Prévision des prix spot de l'électricité sur le marché allemand

Éclairage sur les méthodes de prévision par une approche statistique

Réel enjeu technique et financier, la prévision des prix de l'électricité est une information capitale pour les choix d'investissements liés à l'énergie. Il existe une multitude d'outils de prédiction plus ou moins adaptés au caractère des données à prédire. Présentation de quelques-uns d'entre eux.

Stéphane Genoud et al.

La prévision des prix de l'électricité est une information décisionnelle pour les choix d'investissements liés à l'énergie, que ce soit à court terme (jour pour le lendemain), à moyen terme (1 à 3 mois) ou à long terme (3 ans). Dans un précédent article [1], nous avons fourni un aperçu financier des principales caractéristiques du prix de l'électricité. Nous vous présentons ici un éclairage sur les méthodes de prévision à court terme et à moyen terme en se basant sur une approche purement statistique basée sur des modèles linéaires et les prix spot du marché allemand en 2013.

Pourquoi prévoir le prix de l'électricité ?

Le marché spot permet d'acheter l'électricité la veille pour une livraison le lendemain à un prix fixé par heure. 25 % [2] de l'électricité en Europe sont échangés sur les marchés de gros entre différents acteurs, qu'ils soient producteurs, fournisseurs ou industriels. L'estimation des prix pour les trois prochaines années permet à un industriel de sécuriser ses contrats d'approvisionnement à long terme. À l'échelle du mois, un producteur pourra optimiser ses prévisions d'achat ou de vente. La veille pour le lendemain, un gestionnaire de réseau pourra réaliser un arbitrage entre l'achat sur le marché et l'utilisation des réserves afin d'équilibrer le réseau. Nous utilisons les prix spot cotés sur le marché EEX allemand de l'année 2013 car il s'agit du second marché le plus liquide (1) et des données complètes les plus récentes.

Principe de la prévision par une approche statistique

Notre approche se base sur l'observation des données passées afin de prédire les valeurs futures. Le principe général consiste à trouver une fonction qui caractérise au mieux les données observées et qui, par extrapolation, permettra de déterminer le devenir de la série. À titre d'exemple, l'évolution de la population mondiale dans le temps est une fonction croissante qui, graphiquement, peut être extrapolée par une fonction affine alors que l'évolution de la tension en un point du réseau serait mieux modélisée par une fonction sinusoïdale. Une simple analyse visuelle du tracé des données permet ainsi d'orienter sur le ou les outils susceptibles de fournir les meilleures prévisions. Cela suppose en effet qu'il existe un lien entre les valeurs qui, une fois identifié, nous renseigne sur l'évolution de la série : dans notre cas, le prix spot de demain est lié aux prix des heures précédentes et des jours précédents. Il existe une multitude d'outils de prédiction plus ou moins adaptés au caractère des données à prédire. L'observation graphique des données nous renseigne sur un certain nombre de caractéristiques importantes :

