on aura besoin pour répondre à notre problématique se distinguent entre 4 Catégories :

- Données du régime
- Hypothèses actuarielles
- Hypothèses démographiques
- Données Financières

Données du régime :

Dans notre cas, Le régime englobe plusieurs catégories professionnelles. Chacune de ces catégorie a ses propres caractéristiques comme :

- La distribution de la population par âge et catégorie.

Pour ce point, on avait plusieurs catégories dans la moyenne des âges est très grandes dont les âges les plus jeunes tend à avoir des nombres très faibles. Grâce à cette propriété, la durée de construction de la réserve devient plus courte et par conséquent une date de pérennité plus proche.

— Le nombre de nouveaux actifs chaque année.

Sauf au cas où nous visons avoir une projection des cohortes constituant la population de la même année, le nombre des nouveaux actifs est important pour introduire plus de réalisme sur la procédure puisque toute population ne disparaît pas entre une année et la suivante.

— L'âge de retraite

Hypothèses actuarielles :

Les hypothèses actuarielles sur lesquelles nous nous basons pour nos projections se résument principalement dans le tarif utilisé pour calculer le nombre de points acquis et à liquider chaque année.

Hypothèses démographiques :

Le taux de développement de la population

Nous avons déjà une population de 1,6 million d'individus, soit la moitié de la capacité totale du régime. Donc pour les entrées, nous supposons que le taux est de 10% pour les deux premières années, puis 5% puis 3% pour les cinq suivantes et ensuite le taux sera fixé à 1% comme le taux d'accroissement de la population du Maroc.

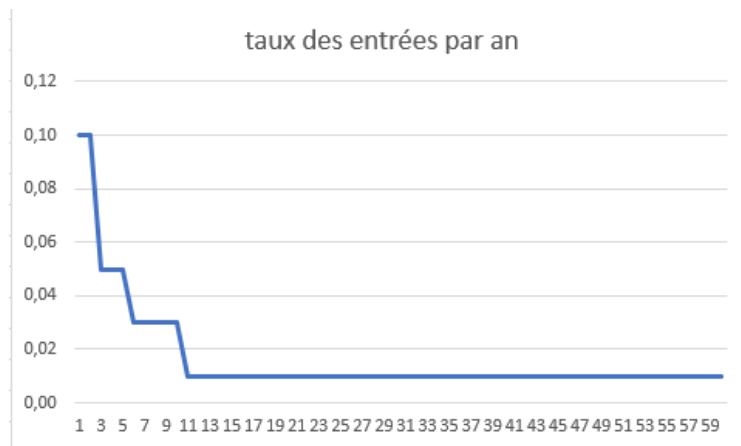


FIGURE 5.1 – Évolution du taux d’entrée par année de projection

Le taux de mortalité

La réglementation n’a pas précisé la table de mortalité à utiliser pour les projections. Alors, on peut modéliser une qui sera plus proche à la réalité des taux de mortalité.

Cette échappatoire nous permettra d’avoir plus de précisions quant à la valeur d’acquisition d’un point du régime en analysant la sensibilité de la durée avant la pérennité du régime envers ses paramètres.

Mais dans notre étude, on utilisera la table de mortalité conventionnelle TV 88-90.

Le taux de Turnover

Le taux de Turnover est inclus pour le seul but de ne pas surestimer les cotisations puisque Les individus peuvent quitter le régime pour une cause sauf la mortalité. Les valeurs sont supposées telles que la somme soit autour de 1%.

Données Financières :

L’un des revenus principaux d’un régime de retraite est le retour sur investissements à long terme. Pour cette raison qu’on aura besoin des taux courts pour modéliser la courbe des taux zéro-coupons.

5.1.2 Structure Financière :

Les engagements se constituent principalement :

- Les prestations envers les retraités ou les ayants-droits. ils se présentent sous forme de points à Liquider.

Principalement les cotisations sont fournies par les affiliés de moins de 65 ans tandis que les prestations appartiennent aux affiliés en retraite ou les ayants-droit des décédés.

Alors, en faisant des projection sur la population, on introduit implicitement des projections sur les résultats financières du régime.

Pour résumé, la structure financière du régime se compose des acquis qui se cumulent des cotisations des actifs et des résultats financiers des placements.

