

EUGÈNE KOUASSI

MICHEL TERRAZA

**Modélisation par le filtre de Kalman de la chronique mensuelle du prix du gas-oil en France de 1960 à 1992**

*Journal de la société statistique de Paris*, tome 135, n° 4 (1994), p. 25-46

[http://www.numdam.org/item?id=JSFS\\_1994\\_\\_135\\_4\\_25\\_0](http://www.numdam.org/item?id=JSFS_1994__135_4_25_0)

© Société de statistique de Paris, 1994, tous droits réservés.

L'accès aux archives de la revue « Journal de la société statistique de Paris » (<http://publications-sfds.math.cnrs.fr/index.php/J-SFdS>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme  
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

## ARTICLES

**MODÉLISATION PAR LE FILTRE DE KALMAN  
DE LA CHRONIQUE MENSUELLE DU PRIX  
DU GAS-OIL EN FRANCE DE 1960 A 1992**

Eugène KOUASSI

*Université d'Abidjan*

et

Michel TERRAZA

*Université de Montpellier I***Résumé**

Une expression exacte de la vraisemblance du Modèle Espace d'État de la série stationnaire en différence fractionnaire du prix HT du gas-oil en France est évaluée sous les hypothèses (i) de distribution préalable diffuse, (ii) de vecteur d'état initial fixé et (iii) d'état d'équilibre stationnaire. La Stabilité Asymptotique (Lyapunov) du Modèle Espace d'État contribue à l'efficacité du Filtre de Kalman dans l'analyse, la prévision et la mise à jour de la série du prix HT du gasoil.

**Mots-clés :** Filtre de Kalman – Maximum de Vraisemblance – Modèle Espace d'État – Série Temporelle – Conditions Initiales – Stabilité Asymptotique au sens de Lyapunov.

**Abstract**

An exact expression for the likelihood of gas-oil prices State Space Model in France is evaluated under the assumptions of (i) diffuse prior, (ii) fixed initial state vector and (iii) steady state Kalman Filter concerning the structure of initial conditions. Lyapunov Asymptotic Stability of the State Space Model contributes to the efficiency of the Kalman filter in the analysis, prediction and updating of HT time series of gas-oil prices in France.

**Key words :** Kalman Filter – Maximum Likelihood – State Space Model Time series – Initial Conditions – Lyapunov Asymptotic Stability.

## 1. Introduction

La plupart des décisions qui concernent les importations, les exportations, la production ou plus généralement le marché de l'énergie et qui ont trait aux variables clefs de ce secteur d'activité relèvent de la connaissance du futur et s'effectuent dans une situation d'incertitude totale ou partielle.

Pourtant, de telles décisions sont essentielles dans l'élaboration des budgets nationaux, des plans de développement à moyen-long terme ainsi que dans l'établissement des stratégies et/ou des scénarios de prix prospectifs, etc.

Dans ce papier, nous utilisons le Filtre de Kalman (e. g. Kalman R. E. (1960), Kalman R. E. and Bucy R. S. (1961) etc.) pour analyser et prévoir le prix Hors Taxes (HT) du gasoil en France qui est l'une des principales variables du marché énergétique de ce pays. Plus précisément, nous réalisons :

(i) L'identification récursive du prix HT conditionné par les prix Toutes Taxes Comprises (TTC) observés du gas-oil à l'aide d'un algorithme récursif. En effet, les prix HT semblent être gouvernés par un mécanisme incomplet et imprédictible.

(ii) L'estimation du modèle Espace d'État du prix HT du gasoil à l'aide d'un algorithme récursif puis sa validation à l'aide de tests sur les résidus et du critère de stabilité au sens de Lyapunov.

(iii) La prévision et la mise à jour du prix HT du gasoil à partir d'un prédicteur évolutif.

Pour atteindre ces différents objectifs, notre méthodologie fait appel à l'approche de Kalman R. E. (1960) et plus généralement aux méthodes récursives (e. g. A. C. Harvey and S. Peters (1990), C. N. Ng and Young P. C. (1990), Goodwin G. C. and Sin K. S. (1984), Anderson B. O. and Moore J. B. (1979), Ljung L. and T. Söderström (1983), etc.).

Le modèle Espace d'État du prix du gas-oil est utilisé ici pour obtenir les estimations du maximum de vraisemblance du prix HT. Notons que cette approche plus générale inclut la régression traditionnelle, les modèles ARIMA, l'approche Bayésienne (e. g. Harrison P. J. and Stevens C. F. (1976)) et les modèles à coefficients dépendant du temps comme cas particuliers. Un ensemble de deux équations décrivent le processus étudié. La première (équation du système) décrit le changement dans les prix HT du gasoil à travers le temps en utilisant une matrice de transition et les prix sur le marché de Rotterdam dans la période précédente comme input ou variable exogène. La seconde (équation de mesure) décrit la relation entre les prix HT et TTC du gasoil. La combinaison de ces deux équations donne un modèle Espace d'État qui est utilisé pour étudier le prix HT du gasoil.

La structure conditionnelle de la fonction de vraisemblance conduit à analyser l'incidence de conditions initiales plus complexes que celles habituellement étudiées.

L'article est structuré de la façon suivante :

Un modèle Espace d'État est développé afin d'analyser l'évolution du prix du gas-oil (Section 2). L'estimation de ce modèle est alors envisagée en examinant la représentation conditionnelle des moyennes et variances de l'équation d'état et de celle de la mesure.

Sous l'hypothèse de normalité, cette représentation conditionnelle est utilisée pour construire la vraisemblance définie comme le produit des probabilités conditionnelles du processus. Les paramètres à estimer sont spécifiés en utilisant différentes conditions initiales (e. g. Ansley C. F. and Kohn R. (1985), Bell W. and Hillmer S. (1991), Marshall T. (1992), Shumway R. H., Olsen D. E. and Levy L. J. (1981), Smith J. Q. (1990) etc.) pour l'estimation (Section 3).

La section 4 concerne l'application de la procédure d'estimation à la série du prix du gas-oil en France de 1960 à 1992. Plus précisément, on aborde l'estimation des équations du Filtre de Kalman, l'analyse de la stabilité asymptotique au sens de Lyapunov du modèle Espace d'État, la validation des résultats, la prédiction et la mise à jour de la série du prix du gas-oil. La dernière section a trait à la discussion des résultats et à l'analyse de leur portée.

## 2. Le modèle Espace d'État

Le modèle Espace d'État du prix HT du gas-oil repose sur la propriété markovienne des processus. Ce modèle est défini par les équations suivantes :

$$P_t^{\text{HT}} = FP_{t-1}^{\text{HT}} + G \text{Prot}_{t-1} + \Gamma \omega_t \quad (2.1)$$

$$P_t^{\text{TTC}} = HP_t^{\text{HT}} + v_t, \quad t = 0, 1, 2, \dots, T \quad (2.2)$$

où :  $P_t^{\text{HT}}$  est un  $p \times 1$  vecteur d'état des prix HT non-observés du gasoil,  $P_t^{\text{TTC}}$  est un  $q \times 1$  vecteur d'outputs observés des prix TTC du gas-oil, et  $\text{Prot}_{t-1}$  est un  $r \times 1$  vecteur d'inputs observés ou exogènes représentant les prix spot sur le marché de Rotterdam. Les matrices  $F$ ,  $G$ ,  $H$  et  $\Gamma$  représentent les coefficients du modèle. Ces matrices sont supposées de dimensions compatibles avec les opérations matricielles requises dans (2.1) et (2.2). Les deux termes  $\omega_t$  et  $v_t$  représentent des processus aléatoires de moyenne nulle, indépendants et identiquement distribués avec :

$$E\{\omega_t, \omega_t^T\} = \Sigma^\omega, \quad E\{v_t, v_t^T\} = \Sigma^v, \quad \text{et} \quad E\{\omega_t, v_t^T\} = 0 \quad (2.3)$$

où :  $\Sigma^\omega$  est une  $p \times p$  matrice symétrique définie positive ;  $\Sigma^v$  est une  $q \times q$  matrice symétrique positive.  $\Sigma^v$  est en général supposée définie positive (e. g. A. C. Harvey and R. G. Pierce (1984) ; Jones R. H. (1980) ; D. S. Stoffer and K. D. Wall (1991) ; Caines P. E. (1988) ou Anderson B. O. and Moore J. B. (1979)). L'état initial  $P_0^{\text{HT}}$  est supposé être normalement distribué avec pour moyenne  $P_0^{\text{HT}}$  et covariance  $\Sigma^{P_0^{\text{HT}}}$ .

