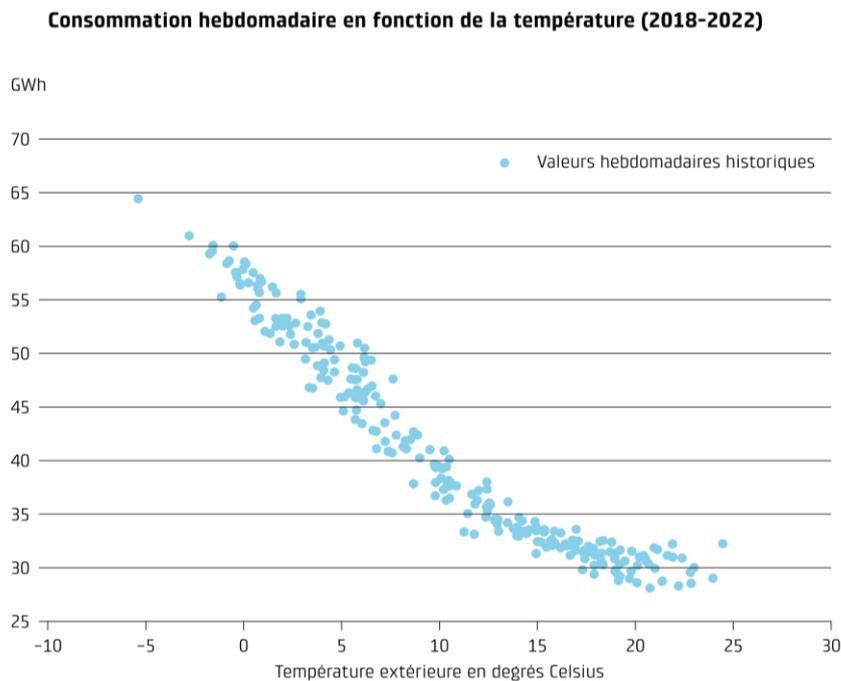


Informations générales sur la modélisation et l'analyse statistique

Les personnes intéressées trouveront ci-dessous des informations techniques plus détaillées.

Prise en compte de la température

La consommation d'électricité de l'approvisionnement de base montre une forte dépendance à la température extérieure, notamment en raison des pompes à chaleur et des chauffages électriques. La relation est représentée sur ce graphique à l'aide de valeurs hebdomadaires historiques pour la consommation et la température.



Graphique 1 : Consommation hebdomadaire en fonction de la température (2018-2022)

Il faut tenir compte de cette relation pour analyser les éventuels effets d'économie découlant des initiatives mises en place en la matière. Un temps doux réduit les besoins en chauffage, ce qui peut également se traduire par une plus faible consommation énergétique sans pour autant être associée à un changement de comportement des consommateurs et consommatrices. Le modèle décrit dans le paragraphe suivant calcule donc le volume de consommation attendu sur la base des températures de la semaine et de la corrélation entre la température et la consommation.

Évaluation de la consommation propre de courant photovoltaïque

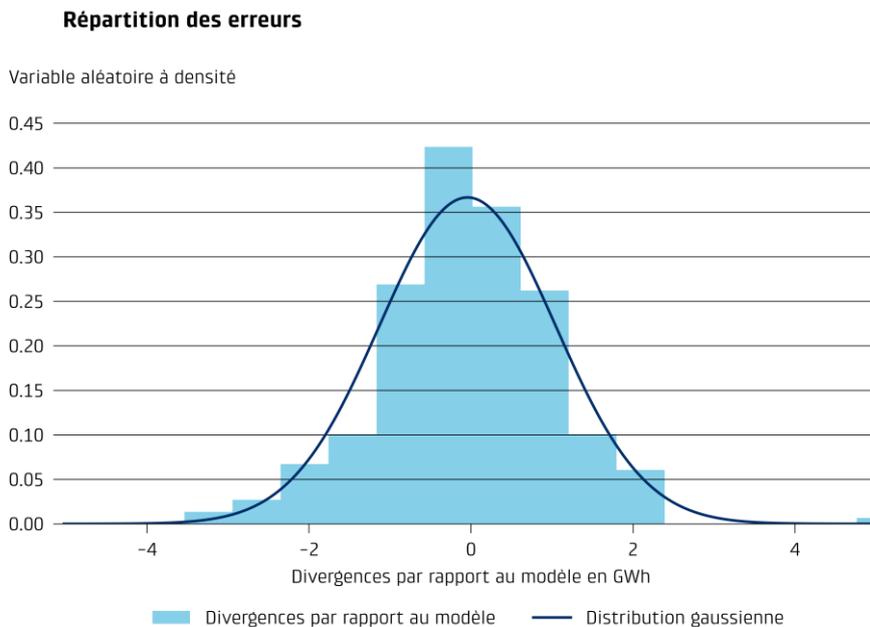
Le secteur de BKW, compte beaucoup d'installations photovoltaïques (PV) qui produisent de l'électricité, laquelle est en partie consommée directement sur place (consommation propre). Ces volumes ne peuvent cependant pas être mesurés directement et en temps réel. C'est le cas pour toutes les petites installations sans établissement de la courbe de charge. En premier lieu, BKW peut mesurer le prélèvement dans le réseau électrique. Pour évaluer la consommation complète, il faut donc simuler la production de ces installations PV à l'aide d'un modèle, puis l'ajouter au prélèvement d'électricité dans le réseau. Les valeurs de consommation affichées correspondent ainsi à la consommation complète des clientes et des clients de l'approvisionnement de base, consommation propre incluse.

