Revue de statistique appliquée

E. DOURILLE

Un exemple d'élaboration de modèles à coefficients techniques

Revue de statistique appliquée, tome 11, n° 1 (1963), p. 87-93
<http://www.numdam.org/item?id=RSA_1963__11_1_87_0>

UN EXEMPLE D'ÉLABORATION DE MODÈLES A COEFFICIENTS TECHNIQUES

E. DOURILLE
Directeur Adjoint du Centre d'Études Régionales sur l'Économie de l'Énergie

INTRODUCTION

Les modèles économétriques mesurent une grandeur en fonction de variables explicatives qui sont, a priori, les principaux facteurs déterminant cette grandeur ; la relation est explicitée par une formule algébrique entre ces variables à l'aide de coefficeints et de paramètres dont les valeurs sont tirées de l'observation.

Ces modèles sont généralement destinés à être opératoires : on se donne, a priori, les valeurs des variables explicatives en un futur couvert par le modèle et on en déduit ce que sera la variable expliquée, ce qui permet de prendre une décision en conséquence.

L'étude des séries statistiques, chroniques ou coupes instantanées, permet de tester les modèles. En choisissant convenablement les coefficients et paramètres respectifs, il y a généralement un grand nombre de modèles concevables s'ajustant aux données statistiques disponibles avec une approximation jugée suffisante.

Comment effectuer un choix parmi ces modèles ?

Quand on ne dispose que de séries statistiques donnant les valeurs des variables explicatives et expliquées à certaines dates (chronique) ou dans certains secteurs à une date donnée (coupe instantanée), il est commode de s'en tenir à la formulation mathématique la plus simple parmi celles qui donnent pratiquement satisfaction.

Quand on dispose de plus d'informations déterminant l'avenir de la grandeur étudiée (par exemple des réalisations techniques futures qui peuvent être prévues avec une probabilité convenable), il peut être préférable d'adopter un modèle tenant compte de ces informations, au prix d'une formulation moins simple et d'hypothèses complémentaires limitant le champ d'application de ce modèle. Il s'agira alors d'élargir au maximum ce champ d'application en améliorant l'information, en effectuant une analyse plus fine de ce phénomène, en mesurant autant que possible les nouveaux facteurs mis en évidence, même au pris d'une formulation plus complexe entre cette grandeur que l'on veut apprécier et ses variables explicatives.

L'étude des perspectives de consommations d'énergie dans l'Industrie fournit un vivant exemple de l'élaboration de tels modèles. Les "modèles à coefficients techniques" sont effectivement utilisés au Centre d'Etudes Régionales sur l'Economie de l'Energie et il nous a paru intéressant de les présenter en suivant le processus de raisonnement itératif qui, de l'hypothèse à l'observation et vice versa, a permis de bâtir ces modèles successifs, qui trouvent leur utilisation en fonction des informations disponibles dans chaque cas particulier.

HYPOTHESES PREMIERES

- La consommation d'énergie C d'un secteur industriel (ou d'un usage industriel dans ce secteur) n'est fonction que de deux variables : la production I de ce secteur, et le temps t.

- L'élasticité a de C par rapport à I est constante.

- Hors l'effet de la variation de I, le taux annuel de variation de C est constant.

Ces hypothèses conduisent à soumettre à l'observation des modèles des types suivants :

$$\frac{C}{C_o} = \left(\frac{I}{I_o}\right)^a . \; b^t \text{ (Ailleret)}^{(1)} \text{ ou } \frac{C}{C_o} = \left(\frac{I}{I_o}\right)^a . \; e^{-r \cdot t} \text{ (Ventura)}^{(2)}$$

où a, b et r sont des constantes, et e la base des logarithmes népériens.

Comme l'a montré G. Morlat [3], il est pratiquement préférable d'effectuer les calculs d'ajustement sur les séries chronologiques disponibles, au moyen des accroissements annuels ΔC et ΔI par le modèle $\frac{\Delta C}{C} = \frac{a \Delta I}{I} - r$, formule dérivée de la différentielle des modèles précédents : $\frac{dC}{C} = a \frac{dI}{I} - rdt$. Il est bon cependant de souligner que dans ce cas on admet l'hypothèse complémentaire suivante :

- l'évolution du phénomène étudié est assez lente pour que les accroissements annuels ΔC et ΔI puissent être assimilés aux dC et dI de l'équation différentielle, avec un dt égal à l'unité pour une année.

OBSERVATIONS

P. Ailleret, puis G. Morlat ont trouvé d'excellentes corrélations en ajustant leurs modèles aux séries chronologiques de consommations d'électricité de différents secteurs industriels.