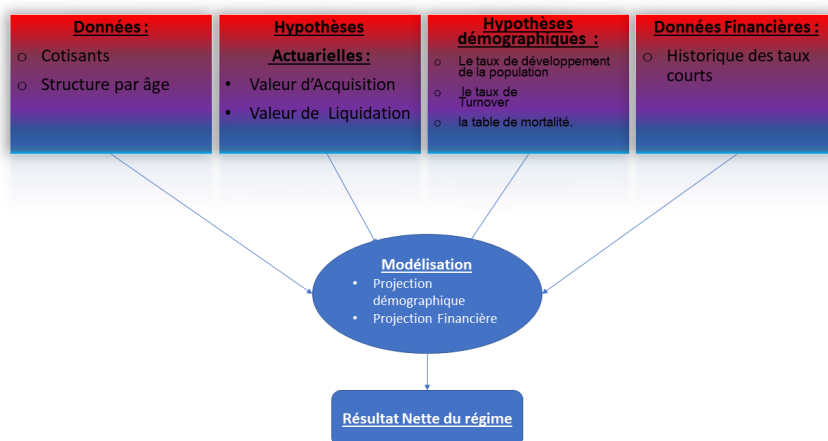


FIGURE 5.2 – Illustration de la projection démographique

5.2 Informations et Requisites :

Dans cette section, nous présentons les informations/Hypothèses dont on est besoin telles que :

— Nuptialité

le taux de nuptialité est la probabilité qu'un affilié soit marié. L'importance de ce paramètre se présente dans la moitié des prestations payée aux conjoints aux cas du décès de l'assuré. Pour prendre en compte le risque que cette partie de réglementation introduit, nous aurons besoin de taux de nuptialité. Les taux que nous utiliserons sont fournis par la CNSS de leurs études sur le population affiliée au système d'Assurance Maladie Obligatoire AMO dont la répartition est faite par âge. La figure suivante introduit ses valeurs :

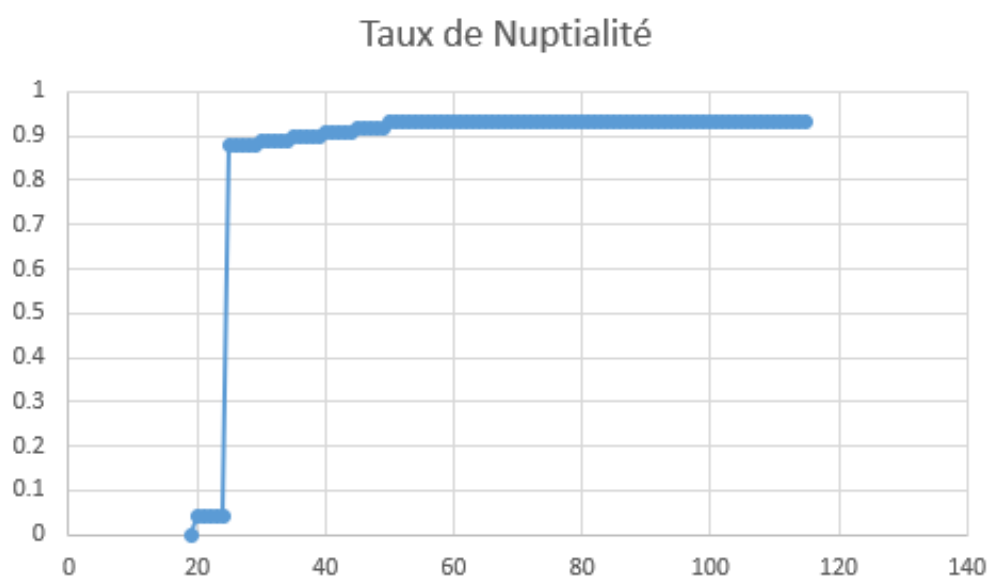


FIGURE 5.3 – Taux de nuptialité par âge

Cette répartition apparaît moins réelle, mais tant que la population n'a pas encore intégré le régime de façon complète, nous ne pouvons pas estimer ses paramètres et encore plus s'ils dépend des caractéristiques de chaque individu.

— Différence d'âge par défaut affilié/conjoint

La différence d'âge nous servira à définir la prestation du conjoint. elle est prise au niveau de 5 ans pour toute la population.

— Probabilité d'avoir un enfant normal

Tandis que le taux brut de mortalité de la population marocaine est en décroissance, la population dans sa globalité est en croissance et cela dû au taux brut de natalité qui dépasse le taux brut de mortalité. La figure suivante introduit l'évolution de ces paramètres pour la population globale dans les prochaines années partant de 2020 :