Analyse statistique

Le modèle statistique est basé sur une régression linéaire multiple. Cela signifie que la consommation d'électricité hebdomadaire est modélisée comme la somme de plusieurs variables d'influence pondérées (appelées régresseurs). La pondération optimale est essentielle pour la qualité du modèle et peut être calculée à l'aide de données historiques et de différents algorithmes (la « méthode des moindres carrés » étant la plus utilisée).

Le choix des régresseurs est tout aussi important pour la pertinence du modèle : un facteur comme la température extérieure peut être traduit dans différents régresseurs. Par exemple, la température moyenne hebdomadaire, la température maximale et minimale ou bien les changements de température peuvent être utilisés comme régresseurs dans le modèle. Nous avons également désigné la construction et la sélection des régresseurs par le terme de « feature engineering ».

La qualité d'un modèle de régression linéaire doit impérativement être vérifiée à l'aide de plusieurs facteurs. La distribution statistique de l'erreur de modèle est un critère important lorsque le modèle est appliqué à des données historiques. Le choix correct des régresseurs et leur pondération optimale sont caractérisées par le fait que cette erreur suit grossièrement une distribution gaussienne (en forme de cloche).



Graphique 2 : Répartition des erreurs

Un autre critère de qualité réside dans l'indépendance statistique de l'erreur de modèle. Cette condition n'est pas automatiquement remplie dans le modèle de consommation présenté. Une consommation supérieure à celle attendue par le modèle au cours d'une semaine donnée augmente la probabilité que la consommation de la semaine suivante soit également supérieure à celle prévue par le modèle. (Causes possibles : effets météorologiques, ou facteurs non modélisés qui s'étendent sur plusieurs semaines.) L'effet pour la période 2018–2022 a été quantifié et corrigé par une estimation dite de Cochrane-Orcutt.

L'un des grands avantages des modèles de régression linéaires par rapport aux méthodes d'apprentissage automatique plus modernes est la facilité de calcul des incertitudes sous forme d'intervalles de confiance (appelés « fourchette de la consommation attendue » jusqu'ici dans le texte). Les intervalles de confiance reflètent non seulement les fluctuations naturelles qui ne sont pas expliquées par les régresseurs modélisés, mais aussi

l'incertitude du modèle. Cela signifie que la pondération optimale des facteurs à l'aide de données historiques comporte elle-même une incertitude. Le résultat de cette modélisation est par conséquent une fourchette de consommation hebdomadaire dans laquelle on s'attend à trouver différentes valeurs hebdomadaires dans 95% des cas si aucun effort d'économie n'a été entrepris.

De plus, des modèles de régression linéaires permettent réaliser de multiples analyses. Par exemple, dans le présent modèle de consommation, les valeurs hebdomadaires et leurs incertitudes peuvent être agrégées en valeurs mensuelles ou trimestrielles. Observer des périodes plus longues permet de réduire les incertitudes, car les fluctuations statistiques sont pondérées dans le temps (Cela n'est cependant valable que sous certaines conditions d'indépendance statistique.) Donc, même si différentes valeurs ne divergent que de manière non significative de celles prévues sur le modèle, une agrégation peut mettre en évidence un écart significatif sur une période plus longue. C'est le cas des données cumulées depuis septembre. Les incertitudes indiquées se réfèrent à une probabilité de 68% que la sous-consommation réelle au cours de cette période se situe dans la plage d'incertitude.

Pourquoi un modèle de régression aide-t-il à identifier une réduction de la consommation énergétique ? De nouvelles valeurs de consommation hebdomadaires étudiées qui divergent significativement de celles attendues par le modèle pendant une période prolongée indiquent que le modèle de régression n'est plus pertinent. Étant donné que le modèle est exclusivement basé sur des données historiques allant de 2018 à août 2022, il en résulte que depuis septembre 2022, un nouveau facteur influence considérablement la consommation énergétique hebdomadaire. Il paraît évident que les appels aux économies d'énergie en sont la cause, mais ce type d'analyse ne permet pas de prouver un lien de causalité.