



## Apports de Google Trends pour le suivi de l'activité économique au Maroc en temps réel

Les crises successives qui ont secoué le monde au cours de ces dernières années ont mis à rude épreuve les capacités des Etats à concevoir et à déployer efficacement des dispositifs de soutien, à fort impact et adaptés aux priorités sectorielles, sociales et territoriales. Ce contexte particulier a remis à l'ordre des priorités l'importance stratégique de développer des outils de monitoring de l'activité économique capables d'informer les décideurs de manière anticipée et en temps réel sur les tendances économiques à très court terme, les retournements cycliques, les risques potentiels à entrevoir... Ceci dans l'objectif ultime d'orienter la conception des mesures à entreprendre et d'optimiser l'intervention publique dans un contexte d'incertitude persistante.

Ainsi et dans l'objectif de combler le retard de publication des indicateurs économiques conventionnels, plusieurs institutions se sont intéressées au développement d'outils et de plateformes de Nowcasting (estimateurs en temps réel de l'activité économique), profitant du développement massif des informations dites « non conventionnelles à haute fréquence ».

Dans le cadre des travaux analytiques de la DEPF mobilisant les Big Data, le présent Policy Brief explore une approche innovante permettant d'utiliser les indices de recherche de Google Trends pour établir des prédictions économiques à court-terme, trimestrielles ou même mensuelles.

S'articulant autour de trois parties, le Policy Brief donne, dans un premier temps, un aperçu synthétique et conceptuel du contenu et des opportunités qu'offre l'utilisation des données de la plateforme Google Trends (GT). La deuxième partie est consacrée, quant à elle, à la présentation des méthodes de traitement statistique appliquées aux séries de données dans l'objectif de corriger les biais techniques, notamment les ruptures méthodologiques, les données manquantes ainsi que les effets saisonniers. Enfin, la troisième partie est focalisée sur l'application du modèle de tracking retenu et à l'analyse de la pertinence et de la robustesse des estimations obtenues.



**Rachida KHTIRA**  
Data Scientist



**Ahlam ER-RAHMANI**  
Chargée d'études  
socio-économiques  
des territoires



## 1. Méthodologie du travail : données utilisées

Depuis quelques années, plusieurs études se sont intéressées aux tendances des recherches des utilisateurs de Google comme sources d'information susceptibles d'améliorer les prévisions<sup>1</sup>. Leur disponibilité, leur fréquence ainsi que leur richesse informationnelle les ont rendues d'une grande utilité pour les conjoncturistes. L'exploitation des indices de recherches de Google Trends a gagné en termes de pertinence, notamment, avec la généralisation progressive de l'accès à Internet dans le monde entier et l'affirmation du moteur de recherche de Google comme leader mondial dans le domaine. En effet, les requêtes des internautes ont été utilisées comme reflet de l'activité des acteurs économiques et pour améliorer la production des statistiques existantes.

### 1.1. Indices de recherche de Google Trends

Lancé en 2006, Google Trends est un outil qui permet de générer des indicateurs d'intensité de recherche pour un mot-clé donné au niveau des moteurs de recherche de Google. Les résultats sont présentés sous forme d'indices de recherche normalisés en fonction de la période choisie et du lieu spécifié dans la requête. En effet, le nombre de recherches pour un mot clé donné est rapporté au nombre total de recherches effectuées pour la période et le lieu précisé, et ce dans un souci de prendre en considération des facteurs exogènes qui auraient pu impacter les opérations de recherche d'une manière générale (augmentation du nombre d'internautes, lenteurs occasionnelles des serveurs...). Les chiffres obtenus sont, par la suite, transformés en un score de 0 à 100 en rapportant chaque valeur à la valeur maximale de la série<sup>2</sup>.

La plateforme Google Trends permet, également, de générer des indices de recherche agrégés par catégorie de mots clés ou par sujet. En effet, les requêtes basées sur des mots clés sont spécifiques à chaque langue, et pourraient regrouper des recherches avec des significations et

objectifs différents. Par exemple, les séries de Google Trends pour le mot-clé "Apple" peuvent confondre les recherches sur le fruit et celles faites sur l'entreprise. Par contre, les catégories et les sujets sont harmonisés dans toutes les langues gérées par Google et fournissent une représentation de l'intérêt général des recherches pour des événements et domaines de recherche plutôt que pour des mots spécifiques.

Ainsi, l'outil Google Trends classe les requêtes de recherche dans environ 1200 catégories. L'affectation de la recherche à une catégorie est probabiliste pour prendre en considération que des mots peuvent avoir plusieurs significations. Concernant les sujets, Google Trends crée des sujets qui regroupent plusieurs requêtes effectuées sur le moteur de recherche Google en fonction de leur objet et de leur signification.

Les indices de Google Trends sont calculés à l'aide d'une méthode d'échantillonnage propre à l'outil et ce, dans un souci d'améliorer le temps de réponse des requêtes<sup>3</sup> et seules les requêtes avec un volume significatif sont représentées. Les données sont mises à jour quotidiennement pour toutes les langues gérées par Google et sont disponibles sur plusieurs fréquences selon la période spécifiée (mensuelles à partir de 2014, hebdomadaire à partir de 2016 et en temps réel pour les 7 derniers jours).