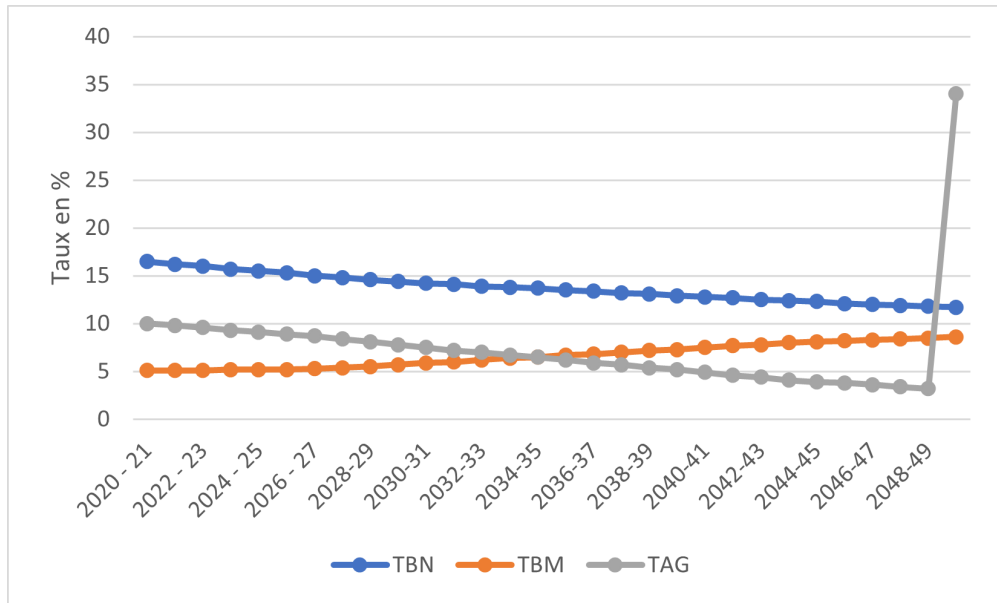


FIGURE 5.4 – Taux d'évolution globale et de naissance

Alors pour la même raison pour laquelle on a besoin du taux de nuptialité, on aura ce type de probabilité Mais par âges.

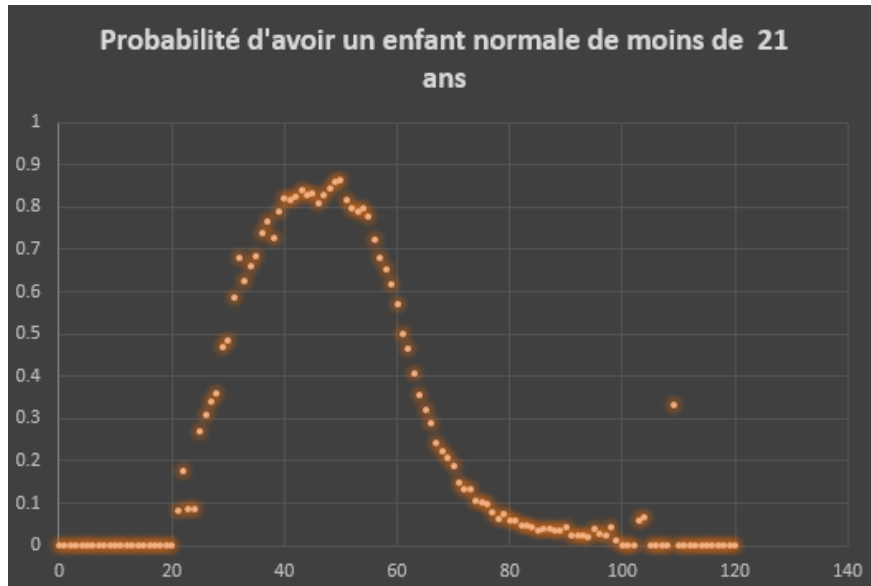


FIGURE 5.5 – Probabilité d’avoir un enfant normale de moins de 21 ans

La distribution apparaît normale et plus conforme avec la réalité. les valeurs nulles pour les âges de moins de 20 ans sont évidentes puisque l’âge légal de mariage et de 18 ans sauf pour des cas particuliers.

Et comme pour le cas du conjoint, on aura besoin de la différence d’âge entre les parents et les enfants, mais cette fois-ci, la distribution et fournis par âge.

