Etude profils dynamiques

Ce travail a pour but d’étudier la possibilité de calculer des profils dynamiques à la place de ceux statiques actuellement utilisés pour les profils de consommation dont la puissance souscrite est inférieure ou égale à 36kVA (appelés profils « bleus »).

Les profils dynamiques, calculés à partir des données mensuelles fournies par les compteurs Linky, permettent une prévision de la consommation bien plus précise que celle obtenue par les profils statiques.

La CRE a effectué et partagé1 son calcul de profils dynamiques qui sont utilisés dans sa méthode de construction du TRVE.

# A quoi correspondent les données de la CRE ?

L'onglet "profils dynamiques" de l’Excel de données de la CRE est composé des courbes de charge prévisionnelles à température normale et des gradients de température (=coefficients qui permettent de prendre en compte les variations de température). Ces grandeurs s'expriment au pas horaire et s'étalent du 1er janvier 2022 au 31 décembre 2022.

1 <https://www.cre.fr/Pages-annexes/open-data>

2 <https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.fsolve.html>

Pour chaque profil bleu, il est donné deux ensembles de paires de valeurs correspondants chacun à un lissage différent. Chaque ensemble est distingué par les appellations "1ère année de lissage" et "2ème année de lissage". A priori, ces années correspondent à deux années où la courbe de température a été lissée.

# Comment ces données sont-elles calculées ?

Partant des historiques de sous-profils dynamiques publiés sur le site d’Enedis, le modèle de la CRE procède en trois étapes :

* Etape 1 : calcul des gradients de température
* Etape 2 : retraitement des données historiques pour les ramener à température normale et estimation des paramètres du profil
* Etape 3 : calibration de modèles stochastiques permettant de générer des scénarios d’écart entre consommation réalisée à température normale et profil à température normale.

## Gradient de température

Sur une plage de température majorée par une température seuil (15°C), RTE a remarqué une relation de proportionnalité entre la variation de la température et celle de la consommation. Le coefficient exprimant cette relation est appelé "gradient de température". Comme il est indiqué en annexe de la section F des règles MA-RE, il traduit ainsi la variation de consommation liée à une variation de 1°C de température. Il faut également prendre en compte "l'inertie" des variations de consommation sur les variations de température (principalement due à l'inertie thermique des bâtiments). Cet effet est modélisé par un décalage et une atténuation de l'amplitude des variations de température appelés par RTE "lissage des températures".

Le calcul du gradient pour chaque demi-heure se base sur la méthode "par delta" utilisée par RTE et détaillée dans les règles du mécanisme de capacité. Il consiste à effectuer une régression linéaire sur une consommation en fonction d'une température lissée seuillée qui ne prennent pas en compte certains points particuliers (jours fériés, vacances scolaires...).

## Coefficients de consommation des années précédentes à température normale hebdomadaires et maj du gradient

Le calcul des coefficients de consommation à température normale hebdomadaires se base sur les coefficients dynamiques fournis par ENEDIS. L'idée est d'effectuer la moyenne des valeurs historiques auxquelles on a retranché la consommation correspondant à la différence entre la température "réalisée" (= température mesurée par ENEDIS) et la température normale seuillée. Un nouveau gradient est ensuite déterminé pour chaque semaine et chaque demi-heure en calculant le produit entre le gradient précédent et le coefficient hebdomadaire tout juste calculé.

## Courbe de charge prévisionnelle et écarts

La courbe de charge historique à température normale va maintenant être normalisée pour avoir une moyenne égale à 1 puis moyennée pour avoir un profil hebdomadaire, journalier et demi-horaire. On obtient alors les coefficients de ce profil en le multipliant aux coefficients hebdomadaires calculés précédemment. Grâce aux différents paramètres calculés, la consommation (ou courbe de charge) prévisionnelle d'abord à température normale puis à température réalisée peut-être déterminée.

La CRE propose ensuite d'évaluer l'écart entre la courbe de charge prévisionnelle et celle historique à température normale. L'étude de cet écart appelé "résidu" permet de générer des scénarios de courbes de charge pour chaque sous-profil de consommation. La CRE propose également un recalibrage de la courbe de charge en prenant en compte la relève mensuelle des données de consommations grâce aux compteurs Linky. Cela permet d'intégrer au calcul le risque porté par les fournisseurs à chaque relève.