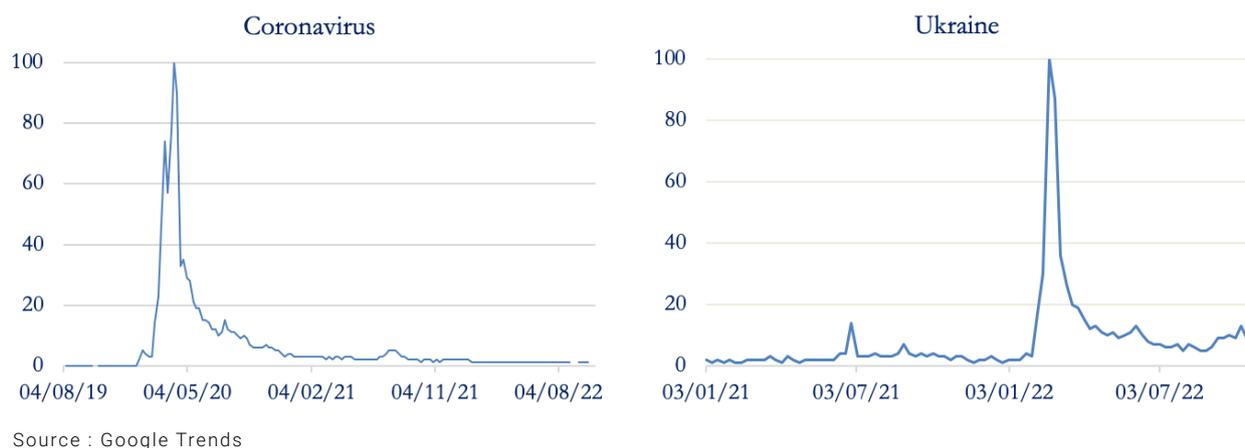
La figure 1 illustre deux exemples d'indices de recherche pour les termes «Coronavirus» et «Ukraine» au Maroc durant les trois dernières années. On constate que l'indice de recherche concernant le terme «Coronavirus» a considérablement augmenté durant la période du confinement et a atteint son pic durant la semaine du 15/03/2020 qui correspond exactement à la première semaine de confinement au Maroc. En revanche, le terme relatif à l'«Ukraine» a enregistré son plus haut niveau le 20 février correspondant au déclenchement de la crise en Ukraine.

<sup>1</sup> Voir chapitre 1.2. Opportunités d'utilisation des données de Google Trends.

<sup>2</sup> <https://support.google.com/trends/answer/4365533?hl=fr>

<sup>3</sup> Selon Google Trends, les indices de recherche sont issus d'un échantillon représentatif. Des milliards de recherches effectuées chaque jour sur le moteur de recherche de Google dans un souci d'améliorer le temps de réponse des requêtes.

Figure 1 : Indices de recherche sur le Web au Maroc correspondant aux termes « Coronavirus » et « Ukraine »



Source : Google Trends

## 1.2. Opportunités d'utilisation des données de Google Trends

Au cours des dernières années, divers travaux ont suggéré la possibilité d'enrichir le diagnostic conjoncturel en exploitant la fréquence de certains termes dans les recherches des internautes. En effet, et compte tenu de la généralisation d'Internet, les recherches sur le web peuvent refléter l'état de l'activité économique d'une manière globale ou pour certains acteurs économiques. D'où l'opportunité d'exploiter ces données pour des fins de prédiction en temps réel (appelé nowcasting).

Le nowcasting est défini comme la prédiction du présent, du futur très proche et du passé récent. Son principe de base est l'exploitation des informations disponibles en temps réel et éventuellement à haute fréquences afin d'obtenir une "estimation précoce" de certains agrégats économiques avant même que les chiffres officiels ne soient disponibles<sup>4</sup>.

En 2009, Choi et Varian ont exploité les recherches Google pour prédire plusieurs indicateurs économiques, notamment les allocations de chômage, la demande d'automobiles et les destinations de vacances aux Etats-Unis. Ils ont constaté que ces indices de recherche sont souvent corrélés avec divers indicateurs économiques et peuvent être utiles pour la

prédiction économique à court terme. Par exemple, le volume de requêtes de recherche sur les ventes d'automobiles au cours de la deuxième semaine de juin de 2009 peut être utile pour prédire les ventes d'automobiles au mois de juin qui ne seront publiées officiellement que plusieurs semaines plus tard en juillet<sup>5</sup>.

En 2015, dans une note de conjoncture de l'INSEE, Bortoli et Combes ont testé ce pouvoir prédictif des séries issues de Google Trends pour la prédiction des dépenses mensuelles des ménages en biens ou en services. Ils ont conclu qu'il est possible de mettre en évidence des corrélations positives entre la fréquence de recherche de certains termes et les comportements d'achat finalement observés pour quelques postes tels que l'habillement, les articles de sport ou l'équipement du logement<sup>6</sup>.

Par ailleurs, dans un papier intitulé « Suivi de l'activité en temps réel avec Google Trends » publié en 2020<sup>7</sup> par le département économique de l'OCDE, Nicolas Woloszko a proposé ce qu'il a appelé le « weekly tracker » pour le suivi hebdomadaire de l'activité économique pour 46 pays de l'OCDE, du G20 et des pays partenaires en utilisant les indices de recherche de Google Trends. Dans la plupart des pays, le modèle développé a montré sa capacité à prédire les cycles économiques.

<sup>4</sup> Marta Bañbura, Domenico Giannone, Michele Modugno and Lucrezia Reichlin, " Now-caSting and the real-time data flow ", 2013

<sup>5</sup> Choi, H., and H. Varian. "Predicting the present with Google trends", 2012.

<sup>6</sup> Didier Blanchet, Pauline Givord, " Big data, official statistics and measuring the economy", INSEE Références, édition 2017.

<sup>7</sup> Nicolas Woloszko "Tracking activity in real time with Google Trends", 2020.

