

ROYAUME DU MAROC

المملكة المغربية



المنذوبية السامية للتخطيط

HAUT COMMISSARIAT AU PLAN

DIRECTION DE LA STATISTIQUE

مديرية الإحصاء

Séminaire sur le traitement des séries chronologiques, Octobre 2010

Application des modèles espace-état et du filtre de Kalman pour la décomposition tendance/cycle

Préparé par : Mr Amine AMAR

Préliminaire

Les modèles dynamiques à facteurs linéaires ou modèles espace-état constituent une classe particulière des modèles dynamiques à facteurs ou à variables cachées. Etant donné la généralité de la formalisation espace-état, de nombreux et divers modèles ou méthodes économétriques peuvent être reformulés et estimés dans un cadre probabiliste : les modèles à facteurs, les filtre HP et les modèles à coefficients aléatoires.

En particulier, la formulation espace-état est appliquée pour les décompositions tendance-cycle. Ainsi, au lieu de recourir à un lissage par un filtre HP, le modèle à composantes inobservables sous la formulation espace-état détermine une tendance douce et un cycle, dont l'interprétation est facilitée par des paramètres de lissage.

Nous essayons dans le présent papier, d'appliquer les modèles espace-état pour constituer un indicateur composite de l'industrie de transformation. Cet indice permet d'expliquer la plus grande partie des évolutions de chacune des soldes d'opinion des industriels. Les soldes d'opinion sont extraits des enquêtes de conjoncture concernant l'industrie de transformation, effectuées par la Direction de la Statistique. Dans un premier point, nous rappelons les principaux concepts liés aux modèles espace-état et les algorithmes, fondamentalement le filtre de Kalman, permettant d'estimer les variables cachées correspondantes aux modèles présentés.

Le plan du présent papier est structuré comme suit :

- Une première partie concernant la présentation du contexte et de la problématique de l'étude ;
- Une deuxième partie qui présente les fondements théoriques utilisés ;
- Une application pratique.

Partie I : Présentation du contexte et de la problématique de l'étude

La conjoncture revêt un aspect fondamental tant par ses incidences sur l'activité économique que par son influence sur le comportement des différents agents économiques. Son étude devient, avec les changements rapides de l'économie, un instrument indispensable pour la gestion des affaires publiques et pour l'entreprise privée qui doit tenir compte des indicateurs de conjoncture pour la conduite de sa gestion et pour le suivi des facteurs qui président à son évolution.

En particulier, le grand intérêt de l'analyse de la conjoncture réside dans le suivi régulier et permanent de l'évolution de l'activité économique, constituant ainsi une sorte de veille et d'alerte économique qui permet aux opérateurs publics et privés de comprendre le présent (ou le passé immédiat), de prévoir le proche avenir et d'agir en conséquence.

Par ailleurs, les enquêtes de conjoncture offrent un type d'information particulièrement utilisé pour l'analyse conjoncturelle. Cependant, le nombre et la diversité des questions posées rendent souvent délicate l'interprétation des résultats obtenus. C'est pourquoi il semble utile de chercher à construire un indicateur unique à partir de l'ensemble des réponses fournies par

une enquête, et qui pourrait constituer une sorte de résumé de l'information qu'elle contient. Une façon simple de le faire consiste à calculer la moyenne simple de plusieurs soldes d'opinion relatifs à l'état de l'activité économique. Cette méthode présente l'avantage de la simplicité, mais le choix des questions retenues comme celui des pondérations attribuées à chaque question reste délicat.

Une approche plus satisfaisante consiste à supposer que chaque variable peut être décomposée en deux composantes orthogonales entre elles : l'une commune à l'ensemble des séries, et l'autre spécifique à la variable considérée. Les modèles à facteurs dynamiques constituent le cadre d'analyse approprié pour formaliser de telles hypothèses. En plus, deux méthodes peuvent être utilisées pour estimer ce genre de modèles. La première se place dans le domaine des fréquences. Dans ce cas, la dynamique du modèle n'a pas besoin d'être spécifiée : les méthodes standards d'analyse factorielle peuvent être utilisées, de façon à décomposer la matrice de densité spectrale. La deuxième méthode relève plus directement du domaine temporel : dès lors que la dynamique des différentes composantes a été spécifiée, le modèle peut être mis sous une forme espace-état et estimé par le filtre de Kalman.

Partie II : Fondements théoriques

➤ Modèles à composantes inobservables :

Les modèles univariés ou multivariés à composantes inobservables proposés par Andrew Harvey (Harvey, 1989 ; Koopman et Harvey, 1997) sont des modèles statistiques qui ont l'avantage sur d'autres méthodes plus descriptives de fournir une modélisation dynamique temporelle de la tendance et des cycles et d'être soumis à l'estimation et à des tests.

Dans les modèles à composantes inobservables de Harvey, une série y_t est décomposée de manière additive en une tendance μ_t , un cycle ψ_t et une composante irrégulière ε_t . Chacune de ces composantes est stochastique et elles sont mutuellement non corrélées entre elles. Les deux composantes cyclique et irrégulière respectent la propriété statistique de stationnarité. La composante irrégulière ε_t est un bruit blanc.

$$y_t = \mu_t + \psi_t + \varepsilon_t \quad t = 1, \dots, T$$

Modélisation de la tendance :

La tendance μ_t est modélisée sous la forme :

$$\begin{aligned} \mu_t &= \mu_{t-1} + \beta_{t-1} + \eta_t \\ \beta_t &= \beta_{t-1} + \zeta_t \end{aligned}$$

Où η_t et ζ_t sont deux bruits blancs indépendants, respectivement de variance σ_η^2 et σ_ζ^2 . Le bruit η_t permet au niveau de la tendance de fluctuer, tandis que ζ_t fait varier la pente β_t . Ce modèle général est dénommé « Tendance localement linéaire ».

La modélisation de la tendance proposée par Harvey permet d'engendrer plusieurs types de tendance, la forme déterministe étant un cas limite. La possibilité d'un choix concernant la localisation des aléas, sur le niveau ou la pente de la tendance, permet d'élargir les formulations disponibles de la tendance stochastique.

Modélisation du cycle :

La composante cyclique est un processus linéaire stationnaire susceptible de faire apparaître une alternance relativement régulière de pics et de creux, tout en admettant une certaine persistance des phases du cycle et d'éventuelles dissymétries entre elles. Le cycle ψ_t est modélisé sous la forme :

$$\begin{bmatrix} \psi_t \\ \psi_t^+ \end{bmatrix} = \phi \begin{bmatrix} \cos\lambda & \sin\lambda \\ -\sin\lambda & \cos\lambda \end{bmatrix} \begin{bmatrix} \psi_{t-1} \\ \psi_{t-1}^+ \end{bmatrix} + \begin{bmatrix} \kappa_t \\ \kappa_t^+ \end{bmatrix}$$

Où les perturbations κ_t et κ_t^+ sont deux bruits blancs indépendants entre eux et de même variance σ_κ^2 . Le paramètre ϕ , compris entre 0 et 1, désigne le facteur d'amortissement du cycle au cours du temps, tandis que la fréquence de ce dernier est donnée par le paramètre λ , qui appartient à l'intervalle $[0, \pi]$. La période du cycle (exprimée par exemple en nombre de trimestres s'il s'agit d'une série trimestrielle) correspond à $2\pi/\lambda$: c'est la durée théorique du cycle, c'est-à-dire le temps nécessaire pour qu'une oscillation cyclique complète ait lieu à la suite d'une impulsion initiale.

ψ_t^+ est une variable « duale » du cycle ψ_t , qui est nécessaire à la formulation récursive de ce dernier.

Le modèle « tendance-cycle », présenté dans un cadre univarié, peut faire l'objet d'une généralisation multivariée, où les décompositions propres à plusieurs variables y_t sont estimées simultanément. L'intérêt de la dimension multivariée est notamment de déterminer si les composantes stochastiques propres à chacune des variables y_t dépendent de facteurs communs, c'est-à-dire si les innovations qui donnent naissance à ces composantes sont étroitement corrélées.

L'estimation du modèle à composantes inobservables est effectuée selon la méthode du maximum de vraisemblance, en utilisant la représentation appelée espace-état, qui décompose le modèle en une équation d'observation et une équation d'état.

L'extraction des composantes inobservables ; tendance et cycle, peut se faire à l'aide du filtre de Kalman. Ce dernier donne une estimation récursive du vecteur d'état à l'instant t conditionnellement à l'information disponible jusqu'en $t-1$.

➤ **Utilisation des filtres pour la décomposition tendance/cycle :**

L'extraction du cycle économique et la détermination de la position d'une économie dans son cycle sont cruciales, notamment pour porter un diagnostic sur les causes et pour proposer des traitements. Il s'agit bien de vérifier est ce que les fluctuations observées dans les séries traitées sont dues aux phénomènes conjoncturels, tendant à se résorber d'eux-mêmes au cours d'une autre phase du cycle économique, ou de phénomènes de long terme dont la correction requiert des mesures structurelles ?

Sur le plan théorique, l'extraction de la tendance et du cycle d'une série temporelle consiste à décomposer cette série en somme d'une composante intégrée d'ordre égal ou supérieur à 1 ; la tendance, et d'une composante stationnaire ; le cycle. Le cycle lui-même peut être décomposé entre cycle proprement dit et composante irrégulière, tous deux stationnaires.

Par ailleurs, les agrégats économiques ont été souvent considérés stationnaires autour d'une tendance déterministe. C'est ce qui explique le recours aux moyennes mobiles pour filtrer le cycle des séries. La moyenne mobile choisie doit conserver les tendances déterministes, réduit au minimum la composante irrégulière et annule les composantes périodiques. Les développements récents des séries temporelles ont permis de comprendre que les principaux agrégats macroéconomiques ne sont pas toujours représentés comme des séries stationnaires autour d'une tendance déterministe. Ainsi, les mesures empiriques des cycles à partir des moyennes mobiles deviennent parfois caduques pour caractériser le cycle d'affaires. Il apparaît nécessaire de prendre en compte le degré de stationnarité des séries et d'étudier la pertinence des chocs stochastiques. Le modèle de Beveridge et Nelson et le filtre de Hodrick et Prescott (HP) s'inscrivent parmi les modèles qui prennent en compte cette réalité.

Hodrick et Prescott ont proposé une méthode qui dégage une tendance fluctuante, tenant compte notamment des évolutions récentes de la série, et un cycle stationnaire. Cette méthode requiert le choix d'un "paramètre de lissage" permettant de régler l'importance des fluctuations de la tendance. Le choix de ce paramètre est arbitraire et cette méthode est peu satisfaisante au voisinage des extrémités de la série. Enfin, le cycle ainsi dégagé inclut la composante irrégulière, ce qui obscurcit sa lecture.

La méthode de Beveridge et Nelson sépare un cycle et une tendance présentant une forte corrélation négative.

Baxter et King (1998) (BK) ont proposé une méthode d'extraction de la tendance et du cycle simple à comprendre et à mettre en œuvre. Cette méthode isole également une composante irrégulière. Le cycle et l'irrégulier ainsi extraits sont stationnaires. La méthode proposée repose sur le choix a priori des fréquences définissant le cycle.

Les filtres BK et HP présentent l'avantage d'être beaucoup plus aisée à utiliser que les modèles à composantes inobservables. En revanche, ils ne s'inscrivent pas dans un cadre probabiliste et ils ne fournissent pas d'intervalle de confiance autour des composantes estimées.

➤ **Présentation des modèles espace-état :**

L'étude de systèmes physiques émettant au cours du temps des signaux déterminés par des états internes non observés, a conduit à développer en traitement du signal les modèles dits espace-état. Ces modèles intègrent la distinction entre les variables observées (le signal) et les variables cachées (l'état interne). Ils sont constitués d'une ou plusieurs :

- Equation(s) de mesure décrivant la manière dont les variables observées sont générées par les variables cachées et les résidus.
- Equations (s) d'état décrivant la manière dont les variables cachées sont générées à partir de leur retard et d'innovations.

Les modèles espace-état peuvent s'écrire sous une forme dite développée qui implique, sous certaines hypothèses, que la variable de mesure s'écrit sous la forme d'une moyenne mobile.

L'intérêt de recourir aux modèles espace-état pour décomposer ou prévoir une série temporelle relativement aux techniques classiques développées par Box et Jenkins, réside dans le fait d'une part, qu'on s'épargne des problèmes de stationnarité et de racine unitaire qui

se posent préalablement à l'estimation d'un modèle ARMA. En particulier, les résultats du filtre de Kalman restent valides en présence de séries non-stationnaires.

Par ailleurs, ce cadre permet également de relâcher l'hypothèse d'une distribution gaussienne pour les bruits. D'autre part, l'estimation optimale prend en compte l'information disponible à partir de la date initiale $t = 0$, alors que les estimateurs optimaux ARMA prennent en compte l'information à partir de $t = -\infty$. Finalement, les coefficients du modèle peuvent évoluer et ne sont pas obligatoirement considérés comme invariants au cours de la période d'estimation.