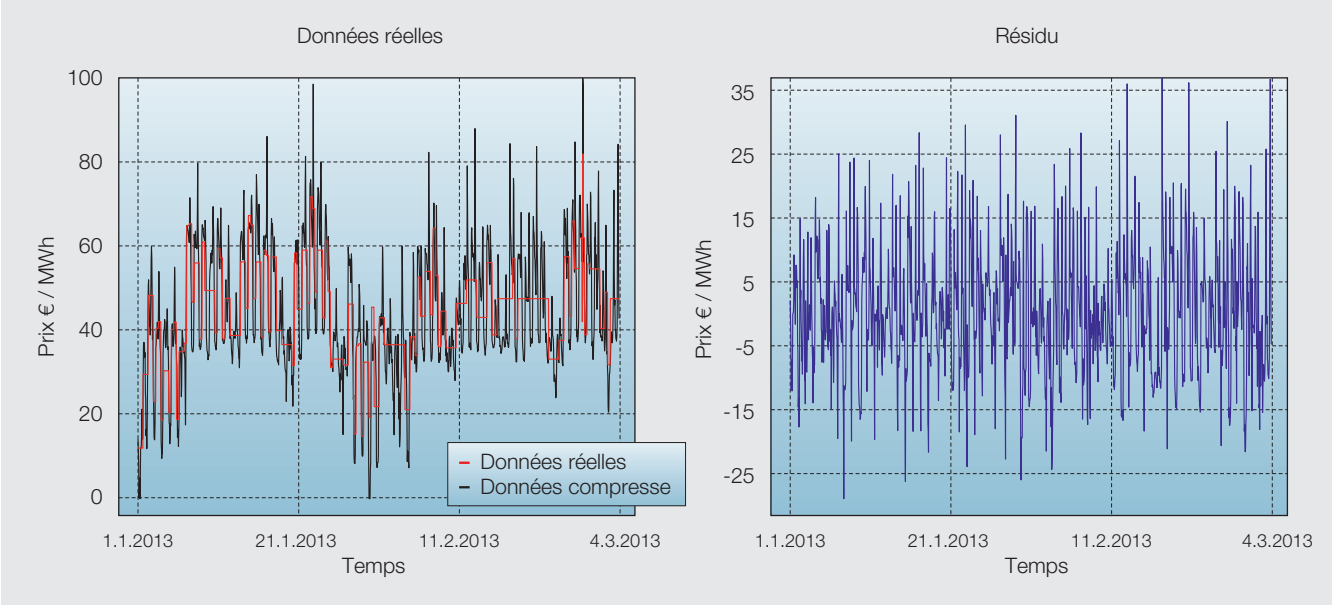

Figure 1 Visualisation des données compressées du 01/01 au 04/03/2013.

la présence d'outliers, la saisonnalité, la variance, la tendance et la stationnarité (2) de la série. La présence ou non de ces éléments permet d'identifier dans le bouquet des outils statistiques ceux qui seront adaptés à la modélisation des données. Afin d'appliquer les modèles de prédiction, les données sont partagées en un échantillon de calibrage qui sert à déterminer les paramètres du modèle et un échantillon test qui permet d'estimer la pertinence du modèle.

Facteurs influant sur le prix de l'électricité

Comme expliqué dans l'article précédent [1], le prix est le croisement entre l'offre et la demande, soit le croisement entre les coûts marginaux de production et les prix auxquels les acteurs sont prêts à acheter l'électricité. Ils seront donc dépendants de plusieurs facteurs tels que les besoins en électricité, le prix des combustibles et la météo. La météo est un facteur clé car elle influence à la fois l'offre via la production des énergies renouvelables et la demande via les besoins en chauffage ou en climatisation. On peut ainsi comprendre l'autocorrélation des prix, étant donné que d'une heure à l'autre ou d'un jour à l'autre, les conditions climatiques sont assez proches et les prévisions meilleures, sauf un changement drastique de la météo qui conduit à des valeurs extrêmes correspondant aux outliers.

Extraction des outliers

Le coté stochastique et ponctuel des outliers les rend non seulement imprévisibles avec les outils classiques, mais ils constituent également des valeurs pouvant altérer l'analyse statistique de l'ensemble des données. Dans un souci d'appliquer les outils adaptés aux phénomènes adéquats, nous traiterons les outliers dans un prochain article. Les outliers sont extraits de la série en utilisant la décomposition en ondelettes [3]. Sur la **figure 1**, les prix réels sont représentés en noir et le signal compressé en rouge. Celui-ci est constitué de valeurs qu'on ne peut prédire telles que les outliers et représente 5 % des données. La série résiduelle en bleu est constituée des prix réels auxquels on a extrait le signal compressé. On obtient alors une série stationnaire qui représente 95 % de l'information initiale. Les modèles de prévision que nous décrirons dans la suite de l'article seront appliqués à la série résiduelle.