### 1.3. Extraction des données

Le présent travail exploite les recherches basées sur les catégories et les sujets. Ainsi, 85 catégories ont été sélectionnées, parmi 1200 catégories incorporées dans l'outil Google Trends, additionnées à 26 sujets de recherche. Les thèmes choisis regroupent :

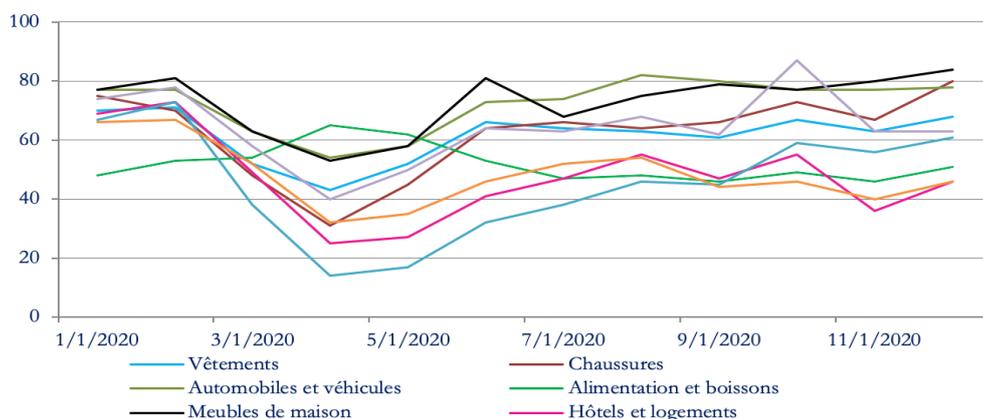
- Des informations sur les requêtes liées aux biens de consommation (les aliments et boissons, les véhicules, les appareils ménagers, ...), aux services (les voyages, les sports, les restaurants, les divertissements, ...) qui représentent une part importante des emplois du PIB.
- Les intensités de recherche sur l'emploi et le chômage.
- Les recherches effectuées dans le cadre de certaines activités industrielles : le transport maritime, les équipements agricoles, .... Ces types de recherches peuvent fournir des informations sur l'évolution de l'offre.

- D'autres variables liées à la situation économique et financière : crise, prêts, faillite, ... pour capturer les fluctuations économiques et financières.

L'extraction a été effectuée grâce à un programme python<sup>8</sup> qui permet d'extraire des données à partir de Google Trends de manière automatisée. Ledit programme permet d'extraire l'ensemble des indices mensuels de recherches correspondant aux sujets et catégories pour la période spécifiée remontant à 2004.

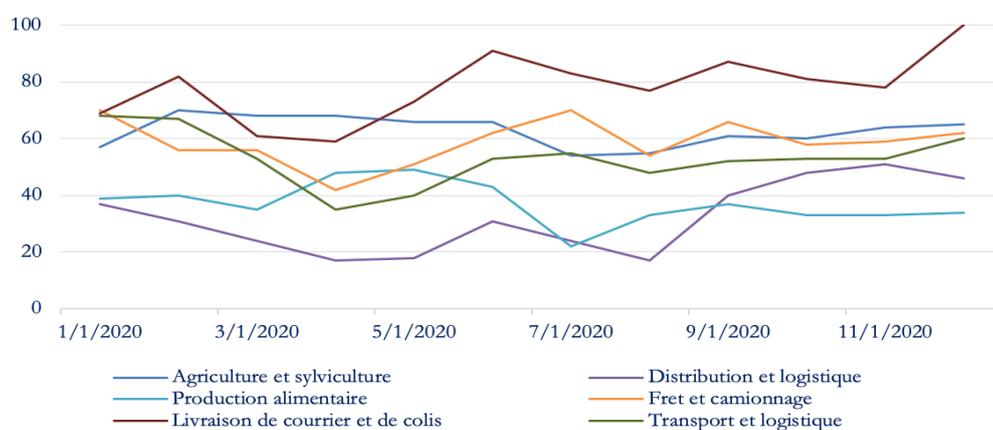
Les figures 2 et 3 illustrent des exemples d'indices de recherches liés à deux groupes de thématiques : la consommation et l'activité industrielle. Ces deux graphiques attestent d'une baisse considérable des recherches liées à la consommation et à l'activité industrielle durant la période de confinement en 2020 à l'exception des recherches liées aux produits alimentaires (alimentations et boisson, production alimentaire) qui ont maintenu un trend haussier durant ladite période.

Figure 2 : Les indices de recherches liées à la consommation au Maroc durant l'année 2020



Source : Google Trends

Figure 3 : Les indices de recherches liées à l'activité industrielle au Maroc durant l'année 2020



Source : Google Trends

<sup>8</sup> Le programme python utilise une pseudo-API PyTrends <https://pypi.org/project/pytrends/>

## 2. Traitement des données

Les données recueillies de Google Trends ont été soumises à un processus de requalification dans l'objectif de préparer les données pour la modélisation. Il s'agit, de correction des biais statistiques éventuels ainsi que l'élimination des imperfections entraînées par des changements méthodologiques introduits au niveau de la plateforme.

### 2.1. Considérations méthodologiques pour l'utilisation des données de Google trends

L'analyse temporelle des données brutes issues de GT permet de constater une tendance à la baisse des séries chronologiques ainsi que des ruptures en 2011, 2016 et 2022<sup>9</sup> dues probablement aux changements effectués au niveau des moteurs de recherches de Google opérés en ces deux années pour améliorer le service aux utilisateurs<sup>10</sup>.

Les séries temporelles sont générées, également, à partir d'un algorithme d'échantillonnage automatique sur l'ensemble des termes de recherche et sont soumises par la suite à des procédures de retraitement. Ce processus peut produire des séries différentes à chaque opération d'extraction de données, notamment à intervalles de temps très espacés. De plus, la liste des thèmes ainsi que les requêtes correspondantes sont mises à jour en permanence en fonction des recherches effectuées au fil du temps.

Ces sources de rupture de données pourraient impacter la stabilité des modèles développés. De ce fait, les indices de recherches nécessitent une attention particulière lors de la phase de traitement afin de minimiser les effets indésirables et réduire l'impact de ces imperfections sur les modèles d'estimation développés. Dans ce sens, plusieurs solutions ont été mises à l'œuvre pour l'ajustement des séries de GT à savoir la correction de la tendance à la baisse, le traitement des années de rupture et l'élimination de la saisonnalité.

### 2.2. Traitement des points de rupture (break points)

Les données de Google Trends ont fait l'objet d'un changement dans le processus de localisation géographique et d'une amélioration du système de collecte des données courant les années 2011, 2016 et 2022. Ces changements génèrent des ruptures de tendance générale pour l'ensemble des séries de données pour ces dates.