### 3. Procédure d'estimation

Plusieurs méthodes d'estimation sont discutées dans la littérature (e. g. Harrison P. J. and Stevens C. F. (1976) and Mehra R. K. (1979) etc.). Nous considérons ici la technique du maximum de vraisemblance (e. g. Schweppe F. C. (1965)). Dans le but d'estimer les paramètres, la fonction de vraisemblance des variables d'observation et d'état est construite. Lorsque celle-ci est obtenue, elle est alors maximisée sur l'Espace des paramètres pour un ensemble donné d'observations afin d'obtenir les estimations du maximum de vraisemblance des différents paramètres. Par la suite, lorsque ces estimations sont obtenues, les équations (2.1) et (2.2) sont utilisées pour prévoir les valeurs futures des variables d'état et d'observation. Celles-ci sont plus tard mises à jour avec l'addition de nouvelles observations du prix HT du gas-oil prises dans le temps.

#### 3.1. Estimation conditionnelle

Intéressons nous à présent à l'estimation du vecteur non-observable du prix HT du gas-oil. Évidemment nous supposons que les coefficients du modèle et la structure de corrélation peuvent être uniquement paramétrés par un  $(k \times 1)$  vecteur  $\vartheta$  ; c'est-à-dire :

$$F = F(\vartheta), G = G(\vartheta), H = H(\vartheta), \Gamma = \Gamma(\vartheta), \Sigma \omega_i = \Sigma \omega_i(\vartheta), \Sigma v_i = \Sigma v_i(\vartheta)$$

Le vecteur  $\vartheta$  est supposé être un élément d'un espace compact  $P$ , habituellement un sous-ensemble de  $\mathbb{R}^K$ . La paramétrisation est telle que le modèle est complètement identifié (e. g. Wall (1987), Pagan A. (1980)). Par ailleurs, les observations sont supposées obtenues à des instants équidistants par exemple :  $(t-1, t, t+1, \dots)$ .

En outre, les termes d'erreur dans les équations (2.1) et (2.2) sont supposés normaux. De plus, nous supposons que l'état initial  $P_{00}^{HT}$  possède une distribution préalable  $f_{P_{00}^{HT}}$  normale, de moyenne  $\tilde{P}_{00}^{HT}$  et de matrice de covariance  $\Sigma_{00}^{HT}$ . La distribution conditionnelle du prix HT du gas-oil  $P_t^{HT}$  à  $t-1$  sachant  $P_{t-1}^{TTC}$ , suit une loi normale multivariée (MVN) :

$$(P_{t-1}^{HT} | P_{t-1}^{TTC}) \sim \text{MVN}(\tilde{P}_{t-1}^{HT}, \Sigma_{t-1}^{HT} | P_{t-1}^{TTC})$$

où  $\tilde{P}_{t-1}^{HT} = E(P_{t-1}^{HT} | P_{t-1}^{TTC})$  est la valeur moyenne de  $P_{t-1}^{HT}$  conditionnée par  $P_{t-1}^{TTC}$ .  $\Sigma_{t-1}^{HT} | P_{t-1}^{TTC}$  est la matrice de variance-covariance de  $P_{t-1}^{HT}$  sachant  $P_{t-1}^{TTC}$ . Notons que  $(X \setminus Y) \sim f$  désigne le fait que la distribution conditionnelle de  $X$  sachant  $Y$  est  $f$ . De surcroit, puisqu'une transformation linéaire d'une distribution gaussienne multivariée ( $P_{t-1}^{TTC}$  sachant  $P_{t-1}^{HT}$ ) i.e  $f_{P_{t-1}^{TTC} | P_{t-1}^{HT}}$  est aussi une distribution multivariée gaussienne, on peut aussi déterminer la distribution conditionnelle de  $P_{t-1}^{HT}$  sachant  $P_{t-1}^{TTC}$  i.e  $(f_{P_{t-1}^{HT} | P_{t-1}^{TTC}})$  par le Théorème de Bayes. Par conséquent, la distribution conditionnelle de  $P_{t-1}^{HT}$  sachant  $P_{t-1}^{TTC}$ ,  $f_{P_{t-1}^{HT} | P_{t-1}^{TTC}}$ , qui est proportionnelle à  $(f_{P_{t-1}^{TTC} | P_{t-1}^{HT}}) * f_{P_{t-1}^{HT}}$  en vertu du Théorème de Bayes est aussi multivariée gaussienne.

La dérivation (3.1) et les équations subséquentes sont récursives au sens où si le vecteur aléatoire conditionnel de la variable (3.1) est donné, l'évaluation de ce dernier aux étapes subséquentes, ne requiert qu'une incrémentation de 1 aux instants suivants.

Supposons pour l'instant que la variable  $(P_{t-1}^{HT} | P_{t-1}^{TTC})$  est donnée et que les paramètres dans les matrices de projection en (2.1) sont connus.

L'algorithme de Filtrage de Kalman dans sa forme générale « prévision-correction » (e. g. Young P. C. (1984)) se présente comme suit :

(i) *Équation du Filtre de Kalman*

Désignons la moyenne et la covariance de la fonction de densité gaussienne  $f_{P_t^{HT} | P_{t-1}^{TTC}}$  par  $\tilde{P}_{t|t-1}^{HT}$  et  $\Sigma_{P_t^{HT} | P_{t-1}^{TTC}}$  respectivement.

Ce que l'on peut encore écrire :

$$(P_t^{HT} | P_{t-1}^{TTC}) \sim \text{MVN} (\tilde{P}_{t|t-1}^{HT}, \Sigma_{P_t^{HT} | P_{t-1}^{TTC}}) \quad (3.2a)$$

et 
$$(P_t^{HT} | P_t^{TTC}) \sim \text{MVN} (\tilde{P}_{t|t}^{HT}, \Sigma_{P_t^{HT} | P_t^{TTC}}) \quad (3.2b)$$

l'équation (3.2) conduit aux équations suivantes :

(ii) *Équation de Prévision*

$$\tilde{P}_{t|t-1}^{HT} = F\tilde{P}_{t-1|t-1}^{HT} + G \text{Prot}_{t-1} \quad (3.3a)$$

avec 
$$\Sigma_{P_t^{HT} | P_{t-1}^{TTC}} = F\Sigma_{P_{t-1}^{HT} | P_{t-1}^{TTC}} F^T + \Gamma \Sigma \omega_i \Gamma^T \quad (3.3b)$$

(calculée à partir de l'équation de Riccati (e. g. Kalman R. E. (1960) Kalman R. E. and Bucy R. S. (1961), Anderson B. O. and Moore (1979) p. 156-8).

(iii) *Équation de Mise à jour ou de Correction*

$$\tilde{P}_{t|t}^{HT} = \tilde{P}_{t|t-1}^{HT} + K(t)V(t) \quad (3.4)$$

où :  $V(t) = P_t^{TTC} - H\tilde{P}_{t|t-1}^{HT}$  représentent les innovations pour  $t=0, 1, 2, \dots, T$ , la matrice de gain de Kalman  $K(t)$  étant :

$$K(t) = \left[ \Sigma_{P_t^{HT} | P_{t-1}^{TTC}} H^T \right] \left[ H \Sigma_{P_t^{HT} | P_{t-1}^{TTC}} H^T + \Sigma v_t \right]^{-1}$$

$$P_{t|t}^{HT} = F P_{t-1|t-1}^{HT} + G \text{Prot}_{t-1} + \left[ \Sigma_{P_t^{HT} | P_{t-1}^{TTC}} H^T \right] \cdot \left[ H \Sigma_{P_t^{HT} | P_{t-1}^{TTC}} H^T + \Sigma v_t \right]^{-1} \cdot \left\{ P_t^{TTC} - H(F P_{t-1|t-1}^{HT} + G \text{Prot}_{t-1}) \right\} \quad (3.5a)$$