On notera que la méthode de calcul ci-dessus renvoie des valeurs au pas demi-horaire et non horaire comme celles fournies par la CRE.

# Ces données sont-elles exploitables ?

Il est tout fait possible en théorie d’utiliser ces données pour créer un jeu de coefficient type à adapter au calendrier de l’année en prévision. Cependant, ces données contiennent les courbes de charge et non les coefficients dynamiques des profils. Il faut donc recalculer à partir de ces valeurs les coefficients (voir notebook calcul : <https://github.com/CastanieAlexandre/Analyse-coefficients-dynamiques.git> ). Pour cela, il faut résoudre un système d’équation de taille 8760x8760 et ce, 7 fois (pour chaque sous-profil). Cela revient à chercher les racines d’une fonction f détaillée dans le notebook. Il faut également préciser que la CAR (consommation annuelle de référence) n’est pas donnée. L’onglet « BDD » de l’Excel fourni par la CRE donne bien des valeurs de consommations totales mais elles ne sont pas cohérentes avec les valeurs données dans l’onglet « Profil dynamique ». Il a été décidé de prendre comme CAR la somme totale des valeurs de Courbe de Charge en attendant d’avoir plus d’informations.

Pour aider l’algorithme de résolution, je lui ai fourni la matrice jacobienne. Pour calculer cette dernière, j’ai d’abord créé un algorithme de calcul formel qui détermine l’expression littérale de la jacobienne puis remplace les inconnues par les valeurs des coefficients trouvés par le solveur. Cette méthode était trop lourde et rallonge le temps de calcul. A la place, j’ai codé un algorithme qui reconstruit une matrice directement avec les valeurs en reprenant la matrice jacobienne trouvée par le calcul formel. Finalement, cette méthode permet de calculer la matrice jacobienne de taille 8760\*8760 en environ 1s.

Il est également possible d’aider l’algorithme en lui précisant un jeu de coefficient initial qui estime la solution finale. Pour l’instant, les calculs sont faits avec un vecteur composé de 1.

Malgré cela et d’autres tentatives d’optimisation, le temps de calcul reste très long. En utilisant un algorithme hybride de Powell (fsolve 2) pour le profil PRO1 P1 et une tolérance de 10-8, le temps de calcul s’élève à 1h06min. Les résultats sont automatiquement sauvegardés dans un Excel appelé « ResultatsCalculPRO1P1.xlsx ». Voici la courbe des coefficients obtenus :

A la suite de ce résultat, on observe deux choses : D’une part l’allure de la courbe respecte celle de la courbe de charge et d’autre part, les valeurs évoluent approximativement de 11 à 24. Ceci est bien plus élevé que pour les coefficients statiques utilisés pour l’instant comme en témoigne le graphique ci-dessous :

Une possible explication est l’existence de plusieurs racines. A creuser.

* Influence de la tolérance

La tolérance influe bien entendu sur le temps de calcul (env. 40min au lieu d’1h) mais également sur les valeurs des racines trouvées. Par exemple, voici le tracé des coefficients calculés à une tolérance de 10-3 et non 10-8 :

* Influence de l’année de lissage.

Voici le tracé pour la même tolérance qu’auparavant (10-8) mais cette fois pour les données de la 2ème année de lissage (toujours pour le profil PRO1 P1).