Le traitement des années post-rupture s'est fondé sur le raccordement des séries sur la base de coefficients de redressement pour chaque série de données. Ainsi, un premier coefficient est appliqué à l'ensemble des valeurs qui précèdent 2011. Un deuxième coefficient est appliqué aux valeurs comprises entre 2011 et 2016. Pour la rupture générée à partir de 2022, la profondeur temporelle correspondante (9 mois) ne permet pas d'établir de correction. L'opération de correction a permis de lisser les points de rupture de l'ensemble des séries. Les séries résultantes sont par la suite normalisées à une valeur maximale de 100.

### 2.3. Trimestrialisation des séries traitées

Le modèle à établir pour la variable cible « le taux de croissance du Produit Intérieur Brut » est à fréquence trimestrielle. Ainsi, dans l'objectif de mettre en cohérence les séries mensuelles extraites de Google Trends avec le besoin du modèle, ces indices ont fait objet d'opération de trimestrialisation basée sur la moyenne arithmétique des données de chaque trimestre.

### 2.4. Détermination de la composante temporelle commune : Transformation logarithmique

La composante temporelle correspond aux effets des différents facteurs exogènes qui pourraient, probablement, impacter les intensités des recherches mesurées par Google Trends à une date donnée hors l'intérêt

<sup>9</sup> Les changements opérés correspondent à :

- en janvier 2011 : changement du processus de localisation géographique
- en janvier 2016 et janvier 2022 : amélioration du système de collecte de données.

<sup>10</sup> Björn Bokelmann, , Stefan Lessmann " Forecasting tourism in Germany using Google Trends", 2019.

à un sujet donné ou à un domaine d'analyse. Cette composante pourrait découler de l'augmentation exponentielle du nombre d'utilisateurs du moteur de recherche de Google, des dysfonctionnements techniques occasionnels de la plateforme ou bien d'autres facteurs. Ce biais est commun à l'ensemble des séries extraites. La démarche suivante a été adoptée pour éliminer cette composante.

Ainsi, on considère la construction des séries sous la forme suivante :

$$SVI_{ct} = \frac{SV_{ct}}{SVT_t} * C_c \quad \text{Avec} \quad C_c = \left( \frac{\text{Max } SV_{ct}}{SVT_t} \right)^{-1}$$

$SVI_{ct}$  : volume de recherche initial pour une catégorie donnée à l'instant t.

$SV_{ct}$  : volume de recherche pour une catégorie donnée à l'instant t.

$SVT_t$  : volume de recherche totale à l'instant t.

$C_c$  : composante commune.

La tendance commune peut être décomposée linéairement en procédant à

une transformation logarithmique des séries de volumes de recherche.

$$svi_{ct} = \log(SVI_{ct}) = sv_{ct} - svt_t + C_c$$

Les tendances issues de chaque série sont ensuite redimensionnées par la méthode de standardisation (centrée réduite) pour qu'elles soient comparables sur le temps, avec une moyenne et écart type identiques.

La composante commune peut être, ainsi, générée en deux étapes : extraction de la tendance centrale des séries filtrées par le biais du filtre Hodrick et Prescott HP et l'agrégation des différentes tendances résultantes en une principale composante à travers l'analyse en composante principale. Ce facteur permet de restituer le maximum de l'information contenue dans l'ensemble des séries.

#### Estimation de la tendance globale des séries

Le facteur commun est soustrait des variables logarithmiques par le biais du filtre Hodrick et Prescott (HP) afin d'éliminer les tendances de long terme.

### Le filtre de Hodrick et Prescott (HP)

Le filtre Hodrick Prescott permet de décomposer une série temporelle ( $k_t$ ) en deux composantes : une tendance permanente non stationnaire ( $k_t^T$ ) qui retrace le trend à long terme et un cycle transitoire stationnaire ( $k_t^C$ ) qui relate les variations à court terme respectivement

$$k_t = k_t^T + k_t^C$$

Le filtre HP est fondé sur l'optimisation dynamique suivante :

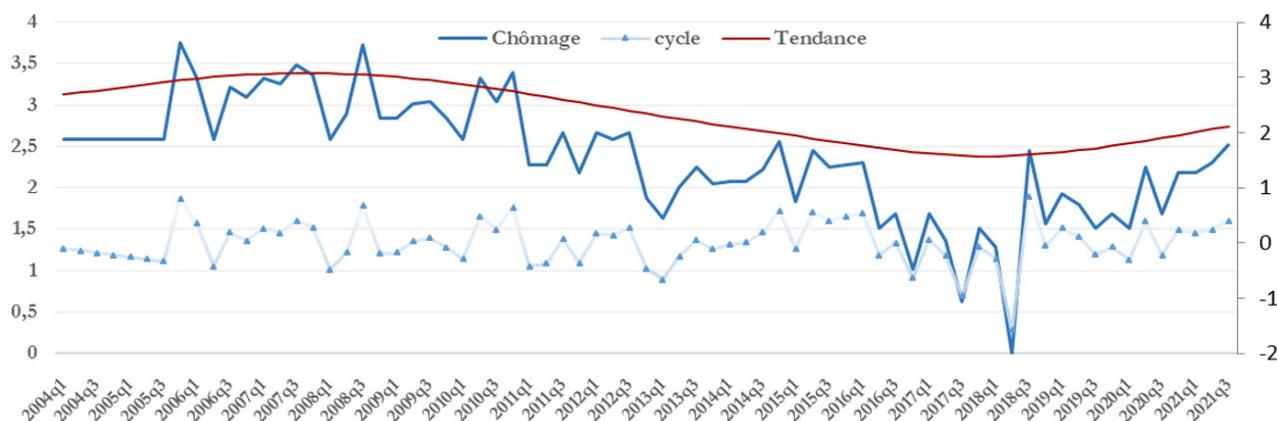
$$\text{Min}_{\{K_t^T\}} \sum_t^T (K_t - K_t^T)^2 + \lambda \sum_t^T [(K_{t+1}^T - K_t^T) - (K_t^T - K_{t-1}^T)]^2$$

$\lambda$  : est le paramètre de lissage qui partage les variations entre la tendance et la composante cyclique.