Ou en utilisant le lemme 1 (Annexe 1) :

$$P_{\hat{t}}^{HT} = F P_{\hat{t}-1|t-1}^{HT} + G \text{Prot}_{t-1} +$$

$$\left[ \sum P_i^{HT} | P_{i-1}^{TTC} H^T \right] \left\{ (\sum v_i)^{-1} - (\sum v_i)^{-1} H \left[ H^T (\sum v_i)^{-1} H + (\sum P_i^{HT} | P_{i-1}^{TTC})^{-1} \right]^{-1} H^T (\sum v_i)^{-1} \right\} \cdot \left\{ P_i^{TTC} - H (F P_{\hat{t}-1|t-1}^{HT} + G \text{Prot}_{t-1}) \right\} \quad (3.5b)$$

$$\text{et} \quad \sum P_i^{HT} | P_i^{TTC} = \sum P_i^{HT} | P_{i-1}^{TTC} - \sum P_i^{HT} | P_{i-1}^{TTC} H^T (H \sum P_i^{HT} | P_{i-1}^{TTC} + \sum v_i)^{-1} H \sum P_i^{HT} | P_{i-1}^{TTC} \quad (3.5c)$$

#### (iv) Conditions Initiales

Nous supposons que :

$\tilde{P}_{00}^{HT} = P_0$  et  $\sum P_{00}^{HT} = \sum_0$  (3.6). (3.3a) et (3.3b) sont les équations de prévision qui spécifient les erreurs de prévision du vecteur d'état et leur matrice de covariance ; tandis que les équations (3.5a), (3.5b) et (3.5c) permettent la mise à jour de la moyenne et de la matrice de covariance du vecteur d'état lorsque de nouvelles observations  $P_i^{TTC}$  deviennent disponibles. L'estimation de  $P_{\hat{t}}^{HT}$  (3.4) révisée, est la somme de l'estimation projetée obtenue à partir des observations à  $t-1$ ,  $P_{\hat{t}-1}^{HT}$  et d'une combinaison linéaire des erreurs de prévision les plus récentes.  $K(t)$  représente le gain de Kalman ; il détermine les pondérations à attribuer aux erreurs de prévision les plus récentes.

### 3.2. Construction de la vraisemblance

Pour déterminer les estimations du maximum de vraisemblance des matrices  $F$ ,  $G$ ,  $H$  et  $\Gamma$ , nous examinons tout d'abord la distribution conditionnelle de  $P_i^{TTC}$  étant donné  $P_{i-1}^{TTC}$ . En utilisant une argumentation similaire à celle déjà utilisée, nous constatons que la distribution conditionnelle du vecteur des observations  $P_i^{TTC}$  étant donnée  $P_{i-1}^{TTC}$  est multivariée gaussienne avec :

$$\tilde{P}_{\hat{t}-1}^{TTC} = E(P_i^{TTC} | P_{i-1}^{TTC}) = H(F P_{\hat{t}-1}^{HT} + G \text{Prot}_{t-1}) \quad (3.7)$$

$$\text{et} \quad \text{Var}(P_i^{TTC} | P_{i-1}^{TTC}) = \left[ H \sum P_i^{HT} | P_{i-1}^{TTC} H^T + \sum v_i \right] \quad (3.8)$$

Sous l'hypothèse de distribution gaussienne multivariée, la densité de probabilité conditionnelle du vecteur des observations  $P_i^{TTC}$  peut s'écrire comme :

$$p(P_i^{TTC} | P_{i-1}^{TTC}) = (2\pi)^{-n/2} |\sum P_i^{TTC}|^{-1/2} \cdot$$

$$\exp \left\{ -\frac{1}{2} (P_i^{TTC} - \tilde{P}_i^{TTC})^T (\sum P_i^{TTC})^{-1} (P_i^{TTC} - \tilde{P}_i^{TTC}) \right\} \quad (3.9)$$

où :  $P_i^{TTC}$  et  $\sum P_i^{TTC}$  représentent le vecteur-moyen et la matrice de variance-covariance respectivement, qui sont donnés par les équations (3.7) et (3.8).

Étant donné  $P_0^{\text{HT}}$ , en considérant que cette variable possède une distribution connue (distribution préalable Bayésienne) ou en la considérant comme un vecteur de paramètres à estimer ( $\Sigma^{P_0^{\text{HT}}}$  identiquement nulle), la distribution initiale du prix TTC du gasoil devient :

$$(P_0^{\text{TTC}} | P_0^{\text{HT}}) \sim \text{MVN} (H\tilde{P}_0^{\text{HT}}, \Sigma) \quad (3.10)$$

où 
$$\Sigma = H\Sigma^{P_0^{\text{HT}}} H^T + \Sigma^v,$$

La densité du prix TTC du gas-oil est donnée par :

$$p(P_0^{\text{TTC}} | P_0^{\text{HT}}) = (2\Pi)^{-n/2} |\Sigma|^{-1/2}.$$

$$\exp \left\{ -\frac{1}{2} \left[ P_0^{\text{TTC}} - (H\tilde{P}_0^{\text{HT}}) \right]^T (\Sigma)^{-1} \left[ P_0^{\text{TTC}} - (H\tilde{P}_0^{\text{HT}}) \right] \right\} \quad (3.11)$$

Par conséquent, la vraisemblance du vecteur des paramètres  $\vartheta = (F, G, H, \Gamma, P_0^{\text{HT}})$  étant donné les prix HT du gas-oil à l'instant  $t$ , peut être exprimée comme suit :

$$L(\vartheta) = \prod_{\tau=1}^t p(P_\tau^{\text{TTC}} | P_{\tau-1}^{\text{TTC}}) p(P_0^{\text{TTC}} | P_0^{\text{HT}}) \quad (3.12)$$

et la log - vraisemblance devient :

$$\ln L(\vartheta) = \prod_{\tau=1}^t \ln p(P_\tau^{\text{TTC}} | P_{\tau-1}^{\text{TTC}}) + \ln p(P_0^{\text{TTC}} | P_0^{\text{HT}}) = \quad (3.13)$$

$$\text{Constante} - \frac{1}{2} \sum_{\tau=0}^t \ln \left| \sum P_\tau^{\text{TTC}} \right| - \frac{1}{2} \sum_{\tau=0}^t (P_\tau^{\text{TTC}} - \tilde{P}_\tau^{\text{TTC}})^T \left| \sum P_\tau^{\text{TTC}} \right|^{-1} (P_\tau^{\text{TTC}} - \tilde{P}_\tau^{\text{TTC}})$$

Cette dernière expression est maximisée pour obtenir le maximum de vraisemblance du vecteur des paramètres  $\vartheta$ .

$$\hat{\vartheta}_{mv} = \text{Arg Max}_{\vartheta} \prod_{\tau=1}^t f(P_\tau^{\text{TTC}} | P_{\tau-1}^{\text{TTC}}, \tau-1) \cdot f(P_0^{\text{TTC}} | P_0^{\text{HT}}) \quad (3.14)$$

### 3.3. Structure des conditions initiales

Pour un algorithme récursif, les valeurs initiales  $\tilde{P}_{0|0}^{\text{HT}}$ ,  $\Sigma^{P_{0|0}^{\text{HT}}}$  la variance de l'équation d'observation  $P_{0|0}^{\text{HT}}$  et  $\Sigma^{P_{0|0}^{\text{HT}}}$  la matrice des variances-covariances du vecteur d'état doivent être prédéterminées. Le choix de  $\tilde{P}_{0|0}^{\text{HT}}$  et  $\Sigma^{P_{0|0}^{\text{HT}}}$  est discuté dans cette sous-section. Il existe en réalité différentes manières de définir la fonction de vrai-

semblance, liées aux hypothèses faites sur la structure des conditions initiales (e. g. A. G. Harvey 1981a and 1981b, Ansley C. F. and Newbold P. (1980)). Dans notre application concernant le prix du gas-oil, nous maximisons la fonction de vraisemblance en prenant comme hypothèse : tout d'abord, une distribution préalable diffuse pour  $P_{00}^{HT}$ , ensuite que  $P_{00}^{HT}$  peut-être fixé (mais inconnu), et enfin qu'il existe un état d'équilibre stationnaire pour le Filtre de Kalman (e. g. A. C. Harvey and S. Peters (1990), T. Marshall (1992), Shumway R. H., Olsen D. E. and Levy L. J. (1981), L. Ljung and T. Söderström (1983) etc.).

(i) *Distribution Préalable Diffuse*

Si  $P_{00}^{HT}$  possède une distribution préalable diffuse, on peut construire les valeurs initiales à partir des  $k$  premières observations i.e.  $\tilde{P}_k^{HT}$  et  $\sum P_k^{HT}$

La fonction de vraisemblance des observations  $P_{k+1}^{TTC}, \dots, P_T^{TTC}$  peut être définie conditionnellement aux  $P_1^{TTC}, \dots, P_k^{TTC}$ . La décomposition de l'erreur de prévision de la vraisemblance est :

$$\ln L = \text{Constante} - \frac{1}{2} \sum_{\tau=k+1}^T \ln |\sum P_{\tau}^{TTC}|$$

$$- \frac{1}{2} T \sum_{\tau=k+1}^T (P_{\tau}^{TTC} - \tilde{P}_{\tau}^{TTC})^T (\sum P_{\tau}^{TTC})^{-1} (P_{\tau}^{TTC} - \tilde{P}_{\tau}^{TTC}) \quad (3.15)$$

(cf. P. De Jong (1988)).

