

## Fondements de la théorie des valeurs extrêmes, ses principales applications et son apport à la gestion des risques du marché pétrolier

*Fundamentals of extreme value theory, its principal applications and its contribution to the risk management of the oil market*

**Bechir Raggad**

---



**Édition électronique**

URL : <http://journals.openedition.org/msh/11069>  
DOI : 10.4000/msh.11069  
ISSN : 1950-6821

**Éditeur**

Centre d'analyse et de mathématique sociales de l'EHESS

**Édition imprimée**

Date de publication : 12 octobre 2009  
Pagination : 29-63  
ISSN : 0987-6936

**Référence électronique**

Bechir Raggad, « Fondements de la théorie des valeurs extrêmes, ses principales applications et son apport à la gestion des risques du marché pétrolier », *Mathématiques et sciences humaines* [En ligne], 186 | Été 2009, mis en ligne le 15 octobre 2009, consulté le 23 juillet 2020. URL : <http://journals.openedition.org/msh/11069> ; DOI : <https://doi.org/10.4000/msh.11069>

---

## FONDEMENTS DE LA THÉORIE DES VALEURS EXTRÊMES, SES PRINCIPALES APPLICATIONS ET SON APPORT À LA GESTION DES RISQUES DU MARCHÉ PÉTROLIER

Bechir RAGGAD<sup>1</sup>

RÉSUMÉ – Depuis quelques années, la théorie des valeurs extrêmes a reçu beaucoup d'attention aussi bien sur le plan théorique que sur le plan pratique. Les domaines d'applications sont en effet très variés : hydrologie, météorologie, biologie, ingénierie, gestion de l'environnement, finance, assurance, sciences sociales, etc. ; en effet la gestion des risques est devenue aujourd'hui fondamentale dans tous ces domaines. Cependant, ce type d'analyse basé sur la théorie des valeurs extrêmes est limité, voire très limité pour analyser les risques encourus sur le marché pétrolier alors que ces techniques peuvent sans doute aider à l'étude quantitative des risques. Le succès croissant de cette théorie nous a incité à mieux explorer son apport en matière de la gestion des fluctuations extrêmes sur le marché pétrolier.

MOTS-CLÉS – Analyse de sensibilité, Gestion des risques, Loi de Pareto généralisée, Pétrole, Théorie des valeurs extrêmes, « Value at Risk »

SUMMARY – Fundamentals of extreme value theory, its principal applications and its contribution to the risk management of the oil market

For many years, Extreme value theory has received much attention on the theoretical level as well as on the practical one. The scope of its applications are indeed very varied and include hydrology, meteorology, biology, engineering, environmental management, finance, insurance, social life, etc. since risk management has become fundamental today in all these fields. However, this type of analysis based on the extreme value theory is little, even very limited to analyze the risks incurred on oil market whereas these techniques can undoubtedly help in a quantitative study of the risks. The increasing success of this theory in many fields encouraged us to better investigate its contribution to risk management of extreme fluctuations on crude oil market.

KEYWORDS – Crude oil, Extreme value theory, Generalized Pareto law, Risk management, Sensitivity analysis, Value at Risk

### 1. INTRODUCTION

La modélisation des événements extrêmes (ouragan, tremblement de terre ou inondation, crues, crises financières, krachs, chocs pétroliers) est aujourd'hui un champ de recherches particulièrement actif, notamment par l'importance de leurs impacts économiques et sociaux. En particulier, depuis quelques années, on note un intérêt croissant pour l'application de la Théorie des Valeurs Extrêmes (TVE) pour la modélisation de tels événements. Pour une présentation assez complète du sujet, nous renvoyons à l'ouvrage de référence de Embrechts, Klüppelberg et Mikosch [1997]

---

<sup>1</sup> Laboratoire Business & Economic SStatistics MODling (BESTMOD), Institut Supérieur de Gestion, 41 rue de la Liberté, Cité Bouchoucha, Le Bardo 2000, Tunis, Tunisie, raggadbechir@yahoo.fr

rappelant les principaux résultats théoriques sur la TVE et à Reiss et Thomas [2001] qui proposent un certain nombre d'exemples pratiques, en finance, en assurance et en sciences environnementales.

Les domaines d'applications utilisant les modèles de la TVE n'ont cessé de se développer ces dernières années touchant des domaines variés. En hydrologie, domaine dans lequel la prévision des crues par exemple est particulièrement importante [Davison et Smith, 1990 ; Katz, 2002], en assurance dont l'une des préoccupations est la prise en compte des grands sinistres [McNeil *et al.*, 1997 ; Rootzen et Tajvidi, 1997]. Leur introduction en finance [Embrechts *et al.*, 1997 ; Danielsson et de Vries, 1997 ; McNeil, 1998 ; Longin, 1998, 2000 ; Embrechts, 1999 ; Gençay et Selçuk, 2004]) est une réponse immédiate à la remise en cause de l'hypothèse de normalité surtout avec les observations en hautes fréquences. En météorologie [Coles et Walshaw, 1994 ; Smith, 2001 ; Klajnmanic, 2003] où l'étude de la vitesse du vent, par exemple, permet d'évaluer le degré de résistance des matériaux face à la pression exercée par le vent (au cours d'une tempête par exemple) sur les bâtiments ou les structures de génie civil.

Dans les domaines des sciences humaines et sociales, la théorie des valeurs extrêmes pourra également contribuer à la compréhension de nombreux problèmes sociaux. À titre d'exemple, dans le domaine de la démographie, tout un débat qui a été initié par Gumbel [1937], auquel Fréchet a pris une part active, sur la notion de « durée extrême de la vie humaine» et sur sa mesure [Thatcher, 1999]. Des études récentes ont mis en lumière l'allure remarquable de la mortalité aux grands âges, à savoir une décroissance du taux de croissance de la mortalité à partir d'un certain âge. Horiuchi et Wilmoth [1998] ont ainsi montré que sur l'échelle logarithmique, la courbe des taux de mortalité présentait une allure concave aux grands âges. Ceci a conduit les démographes à rechercher des modèles compatibles avec cette réalité. Les démographes peuvent également se demander s'il existe un âge limite au-delà duquel nul ne survit ? Si oui, quel est cet âge ultime, auquel chacun peut rêver de parvenir ? La TVE permet de prévoir et d'extrapoler la distribution de probabilité de l'âge maximum que l'être humain pourrait atteindre [Ledford et Robinson, 1999 ; Han, 2005]. La TVE va contribuer vivement au débat sur l'existence d'un âge limite de longévité de l'être humain. Une telle étude est souvent utile pour les organismes de sécurité sociale (dépenses destinées aux personnes âgées, coût de traitement, ...).

Prévoir certains événements ou comportements, à partir d'une étude des valeurs extrêmes d'une série, est donc un des principaux objectifs pour ceux qui tentent d'appliquer la théorie de ces valeurs. Cette théorie est apparue entre 1920 et 1940, grâce à Fréchet, Fisher et Tippett, Gumbel et Gnedenko. Lorsque l'on modélise le maximum d'un ensemble de variables aléatoires, alors, sous certaines conditions que nous préciserons plus loin, la distribution ne peut appartenir qu'à l'une des trois lois suivantes: Weibull (à support borné), Gumbel et Fréchet (à support non borné). Ces trois lois définissent une famille de distributions statistiques appelée « famille paretienne », dont les applications en sciences sociales sont innombrables et très diverses. Sur ce vaste thème, nous renvoyons à l'article de Barbut [1998] portant sur l'application de ces lois à l'étude de la concentration urbaine et de son évolution.

L'introduction de la TVE est devenue une exigence dans de nombreux domaines. En effet, les outils probabilistes traditionnels (développés dans un univers gaussien) sont inadaptés à l'appréhension de ces comportements extrêmes : la moyenne n'existe plus dans le nouvel univers des risques. Pourtant, il existe des situations dans lesquelles cette opération de moyennisation est légitime pour des variables concentrées autour de leur moyenne, avec des variations de faible amplitude, dont toutes les valeurs

pratiquement possibles sont du même ordre de grandeur : celles de « l'homme moyen » de Quetelet<sup>2</sup>. La présence des extrêmes dans une population statistique rend donc caduque l'approche de Quetelet et nécessite une autre approche, celle inventée par Pareto fin du XIX<sup>e</sup> siècle. Pareto s'intéresse à la distribution des revenus dans une économie donnée. Il a constaté que les revenus se répartissent toujours selon une loi mathématique décroissante d'allure entre une petite minorité de riches, une classe moyenne minoritaire, et une large majorité de pauvres. Il a montré que plus que plus les revenus sont élevés, plus la queue de distribution des revenus sera hyperbolique. Cette forme d'hyperbole permet une modélisation précise de la queue de distribution des revenus, et donc permet de bien décrire l'impact des grands revenus sur le revenu moyen. Pour plus de détails sur les lois de type Pareto, nous renvoyons à Barbut [1988]. Il est à noter également que les distributions  $\alpha$ -stables de Lévy, introduites pour la première fois en finance par Mandelbrot [1962], constituent une classe très riche de lois de probabilité capables de représenter différentes asymétries et des queues très lourdes. Une distribution stable possède la propriété suivante : la queue de distribution d'une loi stable décroît selon une loi de Pareto. De nombreuses études statistiques ont été réalisées sur les marchés depuis les travaux de Mandelbrot [1962, 1963] et de Fama [1965]. Toutefois, l'absence de formules explicites des densités de ces distributions a limité leur utilisation. Les distributions de cette classe ont un comportement asymptotique de Pareto avec  $0 < \alpha \leq 2$ . Lorsque  $\alpha = 2$  on retrouve la loi Normale, cependant pour  $\alpha < 2$  le moment d'ordre  $r \geq 2$  n'est pas fini, ces distributions ont donc une variance infinie, et par conséquent une queue très lourde.

