

# Modélisation du CCF Forward Looking IFRS9

Emma Eberle, Morgan Jowitt, Tonin Rivory, Gaétan Dumas

Master MOSEF 2025 – Université Paris 1

## 1 Introduction

Ce projet s'inscrit dans la réforme Bâle IV et les exigences IFRS 9. Il vise à modéliser un **CCF** (*Credit Conversion Factor*) **Forward Looking**, afin de rendre l'estimation de l'EAD (*Exposure at Default*) sensible aux variables macroéconomiques, à l'instar de la PD et de la LGD.

À partir de séries historiques de CCF observés et de données macroéconomiques (PIB, chômage, inflation, immobilier), nous avons construit plusieurs modèles économétriques pour chacun des segments disponibles, afin de projeter le CCF à un horizon de 3 ans.

Ce travail repose sur trois piliers :

- Analyse exploratoire et stationnarisation des séries segmentées et macroéconomiques,
- Modélisation linéaire du CCF par segment,
- Projection à horizon 2024–2026 selon divers scénarios économiques.

## 2 Données et préparation

Nous utilisons trois sources principales :

- **Données CCF segmentées** : séries trimestrielles par note de risque (1 à 6), avec les CCF moyens bruts et les proportions d'effectifs associés. Ce sont nos principales variables cibles.
- **Série globale CCF** : agrégats simplifiés sur les segments 1 à 5 (1\_5) ou 1 à 6 (1\_6), utilisés pour des analyses exploratoires.
- **Variables macroéconomiques** : PIB, taux de chômage, inflation, indice des prix des logements, récupérés de l'INSEE et disponibles de 1980 à 2023.

Les données ont été nettoyées, harmonisées, puis enrichies via des transformations classiques : différences, lags, standardisation.

## 3 Analyse exploratoire et stationnarisation

### 3.1 Structure et tendances générales

Les données CCF se déclinent en deux formats : global (agrégé) et segmenté. Les séries globales montrent une tendance décroissante, tandis que les segments individuels révèlent des dynamiques différenciées, visibles en figure 1.

Le segment 6, constant sur toute la période, a été écarté de la modélisation.

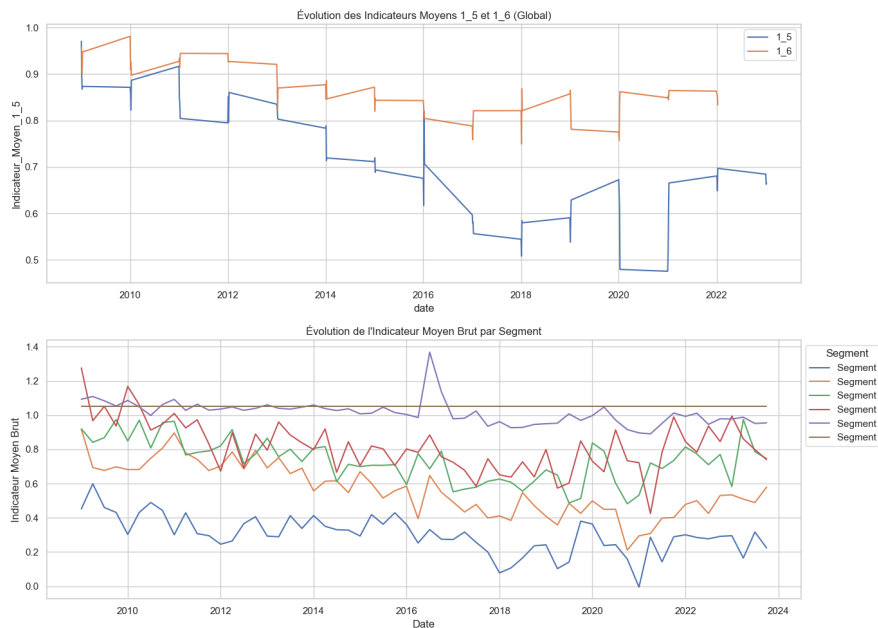


Figure 1: Évolution des indicateurs CCF par segment (2009–2023)

3.2 Tests de stationnarité et transformations

Les tests de racine unitaire (ADF) ont été réalisés sur les séries CCF par segment. Le tableau 1 ci-dessous présente les résultats pour les cinq premiers segments :