Pour calculer (3.10), il est possible d'initialiser le Filtre de Kalman à  $t = 0$  avec  $P_0^{HT} = (0,0)$  et  $\sum P_0^{HT} = \chi I$ , où  $\chi$  est un nombre très grand mais fini (e. g. Ansley C. F. and Newbold P. (1980), Anderson B. O. and Moore J. B. (1979), De Jong (1988), Ansley C. F. and Kohn R. (1985)). Les approximations de  $\tilde{P}_k^{HT}$  et  $\sum P_k^{HT}$  sont alors obtenues après  $k$  itérations du Filtre de Kalman. Alternativement,  $\tilde{P}_k^{HT}$  et  $\sum P_k^{HT}$  peuvent être calculées explicitement en écrivant les  $k$  premières observations en terme de  $P_k^{HT}$  (e. g. GLS ; cf. A. C. Harvey (1982)).

(ii) *Vecteur d'État Initial Fixé*

Par ailleurs, si on suppose que le vecteur d'état initial  $P^{HT}$  est fixé ; les observations dans (2.2) peuvent être exprimées en termes de  $P_0^{HT}$  par substitutions répétées. On obtient :

$$P_t^{TTC} = HF^t P_0^{HT} + H \sum_{i=1}^t F^{i-1} G \text{Prot}_{t-i}(\cdot) + H \sum_{i=0}^{t-1} F^i \Gamma \omega_{t-i} + v_t, \quad t = 1, \dots, T$$

(3.16)

avec  $\zeta_t = H \sum_{i=0}^{t-1} F^i \Gamma \omega_{t-i} + v_t$ ; la fonction de vraisemblance est :

$$\ln L = \frac{-T}{2} \ln 2\pi - (1/2) \ln |\Omega_0| -$$

$$- (1/2) \left( P_i^{\text{TTC}} - H(F^i \tilde{P}_0^{\text{HT}} + \sum_{i=1}^t F^{i-1} G \text{Prot}_{t-1}(\cdot)) \right)^T \Omega_0^{-1} \cdot \left( P_i^{\text{TTC}} - H(F^i \tilde{P}_0^{\text{HT}} + \sum_{i=1}^t F^{i-1} G \text{Prot}_{t-1}(\cdot)) \right);$$

et où :  $V(\zeta_t) = \Omega_0$

Il est important de noter que la substitution répétée et l'inversion de la matrice  $TX T$ ,  $\Omega_0$ , peuvent être évitées si l'on utilise les méthodes de Rosenberg B. (1973) ou de Wecker W. E. and Ansley C. F. (1983).

### (iii) État d'Équilibre Stationnaire du Filtre de Kalman

Dans ce cas, le Filtre de Kalman est initialisé à partir des  $k$  premières observations comme dans la méthode I. La seule différence réside dans le fait que la matrice de covariance des erreurs,  $\Sigma^{P^{\text{HT}}}$  est égale à la matrice de covariance de l'état stationnaire  $\overline{\Sigma^{P^{\text{HT}}}}$  (supposée fixe). Le Filtre de Kalman peut être alors appliqué sans la matrice de covariance  $\overline{\Sigma^{P^{\text{HT}}}}$  de l'état  $P_t^{\text{HT}}$  et la fonction de vraisemblance peut être exprimée en termes de décomposition de l'erreur de prévision (cf. (3.11)). Pour obtenir le Filtre de Kalman de l'état stationnaire, il est nécessaire de résoudre au préalable l'équation de Riccati pour  $\overline{\Sigma^{P^{\text{HT}}}}$  (e. g. Anderson B. O. and Moore J. B. (1979) p. 156-8 et Annexe 2 – Théorème 1 tiré de I. N. Roïtenberg (1974)).

## 4. Application

L'application que nous développons concerne le prix du gas-oil en France sur deux périodes en raison du choc pétrolier de 1972 : 1960-1972 et 1973-1985. La période 1986-1992 est utilisée pour la validation du modèle.

## MODÉLISATION PAR LE FILTRE DE KALMAN

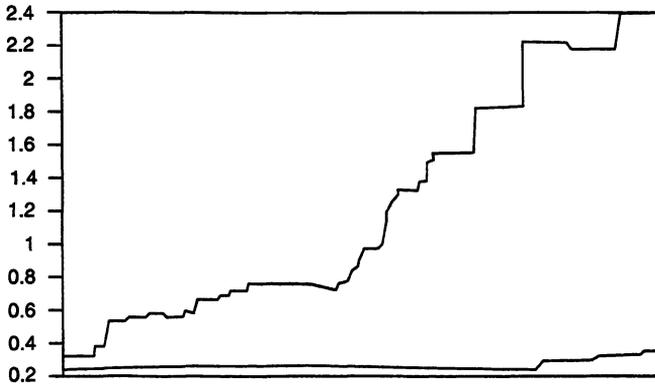


Figure 1a : Prix HT du gas-oil de 1965 à 1985 en France

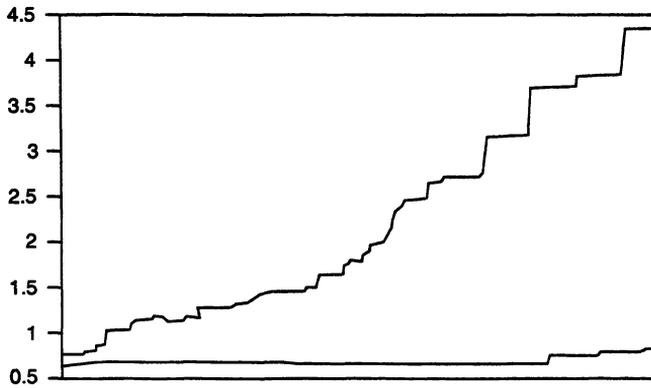


Figure 1b : Prix TTC du gas-oil de 1965 à 1985 en France

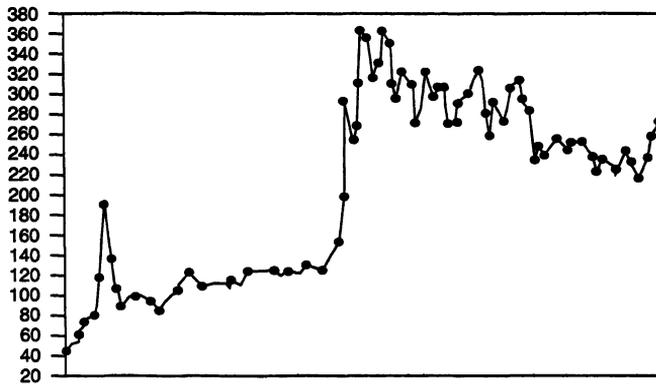


Figure 1c : Prix internationaux du marché de Rotterdam de 1960 à 1985

**4.1. Présentation des Données**

Pour illustrer la procédure d'estimation, nous analysons les données mensuelle de 1960 à 1985 concernant la France. Ces données relatives aux prix HT, TTC et internationaux (Marché de Rotterdam) du gas-oil sont représentées dans la Figure 1.

Étant donné la nature des séries du prix du gas-oil en France (mémoire longue), une différenciation fractionnaire dépendant du degré de différenciation s'imposait. L'estimation du modèle dépendant de  $d$  est nécessairement parcimonieuse (e. g. Hosking M. (1981), Granger C. W. J. and Hallman J. (1991), Granger C. W. J. and R. Joyeux (1980), Granger C. W. J. (1980)).