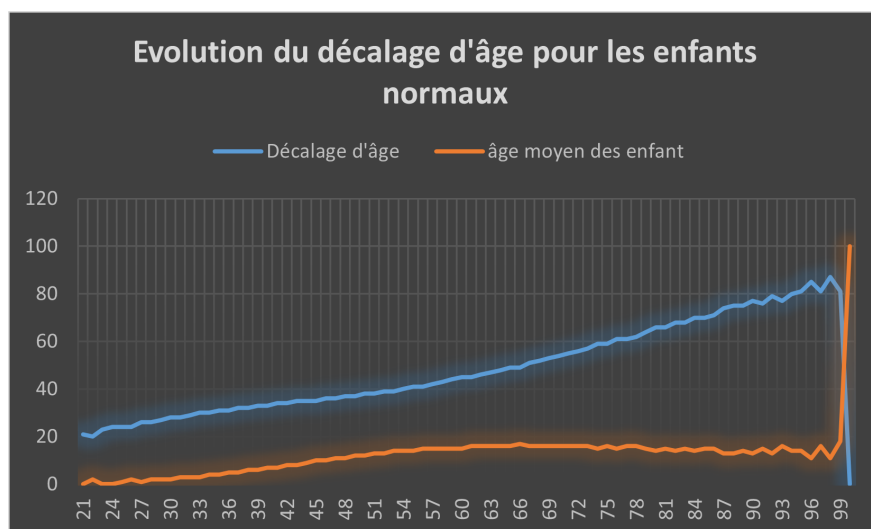


FIGURE 5.6 – Evolution du décalage d’âge pour les enfants normaux

l'âge moyen qui ne dépasse pas les 21 ans montre que pendant la construction des valeurs du décalage d'âge, la contrainte appliquée aux enfants en matière d'âge pour qu'ils bénéficient de leur prestation sont pris en considération.

Pour les enfants handicapés, la contrainte d'âge n'existe pas. pour cette raison qu'ils sont traité à vie sauf que la pension est distribuée sur tous les pensionnés des enfants. Alors voici les informations relatives à cette catégorie quant à la probabilité d'être.

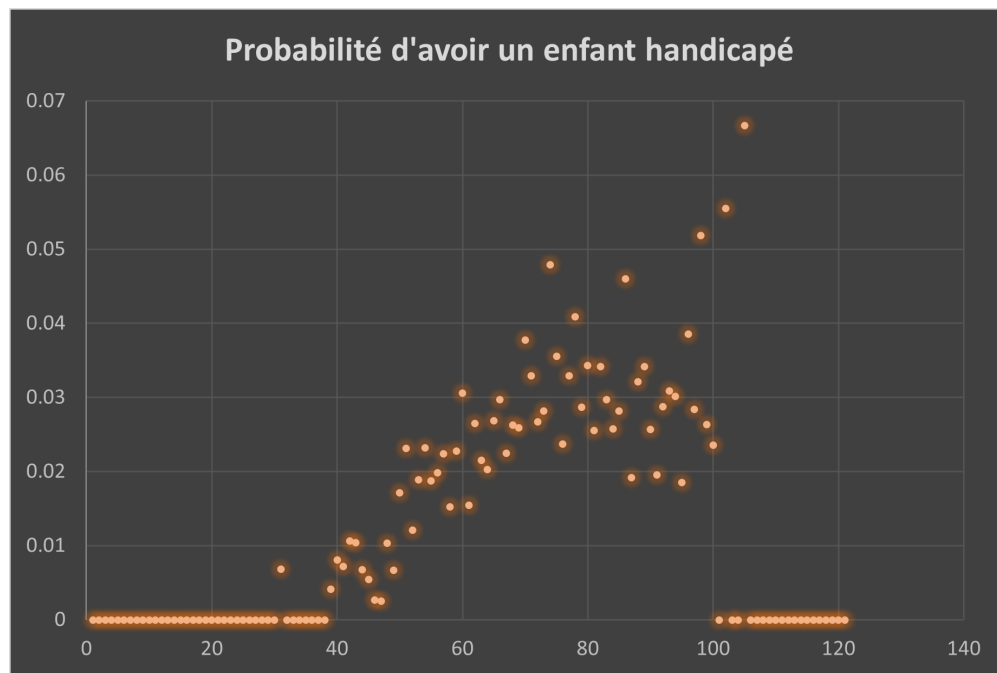


FIGURE 5.7 – Probabilité d'avoir un enfant handicapé

Les probabilités apparaissent moins réelles que nos attentes mais ils suivent un chemin . La raison principale peut tout être simplement le manque d'informations quant à cette partie de la population ou leur faible nombre.

