

UNE MODELISATION SEQUENTIELLE DE LA VaR

Alain MONFORT ⁽¹⁾

Mars, 2008

¹CNAM and CREST.

1 INTRODUCTION

La mise en place de règles prudentielles dans les institutions financières et, plus récemment, dans les compagnies d'assurance a fait apparaître une notion centrale : la valeur en risque ou VaR. Ce que la VaR cherche à mesurer est clair : étant donné un champ de sources de risque, un horizon et un niveau de confiance, la VaR est le seuil au dessous duquel la perte globale doit rester, à l'horizon choisi et avec le niveau de confiance retenu. Cependant, bien que l'objectif poursuivi soit bien défini, la mise en oeuvre pratique du calcul de la VaR soulève une multitude de questions parmi lesquelles :

- comment estimer la loi de probabilité de la perte générée par chaque source de risque ?
- comment tenir compte du fait que cette loi doit être estimée conditionnellement à une information variant dans le temps ?
- comment isoler, pour chaque risque, une partie générale (ou systématique) et une partie spécifique (ou idiosyncratique)?
- comment agréger les différents risques ?
- comment tenir compte de la diversification ou, au contraire, de la contagion des risques ?
- comment prendre en compte les corrélations des risques a une date donnée et de l'évolution temporelle de ces corrélations instantanées ?
- comment introduire les effets de corrélation dans le temps des divers risques?
- comment mesurer les effets de segmentation de la volatilité des risques ?
- comment analyser la propagation des chocs passés sur le niveau ou la volatilité des risques ?
- comment rendre compte des phénomènes d'asymétrie et de queues épaisses dans les lois des pertes ?
- comment ne pas exclure les risques extrêmes ?

Pour tenter de répondre, à cette liste non exhaustive de questions il existe des pratiques très différentes conduisant à un manque de transparence dans l'interprétation des chiffres. D'où vient cette absence de consensus? Après tout on pourrait se dire que les outils puissants de l'économétrie moderne permettent d'apporter des réponses rigoureuses à toutes ces questions et qu'il suffirait, pour chacune d'elles, d'utiliser la méthode reconnue comme optimale par la théorie. Les difficultés sont en fait de trois types. Tout d'abord les méthodes économétriques, aussi sophistiquées soient-elles, sont optimales dans le cadre d'un modèle statistique donné et c'est précisément la diversité des modèles qui pose problème. La deuxième difficulté est que, pour chaque type de modèle, il faut trouver un équilibre entre sa rigueur (et donc sa complexité) et son interprétabilité (et donc sa simplicité). Enfin les méthodes économétriques peuvent se heurter à une faiblesse des données en terme de quantité ou de qualité.

L'objectif de cette note est de se concentrer sur la construction de modèles et non sur leur traitement économétrique car, comme mentionné plus haut, celui-ci est le moins discutable, même s'il peut être complexe. Plus précisément, on va proposer une séquence de modèles, emboîtés les uns dans les autres, prenant progressivement en compte les questions évoquées ci-dessus. Pour chaque question on introduira une formalisation minimale, très faiblement paramétrée, de façon à maintenir une interprétation simple de chaque paramètre. On se limitera donc à la modélisation probabiliste et statistique strictement nécessaire à cette interprétation. On illustrera les diverses étapes par des simulations. La maquette la plus avancée décrite dans cette note pourrait être considérée comme une proposition de "colonne vertébrale" pouvant être enrichie dans diverses directions ou, au contraire, partiellement ignorée. Le but ultime est de faire un pas dans la direction d'une harmonisation des pratiques et une normalisation des procédures.

La note est organisée de la façon suivante. Dans le paragraphe 2 on va introduire plusieurs définitions équivalentes de la VaR. Le paragraphe 3 décrit une première maquette permettant d'appréhender les phénomènes liés à l'agrégation de diverses sources de risque, en particulier la contagion, la diversification, la corrélation instantanée; les diverses sources de risques peuvent provenir de différentes activités d'une même entité, elles peuvent alors apparaître à l'actif ou au passif, ou elles peuvent provenir de l'activité de plusieurs entités que l'on souhaite analyser à un niveau macroéconomique. Le paragraphe 4 propose une maquette plus générale qui introduit les notions de corrélation temporelle, de conditionnement par l'information présente et

d'évolution des corrélations instantanées. Le paragraphe 5 propose une maquette qui contient les deux premières comme cas particuliers et qui ajoute une modélisation de phénomènes importants comme la segmentation de la volatilité, l'hétéroscédasticité conditionnelle et la persistance des chocs de volatilité. Les trois maquettes sont comparées à l'aide de simulations utilisant le filtre de Kalman et le filtre de Kalman étendu. Le paragraphe 6 propose des extensions et des conclusions.

2 DEFINITIONS DE LA VaR.

2.1 VaR et quantiles.

Si on note π_{t+1} la perte globale (positive ou négative) engendrée, entre la date t et $t+1$, par le champ des sources de risque considéré, la VaR de niveau de confiance $1 - \alpha$ (ou de niveau de risque α), notée $VaR_t(\alpha)$, est définie par :

$$\begin{aligned} P_t[\pi_{t+1} \geq VaR_t(\alpha)] &= \alpha \\ \text{ou} & \\ P_t[\pi_{t+1} < VaR_t(\alpha)] &= 1 - \alpha \end{aligned} \tag{1}$$

P_t étant la loi de π_{t+1} , conditionnelle à l'information de la date t . Donc $VaR_t(\alpha)$ s'interprète comme le quantile conditionnel d'ordre $1 - \alpha$ de la perte globale. En pratique α sera faible, par exemple égal à 0.01.

2.2 VaR et réserves.

Un autre angle de vue de la VaR peut être utile. Si on note w_t la valeur des sources de risques en t on a $\pi_{t+1} = w_t - w_{t+1}$. Si on note res_t les réserves en t (non rémunérées entre t et $t+1$) permettant la survie en $t+1$ de l'entité définie par le champ des sources de risques, avec la probabilité $1 - \alpha$ on a :

$$\begin{aligned} P_t[w_{t+1} + res_t > 0] &= 1 - \alpha \\ \text{ou } P_t[w_t - \pi_{t+1} + res_t > 0] &= 1 - \alpha \\ \text{ou } P_t[\pi_{t+1} < w_t + res_t] &= 1 - \alpha \end{aligned}$$

ce qui implique :

$$VaR_t(\alpha) = w_t + res_t \tag{2}$$

Donc $VaR_t(\alpha)$ s'interprète aussi comme la somme de la valeur en t de l'entité et des réserves, c'est-à-dire comme le capital exposé ou la "valeur en risque" en cas de ruine.

2.3 VaR et capital initial optimal

Une troisième définition équivalente de la VaR est aussi éclairante.

Supposons que l'on cherche le capital z_t minimisant l'espérance d'une fonction l_{t+1} du type :

$$\begin{aligned} l_{t+1} &= (1 - \alpha)(\pi_{t+1} - z_t), \text{ si } \pi_{t+1} \geq z_t \\ &= \alpha(z_t - \pi_{t+1}), \text{ si } \pi_{t+1} < z_t \\ \text{ou } l_{t+1} &= (1 - \alpha)(\pi_{t+1} - z_t)^+ + \alpha(z_t - \pi_{t+1})^+ \end{aligned} \quad (3)$$

Autrement dit on attribue un coût élevé, $1 - \alpha$, à la situation dans laquelle la perte π_{t+1} est supérieure au capital et un coût faible, α , à la situation dans laquelle un coussin de garantie $z_t - \pi_{t+1}$ est inutilisé. On peut montrer (voir annexe 1) que le capital z_t réalisant le minimum de l'espérance conditionnelle de l_{t+1} en t est précisément $VaR_t(\alpha)$. Donc $VaR_t(\alpha)$ est le capital en t réalisant le compromis optimal, au sens de la fonction de coût choisie, entre une perte excessive et un capital inutilisé excessif.