Composante saisonnière

L'effet saisonnier se traduit par la répétition d'une valeur à une certaine fréquence. Dans notre cas, celle-ci peut être de 24 h, 7 jours ou de l'ordre des saisons météorologiques.

La **figure 2** illustre la distribution des prix par heure, par semaine (lundi à dimanche) et par mois pour l'année 2013. La courbe bleue est la moyenne des valeurs et les courbes rouges définissent un intervalle contenant 95 % des données. Selon le premier graphique de la figure, à 1 h du matin, les prix sont compris entre 28 et 30 €/MWh alors qu'à 18 h, on achète le MWh entre 43 et 48 €. Sur le deuxième graphique, on constate que le motif horaire se répète avec une amplitude plus ou moins élevée en fonction du jour. Les prix sont plus élevés la semaine

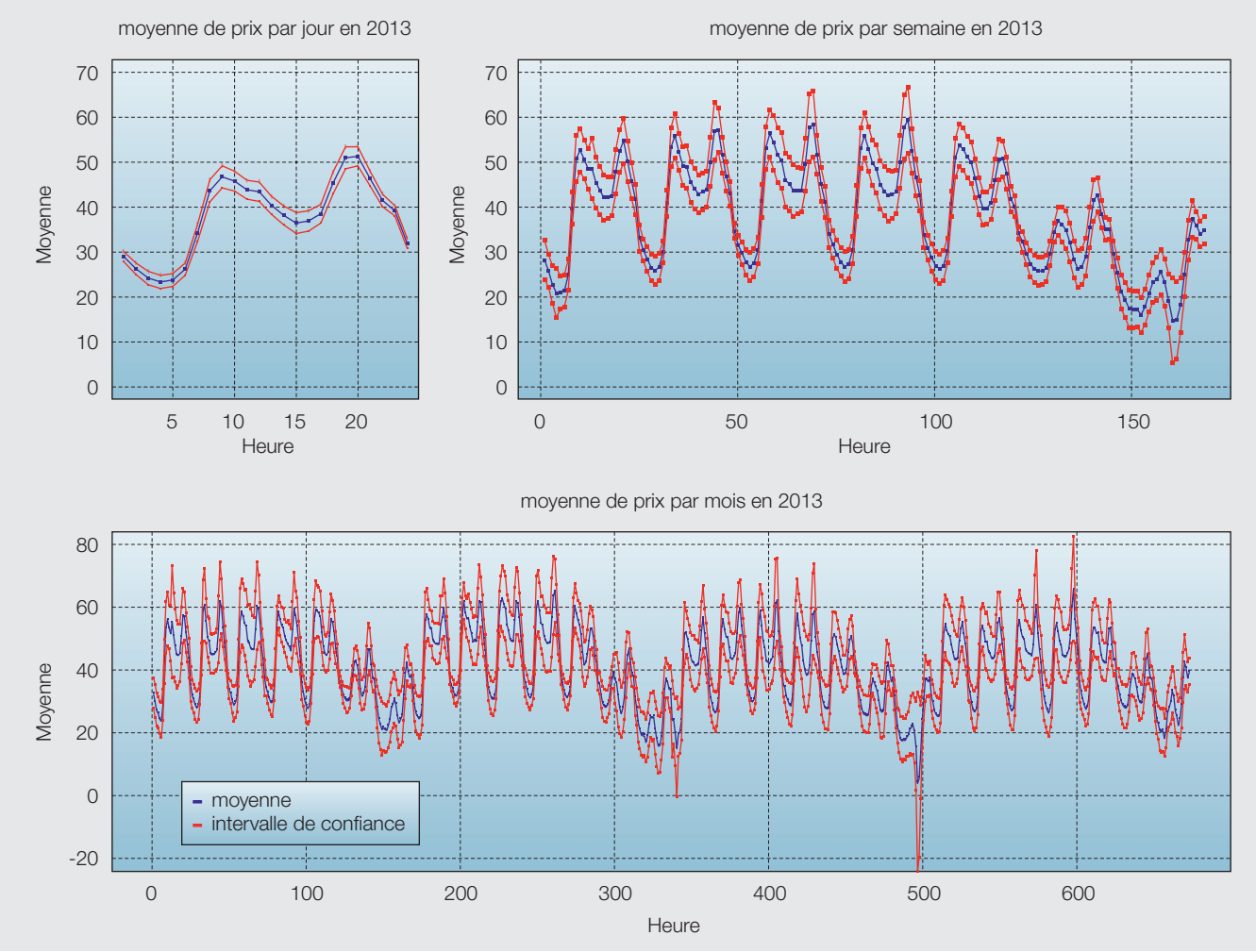

Figure 2 Aperçu de la saisonnalité par heure et par semaine à partir de la distribution des prix spot en 2013.

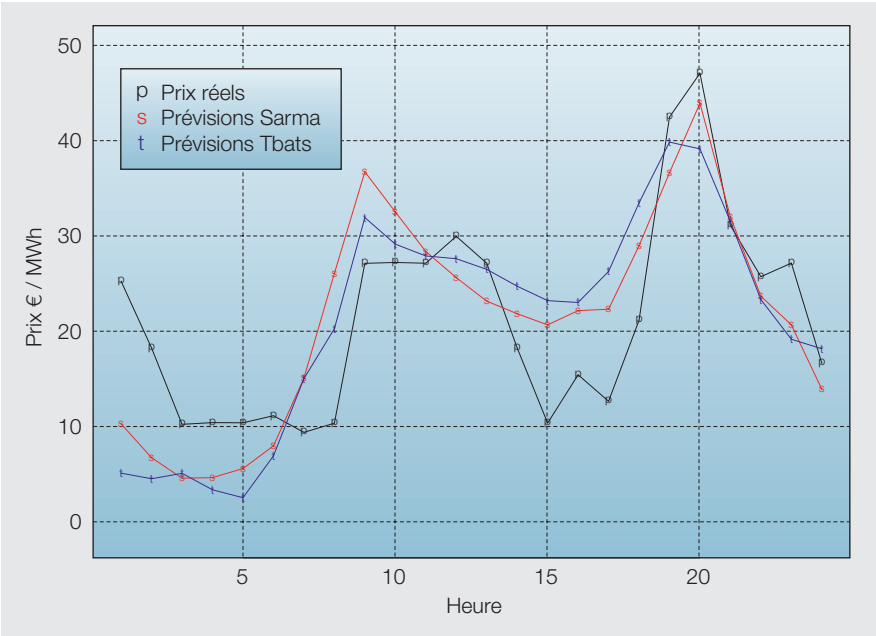

Figure 3 Résultat de la prévision des prix spot du 23/04 par les processus TBATS et SARMA.

que le week-end, ce qui retranscrit les fluctuations de l'activité économique. De même, sur le troisième graphique, on visualise une nette répétition du motif hebdomadaire. Ce sont en effet l'activité économique, nos habitudes de consommation et les cycles météorologiques qui vont rythmer les effets de saisonnalité observés sur les prix de l'électricité. On peut donc extraire une composante saisonnière telle que la série soit égale à une fonction périodique plus ou moins un écart. L'évolution de la fonction périodique étant connue, le cœur de la prédiction est l'estimation de l'écart obtenu après extraction de la saisonnalité.

Méthode de prévision à court terme

SARMA

On utilisera les modèles linéaires de type ARMA et de type TBATS. Les processus ARMAp,q soit AR (Auto-Régressif) d'ordre p et MA (Moyenne Mobile) d'ordre q sont adaptés aux séries stationnaires et à variance constante, ce qui est une bonne approximation de notre série résiduelle. En supposant que la série des valeurs est centrée (moyenne proche de 0), la valeur à prédire est estimée par la formule suivante [4] :

$$X_{t+1} = \underbrace{\sum_{i=1}^{p} a_i X_{t-i}}_{AR_q} + \underbrace{\sum_{j=1}^{q} b_j E_{t-j}}_{MA_q}$$