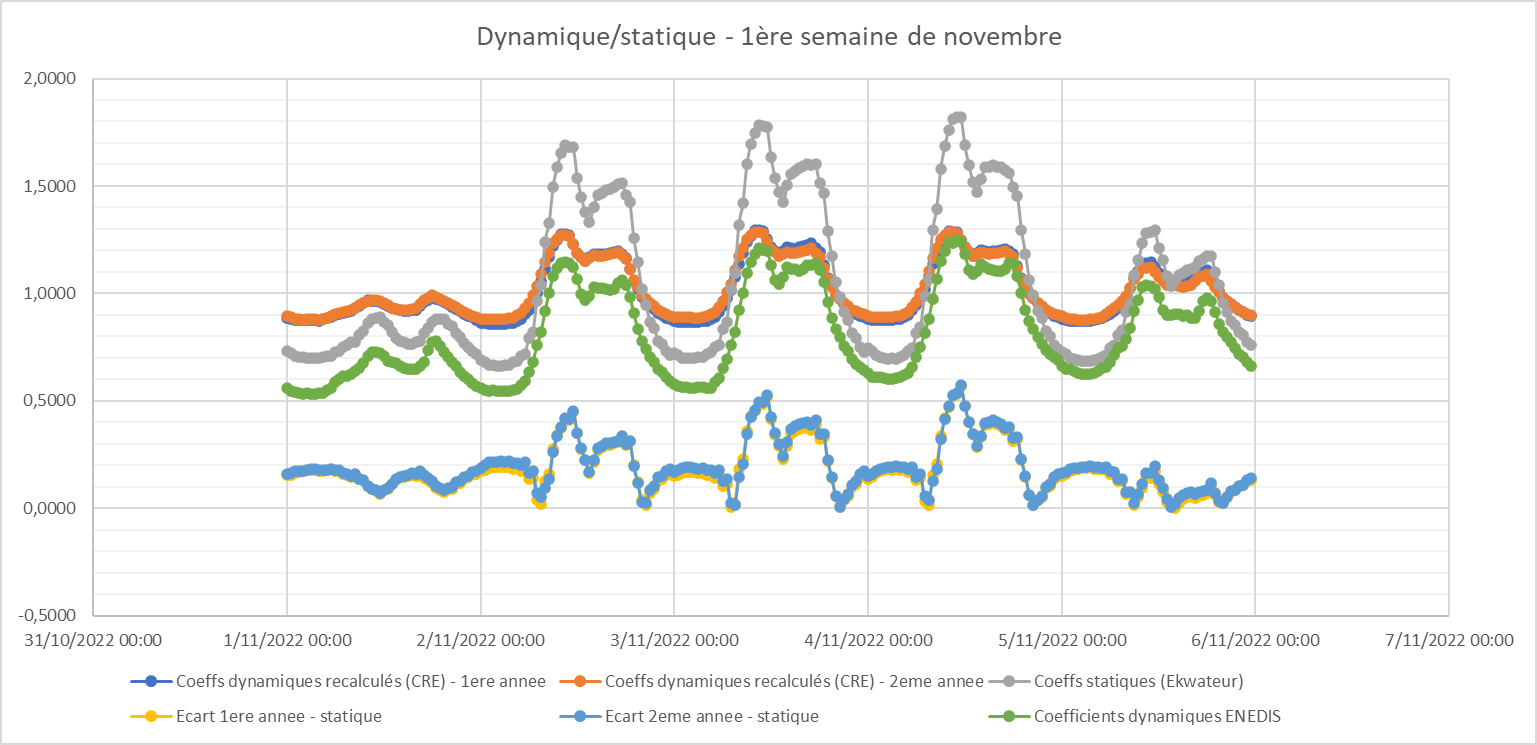
L’année de lissage possède une petite influence sur la valeur des coefficients mais l’allure ne change pas. On note que la CAR est la même d’une année à l’autre.