Dans leur version élémentaire, les modèles espace-état reposent sur un certain nombre d'hypothèses principales : les équations de mesure et d'état sont linéaires ; les bruits d'observation et d'innovation sont des bruits blancs ; les variables cachées suivent à un instant initial donné une loi gaussienne. À ces dernières, se sont ajoutées des hypothèses secondaires permettant de déterminer la forme canonique: l'indépendance entre les bruits d'observation et d'innovation et l'indépendance entre la variable cachée initiale et ces bruits (condition de causalité). Toutes ces hypothèses sont destinées à simplifier les procédures d'estimation.

Soit :

- \mathbf{Y}_t : l'observation ou la variable de mesure ;
- \mathbf{Z}_t : la variable d'état à la date t ;
- $\boldsymbol{\varepsilon}_t$: le vecteur des innovations à la date t ;
- $\boldsymbol{\eta}_t$: le vecteur des erreurs de mesures à la date t ;
- \mathbf{A}_t : la matrice de transition ;
- \mathbf{C}_t : la matrice de mesure ;
- $\mathbf{X}_{1,t}, \mathbf{X}_{2,t}$: des variables exogènes prédéterminées ;
- $\mathbf{C}_t \mathbf{Z}_t$: le signal à la date t .

Soit un processus multidimensionnel \mathbf{Y}_t , on appelle modèle **espace-état** de ce processus, le système (I) décrit par les équations matricielles (1) et (2) :

$$\begin{cases} (1): Z_{t+1} = A_t Z_t + B_t X_{1,t} + \varepsilon_t \\ (2): Y_t = C_t Z_t + D_t X_{2,t} + \eta_t \end{cases} \text{ où } \begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix} \sim NID \left(0, \begin{pmatrix} Q_t & S_t \\ S_t' & R_t \end{pmatrix} \right) \quad (I)$$

où les matrices \mathbf{A}_t , \mathbf{C}_t sont de taille $K \times K$ et $n \times K$, \mathbf{B}_t et \mathbf{D}_t sont des matrices déterministes de taille $K_1 \times K$ et $K_2 \times K$ et \mathbf{Z}_0 est un vecteur aléatoire de loi $N(m, P)$ indépendant du bruit blanc normal.

Le système (I) peut s'écrire sous la **forme développée** suivante :

$$\begin{cases} (1): Z_{t+1} = \prod_{j=0}^t A_{t-j} Z_0 + \sum_{j=1}^t (A_{t-j} \prod_{k=0}^{j-1} B_{t-k}) X_{1,t-j} + \sum_{j=1}^t (\prod_{k=0}^{j-1} A_{t-k}) \varepsilon_{t-j} + \varepsilon_t \\ (2): Y_t = C_t [\prod_{j=0}^t A_{t-j} Z_0 + \sum_{j=1}^t (A_{t-j} \prod_{k=0}^{j-1} B_{t-k}) X_{1,t-j} + \sum_{j=1}^t (\prod_{k=0}^{j-1} A_{t-k}) \varepsilon_{t-j} + \varepsilon_t] + D_t X_{2,t} + D_t X_{2,t} + \eta_t \end{cases}$$

Les variables d'état et de mesure s'écrivent donc en fonction de la variable d'état initiale, du passé des erreurs de mesure et des innovations ainsi que des variables exogènes. Cette forme

est particulièrement utile lorsqu'on s'intéresse à l'estimateur des moindres carrés généralisés du vecteur d'état ou à l'initialisation du filtre de Kalman.

Le système (I) est dit sous **forme canonique** si et seulement si :

$$E(\eta_s) = E(\eta Z) = E(\eta T Z) = 0 \quad t, s=1, \dots, T$$

Le modèle espace-état est alors dit causal et inversible.

On appelle dimension minimale d'un système admettant une représentation espace-état, la plus petite dimension possible du vecteur d'état, K^* . En particulier, la représentation est dite minimale si A_t est de taille $(K^* \times K^*)$.

Enfin, le modèle espace d'états (I) est dit invariant par rapport au temps si les matrices A_t, B_t, C_t et D_t ne dépendent pas de t .

➤ Présentation du filtre de Kalman :

Le filtre de Kalman est un estimateur récursif optimal. Son origine se trouve dans le problème de l'estimation de la trajectoire des satellites et des aéronefs. Dans un tel problème, on cherche à estimer la position (x, y, z) et la vitesse (v_x, v_y, v_z) à partir de l'observation de deux angles : site et azimut (θ, φ) .

Dans la terminologie de Kalman : la position et la vitesse forment le vecteur d'état.

$$\vec{X} = \begin{pmatrix} x \\ y \\ z \\ v_x \\ v_y \\ v_z \end{pmatrix} \quad \text{et} \quad Y = \begin{pmatrix} \theta \\ \varphi \end{pmatrix}$$

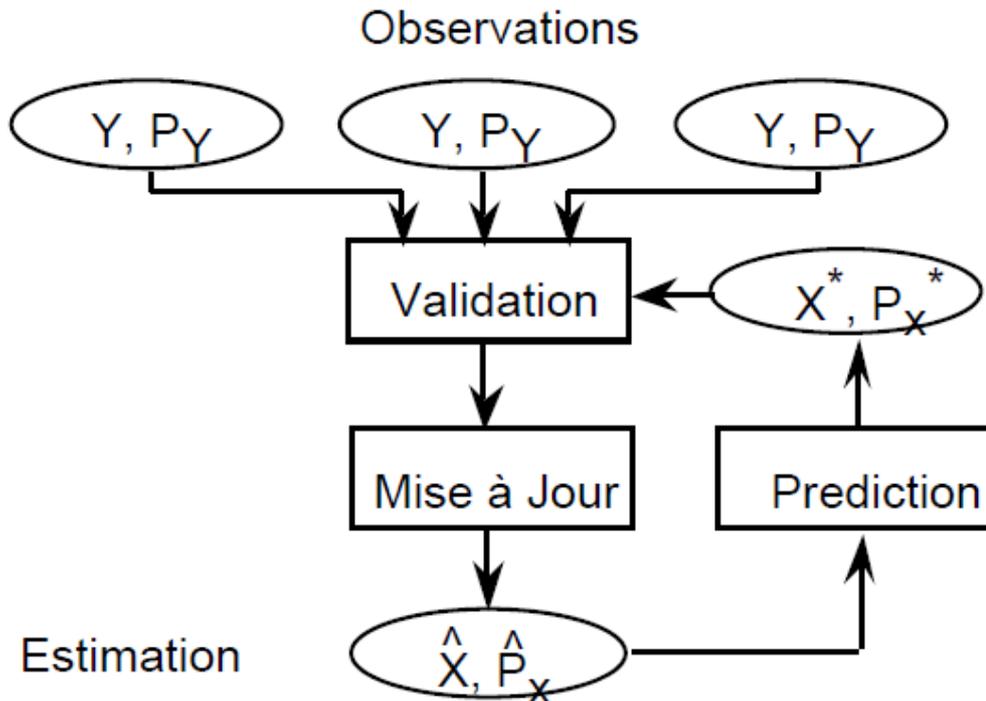
Pour chaque instant t , le filtre donne une estimation \hat{X}_t ainsi que sa précision P_t , à partir de l'observation et l'estimation précédente, \hat{X}_{t-1}, P_{t-1} et l'observation \vec{Y} , et sa précision P_y . Les précisions sont des covariances.

$$\hat{X}_t, P_t := F\{\hat{X}_{t-1}, P_{t-1}, \vec{Y}, P_y\}$$

Ceci permet d'estimer un vecteur de N variables aléatoires à partir de $M < N$ observations. Les précisions sont des 2^{ème} moments de l'erreur.

Le filtre récursif est un processus cyclique avec 5 phases :

- 1) Prédiction de l'évolution temporelle de l'état ;
- 2) Observation ;
- 3) Prédiction de l'observation ;
- 4) Validation de l'observation ;
- 5) Mise à jour de l'estimation.



➤ **Estimation des variables d'état par le filtre de Kalman :**

Afin de simplifier l'analyse, les paramètres définissant le modèle espace-état sont supposés connus (il s'agit principalement des matrices A_t , C_t , R_t , Q_t et P et du vecteur m) et le modèle est supposé dépourvu de variables exogènes (ce qui revient à supposer que B_t et D_t sont nulles.). La question consiste alors à estimer à chaque instant t les variables cachées (le vecteur d'état) conditionnellement aux variables observées jusqu'à la date t (le vecteur de mesure).

Pour calculer des estimations filtrées du vecteur d'état, l'algorithme optimal appelé filtre de Kalman est utilisé (on fait référence à deux critères d'optimalité qui s'avèrent être équivalents dans le cas gaussien : la maximisation de la vraisemblance du vecteur d'état conditionnellement au vecteur de mesure ou la minimisation des carrés des erreurs réalisées sur le vecteur d'état. Dans le cas non-gaussien, le filtre de Kalman reste uniquement optimal parmi les estimateurs linéaires.).

L'algorithme est structuré en deux étapes reprises d'itération en itération. Les deux premières équations (1 et 2) sont des équations de « mises à jour des mesures » (actualisation) et les deux suivantes (3 et 4) de « mise à jour du temps » (prévision). La première étape concerne les lois de probabilité a posteriori qui tiennent compte de l'information à la date t . La seconde étape, à la différence de la première, ne dépend pas des observations à la date t : le calcul peut être fait « hors-ligne », c'est-à-dire sans utiliser les signaux Y_t . Enfin, la dernière équation (5) actualise la matrice de gain K_t qui intervient dans les équations précédentes. Chaque itération se résume par les cinq équations suivantes :

$$(1) Z_{t,t}^* = Z_{t-1,t}^* + K_t(Y_t - C_t Z_{t-1,t}^*)$$

$$\begin{aligned}
(2) \Sigma_{t,t} &= (I - K_t C_t) \Sigma_{t-1,t} \\
(3) Z_{t,t}^* &= A_t Z_{t,t}^* \\
(4) \Sigma_{t,t+1} &= A_t \Sigma_{t,t} A_t' + Q_t \\
(5) K_t &= \Sigma_{t-1,t} C_t' (C_t \Sigma_{t-1,t} C_t' + R_t)^{-1}
\end{aligned}$$

et par l'initialisation : $Z_{-1,0}^* = m, \Sigma_{-1,0} = P$

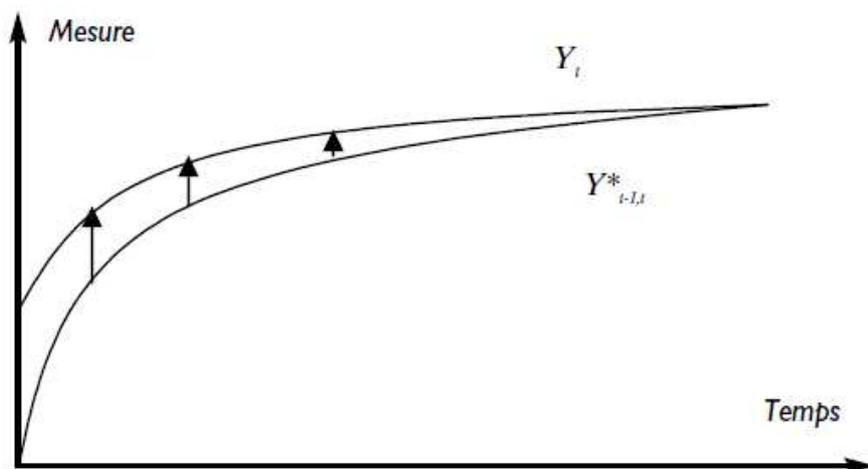
Où :

- $Z_{t,t}^*$ est l'estimation courante du vecteur d'état.
- $\Sigma_{t,t} = V(Z_{t,t} - Z_{t,t}^*)$ est l'erreur quadratique moyenne sur Z_t ,
- $Z_{t-1,t}^*$ est la prévision du vecteur d'état faite à la date t-1 ;
- $\Sigma_{t-1,t} = V(Z_{t-1,t} - Z_{t-1,t}^*)$ est l'erreur quadratique moyenne de prévision correspondante
- K_t est la matrice de gain de Kalman.

➤ Interprétations graphiques du filtre de Kalman :

Si le filtre de Kalman vise d'abord à estimer la variable cachée Z_t , il permet également de réaliser une prévision de Y_t , connaissant son passé. Il suffit pour cela de calculer $Y_{t-1,t}^* = C_t Z_{t-1,t}^*$. En raisonnant sur la variable d'observation, le filtre calcule en fait une prévision $Y_{t-1,t}^*$ qui est ensuite corrigée par comparaison avec la mesure réelle Y_t . Le filtre fait donc dévier progressivement la trajectoire théorique $Y_{t-1,t}^*$ (engendrée par les conditions initiales et l'équation d'état) vers la trajectoire réelle Y_t , comme le montre le schéma 1. En raisonnant sur l'état Z_t , son estimation actualisée $Z_{t,t}^*$ est décomposable entre la prévision $Z_{t-1,t}^*$ et un terme correctif proportionnel à l'erreur relative aux dernières observations Y_t .