L'estimation du paramètre essentiel  $d$  repose en réalité sur plusieurs méthodes :

- Une de celles-ci concerne l'approche développée par Lawrance et Koltega (1977) et qui consiste à utiliser les séries ou les variables mais en les réajustant de manière périodique.
- Nous considérons ici l'approche décrite par Granger C. W. J. et R. Joyeux (1980) qui consiste à estimer le logarithme du spectre de la série du prix du gas-oil, différencié à l'ordre  $d$ , aux basses fréquences à partir d'un modèle AR d'ordre  $m$  ; où  $d$  varie de  $-0.9$  à  $0.4$  ( $d = 0$  étant exclu). Le degré de différenciation conduisant aux meilleures prévisions en termes de *RMSE* est alors retenu. Nous justifions dans notre étude le choix de  $d = 0.4$  par des considérations liées à la précision du modèle Espace d'État du prix du gas-oil puisque le *RMSE* est plus faible (e. g. Granger C. W. J. and R. Joyeux (1980)) par le fait que cette transformation permet d'avoir une variance finie du modèle Espace d'État du prix du gas-oil et par la comparaison des résultats issus de ce modèle à ceux d'un modèle ARMA d'ordre peu élevé (e. g. Granger C. W. J. and R. Joyeux (1980)).

Quant à la série Prot des prix internationaux du gas-oil sur le marché de Rotterdam (prix Spot du Brent), elle a été lissée (lissage exponentiel simple) pour tenir compte de son évolution erratique.

Le modèle des prix du gas-oil retenu est un modèle à deux états décrit sous forme canonique observable. Il comporte une perturbation aléatoire dans l'équation d'état, un terme d'erreur aléatoire dans l'équation de la sortie et un bruit gaussien ; il est défini sur deux périodes :

- Période I : 1960-1972 (relative tranquillité) pour  $t \leq 1972$ .
- Période II : 1973-1985 (turbulence) pour  $t > 1972$ .

Sous forme matricielle ce modèle s'écrit :

$$\begin{bmatrix} P_{1,t}^{HT} \\ P_{2,t}^{HT} \end{bmatrix} = \begin{bmatrix} f_{11} & 0 \\ f_{21} & f_{22} \end{bmatrix} \begin{bmatrix} P_{1,t-1}^{HT} \\ P_{2,t-1}^{HT} \end{bmatrix} + \begin{bmatrix} g_1 \\ g_2 \end{bmatrix} [Prot_{t-1}] + \begin{bmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{bmatrix} \begin{bmatrix} \omega_{1t} \\ \omega_{2t} \end{bmatrix}$$

i.e 
$$P_t^{HT} = F P_{t-1}^{HT} + G Prot_{t-1}(\cdot) + \Gamma \omega_t$$

et 
$$P_t^{TTC} = \begin{bmatrix} \gamma_1 & \gamma_2 \end{bmatrix} \begin{bmatrix} P_{1,t}^{HT} \\ P_{2,t}^{HT} \end{bmatrix} + v_t ; v_t \sim (0, \sigma_v^2)$$

Ce modèle Espace d'État est minimal, puisqu'il est à la fois contrôlable et observable.

En effet à partir de (2.1) et (2.2), on peut écrire :

$$\begin{aligned} M &\triangleq [G, FG] & \text{rang}(M) &= 2 \\ N &\triangleq [H', F'H'] & \text{rang}(N) &= 2 \end{aligned}$$

Ces deux conditions sont suffisantes pour identifier complètement (2.1) et (2.2) (e.g. Anderson B. O. and Moore J. B. (1979), Aoki (1990), Burmeister, Wall and Hamilton (1986), Wall K. D. (1987)) puisque  $M$  et  $N$  sont de rang plein et  $\text{rang}(M) = \dim(M) = \text{rang}(N) = \dim(N) = 2$ .

#### 4.2. Estimation et Tests

L'estimation du modèle sur la période : 1960-1985 donne des résultats qui figurent dans la Table 1. Les résultats obtenus font intervenir trois hypothèses concernant la structure des conditions initiales : les premières estimations reposent sur une fonction de vraisemblance de Kalman avec une distribution préalable diffuse. Le Filtre de Kalman est initialisé avec  $P_0^{HT} = (0,0)$  et  $\Sigma_0^{HT} = 10^6 I_4$  (e. g. P. De Jong (1988), A. C. Harvey and Peters S. (1990)). Les estimations obtenues sont comparées à un second ensemble d'estimations où l'on considère que :  $P_0^{HT}(x, y)$  est fixée mais  $x$  et  $y$  ( $x$  et  $y$  étant les deux états initiaux liés à  $P_{1,t}^{HT}$  et  $P_{2,t}^{HT}$  respectivement) sont inconnues et  $\Sigma_0^{HT} = 0$  (e. g. P. De Jong (1988) ou Ansley C. F. and Kohn R. (1985) ou Rosenberg B. (1973)). Néanmoins, pour éviter d'imposer des contraintes de non-singularité, nous prenons comme matrice de covariance de l'état initial, l'inverse de celle utilisée en  $I$  i.e.  $\Sigma_0^{HT} = (10^6 I_4)^{-1}$ .

L'estimation en terme de substitutions répétées donne pour ce couple d'éléments :  $(x, y) = (0,000 15 ; 0,000 12)$ . Le Filtre de Kalman obtenu dans ce cas suppose l'existence d'un processus déterministe pour l'état considéré. Un dernier ensemble d'estimations suppose l'existence d'un état d'équilibre stationnaire et est comparé aux résultats précédemment obtenus. La matrice initiale  $P_0^{HT}$  est donnée par :  $P_0^{HT} = (0,0)$  et la résolution de l'équation de Riccati (cf. Théorème 1) donne :

$$\overline{\Sigma_0^{HT}} = \begin{bmatrix} 0,069 87 & 0 \\ 0 & 9,046 1 \end{bmatrix}$$

Notons que la matrice initiale  $P_0^{HT}$  est construite à partir des  $k$  premières observations avec :

$$\lim_{i \rightarrow \infty} \sum_0^{P_i^{HT}} = \overline{\Sigma_0^{HT}}$$

( $\overline{\Sigma_0^{HT}}$  est obtenue en utilisant les propriétés des matrices symplectiques).

## MODÉLISATION PAR LE FILTRE DE KALMAN

*Table 1 : Paramètres estimés sous les trois hypothèses du Filtre de Kalman*

	Paramètres	S. E.
$f_{11}$	0,844 30 (*)	0,078 52
$f_{12}$	0	—
$f_{21}$	0,097 67	0,689 53
$f_{22}$	0,968 69 (*)	0,034 50
$g_1$	0,000 00	—
$g_2$	0,000 174 (*)	0,000 09
$\Gamma_{11}$	1,000 a	—
$\Gamma_{12}$	0,000 a	—
$\Gamma_{21}$	0,000 a	—
$\Gamma_{22}$	1,000 a	—
$L_\infty$	602,698 20	
SSR	0,333 77	
SEE	0,052 29	
Q-Stat (36)	50,303 22	
LM	128,451 00	
JB	0,095 56	
Constant	-0,015 14	

$L_\infty$  : Log - vraisemblable.

JB : Jarque Bera Normality Test.

SEE : Somme des Carrés des erreurs.

SE : Standard - Error.

(\*) : Significatif à 1 %.

LM : Lagrange Multiplier Test.

SSR : Somme des Carrés des résidus.

Q-(36) : Ljung - Box Q-Statistic.

a : Paramètre fixé lors de l'estimation.

(\*\*) : Significatif à 5 %.

Les résultats de nos applications montrent que les modèles obtenus sont satisfaisants et les termes de perturbation peuvent être considérés comme la réalisation de bruits blancs indépendants ainsi qu'en témoigne l'examen des spectres des résidus.

L'estimation de l'équation de la sortie, quant à elle, indique que  $P_i^{TTC}$  est obtenue avec  $H = [\gamma_1 \ \gamma_2] = [0,042 \ 5 \quad 1,197]$ ; la variance de l'erreur dans l'équation de sortie étant égale à  $\sigma_v^2 = 0,004 \ 61$ .

La décomposition de l'erreur de prévision au niveau de la méthode I (condition initiale diffuse) correspond à la définition usuelle de la fonction de vraisemblance d'un modèle ARIMA.

La méthode II (vecteur d'état initial fixé) et la méthode III (état d'équilibre stationnaire) donnent des sommes résiduelles identiques à celles de la méthode I.

De surcroît, les estimateurs définis par ces différentes méthodes d'initialisation possèdent tous les mêmes propriétés asymptotiques.

Les estimations qui en découlent sont aussi identiques. Ce résultat important pourrait s'expliquer par le fait que dans les échantillons de grandes tailles, et sous l'hypothèse de stabilité asymptotique du modèle étudié, les conditions initiales ne peuvent perturber les trajectoires des processus. Le modèle Espace d'État étudié ici étant linéaire dans les paramètres et dans les variables, il y a donc nécessairement retour à l'équilibre dès lors qu'une perturbation transitoire ou temporaire survient.