Dans certaines situations, que l'on rencontre souvent dans divers domaines, telles les sciences sociales ou humaines, on est concerné plus par « l'homme extrême » de Pareto (cf. [Barbut, 2003]) que par « l'homme moyen ». En particulier, le cas pour la distribution parétienne : Barbut [1988] retrouve ainsi que « l'homme de Pareto » est bien celui des valeurs extrêmes, au contraire de « l'homme moyen » de Quetelet. Les théories probabilistes applicables sont dès lors fondées sur le théorème « des valeurs extrêmes » dont la principale caractéristique est l'absence d'échelles<sup>3</sup> connues et maîtrisées – situations, par exemple, dans lesquelles les événements sont d'autant plus chers qu'ils sont peu probables, ce qui signifie que l'espérance de la loi de probabilité est infinie.

L'abondante littérature et le succès croissant de la TVE dans de nombreux domaines nous a incité à mieux examiner l'apport de cette théorie pour l'analyse des risques encourus sur le marché pétrolier. En appliquant par exemple la TVE au marché pétrolier nous pouvons répondre aux types de questions suivantes : quelle est la probabilité d'occurrence d'un krach du type de celui de 1973 ou encore d'une baisse de 50 % en journée sur le brut, un mouvement n'ayant encore jamais été enregistré ? L'étude et la modélisation des ces extrêmes est d'importance capitale. En fait la performance d'un investisseur, d'une banque ou d'une entreprise sur une période donnée est souvent le fait de quelques journées exceptionnelles, la plupart des journées d'activité ne contribuant que marginalement au résultat. Ainsi on peut tenter de modéliser l'incidence de krachs pétroliers ou de tensions extrêmes sur les portefeuilles des investisseurs sur le marché.

---

<sup>2</sup> L'« homme moyen » d'une population est, selon Quetelet, un individu dont les caractéristiques physiologiques sont chacune égales à la moyenne des autres caractéristiques physiologiques de la population.

<sup>3</sup> L'absence d'échelle veut dire que l'on ne peut pas repérer facilement à quel niveau de risque l'on se trouve.

Les premières données dont nous avons besoin sont les historiques de cours Spot<sup>4</sup> de deux bruts internationaux de référence à savoir le Brent (brut de Mer du Nord, constitue la référence pour le marché européen) et le WTI (West Texas Intermediate, il est traité à New York au NYMEX<sup>5</sup>) (cf. Figure 1). Ces chroniques observées sur la période 1987-2008 montrent que les cours du brut ont subi, en particulier au cours de ces dernières années, de très fortes fluctuations. Ces mouvements spectaculaires ont ravivé l'intérêt porté à la question de la volatilité des marchés pétroliers. D'importantes fluctuations ont été enregistrées telles que le pic éphémère de fin 1990, suite à l'invasion du Koweït par l'Irak, et la chute drastique observée en 1998, consécutive à une erreur d'anticipation<sup>6</sup>. Entre 1999 et 2000, les cours du pétrole se sont inscrits en hausse, tirés par la forte croissance mondiale. Notons également une certaine période de stabilité des cours jusqu'en 2002, suivi une flambée des prix pétroliers qui s'est déclenchée en 2003 avec une accélération dès 2004 qui les ont fait atteindre les 78 \$ le baril avec des fluctuations de plus ou moins 5 à 10 % d'une semaine à l'autre en 2005. Néanmoins depuis septembre 2005, la hausse tendancielle des cours a marqué une pause mais le marché reste tendu. Ainsi le prix du baril de Brent est passé de 10 \$ au début de l'année 1999 à 44 \$ en janvier 2005 et 73 \$ en août 2006. En 2007, on a assisté à une accélération de la hausse, mais depuis le début de 2008 on constate une accélération assez rapide de la tendance haussière du prix moyen du baril de brut. Depuis le milieu du mois de mai de cette année 2008, la hausse du prix du pétrole brut s'est dangereusement accélérée, le prix moyen du baril de pétrole est passé de 119 \$ le 16 mai à 140 \$ le 28 juin, soit une augmentation de 15 % sur une période de 6 semaines. Plus récemment, la crise sur les marchés financiers « crise de liquidité » déclenchée en septembre 2008 a plutôt eu pour effet de provoquer la chute des prix énergétiques, en anticipation d'un fort ralentissement de la demande mondiale sur fond de récession économique. Les prix du brut ont ainsi perdu plus de la moitié de leur valeur en moins de quatre mois : ils évoluaient à environ 65 \$ le baril, en octobre, après être montés à plus de 147 \$ début juillet 2008.

La Figure 2 montre les chroniques des rendements journaliers, définis par  $r_t = 100 * \log(p_t) - \log(p_{t-1})$  où  $p_t$  est le cours du baril au jour  $t$ . Un examen rapide de ce graphique permet de constater que les rendements sont caractérisés par des périodes de faibles turbulences suivies par d'autres de fortes turbulences. On assiste ainsi à un regroupement des extrêmes en classes « clusters » ou paquets de volatilités. Ce phénomène, typique des données financières et macroéconomiques lorsque les fréquences sont élevées, remet en cause l'hypothèse d'homoscédasticité et montre l'existence d'effets ARCH. L'année 1991 fait apparaître un pic qui coïncide avec la guerre du Golfe, synonyme d'une forte volatilité. D'autres pics peuvent être également constatés surtout au cours de la dernière décennie.

---

<sup>4</sup> C'est le cours du marché physique au comptant, c'est-à-dire celui des transactions physiques à livraison immédiate ou quasi-immédiate, compte tenu des délais d'acheminement des produits. Les principaux opérateurs sur ce marché Spot sont les compagnies productrices privées et publiques, les raffineurs et les négociants.

<sup>5</sup> New York Mercantile Exchange

<sup>6</sup> L'OPEP (Organisation des pays exportateurs de pétrole) a augmenté ses quotas dans un contexte de baisse de la demande sur fond de crise financière Sud-Est Asiatique qui a entraîné une chute brutale des prix jusqu'en février 1999 où ils atteignent 10 dollars le baril.

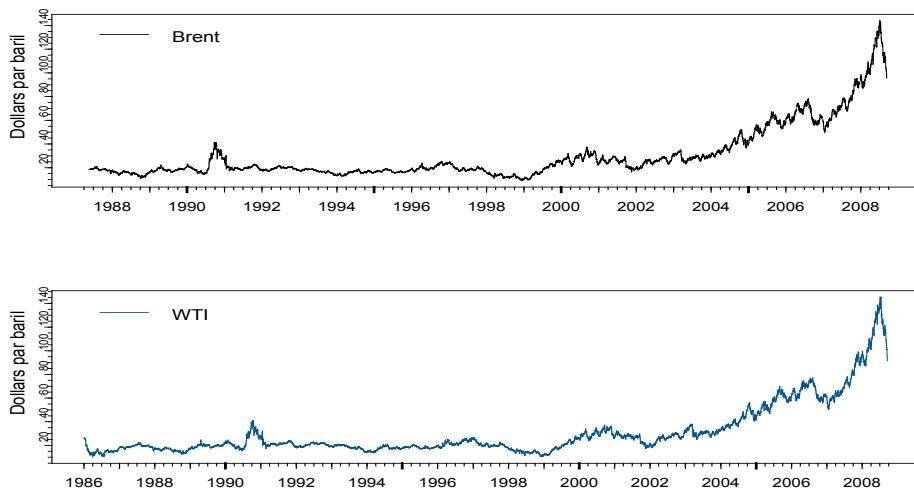


FIGURE 1. Cours Spot du baril Brent et WTI

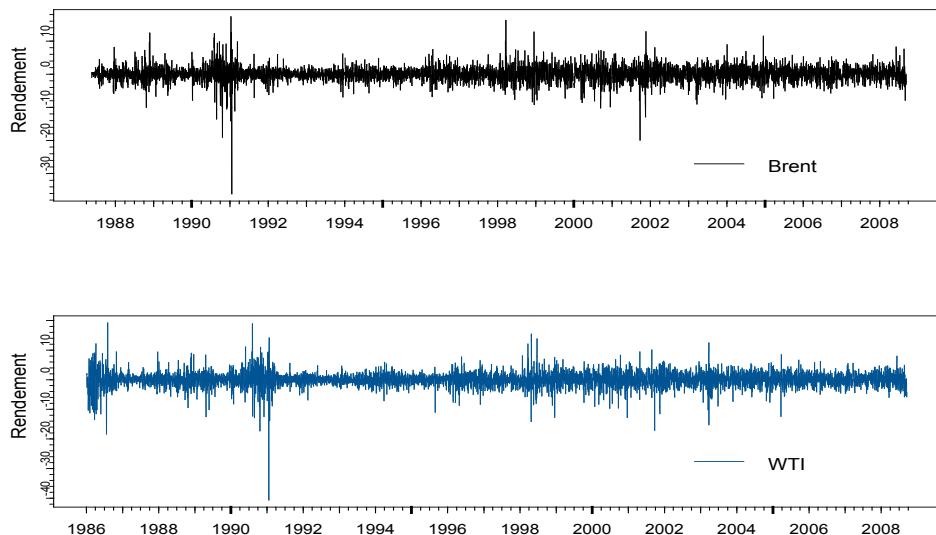


FIGURE 2. Évolutions des rendements Brent et WTI.

Cet aperçu historique montre que les prix du pétrole sont fondamentalement volatiles dans la mesure où ils sont fortement dépendants des aléas politiques, climatiques et économiques : la moindre variation de ces aléas peut provoquer des changements brusques sur le niveau des prix. Cela est dû au fait que l'offre et la demande sont peu élastiques aux prix, de sorte que de petites variations de l'une ou de l'autre peuvent causer de grandes fluctuations des prix.