Schéma 1 : Trajectoire théorique et trajectoire réelle



➤ Estimation des paramètres par la méthode de maximum de vraisemblance :

En pratique, les matrices A_t , C_t , Q_t , R_t et P (introduites dans un modèle espace-état) sont inconnues et doivent être estimées. Un algorithme itératif a été introduit par Dempster et al. (1977) pour estimer le maximum de vraisemblance des modèles stochastiques à variables cachées.

Pour procéder à une estimation par maximum de vraisemblance des paramètres d'un modèle espace-état, il est nécessaire d'avoir l'expression de la fonction de vraisemblance. Pour chaque jeu de paramètres θ , la log-vraisemblance associée à un échantillon Y_1, \dots, Y_T d'un modèle espace-état s'exprime à partir des valeurs prévues de l'état $Z_{t-1,t}^*$ et des matrices de covariance associées $\Sigma_{t-1,t}$:

$$\ln l(Y_{0:T}; \theta) = cte - \frac{1}{2} \sum_{t=0}^T \ln \det M_{t-1,t}(\theta) - \frac{1}{2} \sum_{t=0}^T \tilde{Y}'_{t-1,t}(\theta) M_{t-1,t}^{-1}(\theta) \tilde{Y}_{t-1,t}(\theta)$$

Avec :

$$\tilde{Y}_{t-1,t}(\theta) = Y_t - C_t Z_{t-1,t}^*$$

$$\text{et } M_{t-1,t}(\theta) = C_t \Sigma_{t-1,t} C_t' + R_t$$

Cet algorithme itératif génère une séquence d'estimations $(\theta_i)_{i=1,2,\dots}$ à partir d'une condition initiale θ_0 .

Chaque itération se décompose en deux étapes qui s'écrivent :

- ✓ Étape E : $\ln l(Y_{0:T}; \theta)$ se déduit de $Z_{t-1,t}^*(\theta_i)$ et de $\Sigma_{t-1,t}(\theta_i)$, calculés par un filtre de Kalman.
- ✓ Étape M: la maximisation de $\ln l(Y_{0:T}; \theta)$ par rapport à θ conduit à Σ_{i+1} .

➤ Décomposition tendance/cycle par le moyen d'un modèle espace-état :

Dans le modèle proposé par Harvey, une série Y_t est décomposée, de manière additive, en une tendance T_t , un cycle C_t et une composante irrégulière ε_t . Chacune de ces composantes est stochastique et elles sont supposées mutuellement non corrélées entre elles. La tendance suit un modèle localement linéaire. La composante cyclique est un processus linéaire stationnaire susceptible de faire apparaître une alternance relativement régulière de pics et de creux, tout en admettant une certaine persistance des phases du cycle et d'éventuelles dissymétries entre elles. Sa modélisation requiert l'introduction d'une variable C_t^* qui n'a pas d'interprétation particulière. La composante irrégulière ε_t est un bruit blanc.

$$\left\{ \begin{array}{l} Y_t = T_t + C_t + \varepsilon_t \\ T_t = T_{t-1} + \beta_{t-1} + \eta_t \\ \beta_t = \beta_{t-1} + \varepsilon_t \\ C_t = \rho (\cos \lambda \cdot C_{t-1} + \sin \lambda \cdot C_{t-1}^*) + \kappa_t \\ C_t^* = \rho (-\sin \lambda \cdot C_{t-1} + \cos \lambda \cdot C_{t-1}^*) + \kappa_t^* \end{array} \right.$$

Où $\varepsilon_t, \eta_t, \kappa_t$ et κ_t^* sont des bruits blancs indépendants, respectivement de variance σ_η^2 et $\sigma_\zeta^2, \sigma_\kappa^2$ et à nouveau σ_κ^2 . Le paramètre ρ , compris entre 0 et 1, est un paramètre de lissage : plus il est proche de 0, plus le cycle est irrégulier ; plus il est proche de 1, plus le cycle est lisse.

Pour estimer les paramètres $(\sigma_\varepsilon^2, \sigma_\eta^2, \sigma_\zeta^2, \sigma_\kappa^2, \lambda, \rho)$, la tendance et le cycle, un modèle espace-état est formulé, avec l'équation d'observation :

$$Y_t = (1 \ 0 \ 1 \ 0)X_t + \varepsilon_t$$

Où X_t est le vecteur d'état défini par $X_t = (T_t \ \beta_t \ C_t \ C_t^*)$ qui suit l'équation d'état :

$$X_t = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \rho \cos \lambda & \rho \sin \lambda \\ 0 & 0 & -\rho \sin \lambda & \rho \cos \lambda \end{pmatrix} X_{t-1} + \begin{pmatrix} \eta_t \\ \zeta_t \\ \kappa_t \\ \kappa_t^* \end{pmatrix}$$

➤ Présentation de l'analyse factorielle statique dans un cadre dynamique :

L'analyse factorielle « classique » a pour but d'offrir une description parcimonieuse d'un ensemble d'observations. Elle tente de représenter les variables étudiées dans un cadre linéaire, en fonction d'un ensemble de variables latentes, appelées facteurs, ou facteurs communs. C'est une technique qui est appropriée lorsqu'un petit nombre de facteurs peut rendre compte d'une part importante de l'information contenue dans l'ensemble des variables initiales.

Cette méthode a été initialement créée pour l'étude de données individuelles. Elle repose sur la modélisation suivante : on note n le nombre de variables étudiées, T le nombre d'observations dont on dispose pour chaque variable, et y_{it} la valeur de la t -ième observation de la variable y_i ; le modèle décrivant les variables y_1, \dots, y_n en fonction de p facteurs communs F_1, \dots, F_p ; $p < n$, s'écrit :

$$y_{it} = m_i + \lambda_{i1}F_{1t} + \dots + \lambda_{ip}F_{pt} + u_{it} \text{ pour } i = 1 \text{ à } n \text{ et pour tout } t.$$

On suppose que les $(u_{it})_{t \in T}$ sont indépendants entre eux et indépendants des facteurs :

$$\forall i \neq j \forall (t, \tau) E(u_{it}u_{j\tau}) = 0 \text{ et } \forall i \neq j \forall (t, \tau) E(u_{it}F_{j\tau}) = 0$$

Comme le modèle est destiné à l'étude de données individuelles, on suppose en outre que :

$$\forall t \neq \tau \forall (t, \tau) E(u_{it}u_{i\tau}) = 0 \text{ et } E(F_{it}F_{j\tau}) = 0$$

Un tel modèle n'est donc applicable à l'étude de données temporelles (y_{it}) que dans le cas très restrictif où tous les processus (F_{jt}) et (u_{it}) sont supposés sans autocorrélation (c'est-à-dire sont des bruits blancs au sens faible).

En utilisant des notations matricielles :

$$y_t = (y_{1t}, \dots, y_{nt})', F_t = (F_{1t}, \dots, F_{pt})', \\ u_t = (u_{1t}, \dots, u_{nt})', \Lambda = (\lambda_{ij})_{1 \leq i \leq n, 1 \leq j \leq p}$$

Le modèle s'écrit de façon plus concise :

$$y_t = m + \Lambda F_t + u_t$$

Avec : $EF_t = 0, Eu_t = 0, E(u_t u_t') = D = \text{diag}(d_1, \dots, d_n),$

$$\forall(t, \tau) E(F_t u_\tau') = 0,$$

$$\forall(t, \tau), t \neq \tau, E(F_t F_\tau') = 0,$$

$$\forall(t, \tau), t \neq \tau, E(u_t u_\tau') = 0$$

Un tel modèle n'a d'intérêt que lorsque les variables présentent une forte corrélation car il permet dans ce cas d'analyser la structure de cette corrélation. Plus précisément, les facteurs rendent compte des corrélations entre variables, alors que chaque u_{it} représente une source de variation affectant la seule variable y_{it} : chaque u_{it} est appelé composante spécifique de la variable y_{it} . Les λ_{ij} sont appelés les pondérations des facteurs : chaque λ_{ij} représente la contribution du facteur F_j au comportement de la variable y_{it} .

Il existe deux types de méthode pour estimer le modèle présenté ci-dessus : l'analyse factorielle principale et la méthode du maximum de vraisemblance sous hypothèse de normalité. La première ne nécessite aucune connaissance sur le nombre de facteurs à retenir, alors que ce nombre doit être donné a priori pour mettre en œuvre la seconde. En revanche, la méthode de maximum de vraisemblance fournit des estimateurs asymptotiquement efficaces des paramètres, ce qui n'est pas le cas de l'analyse factorielle principale.

En pratique, les deux méthodes sont utilisées. Lors d'une première étape, l'analyse factorielle principale fournit un critère permettant de choisir un nombre de facteurs qui semble pertinent. En fait, sa mise en œuvre revient à effectuer une analyse en composantes principales, mais cette ACP porte non pas sur la matrice de corrélation mais sur la matrice dite de corrélation réduite : les termes hors diagonale de cette matrice sont égaux aux corrélations entre variables, mais les éléments diagonaux sont fixés, en première approximation, aux valeurs des corrélations canoniques de chaque variable avec l'ensemble des autres. On est alors conduit à calculer les valeurs propres de cette matrice, et le nombre de facteurs à retenir est choisi en fonction de la taille de ces valeurs propres. Dans une deuxième étape, il est possible de mettre en œuvre la méthode de maximum de vraisemblance. En outre, un test du rapport de vraisemblance permet, dans ce cadre statique, de contrôler que le nombre de facteurs retenu est correct.

Partie III : Application aux données de l'enquête Conjoncture

➤ Présentation de l'enquête conjoncture et des données utilisées dans le cadre de l'étude :

A-Présentation de l'enquête conjoncture

Le Haut Commissariat au Plan réalise deux enquêtes de conjoncture. La première concerne les entreprises industrielles, énergétiques et minières, et la deuxième cible celles du bâtiment et travaux publics. Le but de ces enquêtes est d'appréhender, à travers les opinions des

responsables d'entreprises, les tendances d'évolution d'un ensemble d'indicateurs, à court terme, dont notamment ceux relatifs à la production, l'utilisation des capacités productives, les investissements, les carnets de commandes et la main-d'œuvre.

L'enquête réalisée trimestriellement par le Haut Commissariat au Plan couvre les grandes entreprises relevant des secteurs de l'Industrie de transformation, de l'Energie, des Mines et du Bâtiment et Travaux Publics. Elle porte sur un échantillon représentatif de l'ensemble des branches et des sous-branches des domaines concernés. Son objectif est de saisir trimestriellement, les tendances de la conjoncture à travers les opinions émises par les chefs d'entreprises concernant les aspects suivants :

- ✓ la production ;
- ✓ la demande ;
- ✓ l'utilisation des capacités productives;
- ✓ les investissements ;
- ✓ les prix de vente ;
- ✓ la trésorerie ;
- ✓ les carnets de commandes ;
- ✓ la main-d'œuvre ;
- ✓ l'environnement économique.

Les autres enquêtes de conjoncture réalisées périodiquement par les différents départements sectoriels sont effectuées auprès d'un panel d'entreprises relevant uniquement de certains secteurs tels que celui de l'Industrie de transformation (enquête semestrielle du département du Commerce et de l'Industrie et enquête mensuelle de Bank Al-Maghrib) ou celui du Bâtiment et Travaux Publics (enquête semestrielle du département de l'équipement).

B-Présentation des données utilisées

Les données utilisées représentent les soldes d'opinion déduits des réponses aux questions, concernant l'industrie de transformation, de l'enquête de conjoncture sur la période allant du premier trimestre 1997 au quatrième trimestre 2009.

Les variables pertinentes retenues sont présentées dans le tableau ci-dessous :

Tableau1 : Liste des variables sélectionnées pour la construction de l'indicateur synthétique de conjoncture industrielle

N° d'ordre	Variables	libellés
1	Prod_obs	Solde d'opinion associé à la production observée
2	Prod_prévue	Solde d'opinion associé à la production prévue
3	S_p_f	Solde d'opinion associé au stock en produits fabriqués
4	D_g	Solde d'opinion associé à la demande globale
5	V_m_loc	Solde d'opinion associé aux ventes sur le marché local
6	V_m_ext	Solde d'opinion associé aux ventes sur le marché extérieur

➤ **Mise en pratique de la méthode :**

A- Analyse préliminaire

✓ **L'indice d'adéquation de Kaiser-Meyer-Olkin (KMO)**

L'indice KMO permet de mesurer l'adéquation des variables utilisées dans une analyse factorielle. Il est calculé pour l'ensemble des variables (KMO) et pour chaque variable prise individuellement (MSA : Kaiser's Measure of Sampling Adequacy). L'indice d'adéquation permet de juger de la pertinence ou non de recourir à l'analyse en composante principale. C'est un indicateur résumé qui permet de vérifier à la fois pour l'ensemble des variables et pour chaque variable prise individuellement, si les corrélations originales sont supérieures aux corrélations partielles. Ainsi, une variable qui ne serait corrélée à aucune autre devrait certainement être retirée de l'analyse, puisque nous nous intéressons à la variance commune partagée entre les variables. Les valeurs de KMO et des MSA supérieures à 0,8 sont jugées bonnes alors que celles qui sont inférieures à 0,5 sont inacceptables.