2.4 VaR et capital économique.

Supposons que la loi conditionnelle de π_{t+1} à la date t soit celle de $\mu_t + \sigma_t U$ $\sigma_t > 0$, où U est une variable aléatoire de moyenne nulle, de variance 1 et de fonction de répartition F ; μ_t est alors l'espérance conditionnelle de π_{t+1} , σ_t son écart type conditionnel et on a :

$$\begin{aligned} P_t[\mu_t + \sigma_t U < VaR_t(\alpha)] &= 1 - \alpha \\ \text{ou } P_t[U < \frac{VaR_t(\alpha) - \mu_t}{\sigma_t}] &= 1 - \alpha \\ \text{ou } F\left[\frac{VaR_t(\alpha) - \mu_t}{\sigma_t}\right] &= 1 - \alpha \\ \text{ou } VaR_t(\alpha) &= \mu_t + \sigma_t F^{-1}(1 - \alpha) \end{aligned} \quad (4)$$

Donc $VaR_t(\alpha)$ se décompose en une "perte attendue" μ_t , c'est-à-dire la moyenne conditionnelle des pertes et une perte "perte inattendue" $\sigma_t F^{-1}(1 - \alpha)$, aussi appelée "capital économique". On voit également que la mesure de risque constituée par $VaR_t(\alpha)$ est une combinaison linéaire de

μ_t et σ_t , le poids $F^{-1}(1 - \alpha)$ affecté à la volatilité conditionnelle σ_t étant d'autant plus fort que le niveau de risque α est petit et que la loi de fonction de répartition F a une queue droite épaisse.

3 MODELE M1 : agrégation, diversification, contagion, corrélations instantanées

3.1 Définition

Supposons qu'il y ait n sources de risque indicées par $i = 1, \dots, n$. On note $\pi_{i,t+1}$ la perte engendrée par la source i entre t et $t + 1$, on note Π_{t+1} le vecteur de composantes $\pi_{i,t+1}, i = 1, \dots, n$, et $\pi_{t+1} = \sum_{i=1}^n \pi_{i,t+1}$ la perte globale.

Pour appréhender les problèmes d'agrégation et les phénomènes de diversification, contagion, dus aux corrélations instantanées entre les pertes $\pi_{i,t+1}$ on introduit la maquette simple suivante :

$$\pi_{i,t+1} = a_i + b_i F_{t+1} + u_{i,t+1}, i = 1, \dots, n \quad (5)$$

On suppose que les F_{t+1} sont inobservables, indépendants entre eux, de loi $N(0, 1)$, loi normale centrée réduite, que les $u_{i,t+1}$ sont inobservables, de loi $N(0, \sigma_i^2)$ indépendants entre eux et des F_{t+1} , et que les a_i, b_i sont des paramètres. F_{t+1} s'interprète donc comme un facteur de risque commun tandis que les $u_{i,t+1}$ sont des risques idiosyncratiques.

Les pertes $\pi_{i,t+1}$ sont normales de moyenne a_i et de variance $b_i^2 + \sigma_i^2$.

En outre deux pertes de la même période $\pi_{i,t+1}$ et $\pi_{j,t+1}$ ont une covariance égale à $b_i b_j$ et donc une corrélation $r_{ij} = \frac{b_i b_j}{(b_i^2 + \sigma_i^2)^{1/2} (b_j^2 + \sigma_j^2)^{1/2}}$. Les pertes de dates différentes sont ici supposées indépendantes.

Les paramètres a_i sont donc des paramètres de niveaux moyens des pertes, les b_i sont des paramètres de sensibilité au facteur de risque commun F_{t+1} et les σ_i^2 sont les variances des risques idiosyncratiques.

La covariance entre $\pi_{i,t+1}$ et F_{t+1} est b_i et la corrélation est $b_i / (b_i^2 + \sigma_i^2)^{1/2}$.

La perte globale π_{t+1} est de loi normale $N[e'a, (e'b)^2 + e'\sigma^2]$, en notant e le vecteur de taille n dont les composantes valent 1, et σ^2 le vecteur dont les

composantes sont les σ_i^2 ou $N(e'a, c^2 + e'\sigma^2)$, en notant $c = e'b = \sum_{i=1}^n b_i$.

La $VaR_t(\alpha)$ est donc en appliquant (4) :

$$VaR_t(\alpha) = e'a + (c^2 + e'\sigma^2)^{1/2}\Phi^{-1}(1 - \alpha) \quad (6)$$

Φ étant la fonction de répartition de loi $N(0, 1)$.

c dépend de la structure des covariances b_i entre $\pi_{i,t+1}$ et F_{t+1} . Plus c^2 ou $|c|$ est grand plus $VaR_t(\alpha)$ est grande : c^2 peut être vu comme une mesure de contagion. En particulier si $c = 0$, c'est-à-dire si la somme des sensibilités ou des covariances b_i entre les $\pi_{i,t+1}$ et F_{t+1} est nulle, le champ de risques est immunisé contre le facteur de risque commun. C'est évidemment le cas si tous les b_i sont nuls, c'est-à-dire si les $\pi_{i,t+1}$, $i = 1, \dots, n$, sont indépendants

Notons que les $VaR_{it}(\alpha)$ associées à chaque risque sont égales à :

$$VaR_{it}(\alpha) = a_i + (b_i^2 + \sigma_i^2)^{1/2}\Phi^{-1}(1 - \alpha) \quad (7)$$

et leur somme est :

$$\sum_{i=1}^n VaR_{it}(\alpha) = e'a + \sum_{i=1}^n (b_i^2 + \sigma_i^2)^{1/2}\Phi^{-1}(1 - \alpha) \quad (8)$$

Comparant (6) et (8) on vérifie que cette somme est toujours supérieure ou égale à $VaR_t(\alpha)$, l'égalité n'ayant lieu que si $\sigma_i^2 = 0, \forall i$ et si tous les b_i sont de même signe (voir Annexe 2), c'est-à-dire si les π_{it} sont des fonctions linéaires toutes croissantes (ou décroissantes) de F_t . Si on exclut ce cas limite, il y a diversification partielle.

Notons aussi que si on supposait, à tort, que les $\pi_{i,t+1}$ sont indépendants on en déduirait une $VaR_t(\alpha)$ erronée :

$$VaR_t^e(\alpha) = e'a + \left[\sum_{i=1}^n (b_i^2 + \sigma_i^2) \right]^{1/2} \Phi^{-1}(1 - \alpha) \quad (9)$$

Comparant (6) et (9) on voit que cette $VaR_t^e(\alpha)$ erronée peut sous-estimer ou sur-estimer la $VaR_t(\alpha)$ (voir Annexe 3). Elle la sous-estime si et seulement si $c^2 > \sum_{i=1}^n b_i^2$, c'est-à-dire si la contagion est suffisamment forte; dans ce cas $VaR_t^e(\alpha)$ sur-estime la diversification. La valeur $VaR_t^e(\alpha)$ est correcte si $c^2 =$

$\sum_{i=1}^n b_i^2$, c'est-à-dire $(\sum_{i=1}^n b_i)^2 = \sum_{i=1}^n b_i^2$; c'est le cas si les π_{it} sont effectivement indépendants, c'est-à-dire si les b_i sont nuls sauf au plus un, mais c'est le cas plus généralement si la somme des covariances $\sum_{i<j} b_i b_j$ est nulle.

3.2 Risques homogènes.

Si le champ de sources de risque est homogène, c'est-à-dire si $a_i = a_0$, $b_i = b_0$, $b_0 > 0$, $\sigma_i^2 = \sigma_0^2$ pour $i = 1, \dots, n$, on a :

$$VaR_t(\alpha) = na_0 + (n^2 b_0^2 + n\sigma_0^2)^{1/2} \Phi^{-1}(1 - \alpha) \quad (10)$$

Les VaR_{it} associées à chaque risque sont égales et valent :

$$VaR_{it}(\alpha) = a_0 + (b_0^2 + \sigma_0^2)^{1/2} \Phi^{-1}(1 - \alpha)$$

La VaR moyenne par source de risque est

$$\begin{aligned} \overline{VaR}_t(\alpha) &= \frac{1}{n} VaR_t(\alpha) \\ &= a_0 + (b_0^2 + \frac{\sigma_0^2}{n})^{1/2} \Phi^{-1}(1 - \alpha) \\ &< VaR_{it}(\alpha) \end{aligned}$$

Si n tend vers l'infini, $\overline{VaR}_t(\alpha)$ tend vers $a_0 + b_0 \Phi^{-1}(1 - \alpha)$, la partie due à la perte non anticipée tend donc vers $b_0 \Phi^{-1}(1 - \alpha)$, qui est non nulle si $b_0 \neq 0$; il y a alors diversification partielle même asymptotiquement car l'effet du risque commun F_t ne peut pas être diversifié. Si $b_0 = 0$, les $\pi_{i,t+1}$ sont indépendants de même loi, $\overline{VaR}_t(\alpha)$ tend vers a_0 , lorsque n tend vers l'infini, la partie non anticipée est nulle, et il y a diversification totale asymptotiquement.