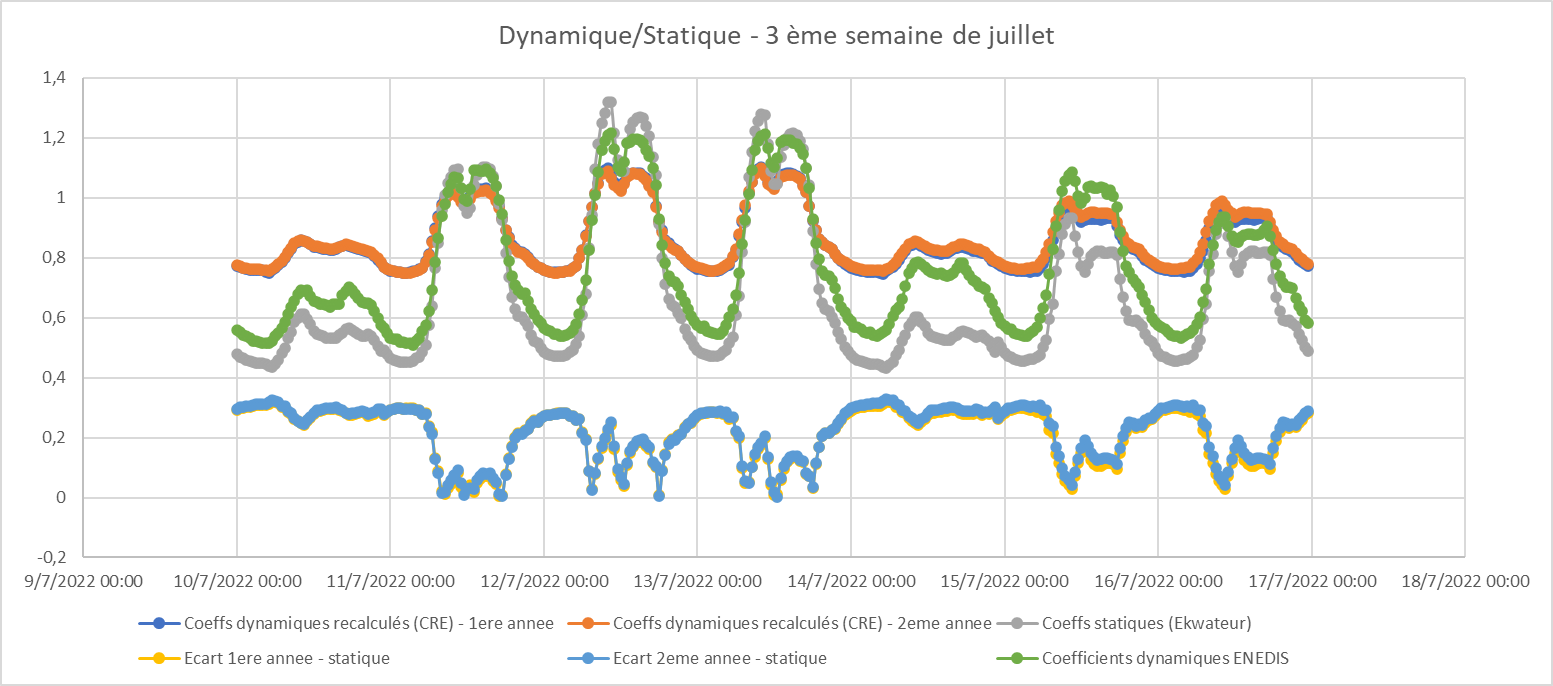
D’autres méthodes de calcul ont été explorées pour résoudre le système sans succès pour l’instant :

* Calcul via pivot de Gauss en écrivant le problème sous la forme Ax=b
* Ecriture du problème sous la forme d’un problème d’optimisation afin de profiter de la puissance et de la vitesse des algorithmes d’optimisation.

Si nous avons bien l’allure, il nous manque l’amplitude adéquate. Il a alors été pensé de se ramener à la moyenne des coefficients statiques :

Avec ce jeu de coefficient il est maintenant possible d’identifier une semaine type et déterminer un profil pour les années suivantes. Or, ce travail a déjà été fait pour les coefficients statiques. L’étape suivante est de le reprendre.

Comparaison avec les coefficients statiques et les données d’ENEDIS:



# Questions à poser à la CRE

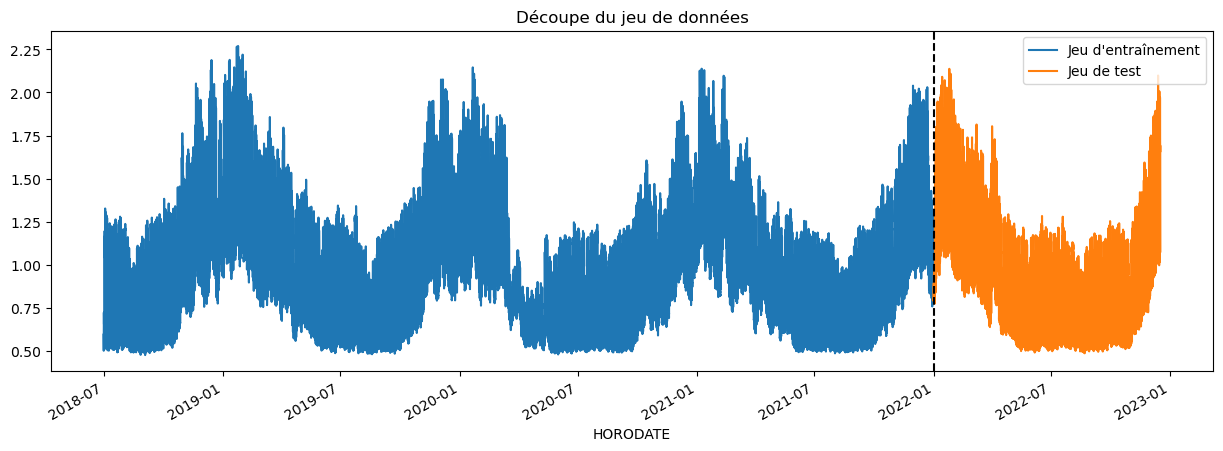
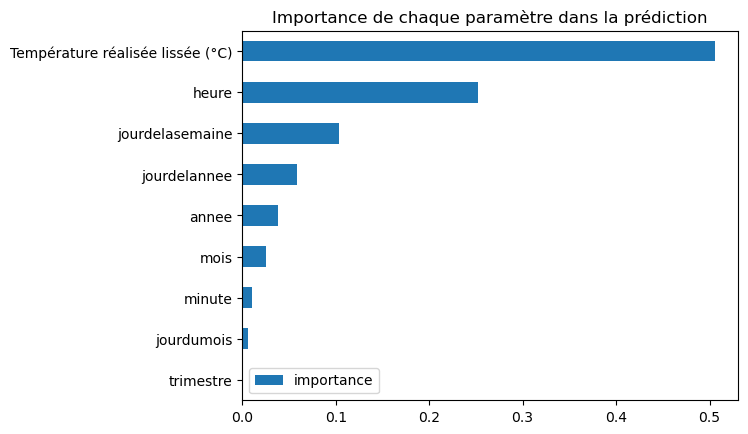
Après avoir effectué ce travail, plusieurs points d’ombre subsistent. Ils pourraient être potentiellement effacés en posant des questions à la CRE directement. Une adresse mail a été mise à disposition à cet effet : [opendata@cre.fr](mailto:opendata@cre.fr)

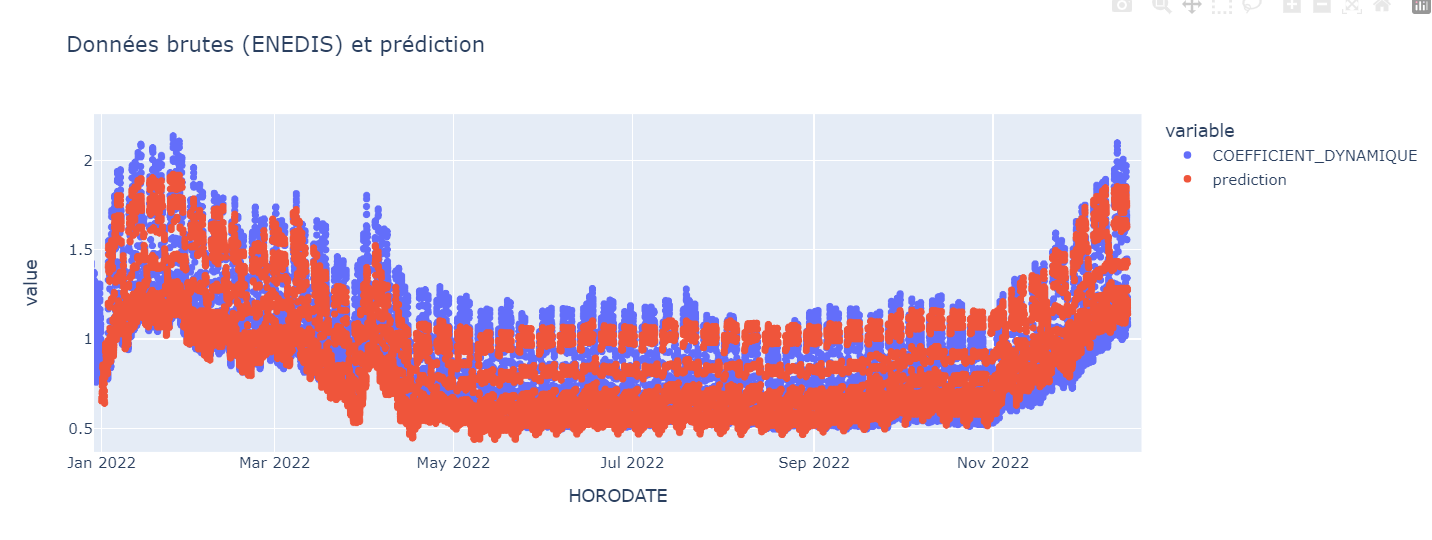
Les explications à fournir portent sur :