### 4.3. Analyse de la L-Stabilité

Les matrices carrées d'ordre 2  $F_i$  ( $i = 1, 2, 3$ ) obtenues de nos estimations vérifient la double condition :

$$f_{11}f_{22} - f_{12}f_{21} < 1 \quad \text{et} \quad f_{11} + f_{22} < 1 + f_{11}f_{22} - f_{12}f_{21}$$

De surcroît, les  $F_i$  sont séquentiellement stables (s-stables) car les matrices  $F_i$  sont productives. Par ailleurs, puisque les  $F_i$  sont s-stables, l'équilibre 0 de l'évolution linéaire :

$$P_t^{\text{HT}} = F P_{t-1}^{\text{HT}}$$

est stable au sens de Lyapunov ou tout simplement L-Stable c'est-à-dire que les effets des conditions initiales disparaissent asymptotiquement.

Les valeurs propres des matrices de transition  $F_i$  du modèle Espace d'État sont donc en module strictement inférieures à 1.

### 4.4. Prévision et mise à jour

Sous l'hypothèse de stabilité asymptotique des différents modèles et compte tenu des tests de validation portant sur les résidus, les prix HT du gasoil ont été projetés puis mis à jour sur la période 1986-1992.

## MODÉLISATION PAR LE FILTRE DE KALMAN

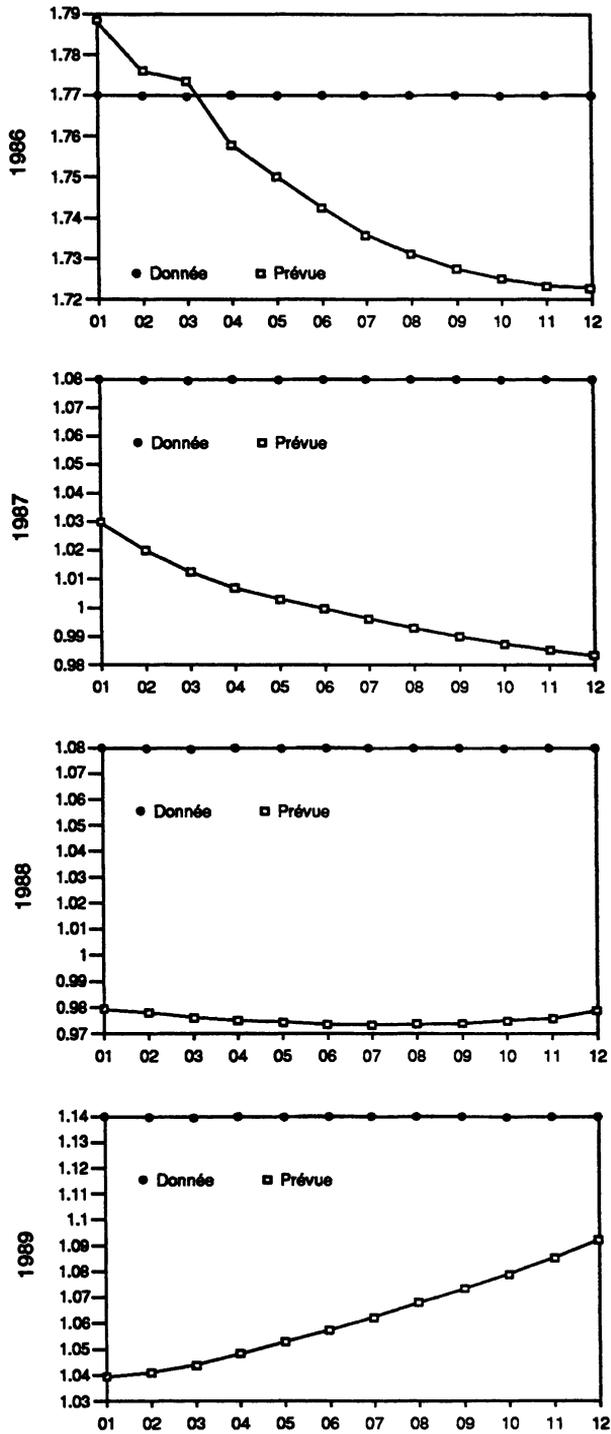


Figure 2 : Prévisions et corrections du prix HT du gas-oil

## MODÉLISATION PAR LE FILTRE DE KALMAN

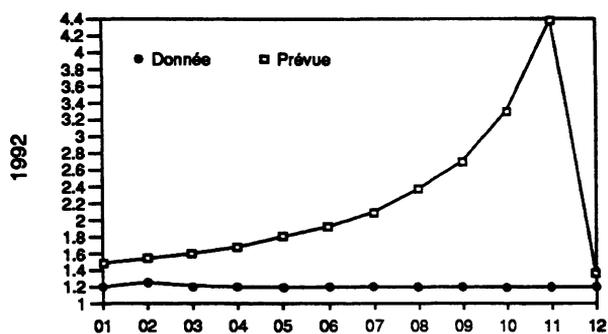
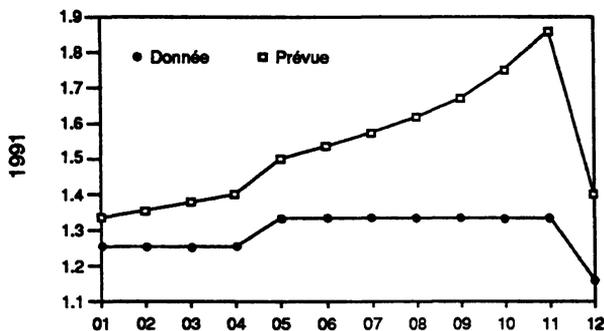
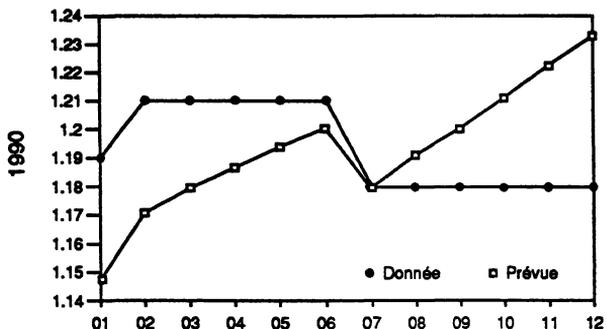


Figure 2 (suite) : Prévisions et corrections du prix HT du gas-oil

MODÉLISATION PAR LE FILTRE DE KALMAN

Table 2 : Performances des Modèles Retenus (Prédicteur Diffus ou Vecteur d'État Fixé ou État d'Équilibre Stationnaire)

Steps	Mean Error	Mean Absolute Error	RMS Error	Theil U	N. Obs.	Time
1	-.18532007E-01	.46270663E-01	.12144715	1.0747515	84	1986/01
2	-.52299457E-02	.43640129E-01	.10122430	.91943761	83	1986/02
3	.40935964E-02	.56760927E-01	.10027760	.83319195	82	1986/03
4	.12484426E-01	.7079944E-01	.10423616	.79884827	81	1986/04
5	.19993594E-01	.83027588E-01	.11283987	.80727277	80	1986/05
6	.26646468E-01	.96611151E-01	.12496901	.83640443	79	1986/06
7	.32462098E-01	.10692167	.13905962	.87694676	78	1986/07
8	.37426579E-01	.11475606	.15430518	.92790290	77	1986/08
9	.41468964E-01	.12136305	.17093032	.98017105	76	1986/09
10	.44792289E-01	.12804347	.18816609	1.0316958	75	1986/10
11	.47252862E-01	.13443866	.20554616	1.0811446	74	1986/11
12	.48501828E-01	.14826359	.27227003	1.4694065	73	1986/12
13	.49723263E-01	.16285630	.28297458	1.3166420	72	1987/01
14	.59855373E-01	.16850953	.29440572	1.4796444	71	1987/02
15	.65586133E-01	.17777360	.31393501	1.5668327	70	1987/03
16	.70818528E-01	.18733685	.33356059	1.6547059	69	1987/04
17	.75608400E-01	.19669279	.35329014	1.7412343	68	1987/05
18	.79996701E-01	.20617132	.37304690	1.8247623	67	1987/06
19	.84077572E-01	.21542611	.39286168	1.9077632	66	1987/07
20	.87726446E-01	.22465380	.41264408	1.9854337	65	1987/08
21	.91064830E-01	.23349832	.43230013	2.0609364	64	1987/09
22	.94082133E-01	.24205440	.45179613	2.1332636	63	1987/10
23	.96756170E-01	.25024154	.47120038	2.2024543	62	1987/11
24	.99071793E-01	.25657197	.48502637	2.2346487	61	1987/12
25	.10115642	.26586711	.50547946	2.3340806	60	1988/01
26	.10288800	.27479708	.52394406	2.4081949	59	1988/02
27	.10427083	.28384777	.54239691	2.4812058	58	1988/03
28	.10533317	.29277390	.56104486	2.5551694	57	1988/04
29	.10608449	.30166427	.57979458	2.6300519	56	1988/05
30	.10658942	.31052280	.59853228	2.6984438	55	1988/06
31	.10672716	.31923443	.61738982	2.7680952	54	1988/07
32	.10656877	.32784411	.63629875	2.8367137	53	1988/08
33	.10610250	.33638928	.65535222	2.9046682	52	1988/09
34	.10530515	.34484969	.67462822	2.9722489	51	1988/10
35	.10406670	.35312928	.69419769	3.0399096	50	1988/11
36	.10250379	.35658369	.70261078	3.0192953	49	1988/12
37	.10086336	.36658682	.72486892	3.1646675	48	1989/01
38	.97589163E-01	.37494959	.74549987	3.2254407	47	1989/02
39	.94418567E-01	.38346353	.76553808	3.2969592	46	1989/03
40	.90812869E-01	.39185175	.78576048	3.3702532	45	1989/04
41	.86789914E-01	.40026941	.80609070	3.4421528	44	1989/05
42	.82314441E-01	.41003443	.82974402	3.5392318	43	1989/06