3.3 Reconstitution du facteur de risque commun

Il peut être intéressant de reconstituer aux mieux le facteur de risque commun F_t au vu des observations des vecteurs de pertes Π_t, \dots, Π_1 . Dans le modèle statique $M1$ considéré dans ce paragraphe l'approximation optimale

\hat{F}_t de F_t ne dépend que du vecteur Π_t de la même date; plus précisément on a (voir annexe 4) :

$$\begin{aligned}\hat{F}_t &= b'(\Sigma + bb')^{-1}(\Pi_t - a) \\ &= \sum_{i=1}^n b_i \frac{\pi_{it} - a_i}{\sigma_i^2} / (1 + \sum_{i=1}^n \frac{b_i^2}{\sigma_i^2})\end{aligned}\quad (11)$$

(Σ étant la matrice diagonale de termes diagonaux σ_i^2).

La contribution de π_{it} à \hat{F}_t est positive si $\pi_{it} - a_i$ et b_i sont de même signe et négative sinon; elle est d'autant plus forte, en valeur absolue, que $|b_i|$ et $|\pi_{it} - a_i|$ sont grands et que σ_i^2 est petit.

L'espérance mathématique de \hat{F}_t est nulle comme celle de F_t et la variance de l'erreur $F_t - \hat{F}_t$ est :

$$V(F_t - \hat{F}_t) = \frac{1}{1 + \sum_{i=1}^n \frac{b_i^2}{\sigma_i^2}}\quad (12)$$

Elle est d'autant plus petite que les ratios de type "signal-bruit" $\frac{b_i^2}{\sigma_i^2}$ sont grands.

Il faut noter que le signe de F_t n'a pas de sens (car on peut toujours remplacer F_t par $-F_t$ et b par $-b$) de même que son amplitude, liée à la valeur arbitraire 1 de sa variance; en revanche ses valeurs relatives à deux dates sont interprétables. Par exemple si $\frac{\hat{F}_t}{\hat{F}_{t-1}}$ est positif et grand, cela signifie que le facteur de risque commun a eu un impact de même sens aux deux dates t et $t - 1$ pour chacune des pertes individuelles, que cet impact est beaucoup plus fort en t qu'en $t - 1$ et que le sens de cet impact sur π_{it} et $\pi_{i,t-1}$ est positif si b_i est positif, négatif sinon.

3.4 Un exemple

On considère les valeurs numériques $t = 1, \dots, 100$, $a_i = -0.2$, $\sigma_i^2 = 1\forall i$, $\alpha = 1\%$. Les b_i sont égaux à 1 ou -1 et on a n_1 valeurs égales à 1 ou donc $n - n_1$ valeurs égales à -1.

La mesure de contagion est $c^2 = (2n_1 - n)^2$. La VaR_t^e erronée calculée sous l'hypothèse d'indépendance des risques est supérieure à VaR_t si $c^2 < \sum_{i=1}^n b_i^2$, c'est-à-dire $(2n_1 - n)^2 < n$ ou $\frac{n}{2} - \frac{\sqrt{n}}{2} < n_1 < \frac{n}{2} + \frac{\sqrt{n}}{2}$. Si on prend $n = 16, n_1 = 12$, l'intervalle précédent est $\{6, 10\}$, donc $c^2 > \sum_{i=1}^n b_i^2$ et on a $VaR_t^e < VaR_t$: l'indice de contagion est suffisamment élevé pour entraîner une VaR_t^e inférieure à VaR_t .

La figure 1 montre une trajectoire simulée du facteur de risque commun et sa reconstitution par (11).

La reconstitution est bonne car, pour chaque perte, l'indice signal bruit $\frac{b_i^2}{\sigma_i^2}$ vaut 1, la variance de l'erreur de prévision donnée par (12) vaut $1/17 \simeq 0.059$ et l'écart type de l'erreur vaut 0.24

La figure 2 montre une trajectoire de la perte totale $\pi_t, t = 1, \dots, 100$, la VaR_t , la VaR_t^e erronée, calculée sous l'hypothèse d'indépendance instantanée des sources de risque et la perte globale attendue. La différence entre VaR_t et la perte attendue fournit le capital économique. On constate que la valeur de π_t dépasse une fois VaR_t , ce qui est conforme à la valeur du niveau de risque $\alpha = 1\%$. En revanche π_t dépasse VaR_t^e sept fois, car, comme on l'a vu, VaR_t^e est trop faible.

La somme de toutes les pertes sur la période est -225 (le gain est donc de 225) la somme des pertes dépassant VaR_t est 1,2 soit environ 0,5% des gains, alors que la somme des pertes dépassant VaR_t^e est 26 soit environ 11,5% des gains. On voit, dans cet exemple, que la non prise en compte des corrélations instantanées des sources de risque a des conséquences importantes.

4 MODELE M2: corrélation temporelle, conditionnement par l'information présente, évolution des corrélations instantanées.

Dans le modèle M1 toutes les variables aléatoires sont indépendantes d'une période à une autre : le modèle est statique. Dans le modèle M2 on introduit de la dynamique dans le comportement du facteur de risque commun F_t .

4.1 Dynamique du facteur de risque commun.

On suppose que le processus inobservable F_t est autorégressif d'ordre 1, ($AR(1)$) de variance marginale 1 :

$$F_{t+1} = \rho F_t + (1 - \rho^2)^{1/2} \varepsilon_{t+1} \quad (13)$$

où les ε_t sont indépendants entre eux, indépendants des $u_{i,t}$ et de loi $N(0, 1)$, et ρ est un paramètre vérifiant $|\rho| < 1$.

Dans ce modèle la corrélation entre F_{t+1} et F_t est ρ , la meilleure prévision de F_{t+1} en t , si F_t était observable, serait ρF_t et la variance conditionnelle de l'erreur de prévision serait $1 - \rho^2$. Il faut noter que, si ρ^2 proche de 1, cette variance conditionnelle est beaucoup plus faible que la variance non conditionnelle (ou marginale) qui vaut 1 : autrement dit la connaissance de F_t apporte une grande information sur F_{t+1} , contrairement à ce qui se passait dans le modèle M1. Le modèle M1 est emboîté dans M2 puisqu'il est obtenu en mettant le paramètre de corrélation ρ , à zéro.

4.2 Dynamique du vecteur des pertes et calcul de VaR_t .

La dynamique de F_t engendre une dynamique du vecteur Π_t des pertes π_{it} qui satisfont toujours les mêmes equations que dans M1 :

$$\pi_{i,t+1} = a_i + b_i F_{t+1} + u_{i,t+1}$$

ou, vectoriellement :

$$\Pi_{t+1} = a + b F_{t+1} + u_{t+1} \quad (14)$$

et on a : $E(\Pi_{t+1}) = a, V(\Pi_{t+1}) = \Sigma + bb', cov(\Pi_{t+1}, \Pi_t) = \rho bb', \Sigma$ étant la matrice diagonale de termes diagonaux σ_i^2 .

L'information disponible à la date t est constituée de $(\Pi_t, \Pi_{t-1}, \dots, \Pi_1)$, notée Π^t , et la loi conditionnelle pertinente pour le calcul de VaR_t est la loi conditionnelle de π_{t+1} sachant Π^t , où $\pi_{t+1} = \sum_{i=1}^n \pi_{i,t+1} = e' \Pi_{t+1}$ (e étant le vecteur de taille n dont les composantes valent 1).