De façon assez répétitive les activités de marché sont donc affectées par des mouvements relativement violents et auxquels on attribue les qualificatifs de crise ou de krack. L'outil de mesure des risques de marché communément reconnu, surtout en finance, est la Value at Risk (VaR). La VaR, ou montant exposé au risque, définit pour

une période et un niveau de confiance pré-spécifiés, la diminution potentielle maximale de la valeur d'un portefeuille ou d'un ensemble des portefeuilles. Elle a été introduite pour répondre à une question assez simple mais extrêmement précise : « Pour une valeur de probabilité donnée, à combien peut s'élever la perte potentielle associée à un portefeuille donné, au cours d'un horizon T fixé (1 jour, 10 jours,...) ? ». Malheureusement, par construction, la VaR ne permet pas d'appréhender les conditions extrêmes des marchés et peut présenter un aspect faussement rassurant pour un intervenant sur le marché, notamment lorsque le passé récent n'a enregistré aucun accident particulier et les approches usuelles de la statistique classique prennent difficilement en compte les variations extrêmes de prix observées sur les marchés. L'estimation<sup>7</sup> de la VaR devra donc passer par l'utilisation des techniques issues de la théorie des extrêmes qui s'est particulièrement développée depuis le milieu des années 70 (parmi d'autres, on peut citer les travaux de [Pickands, 1975 ; Hill, 1975 ; Smith, 1987 ; ou encore Dekkers *et al.*, 1989]). Récemment, cette théorie a été rapidement appliquée surtout aux problématiques financières et assurantielles (parmi d'autres, [McNeil, 1997; McNeil et Frey, 2000]).

La théorie des valeurs extrêmes permet d'évaluer les événements rares et les pertes associées à leur apparition. En d'autres termes lorsqu'une perte importante survient, cette théorie permet d'en évaluer l'ampleur. Les résultats théoriques sur le comportement stochastique des extrêmes d'échantillon qu'offre la théorie des valeurs extrêmes permettent de proposer un cadre mathématique rigoureux pour réaliser de telles extrapolations. De plus cette théorie acquiert une importance particulière du fait qu'elle s'intéresse directement à la queue de la loi. En fait seules les données extrêmes sont utilisées pour estimer les paramètres des modèles de la TVE ce qui assure un meilleur ajustement du modèle à la queue de la loi et donc une meilleure estimation de la VaR.

Notre étude va porter sur l'exploration de l'apport de la théorie des valeurs extrêmes pour la gestion des risques sur le marché pétrolier. Afin d'évaluer un tel apport les modèles de la TVE seront comparés avec les modèles les plus fréquemment rencontrés dans les travaux sur la VaR. De plus, cette contribution est destinée à combler le déficit de travaux dédiés à la mesure et à l'analyse quantitative des risques sur le marché du brut, en terme de la VaR. Les résultats indiquent que le modèle GPD (Generalized Pareto Distribution) conditionnel est le plus performant comparé aux autres modèles utilisés. Le modèle GARCH à erreurs Student (GARCH-T) peut aboutir à des résultats également satisfaisants. Ces deux modèles produisent une VaR qui reflète l'aspect dynamique de la volatilité et ils sont capables de capturer la lourdeur des queues. La VaR Normale est une mesure qui nous semble inadaptée pour évaluer le risque extrême des rendements du brut dont les queues de loi sont les plus épaisses. Une analyse de sensibilité des résultats trouvés par rapport au choix de la fenêtre d'estimation montre qu'en général, il y a un arbitrage entre le biais et la robustesse du modèle, et que la performance de l'ensemble des modèles se détériore pour des fenêtres de tailles inférieures à 1000 observations.

Cet article est organisé comme suit : la Section 2 est un rappel des théories probabilistes des valeurs extrêmes dans le cas des variables aléatoires réelles indépendantes et identiquement distribuées (i.i.d) avec des exemples simples. La

---

<sup>7</sup> Dans la littérature, trois grandes familles de méthodes ont été généralement adoptées pour estimer la VaR à savoir : les approches paramétriques (RiskMetrics et GARCH), non paramétriques (simulation historique) et récemment les approches basées sur la théorie des valeurs extrêmes. Pour une discussion très complète sur les différentes approches d'estimation de la VaR, nous renvoyons à l'article de Manganelli *et al.* [2001].

troisième section présente un aperçu sur la procédure d'estimation des paramètres et des quantiles de la GPD. L'illustration de l'approche basée sur la TVE, utilisée pour la gestion des fluctuations extrêmes sur le marché pétrolier, est présentée dans la Section 4.

## 2. THÉORIE DES VALEURS EXTRÊMES : QUELQUES RÉSULTATS THÉORIQUES

### 2.1. LOI GÉNÉRALISÉE DES VALEURS EXTRÊMES

La théorie des valeurs extrêmes a pour but d'étudier la loi du maximum d'une suite des variables aléatoires réelles même si, et spécialement si, la loi du phénomène n'est pas connue. Formellement, considérons  $(X_1, X_2, \dots, X_n)$  une suite de  $n$  variables aléatoires (v.a.) indépendantes et identiquement distribuées (i.i.d) de fonction de répartition  $F$  définie par

$$F(x) = \Pr(X_i \leq x) \quad \text{pour } i = 1, \dots, n \quad (1)$$

Pour étudier le comportement extrême des événements, on considère la variable aléatoire  $M_n = \max(x_1, x_2, \dots, x_n)$  le maximum<sup>8</sup> d'un échantillon de taille  $n$ .

Comme les variables aléatoires sont i.i.d., alors la fonction de répartition de  $M_n$  est donnée par

$$F_{M_n}(x) = \Pr(M_n \leq x) = (F(x))^n \quad (2)$$

La formule (2) présente un intérêt très limité<sup>9</sup>. De plus, la loi d'une variable aléatoire parente  $X$  est rarement connue avec précision et, même si la loi de cette variable parente  $X$  est connue avec exactitude, la loi du terme maximum n'est pas toujours facilement calculable. Pour ces raisons, il est intéressant de considérer les comportements asymptotiques du maximum convenablement normalisé.

Un des résultats fondamentaux de la TVE est le théorème suivant établi en 1928 par Fisher et Tippett.

**DÉFINITION 1** (lois de même type). *On dit que deux variables aléatoires réelles  $X$  et  $Y$  sont de même type s'il existe des constantes réelles  $a > 0$  et  $b \in \mathbb{R}$  tels que  $Y = aX + b$  i.e. Si  $F$  et  $H$  sont des lois respectives des variables  $X$  et  $Y$  alors on a  $F(ax + b) = H(x)$ .*

*Autrement dit, les variables « de même type » ont la même loi, à un facteur de localisation et d'échelle près.*

De façon analogue au théorème central limite, peut-on trouver des constantes de normalisation :  $a_n$  et  $b_n$  avec  $a_n > 0$  et  $b \in \mathbb{R}$  et une loi non-dégénérée de loi  $H$  telle que :

---

<sup>8</sup> On obtient la correspondance entre minimum et maximum par la relation suivante :  $\min(X_1, X_2, \dots, X_n) = -\max(-X_1, -X_2, \dots, -X_n)$ . Ainsi, tous les résultats que nous allons présenter pour les maxima pourront être transposés pour les minima.

<sup>9</sup> La distribution asymptotique du maximum, déterminée en faisant tendre  $n$  vers l'infini, donne une loi dégénérée. En effet  $\lim_{n \rightarrow +\infty} F_{M_n}(x) = \begin{cases} 0 & \text{si } F(x) < 1 \\ 1 & \text{si } F(x) = 1 \end{cases}$ .

$$\Pr\left\{\frac{M_n - b_n}{a_n} \leq x\right\} = (F(a_n x + b_n))^n \rightarrow H(x) \quad (3)$$

Fisher et Tippett trouvent en 1928 une solution à ce problème au moyen d'un théorème qui porte leur nom et qui est l'un des fondements de la théorie des valeurs extrêmes.

**THÉORÈME 1** (de Fisher-Tippett ou théorème des 3 types extrêmes). *S'il existe deux suites de constantes de normalisation avec  $a_n > 0$  et  $b \in \mathbb{R}$  et une loi non-dégénérée de loi  $H$  telle que :  $\frac{M_n - b_n}{a_n} \xrightarrow{\text{loi}} H$ , alors  $H$  est l'une des trois lois limites :*

$$\Lambda(x) = \exp\{-\exp[-x]\}, \quad x \in \mathbb{R} \quad (\text{distribution de Gumbel})$$

$$\Phi_\alpha(x) = \begin{cases} \exp\{-x^{-\alpha}\}, & x > 0 \\ 0, & x \leq 0 \end{cases} \quad (\text{distribution de Fréchet})$$

$$\Psi_\alpha(x) = \begin{cases} \exp\{-(-x)^\alpha\}, & x \leq 0 \\ 1 & \text{si } x > 0 \end{cases} \quad (\text{distribution de Weibull})$$

Ce théorème<sup>10</sup> donne un résultat très intéressant : quelle que soit la loi limite de la variable parente, la loi limite des extrêmes a toujours la même forme. Bien que le comportement de ces lois soit complètement différent, elles peuvent être combinées en une seule paramétrisation contenant un unique paramètre  $\xi$  qui contrôle la « lourdeur » de la queue de loi appelé indice des valeurs extrêmes (ou indice de queue) :

$$H_\xi(x) = \begin{cases} \exp\left(-(1 + \xi x)^{-\frac{1}{\xi}}\right), & \xi \neq 0, \quad 1 + \xi x > 0 \\ \exp(-\exp(-x)), & \xi = 0, \quad -\infty \leq x \leq +\infty \end{cases} \quad (4)$$

où  $H$  est une fonction de répartition non-dégénérée. Cette loi est appelée loi de valeurs extrêmes généralisée (Generalized Extreme Value), que l'on notera GEV. En introduisant les paramètres de localisation<sup>11</sup>  $\mu$  et de dispersion<sup>12</sup>  $\sigma$  dans la paramétrisation (4), on obtient la forme la plus générale de la GEV :

$$H_{\xi, \mu, \sigma}(x) = \exp\left(-\left(1 + \xi \frac{x - \mu}{\sigma}\right)^{-\frac{1}{\xi}}\right), \quad \xi \neq 0, \quad 1 + \xi \frac{x - \mu}{\sigma} > 0$$

Le théorème de Fisher-Tippett fournit, en quelque sorte, la contrepartie du Théorème Central Limite (TCL) dans le cas d'événements extrêmes. Cependant, contrairement au TCL, où la loi normale est la seule loi limite possible, dans le cas des extrêmes, trois types de loi limite sont possibles (cf. Figure 3) :

Le cas  $\xi = 0$  correspond à la loi Gumbel.