Table 1: Test ADF sur les séries CCF par segment

Série	ADF Statistic	p-value	Crit. 1%	Crit. 5%	Crit. 10%
Indicateur_Moyen_Brut_1	-3.983	0.0015	-3.546	-2.912	-2.594
Indicateur_Moyen_Brut_2	-1.834	0.364	-3.548	-2.913	-2.594
Indicateur_Moyen_Brut_3	-2.229	0.196	-3.566	-2.920	-2.598
Indicateur_Moyen_Brut_4	-5.325	0.00001	-3.546	-2.912	-2.594
Indicateur_Moyen_Brut_5	-4.431	0.00026	-3.546	-2.912	-2.594

Les segments 1, 4 et 5 sont stationnaires au seuil de 5%, tandis que les segments 2 et 3 affichent des p-values largement supérieures, indiquant une non-stationnarité.

Pour garantir la validité des modèles, les transformations suivantes ont été appliquées :

- **Segment 2** : taux de variation (pct\_change),
- **Segment 3** : log-différence.

La stationnarité des séries transformées a été confirmée par les ADF post-transformation et par l'analyse des fonctions d'autocorrélation (ACF).

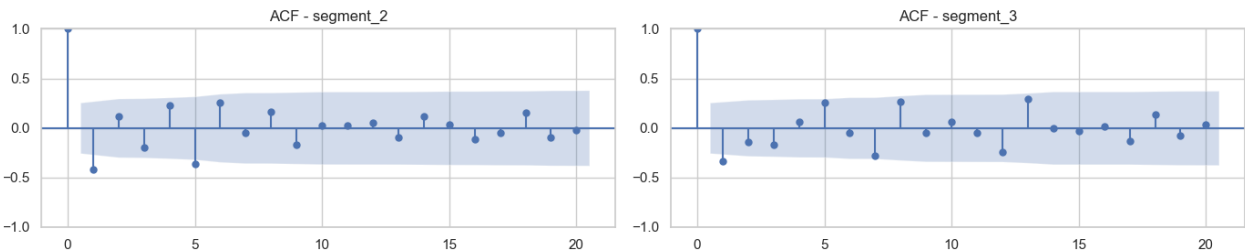


Figure 2: ACF des séries transformées (segments 2 et 3)

L'ACF post-transformation confirme la stationnarité des séries utilisées : les autocorrélations chutent rapidement et deviennent non significatives après quelques lags, ce qui suggère l'absence de dépendance persistante dans le temps.

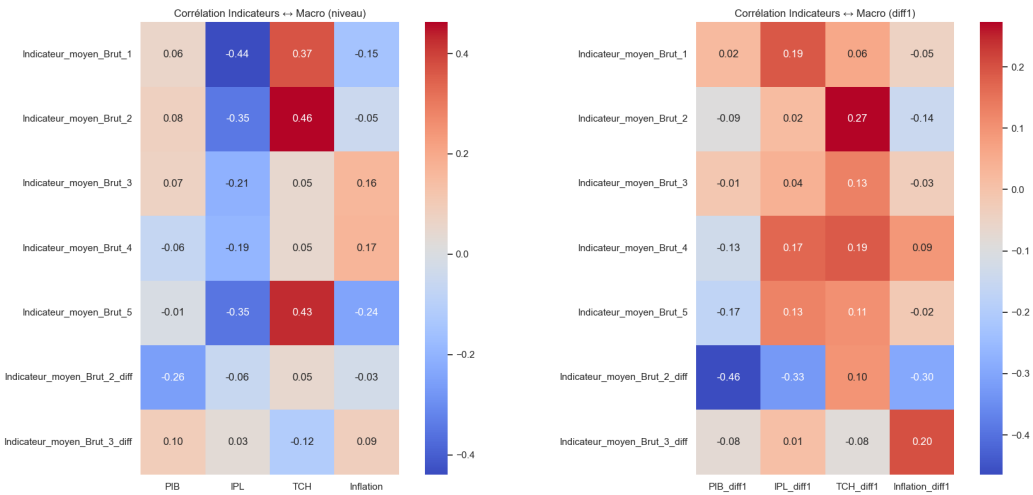


Figure 3: Corrélations entre les CCF par segment et les variables macroéconomiques (observées et différenciées)

Les corrélations observées confirment le lien explicatif potentiel des variables macroéconomiques dans la modélisation. De plus, la différenciation des variables permet d'augmenter certaines corrélations, suggérant une meilleure dynamique commune après transformation.