La loi conditionnelle de Π_{t+1} sachant Π^t est la loi normale :

$$N(a + b\hat{F}_{t+1/t}, \Sigma + \omega_{t+1/t}^2 bb') \quad (15)$$

où $\hat{F}_{t+1/t} = E(F_{t+1}/\Pi^t)$ (espérance conditionnelle de F_{t+1} sachant Π^t) et où $\omega_{t+1/t}^2 = V(F_{t+1}/\Pi^t)$ (variance conditionnelle de F_{t+1} sachant Π^t).

Les quantités $\hat{F}_{t+1/t}$ et $\omega_{t+1/t}^2$ peuvent être calculées récursivement en utilisant le filtre de Kalman (voir Gourieroux-Monfort, 1995, chapitre 15) :

$$\left\{ \begin{array}{l} i) \quad \hat{F}_{t/t} = \hat{F}_{t/t-1} + b'(\omega_{t/t-1}^{-2} \Sigma + bb')^{-1} + (\Pi_t - a - b\hat{F}_{t/t-1}) \\ ii) \quad \hat{F}_{t+1/t} = \rho \hat{F}_{t/t} \\ iii) \quad \omega_{t/t}^2 = \omega_{t/t-1}^2 [1 - b'(\omega_{t/t-1}^{-2} \Sigma + bb')^{-1} b] \\ iv) \quad \omega_{t+1/t}^2 = \rho^2 \omega_{t/t}^2 + 1 - \rho^2 \\ \hat{F}_{1/0} = 0; \omega_{1/0}^2 = 1 \end{array} \right. \quad (16)$$

Les quantités $\hat{F}_{t/t}$ et $\omega_{t/t}^2$ intervenant dans (16) sont :

$$\begin{aligned} \hat{F}_{t/t} &= E(F_t/\Pi^t) \\ \hat{\omega}_{t/t}^2 &= V(F_t/\Pi^t) \end{aligned} \quad (17)$$

appelées respectivement "filtrage" de F_t et variance de l'erreur de filtrage.

$\hat{F}_{t/t}$ est la meilleure approximation de F_t connaissant Π^t , tandis que $\hat{F}_{t/t-1}$ meilleure approximation de F_t connaissant Π^{t-1} , c'est-à-dire la prévision optimale de F_t en $t-1$.

Dans le cas du modèle M1, on a $\rho = 0$ et $\hat{F}_{t+1/t} = 0, \omega_{t+1/t}^2 = 1$, et d'après 16 i)

$$\hat{F}_{t/t} = b'(\Sigma + bb')^{-1}(\Pi_t - a)$$

on retrouve l'équation (11).

Dans le cas général, la loi de Π_{t+1} sachant Π^t , donnée par (15), montre que la loi pertinente pour calculer VaR_t dépend des valeurs présentes et passées de π_t . Les équations (16) impliquent que la moyenne conditionnelle dépend linéairement des observations passées des pertes, tandis que la matrice de variance-covariance conditionnelle $\Sigma + \omega_{t+1/t}^2 bb'$ évolue dans le temps de façon indépendante des observations et converge rapidement vers une limite.

La loi de $\pi_{t+1} = e'\Pi_{t+1}$ sachant Π^t est la loi normale :

$$\begin{aligned} & N[e'a + e'b\hat{F}_{t+1/t}, \omega_{t+1/t}^2(e'b)^2 + e'\sigma^2] \\ & \text{avec : } \sigma^2 = (\sigma_1^2, \dots, \sigma_n^2)' \end{aligned} \quad (18)$$

ou en utilisant la notation $c = e'b$,

$$N[e'a + c\hat{F}_{t+1/t}, \omega_{t+1/t}^2 c^2 + e'\sigma^2] \quad (19)$$

On voit en particulier que la mesure de contagion c^2 a un impact sur l'évolution temporelle de la variance conditionnelle des pertes globales.

Le calcul de VaR_t découle de (4) et on a :

$$VaR_t = e'a + c\hat{F}_{t+1/t} + (\omega_{t+1/t}^2 c^2 + e'\sigma^2)^{1/2} \Phi^{-1}(1 - \alpha) \quad (20)$$

4.3 Calculs erronés de la VaR

Si on ne tient pas compte des effets de corrélation instantanée ou de corrélation temporelle on est amené à proposer des valeurs erronées de la VaR.

Ainsi si on oublie la corrélation temporelle, c'est-à-dire si on suppose à tort que $\rho = 0$, on se place dans le cas où les π_{t+1} sont indépendants, gaussiens de loi $N(e'a, c^2 + e'\sigma^2)$ et la VaR erronée obtenue est :

$$VaR_t^{(1)} = e'a + (c^2 + e'\sigma^2)^{1/2} \Phi^{-1}(1 - \alpha) \quad (21)$$

Si on oublie en outre la corrélation instantanée c'est-à-dire si on suppose, à tort, que les $\pi_{it}, i = 1, \dots, n$ sont indépendants, de lois $N(a_i, b_i^2 + \sigma_i^2)$, on aboutit à la VaR :

$$VaR_t^{(2)} = e'a + (e'b^2 + e'\sigma^2)^{1/2}\Phi^{-1}(1 - \alpha) \quad (22)$$

avec $b^2 = (b_1^2, \dots, b_n^2)'$

Ces valeurs de VaR erronées sont constantes dans le temps, contrairement à VaR_t , et peuvent être très différentes de VaR_t , comme le montre l'exemple suivant.

4.4 Exemple

On prend les mêmes valeurs numériques que dans 3.4 pour tous les paramètres autre que ρ et on prend $\rho = 0.8$. La figure 3 montre une simulation du facteur de risque commun F_t ainsi que son filtrage $\hat{F}_{t/t}$ et sa prévision $\hat{F}_{t/t-1}$. Conformément à la théorie le filtrage $\hat{F}_{t/t}$ fournit une meilleure approximation de F_t que la prévision $\hat{F}_{t/t-1}$, en particulier dans les périodes de retournement.

La figure 4 montre la simulation correspondante de la perte globale π_t ainsi, que sa prévision en $t-1$, à savoir $e'a + e'b\hat{F}_{t/t-1}$, là encore les prévisions les moins bonnes correspondent aux périodes de retournement.

La figure 5 montre à nouveau la simulation de la perte globale, la VaR_t exacte, les $VaR_t^{(1)}$ et $VaR_t^{(2)}$ erronées et la perte attendue.

Dans cet exemple π_t dépasse une fois (sur cent) VaR_t , ce qui est en cohérence avec $\alpha = 1\%$. La perte π_t dépasse également une fois $VaR_t^{(1)}$, mais $VaR_t^{(1)}$ est très souvent trop grande (dans les périodes 20-30, ou 75-100), ce qui implique un fort capital inutilisé. Enfin π_t dépasse treize fois $VaR_t^{(2)}$, qui est manifestement beaucoup trop faible. L'interprétation intuitive de ces résultats est simple. La $VaR_t^{(1)}$ est très souvent trop grande car l'oubli de la corrélation temporelle revient à remplacer la variance conditionnelle de F_t , à savoir $1 - \rho^2$, par sa variable non conditionnelle, à savoir 1, et donc à majorer la VaR. Il faut noter cependant que cette surestimation moyenne n'évite pas un dépassement de $VaR_t^{(1)}$, comme pour la VaR_t correcte. La $VaR_t^{(2)}$ avec oubli de la corrélation temporelle et de la corrélation instantanée subit deux influences opposées : le premier oubli, comme on vient de le voir, sur-estime la VaR, tandis que le deuxième la sous-estime, avec les valeurs

numériques retenues, comme cela a été constaté en 3.4. Cette deuxième influence l'emporte nettement dans cet exemple.

On sait d'ailleurs d'après 2.3 que pour chaque date t , VaR_t réalise le minimum en z_t de

$$E_t[(1 - \alpha)(\pi_{t+1} - z_t)^+ - \alpha(z_t - \pi_{t+1})^+]$$

donc, pour n'importe quelle séquence z_t on a :

$$E_t[(1 - \alpha)(\pi_{t+1} - z_t)^+ + \alpha(z_t - \pi_{t+1})^+ - (1 - \alpha)(\pi_{t+1} - VaR_t)^+ + \alpha(VaR_t - \pi_{t+1})^+] \geq 0$$

Cette inégalité est aussi vraie en remplaçant l'espérance conditionnelle par l'espérance, ce qui veut dire que la moyenne empirique de la somme entre crochet est forcément positive quand elle est prise sur un nombre de dates suffisamment grand.