<sup>10</sup> Pour la preuve, nous renvoyons à Embrechts *et al.*, 1997, chapitre 3, p. 122.

<sup>11</sup> Un paramètre de localisation (ou de position) est, comme son nom l'indique, un paramètre qui régit la position d'une densité de probabilité. Si ce paramètre (scalaire ou vectoriel) est noté  $\mu$ , la densité se présente formellement comme :  $f_\mu(x) = f(x - \mu)$ .

<sup>12</sup> Un paramètre de dispersion (ou d'échelle) est un paramètre qui régit l'aplatissement d'une famille paramétrique de lois de probabilités. Il dirige l'échelle ou encore la dispersion de la distribution.

Le cas  $\xi > 0$  correspond à la loi de Fréchet de paramètre  $\alpha = \frac{1}{\xi}$ .

Le cas  $\xi < 0$  correspond à la loi de Weibull de paramètre  $\alpha = -\frac{1}{\xi}$ .

Quand  $\xi = 0$ , la loi de  $X$  présente une décroissance de type exponentiel dans la queue de la loi, on dit que l'on est alors dans le domaine d'attraction<sup>13</sup> de Gumbel (c'est le cas par exemple de la loi normale, exponentielle, gamma ou lognormale) ; le domaine de Fréchet, correspondant à une loi de  $X$  non bornée, est le domaine où  $\xi > 0$  (c'est le cas par exemple de la loi de Cauchy ou Pareto) et enfin quand  $\xi < 0$ , la loi de  $X$  est bornée et on dit que l'on est dans le domaine d'attraction de Weibull (c'est le cas par exemple de la loi uniforme ou bêta).

La loi la plus courante rencontrée sur des données financières ou macroéconomiques est la loi de Fréchet. La loi de Weibull étant écartée car une loi de Weibull ne peut être rare or les variations en finance ou en économie sont généralement non bornées dont la mesure où on peut toujours rencontrer des variations extrêmes. La loi de Gumbel est extrêmement rare également. Néanmoins, cette dernière loi sert de référence et permet ainsi de mesurer, notamment graphiquement, l'écart entre les lois empiriques et la loi normale.

L'approche basée sur la GEV a été critiquée dans la mesure où l'utilisation d'un seul maxima conduit à une perte d'information contenue dans les autres grandes valeurs de l'échantillon. Pour pallier ce problème, la méthode POT (Peaks-over-Threshold) ou méthode des excès au-delà d'un seuil élevé a été introduite par Pickands [1975].

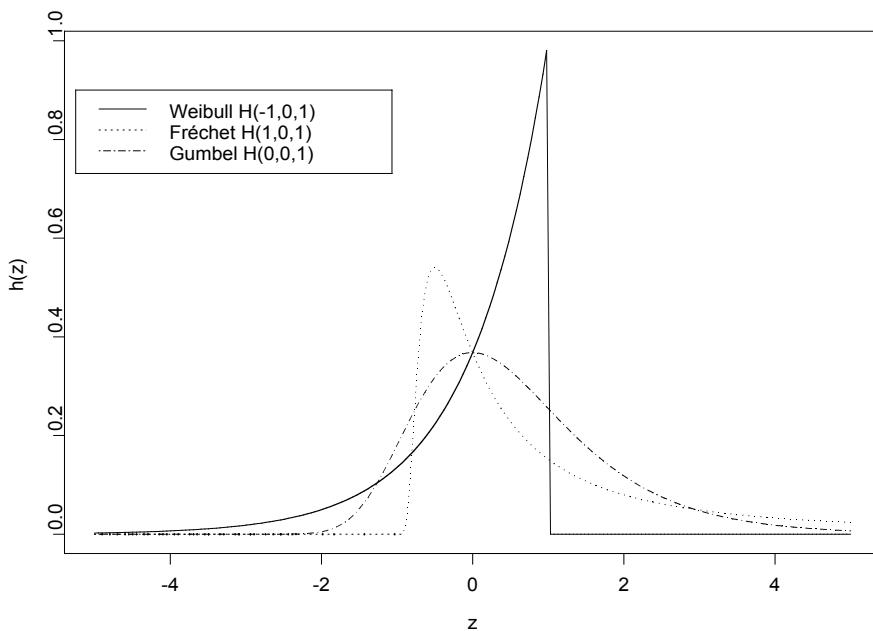


FIGURE 3. Densité des lois de valeurs extrêmes, avec  $\xi = 1$  pour la loi de Fréchet,  $\xi = 0$  pour la loi de Gumbel et  $\xi = -1$  pour la loi de Weibull.

<sup>13</sup> On appelle domaine d'attraction de  $G$  (ou domaine d'attraction maximal), l'ensemble des lois  $F$  pour lesquelles le maximum normalisé suit asymptotiquement la loi  $G$ .

## 2.2. EXCÈS AU-DELÀ D'UN SEUIL

La méthode des excès au-delà d'un seuil (ou Peak Over Threshold, POT) repose sur le comportement des valeurs observées au-delà d'un seuil donné. En d'autres termes, elle consiste à observer non pas le maximum ou les plus grandes valeurs mais toutes les valeurs des réalisations qui excèdent un certain seuil élevé. L'idée de base de cette approche consiste à choisir un seuil suffisamment élevé et à étudier les excès au-delà de ce seuil. Cette méthode initialement développée par Pickands [1975] et abondamment étudiée par divers auteurs tels que de Smith [1987], Davison et Smith [1990], ou Reiss et Thomas [2001] pour de plus amples références.

On définit un seuil  $u \in \mathbb{R}$ ,  $N_u = \text{card}\{i : i = 1, \dots, n, X_i > u\}$  et  $Y_i = X_i - u > 0$  pour  $1 \leq j \leq N_u$  où  $N_u$  est le nombre des dépassements du seuil  $u$  par les  $(X_i)_{1 \leq i \leq n}$  et  $Y_1, \dots, Y_{N_u}$  les excès correspondants (cf. Figure 4).

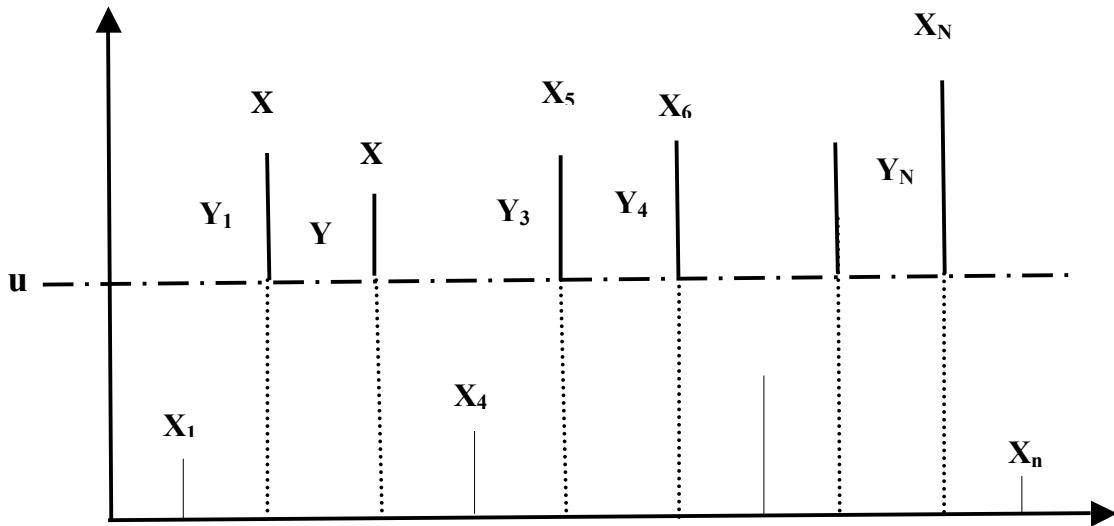


FIGURE 4. Méthode des excès :  $u$  réel suffisamment élevé appelé seuil,  
 $Y$  : excès de  $X$  au-delà de  $u$ .

On cherche à partir de la loi  $F$  de  $X$  à définir une loi conditionnelle  $F_u$  par rapport au seuil  $u$  pour les variables aléatoires dépassant ce seuil.

On définit alors la loi conditionnelle des excès  $F_u$  par :

$$F_u(y) = \Pr(X - u \leq y \mid X > u) = \frac{F(u + y) - F(u)}{1 - F(u)}, \quad y \geq 0 \quad (5)$$

Le théorème de Pickands-Balkema-de Haan ci-après donne la forme de la loi limite pour les valeurs extrêmes : sous certaines conditions de convergence, la loi limite est une loi de Pareto généralisée que l'on notera GPD.