Dès 1956, E. Ventura ajusta son modèle sur la série des consommations totales annuelles d'énergie dans l'Industrie française : pour diverses raisons dont la principale était l'insuffisance des données d'observation tant en nombre qu'en qualité (les statistiques de base des charbons sont des statistiques de livraisons et non de consommations), la corrélation fut nettement moins bonne. En admettant une distribution normale des valeurs r et a, données par l'étude de cette série chronologique, autour de leurs valeurs moyennes, il a même observé que les valeurs admissibles de r comprenaient 0 dans l'intervalle de confiance à un seul écart-type, et que par conséquent il était plus simple de prendre le modèle à élasticité constante

----

(1) "Estimation des besoins énergétiques" par P. Ailleret. Revue Française de l'Energie de Septembre 1955.

(2) "Prévisions de consommation d'énergie de l'Industrie Française" par E. Ventura. Annales des Mines de Février 1957.

(3) Communication de G. Morlat au Congrès international à la Statistique (Paris 1961).

$\frac{\Delta C}{C_0} = \left(\frac{\Delta I}{I_0}\right)^a$, bien entendu a étant judicieusement choisi en fonction de r = 0.

Que le coefficient de corrélation soit fort ou faible, on n'échappe pas à une certaine indétermination des couples de valeurs a et r à choisir à l'intérieur de l'intervalle de confiance que l'on s'est donné.

Or, G. Morlat a justement fait remarquer que, pour la prévision à long terme, on aboutit à des résultats fort différents selon le couple de valeur a et r adoptés ; aussi souhaitait-il que l'on puisse donner à l'un de ces coefficients une signification technique permettant de le calculer, ou tout au moins de l'estimer, par une toute autre méthode.

HYPOTHESES

Gardons un modèle de type $\frac{\Delta C}{C} = a \frac{\Delta I}{I} - r.\ t$, mais convenons que les accroissements concernent tout l'intervalle de temps (0, n) de sorte que :

$$\Delta C\ (\%) = \alpha .\ \Delta I\ (\%) - \beta .\ n$$

Peut-on donner une signification "technique" à ces coefficients α et β, c'est-à-dire utiliser des informations sur l'évolution technique des équipements énergétiques ?

Si ρ_n et ρ_0 mesurent les consommations unitaires moyennes générales de tous les équipements du secteur industriel étudié au cours des années n et 0, on montre facilement que :

$$\alpha = \frac{\rho_n}{\rho_0} \quad \text{et} \quad \beta = \frac{\rho_n - \rho_0}{\rho_0 \times n} \times 100$$

ce qui veut dire que α mesure le rapport des consommations unitaires aux temps n et 0 des équipements énergétiques étudiés, et que β est la valeur moyenne annuelle de la variation de consommation unitaire entre les temps n et 0. Les coefficients "techniques" α et β ainsi définis dépendent non seulement du temps (ρ_n est certainement fonction du progrès technique donc du temps) mais aussi de la production I (plus il y aura d'équipements nouveaux, plus ρ_n doit différer de ρ_0).

Le modèle simple qui vient d'être décrit est le premier "modèle à coefficients techniques" (appelé modèle A) du CEREN : il est très utile car ce que l'on observe le plus facilement ce sont les variations globales de la consommation unitaire ρ sur des périodes passées, et les techniciens donnent volontiers des estimations de l'évolution future de ce paramètre ρ en fonction de cette évolution passée. Ce faisant, ils admettent implicitement une hypothèse d'évolution de la production : ce sera généralement une extrapolation de l'évolution constatée au cours des quelques années les plus récentes.

Mais une telle hypothèse implicite n'est pas admissible quand précisément on se donne a priori une évolution nettement différente. D'où la nécessité de rechercher un modèle dans lequel le couple des coefficients techniques α et β est pratiquement indépendant de I.

Voici comment il a été opéré.

On a divisé l'ensemble des équipements énergétiques, tel qu'il se présentera au cours de l'année n de la prévision, en trois groupes :

- les équipements "anciens" en service au cours de l'année 0 et restant en vie en année n de consommation unitaire moyenne ρ'_n en année n avec une production totale I'_n,

- les équipements de "renouvellement" qui ont remplacé les équipements déclassés des années 0 à n pour assurer la même production initiale I_0 (production I''_n et consommation unitaire moyenne ρ''_n au cours de l'année n),

- les équipements "de développement" qui ont assuré la variation de production ΔI, avec une consommation unitaire moyenne ρ'''_n au cours de l'année n.