En ce qui concerne le cas de notre étude, l'indice KMO (Kaiser-Meyer-Olkin) vaut 0,7 ; cette mesure supérieure à 0,5 confirme l'acceptation de l'échantillon de variables obtenues.

Tableau2 :L'indice MSA des variables retenues

N°d'ordre	VARIABLES	LIBELLES	MSA
1	Prod_obs	Solde d'opinion associé à la production observée	0,7
2	Prod_prévue	Solde d'opinion associé à la production prévue	0,7
3	S_p_f	Solde d'opinion associé au stock en produits fabriqués	0,5
4	D_g	Solde d'opinion associé à la demande globale	0,7
5	V_m_loc	Solde d'opinion associé aux ventes sur le marché local	0,7
6	V_m_ext	Solde d'opinion associé aux ventes sur le marché extérieur	0,6
Ensemble des variables		Mesure de l'échantillonnage de Kaiser-Meyer-Olkin (KMO)=0,7	

Les valeurs des MSA confirment que les variables sont corrélées entre elles. Elles seront retenues dans la suite de l'analyse.

✓ **Stationnarité des séries retenues**

Les variables utilisées dans notre étude sont des suites d'observations indexées par le temps. Il est important, pour la suite de l'analyse, de savoir si les valeurs prises par les variables à chaque date sont corrélées entre elles. D'où l'étude de la stationnarité de séries retenues.

Le test de stationnarité mis en œuvre est le test PP (Philips Perron). Ce test permet de tester l'hypothèse H_0 : la série est intégrée d'ordre au moins 1 c'est-à-dire qu'elle n'est pas stationnaire ; contre l'hypothèse alternative H_1 : la série est stationnaire. Les résultats de ce test sont présentés dans le tableau ci-dessous.

Tableau 3 : Résultat du test PP sur les variables retenues

Variables	PP _c	PP _{tabulé}	Conclusion
Prod_obs	-10,40	-1,94	Stationnaire
Prod_prévue	-12,27	-1,94	Stationnaire
S_p_f	-5,62	-1,94	Stationnaire
D_g	-9,60	-1,94	Stationnaire
V_m_loc	-7,37	-1,94	Stationnaire
V_m_ext	-7,97	-3,49	Stationnaire

Les résultats du tableau ci-dessus montrent que toutes les variables retenues sont stationnaires au seuil de 5% et de 10%.

B- Application de l'analyse factorielle

Le tableau ci-dessous présente la proportion de variance commune de chaque variable.

Tableau 4 : Variance totale et commune des variables retenues

Variables	Variance totale	Variance commune
Prod_obs	1,000	0,841
Prod_prévue	1,000	0,324
S_p_f	1,000	0,771
D_g	1,000	0,748
V_m_loc	1,000	0,616
V_m_ext	1,000	0,643

Les valeurs 0,841 ; 0,324 ; 0,771 ; 0,748 ; 0,616 et 0,643 correspondent à la proportion de la variance totale expliquée par chaque variable.

Autrement dit, 0,841 unité de variance de la variable **Prod_obs** (Production observée) correspond à la variance importante partagée par les autres variables, alors que 0,159 unités, c'est-à-dire : $1 - 0,841$ serait de la variance négligeable, puisque correspondant soit à la variance spécifique de cette variable, soit à son erreur de mesure.

La somme des variances communes donne un total de 3,943 correspondant à 65,716 % de la variance totale. C'est cette proportion de variance commune qui sera éventuellement répartie de façon décroissante entre les différents facteurs retenus.

Le tableau ci-après présente les valeurs propres de la matrice de corrélation réduite obtenue.

Tableau 5 : valeurs propres des facteurs

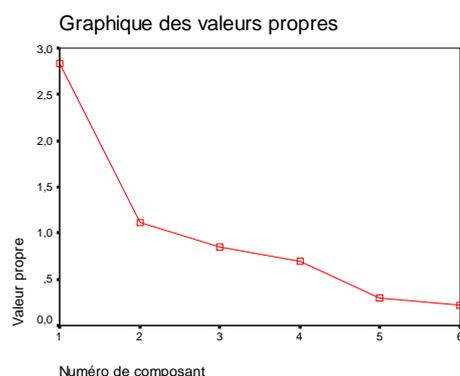
Composantes	Valeurs propres initiales			Extraction somme des carrés des facteurs retenus		
	Total	% de la variance	% cumulé	Total	% de la variance	%cumulé
1	2,83	47,22	47,22	2,83	47,22	47,22
2	1,11	18,52	65,74	1,11	18,52	65,74
3	0,85	14,20	79,95	0,85	14,20	79,95
4	0,69	11,51	91,46	0,69	11,51	91,46
5	0,29	4,94	96,41	0,29	4,94	96,41
6	0,21	3,58	100	0,21	3,58	100

Méthode d'extraction : Analyse des principaux composants

La détermination du nombre de facteurs communs à extraire se fait en considérant le critère de Kaiser indiquant la limite inférieure du nombre de facteurs à extraire. Selon ce critère, seul les facteurs ayant une valeur propre supérieure à 1 sont à retenir. Selon le tableau ci-dessus, le premier et le deuxième axes factoriels sont à retenir ; car ayant leurs valeurs propres égales à 2,83 et 1,11 respectivement. Ainsi, les deux premiers facteurs suffisent pour expliquer le nuage des points étudié, puisqu'ils expliquent plus de 65% de l'information contenue dans ces variables, tandis que si nous tenons compte des autres facteurs nous n'aurons presque aucune information supplémentaire car ils n'expliquent que 34,26% de la variance totale. Nous sommes amenés donc à retenir ces deux premiers facteurs.

Par ailleurs, le graphe des valeurs propres justifie bien le nombre de facteurs retenus, puisque à partir de la deuxième valeur propre, nous constatons que des faibles variations :

Figure2 : Graphes des valeurs propres



Nous utilisons la méthode de maximum de vraisemblance pour tester si le nombre de facteurs retenu est correct ou non. Les valeurs propres associées aux deux facteurs retenus et extraits selon la méthode de maximum de vraisemblance sont présentées dans le tableau ci-dessous :

Tableau 6 : Les valeurs propres à partir de la MMV

Facteurs	Valeurs propres initiales	Somme des carrés chargés	
	Total	% de la variance	% cumulé
1	2,8333294	26,3883905	26,3883905
2	1,11136475	28,8972014	55,2855919

Méthode d'extraction : Maximum de vraisemblance

La méthode de maximum de vraisemblance confirme le choix des deux facteurs.

C- Construction et pertinence de l'indicateur synthétique

L'objectif de l'étude est de chercher la variable non observée qui permet de résumer l'information contenue dans les différentes variables. Par la suite, nous allons tester l'hypothèse si nous pouvons retenir un seul facteur.

- **Le cas statique**

Le tableau suivant présente les pondérations du facteur retenu :

Tableau 7: Les pondérations du facteur retenu par la MMV

Variables	Facteur
	1
Prod_obs	0,50379695
Prod_prévue	-0,16379627
S_p_f	0,10171096
D_g	0,51267046
V_m_loc	0,17464066
V_m_ext	0,99949355

Le test de khi-deux permet de tester l'hypothèse de garder un seul facteur (l'hypothèse nulle) :

Tableau 8: Le test de Khi-deux

Khi-deux	ddl	Qualité d'ajustement de la signification
6,49180321	4	0,16530758

D'après le tableau ci-dessus, la valeur observée de la loi Khi-deux est environ 6,49. Or pour un niveau de signification de 5 % et un ddl de 4, la table de la loi de Khi-deux affiche une valeur de 0,711.

Puisque **6,49 > 0,711**, on rejette alors le nombre de facteurs retenus (l'hypothèse nulle).

- **Le cas dynamique**

Nous testons si nous pouvons garder un seul facteur tout en tenons compte de la dynamique des variables.

Nous fixons maintenant le nombre de facteurs à 1.

Pour faire ce test nous utilisons la statistique du test de pseudo score (qui est identique à la statistique de Hausmann Wald) :

$$\xi_s = \frac{T}{2} \text{vech}(\hat{\Sigma}_T^0 - S)' (\hat{M} D_{n_j}^+ \hat{\Omega}_0 D_n^+ \hat{M}')' \text{vech}(\hat{\Sigma}_T^0 - S)$$

Sous H_0 , on a :

$$\xi_s \xrightarrow{L} \chi^2(r) \text{ avec } r = \frac{1}{2}((n-p)^2 - (n+p))$$

L'utilisation du logiciel STATA permet de calculer la statistique de Hausmann-Wald, ainsi pour notre cas, on a : $\xi_s = -108,33$

Par ailleurs, pour un seuil de 5% et un degré de liberté égal à 9, la table de la loi de Khi-deux nous donne : $\chi_{5\%}^2(9) = 3,32$.

Donc nous acceptons, dans le cas dynamique, l'hypothèse d'avoir retenu un seul facteur.

▪ Convergence des estimateurs dans un cadre dynamique

La convergence, dans un cadre dynamique, des estimateurs obtenus dans l'analyse factorielle statique est assurée sous deux hypothèses, à savoir la stationnarité et la normalité.

La stationnarité :

Les tests de stationnarité mises en œuvre confirment la stationnarité des processus retenus (voir tableau 3) ainsi que les représentations graphiques des différents soldes d'opinion.