4 Modélisation économétrique

4.1 Analyse des résidus

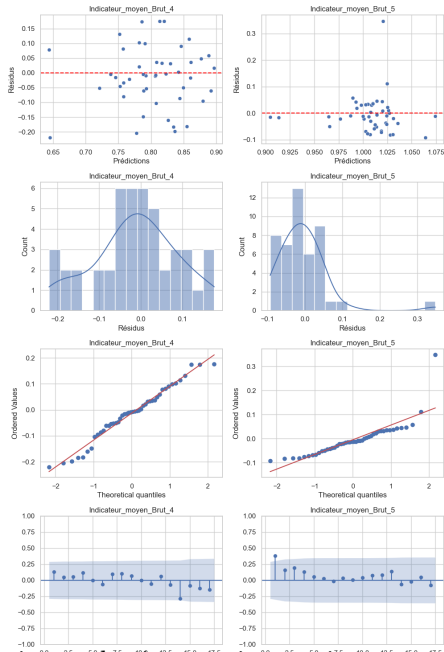
Dans le cadre de la vérification des hypothèses du modèle MCO, nous avons analysé l'autocorrélation des résidus pour chacun des cinq segments de CCF. La figure 4 présente les fonctions d'autocorrélation (ACF) des résidus issus des modèles ajustés ainsi que leurs distributions.

Pour la majorité des segments, les résidus restent contenus dans l'intervalle de confiance à 95%, suggérant une indépendance globale. Quelques pics (notamment sur les segments 2 et 5) laissent penser à une autocorrélation faible, pouvant altérer l'efficacité des estimateurs.

Ces résultats confirment globalement l'hypothèse d'indépendance, mais justifient l'exploration de modèles alternatifs pour les segments sensibles.

4.2 Sélection des variables explicatives

Une base macro enrichie a été construite à l'aide de dérivations économiques classiques : interactions, moyennes glissantes, accélérations, ratios structurels et indicateurs d'environnement (crise, post-COVID). Après nettoyage, 13 variables explicatives ont été retenues pour l'analyse :



Variable	Interprétation économique
<b>Ratios macroéconomiques</b>	
PIB_to_IPL	Efficacité économique : PIB par unité d'investissement
TCH_to_Inflation	Gains de productivité face à l'inflation
IPL_to_TCH	Rendement du capital par productivité du travail
<b>Dynamiques temporelles</b>	
TCH_ma3	Moyenne mobile de TCH sur 3 trimestres
TCH_volatility	Volatilité de la productivité : risque macro
Inflation_acc	Accélération de l'inflation (ruptures de rythme)
Inflation_growth_3q	Hausse glissante de l'inflation sur 3 trimestres
<b>Interactions économiques</b>	
PIBxInflation	Surchauffe potentielle : PIB et inflation croisés
InflationxTCH	Chocs d'offre conjoints entre prix et productivité
<b>Indicateurs de contexte</b>	
recession_flag	Trimestres de récession (PIB < 0)
pre_covid	Périodes avant mars 2020 (pré-crise Covid)

Table 2: Variables sélectionnées et leur sens économique

4.3 Analyse de colinéarité

L'analyse du facteur d'inflation de variance (VIF) révèle une colinéarité modérée mais acceptable pour l'ensemble des variables (aucune VIF > 5 hors constante). Les interactions TCH\_ma3, TCH\_volatility ou PIBxInflation sont informatives mais corrélées.

Table 3: Extrait des scores VIF (variables explicatives)

Variable	VIF	Remarque
TCH_diff1	4.71	corrélation à TCH_ma3
TCH_volatility	3.00	dynamique utile
PIBxInflation	2.12	interaction logique
pre_covid	1.79	rupture structurelle
PIB_to_IPL	1.10	ratio explicatif

4.4 Régression pondérée WLS

Des modèles WLS ont été estimés par segment à partir des variables macro transformées, pondérés par les poids de notes (PourcNoteCohorte5). Les résultats montrent des performances variables selon les segments :

- **Segment 1** :  $R^2 = 0.32$  (faible, peu de signifiante)
- **Segment 2** :  $R^2 = 0.47$  (meilleur modèle pondéré)
- **Segment 3** :  $R^2 = 0.44$ , modèle stable
- **Segment 4** :  $R^2 = 0.47$ , bonnes performances

- **Segment 5** :  $R^2 = 0.31$ , améliorabile

La majorité des segments présentent des résidus centrés, faiblement autocorrélés, mais globalement non normaux, justifiant le recours à d’autres approches robustes.

