# La modélisation et la prévision macroéconomique en Outre-mer : Synthèse du cas d'utilisation du package *Tresthor* en Polynésie française

## Introduction

**Accroche**  Un modèle quasi comptable est actuellement en place en Polynésie, le modèle TABLO. Ce type de modèle est très utilisé dans les Outre-mer et les pays en développement de par sa simplicité d’utilisation. Il permet en effet d’être exécuté à partir d’un tableur et ouvre donc la possibilité d’appréhender la structure de l’économie à un large public. TABLO est de plus adaptable à des économies qui ne sont pas encore inscrites dans un cadre totalement unifié et structuré. Une des utilisations phare de ce modèle est d’engendrer des comptes rapides pour les années récentes. Cette tâche est importante dans des pays qui ne possèdent pas encore la capacité de fournir des comptes provisoires sur les années récentes dans des délais acceptables. TABLO permet ainsi un développement rapide du système statistique et de son utilisation effective. L’idée derrière TABLO est d’éviter d’estimer des relations de comportements, faute de disposer de séries longues et homogènes, et donc de favoriser des relations comptables et d’utiliser des ratios constant (notamment pour les variables suivantes coefficients techniques, coefficients budgétaires, taux de marge bénéficiaires, propension à consommer) ainsi que des relations sans validation empirique (fonction de demande de monnaie, comportement de détermination des prix par les entrepreneurs individuels).

**Elément déclencheur** : Les limites des systèmes statistiques en Outre mer justifiait donc complètement l’implémentation de modèle quasi comptable comme TABLO dans les années 90 et 2000. Cependant, étant donné le rapide développement de ces territoires, il est important d’étudier si ces revendications sont toujours acceptables. Il se pourrait en effet que le système statistique soit assez mature pour utiliser des modèles économétriques. L’occasion a été donné à l’ISPF de s’investir dans un projet de refonte du système de modélisation à travers une mission de prévision des agrégats macroéconomique à court terme. L’utilisation d’un modèle économétrique a donc été mis sur le devant de la scène pour palier aux insuffisances de TABLO qui est contraint par la multiplicité de ses hypothèses simpliste au regard de la théorie économique. L’intention de l’ISPF n’est pas d’éradiquer le modèle TABLO pour un nouveau modèle plus complexe mais de faire collaborer les deux en favorisant TABLO pour la définition des comptes provisoire et un modèle plus sophistiqué pour la prévision de court terme.

**Définition** : L’objectif est de modéliser l’économie polynésienne afin d’obtenir une prévision macro-économique robuste et parcimonieuse. L’ISPF a décidé de tirer profit des travaux menés récemment à la Direction Générale du Trésor (DG Trésor) sur le modèle Opale et son utilisation pratique. Opale est un système d’équations comptables et économétriques utilisé par la DG Trésor afin de réaliser des prévisions macroéconomiques sur l’économie française à l’horizon de 1 à 2 ans. Son utilisation est relativement complexe par rapport aux moyens technique d’un petit territoire comme la Polynésie française. Toutefois, alors que la DG Trésor utilisait ce modèle sur le logiciel économétrique Troll soumis à l’acquisition d’une licence payante et difficile d’accès, ils ont décidé de muter son exploitation sur le langage de programmation R, utilisable gratuitement sur le logiciel R studio. Par-dessus tout, la *DG Trésor* a construit le package *Tresthor* disponible en open source permettant de traiter tout type de modèle économique de type Opale répondant aux critères de construction exigé par l’outil. De plus, la mise en place de ce package dans le domaine publique a été accompagné par 3 manuels d’utilisation clairs, concis et relativement accessibles. Alors que l’ISPF forme activement ses agents à la programmation sur R et cherche un modèle de prévision économique accessible d’utilisation et adaptable à la spécificité de l’économie polynésienne et des comportements macroéconomiques qui s’y exercent, le choix d’exploitation du package *Tresthor* s’est trouvé être un choix logique et opportun. En bref, *Tresthor,* créé sur la base conceptuelle de la maquette *Opale*, permet de créer des modèles économiques, d’estimer les équations économétriques que le modèle contient sur la base des observations passées, de simuler les comportements des variables endogènes et de résoudre le système d’équation pour effectuer des prévisions pour chacune des variables d’intérêt à l’aune de leur comportement moyen sur le passé.