**THÉORÈME 2** (de Pickands-Balkema-de Haan). *Une fonction de répartition  $F$  appartient au domaine d'attraction maximale de  $H_\xi$ , si et seulement si, il existe une fonction positive  $\beta(u)$  telle que :*

$$\lim_{u \rightarrow x_F} \sup_{0 \leq y \leq x_F - u} |F_u(y) - G_{\xi, \beta(u)}(y)| = 0 \quad (6)$$

où  $F_u(y)$  est la fonction de répartition conditionnelle des excès pour  $u$  élevé,  $x_F$  est le point terminal de  $F$ ,  $x_F = \{x \in \mathbb{R} : F(x) < 1\}$  et  $G_{\xi, \beta(u)}(y)$  est la GPD donnée par

$$G_{\xi, \beta(u)}(y) = \begin{cases} 1 - \left(1 + \xi \frac{y}{\beta(u)}\right)^{-\frac{1}{\xi}}, & \xi \neq 0 \\ 1 - \exp\left(-\frac{y}{\beta(u)}\right), & \xi = 0 \end{cases} \quad (7)$$

où :  $y \geq 0$  pour  $\xi \geq 0$  et  $0 \leq y \leq -\frac{\beta(u)}{\xi}$  pour  $\xi < 0$

Ce théorème montre l'existence d'une relation étroite entre la GPD et la GEV ; Pickands [1975] a montré que pour n'importe quelle loi  $F$ , l'approximation GPD définie par (6) n'est vérifiée que s'il existe des constantes de normalisation et une loi non-dégénérée telle que le résultat donné par (3) est vérifié. Dans ce cas, si  $H$  est écrite sous la forme d'une GEV équation (4), alors l'indice de queue  $\xi$  est le même que celui de la GPD donnée dans (7).

De même pour la GPD, le cas où  $\xi > 0$  correspond aux lois à queues épaisses, pour lequel  $1 - G$  se comporte comme une puissance,  $x^{-\frac{1}{\xi}}$  pour  $x$  assez élevé. Si  $\xi = 0$ , on trouve  $1 - \exp\left(-\frac{y}{\beta(u)}\right)$ , c'est la loi exponentielle de paramètre  $\beta$  et enfin pour  $\xi < 0$  c'est la loi de Pareto de type II à support borné (cf. Figure 5).

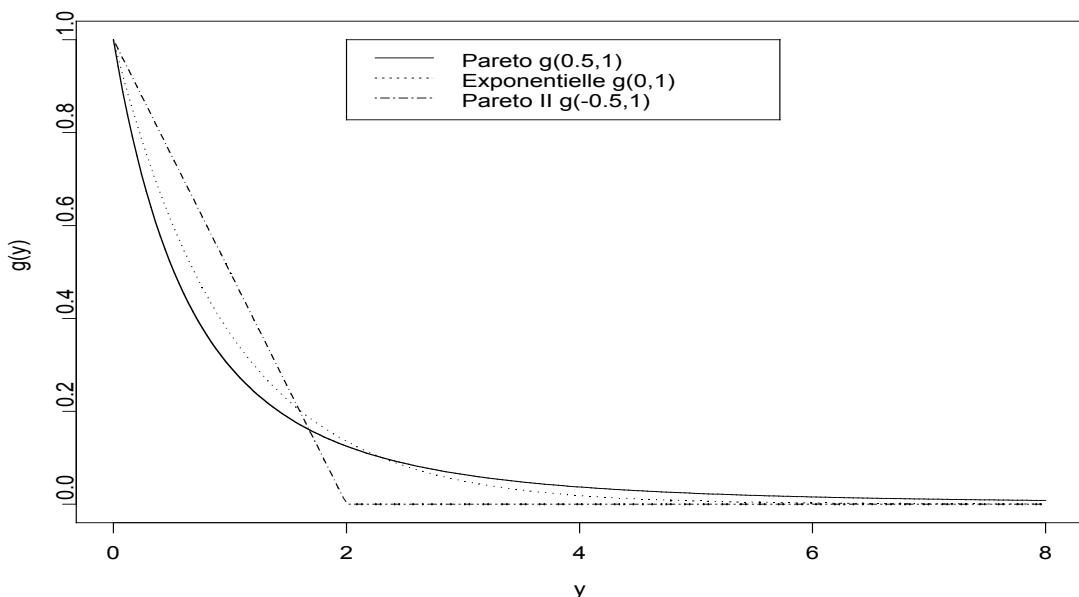


FIGURE 5. Densité des lois de valeurs extrêmes, avec  $\xi = 0,5$  pour la loi de Pareto,  $\xi = 0$  pour la loi exponentielle et  $\xi = -0,5$  pour la loi de Pareto II.

La GPD a les propriétés suivantes :

$$E(Y) = \frac{\beta}{1-\xi}, \quad (\xi < 1) \quad \text{et} \quad V(Y) = \frac{\beta^2}{(1-\xi)^2(1-2\xi)}, \quad (\xi < \frac{1}{2}).$$

En pratique le choix du seuil constitue une difficulté. En fait,  $u$  doit être assez grand pour que l'approximation GPD définie par (6) soit valide, mais pas trop élevé pour garder un nombre suffisant de dépassements pour estimer les paramètres du modèle. Le seuil doit être choisi de façon à faire un arbitrage, traditionnel en statistiques, entre le biais et la variance. Généralement,  $u$  est déterminé graphiquement en exploitant la linéarité de la fonction<sup>14</sup> d'excès moyenne  $e(u)$  pour la GPD (cf. [Embrechts *et al.*, 1997]). Cette technique fournit une aide précieuse, cependant, il ne faut pas attendre d'elle la bonne valeur de  $u$ . En pratique, plusieurs valeurs de  $u$  doivent être testées. Ce problème du choix a suscité de nombreux travaux dans la littérature. Beirlant *et al.* [1996] ont suggéré de choisir le seuil  $u$  qui minimise l'erreur quadratique moyenne asymptotique de l'estimateur de Hill de l'indice de queue, tout en supposant que  $F$  appartient au domaine d'attraction maximale de Fréchet. Parmi d'autres, Hall [1990] ; Danielsson *et al.* [2001], ont introduit des approches Bootstrap pour trouver le seuil optimal. Une autre méthode utilisée par plusieurs auteurs (cf. par exemple, [McNeil et Frey, 2000 ; Raggad, 2007 ; Marimoutou *et al.*, 2009]) consiste à choisir un seuil aléatoire. En fait, ils prennent  $k$  observations excédentaires et le  $k$  peut être déterminé par des méthodes Monte-Carlo. Pour une discussion des principales approches, récemment développées, permettant le choix du seuil optimal, nous renvoyons à Matthys et Beirlant [2000].

### 2.3. EXEMPLES DE COMPORTEMENTS LIMITES DANS LES QUEUES

Dans cette section, nous proposons trois exemples illustrant comment se manifestent les lois limites de la GEV et la GPD en pratique, compte tenu de différentes hypothèses sur la loi  $F$  de la variable parente  $X$  de l'échantillon.

**EXEMPLE 1 : Cas de la loi exponentielle**

Pour la loi exponentielle de paramètre  $\lambda = 1$ , la fonction de répartition de cette loi est  $F(x) = 1 - e^{-x}$  pour  $x \geq 0$ . En posant  $b_n \log(n)$  et  $a_n = 1$  alors,

$$\begin{aligned} F^n(a_n x + b_n) &= \left(1 - e^{-x-\log n}\right)^n \\ &= \left(1 - \frac{e^{-x}}{n}\right)^n \\ &\rightarrow \exp\left\{-e^{-x}\right\} \\ &= \Lambda(x) \end{aligned}$$

c'est-à-dire que le maximum convenablement normalisé de la loi exponentielle converge vers la loi de Gumbel. Cette loi appartient au domaine d'attraction maximal de Gumbel.

---

<sup>14</sup> Pour une v.a. qui suit une  $G_{\xi\beta}$  de paramètre  $\xi < 1$ ,  $e(u)$  est linéaire avec  $u$  :

$$e(u) = E(X - u | X > u) = \frac{\beta + \xi u}{1 - \xi}, \quad (\xi < 1).$$

Aussi, les lois dans le domaine d'attraction de la loi de Gumbel sont parfois appelées lois de type exponentiel.

En ce qui concerne la méthode des excès au-delà d'un seuil  $u$ , en prenant  $\beta_u = 1$ , alors

$$\begin{aligned} F_u(y) &= \frac{F(u+y) - F(u)}{1 - F(u)}, \quad y \geq 0 \\ &= \frac{e^{-u} - e^{-u-y}}{e^{-u}} \\ &= 1 - e^{-y} \end{aligned}$$

pour tout  $y > 0$ . Aussi, la loi limite est la loi GPD de paramètre  $\xi = 0$  avec  $\beta_u = 1$ . Notons que dans ce cas, la loi GPD n'est pas simplement la loi limite, mais il s'agit de la loi exacte pour tout  $u$ .

#### EXEMPLE 2 : Cas de la loi de Pareto

Pour la loi de Pareto de fonction de répartition  $F(x) = 1 - cx^{-\alpha}$ , où  $c > 0$  et  $\alpha > 0$ . En posant,  $b_n = 0$  et  $a_n = (nc)^{\frac{1}{\alpha}}$  alors on a pour  $x > 0$

$$\begin{aligned} F^n(a_n x + b_n) &= (1 - c(a_n x)^{-\alpha})^n = (1 - ca_n^{-\alpha} x^{-\alpha})^n \\ &= \left(1 - \frac{x^{-\alpha}}{n}\right)^n \\ &\rightarrow \exp\left\{-x^{-\alpha}\right\} \\ &= \Phi_\alpha(x) \end{aligned}$$

qui est la loi de Fréchet. La loi de Pareto appartient au domaine d'attraction maximal de Fréchet. Aussi, les lois dans le domaine d'attraction de la loi de Fréchet sont parfois appelées lois de type Pareto.