La valeur future, notée X_(t+1), est exprimée comme la somme d'une combinaison linéaire des p valeurs précédentes et d'une combinaison linéaire des q valeurs d'un bruit blanc (3) de même taille que notre échantillon. Les coefficients a_i et b_j sont constants et estimés à partir de l'échantillon de calibrage. Ces processus étant adaptés aux séries sans tendance ni saisonnalité, il existe une extension SARMAp,q applicable aux séries avec saisonnalité qui élimine la composante saisonnière, puis applique un ARMAp,q aux valeurs résiduelles.

TBATS

Le modèle TBATS (Trigonométrique-Box Cox-Arma-Trend-Saisonnality) permet d'estimer X_(t+1) au voisinage de T par la formule suivante [5] :

$$X_{t+1} = \underbrace{A_{t-1} + B_{t-1}(t-T)}_{Linéarité} + \underbrace{S_{m1} + \cdots + S_{mn}}_{Saisonnalité} + \underbrace{d_t}_{arma}$$

L'approche est différente de celle des processus ARMA, l'idée étant de définir une droite au voisinage de l'instant T qui nous donnera les valeurs de la série à l'instant t+h. Les paramètres de la fonction affine sont déterminés en minimisant l'écart entre les valeurs prises par la fonction et celles de l'échantillon de calibrage. Ces paramètres sont représentés par les termes A et B dans la formule. Les termes S_m1…S_mn de la formule correspondent aux composantes saisonnières de période m_1..m_n estimées par une décomposition en séries de Fourier. Dans notre cas, l'observation des prix sur une année présente deux composantes saisonnières : m_1=24h et m_(2)=7j. Le terme d_t est la prédiction par un processus ARMA de l'écart entre le prix réel et les autres termes de la formule [6].

Méthode de prévision à moyen terme

On s'intéresse ici aux prix par heure des 30 prochains jours. Parce qu'on garde une précision de prédiction par heure et qu'on étend à la fois l'horizon de prévision à un mois, la prédiction se fait heure par heure en considérant par exemple que le prix à 7 h du lendemain est corrélé aux prix observés à 7 h les jours précédents. La série résiduelle est de ce fait éclatée en 24 séries contenant chacune les prix observés à une heure donnée. On va donc utiliser des processus légèrement différents afin d'intégrer l'influence croisée des horaires : les modèles nested-ARMA et VAR (AR multivarié) [6]. Le processus nested-SARMAp,q est un SARMAp,q appliqué à la série des prix à 7 h auquel on applique les prix à 6 h comme régresseur externe. L'utilisation des régresseurs externes est très fréquente en statistique afin de prendre en compte l'évolution d'un phénomène externe sur celui que l'on veut prédire. Afin de prédire le débit d'eau d'une rivière, on peut par exemple utiliser les données de précipitation de la région. Le VARp est un modèle ARp multivarié, c'est-à-dire qu'on va utiliser d'autres variables ayant des effets croisés afin de réaliser la prévision. En finance par exemple, la prévision de l'inflation prend en compte des facteurs tels que les taux d'intérêt ou la valeur des actifs cotés en Bourse en tant que modèles multivariés. Dans notre cas, le prix à 7 h est prédit en utilisant les prix à 6 h et à 8 h du jour précédant.

Résultats de la prévision à court terme

À court terme, notre objectif est d'estimer la veille les prix de l'électricité des 24 heures du lendemain. La **figure 3** représente la prévision des prix du 23/03/2013 en appliquant un SARMA1,1 et un TBATS à la série résiduelle des prix du 01/01/2013 au 23/03/2013. Afin de pouvoir confronter les résultats aux prix réels, le signal compressé est rajouté au résidu et aux prévisions. Rappelons que le signal compressé n'a pas été prédit.