MODÉLISATION PAR LE FILTRE DE KALMAN

Table 2 (suite) : Performances des Modèles Retenus (Prédicteur Diffus ou Vecteur d'État Fixé ou État d'Équilibre Stationnaire)

Steps	Mean Error	Mean Absolute Error	RMS Error	Theil U	N. Obs.	Time
43	.77383433E-01	.41805271	.84939852	3.5818911	42	1989/07
44	.72095562E-01	.42716183	.87100086	3.6430183	41	1989/08
45	.66599993E-01	.43694374	.89377376	3.7067736	40	1989/09
46	.61015938E-01	.44757027	.91793149	3.7738261	39	1989/10
47	.55374002E-01	.45899666	.94337877	3.8435406	38	1989/11
48	.49352728E-01	.46384689	.95406031	3.7929077	37	1989/12
49	.44131181E-01	.47894826	.98429646	3.9833341	36	1990/01
50	.37525475E-01	.49233296	1.0130694	4.0699690	35	1990/02
51	.30635701E-01	.50575164	1.0412944	4.1418926	34	1990/03
52	.23502642E-01	.51984341	1.0703945	4.2250725	33	1990/04
53	.15821812E-01	.52947846	1.0905093	4.2255664	32	1990/05
54	.74491662E-02	.54847979	1.1286964	4.4422119	31	1990/06
55	-.16974714E-02	.56344389	1.1581081	4.5315833	30	1990/07
56	-.10675451E-01	.58100834	1.1905104	4.7026333	29	1990/08
57	-.20888920E-01	.59922279	1.2242663	4.8617774	28	1990/09
58	-.31874415E-01	.61883991	1.2595630	5.0309411	27	1990/10
59	-.43631623E-01	.63999376	1.2961937	5.2099719	26	1990/11
60	-.56428448E-01	.66966927	1.3445828	5.5398580	25	1990/12
61	-.70353366E-01	.68936187	1.3792500	5.5126797	24	1991/01
62	-.89206205E-01	.71160130	1.4214362	5.4532327	23	1991/02
63	-.10898533	.73871227	1.469	5.5114122	22	1991/03
64	-.13156861	.76421257	1.5156553	5.4851579	21	1991/04
65	-.15832384	.78299802	1.5509456	5.3643784	20	1991/05
66	-.18969911	.81365464	1.6095332	5.5065691	19	1991/06
67	-.23134406	.83721232	1.6650683	5.3894394	18	1991/07
68	-.27930975	.86174273	1.7241559	5.3384510	17	1991/08
69	-.33778082	.88449362	1.7867751	5.2855882	16	1991/09
70	-.41053972	.90354278	1.8530081	5.2291754	15	1991/10
71	-.50378092	.91650874	1.9220483	5.1650160	14	1991/11
72	-.23548151	.69505470	1.6989519	4.5672070	13	1991/12
73	-.30668147	.69501013	1.7506546	4.3936076	12	1992/01
74	-.37041801	.73085726	1.8381475	4.7367671	11	1992/02
75	-.44782481	.77806822	1.9475208	4.9833594	10	1992/03
76	-.53900148	.83202603	2.0633530	5.1669726	9	1992/04
77	-.64459025	.90980132	2.2033936	5.4698950	8	1992/05
78	-.77659950	1.0135498	2.3719397	5.7948101	7	1992/06
79	-.94769305	1.1566467	2.5799685	6.1746830	6	1992/07
80	-.1.1791948	1.3645179	2.8459052	6.6267184	5	1992/08
81	-.1.5176883	1.6845901	3.2037687	7.1779451	4	1992/09
82	-.2.0732532	2.2260385	3.7245159	7.8721476	3	1992/10
83	-.3.1760684	3.3162849	4.5918576	8.7865105	2	1992/11
84	.14412811	.14412811	.14412811	.1968963	1	1992/12

Les prévisions obtenues portent sur un horizon très éloigné (1986 à 1992), soit au total 84 prévisions réalisées.

La moyenne des erreurs qui mesure les biais de prévision est proche de 0 pour un horizon de prévision ne dépassant pas  $h = 10$ .

Les autres statistiques (Mean Absolute Error, RMS Error, Theil U) sont des mesures de précision. Ainsi, on remarque qu'au-delà de l'horizon  $h = 10$ , la précision du modèle est également très sujette à caution. Une solution à ce problème serait de limiter l'horizon de prévision à  $h = 10$  et de procéder à une re-estimation périodique du modèle de base au fur et à mesure que de nouveaux échantillons deviennent disponibles (avant l'utilisation de l'algorithme « prévision-correction » de Kalman).

Au total, on remarque qu'à court terme ( $h \leq 10$ ) le modèle Espace d'État du prix du gasoil présente des qualités statistiques très satisfaisantes dans la prévision du prix du gasoil en France.

### Discussion et conclusion

Les coefficients des différents modèles obtenus sont identiques ; de même que les termes résiduels SSR et SEE et celles des Log-Vraisemblances ; cela tient certainement à la nature des conditions initiales dont l'effet disparaît asymptotiquement. L'étude de la performance du modèle Espace d'État du gasoil (ME, MAE, RMSE, Theil U) indique, à très court terme, une absence de biais de prévision si l'on tient compte des deux premiers critères d'évaluation (Mean Error et Mean Absolute Error).

Les autres statistiques font ressortir la très bonne précision, à court terme, du modèle de prévision utilisé. Le modèle retenu présentant en moyenne moins de 8 % d'erreur (*cf.* RMSE). Ce résultat s'avère important car il montre qu'asymptotiquement et sous réserve de la L-Stabilité, le choix des conditions initiales n'influence pas la qualité de la prévision et la mise à jour des prix HT du gasoil.

Certes, l'application du Filtre de Kalman à l'analyse de l'évolution des prix HT du gasoil en France n'épuise pas totalement l'analyse récursive de l'évolution des prix HT du gasoil ; celle-ci pourrait être généralisée en supposant par exemple l'existence de termes d'erreurs non-normales dans les différentes équations ou la prise en compte de coefficients non-linéaires dans le modèle Espace d'État. Néanmoins notre approche fournit des résultats intéressants quant à la description, à l'analyse et à la prévision du prix HT du gasoil en France et permet ainsi de montrer l'efficacité des modèles Espace d'État.

De surcroît, lorsque ces derniers sont asymptotiquement stables au sens de Lyapunov; leurs trajectoires ne peuvent diverger, dès lors qu'on choisit les conditions initiales dans un voisinage de l'équilibre.

ANNEXE 1

**Lemme 1 :** (Inversion de Matrices) soient  $A, B, C$  et  $D$  des matrices de dimensions compatibles telles que le produit  $BCD$  et la somme  $A + BCD$  existent. Alors :

$$[A + BCD]^{-1} = A^{-1} - A^{-1}B[DA^{-1}B + C^{-1}]^{-1}DA^{-1} \quad (A1)$$

**Preuve :** Multiplions le membre de droite de (A1) par  $A + BCD$ .

Ce qui donne :

$$I + A^{-1}B[DA^{-1}B + C^{-1}]^{-1} \{ [DA^{-1}B + C^{-1}]CD - D - DA^{-1}BCD \} =$$

$$I + A^{-1}B[DA^{-1}B + C^{-1}]^{-1} \{ 0 \} = I$$

ce qui prouve (A1).