**Enjeux :** L’enjeu est de construire un modèle économétrique sur la base théorique du modèle *Opale* mais adapté et simplifié pour l’économie polynésienne et de mettre en place l’outil *Tresthor* à l’ISPF à partir de ce modèle. L’objectif est d’analyser la performance d’un tel outil dans le système statistique actuel relativement au modèle TABLO déjà en place. La modélisation et la prévision en Polynésie française, et plus généralement en Outre-mer, pâtissent à la fois de la faible quantité de données collectée, aussi bien en terme de quantité d’indicateurs que de récurrence de collecte, et du manque d’investissement en capital humain et technique dans le développement d’outils appropriés. Il s’agit ainsi d’étudier si un tel projet peut surmonter ces défis ou bien si le pays n’est encore pas assez développé pour accueillir l’outil.

**Problématique** : Pourquoi les travaux de *la Direction Générale du Trésor* sont une aubaine pour les Outre-mers ? Nous verrons dans cette étude que le système statistique n’est pas encore assez développé pour une utilisation effective et immédiate de *Tresthor* mais son implémentation expérimentale dans les instituts de statistique en Outre-mers peut participer à impulser un effort et un investissement dans la modélisation économétrique et dans le développement de la collecte de données macroéconomique pertinentes à l’utilisation de tels modèles.

**Annonce du plan :** Afin de traiter ces points d’intérêts, nous retracerons d’abord la mise en place d’un programme informatique basé sur l’utilisation de *Tresthor* et d’un modèle économétrique adapté à l’économie polynésienne avec les données de l’ISPF. Nous verrons les difficultés rencontrées et les grands travaux qu’il reste à réaliser. Nous verrons ensuite pourquoi il est important d’investir dans l’outil dès à présent, même pour une utilisation effective non immédiate. Nous finirons par énoncer les principales recommandations utiles pour répliquer ce cas d’utilisation de *Tresthor* en Outre-mer ainsi que des pistes de recherches pour améliorer le projet.

## Mise en place de l’outil

### Création du modèle

Le modèle *Opale,* utilisé par la *DG Trésor*, retrace l’équilibre offre-demande des biens et services des comptes nationaux trimestriels en volume. Le modèle comprend principalement une modélisation comptable de l’offre, la prévision portant essentiellement sur les postes de la demande. Le modèle Opale comprend par ailleurs une modélisation complète de la boucle prix-salaires, qui relie les prix de demande aux prix d’offre, eux-mêmes dépendants des coûts de production. La qualité du modèle en prévision à un ou deux ans prime sur ses propriétés à long terme. Le modèle *Opale* n’est pas utilisé pour construire un scénario de long terme ni pour réaliser des évaluations de politique structurelles. La prévision doit alors se fonder prioritairement sur une approche par la demande au contraire d’un modèle de long terme comme Mesange basé sur la demande. Les économistes considèrent en effet que les capacités de production sont figées à court terme et l’offre s’adapte donc aux variations conjoncturelles de la demande pour ce type de délai. À l’inverse d’un modèle construit pour des scénarios de moyen ou long terme et des évaluations de politique économique, le modèle *Opale* cherche à retracer de manière plus fine certaines informations de court terme prenant en compte des variables de cycle (la modélisation de l’investissement prend par exemple en compte l’impact du taux d’utilisation des capacités de production) et des variables expliquant la volatilité conjoncturelle (la température par exemple). Enfin, ce type de modèle permet d’appréhender l’économie de façon complexe sans nécessiter un nombre d’hypothèses démesurés. Ainsi, la résolution du modèle n’est majoritairement pas basée sur la subjectivité du prévisionniste mais bien sur les données recueillies et les choix de spécifications des équations économétriques réalisés par le modélisateur.