Les deux modèles suivent généralement les variations des prix réels. Les prévisions ont donc une bonne compréhension de l'effet saisonnier. Le test de Spearman, compris entre -1 et 1, est un indicateur de la corrélation entre les prévisions et le prix réel. Lorsqu'il est positif et proche de 1, l'évolution des prévisions et des données sont en phase. On obtient 0.26 pour le TBATS et 0.43 pour le

Figures: HES-SO Valais Wallis

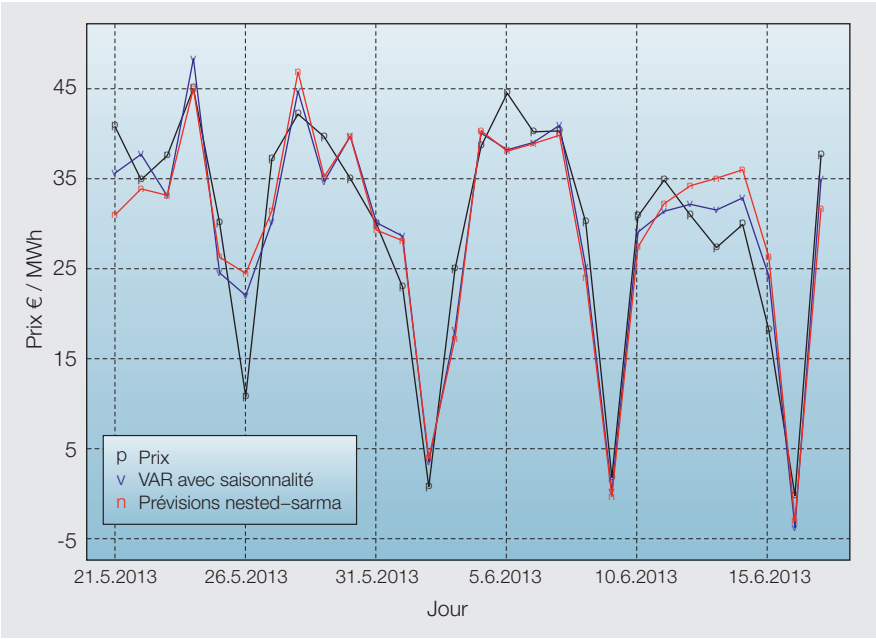

Figure 4 Résultat de la prévision du prix spot à 7 h du 21/05 au 17/06 par les processus VAR1 et nested-SARMA1.

SARMA. Cet indicateur nous confirme que les modèles suivent mieux les tendances de fond que les petites variations. On constate également que la précision des résultats est différente en fonction des heures et du modèle. À 11 h, les prévisions et le prix réel sont quasiment égaux alors qu'à 15 h, on a un écart de 5 € entre le prix réel et le modèle TBATS et de 8 € entre le prix réel et le SARMA. Le TBATS fournit donc une meilleure prévision à 15 h, pourtant on observe qu'à 20 h, c'est le SARMA qui approche au mieux le prix réel. La moyenne des écarts (4) sur les 24 h est de 7 €/MWh pour le SARMA et de 8 €/MWh pour le TBATS.

Résultats de la prévision à moyen terme

L'enjeu à moyen terme consiste à estimer le prix par heure pour les 30 prochains jours. Pour notre exemple, le prix à 7 h du 21/05 au 17/06/2013 est prédit à partir des données du 01/01 au 20/05/2013. Rappelons qu'on utilise un nested-SARMA1,1 avec le prix à 6 h comme régresseur externe et un VAR1 sur les prix à 6 h, 7 h et 8 h. De même que précédemment, le signal compressé est rajouté au résidu et aux prévisions afin d'obtenir les résultats apparaissant à la figure 4.