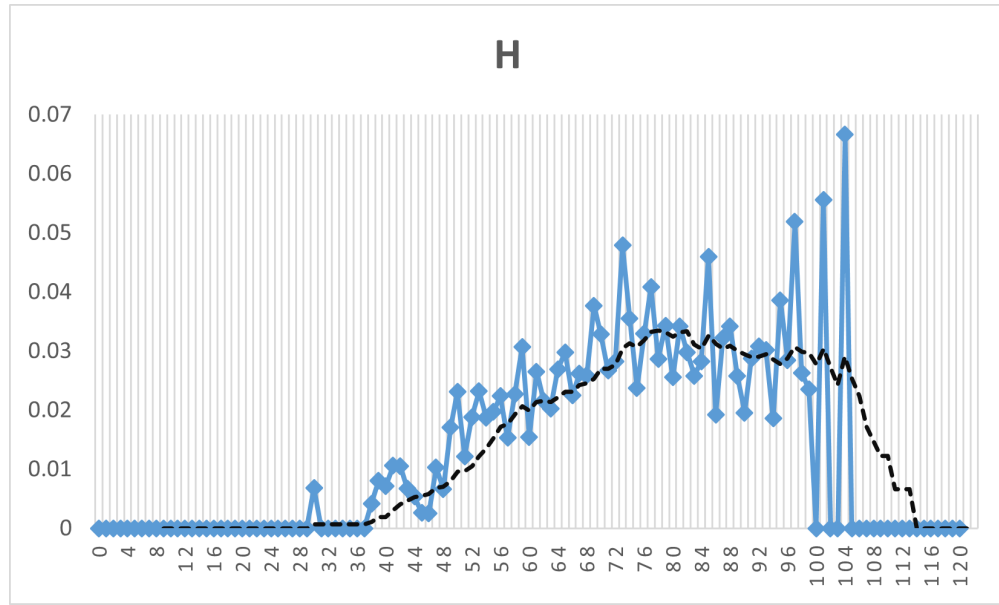


FIGURE 5.8 – Lissage de la courbe de probabilité d’avoir enfant handicapé par moyenne mobile

5.3 Projection Démographique :

Pour préciser Les flux financiers passifs du régime nous devant passer par une projection démographique. cette projection nous servira à déterminer le nombre d’actifs, le nombre de retraités et le nombre des ayants-droit pour chaque année.

Le processus de projection vers une année n se constitue du calcul du nombre d’affiliés de cette année à partir du nombre de l’année n . Le résultat se compose des survivants de l’année $n-1$ (survivants de la mortalité), des sorties pour d’autres raison autre que la mortalité et des nouveaux affiliés sous forme de nouveaux entrés.

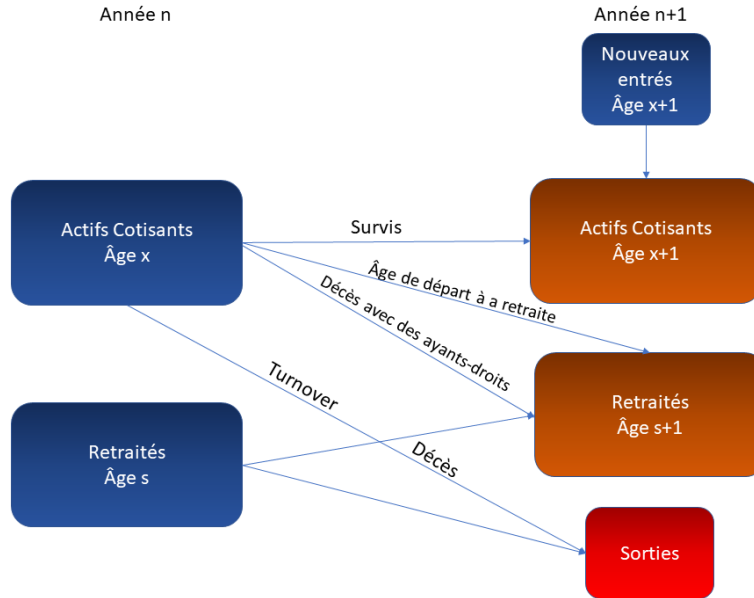


FIGURE 5.9 – Structure de projection démographique

Partant d'une **catégorie** ayant un nombre $N(x, n)$ pour un âge x et année n , Le nombre de points pour cet année est :

$$NP_n = N_n * (c_n / VA_n) = (\sum_{x=18}^{64} N_{x,n}) * c_n / VA_n$$

Avec :

$$N_{x,n} = N_{x-1,n-1} - D_{x-1,n-1} - ST_{x-1,n-1} + E_{x,n}$$

$N_{x-1,n-1}$: Le nombre d'individus affiliés aux régime ayant pour âge $x-1$ à l'année $n-1$.

c_n : est la cotisation tarifaire pour un individu à l'année n .

$D_{x-1,n-1} = N_{x-1,n-1} * (1 - p_{x-1})$: le nombre de décès des individus d'âges $x-1$ à l'année $n-1$

$ST_{x-1,n-1}$: Nombre d'individus quittant le régime sous quelque raison.

$E_{x,n}$: Nombre des nouveaux entrés d'âge x à l'année n .