Si  $\lambda=0$ , le cycle dans ce cas est nul et la tendance absorbe toute la variabilité de la série. Par contre si  $\lambda$  est proche de l'infini, dans ce cas la variance est due principalement au cycle.

$\lambda$  peut prendre plusieurs valeurs selon la fréquence des données. Il est égal à 1600 pour les données trimestrielles.

Figure 4 : Série des Indemnités de chômage décomposée en cycle et tendance



Source : Calcul de la DEPF sur la base des données de Google Trends

### Analyse en composante principale

La composante commune est calculée sur la base de la méthode d'Analyse en Composante Principale. La 1ère composante principale s'accapare de 56% de la variabilité totale des

données. Ainsi, la variabilité représentée par cette variable a été jugée satisfaisante pour représenter la composante commune. Elle a été par la suite soumise au processus de normalisation.

### Analyse en Composante Principale (ACP)

L'ACP est une méthode d'analyse multidimensionnelle des données composées de variables quantitatives continues ou discrètes. Elle vise à décrire les relations existantes entre un grand nombre de variables observées et à réduire la dimensionnalité des données en passant de  $k$  variables à  $m$  facteurs d'interaction entre les variables numériques. La méthode de l'ACP permet, en effet, de synthétiser l'essentiel de l'information contenue dans les variables par la génération de facteurs non redondants classés par ordre d'importance.

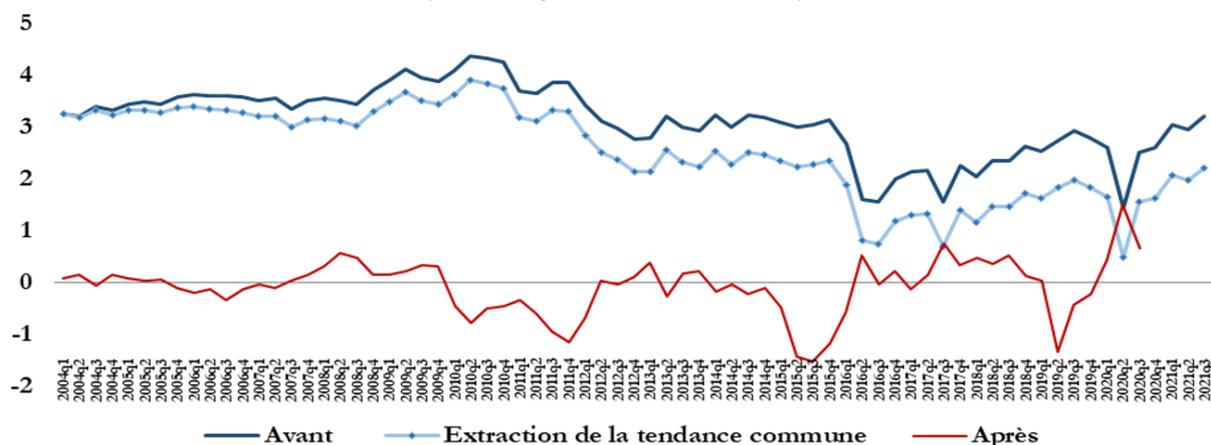
Les facteurs représentent des combinaisons linéaires entre les variables avec la condition de non corrélation. Ceci permet de faciliter l'interprétation et la compréhension des résultats et à éviter le problème de colinéarité entre les variables en cas d'estimation. Ces facteurs sont appelés des composantes principales où la première composante est souvent retenue puisqu'elle contient la majorité de l'information contenue dans les variables.

### 2.5. Dessaisonnalisation des indices de recherche

Contrairement aux sujets de recherche, les catégories de recherche sont plus sensibles aux tendances saisonnières,

d'où la nécessité de les dessaisonnaliser. Ces variables ont été transformées par un glissement annuel (après transformation logarithmique, élimination des tendances de long termes et extraction de la tendance commune).

**Figure 5 : Série de l'indice des recherches relatives à la catégorie « assurances maladie » en logarithme (Avant et après dessaisonnalisation)**



Source : Calcul de la DEPF sur la base des données de Google Trends

### 3. Estimation du modèle

#### 3.1. Méthode d'apprentissage profond : Réseau de neurones

De nos jours, les algorithmes d'apprentissage profond ont acquis une immense popularité dans divers domaines d'application. De nombreux organismes exploitent les forces des réseaux neuronaux pour résoudre des problématiques liées à la modélisation, la classification, la reconnaissance des formes et la prédiction pour différents domaines scientifiques, notamment, en économie, en finance et en médecine. De même, ces algorithmes sont largement utilisés pour les prévisions climatiques et dans le domaine de l'agriculture, en particulier dans la classification des types de cultures et l'estimation de la production agricole<sup>11</sup>.

Les réseaux de neurones se basent sur des modèles statistiques assez complexes, résolus à partir des algorithmes d'apprentissage effectués sur un échantillon de données dit d'entraînement. L'opération d'apprentissage nécessite la mobilisation d'une masse importante de données. Toutefois, la disponibilité des grands échantillons d'entraînement peut être relativement limitée dans plusieurs domaines d'application et même pour certains cas de modélisation le nombre de prédicteurs peut être supérieur au nombre d'observations. Ce cas de figure est récurrent dans le traitement des images, la classification de documents, l'astronomie et les sciences atmosphériques<sup>12</sup>.

Dans le cas du Weekly tracker de l'OCDE, et pour palier au problème de la taille réduite de l'échantillon, il a été procédé à l'enrichissement des données utilisées à travers l'exploitation des indices de recherche relatifs à un échantillon de 40 pays. Toutefois, l'utilisation d'un modèle spécifique à un seul pays semble être plus intuitif puisque les variables choisies pourraient être personnalisées au contexte du pays en question, selon sa culture, ses habitudes et ses institutions.