En admettant que $\rho''_n = \rho'''_n$ (on suppose donc que les équipements de renouvellement ne se distinguent pas techniquement des équipements de développement), un calcul simple analogue au précédent donne laors à α et β les significations techniques suivantes :

$$\alpha = \frac{\rho''_n}{\rho_o} \quad \beta = \frac{\rho_o - (1 - v)\ \rho'_n - v\ \rho''_n}{n.\rho_o} \text{ où } v = \frac{I''_n}{I_o}$$

En première approximation ces ceofficients techniques ne dépendent plus de la production I.

A regarder de plus près, une production industrielle en surdéveloppement induira des effets sur la durée de vie des équipements (donc sur ρ'_n) et sur le progrès technique des équipements neufs (donc sur ρ''_n et ρ'''_n) ; nous supposerons par la suite que ces effets induits sont négligeables compte tenu de la précision actuelle de nos estimations.

Ce modèle à coefficients techniques (modèle B du CEREN) n'est pratiquement pas utilisé car ρ'_n, ρ''_n ou ρ'''_n ne sont pas des données usuelles d'observation d'où la nécessité de recourir à de nouvelles hypothèses afin d'exprimer les coefficients techniques α et β, définis précédemment, à l'aire de paramètres quantifiables.

Appelons :

- p la variation de consommation unitaire annuelle des équipements neufs,

- w la variation de consommation unitaire annuelle des équipements après leur mise en service,

- D la durée de vie des équipements.

Les hypothèses suivantes :

- p, w, D sont des constantes,

- le remplacement des équipements, ainsi que des installations en développement, se fait par tranches égales de productions annuelles m aussi bien avant qu'après l'année 0, permettent d'exprimer α et β en fonction de ces constantes p, w, D.

En effet, en utilisant les notations précédentes, et en appelant z_o la consommation unitaire d'énergie des équipements neufs étudiés au temps 0, les consommations unitaires <u>moyennes</u> ρ_o (au temps 0) et ρ'_n, ρ''_n et ρ'''_n (au temps n) s'expriment comme suit :

$$\rho_o = \frac{\sum_{i=-D}^{i=o} m\, z_o\,(1 + ip - iw)}{I_o}$$

$$\text{ou } \rho_o = \frac{D \times m \times z_o}{I_o}\left(1 + \frac{D+1}{2}\,(w - p\,\right)$$

par application de l'égalité $1 + 2 + 3 + \ldots + n = \frac{n\,(n+1)}{2}$ où $I_o = D \times m$ par hypothèse

$$\text{donc } \rho_o = \left(1 + \frac{D+1}{2}\,(w - p)\right) z_o$$

Des calculs analogues donnent :

$$\rho_n' = \frac{\sum_{i=-D'}^{i=n} m\, z_o\,(1 + ip - iw + nw)}{I_o'} = \left(1 + \frac{(D - n + 1)}{2}\,(w - p) + nw\right) z_o$$

$$\rho_n'' = \rho_n''' = \frac{\sum_{a=o}^{i=n} m\, z_o\,(1 + ip + (n - i)\,w)}{I_n''\ (\text{ou } I_n''')} = \left(1 + \frac{(n+1)}{2}(w + p)\right) z_o$$

En reportant ces valeurs de ρ_o, ρ_n', ρ_n'', et ρ_n''' dans les formules précédemment trouvées pour α et β, on obtient le modèle à coefficients techniques suivant (modèle C du CEREN) : en posant toujours $\Delta C\,(\%) = \alpha.\ \Delta I\,(\%) - \beta.\ n$:

$$\alpha = \frac{1 + \frac{(n+1)}{2}\,(p + w)}{1 + \frac{(D+1)}{2}\,(w - p)}\ ; \quad \beta = \frac{100\, p\left(1 + \frac{1}{D}\right)}{1 + \frac{(D+1)}{2}\,(w - p)}$$

OBSERVATIONS

Il existe des données statistiques prises dans l'information technique permettant d'évaluer p, w, D. Admettre que p, w, D sont des constantes veut dire que l'on ajustera les séries chronologiques de consommations unitaires sur des droites, et que tous les équipements ont eu et auront la même durée de vie.

Une telle simplification peut être suffisante dans certains cas, mais le plus souvent on constate que p, w et D sont des fonctions du temps et leurs valeurs ont trop varié entre les années (- D) et (+ n) pour que l'on puisse admettre l'hypothèse de leur constance.

Il faut donc retourner aux hypothèses pour concevoir un modèle plus proche de la réalité.

Cependant, ce modèle rend de grands services. En premier lieu il permet de prévoir ΔC (ΔI étant donné) en fonction de paramètres nouveaux : progrès de la technique de production et d'utilisation de l'énergie (p), évolution des équipements après leur mise en service (w), durée de vie de ces équipements (D) ; il ouvre une voie nouvelle aux informations techniques à rechercher.