En se basant sur la méthode des excès au-delà d'un seuil  $u$  et en considérant  $\beta_u = ub$  pour  $b > 0$ , on a alors

$$\begin{aligned} F_u(y) &= \frac{F(u+uby) - F(u)}{1 - F(u)}, \quad y \geq 0 \\ &= \frac{cu^{-\alpha} - c(u+uby)^{-\alpha}}{cu^{-\alpha}} \\ &= 1 - (1+by)^{-\alpha} \end{aligned}$$

Pour  $\xi = \frac{1}{\alpha}$  et  $b = \xi$ , la limite est alors la loi GPD de paramètre  $\xi$ .

#### EXEMPLE 3 : Cas de la loi normale

Soit  $F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{y^2}{2}} dy$  la fonction de répartition de la loi  $N(0,1)$ . D'après Feller [1968],  $1 - F(x)$  est équivalente à  $\frac{1}{x\sqrt{2\pi}} e^{-\frac{x^2}{2}}$  au voisinage de l'infini, donc

$$1 - F(x) \sim \frac{1}{x\sqrt{2\Pi}} e^{-\frac{x^2}{2}} \text{ quand } x \rightarrow +\infty, \text{ alors :}$$

$$\lim_{u \rightarrow +\infty} \frac{1 - F\left(u + \frac{z}{u}\right)}{1 - F(u)} = \lim_{u \rightarrow +\infty} \left[ \left(1 + \frac{z}{u^2}\right)^{-1} \exp\left\{-\frac{1}{2}\left(u + \frac{z}{u}\right)^2 + \frac{1}{2} u^2\right\} \right] = e^{-z}$$

Si on suppose dans un premier temps que  $\beta_u = \frac{1}{u}$ , alors

$$1 - \frac{1 - F\left(u + \frac{z}{u}\right)}{1 - F(u)} = \frac{F(u + \beta_u z) - F(u)}{1 - F(u)} \rightarrow 1 - e^{-z} \text{ quand } u \rightarrow +\infty$$

et par la suite la loi limite des excès au-delà d'un seuil est la loi exponentielle.

Dans un second temps si on considère  $b_n$  solution de  $F(b_n) = 1 - \frac{1}{n}$  et  $a_n = \frac{1}{b_n}$ , on obtient

$$n\{1 - F(a_n x + b_n)\} = \frac{\{1 - F(a_n x + b_n)\}}{1 - F(b_n)} \rightarrow e^{-x}$$

et donc

$$\lim_{n \rightarrow +\infty} F^n(a_n x + b_n) = \lim_{n \rightarrow +\infty} \left(1 - \frac{e^{-x}}{n}\right)^n = \exp\{-e^{-x}\}$$

converge vers la loi de Gumbel.

Comme il a été remarqué par Smith [2003], bien que la loi de Gumbel et la loi exponentielle soient respectivement les limites exactes des maxima et des excès au-delà d'un seuil, les meilleures approximations sont obtenues en utilisant la loi GEV et la loi GPD. Ces lois sont très générales et permettent d'inclure même le cas où  $\xi \neq 0$ . En fait, il est généralement préférable de considérer ces deux lois même si on pense que la loi de Gumbel ou la loi exponentielle sont les vraies lois limites.

## 2.4. LES $R$ PLUS GRANDES VALEURS

Dans certaines situations, les seuls maxima ne contiennent pas assez d'informations pour estimer les paramètres du modèle GEV. Le résultat mathématique sur lequel est fondée l'équation (3) peut être généralisé à la loi jointe des  $R$  plus grandes valeurs quand  $n \rightarrow +\infty$ . L'inconvénient majeur de cette approche est qu'elle est peu robuste vis-à-vis de la violation de l'hypothèse i.i.d (lorsqu'il y a des variations saisonnières ou des observations dépendantes) que les résultats classiques (cf. Section 2.1) de la TVE.

Considérons  $X_{n,1} \geq X_{n,2} \geq \dots \geq X_{n,R}$  les  $R$  plus grandes valeurs d'un échantillon de taille  $n$ , et  $a_n$  et  $b_n$  sont les constantes de normalisation définies par (3), alors  $\left(\frac{X_{n,1} - b_n}{a_n}, \dots, \frac{X_{n,R} - b_n}{a_n}\right)$  converge en loi vers le vecteur aléatoire  $(Y_1, \dots, Y_R)$ , de densité jointe

$$h(Y_1, \dots, Y_R) = \sigma^{-R} \exp \left\{ - \left( 1 + \xi \frac{y_R - \mu}{\sigma} \right)^{-\frac{1}{\xi}} - \left( 1 + \frac{1}{\xi} \right) \sum_{j=1}^R \log \left( 1 + \xi \frac{y_j - \mu}{\sigma} \right) \right\}$$

où  $\sigma$  est un paramètre de dispersion. Il est à noter que la vitesse de convergence de la densité jointe des  $R$  plus grandes valeurs décroît rapidement lorsque  $R$  augmente. Ainsi  $R$  devra être choisi suffisamment grand pour avoir assez d'information mais pas trop afin d'éviter le risque de prendre en compte des valeurs qui ne sont pas forcément dans les queues ce qui conduit à un biais dans l'estimation. Pour plus des détails, nous renvoyons à Smith [1986].

### 3. ESTIMATION STATISTIQUE

En se référant à la littérature, on note diverses méthodes qui ont été proposées pour estimer les paramètres des lois GEV et GPD. On trouve la méthode basée sur le maximum de vraisemblance [Smith, 1987], la méthode des moments pondérées [Greenwood *et al.*, 1979], la méthode des moments [Christopeit, 1994] ou encore par des méthodes bayésiennes [Lye *et al.*, 1993]. Il existe également des approches non paramétriques pour l'estimation de l'indice de queue. Les plus utilisées en pratique sont l'estimateur de Pickands [1975], l'estimateur de Hill [1975] (pour le cas des lois de type Fréchet uniquement) et l'estimateur de Dekkers-Einmahl-de Hann [Dekkers *et al.*, 1989]. La méthode la plus populaire et qui sous certaines conditions est la plus efficace est la méthode de maximum de vraisemblance.

Dans ce qui suit, nous allons présenter dans un premier temps cette dernière méthode d'estimation paramétrique pour la GPD. Par la suite, nous présenterons l'estimation de la VaR selon une approche basée sur la méthode POT. Pour un exposé plus complet sur les méthodes d'estimation des modèles de la TVE, nous renvoyons à Embrechts *et al.* [1997].

#### 3.1 ESTIMATION DES PARAMÈTRES DE LA GPD PAR MAXIMUM DE VRAISEMBLANCE

Considérons à nouveau la loi GPD dont l'expression analytique est donnée par l'équation (7). La densité bi-paramétrique de la GPD est donnée par :

$$g(y) = \frac{1}{\beta} \left( 1 + \xi \left( \frac{y}{\beta} \right) \right)^{-\frac{1}{\xi}-1} \quad (8)$$

pour  $y \geq 0$  si  $\xi \geq 0$ , et  $0 \leq y \leq -(\beta/\xi)$  si  $\xi < 0$ .

L'estimation de la GPD, par la méthode de maximum de vraisemblance, porte sur l'indice de queue  $\xi$  ainsi que sur le paramètre d'échelle  $\beta$ . L'expression de la log-vraisemblance est donc :

$$\ln(\beta, \xi; Y) = -N_u \ln \beta - \left( \frac{1}{\xi} + 1 \right) \sum_{i=1}^{N_u} \ln \left( 1 + \frac{\xi}{\beta} y_i \right) \quad (9)$$

pour un échantillon  $Y_1, \dots, Y_{N_u}$  des excès. À partir d'ici on peut dériver par rapport à chaque paramètre, obtenir les équations à résoudre, et dégager l'estimateur de maximum de vraisemblance (EMV)  $\hat{\theta}_{ML}$  de  $\theta = (\beta, \xi)$ .

Pour  $\xi > -0,5$ , Smith [1985, 1987], Hosking et Wallis [1987] prouvent que les conditions de régularités de la fonction de vraisemblance sont remplies et que l'estimateur du maximum de vraisemblance conduit à un estimateur non biaisé, asymptotiquement normal.

### 3.2 ESTIMATION DES QUANTILES EXTRÊMES

Rappelons que la loi des excès au-delà de  $u$  suffisamment élevé est :

$$F_u(y) = \Pr(X - u \leq y \mid X > u) = \frac{F(u + y) - F(u)}{1 - F(u)} = \frac{F(u + y) - F(u)}{\bar{F}(u)}$$

où  $\bar{F} = 1 - F$  est la fonction de survie.

Ceci peut être réécrit comme

$$\begin{aligned} F_u(y) \bar{F}(u) &= F(u + y) - F(u) = -1 + F(u + y) + 1 - F(u) \\ &= -[1 - F(u + y)] + [1 - F(u)] = -\bar{F}(u + y) + \bar{F}(u) \\ &\Leftrightarrow \\ \bar{F}(u + y) &= \bar{F}(u) - F_u(y) \bar{F}(u) = \bar{F}(u) [1 - F_u(y)] = \bar{F}(u) \bar{F}_u(y) \end{aligned}$$

Donc, on aura :

$$\bar{F}(u + y) = \bar{F}(u) \bar{F}_u(y) \quad (10)$$

Rappelons que l'équation (6) du théorème de Balkema et Haan [1974] et Pickands [1975] donne un résultat limite pour  $F_u(y)$ . Ceci nous amène à proposer l'approximation suivante, pour  $u$  assez grand :

$$F_u(y) \approx G_{\xi, \beta(u)}(y), \quad u \rightarrow \infty, \quad y \geq 0$$

ou encore

$$\bar{F}_u(y) \approx \bar{G}_{\xi, \beta(u)}(y), \quad u \rightarrow \infty, \quad y \geq 0 \quad (11)$$

où  $G_{\xi, \beta(u)}(y)$  est la GPD donnée par l'équation (7).