#### 4.5 Modèles alternatifs

Afin de valider la robustesse de notre démarche et d’explorer des pistes complémentaires, nous avons testé plusieurs modèles alternatifs, en particulier sur les segments pour lesquels les régressions pondérées classiques montraient des limites (hétéroscédasticité, multicollinéarité, faible  $R^2$ , etc.).

Les modèles évalués incluent :

- **Régression Ridge** : utile pour stabiliser les coefficients en présence de multicollinéarité, mais peu performante sur nos données (fort biais).
- **Régression robuste (Huber)** : testée pour atténuer l’effet des valeurs extrêmes, notamment pour le segment 5, mais les résultats sont décevants sur la majorité des segments.
- **Random Forest Regressor** : performant en  $R^2$ , mais modèle non explicable, ce qui limite son usage dans un contexte réglementaire IFRS9.
- **Partial Least Squares (PLS)** : bien adapté pour réduire la dimension sous forte colinéarité, avec des  $R^2$  allant de 0.26 à 0.71 selon les segments. Mais, la perte d’interprétabilité limite son usage en production.

Pour le segment 5, dont les résidus étaient non normaux, une évaluation comparative de modèles alternatifs a été menée afin d’identifier ceux susceptibles d’améliorer la normalité des erreurs.

Table 4: Comparaison des performances sur le segment 5

Modèle	$R^2$	Shapiro	Jarque-Bera	Normalité
Random Forest	0.83	< 0.001	< 0.001	
Huber Regressor	-0.12	< 0.001	< 0.001	
Régression Quantile	0.22	< 0.001	< 0.001	

Malgré un score de performance élevé pour le modèle Random Forest, l’absence de normalité des résidus et la faible transparence du modèle ne permettent pas une adoption. Les régressions robustes, quant à elles, n’apportent pas d’amélioration significative. En conséquence, le modèle de régression linéaire pondérée (WLS), bien que plus simple, reste privilégié pour la projection forward-looking, en raison de son interprétabilité et de sa conformité aux exigences méthodologiques d’IFRS 9.

#### 4.6 Comparaison des modèles par score

Enfin, un test systématique a été mené pour sélectionner le meilleur nombre de variables explicatives ( $k$ ) par segment. Le tableau ci-contre synthétise les résultats optimaux.

Les modèles finaux obtenus (WLS ou OLS selon les segments) seront utilisés dans la phase de projection forward-looking.

Table 5: Scores  $R^2$  des meilleurs modèles par segment (OLS vs WLS)

Segment	Modèle	$k$	$R^2$
Segment 1	WLS	11	0.315
Segment 2	WLS	11	0.468
Segment 3	OLS	11	0.471
Segment 4	OLS	11	0.526
Segment 5	WLS	11	0.311

### 5 Prédictions Forward-Looking à 3 ans

#### 5.1 Scénarios économiques projetés (2024–2026)

Trois types de scénarios économiques ont été définis à partir de prévisions de PIB, d’inflation, d’IPL et de TCH :

- **PESSimiste** : croissance ralentie, inflation élevée, tension sur le marché de l’emploi ;
- **CENTral** : hypothèse modérée, correspondant au consensus ;
- **OPTImiste** : croissance dynamique, inflation maîtrisée, repli du chômage.

Chaque scénario est défini trimestre par trimestre pour les années 2024 à 2026.

#### 5.2 Méthodologie de projection

Deux approches ont été comparées :

1. Une approche basée sur les modèles OLS/WLS construits pour chaque segment à partir des variables macro (différences et lags) ;

- Une approche alternative par différents modèles non linéaires comme par exemple le Gradient Boosting (*à titre informatif uniquement*), car ces modèles ne respectent pas les hypothèses statistiques nécessaires à l'interprétation linéaire, mais il présente de meilleures performances prédictives ce qui était selon nous intéressant à observer à titre comparatif.