En second lieu il permet de calculer p, w, et D sur une période passée pendant laquelle on connaît les variations de productions ΔI et de consommations d'énergie ΔC. Pratiquement, on évalue p à l'aide de la durée de vie moyenne D observée et d'une hypothèse sur l'évolution de w en fonction de p.

Mais le champ d'application de ce modèle aurait été trop restreint si nous n'avions pas conçu un modèle nous libérant, au moins partiellement, de l'hypothèse de constance des valeurs de p et w : ce fut l'objet du modèle D du CEREN.

Il n'était pas réaliste non plus de garder l'hypothèse de constance de la durée de vie des équipements : c'est pourquoi nous avons conçu le modèle E que nous décrivons directement ci-dessous.

HYPOTHESES

Supposons, non plus que p, w, D sont des constantes, mais

- que p prend trois valeurs p_1, p_2, p_3 respectivement pour les périodes $(-D_1, 0)$, $[-(D-n), 0]$, $(0, n)$,
- que w prend également trois valeurs w_1, w_2, w_3 pour ces mêmes périodes,
- que D prend deux valeurs : D_1 avant l'année 0 et D_2 après cette année 0, dans ces conditions nous avons calculé que les coefficients techniques α et β s'exprimaient ainsi en fonction de ces nouveaux paramètres :

$$\alpha = \frac{1 + \frac{(n+1)}{2}(w_3 + p_3)}{1 + \frac{(D+1)}{2}(w_1 - p_1)}$$

$$\beta = 100\left[\frac{1 + \frac{(D_1+1)}{2}(w_1 - p_1) - \left(1 - \frac{n}{D_3}\right)\left(1 + \frac{(D_1 - n + 1)}{2}(w_2 - p_2) + nw_2\right) - \frac{n}{D_3}\left(1 + \frac{(n+1)}{2}(w_3 + p_3)\right)}{n\left(1 + \frac{(D_1+1)}{2}(w_1 - p_1)\right)}\right]$$

OBSERVATIONS

On peut observer les valeurs de p, w, D évoluant avec le temps, et leurs ajustements sur des segments de droite donnent pratiquement satisfaction dans l'état actuel de nos connaissances ; ainsi le CEREN a-t-il pu, pour une dizaine de branches industrielles, parfois par usage et par grande Région, établir un tableau des valeurs calculées ou estimées de ces paramètres permettant de calculer, à l'aide des modèles à coefficeints techniques que nous avons présentés ci-dessus, des coefficients techniques α et β qui rendent dès maintenant grand service.

On conçoit aisément qu'il faudra encore beaucoup de patientes recherches afin de recueillir suffisamment de données statistiques issues des recherches techniques pour que le CEREN puisse considérer que ses prévisions à l'aide de ces coefficients techniques ont atteint une précision suffisante.

C'est dire que nous attendrons maintenant les prochains progrès de l'observation.

CONCLUSION

Ce modèle E, qui est effectivement employé pour tous les ensembles d'équipement les mieux connus, donne satisfaction dans l'état actuel de nos connaissances.

Un grand progrès sera accompli dans le domaine de la prévision de la consommation d'énergie dans l'Industrie lorsque tous les usages de chaque branche industrielle pourront être traités par un tels modèle.

Alors, pour les secteurs les mieux connus, on pourra reprendre les définitions des consommations unitaires moyennes en supposant que tous les paramètres, m, z, w, sont des fonctions du temps.

$$\rho_o = \int_{-D}^{o} \frac{m\ (t)\ (z\ (t) + \int_t^o w\ (u)\ du)\ dt}{I_o}$$

$$\rho'_n = \int_o^n \frac{m''\ (t)\ (z\ (t) + \int_t^n w'\ (u)\ du)\ dt}{I'_n}$$

$$\rho''_n = \int_o^n \frac{m''\ (t)\ (z\ (t) + \int_t^n w''\ (u)\ du)\ dt}{I''}$$

et une formule analogue pour ρ'''_n.

Des variations de ces paramètres en fonction du temps sur les périodes observées, on pourra induire des hypothèses d'évolution sur le futur ayant une meilleur probabilité de coïncidence avec la réalité.

Ainsi des hypothèses encore moins restrictives, et des modèles élaborés en conséquence, permettront de mieux analyser l'importance des principaux facteurs influençant les consommations industrielles d'énergie ; ils auront un champ d'application plus étendu, leurs résultats prévisionnels auront une meilleure précision a priori, dans tous les cas, la prévision aura gagné en rationalité.

Comme on le voit, le travail d'itération entre les hypothèses et les observations n'est pas prêt de se terminer dans ce domaine, comme dans bien d'autres.