Les entrées est les sorties ST sont définis par une proportion de la population $N_{x-1,n-1}$

(variable par âge).

fixons n dans $\{1, 2, 3, \dots, 60\}$:

La cotisation totale de l'année n se définit par :

$$C_n = \sum_{x=18}^{64} N_{x,n} * c_n$$

Pour le nombre de prestataires :

$$P_n = 0.5 * P_{f,n} + 0.47 * P_{en,n} + 0.03 * P_{eh,n} + \sum_{i=65}^w N_{i,n}$$

$$P_{f,n} = \sum_{k=0}^n P_{f,n,k}$$

$$P_{en,n} = \sum_{k=0}^n P_{en,n,k}$$

$$P_{eh,n} = \sum_{k=0}^n P_{eh,n,k}$$

$$P_{f,n,k} = (\sum_{i=18}^w D_{i,k} * n-k p_{i-d_c} * p_{m,i})$$

$$P_{en,n,k} = (\sum_{i=18}^w D_{i,k} * n-k p_{i-d_{N,i}} * p_{en,i}) \mathbb{1}_{(i-d_{N,i}+n-k < 21)}$$

$$P_{eh,n,k} = (\sum_{i=18}^w D_{i,k} * n-k p_{i-d_{H,i}} * p_{eh,i})$$

$P_{f,n}$: **Nombre des ayants-droits un conjoint à l'année n**

d_c : est l'écart d'âge entre les conjoints =5ans.

$p_{m,i}$: probabilité qu'un assuré d'âge i ait un conjoint

$P_{en,n}$: **Nombre des ayants-droits des enfants normaux à l'année n**

$d_{N,i}$: est la différence d'âge entre l'assuré et l'enfant normal.

$p_{N,i}$: probabilité qu'un assuré d'âge i ait un enfant normal

$P_{eh,n}$: Nombre des ayants-droits des enfants handicapés à l'année n

$d_{H,i}$: est la différence d'âge entre l'assuré et l'enfant andicapé.

$p_{H,i}$: probabilité qu'un assuré d'âge i ait un enfant andicapé

5.4 Projection Financière :

5.4.1 Introduction :

Le premier objectif de cette partie est de calculer le total des cotisations et des prestations dans les 60 prochaines années, puis nous actualiserons les flux futurs à la date d'aujourd'hui en utilisant le taux zéro coupon modélisé avec le modèle classique de Vasicek.

5.4.2 Formulation :

La cotisation totale de l'année n se définit par :

$$C_n = \sum_{x=18}^{64} N_{x,n} * c_n$$

$$Prestation_n =$$

$$\sum_{k=0}^n NP_k * VL_k * (0.5 * P_{f,n,k} + 0.47 * P_{en,n,k} + 0.03 * P_{eh,n,k}) + (NP_n * VL_n * \sum_{i=65}^w N_{i,n})$$

5.4.3 Automatisation sous R shiny :

Interface R shiny :



FIGURE 5.10 – Interface pour la projection des cotisations et des prestations

les zones 1 à 6 sont pour importer la base de données des TNS et des probabilités, changer les hypothèses comme la valeur d'acquisition et le taux d'augmentation des revenus, et enfin le téléchargement des sorties. Les zones 8 et 9 sont pour afficher les sorties

Projection pure :

Nous allons fixer les hypothèses :

- La valeur d'acquisition : 15
- Turnover : 0
- Taux d'augmentation des revenus : 2%
- taux d'entrée par année de projection

Année	1 à 2	3 à 5	6 à 10	11 à 60
Taux	10%	5%	3%	1%

FIGURE 5.11 – Taux d'entrée par année de projection

Voici les résultats de la projection de la catégorie **Adoules** :