La relation (10) montre que l'estimation de la queue de  $F$  pourrait être obtenue en estimant  $\bar{F}(u)$  et  $\bar{F}_u(y)$  séparément. Un estimateur naturel pour  $\bar{F}(u)$  est la loi empirique

$$\hat{\bar{F}}(u) = \hat{\bar{F}}_n(u) = \frac{1}{n} \sum_{i=1}^n I_{\{X_i > u\}} = \frac{N_u}{n} \quad (12)$$

où  $I_{\{X_i > u\}}$  est une fonction indicatrice ayant l'expression suivante :

$$I_{\{X_i > u\}} = \begin{cases} 1 & \text{si } X_i > u, \quad \forall i = 1, \dots, n \\ 0 & \text{sinon} \end{cases}$$

et  $N_u$  étant le nombre des excès au-delà du seuil  $u$ .

D'autre part, l'approximation (11) permet de proposer un estimateur pour  $\bar{F}_u(y)$ , soit

$$\hat{\bar{F}}_u(y) = \hat{\bar{G}}_{\hat{\xi}, \hat{\beta}}(y) = \left(1 + \hat{\xi} \frac{y}{\hat{\beta}}\right)^{-\frac{1}{\hat{\xi}}} \quad (13)$$

pour des  $\hat{\xi} = \hat{\xi}(N_u)$  et  $\hat{\beta}(u) = \hat{\beta}(N_u)$  adéquats.

L'estimateur résultant pour la queue  $\bar{F}(u+y) = \bar{F}(x)$  (pour  $y > 0$ ) a donc la forme

$$\hat{\bar{F}}(u+y) = \hat{\bar{F}}(x) = \hat{\bar{F}}(u) F_u(y) = \frac{N_u}{n} \left(1 + \hat{\xi} \frac{y}{\hat{\beta}}\right)^{-\frac{1}{\hat{\xi}}} \quad (14)$$

Dans le cas de Gumbel et de Fréchet ( $\xi \geq 0$ ), la restriction au domaine pour (14) est  $y \geq 0$  du fait que l'on estime la partie supérieure de la queue de  $\bar{F}$ .

En inversant (14), on obtient l'estimateur du quantile extrême  $x_p$

$$\hat{x}_p = u + \frac{\hat{\beta}}{\hat{\xi}} \left[ \left( \frac{n}{N_u} (1-p) \right)^{-\frac{1}{\hat{\xi}}} - 1 \right], \quad \text{pour } p > F(u) \quad (15)$$

La VaR n'est rien d'autre qu'un quantile extrême calculé à partir de la loi asymptotique des extrema (Generalized Pareto Loi), obtenue en modélisant les pertes (ou les profits) extrêmes par la méthode des excès. Ainsi, la  $\text{VaR}_p$ , correspondant au modèle GPD inconditionnel, pour un horizon donné  $t$  et avec une certaine probabilité  $p$ , ou de façon équivalente à un niveau de confiance (typiquement 95 % ou 99 %)  $c = 1 - p$  est

$$\overline{\text{Var}}_p = u + \frac{\hat{\beta}}{\hat{\xi}} \left[ \left( \frac{n}{N_u} (1-p) \right)^{-\frac{1}{\hat{\xi}}} - 1 \right] \quad (16)$$

où  $\hat{\xi}$  et  $\hat{\beta}$  représentent les estimateurs des paramètres de la loi GPD s'ajustant au  $N_u$  excès de rendements au-delà d'un seuil  $u$ , délimitant le début de la queue gauche (ou droite) de la loi parente.

Le second modèle testé est celui de McNeil et Frey [2000]. Il s'agit du modèle GPD conditionnel qui permet de combler les inconvénients des modèles standards de la TVE.

La procédure d'estimation de cette approche conditionnelle est à deux étapes :

#### ÉTAPE 1 : Estimation d'un modèle de type GARCH pour les rendements

Les chroniques des rendements Brent et WTI (cf. Figure 2) sont caractérisées par le « clustering » ou regroupement de volatilité, à savoir les périodes de forte volatilité alternent avec les périodes de faible volatilité. Ce phénomène, que nous appelons aussi

l'hétéroscédasticité conditionnelle, est particulièrement fréquent dans les données financières. Or, pour ce type de chroniques, les modèles linéaires utilisés auparavant, tels que les processus autorégressifs moyenne mobile (ARMA), développés par Box et Jenkins durant les années 1970, sont limités : ils ne permettent pas la prise en compte des phénomènes de variabilité de la volatilité en fonction du temps et des mécanismes d'asymétrie. En effet, dans la modélisation ARMA, l'espérance conditionnelle varie au cours du temps alors que la variance conditionnelle ne change pas.

Face aux lacunes des représentations ARMA(p,q) surtout pour les problèmes financiers, Engle [1982] a proposé une nouvelle classe de modèles Autorégressifs conditionnellement hétéroscédastiques (ARCH) apte à capter le comportement de la volatilité dans le temps. Elle a été par la suite généralisée par Bollerslev [1986] pour obtenir la classe des modèles GARCH. Depuis ces travaux précurseurs, de nombreuses spécifications de la volatilité ont été proposées. Pour une revue assez complète sur le sujet nous renvoyons à l'ouvrage de Tsay [2002].

Dans ce qui suit et sans entrer dans les détails de la méthodologie GARCH, nous précisons les principaux éléments utilisés par la suite.

Ainsi, supposons la dynamique de la chronique des rendements  $r_t$  modélisée par :

$$r_t = \mu_t + \sigma_t z_t \quad (17)$$

où  $\mu_t$  est la moyenne conditionnelle à l'ensemble de l'information disponible, à savoir les données de rendements, jusqu'au temps  $t - 1$  et  $z_t$  désigne les innovations supposées i.i.d, d'espérance  $E(z_t) = 0$  et de variance  $V(z_t) = 1$  et de fonction de répartition marginale  $F_z$ . Pour simplifier la présentation, nous supposons que la variance conditionnelle  $\sigma_t^2$  de la chronique ajustée  $\varepsilon_t = r_t - \mu_t$  est un processus GARCH(1,1) :

$$\sigma_t^2 = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \gamma \sigma_{t-1}^2 \quad (18)$$

où  $\beta_0 > 0$ ,  $\beta_1 > 0$  et  $\gamma > 0$ . De plus, le processus GARCH est stationnaire au sens faible si et seulement si la somme des coefficients  $(\beta_1 + \gamma)$  est inférieure à 1.

Sous l'hypothèse de normalité de la loi conditionnelle des innovations, la fonction log-vraisemblance d'un échantillon de  $n$  observations i.i.d est donnée par :

$$L(\theta) = -n \ln(2\Pi) - \frac{1}{2} \sum_{i=1}^n \ln(\sigma_i^2) - \frac{1}{2} \sum_{i=1}^n \frac{(r_i - \mu_i)^2}{\sigma_i^2} \quad (19)$$

La maximisation de  $L$  permet d'obtenir l'estimateur de maximum de vraisemblance (EMV)  $\hat{\theta}_{ML}$  des paramètres inconnus ainsi que d'extraire les résidus standardisés :

$$(Z_{t-n+1}, Z_{t-n+2}, \dots, Z_t) = \left( \frac{r_{t-n+1} - \hat{\mu}_{t-n+1}}{\hat{\sigma}_{t-n+1}}, \frac{r_{t-n+2} - \hat{\mu}_{t-n+2}}{\hat{\sigma}_{t-n+2}}, \dots, \frac{r_t - \hat{\mu}_t}{\hat{\sigma}_t} \right) \quad (20)$$

### ÉTAPE 2 : La GPD pour l'estimation du quantile des résidus

Les résidus standardisés ainsi obtenus sont utiles pour vérifier si la spécification GARCH retenue est adéquate. En particulier, les innovations doivent être i.i.d. En général, cette hypothèse est plus plausible pour les innovations que pour la chronique des rendements. Si c'est le cas, la GPD, fondée sur l'hypothèse i.i.d, permettra d'estimer les queues de la loi de ces innovations. Par la suite, nous déduisons le quantile extrême  $z_p = \overline{\text{VaR}}_p(Z)$  en nous appuyant sur la formule<sup>15</sup> (16), appliquée sur les innovations.

Ainsi, la  $\overline{\text{VaR}}_{p,t+1}$  est donnée par :

$$\overline{\text{VaR}}_{p,t+1} = \hat{\mu}_{t+1} + \overline{\text{VaR}}_p(Z) \hat{\sigma}_{t+1} \quad (21)$$

Cette approche combinée GARCH-GPD a divers avantages : elle permet d'une part de capturer la nature hétérosclélastique des chroniques temporelles étudiées via une spécification GARCH. D'autre part, elle permet une meilleure description des quantiles extrêmes via la TVE. Depuis les travaux précurseurs de McNeil et Frey [2000], différentes études ont adopté cette approche conditionnelle. Bali et Neftçi [2003] ont appliqué l'approche GARCH-TVE sur le taux d'intérêt à court terme des États-Unis. Ils prouvent que l'approche considérée est plus précise en matière d'estimation de la VaR qu'un modèle GARCH à erreurs Student (GARCH-T).

De son coté, Byström [2004] constate également que le modèle GARCH-GPD est plus performant que les modèles paramétriques en terme des prévisions VaR pour différents marchés des actions internationaux. Dans une application au marché de l'énergie à NYYMEX, Krehbiel et Adkins [2005] ont trouvé que ce type de modélisation est statistiquement supérieure au modèle EWMA et au modèle AR(1)-GARCH(1,1) à innovations normales. Plus récemment, Raggad [2007] et Marimoutou *et al.* [2009] ont démontré également le succès de ce type de modélisation.