ANNEXE 2

**Théorème 1 :** (Résolution de l'Équation de Ricatti)

La matrice symétrique définie positive  $\Sigma$ , solution de l'équation matricielle :

$$s(\Sigma) = 0 \quad (A1)$$

est déterminée par la relation :

$$[-\Sigma, E] \Delta = 0 \quad (A2)$$

ou par la relation équivalente :

$$\Delta(-M) \begin{bmatrix} E \\ \Sigma \end{bmatrix} = 0 \quad (A3)$$

avec  $E$ , la matrice unité. Ici,  $\Delta(M)$  est le polynôme dont la variable est la matrice. Le polynôme scolaire correspondant  $\Delta(\lambda)$  est déterminé par la relation :

$$\det(\lambda E - M) = (-1)^n \Delta(\lambda) \Delta(-\lambda) \quad (A4)$$

Le polynôme  $\Delta(\lambda)$  étant de degré  $n$  et ses zéros possédant des parties réelles négatives.

**Preuve :** (cf. I. N. Roïtenberg (1974), pp. 436-439).

RÉFÉRENCES

- ANSLEY C.F. and KOHN R. (1985) "Estimation, Filtering and Smoothing in State Space Models with Incompletely Specified Initial Conditions", *Annals of Statist.*, 1286-1316.
- ANSLEY C.F. and KOHN R. (1989) "Filtering and Smoothing in State Models with Partially Diffuse Initial Conditions", *J. of Time Series Anal.*, 275-293.
- ANSLEY C.F. and NEWBOLD P. (1980) "Finite Sample Properties of Estimation for ARMA Processes", *J. of Econometrics*, 159-183.
- AOKI M. (1991) "A State Space Time Series Modelling Method without Individual Detrending", *J. of Time Series Anal.*, 1-26.
- BELL W. and HILLMER S. (1991) "Initializing the Kalman Filter for Non-Stationary Time Series Models", *J. of Time Series Anal.*, 283-300.
- BOYLES R.A. (1983) "On the Convergence of the EM Algorithm", *JRSS, B*, vol. 45, 47-50.
- BURMEISTER E., WALL K.D. and HAMILTON J.D. (1986) "Estimation of Unobserved Expected Monthly Inflation using Kalman Filtering", *J. of Bus. and Eco. Statist.*, 147-160.
- CAINES P.E. (1988) *Linear Stochastic Filtering*, New York, J. Wiley.
- DAVID S.T. and WALL K.D. (1991) "Bootstrapping State Space Models : Gaussian Maximum Likelihood Estimation and Kalman Filter", *JASA*, 1024-1033.
- DE JONG P. (1988) "The Likelihood for a State Space Model", *Biometrika*, 165-169.
- DE JONG P. (1988) "Stable Algorithms for the State Space Model", *J. of Time Series Anal.*, 143-157.
- EDWARDS C.L. and HOWREY E.P. (1991) "A "True" Time Series and its Indicators: an Alternative Approach", *JASA*, 878-882.
- GRANGER C.W.J. (1980) "Long Memory Relationships and the Aggregation of Dynamic Models", *J. of Econometrics*, 227-238.
- GRANGER C.W.J. and HALLMAN J. (1991) "Long Memory Series with Attractors", *Oxford Bull. of Eco. and Statist.*, 11-26.
- GRANGER C.W.J. and JOYEUX R. (1980) "An Introduction to Long-Memory Time Series Models and Fractional Differencing", *J. of Time Series Anal.*, 15-29.
- GOODWIN G.C. and SIN K.S. (1984) *Adapting Filtering, Prediction and Control*, Englewood, Cliffs NJ: Prentice-Hall.
- HARVEY A.C. (1982) *Estimation Procedures for a Class of Univariate Time Series Models*, LSE Econometrics, Program Discussion Paper, n° A 28.
- HARVEY A.C. (1981a) *Time Series Models*, Oxford, Philipp Allan and Atlantic Highlands, NJ: Humanities Press.
- HARVEY A.C. (1981b) "Finite Sample Prediction and Overdifferencing", *J. of Time Series Anal.*, 221-232.
- HARVEY A.C. and PETERS S. (1990) "Estimation Procedures for structural Time Series Models", *J. of Forecasting*, 89-108.
- HARVEY A.C. and PIERCE R.G. (1984) "Estimation Missing Observations in Economic Time Series", *JASA*, 125-131.
- HOSKING J.R.M. (1981) "Fractional Differencing", *Biometrika*, 165-176.

- JONES R.H. (1980) "Maximum Likelihood Fitting of ARMA Models to Time Series with Missing Observation", *Technometrics*, 389-395.
- KALMAN R.E. (1960) "A New Approach to Linear Filtering and Prediction Problem", *Trans. of the American Soc. of Mechanical Engineers, Series D* 82, 35-45.
- KALMAN R.E. (1987) "Non-Gaussian State Space Modeling of Non-Stationary Time Series", *JASA*, 1032-1041.
- KALMAN R.E. and BUCY R.S. (1961) "New Results in Linear Filtering and Prediction", *Trans. of the American Soc. of Mechanical Engineers, Series D* 82, 95-106.
- KOHN R. and ANSLEY C.F. (1989) "A Fast Algorithm for Signal Extraction, Influence and Cross-validation in State Space Models", *Biometrika*, 65-79.
- LAWRANCE A. and KOTTEGA N.T. (1977) "Stochastic Modelling of Riverflow Time Series", *JRSS, A* 140, 1-47.
- LEDOLTER J. and KAHL D.R. (1982) "An Empirical Evaluation of Adaptive Filtering", *Proceedings of American Institute for Decision Sciences*, San Francisco, 342-344.
- LJUNG L. and SÖDERSTRÖM T. (1983) *Theory and Practice of Recursive Identification*, The MIT Press, A.S. Willkry ed.
- MARSHALL P. (1992) "State Space Models with Diffuse Initial Conditions", *J. of Time Series Anal.*, 411-414.
- MEHRA R.K. (1979) "Kalman Filters and their Applications to Forecasting", *TIMS Studies in the Management Sciences*, 12, 75-94.
- MORRIS N.D. and PFEFFERMANN D. (1984) "A Kalman Filter Approach to the Forecasting of Monthly Time Series Affected to Moving Festivals", *J. of Time Series Anal.*, 255-268.
- PAGAN A. (1980) "Some Identification and Estimation Results for Regression Models with Stochastically Varying Coefficients", *J. of Econometrics*, 341-363.
- ROÏTENBERG I.N. (1974) *Théorie du Contrôle Automatique*, Ed. Mir, Moscou, 1974.
- ROSENBERG B. (1973) "Random Coefficient Models: The Analysis of a Cross Section Time Series by Stochastically Convergent Parameters Regression", *Annals of Econometric and Social Measurement*, 799-820.
- SCHWEPPE F. (1965) "Evaluation of Likelihood Functions for Gaussian Signals", *JEEE Trans. on Inf. Theory*, 61-70.
- SHUMWAY R.H., OLSEN D.E. and LEVY L.J. (1981) "Estimation and Tests of Hypotheses for the Initial Mean and Covariance in the Kalman Filter Model", *Communications in Statistics Theory and Methods*, 1625-41.
- SINGH A.C. and ROBERTS G.R. (1992) "State Space Modelling of Cross-classified Time Series of Counts", *Inter. Statist. Rev.*, 60, 321-335.
- SMITH J.Q. (1990) "Non-Linear State Space Models with Partially Specified Distributions on States", *J. of Forecasting*, 137-150.
- STOFFER D.S. and WALL K.D. (1991) "Bootstrapping State Space Models: Gaussian Maximum Likelihood Estimation and the Kalman Filter", *JASA*, 1021-1033.
- SULLIVAN P.J. (1992) "Kalman Filter Approach to Catch-at-Length Analysis", *Biometrics*, 237-257.
- WECKER W.E. and ANSLEY C.F. (1983) "The Signal Extraction Approach to Non-linear Regression and Spline Smoothing", *JASA*, 81-89.
- YOUNG P.C. and NG C.N. (1990) "Recursive Estimation and Forecasting of Non-stationary Time Series", *J. of Forecasting*, 173-204.