Ce résultat signifie que la somme pondérée des valeurs absolues des écarts entre π_{t+1} et toute suite z_t , $\sum_{t=1}^T [(1 - \alpha)(\pi_{t+1} - z_t)^+ + \alpha(z_t - \pi_{t+1})^+]$, est minimale lorsqu'on prend $z_t = VaR_t$.

5 MODELE M3 : segmentation de la volatilité, hétéroscédasticité conditionnelle, persistance des chocs de volatilité

Dans le modèle M2, le facteur de risque commun F_t , qui pilote la dynamique de Π_t , a une variance conditionnelle au passé constante, et égale à $1 - \rho^2$; le processus F_t est conditionnellement homoscédastique. Une hypothèse plus réaliste consiste à introduire une dynamique avec une variance conditionnelle au passé non constante, c'est-à-dire à introduire de l'hétéroscédasticité conditionnelle, pour prendre en compte les effets de segmentation de la volatilité.

5.1 Modélisation AR(1) - ARCH (1) du facteur de risque commun.

Dans le modèle M3 on suppose la dynamique suivante pour F_t :

$$\begin{aligned}
F_{t+1} &= \rho F_t + (1 - \rho^2)^{1/2} s_{t+1} \varepsilon_{t+1}, |\rho| < 1 \\
s_{t+1}^2 &= 1 - \gamma + \gamma s_t^2 \varepsilon_t^2, 0 < \gamma < 1
\end{aligned} \tag{23}$$

où les ε_t sont indépendants et de loi $N(0, 1)$.

La variance conditionnelle du bruit $\eta_{t+1} = s_{t+1} \varepsilon_{t+1}$ sachant son passé est donc $1 - \gamma + \gamma \eta_t^2$, c'est un processus *ARCH*(1) de variance non conditionnelle 1.

Le modèle M2 est obtenu en faisant $\gamma = 0$ il est donc emboîté dans M3. On voit qu'une valeur élevée de η_t , en valeur absolue, et donc de η_t^2 , entraîne une valeur élevée de la variance de η_{t+1} et donc une grande probabilité de valeur élevée de $|\eta_{t+1}|$: c'est le phénomène de segmentation de la volatilité, qui sera d'autant plus marqué que γ est proche de 1; γ s'interprète donc comme un paramètre de segmentation de la volatilité.

De même la loi conditionnelle de F_{t+1} sachant son passé est :

$$N\{\rho F_t, (1 - \rho^2)(1 - \gamma) + \gamma(F_t - \rho F_{t-1})^2\} \tag{24}$$

les valeurs présentes et passées de F_t ont un impact non seulement sur l'espérance conditionnelle de F_{t+1} mais aussi sur sa variance conditionnelle.

Le vecteur des pertes Π_{t+1} est toujours donné par :

$$\Pi_{t+1} = a + bF_{t+1} + u_{t+1}$$

où les u_t sont indépendants entre eux, indépendants des ε_t et de loi $N(0, \Sigma)$, Σ matrice diagonale de terme diagonaux σ_i^2 .

Dans ce modèle on a toujours les propriétés de M2 :

$$E(F_{t+1}) = 0, V(F_{t+1}) = 1$$

$$E(\Pi_{t+1}) = a, V(\Pi_{t+1}) = \Sigma + bb'$$

$$\text{cov}(\Pi_{t+1}, \Pi_t) = \rho bb'$$

En revanche la loi conditionnelle de Π_{t+1} sachant Π^t est modifiée. Cette loi n'est d'ailleurs plus gaussienne mais elle est peut être approximée par une loi gaussienne, du même type que dans le modèle M2 :

$$N(a + b\hat{F}_{t+1/t}, \Sigma + \omega_{t+1/t}^2 bb') \quad (25)$$

mais avec un calcul récursif de $\hat{F}_{t+1/t}$ et $\omega_{t+1/t}^2$ différent. Il est donné par le filtre de Kalman étendu, obtenu en remplaçant 16 iv) par :

$$\omega_{t+1/t}^2 = \rho^2 \omega_{t/t}^2 + (1 - \rho^2) s_{t+1/t}^2 \quad (26)$$

avec :

$$s_{t+1/t}^2 = 1 - \gamma + \gamma(\hat{F}_{t/t} - \rho\hat{F}_{t-1/t-1})^2(1 - \rho^2)^{-1}$$

Le calcul de VaR_t peut de la même façon être approximé par (20).

On peut également calculer des VaR de façon erronée. La $VaR_t^{(3)}$ est calculée en ignorant la segmentation de la volatilité, c'est-à-dire en mettant $\gamma = 0$ et en utilisant le filtre de Kalman (16). On peut aussi calculer $VaR_t^{(1)}$ en ignorant la segmentation de la volatilité et la corrélation temporelle et $VaR_t^{(2)}$ et ignorant la segmentation de volatilité et les deux types de corrélation.

5.2 Exemple

On prend les mêmes valeurs des paramètres qu'en 4.4 et $\gamma = 0.8$. La figure 6 montre le facteur de risque commun, son filtrage et sa prévision; on constate à nouveau des problèmes de prévision lors des retournements mais aussi lors des périodes de haute volatilité. La figure 7 confirme ses constatations pour la perte globale. La figure 8 montre que π_t dépasse VaR_t une fois (sur cent). $VaR_t^{(3)}$ oscille autour de la VaR_t correcte : elle peut être trop faible ou trop forte selon les périodes, ce qui est cohérent avec le fait que la variance conditionnelle de F_{t+1} , à savoir $(1 - \rho^2)s_{t+1}^2$, a pour espérance $(1 - \rho^2)$, qui est la variance obtenue en prenant $\gamma = 0$. $VaR_t^{(1)}$ est dépassée une fois mais est systématiquement trop grande tandis que $VaR_t^{(2)}$ est dépassée sept fois. On retrouve des résultats analogues à ceux de la section 4.4.

6 EXTENSIONS ET CONCLUSION

Les modèles M1, M2, M3 pourraient être généralisés dans de nombreuses directions tout en conservant la même facilité de traitement grâce à la flexibilité

du filtre de Kalman et du filtre de Kalman étendu :

- introduction de variables exogènes
- introduction de plusieurs retards dans la dynamique des facteurs.
- introduction d'effets asymétriques des chocs de volatilité
- prise en compte de VaR_t à des horizons supérieurs à 1

D'autres développements pourraient être introduits à condition d'utiliser des techniques plus avancées comme l'inférence indirecte, le maximum de vraisemblance simulé ou les algorithmes MCML; avec ces techniques on pourrait par exemple s'affranchir de la normalité conditionnelle en introduisant des mélanges de lois normales et en utilisant les techniques de traitement des variables cachées discrètes.

Mais les maquettes M1, M2, M3 pourraient déjà constituer les outils de base d'une harmonisation des méthodes de calcul de la VaR. Une telle harmonisation pourrait déboucher sur des normes structurées, susceptibles de s'adapter à divers contextes, permettant une meilleure évaluation des risques et donnant des possibilités de comparaison des comportements de diverses entités à un instant donné ou à des périodes différentes.