Pour l'étude en cours, l'information utilisée se rapporte à une centaine d'indices de recherches GT (voir Annexe 1) relatifs au Maroc à partir de janvier 2004 jusqu'à septembre 2022, soit 225 observations mensuelles. La variable cible est le taux de croissance en glissement annuel du PIB trimestriel en volume relatif au Maroc pour la période trimestrielle correspondante (voir la figure 6).

La relation entre les variables de Google Trends et la croissance du PIB est ajustée à l'aide d'un réseau de neurones. L'algorithme de réseau de neurones implémenté est un perceptron multicouche standard initialisé avec les paramètres par défaut dans le logiciel statistique Python scikit-learn. Ce perceptron comprend deux couches cachées et utilise des fonctions d'activation "Relu" et le solveur "Adam". La régularisation L2 est assurée via le paramètre alpha. L'arrêt précoce a été utilisé pour réduire le temps d'apprentissage.

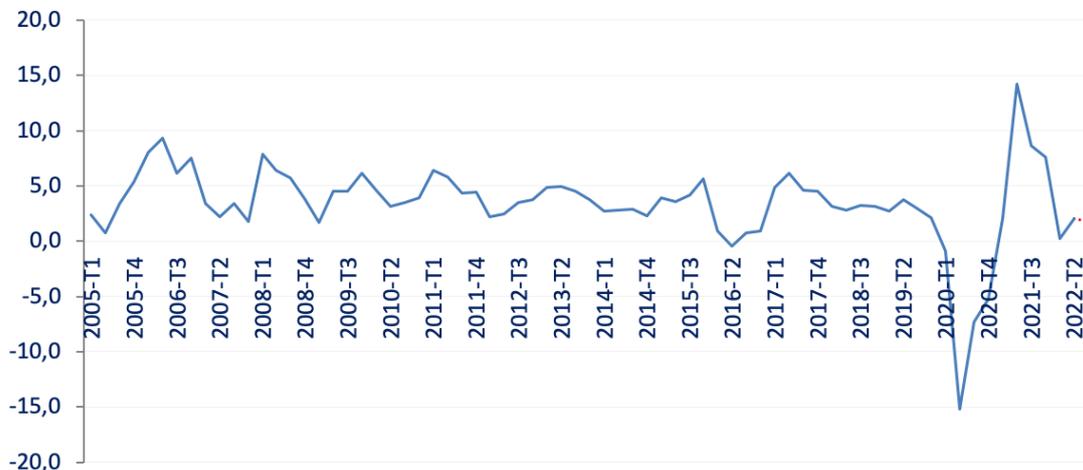
<sup>11</sup> Oludare Isaac Abiodun, Aman Jantan, Abiodun Esther Omolara, Kemi Victoria Dada, Nachaat AbdElatif Mohamed, Humaira Arshad, "State-of-the-art in artificial neural network applications: A survey", Heliyon 4 (2018).

<sup>12</sup> Iain M. Johnstone, D. Michael Titterton, "Statistical challenges of high-dimensional data", 2009.

La définition des hyper paramètres est effectuée manuellement en testant plusieurs architectures et scénarios tout en veillant à avoir un bon niveau d'ajustement avec des performances raisonnables. Les données ont été mises à l'échelle

avant de les soumettre à l'algorithme d'apprentissage. Les performances trouvées sont satisfaisantes comparées avec celles issues de deux algorithmes de régression linéaire régularisée Ridge et Lasso.

**Figure 6 : Taux de croissance, en glissement annuel, du PIB trimestriel marocain en volume**



Source : HCP

### LE PERCEPTRON MULTICOUCHE

Un perceptron multicouche est une structure très particulière des réseaux de neurones. Il consiste en une succession de couches constituées d'unités neuronales. Chaque neurone reçoit des signaux provenant de la couche précédente, effectue un calcul et transmet le résultat à la couche suivante. Il n'existe pas d'interconnexions entre les neurones situés à l'intérieur d'une même couche : les activations des différents neurones sont propagées de la couche d'entrée vers la couche de sortie à travers tous les neurones constitutifs du réseau.

La figure 7 présente un perceptron avec trois couches. La première couche du réseau est la couche d'entrée qui reçoit les variables. Elle contient (n) neurones. La deuxième couche, appelée couche cachée, contient (m) neurones.

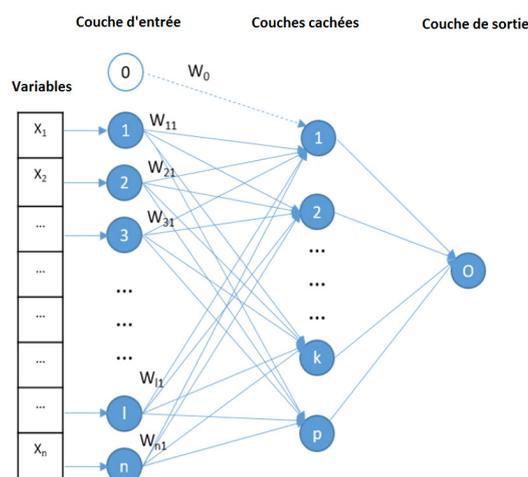
La couche de sortie qui produit le résultat final contient un seul neurone. Chaque neurone j de la couche cachée reçoit une somme pondérée des n entrées, avec n poids  $w_{1j}, w_{2j}, \dots, w_{nj}$ , et avec l'ajout d'un biais  $w_0$  :

$$p_i = w_{1j}x_1 + w_{2j}x_2 + \dots + w_{nj}x_n + w_0$$

Ensuite, une fonction d'activation f est appliquée à cette somme. Un exemple d'une telle fonction est la fonction Relu.