Le modèle Opale étant bien trop complexe pour être exploitable en Outre-mer (plus de 500 équations) et nécessitant de plus un système de collecte de donnée très développé, le choix a été fait de partir de la base du modèle simplifié proposé par la *DG Trésor (footnote :* à titre d’exemple dans un manuel retraçant l’utilisation de *Tresthor* avec des données issues de la Grande Bretagne) pour construire un modèle adapté à la Polynésie française. Toutefois, l’utilisation d’un tel modèle, même dans un cadre simplifié, nécessite de nombreuses modifications d’équation afin de pouvoir être opérationnel avec les données d’un territoire d’Outre-mers. Ainsi, ce projet a nécessité de retenir un nombre d’équations très limité et uniquement comptable dans un premier temps avec des variables dont les observations sont facilement accessibles. Dès que le programme informatique a permis de faire fonctionner cette première version de modèle, des équations comptables plus détaillées ont ensuite été incorporées progressivement. Enfin, il s’agissait de remplacer certaines équations comptables par des équations de comportement (économétriques) afin de donner au modèle sa raison d’être.

### Construction du programme informatique

Le modèle est utilisable en pratique grâce à un programme informatique coder sur le langage de programmation R. Ce programme permet d’utiliser les fonctions disponible dans le package Tresthor mais a également d’autres rôles. Dans ce projet, les best pratices ont été appliqués pour structurer le programme. Un fichier nommé main permet de définir les réglages utilisateurs (période de prévision souhaitée) et d’appeler chaque sous programme successivement. Les sous-programmes permettent de charger les librairies utilisées, de charger les fonctions, d’extraire les bases de données et de les fusionner en une seule, de retraiter et créer des variables ainsi que de définir des hypothèses de scénarios pour les variables exogènes, d’utiliser le package Tresthor pour estimer et prévoir, et enfin d’obtenir des représentations graphiques permettant d’analyser les résultats obtenus. Un dictionnaire de variable semi-automatisé a également été inclue en plus d’un sous-programme retraçant les tests économétriques. Ces sous-programmes font appel à des inputs, i.e le modèle et les bases de données, et il en ressort des output, i.e. la base de donnée complétée des données en prévision ainsi que des sorties graphiques. Une majorité du code est directement inspiré de l’exemple de modèle simplifié proposé par la DG Trésor mais le programme a été entièrement restructuré, adapté et améliorer pour le projet précis de la modélisation en Polynésie française. La librairie data.table a été utilisée pour fluidifier et accélérer l’exécution du programme. L’intégralité du programme est disponible sur gitlab.

### Tests économétriques

Afin de réaliser les tests économétriques, il s’agit d’avoir obtenu en amont les observations de chacune des variables incluent dans les équations économétriques sur la période de prévision. Une méthode de test a été construite à l’occasion de ce projet à l’ISPF où les principaux enjeux statistiques et stratégique exposé ici sont détaillées (footnote : rapport sur git). Les équations économétriques du modèle Opale sont construites sur la base conceptuelle des modèles à correction d’erreur (MCE). L’objectif était ainsi d’employer ce même type d’équation pour le modèle économique Polynésien afin de se reposer sur un cadre théorique qui a fait ses preuves. En bref, ces équations permettent de dissocier la dynamique de court terme de la dynamique de long terme d’une série en explicitant à la fois les relations de long terme, qui sont en faite des corrections d’erreur (footnote : abus de langage car c’est en fait le résidu économétrique qui est corrigé) de la période passée, et à la fois les relations de court terme grâce aux taux de croissance des séries. Les MCE sont un cadre statistique naturel pour traiter des séries pour la plupart non stationnaires. Une estimation par la méthode directe des moindres carrées ordinaires de séries non stationnaires serait fallacieuse. Les MCE permettent d’introduire des décalages temporels dans les relations entre agrégats macroéconomiques transformant des liaisons instantanées en diffusions dont les durées d’amortis excèdent l’horizon d’un cadrage annuel. Les tests économétriques employés ne sont donc pas usuelles, ils consistent effectivement à tester la cointégration des séries entre elles. Si des séries sont cointégrées alors on peut les mettre en relation via une relation économétrique reposant sur un modèle à correction d’erreur. Quand les séries étaient cointégrées (footnote : on verra un peu plus tard qu’on a accepté de passer outre des hypothèses pour mener le projet), il s’agissait de trouver la spécification d’équation adéquat pour la relation considérer. La stratégie a été de tester une première spécification de type MCE pour chaque relation puis d’affiner la spécification afin d’optimiser l’explication de la variance dépendante, la significativité des paramètres, la cohérence statistique et la robustesse. Toutefois, il s’agissait de privilégier la cohérence économique de la mécanique en place en premier lieu. L’objectif est en effet de retracer l’histoire économique de la Polynésie en définissant des spécifications permettant de modéliser les comportements agrégés.