De même que précédemment, la fluctuation des prévisions coïncide généralement avec celle des prix réels. Le test de Spearman est de 0,89 pour le SARMA et de 0,72 pour le TBATS. Les principaux pics des prix réels sont nettement décrits par les prévisions mais sans atteindre les valeurs extrêmes. Les moments où l'évolution des prix réels et prédits sont en contradiction vont générer davantage de perte. Entre le 11/06 et le 14/06 par exemple, le VAR décrit bien une baisse du prix alors que le nested-SARMA prévoit une augmentation, ce qui conduit à une erreur de 9 €/MWh pour le SARMA le 13/06. Le 26/06, c'est le VAR qui s'écarte le plus des variations du prix spot avec un écart de 12 €/MWh. Visuellement, bien qu'il n'y ait pas d'écart trivial entre les deux modèles, on peut tout de même noter que le processus VAR approche les pics des prix spot de façon plus précise. Sur l'ensemble de la période étudiée, la moyenne des écarts est de 4 €/MWh pour le VAR et de 5 €/MWh pour le nested-SARMA.

Conclusion

La prévision des prix de l'électricité est un réel enjeu technique et financier. Nous vous avons présenté une approche de prévision utilisant des modèles linéaires car ils sont plus simples à décrire. Cependant il existe des multitudes approches telles que les fonctions polynomiales ou les réseaux de neurones. Rappelons que la méthode que nous vous avons présentée consiste à extraire les outliers pour laquelle d'autres outils de prévision sont plus adaptés, à analyser la partie déterministe représentée par la saisonnalité et à prédire la partie restante. On s'aperçoit, à partir des résultats, qu'il n'y a pas de modèle idéal, c'est sans doute l'association de plusieurs approches et encore une fois l'œil et l'expérience de l'analyste qui

Résumé

Prognose der Spotpreise für Strom auf dem deutschen Markt

Ein statistischer Ansatz beleuchtet die Prognosemethoden

Die Prognose der Strompreise stellt in technischer und finanzieller Hinsicht eine echte Herausforderung dar. Ein Artikel in der Dezemberausgabe des Bulletins [1] gab einen finanziellen Überblick über die wichtigsten Merkmale des Strompreises, der aktuelle Artikel beleuchtet nun die Methoden für kurz- und mittelfristige Prognosen. Diese Methoden stützen sich auf einen rein statistischen Ansatz, dem wiederum lineare Modelle und die Spotpreise des deutschen Marktes des Jahres 2013 zugrundeliegen.

Vergangene Daten werden beobachtet, um künftige Werte vorherzusagen. So steht der Spotpreis von morgen im Zusammenhang mit den Preisen der vorhergehenden Stunden und Tagen. Es gibt eine Vielzahl von mehr oder weniger auf die jeweiligen vorherzusagenden Daten abgestimmten Prognoseinstrumenten.

Wie im Artikel vom Dezember bereits erläutert, hängen die Preise von verschiedenen Faktoren ab, u. a. vom Strombedarf, vom Brennstoffpreis und vom Wetter. Das Wetter ist ein Schlüsselfaktor, da es sowohl das Angebot über die Produktion erneuerbarer Energien als auch die Nachfrage über den Heiz- oder Kühlbedarf beeinflusst. So kann man nachvollziehen, dass es eine Autokorrelation der Preise gibt. Denn die klimatischen Bedingungen verändern sich von einer Stunde zur nächsten oder von einem Tag zum andern kaum. Dies führt zu besseren Prognosen – ausser bei einem drastischen Wetterumschwung, der Extremwerte und damit Ausreisser bewirkt. Diese Phänomene können mit den klassischen Instrumenten nicht vorhergesagt werden.

Die in diesem Artikel vorgestellte Methode schliesst die Ausreisser aus, für die es andere, geeignetere Prognoseinstrumente gibt, analysiert den deterministischen Teil, der auf die Saisonalität zurückgeht, und sagt den restlichen Teil voraus. Die weitere Arbeit besteht darin, die Ausreisser zu prognostizieren. Der nächste Artikel wird einen Ansatz zur Prognose der Ausreisser vorstellen, der auf Wetterschwankungen beruht sowie auf der räumlichen Verteilung der erneuerbaren Energien in Deutschland. Cr

permettront de décider des meilleures prévisions. La suite du travail consiste à prédire les outliers. Ces phénomènes sont les conséquences de changements parfois imprévisibles, comme par exemple la météo qui influence la production électrique via les énergies renouvelables. Dans le prochain article, nous vous présenterons une approche de prévision des outliers basée sur les fluctuations météorologiques et la répartition territoriale des énergies renouvelables en Allemagne.