ANNEXE 1

On a :

$$\begin{aligned}
 l_{t+1} &= (1 - \alpha)(\pi_{t+1} - z_t) \mathbb{1}_{\pi_{t+1} \geq z_t} + \alpha(z_t - \pi_{t+1}) \mathbb{1}_{\pi_{t+1} < z_t} \\
 \frac{\partial E_t l_{t+1}}{z_t} &= -(1 - \alpha)P_t(\pi_{t+1} \geq z_t) + \alpha P_t(\pi_{t+1} < z_t) \\
 &= -(1 - \alpha)[1 - P_t(\pi_{t+1} < z_t)] + \alpha P_t(\pi_{t+1} < z_t) \\
 &= -(1 - \alpha) + P_t(\pi_{t+1} < z_t)
 \end{aligned}$$

Donc la nullité de la dérivée implique :

$$z_t = VaR_t(\alpha).$$

ANNEXE 2

$$VaR_t = e'a + [(\sum_{i=1}^n b_i)^2 + \sum_{i=1}^n \sigma_i^2]^{1/2} \Phi^{-1}(1 - \alpha)$$

$$\sum_{i=1}^n VaR_{it} = e'a + [\sum_{i=1}^n (b_i^2 + \sigma_i^2)]^{1/2} \Phi^{-1}(1 - \alpha)$$

$\sum_{i=1}^n VaR_{it} - VaR_t$ est du signe de :

$$\sum_{i=1}^n (b_i^2 + \sigma_i^2)^{1/2} - [(\sum_{i=1}^n b_i)^2 + \sum_{i=1}^n \sigma_i^2]^{1/2}$$

ou de

$$\sum_{i=1}^n (b_i^2 + \sigma_i^2) + \sum_{i \neq j} (b_i^2 + \sigma_i^2)^{1/2} (b_j^2 + \sigma_j^2)^{1/2}$$

$$- \sum_{i=1}^n (b_i^2 + \sigma_i^2) - \sum_{i \neq j} b_i b_j$$

ou de

$$\sum_{i \neq j} [(b_i^2 + \sigma_i^2)^{1/2} (b_j^2 + \sigma_j^2)^{1/2} - b_i b_j]$$

d'où le résultat.

ANNEXE 3

$VarR_t(\alpha) - VaR_t^e(\alpha)$ est du signe de

$$(c^2 + e'\sigma^2)^{1/2} - \left[\sum_{i=1}^n (b_i^2 + \sigma_i^2) \right]^{1/2}$$

ou de $c^2 - \sum_{i=1}^n b_i^2$

ANNEXE 4

On a :

$$(\Sigma + bb')^{-1} = \Sigma^{-1} - \frac{\Sigma^{-1}bb'\Sigma^{-1}}{1 + b'\Sigma^{-1}b}$$

et donc :

$$\begin{aligned} b'(\Sigma + bb')^{-1} &= b'\Sigma^{-1} - \frac{b'\Sigma^{-1}bb'\Sigma^{-1}}{1 + b'\Sigma^{-1}b} \\ &= \frac{b'\Sigma^{-1}}{1 + b'\Sigma^{-1}b} \end{aligned}$$

On en déduit :

$$\begin{aligned} \hat{F}_t &= b'(\Sigma + bb')^{-1}(\Pi_t - a) \\ &= \sum_{i=1}^n b_i \frac{\pi_{it} - a_i}{\sigma_i^2} / \left(1 + \sum_{i=1}^n \frac{b_i^2}{\sigma_i^2} \right) \end{aligned}$$

De même :

$$\begin{aligned}V(F_t - \hat{F}_t) &= V(F_t) - V(\hat{F}_t) \\&= 1 - b'(\Sigma + bb')^{-1}b \\&= 1 - \frac{b'\Sigma^{-1}b}{1 + b'\Sigma^{-1}b} \\&= \frac{1}{1 + b'\Sigma^{-1}b} \\&= 1 / \left(1 + \sum_{i=1}^n \frac{b_i^2}{\sigma_i^2}\right)\end{aligned}$$

FIGURE 1: Facteur de Risque Commun (Trait plein)
Reconstitution (Pointilles)

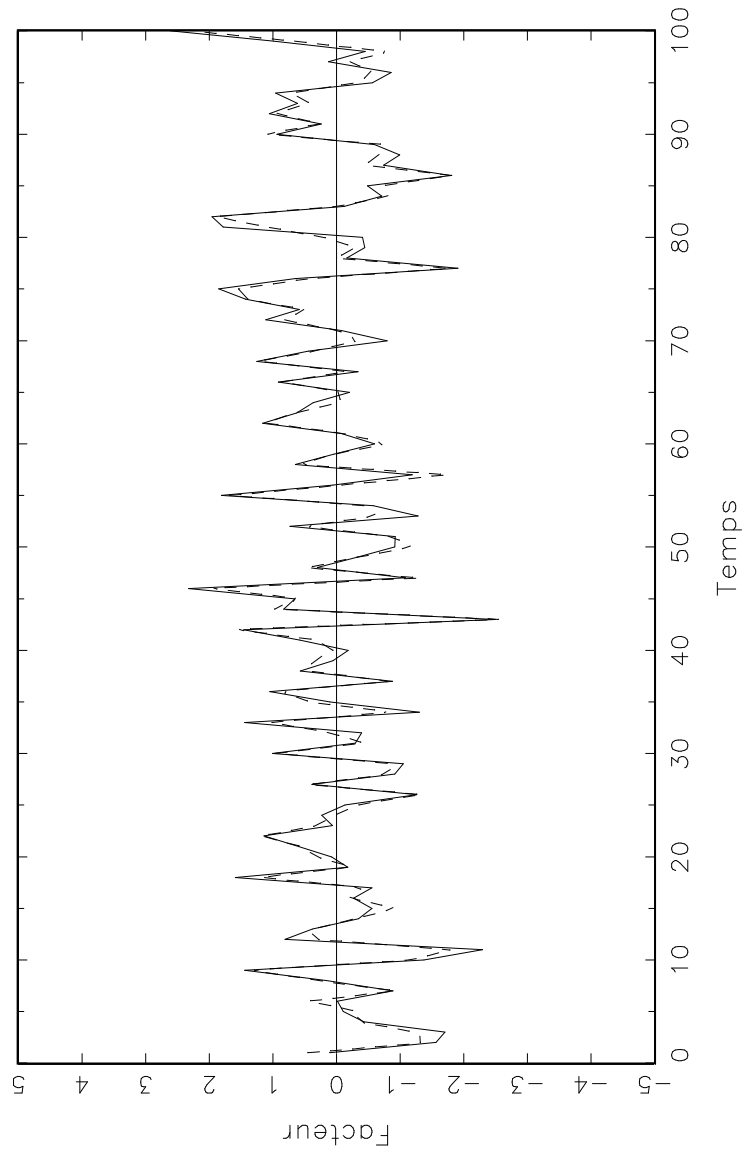
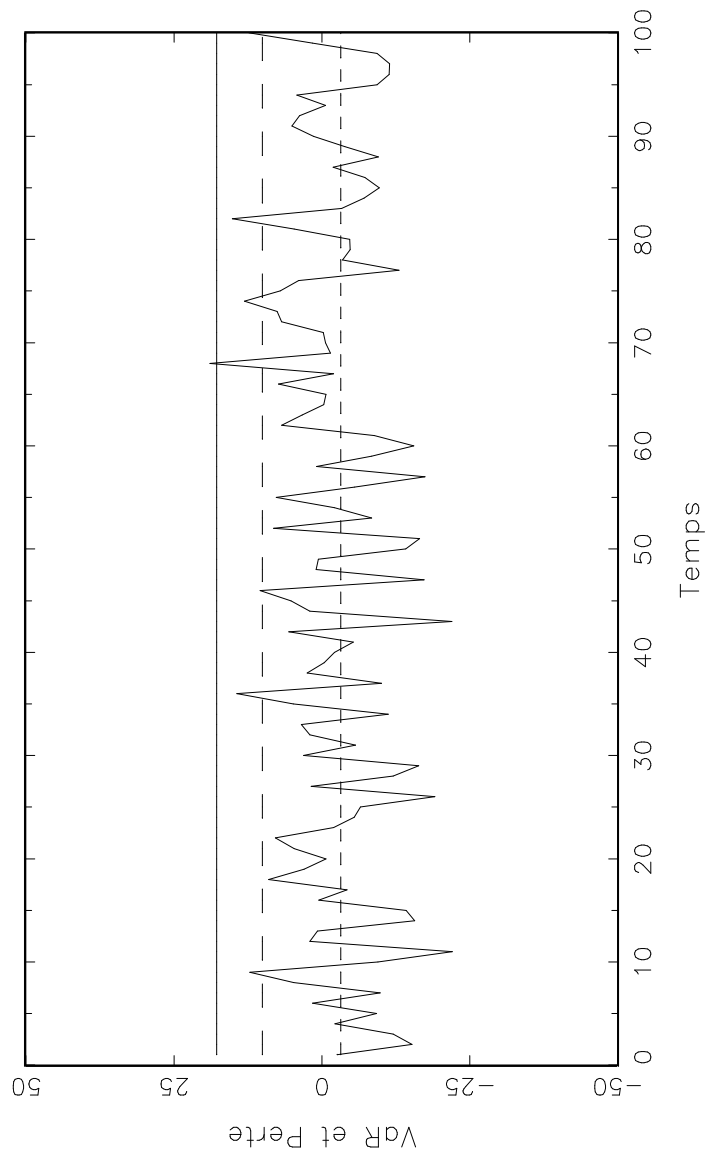