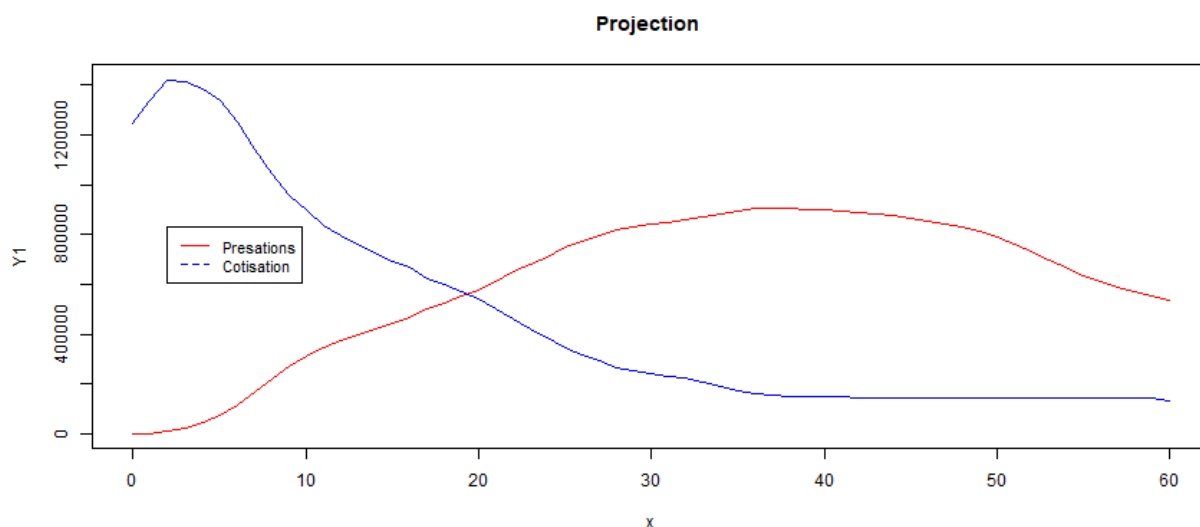


FIGURE 5.12 – Projection pour L'Adoules

Nous remarquons que la courbe des cotisations augmente pendant les deux premières années grâce aux nouvelles entrées de 10% au nombre faible de sorties. Après les deux premières années, nous remarquons que la courbe des cotisations commence à diminuer et la courbe des prestations commence à augmenter. Ceci s'explique par le fait que le nombre d'entrées dans le système est faible et que le nombre de retraités commence à augmenter. Après 20 ans, la courbe des cotisations croise la courbe des prestations. À partir de ce moment, les prestations deviennent supérieures aux cotisations, ce qui s'explique par le fait que l'âge moyen des Adoules est de presque 50 ans, donc après 15 ans, ils prendront tous leur retraite ou laisseront des conjoints ou des orphelins.

Prestations cumulées Vs Cotisations cumulées

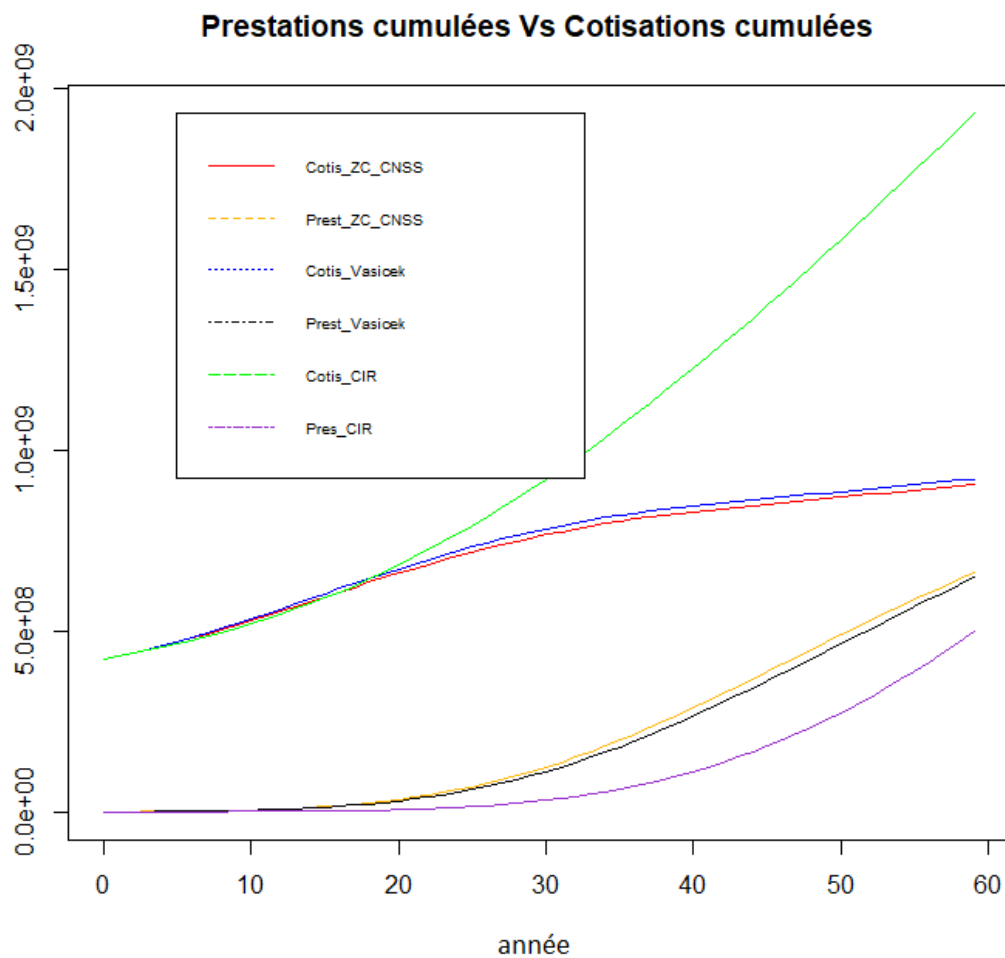


FIGURE 5.13 – Prestations cumulées Vs Cotisations cumulées

On remarque que la projection financière des cotisations et des prestations avec le taux zéro coupon de la CNSS et le taux estimé par le modèle de Vasicek sont proches, mais pour le modèle de CIR ils nous donnent des résultats un peu éloignés de nos attentes.