$$f(p_i) = \max(0, p_i)$$

**Figure 7 : Architecture du perceptron multicouche**

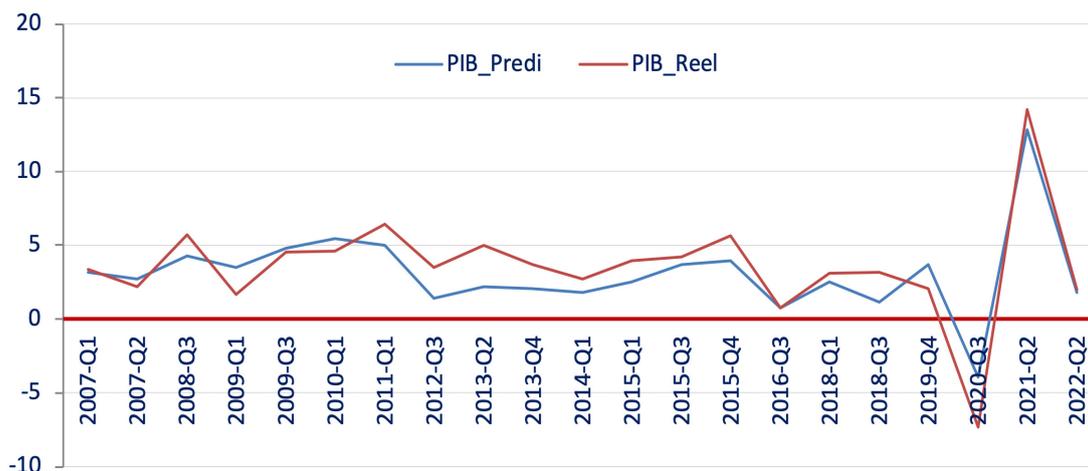


### 3.2. Résultats des estimations pour les données de test

La figure 8 présente les estimations trimestrielles, issues du modèle établi,

appliquées aux données de l'échantillon de test, soit des données non vues par l'algorithme lors de la phase d'apprentissage des données.

**Figure 8 : Estimation immédiate (M-3) de la croissance du PIB trimestriel avec Google trends**



Source : Calcul de la DEPF

Comme le montre la figure 8, le modèle d'estimation du taux de croissance du PIB trimestriel arrive à bien simuler le passé. Il capture aussi bien le signe que la tendance de la croissance du PIB trimestriel pour l'échantillon arrêté.

Le modèle a été, ainsi, utilisé pour l'estimation des trois premiers trimestres de 2022. Les résultats (voir le tableau 1) font état d'un net ralentissement de l'activité économique ce qui s'inscrit dans la tendance des comptes établis pour le

premier et le deuxième trimestre de 2022 et des estimations effectuées par les prévisionnistes pour le troisième trimestre.

Ainsi, pour le premier et le deuxième trimestre la croissance estimée par le modèle est respectivement de 0,8 % et 1,8%, soit des niveaux proches des valeurs arrêtées par les comptes trimestriels publiés par le HCP (0,3% et 2%). Ce net ralentissement de l'activité économique est infiniment lié à la baisse des activités agricoles en lien avec les répercussions de la sécheresse et le

contexte actuel marqué par l'intensification des tensions inflationnistes. Ces chiffres marquent, également, le niveau de résilience affiché par l'économie marocaine notamment celui des activités hors agricoles.

Par ailleurs, la tendance de croissance économique observée au cours de T1-2022 et T2-2022 devrait se poursuivre en T3-2022. Selon les résultats des estimations du modèle, elle s'établirait à 1,3%, marquant ainsi un ralentissement par rapport au

niveau estimé pour le trimestre précédent. Cette tendance reste en ligne avec les résultats enregistrés dans les pays voisins partenaires du Maroc. La zone euro devrait, ainsi, enregistrer une nette détérioration de la croissance économique en T3-2022 pour s'établir à 0,1% contre 0,9% en T2-2022 et 0,7% en T1-2022 en glissement trimestriel. Ce net ralentissement est, notamment prévisible en France (0,2% en T3-2022 contre 0,5% en T2-2022).

**Tableau 1 : Estimation du taux de croissance, en glissement annuel, du PIB trimestriel pour l'année 2022**

Taux de croissance du PIB trimestriel	
1 <sup>er</sup> trimestre 2022(entrainement)	0,8%
2 <sup>ème</sup> trimestre 2022 (test)	1,8%
3 <sup>ème</sup> trimestre 2022(prédiction)	1,3%

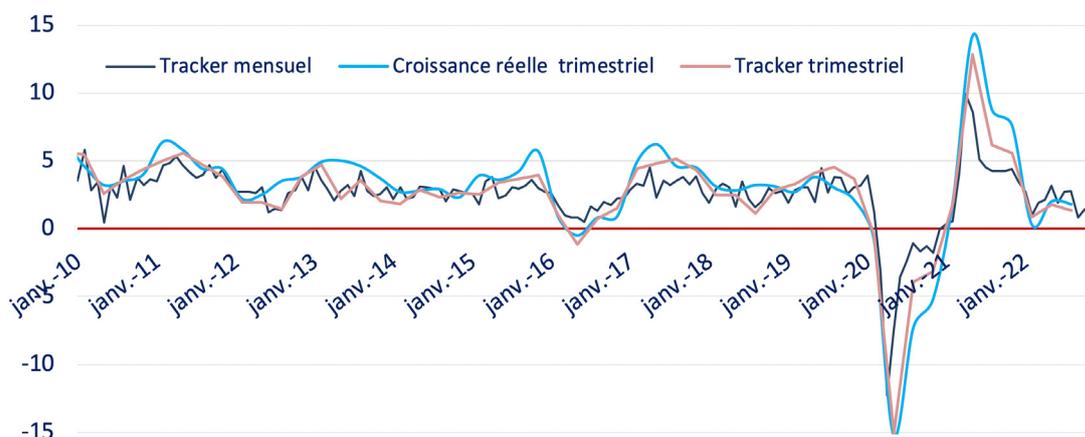
Source : Calcul de la DEPF

### 3.3. Suivi mensuel du PIB

Dans l'objectif de construire un estimateur à haute fréquence de la croissance économique, le modèle trimestriel estimé a été appliqué aux données mensuelles issues de GT. Les données mensuelles ont subi le même processus de traitement appliqué aux données trimestrielles utilisées pour l'estimation du modèle. Il s'agit du traitement des points de rupture, l'élimination de la composante commune et la dessaisonalisation (expliquées dans la partie « traitement des données »).