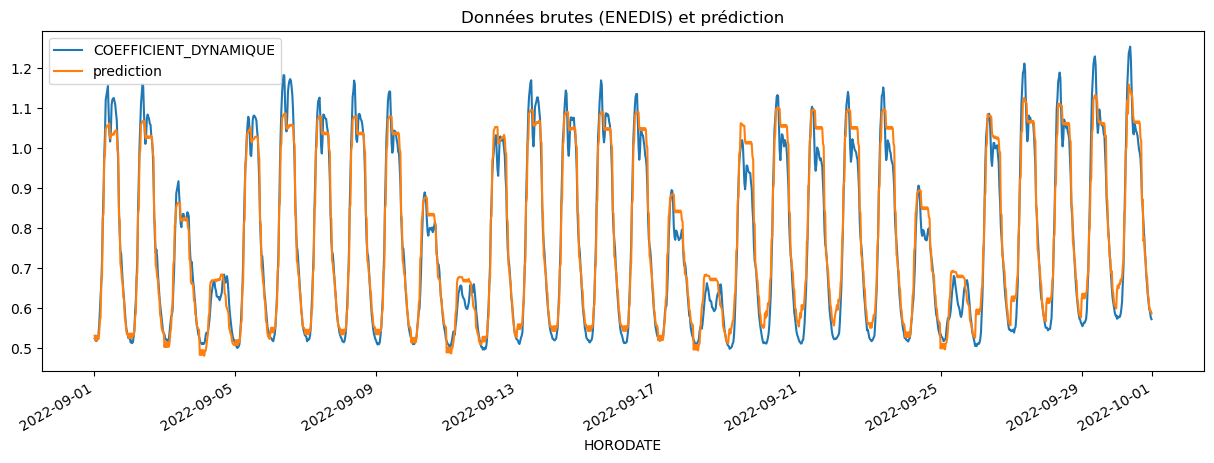
* Les 2 années de lissage
* Les données de l’onglet BDD qui ne permettent pas de retrouver la CAR
* L’absence des coefficients dynamiques
* Les coefficients dynamiques (CRE comme ENEDIS) qui sont bien plus faibles (notamment aux pointes) que les coefficients statiques.

# Piste à explorer

Les données d’Enedis sont très fournies. Il existe aujourd’hui des algorithmes de machine Learning extrêmement puissants qui seraient utiles pour prévoir une série temporelle comme celle des coefficients dynamiques. J’ai effectué un premier essai rapide sans prise en compte de la température ou des jours non-représentatifs (week-ends, jours fériés…). Voici les résultats :







Les résultats ne sont pas parfaits mais il est tout à fait possible d’améliorer le résultat en prenant en compte des paramètres supplémentaires. L’intérêt majeur de cette méthode est sa simplicité, sa rapidité et son automatisation.