À partir du dernier trimestre observé, les prédictions ont été réalisées sur 12 trimestres à venir sous chacun des scénarios.

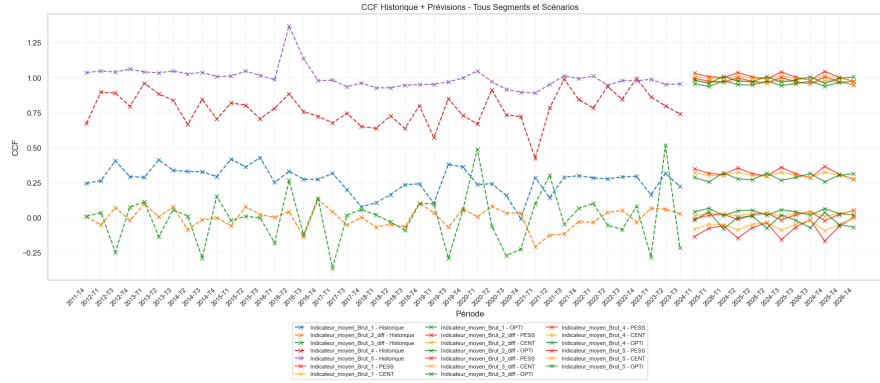


Figure 5: *Prédictions du CCF par segment pour les trois scénarios économiques (OLS/WLS)*

### 5.3 Projections agrégées pondérées par scénario

En utilisant un jeu de poids subjectifs ( $PESS = 25\%$ ,  $CENT = 50\%$ ,  $OPTI = 25\%$ ), une projection forward-looking unique a été construite pour chaque segment.

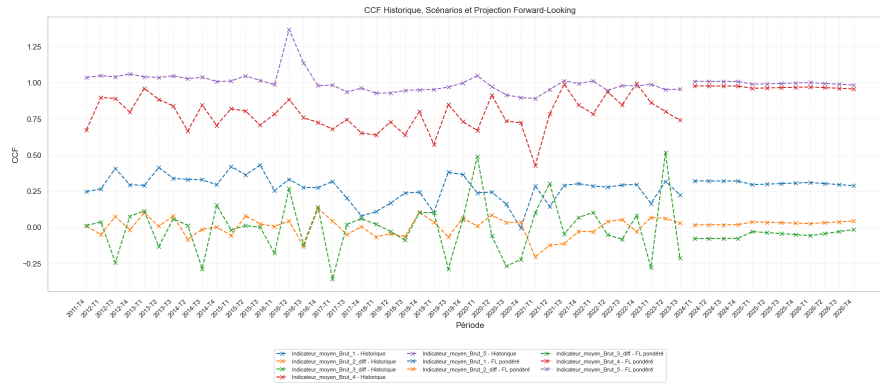


Figure 6: *Projection forward-looking pondérée des CCF (OLS/WLS)*

## 6 Conclusion

Ce projet a permis de mettre en œuvre une méthodologie robuste de modélisation du **Credit Conversion Factor (CCF)** à partir de variables macroéconomiques, dans une optique *forward looking* conforme aux exigences d'IFRS9.

Nous avons combiné :

- une analyse rigoureuse de la stationnarité et des corrélations,
- une modélisation économétrique par segment (OLS, WLS),
- des enrichissements macroéconomiques pertinents (ratios, interactions, effets de crise),
- et une projection sur 3 ans fondée sur des scénarios macroéconomiques multiples.

Les résultats obtenus montrent une capacité raisonnable à expliquer les dynamiques des CCF par segment, avec des performances variables selon la nature des données et les modèles utilisés.

**Perspectives** : une amélioration des performances pourrait être obtenue en :

- affinant la granularité temporelle (mensualisation, rolling windows),
- enrichissant les scénarios macroéconomiques (hypothèses de stress plus complexes),
- intégrant des modèles non linéaires plus avancés (XGBoost, réseaux neuronaux),
- ou en travaillant sur des données de cohortes réelles avec CCF observés.

Ce travail constitue une base solide pour la mise en place d'une approche EAD sensible à la conjoncture dans un cadre réglementaire.