Figure3 : Représentation du solde d'opinion de la production observée

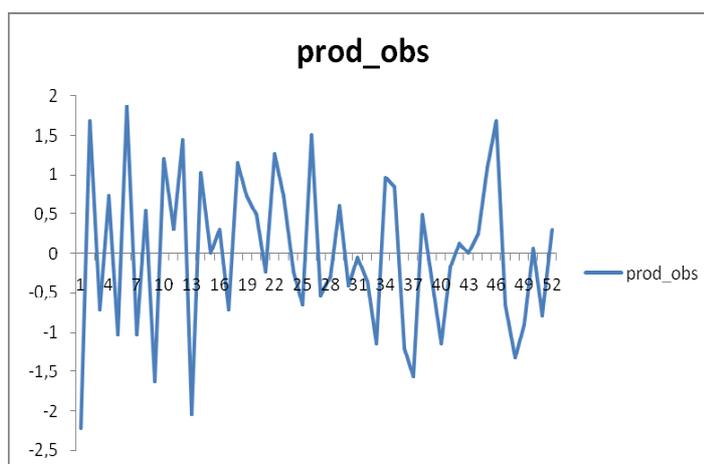


Figure4 : Représentation du solde d'opinion de la production prévue

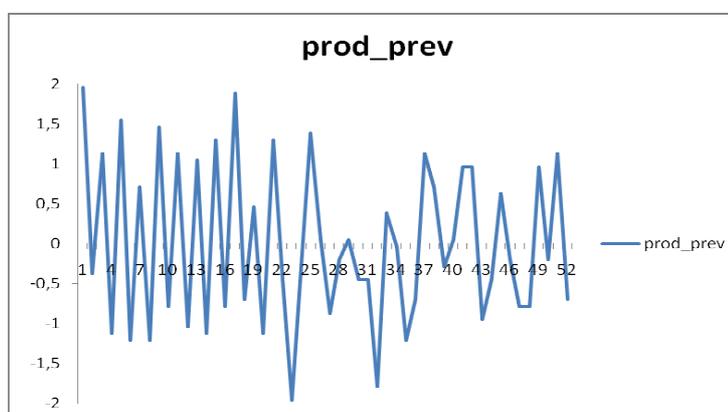


Figure5 : Représentation du solde d'opinion du stock en produits fabriqués

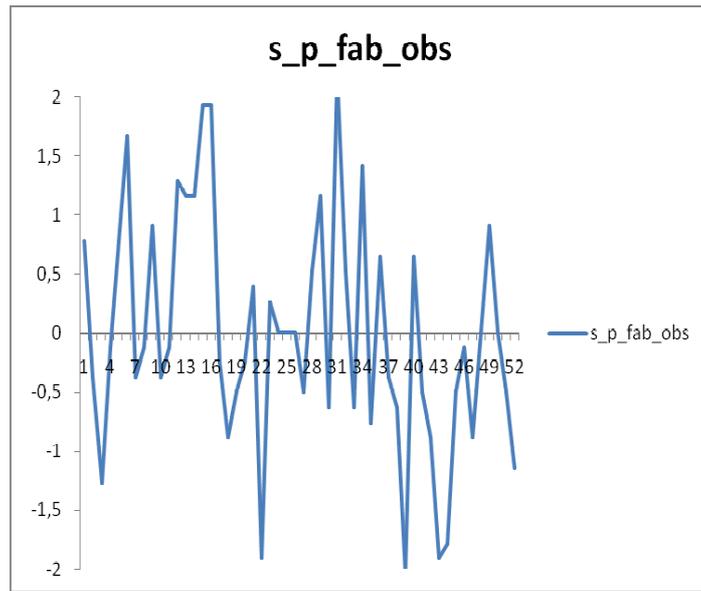


Figure6 : Représentation du solde d'opinion de la demande globale

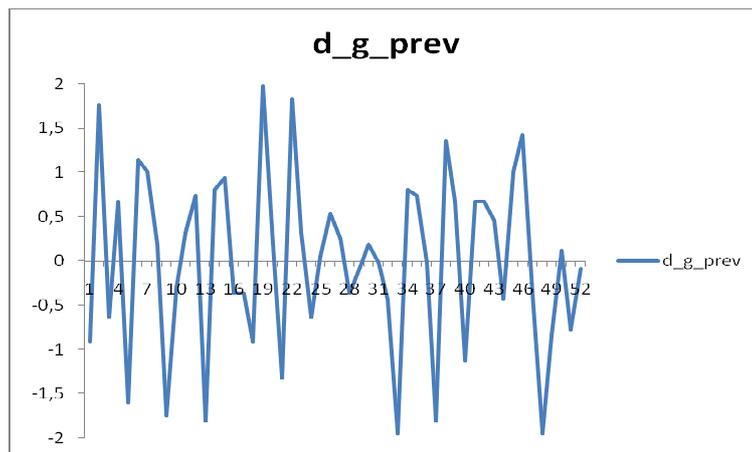


Figure7 : Représentation du solde d'opinion des ventes sur le marché local

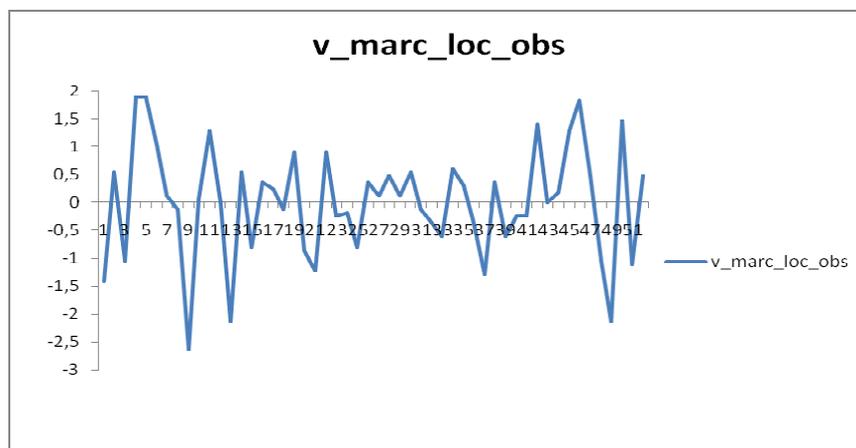
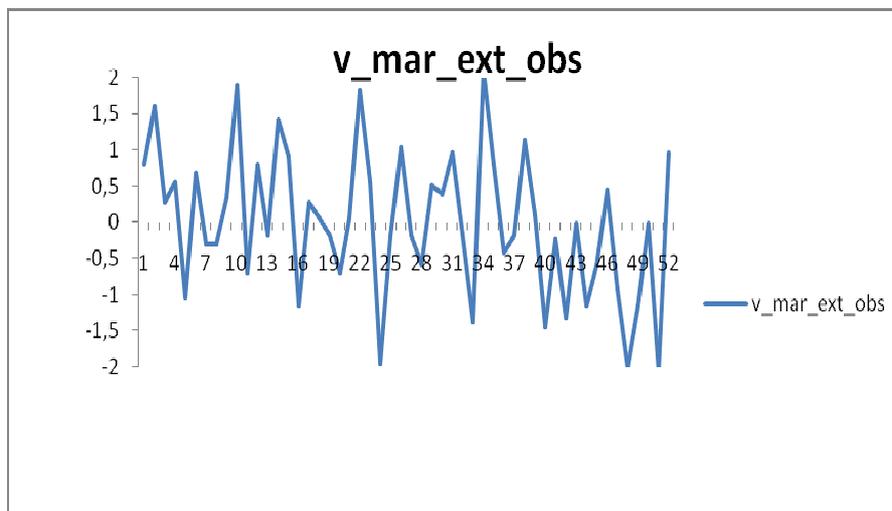


Figure8 : Représentation du solde d'opinion des ventes sur le marché extérieur



La normalité :

Nous utilisons le test non paramétrique de Kolmogorov-Smirnov pour tester la normalité des variables retenues :

Tableau 9: Test de Kolmogorov-Smirnov

		Prod_obs	Prod_prévue	S_p_f	D_g	V_m_loc	V_m_ext
N		52	52	52	52	52	52
Paramètres normaux	Moyenne	7,7013E-10	9,3132E-10	4,4775E-09	1,0746E-09	-1,0227E-08	6,3402E-09
	Ecart-type	1	1	1	1	1	1
Différences les plus extrêmes	Absolue	0,0536972	0,12027273	0,13068292	0,07460643	0,09408499	0,07652126
	Positive	0,05259891	0,11970961	0,13068292	0,06058537	0,0815334	0,05209379
	Négative	-0,0536972	-0,12027273	-0,07211008	-0,07460643	-0,09408499	-0,07652126
Z de Kolmogorov-Smirnov		0,38721603	0,86729896	0,94236791	0,53799456	0,67845654	0,55180269

La signification asymptotique de chacune des variables est supérieure à 0,05. Donc la distribution empirique correspond bien à la distribution théorique. Voire une distribution normale. L'hypothèse de normalité est bien vérifiée.

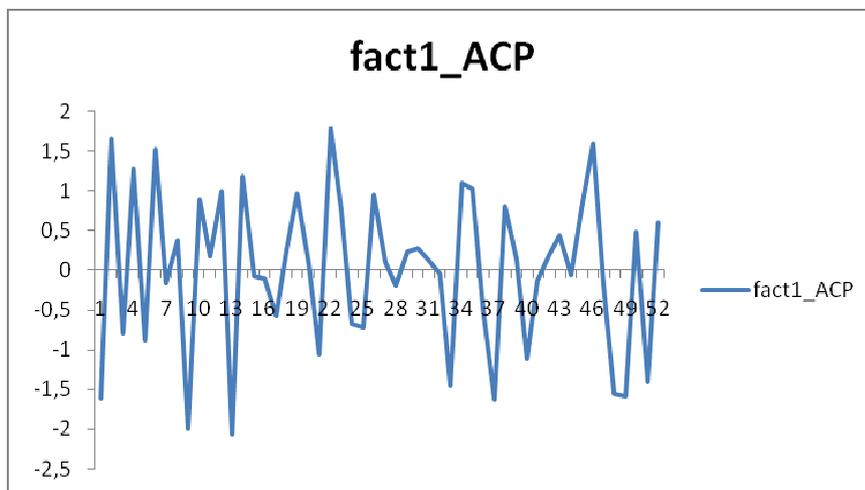
Par la vérification de ces deux hypothèses, nous pouvons dire que les estimateurs obtenus dans l'analyse factorielle statique convergent dans un cadre dynamique.

D- Modélisation du facteur commun et des composantes spécifiques

✓ Identification d'un modèle pour le facteur commun

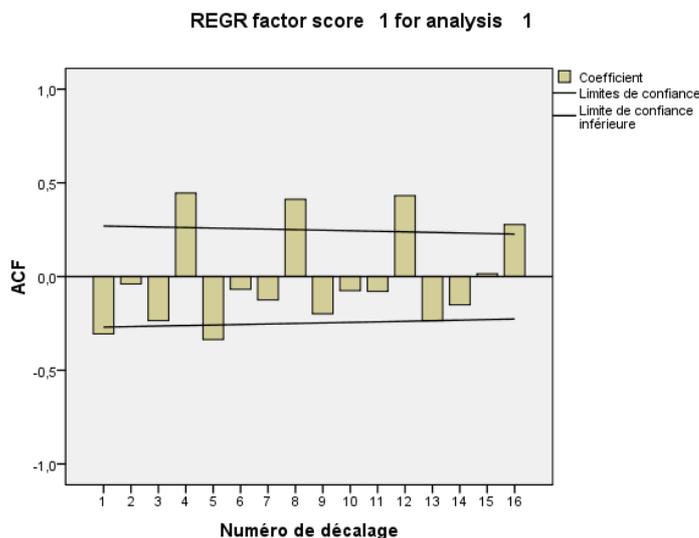
Nous représentons d'abord l'évolution du facteur commun dans le temps :

Figure9 : Evolution de la série du facteur commun dans le temps



D'après le graphe de la série facteur commun, nous constatons qu'il ya une stabilité de la variance ainsi que du niveau, donc nous ne procédons à aucune transformation de la série.

Figure10 : Evolution de la série du facteur commun dans le temps



Ajustement d'un modèle ARMA

Dans cette partie, on a opté pour la méthode du critère d'information AIC qui consiste en le choix de deux modèles purs AR(P) et MA(Q), puis l'établissement du tableau donnant la valeur AIC pour chaque combinaison (p_i , q_j) inférieurs respectivement à P et Q. Le modèle retenu sera celui qui donnera le minimum des valeurs obtenues.

- ✓ Ajustement des modèles AR(P) et MA(Q) purs :

On choisira d'après l'autocorrélogramme P=1 et Q=1.

- ✓ Choix du modèle ARMA (p,q) selon le critère AIC :

P	Q	0	1
0		*	140,90
1		143,22	141,16

D'après ce tableau, la valeur minimale du critère AIC est atteinte pour les valeurs p=0 et q=1. Le modèle à retenir est donc un MA(1).

Le modèle ajusté est :

$$F_t = a_t - 0,458 a_{t-1}$$

Avec :

F_t : La série du facteur commun.

a_t : Un bruit blanc (0, σ_a^2).

- ✓ Ajustement des modèles adéquats pour les erreurs

On a l'équation suivante :

$$Y_{it} = \lambda_i F_t + u_{it} \text{ avec } i : 1 \text{ à } 6 \text{ et } t = 1 \text{ à } 52$$

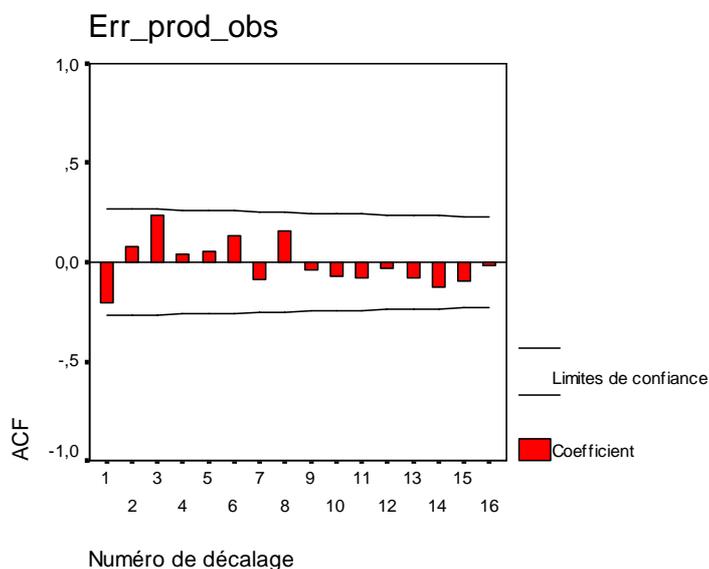
Nous procédons par la suite à la modélisation des erreurs

Figure11 : Evolution de l'erreur associée à la production observée



Le graphe montre la stationnarité du processus.

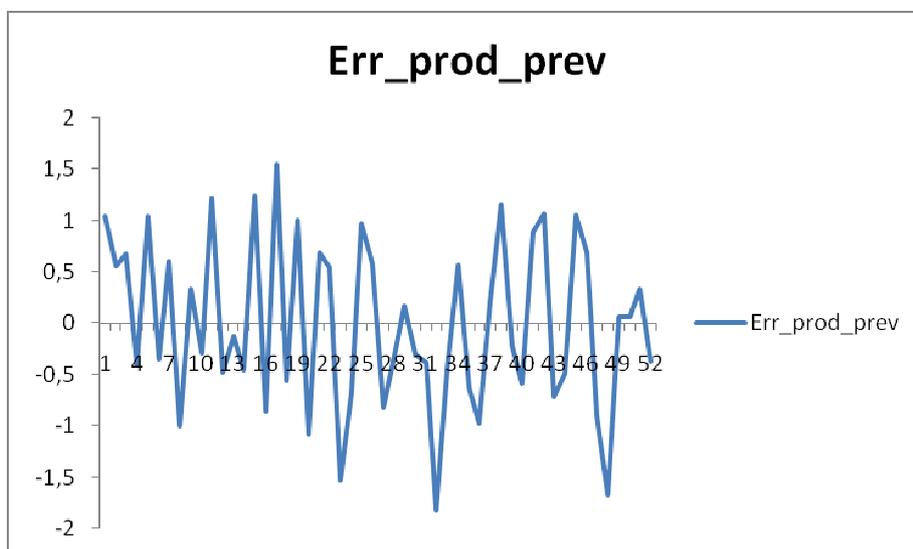
Figure12 : L'autocorrélogramme de l'erreur associée à la production observée



D'après le graphique ci-dessus, on constate qu'à partir du premier ordre, le coefficient d'autocorrélation est inférieur à 2×1 'écart type, donc le modèle qui s'ajuste le mieux à cette série est un bruit blanc.