Transition :

Les 3 étapes précédentes sont en fait 3 projets totalement dépendant qui ne peuvent se développer que mutuellement. Après la mise en place d’un programme informatique permettant d’obtenir une prévision comptable avec un modèle comptable, il s’agit d’ajouter une part une les équations économétriques au modèle selon les spécifications définis à la suite des tests économétriques. Le programme peut être retesté à chaque fois et amender pour s’adapter aux spécificités des variables insérées à chaque étape. Il est en effet nécessaire d’importer les données que requiert l’équation ajoutée ou de créer des variables. Si le modèle est testé en prévision il faut également définir des hypothèses de scénario (plutôt neutre pour obtenir des résultats de référence).

## Bilan de l’outil de prévision

### Performance de l’outil

Afin de juger la performance globale des résultats obtenus il est question de définir les indicateurs qui témoignent de cette performance. La performance globale est en réalité principalement conditionnée à la performance des équations économétriques individuelles. Il suffit d’une équation économétrique dont la relation estimée est hasardeuse pour mettre à mal le modèle dans son intégralité. La phase des tests économétrique est donc cruciale. Les indicateurs de performance des équations économétriques ont déjà été discuté dans la partie précédente mais il s’agit ici de d’évaluer la cohérence globale du système et non plus la cohérence individuelle des équations. Pour ce faire, le modèle peut être résolu sur des exercices de type post mortem ou des scénarios d’hypothèses alternatif sur le passé. (footnote : Le modèle Opale produit des résultats de grande qualité sur ce type d’utilisation avec les données française). Le premier type de résolution consiste simplement à simuler le modèle sur les années observées avec les coefficients estimés. Dans ce cas les variables exogènes prennent les valeurs observées. Les variables simulées des équations économétriques ne prennent pas la même valeur que quand l’équation est simulée individuellement car les variables explicatives peuvent maintenant être le fruit d’autre simulation et non plus d’observation tel que ça l’était dans la phase de test économétrique. La simulation transfert donc l’analyse au niveau global. Le second type de résolution consiste à prévoir sur la période passée en émettant des hypothèses de scénario sur la période. Il est préférable d’être capable de retracer l’histoire économique à travers la simulation avant de tester le modèle sur des exercice de type post mortem. Le programme inclue des graphiques permettant de comparer le simulé à l’observé. Un écart constant entre les courbes témoigne d’un problème de modélisation. L’enjeu est de pouvoir retracer l’histoire économique des principales variables en niveau mais aussi en taux de croissance. Un modèle peut réaliser une simulation acceptable en niveau et pourtant très peu expliquer les taux de croissance des variables. Ceci est contraignant pour un modèle de prévision de court terme qui a pour objectif de prévoir des fluctuations.