FIGURE 2: Perte et VaR(Traits Pleins) ,
Perte attendue(Pointilles) ,
VaR sans Correlation Instantanee(Pointilles longs)



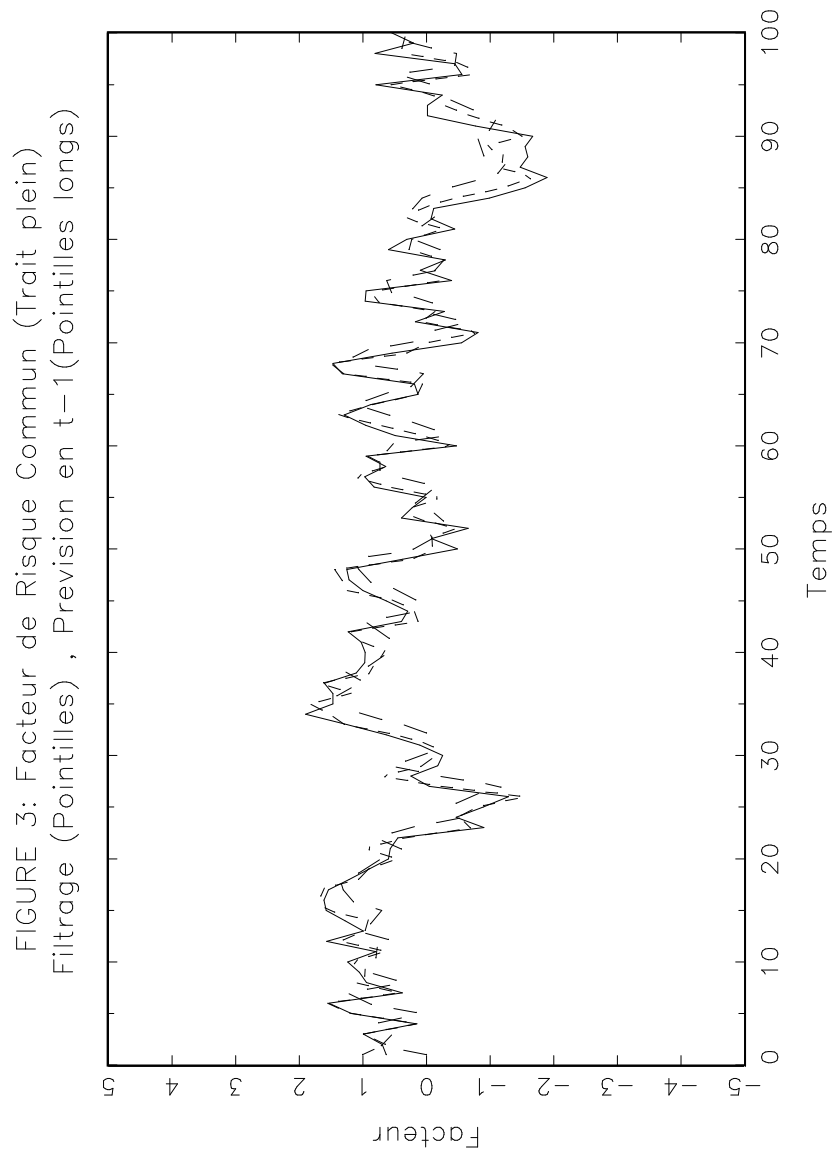


Figure 4: Perte(Trait plein) et Prevision(Pointilles)

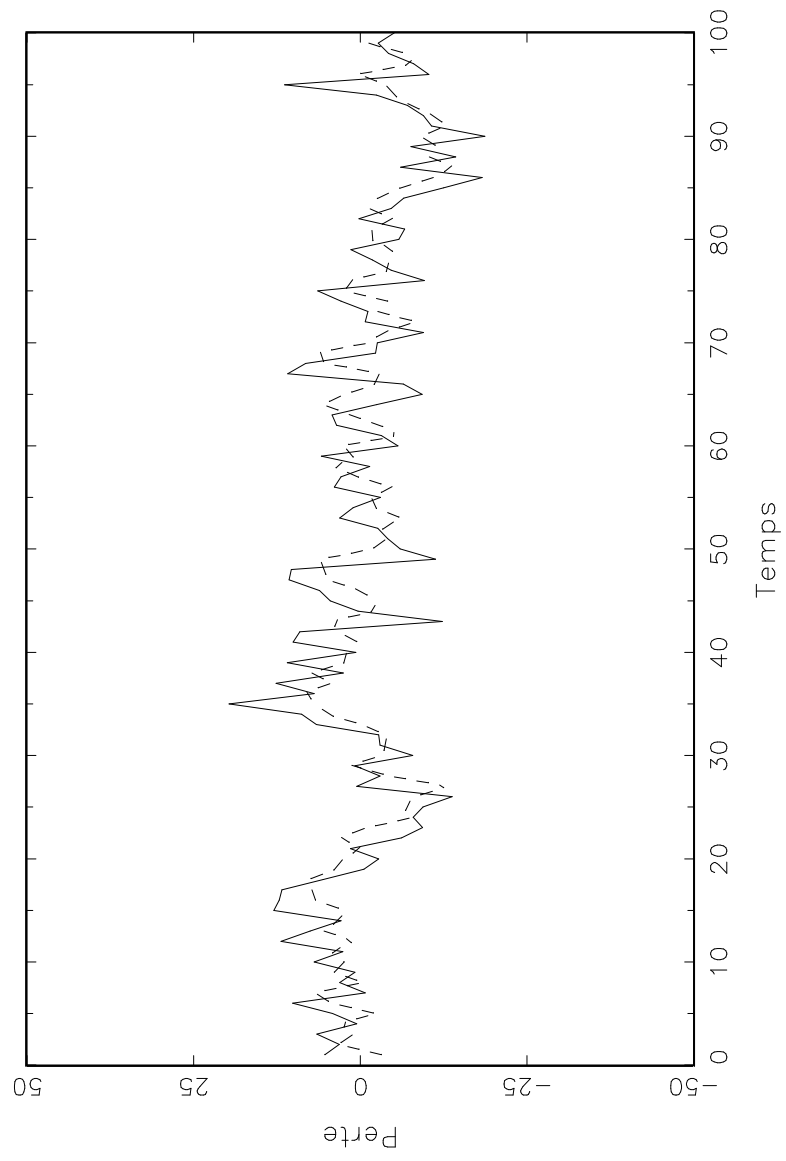
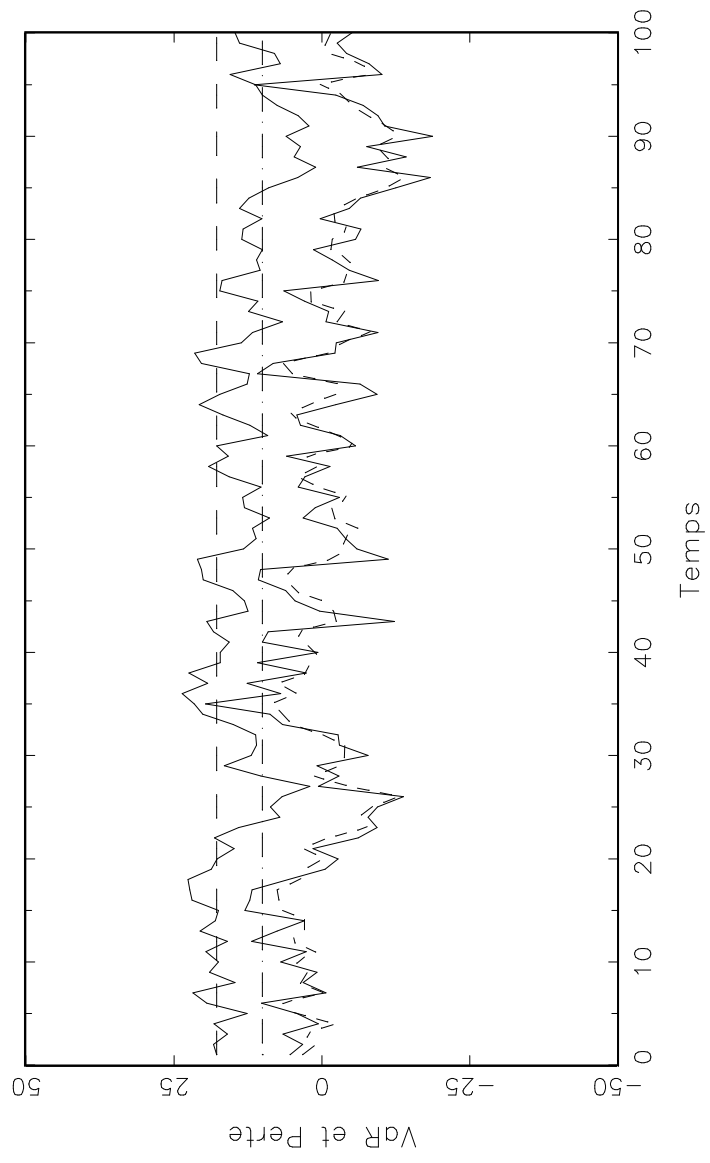