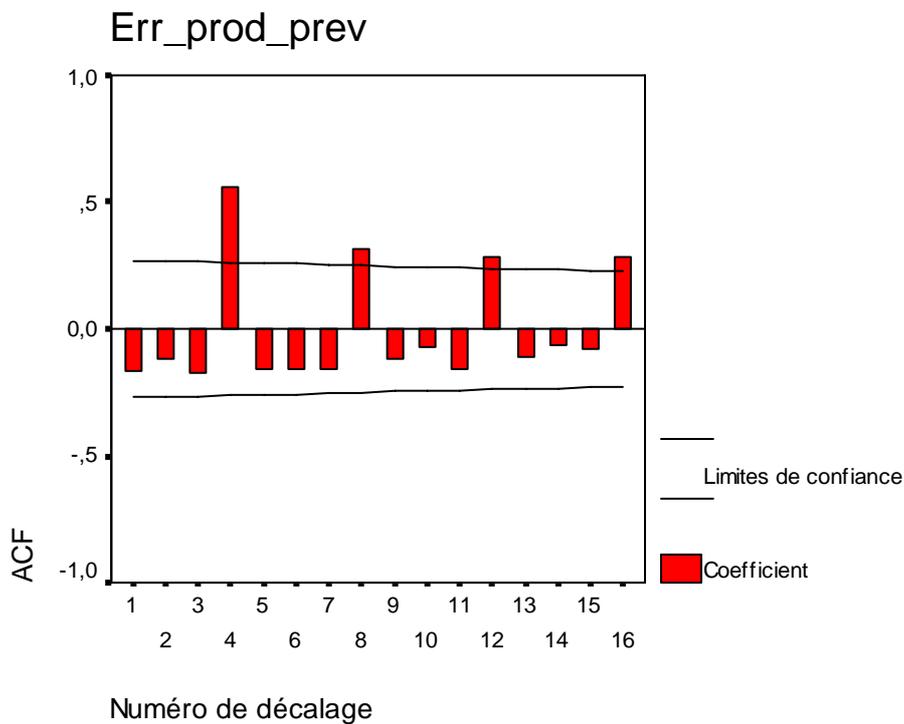
$$U_{1t} = a_{1t} \text{ avec } a_{1t} \text{ un } BB(0, \sigma_a^2)$$

Figure13 : Evolution de l'erreur associée à la production prévue



Le graphe montre la stationnarité du processus.

Figure14 : L'autocorrélogramme de l'erreur associée à la production prévue



✓ Ajustement des modèles AR(P) et MA(Q) purs :

On choisira d'après l'autocorrélogramme P=1 et Q=1.

✓ Choix du modèle ARMA (p,q) selon le critère AIC :

	Q	0	1
P	0	*	126,61
1		127,30	128,30

D'après ce tableau, la valeur minimale du critère AIC est atteinte pour les valeurs p=0 et q=1. Le modèle à retenir est donc un MA(1).

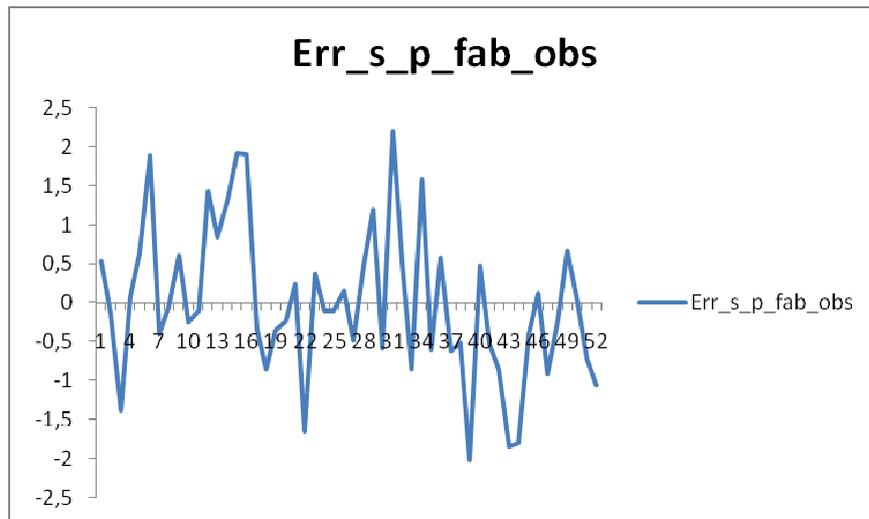
Le modèle ajusté est :

$$U_{2t} = a_{2t} - 0,243 a_{2t-1}$$

Avec :

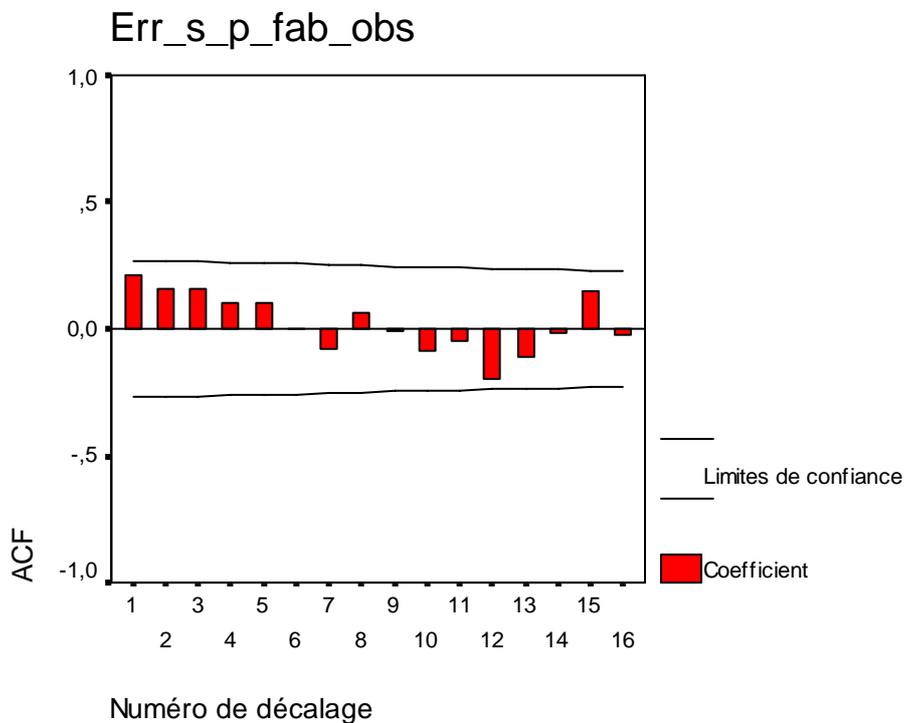
a_{2t} : Un bruit blanc (0, σ_a^2).

Figure15 : Evolution de l'erreur associée au stock en produits fabriqués



Le graphe montre la stationnarité du processus.

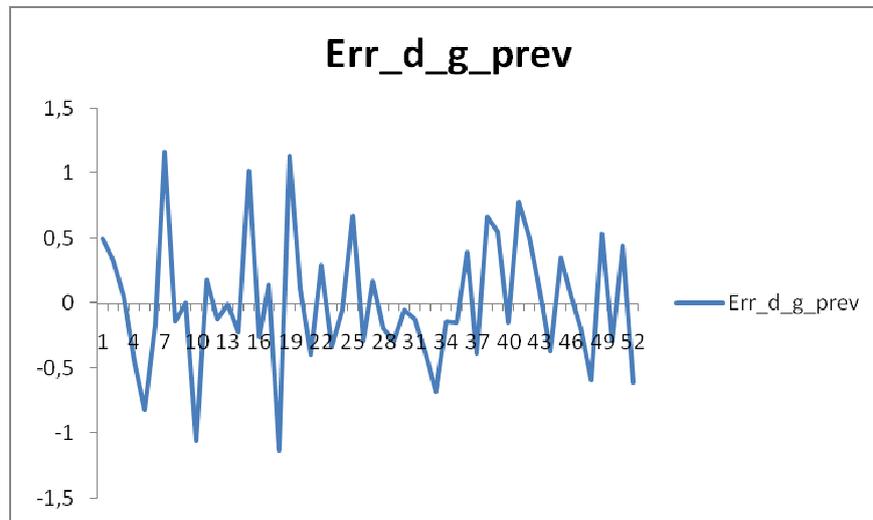
Figure16 : L'autocorrélogramme de l'erreur associée au stock en produits fabriqués



D'après le graphique ci-dessus, on constate qu'à partir du premier ordre, le coefficient d'autocorrélation est inférieur à 2*l'écart type, donc le modèle qui s'ajuste le mieux à cette série est un bruit blanc.

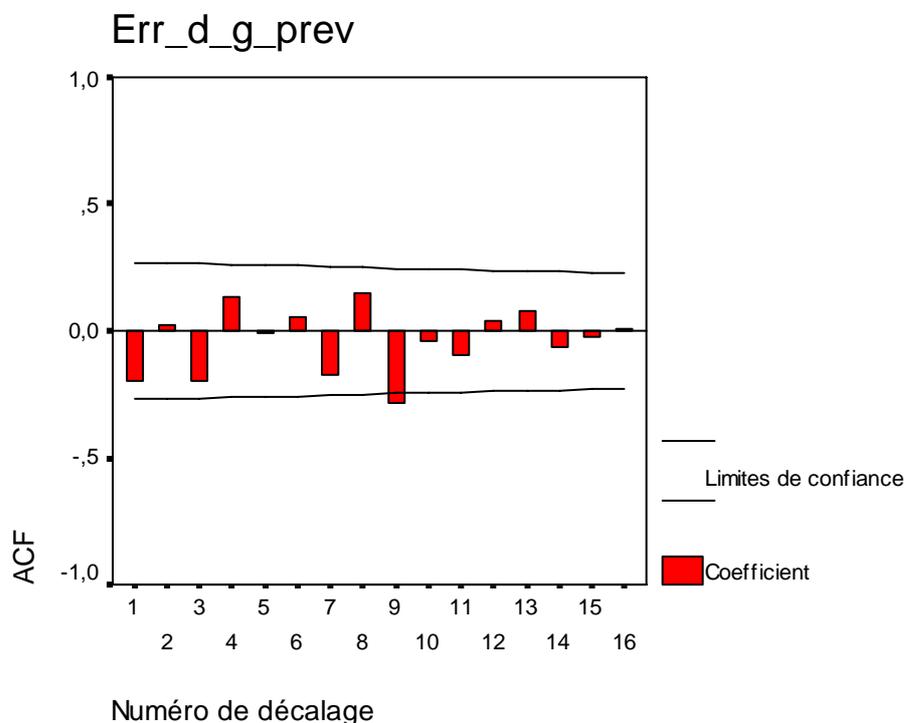
$$U_{3t} = a_{3t} \text{ avec } a_{3t} \text{ un } BB(0, \sigma_a^2)$$

Figure17 : Evolution de l'erreur associée à la demande globale



Le graphe montre la stationnarité du processus.

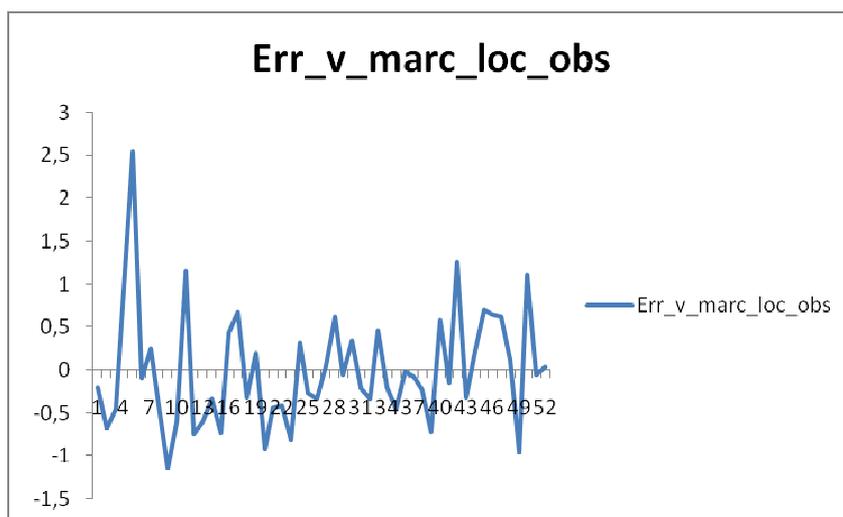
Figure18 : L'autocorrélogramme de l'erreur associée à la demande globale



D'après le graphique ci-dessus, on constate qu'à partir du premier ordre, le coefficient d'autocorrélation est inférieur à $2 \cdot 1$ l'écart type, donc le modèle qui s'ajuste le mieux à cette série est un bruit blanc.