Des allers-retours constants ont été réalisés entre amélioration, ajout de données et tests du modèle au cours du processus d’apprentissage de l’outil. Toutefois, il s’est avérer que les données ne permettaient pas pour toutes les relations économétriques essentielles de trouver des spécifications dont les séries étaient cointégrées. De plus, certaines spécifications expliquent trop peu la variance de la variable dépendante. Ces problèmes sont en partie causés par la structure des données collectées. Des variables cruciales sont manquantes, ou ont été construites artificiellement car le système de collecte de données économique en Polynésie ne permet d’obtenir les données adéquates. Les observations de la consommation des ménages et du revenu disponible brut sont par exemple imprécises. En outre, le package Tresthor requiert des données trimestrielles or les données économiques sont collectées de façon annuelle en Outre-Mer. La trimestrialisation nécessite un important retraitement des données en amont impliquant beaucoup d’imprécisions et n’apportant pas d’informations statistiques supplémentaire. Dans le cas de trimestrialisation artificiel, la régression de variables à partir de 60 observations (trimestrielles) revient en fait grossièrement à utiliser seulement 15 observations (annuels). Or, 60 observations s’avèrent déjà être la limite basse pour effectuer une régression en termes d’information statistiques.

Certaines conditions nécessaires ont donc été omise pendant ce cas d’utilisation afin de continuer l’exploration du modèle en conscience de cause. Il était en effet admis qu’un important travail d’amélioration de la collecte de donnée restait à faire en amont. La poursuite du projet est effectivement l’occasion de définir précisément quelles sont les données impératives à collecter (données économiques et indicateurs externes) et sous quelles formes (récurrence, échelle, période).

Dans la suite du projet les simulations du modèle ont permis de montrer qu’il était possible d’expliquer la dynamique des variables en niveau et les principaux chocs pour une majorité de variable. La précision de cette simulation était cependant à remettre en question. Même si un modèle ne se doit pas de retracer la réalité, il est tout de même nécessaire de pouvoir d’obtenir un simulée très proche du réalisée sinon quoi, la prévision (basée sur des hypothèses contrairement à la simulation) aura beaucoup de mal a anticipé des dynamiques économiques. De plus, le simulée en taux de variation est décorrélé du taux de variation réel. Finalement, des prévisions ont été performés mais davantage à titre de test informatique étant les problèmes soulevés.

Il faut donc retenir qu’il est important de sauter des étapes afin d’explorer l’outil de façon dynamique mais il n’en demeure pas moins que des problèmes de base de données, premier input du projet, contraignent totalement la poursuite du projet. L’exploration de l’outil est donc dépendante de la collecte des données qui peut elle-même faire l’objet d’évolution en fonction de la direction du projet. L’enjeu réel est de pouvoir investir dans la mise en place de l’outil en sachant qu’il requerra une importante mise de moyens dans la collecte de données et un délai d’attente en termes d’objectif probablement s’évaluant dans le long terme. Cependant, il ne faut pas renoncer à la poursuite d’un tel projet à cause du cout reel qu’il engendre (relativement aux bénéfices actualisées) car l’étude et la mise en place de l’outil est plutôt rapide et accessible et le développement de la collecte de données peut se faire naturellement et conjointement au développement de la statistique en Outre-mer à partir de l’impulsion initié par l’étude de l’outil. En effet, sans cette investissement initiale, le développement de la statistique dans le territoire risquerait de prendre un retard conséquence du fait de la non connaissance des conditions d’entrée des modèles économétriques.

post mortem

flex avec les contributions

revoir un peu les grandes étapes à respecter

### Travaux à venir

### Rentabilité de l’investissement et valeur ajoutée du travail

## Réplication du projet

### Recommandation

### Pistes de recherche

## Conclusion

Ce type de modèle, compact et souple d’utilisation, couplé à la mise à disposition en open source du package Tresthor, permet une utilisation accessible aux acteurs de la statistique en outre-mer. Même si les données disponible dans ce type de territoire ne permette pas d’avoir la profondeur d’analyse du modèle Opale, le package Tresthor permet en fait de traiter n’importe quel modèle du même type tant qu’il respecte les contraintes établies.

NB : Ajouter un diagramme simplifié