La figure 9 présente la courbe des données trimestrielles ainsi que celles du tracker mensuel et met en exergue la croissance économique trimestrielle en comparaison avec les deux trackers (mensuel et trimestriel). De surcroît, des points d'interaction entre les prédictions mensuelles et la croissance trimestrielle du PIB sont assez visible dans la figure, ce qui permet de juger de la pertinence de ce Tracker. Le modèle arrive à restituer les fluctuations économiques sur toute la période observée. La chute observée au deuxième trimestre 2020 est bien saisie par le modèle bien qu'avec moindre ampleur.

**Figure 9 : Estimation immédiate (Fin du mois en question) de la croissance du PIB mensuel avec Google Trends**



Source : Calcul de la DEPF

On constate ainsi une tendance à la baisse de l'économie nationale au troisième trimestre 2022 après 2% au deuxième trimestre 2022. Cette baisse est confirmée par la régression

du tracker mensuel qui a diminué de 0,8% en mois de Septembre 2022, après une hausse de 2,8% au mois d'août 2022.

## » Conclusions et pistes d'amélioration

Le modèle du tracking de l'activité économique au Maroc développé dans le cadre de ce travail a permis de capter le signe et la tendance du taux de croissance du PIB au niveau trimestriel sur toute la période observée et témoigne de la capacité des indices de Google Trends (GT) à modéliser d'une manière pertinente les fluctuations économiques passées. Il a permis, en effet, de reproduire les cycles économiques observés au Maroc depuis 2004. A juste titre, le modèle trimestriel a démontré, avant la publication des estimations du HCP, une tendance de ralentissement de la croissance de l'économie nationale au troisième trimestre 2022 après avoir progressé de 2% au deuxième trimestre 2022.

De même et dans l'objectif de construire un estimateur à haute fréquence de la croissance économique, le modèle trimestriel estimé a été appliqué aux données mensuelles issues de GT permettant de restituer d'une manière pertinente les fluctuations économiques

sur toute la période observée offrant ainsi une opportunité d'anticiper l'évolution de l'activité économique à fréquence mensuelle.

En perspective de l'amélioration de ce modèle, plusieurs pistes pourraient être entreprises. Il s'agit en premier lieu de palier au problème de la taille réduite de l'échantillon utilisé, lors de la phase d'apprentissage du modèle, en adoptant un panel de pays qui présentent des similitudes structurelles à l'économie marocaine, notamment en matière de niveau de développement économique, de taux de pénétration d'internet .... Cela permettrait d'améliorer la robustesse des corrélations du modèle et d'affiner les résultats obtenus.

Dans un deuxième temps, il s'agit d'enrichir le dictionnaire des sujets et catégories utilisés dans les estimations à travers l'intégration de nouvelles variables explicatives, notamment celles adaptées aux contextes des pays retenus.

## BIBLIOGRAPHIE

- Nicolas Woloszko "Tracking activity in real time with Google Trends", 2020
- Björn Bokelmann,\* , Stefan Lessmann " Forecasting tourism in Germany using Google Trends", 2019
- Oludare Isaac, Abiodun,Aman, Jantan,Abiodun, Esther Omolara,,Kemi Victoria, Dada,Nachaat, AbdElatif Mohamed, Humaira Arshad , "State-of-the-art in artificial neural networkapplications: A survey", Heliyon 4 (2018).
- Didier Blanchet, Pauline Givord, " Big data, official statistics and measuring the economy", Insee Références, édition 2017
- Marta Bańbura, Domenico Giannone, Michele Modugno and Lucrezia Reichlin, " Now-caSting and the real-time data flow ", 2013
- Choi, H., and H. Varian. "Predicting the present with Google trends", 2012
- Iain M. Johnstone, D. Michael Titterington, "Statistical challenges of high-dimensional data", 2009.
- <https://machinelearningmastery.com/train-neural-networks-with-noise-to-reduce-overfitting/>
- <https://machinelearningmastery.com/how-to-handle-big-p-little-n-p-n-in-machine-learning/>
- <https://support.google.com/trends/answer/4365533?hl=fr>
- <https://github.com/pat310/google-trends-api/wiki/Google-TrendsCategories>
- <https://trends.google.com/trends/explore?date=2019-01-01%202021-11-17&geo=MA&q=Covid>
- <https://pypi.org/project/pytrends/>



## CONTACT

### Adresse

### DEPF

Boulevard Mohamed V. Quartier  
Administratif,  
Rabat-Chellah Maroc



### Téléphone

(+212) 5 37.67.74.15/16

### Online

Email : [depf@depf.finances.gov.ma](mailto:depf@depf.finances.gov.ma)

Site web: [depf.finances.gov.ma](http://depf.finances.gov.ma)

### Impression

Centre d'impression de la DAAG

## DERNIERS NUMÉROS PARUS

### 2022

- N° 37 - Novembre** Contribution des subventions à la lutte contre l'inflation
- N° 36 - Juillet** Enjeux du Green Deal européen pour le secteur agroalimentaire marocain
- N° 35 - Juillet** Les déterminants du sous-emploi au Maroc
- N° 34 - Avril** Performances des secteurs agricole et agroalimentaire sous le prisme de l'export : marges intensive et extensive
- N° 33 - Avril** Enjeux de la crise russo-ukrainienne sur la sécurité alimentaire au Maroc et dans la région MENA
- N° 32 - Avril** Vulnérabilité des importations marocaines : diagnostic et recommandations
- N° 31 - Mars** Impacts potentiels de la crise russo-ukrainienne sur l'économie nationale
- N° 30 - Mars** Analyse genre de la contribution de l'utilisation de la main d'œuvre à l'amélioration du niveau de vie : Analyse retro et prospective à la lumière des recommandations du Nouveau Modèle de Développement
- N° 29 - Février** Potentiel du tourisme interne en tant que levier de relance post covid-19
- N° 28 - Janvier** Positionnement du Maroc dans les chaînes de valeur mondiale