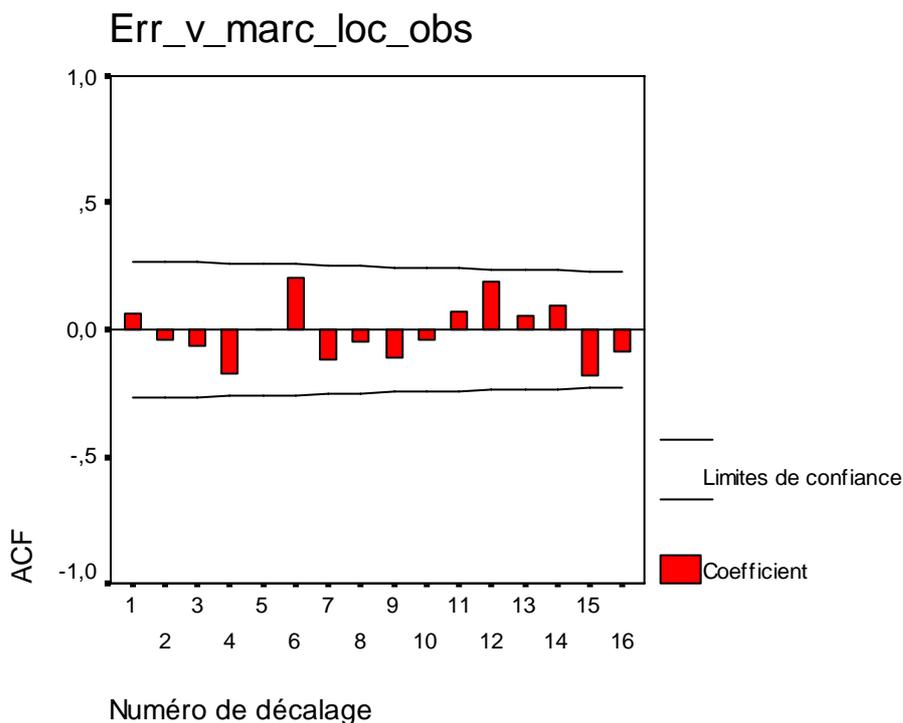
$$U_{4t} = a_{4t} \text{ avec } a_{4t} \text{ un } BB(0, \sigma_a^2)$$

Figure19 : Evolution de l'erreur associée aux ventes sur le marché local



Le graphe montre la stationnarité du processus.

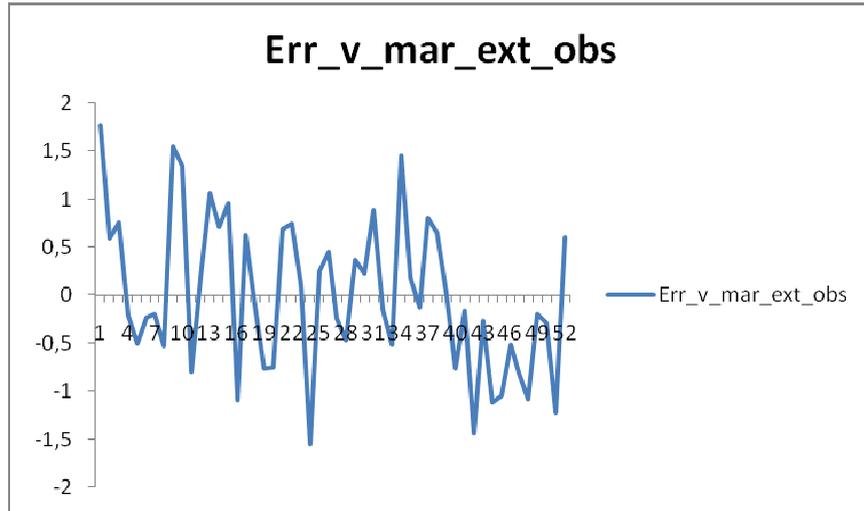
Figure20 : L'autocorrélogramme de l'erreur associée aux ventes sur le marché local



D'après le graphique ci-dessus, on constate qu'à partir du premier ordre, le coefficient d'autocorrélation est inférieur à 2*l'écart type, donc le modèle qui s'ajuste le mieux à cette série est un bruit blanc.

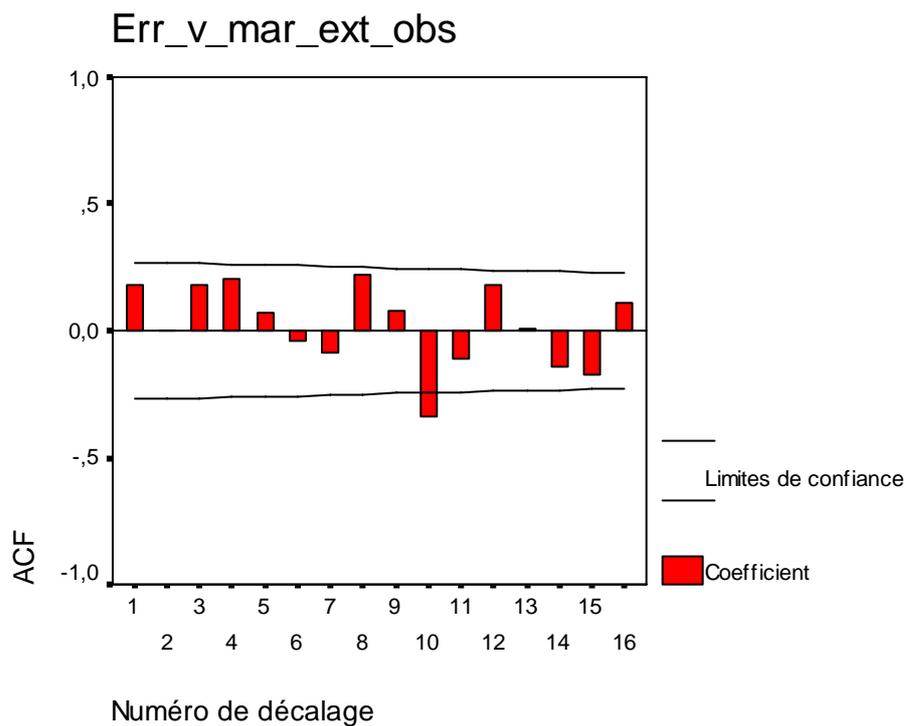
$$U_{5t} = a_{5t} \text{ avec } a_{5t} \text{ un } BB(0, \sigma_a^2)$$

Figure21 : Evolution de l'erreur associée aux ventes sur le marché extérieur



Le graphe montre la stationnarité du processus.

Figure22 : L'autocorrélogramme de l'erreur associée aux ventes sur le marché extérieur



D'après le graphique ci-dessus, on constate qu'à partir du premier ordre, le coefficient d'autocorrélation est inférieur à 2×1 l'écart type, donc le modèle qui s'ajuste le mieux à cette série est un bruit blanc.

$$U_{\epsilon t} = a_{\epsilon t} \text{ avec } a_{\epsilon t} \text{ un } BB(0, \sigma_a^2)$$

✓ **Estimation par le filtre de Kalman**

Dans cette partie, nous tenons compte de la dynamique des variables pour estimer le modèle. Nous allons garder les modèles de type ARMA obtenus dans le cas statique pour le facteur commun et les erreurs associés à chaque variable. Le modèle général à estimer est le suivant :

$$\begin{aligned}
 y_{it} &= \lambda_i F_t + u_{it} && \text{pour } i=1 \text{ à } 6 \text{ et pour tout } t, \\
 F_t &= a_t - \theta_1 a_{t-1} && \text{pour tout } t, \\
 u_{it} &= a_{it} && \text{pour } i=1,3,4,5,6 \text{ et pour tout } t, \\
 u_{2t} &= a_{2t} - \theta_2 a_{2t-1} && \text{pour tout } t.
 \end{aligned}$$

Les processus a_{it} et a_t sont supposés indépendants.

Nous pouvons réécrire le modèle général sous une forme état-mesure en posant le vecteur d'état :

$$\alpha_t = (F_t, a_t, a_{2t}, u_{1t}, u_{2t}, u_{3t}, u_{4t}, u_{5t}, u_{6t})'$$

Sous une forme matricielle nous aurons l'écriture suivante :

$$\begin{pmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \\ y_{4t} \\ y_{5t} \\ y_{6t} \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \lambda_2 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \lambda_3 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ \lambda_4 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ \lambda_5 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ \lambda_6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} F_t \\ a_t \\ a_{2t} \\ u_{1t} \\ u_{2t} \\ u_{3t} \\ u_{4t} \\ u_{5t} \\ u_{6t} \end{pmatrix}$$

Et aussi :

$$\begin{pmatrix} F_t \\ a_t \\ a_{2t} \\ u_{1t} \\ u_{2t} \\ u_{3t} \\ u_{4t} \\ u_{5t} \\ u_{6t} \end{pmatrix} = \begin{pmatrix} 0 & -\theta_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\theta_2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} F_{t-1} \\ a_{t-1} \\ a_{2t-1} \\ u_{1t-1} \\ u_{2t-1} \\ u_{3t-1} \\ u_{4t-1} \\ u_{5t-1} \\ u_{6t-1} \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} a_t \\ a_{1t} \\ a_{2t} \\ a_{3t} \\ a_{4t} \\ a_{5t} \\ a_{6t} \end{pmatrix}$$

Ainsi :

$$\alpha_t = A_t \alpha_{t-1} + B_t \eta_t$$

Avec :

$$\alpha_t = \begin{pmatrix} F_t \\ a_t \\ a_{2t} \\ u_{1t} \\ u_{2t} \\ u_{3t} \\ u_{4t} \\ u_{5t} \\ u_{6t} \end{pmatrix}; A_t = \begin{pmatrix} 0 & -\theta_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\theta_2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}; \alpha_{t-1} = \begin{pmatrix} F_{t-1} \\ a_{t-1} \\ a_{2t-1} \\ u_{1t-1} \\ u_{2t-1} \\ u_{3t-1} \\ u_{4t-1} \\ u_{5t-1} \\ u_{6t-1} \end{pmatrix};$$

$$B_t = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}; \eta_t = \begin{pmatrix} a_t \\ a_{1t} \\ a_{2t} \\ a_{3t} \\ a_{4t} \\ a_{5t} \\ a_{6t} \end{pmatrix}; C_t = \begin{pmatrix} \lambda_1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \lambda_2 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ \lambda_3 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ \lambda_4 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ \lambda_5 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ \lambda_6 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Nous aurons donc la représentation état-mesure suivante :

$$\begin{cases} \alpha_t = A_t \alpha_{t-1} + B_t \eta_t \\ y_t = C_t \alpha_t \end{cases}$$

Dans laquelle le processus (η_t) est non autocorrélé et vérifie :

$$V(\eta_t) = Q_t \text{ et } E(\eta_t) = \mathbf{0}$$

Nous utilisons par la suite une procédure du maximum de vraisemblance par filtre de Kalman pour estimer notre modèle. Pour tout ensemble fixé de valeurs des paramètres, le filtre de Kalman en tant que tel consiste à déterminer la meilleure prévision du vecteur d'état, compte tenu de l'information disponible à cette date.

La procédure du maximum de vraisemblance par filtre de Kalman a été effectuée sous Eviews

```
@state sv1 = [var = exp(c(2))]
@state sv2 = sv3(-1)
@state sv3 = [var = exp(c(3))]
@state sv4 = sv3(-1)

@signal prod_obs = c(4)*(sv1+c(1)*sv2) + [var = exp(c(5))]
@signal prod_prev = c(6)*(sv1+c(1)*sv2) + sv3 + c(7)*sv4
@signal s_p_fab_obs= c(8)*(sv1+c(1)*sv2) + [var = exp(c(9))]
@signal d_g_prev= c(10)*(sv1+c(1)*sv2) + [var = exp(c(11))]
@signal marc_loc_obs= c(12)*(sv1+c(1)*sv2) + [var = exp(c(13))]
@signal mar_ext_obs= c(14)*(sv1+c(1)*sv2) + [var = exp(c(15))]

@param c(1) -0.45 c(4) 0.91 c(6) -0.56 c(7) -0.24 c(8) -0.14 c(10) -0.86 c(12) -0.73 c(14) 0.6
```

Dans un modèle espace-état, les équations de mesure sont précédées de @signal dans le logiciel Eviews et les équations d'état sont précédées de @state. Les commandes (1-4) indiquent à Eviews que les variables dépendantes sont observées tandis que les commandes (5-10) lui signifient que les variables correspondantes sont inobservées et doivent être simulées.

Les variances sont indiquées entre crochets par la commande VAR, ce qui est une procédure Eviews dans les modèles estimés à l'aide du filtre de Kalman.

La commande @param initialise les valeurs de départ des paramètres à estimer.

Les résultats affichés se présentent comme suit :

```
@STATE SV1 = [VAR = EXP(-1.936480749)]
@STATE SV2 = SV3(-1)
@STATE SV3 = [VAR = EXP(-0.2776536041)]
@STATE SV4 = SV3(-1)
@SIGNAL PROD_OBS = 0.3650872837*(SV1-0.3383888376*SV2) + [VAR = EXP(-0.0333100125)]
@SIGNAL PROD_PREV = 0.1077113343*(SV1-0.3383888376*SV2) + SV3 - 0.3134195474*SV4
@SIGNAL S_P_FAB_OBS= 0.3344470061*(SV1-0.3383888376*SV2) + [VAR = EXP(-0.0386195103)]
@SIGNAL D_G_PREV= -1.77188819*(SV1-0.3383888376*SV2) + [VAR = EXP(-1.370257407)]
@SIGNAL MARC_LOC_OBS= -1.281507944*(SV1-0.3383888376*SV2) + [VAR = EXP(-0.5456122824)]
@SIGNAL MAR_EXT_OBS= -1.201644517*(SV1-0.3383888376*SV2) + [VAR = EXP(-0.4276393192)]
```

E- Interprétation des résultats

✓ Pertinence de l'indicateur construit et estimé par filtre de Kalman

L'indicateur construit se présente sous la forme variable cachée d'une qui retrace à chaque date la différence entre la probabilité d'une phase conjoncturelle favorable et celle d'une phase défavorable. L'indicateur évolue entre des valeurs positives et négatives ; un point extrêmement positif signale que l'activité est en période de nette accélération, par contre, un point extrêmement négatif signale que l'activité est en période de nette décélération. Les moments où l'indicateur est proche de 0 sont assimilés à des phases de stabilisation pendant lesquelles les signaux reçus sont très variés et ne font état d'aucun mouvement prononcé à la hausse ou à la baisse. Ces phases de stabilisation traduisent le retour du rythme de croissance de l'activité vers sa moyenne de long terme.