On remarque aussi que la courbe des cotisations est toujours au-dessus de la courbe des prestations ce qui montre qu'il n'y aura pas de déficit des réserves dans les 60 prochaines années.

Projection des cotisations et prestations pour toutes les catégories

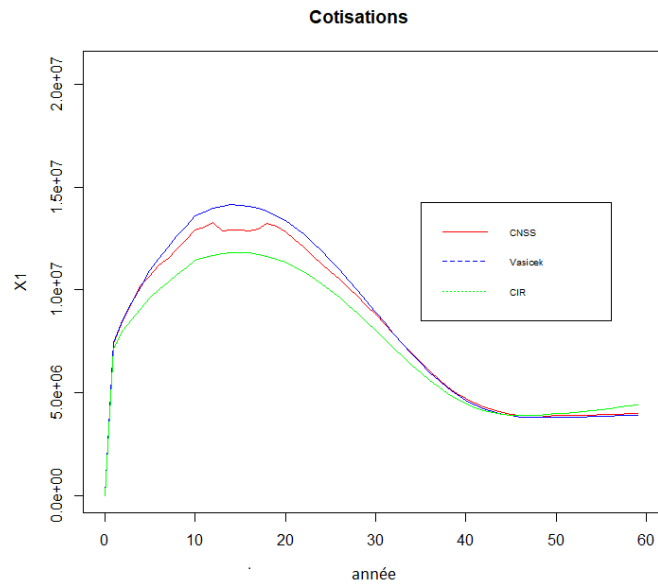


FIGURE 5.14 – projection des cotisations pour toutes les catégories

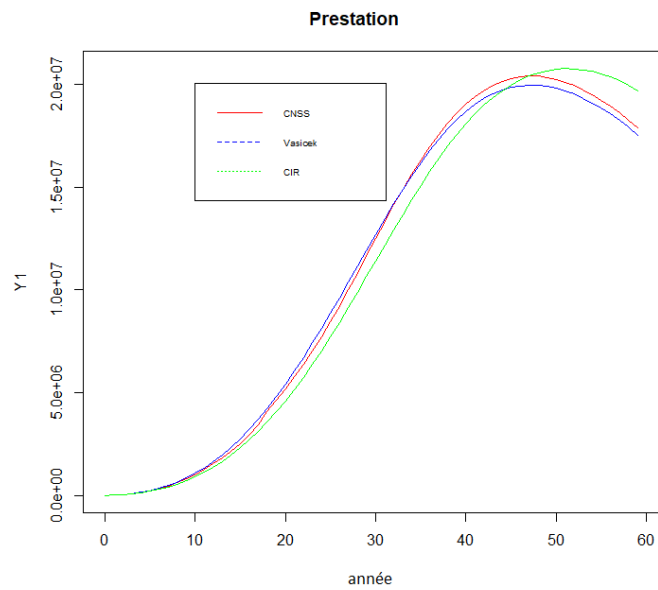


FIGURE 5.15 – projection des prestations pour toutes les catégories

Conclusion

Sous la lumière de ce travail, on a réalisé une étude actuarielle du régime de retraite en points destiné aux travailleurs non salariés, dans laquelle on a trouvé un tarif d'équivalence des engagements qui peut être considéré chère pour certaines catégories ayant un revenu faible et dont l'aide de l'État est souhaitable.

Nous avons mis à l'essai la sensibilité de ce tarif à plusieurs paramètres, dont le résultat correspondait aux attentes.

D'après la projection financière, nous avons déduit que le meilleur modèle illustrant les taux courts est celui de Vasiček, gagnant le défi contre le réseau LSTM qui apparaît mieux adapté aux modélisations court terme de ces données financières.

Avec une projection démographique et la courbe des taux modélisés, la projection des engagements nous a montrés que la pérennité de ce régime ne sera ni un événement soudain ni une catastrophe proche puisque le résultat net des engagements sera encore positif après 60 ans.

Bibliographie

- Mr DEVOLDER Pierre, cours Prévoyance et Assurances de groupe
- Mr MARRI Fouad, cours assurance vie.
- EL QALLI Yassine, cours courbe des taux.
- Y.Mäkinen, J.Kanniainen .., *Forecasting of Jump Arrivals in Stock Prices : New Attention-based Network Architecture using Limit Order Book Data*
- Ashish Vaswani, *Attention Is All You Need*
- Aurélien Géron, *Hands-on Machine Learning with Scikit-Learn, Keras and TensorFlow*
- <https://www.actuarialab.net/>
- www.ressources-actuarielles.net
- <http://wikistat.fr>
- Rapport du nouveau modèle de développement.
- www.cnss.ma