FIGURE 5: Perte et VaR(Traits Pleins) ,
Perte attendue(Pointilles),VaR sans Corr Temporelle(Pointilles longs)
VaR sans Corr Temporelle et Instantanees(Points et tirets)



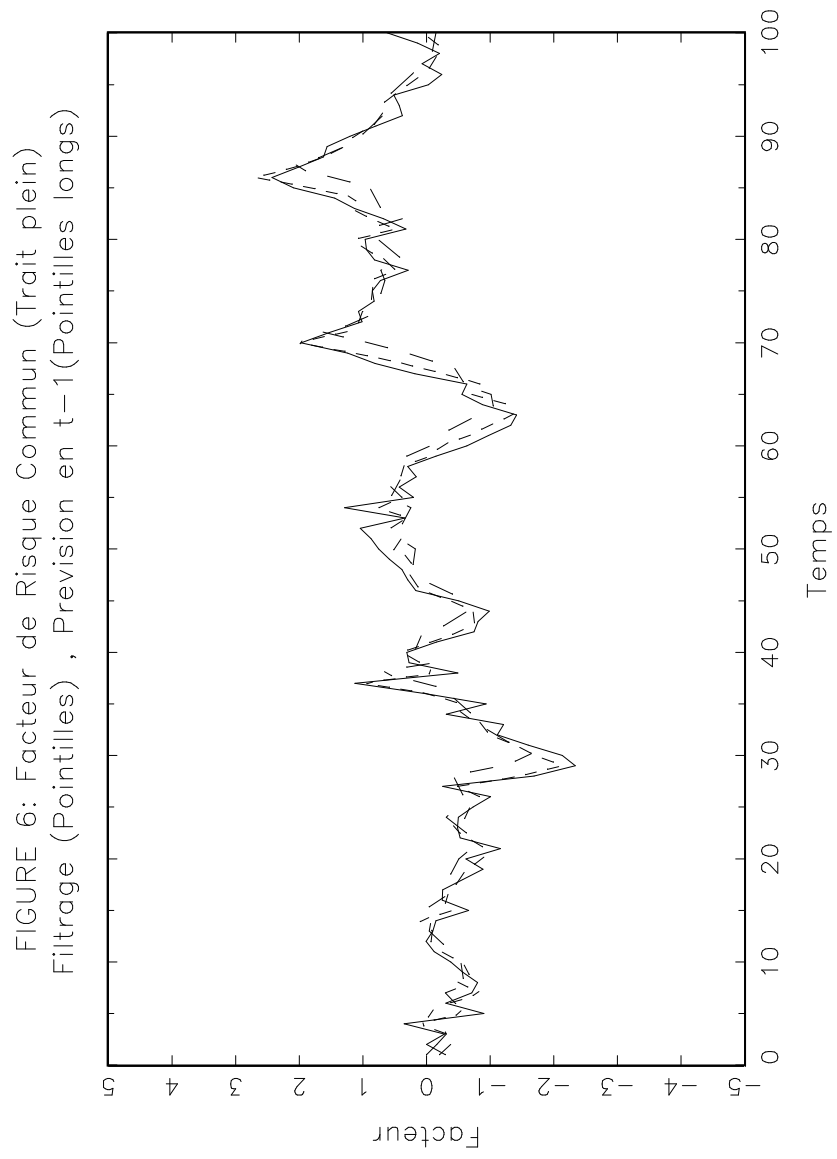


Figure 7: Perte(Trait plein) et Prevision(Pointilles)

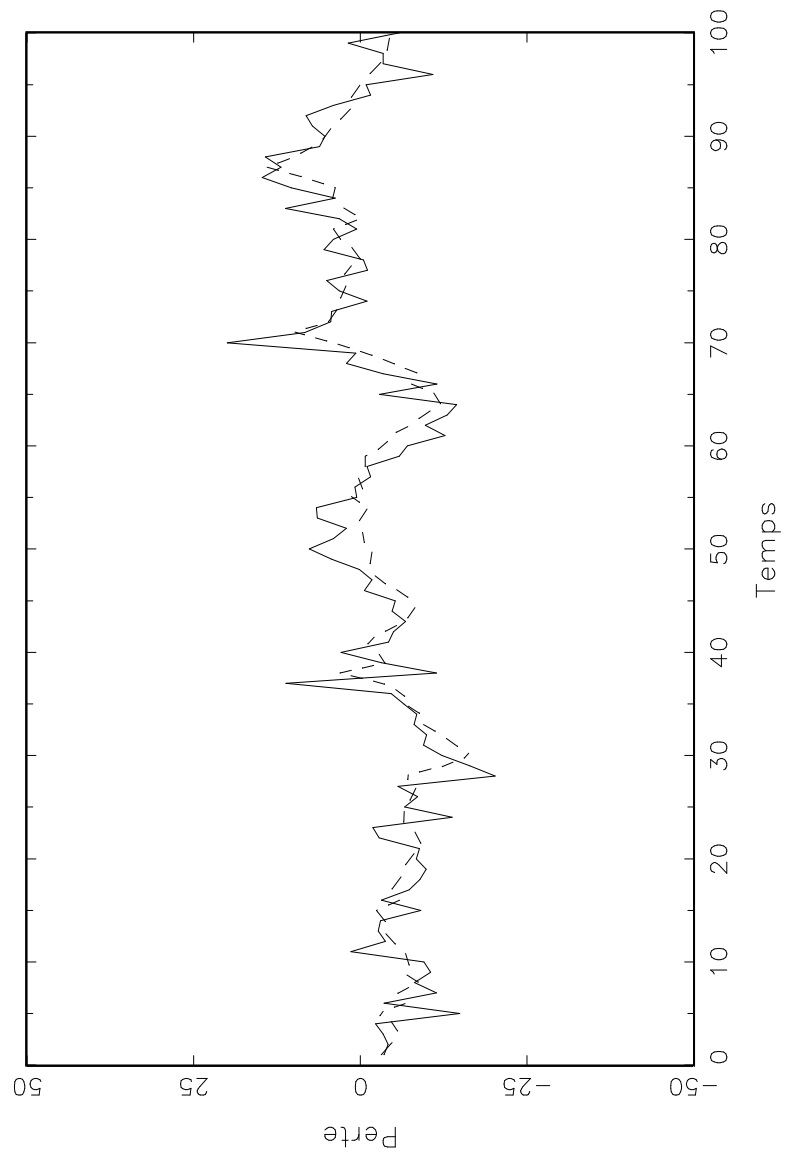


FIGURE 8: Perte et VaR(Traits Pleins),
Perte attendue(Pointilles),VaR1(Pointilles longs),VaR2(Pointes et tirets)
VaR3(Pointilles)