Le modèle adéquat de l'indicateur synthétique construit et estimé par filtre de Kalman est un MA(1) :

$$F_t = a_t - 0,4 a_{t-1}$$

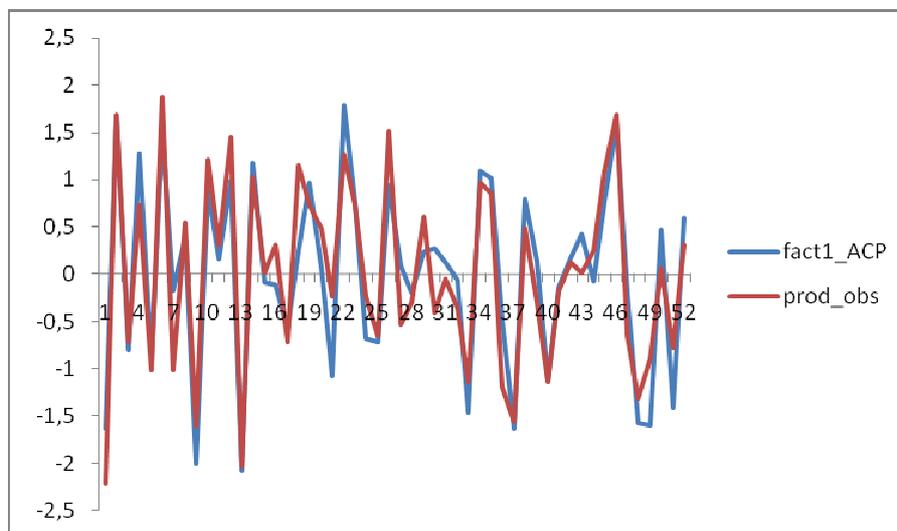
Avec :

a_t : Un bruit blanc $(0, \sigma_a^2)$.

✓ **Grille d'interprétation pour les soldes d'opinions de l'industrie de transformation**

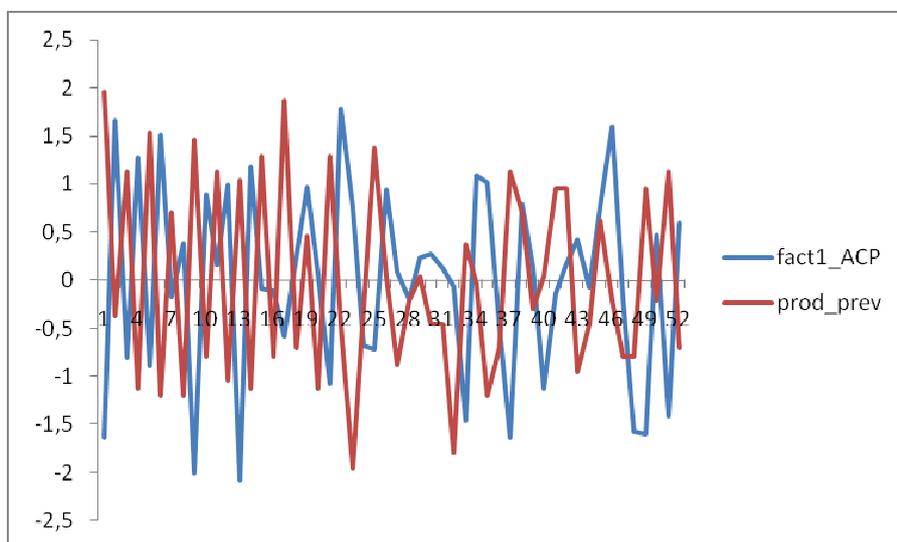
L'analyse factorielle présente l'intérêt d'isoler la part d'information spécifiquement apportée par chaque solde d'opinion. Ce faisant, elle fournit une sorte de grille de lecture complète de l'enquête.

Figure24 : Tendence de la production observée



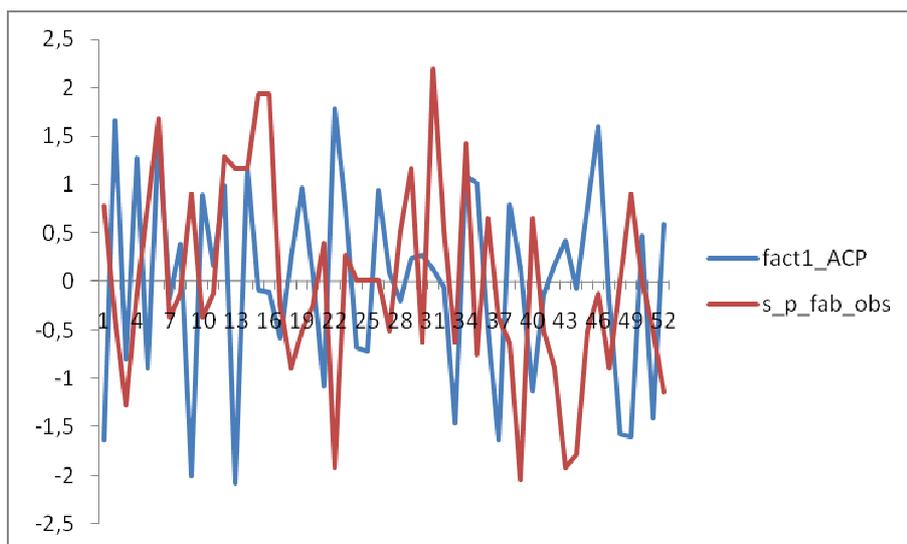
Lorsque nous comparons l'évolution du facteur commun obtenu du solde d'opinion de la production observée, nous constatons que les retournements des deux séries se produisent le plus souvent à des dates très voisines. De façon générale, l'évolution de l'indicateur synthétique sur la période de l'étude est en phase avec celle de la production observée. Ceci montre que l'évolution des perspectives de la production des entreprises industrielles permet de déduire celle de la conjoncture.

Figure25 : Tendence de la production prévue



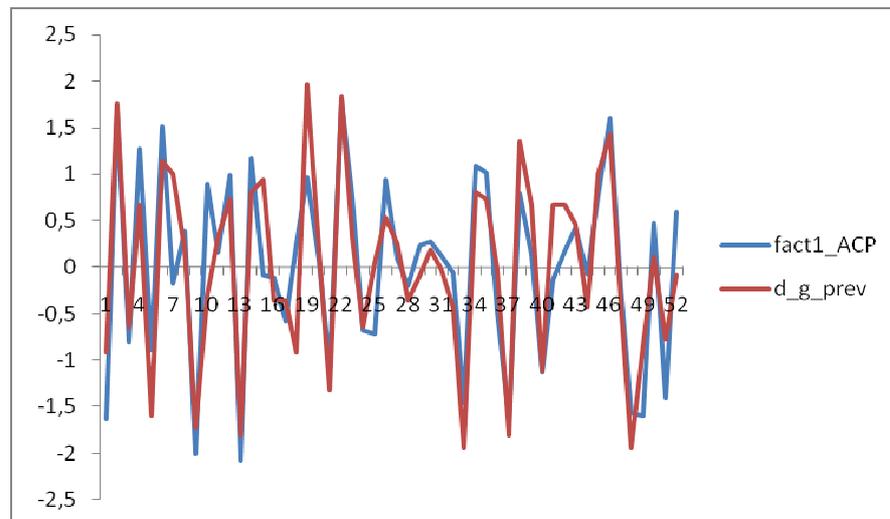
En ce qui concerne la production prévue, il semble, en première approximation, que cette information spécifique tendait à positionner le solde d'opinion au dessous du climat général en période de conjoncture ascendante et en dessus dans le cas contraire. Le croisement des deux courbes signale qu'un véritable retournement était en cours.

Figure26 : Tendence du stock en produits fabriqués



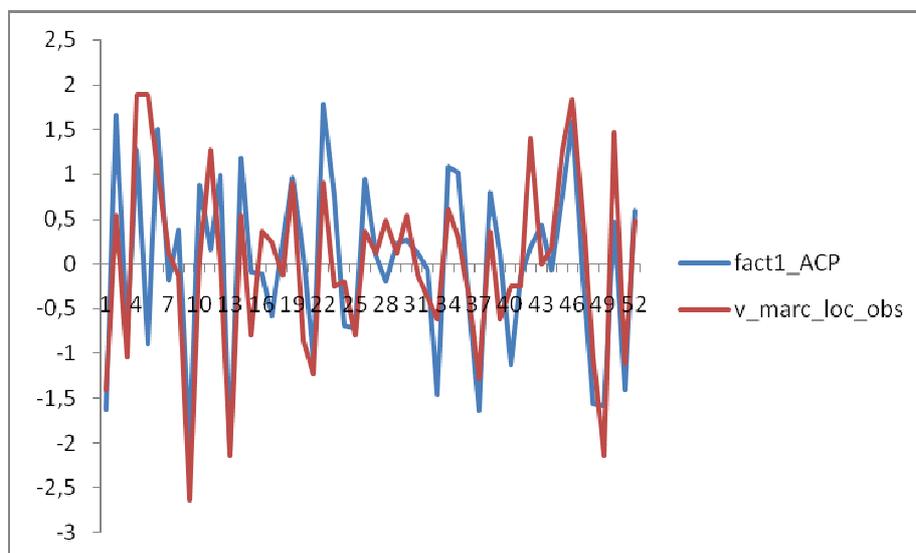
Un intérêt de la décomposition en information commune et spécifique du solde d'opinion relatif au niveau du stock en produits fabriqués pourrait être la proposition d'une lecture plus précise de la réponse. Le graphique ci-dessus montre un déphasage entre les deux courbes, en plus on peut constater que les mouvements dus aux fluctuations du climat général traduiraient le comportement de stockage lié au cycle conjoncturel ; en phase ascendante, pour des raisons tenant à la fois à une activité passée et à des anticipations de demande favorables, le niveau des stocks a tendance à être jugé de plus en plus léger et les industriels ont donc plutôt enclins à reconstituer leurs stocks, alors que les variations inverses se produisent en phase descendante.

Figure27 : Tendance de la demande globale



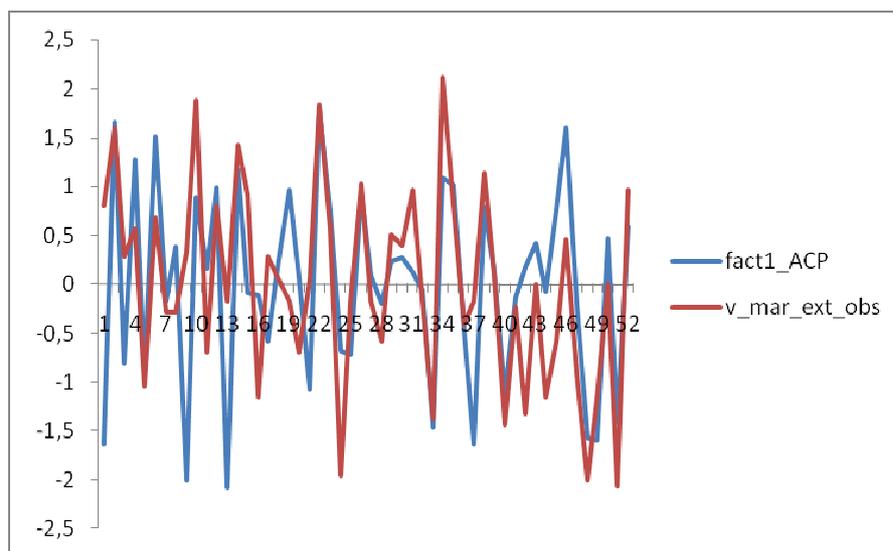
L'information spécifique lie à la question sur la demande globale offre un éclairage intéressant sur les décalages conjoncturels éventuels.

Figure28 : Tendance de vente sur le marché local



Les fluctuations de l'information spécifiquement apportée par la réponse aux ventes sur le marché local sont approximativement proches de celles de la série du facteur commun.

Figure29 : Tendence de vente sur le marché extérieur



L'information spécifique liée à la question aux ventes sur le marché extérieur offre un éclairage intéressant sur des décalages conjoncturels éventuels. La confrontation des deux courbes, montrent que ces derniers évoluent dans le même sens.