Notes

- (1) La liquidité d'un marché financier se traduit par des volumes et des fréquences d'échange importants, ce qui limite l'effet des transactions sur la variation des prix.
- (2) Une série stationnaire a une moyenne constante dans le temps.
- (3) Un bruit blanc est un ensemble de valeurs aléatoires de moyenne nulle et de variance 1.
- (4) La moyenne est obtenue par le RMSE (Root Main Square Error), soit la racine carré de la moyenne du carré des écarts entre le prix réel et la prévision.

Lien

- www.hevs.ch

Références

[1] S. Genoud, F.M. Cimmino, D. Previdoli, Le marché financier de l'électricité en Allemagne, Bulletin SEV/AES, 2014.
[2] European Commission. 2010a. Questions&Answers : Preventing Abuse in Wholesale Energy Markets – Europa.eu : Memo/10/655.
[3] F. Jose, G. Gomez, Wavelet methods for time series analysis, 2013.
[4] J. Jacques, Introduction aux séries temporelles, Polytech Lille, 2013.
[5] A.M. De Livera, R.J. Hyndman, R.D. Snyder, Forecasting time series with complex seasonal patterns using exponential smoothing, 2010, http://www.buseco.monash.edu.au/depts/ebs/pubs/wpapers/
[6] C.H. Aladay, E. Grioglu, Advances in time series, 2010.

Auteurs

Dr **Stéphane Genoud** est professeur en Management de l'énergie à l'Institut Entrepreneurship & Management de la HES-SO Valais Wallis. Il est titulaire d'un CFC d'électricien, un diplôme d'ingénieur HES, une licence en économie, plusieurs masters (finance et énergie) et une Thèse de Doctorat en économie sur l'analyse, d'un point de vue du développement durable, des modes de production de l'électricité. En parallèle à ses activités académiques, il a créé plusieurs sociétés où il travaille aussi comme consultant senior, membre du conseil.

HES-SO Valais Wallis, 3960 Sierre, stephane.genoud@hevs.ch

Francesco Maria Cimmino est actuellement stagiaire à l'Institut Entrepreueurship & Management de la HES-SO Valais Wallis. Il a obtenu un master de deuxième niveau en Statistique et Économie alla Sapienza de Rome (2012) et un master en économétrie pour la banque et la finance à l'Aix-Marseille School of Economics (2014).

HES-SO Valais Wallis, 3960 Sierre, francesco.cimmino@hevs.ch

Mariam Barque est actuellement assistante de recherche à l'Institut d'Informatique et de Gestion de la HES-SO Valais Wallis. Ingénieur généraliste de l'école des mines d'Albi et diplômée du mastère spécialisé Optimisation des Systèmes Energétiques des Mines de Paris en 2011, elle a travaillé pour EDF R&D dans le cadre de son Mastère puis comme consultante à CGI Business Consulting.

HES-SO Valais Wallis, 3960 Sierre, mariam.barque@hevs.ch

Dr **Dominique Genoud** est professeur à l'institut d'informatique et de gestion de la HES-SO Valais Wallis. Il est diplomé de l'EPFL et l'auteur d'une thèse sur la reconnaissance vocale et les systèmes biométriques de l'IDIAP. Ses activités de recherches couvrent plusieurs domaines dont la gestion de l'énergie et l'analyse de données appliquées aux biotechnologies et à l'industrie pharmaceutique. En parallèle à ses activités académiques, il est CEO de sa propre société fondée en 2006.

HES-SO Valais Wallis, 3960 Sierre, dominique.genoud@hevs.ch