## 4. APPLICATION AU MARCHÉ PÉTROLIER

### 4.1 STATISTIQUES PRÉLIMINAIRES ET MODÉLISATION DE LA VOLATILITÉ

Cette étude est réalisée sur des données<sup>16</sup> du marché pétrolier<sup>17</sup>. Nous nous intéressons particulièrement au cours journaliers Spot de deux bruts internationaux de référence à savoir le Brent et le WTI (cf. Figure 1). La première chronique couvre la période allant du 20/05/1987 au 11/09/2008, soit 5478 observations alors que la deuxième s'étale sur une période allant du 02/01/1986 au 16/09/2008, soit 5827 observations. Ce choix est justifié par l'importance des chocs pétroliers, économiques et financiers survenant au cours de cette période. De plus, cette même période est caractérisée par une importante augmentation de la volatilité de cours de brut.

<sup>15</sup> Pour la queue gauche, on applique la même formule après avoir multiplier les résidus standardisés par -1.

<sup>16</sup> Source : Energy Information Agency, <http://www.eia.doe.gov/>

<sup>17</sup> On trouve le marché physique au comptant (dit « Spot ») qui est celui des transactions physiques à livraison immédiate et le marché physique à terme qui correspond aux transactions physiques à livraison différée (on parle encore de marché «forward» où s'échangent des cargaisons de pétrole pour une date ultérieure à un prix prédéterminé et de marché des «futures» qui est également un marché où s'échangent des intentions d'achat ou de vente futurs à un prix immédiatement fixé).

Dans le Tableau 1 sont résumées les principales statistiques concernant les rendements Brent et WTI. En moyenne, le rendement est faible. Le test de normalité est celui de Jarque-Bera, et ceci, par la mesure des coefficients d'asymétrie (Skewness) et d'aplatissement (Kurtosis) qui, sous l'hypothèse de la normalité, prennent respectivement les valeurs 0 et 3. D'après ce tableau, on remarque des coefficients d'asymétrie négatifs ( $\neq 0$ ) et des coefficients d'aplatissement élevés ( $> 3$ ). Ceci dénote que les queues de lois empiriques de ces rendements sont plus épaisses que celles de la loi normale. Tous les skewness sont négatifs prouvant ainsi que la loi des cours est asymétrique, sachant qu'il y a plus de mouvements à la baisse qu'à la hausse. La statistique Jarque-Bera permet de rejeter l'hypothèse nulle de normalité pour n'importe quel niveau de signification. La loi empirique est donc asymétrique et leptokurtique<sup>18</sup>.

	BRENT	WTI
Moyenne (en %)	0.029	0.022
Ecart-type (en %)	2.289	2.493
Minimum (en %)	-36.12	-40.64
Maximum (en %)	17.33	19.15
Skewness	-0.835	-1.016
Kurtosis	19.91	20.81
Jarque-Bera	59069.27** (0.00)	91008.74** (0.00)
Ljung-Box	54.9124** (0.0001)	60.7061** (0.00)
Test-LM	414.2753** (0.000)	241.2764 ** (0.002)

(\*\*) Statistiquement différent de zéro au seuil de 1 %

TABLEAU 1. Quelques statistiques et tests préliminaires pour les rendements de deux chroniques, Brent et WTI

La statistique de Ljung-Box (LB(22)) calculée avec 22 retards permet de détecter l'autocorrélation des rendements. Sous l'hypothèse  $H_0$  de nullité jointe des 22 premières autocorrélations, cette statistique suit une loi Khi-deux à 22 degrés de liberté. Ce test montre la présence d'une certaine corrélation dans les chroniques des rendements. Il en découle que ces dernières sont incompatibles avec l'hypothèse d'indépendance de la TVE. Pour remédier à ce problème, il est nécessaire d'appliquer la TVE sur des rendements préalablement filtrés à travers un modèle de type GARCH. Le résultat d'une recherche de spécification<sup>19</sup> en termes de critère de Akaike (AIC) et le critère d'information bayésien (BIC) pour différents valeurs des retards p et q, montrent que les spécifications AR(1)-GARCH(1,1) et GARCH(1,2) sont raisonnables respectivement pour les deux chroniques Brent et WTI. Un processus AR(1)-GARCH(1,1) peut s'écrire sous la forme des équations suivantes :

<sup>18</sup> L'hypothèse de normalité des rendements est généralement rejetée. Les queues des distributions empiriques des rendements sont généralement plus épaisses que celles d'une loi gaussienne. On parle alors de distribution leptokurtique.

<sup>19</sup> Pour plus des détails sur les critères de sélection des modèles nous renvoyons au livre de Yaffee et McGee [1999].

$$\left\{ \begin{array}{l} X_t = \mu_t + \sigma_t z_t \\ \mu_t = \alpha_0 + \alpha_1 X_{t-1} \end{array} \right. \quad (22.1)$$

$$\left\{ \begin{array}{l} \sigma_t^2 = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \gamma \sigma_{t-1}^2 \end{array} \right. \quad (22.2)$$

$$\left\{ \begin{array}{l} \sigma_t^2 = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \gamma \sigma_{t-1}^2 \end{array} \right. \quad (22.3)$$

où  $z_t$  désignent les innovations indépendantes et identiquement distribuées, telles que  $E(z_t) = 0$  et  $E(z_t^2) = 1$ . L'espérance conditionnelle, à l'ensemble de l'information disponible, à savoir les données de rendements, jusqu'au temps  $t-1$ , de  $X_t$ ,  $\mu_t$ , est modélisée par un processus autorégressif d'ordre un. La variance conditionnelle à l'information disponible à la date  $t-1$ ,  $\Omega_{t-1}$ , du processus est  $\sigma_t^2$ . Ce modèle est estimé par Pseudo-Maximum de Vraisemblance (PML) toute en supposant la normalité conditionnelle du terme d'innovations  $z_t$ .

Le Tableau 2 présente les estimations de la représentation AR(1)-GARCH(1,1) et GARCH(1,2) respectivement pour les deux chroniques des rendements Brent et WTI, lorsque les innovations sont supposées normales. Globalement, tous les paramètres de ce modèle sont largement significatifs. En effet, pour chacune des chroniques étudiées, l'équation de la moyenne (cf. 22.2) montre que le terme constant ( $\alpha_0$ ) ainsi que le coefficient ( $\alpha_1$ ) sont significatifs (à l'exception du terme constant pour la chronique WTI) au seuil de 5 %. De même pour l'équation de la variance (cf. 22.3), on note que la constante ( $\beta_0$ ), le coefficient autorégressif de la volatilité conditionnelle ( $\gamma$ ) ainsi que le coefficient de la dernière innovation observée ( $\beta$ ), la somme des coefficients ( $\gamma + \beta$ ) est

	Paramètre	Ecart-type	T-stat	p-value
Brent	Equation de la moyenne			
Constante	0.05219*	0.023951	2.179	0.014
AR(1)	0.05894**	0.014470	4.074	0.000
Equation de la variance				
Constante	0.07793**	0.009774	7.973	0.000
ARCH(1)	0.08906**	0.004511	19.744	0.000
GARCH(1)	0.899530**	0.005321	169.049	0.000
WTI	Equation de la moyenne			
Constante	0.03371	0.0022166	1.521	0.064
Equation de la variance				
Constante	0.08879**	0.012843	6.914	0.000
ARCH(1)	0.11957**	0.007208	16.589	0.000
GARCH(1)	0.58251**	0.064373	9.049	0.000
GARCH(2)	0.28948**	0.058114	4.981	0.000

(\*) et (\*\*) désignent respectivement la significativité aux seuils de 1 % et 5 %

TABLEAU 2. Résultats d'estimation des modèles AR(1)-GARCH(1,1) GARCH(1,2) respectivement pour les deux chroniques Brent et WTI

inférieure à 1. La dynamique de la volatilité conditionnelle apparaît clairement stationnaire. De plus, elle se caractérise par la présence d'un phénomène de persistance

de la volatilité. Les valeurs estimées de la variance conditionnelle obtenues dans le modèle AR(1)-GARCH(1,1) sont reproduites sur la Figure 6 et rendent bien compte des clusters de volatilité, c'est-à-dire de la succession de périodes de turbulences et de calme sur le marché considéré. Ce dernier constat est visiblement observable, par exemple, au cours de la période de la guerre de Golfe (1991) (un pic des volatilités).

Le Tableau 3 montre l'effet du filtrage de la chronique des rendements du Brent à travers la spécification AR(1)-GARCH(1,1). Pour la chronique des rendements, le résultat du test LM, pour un ordre  $p = 12$ , conclut clairement au rejet de l'hypothèse nulle d'absence d'effets ARCH. Le test de Ljung-Box permet de rejeter l'hypothèse nulle d'absence d'autocorrélation. De même, le test de Jarque Bera conduit au rejet de l'hypothèse nulle de normalité. Toutefois, pour la chronique des résidus standardisés, on constate l'absence d'autocorrélation et l'absence d'effets ARCH. Cependant, on maintient le rejet de l'hypothèse nulle de normalité.

Brent	Jarque-Bera	Ljung-Box	LM-test
Rendements	59069.27** (0.00)	54.9124** (0.000)	414.2753** (0.000)
Résidus standardisés	894.3** (0.00)	13.37 (0.343)	43.2888 (0.220)

(\*\*) statistiquement différent de zéro au seuil de 1%

TABLEAU 3. Tests de diagnostics des effets de filtrage AR(1)-GARCH(1,1)  
des rendements Brent

En résumé, nous pouvons conclure que l'hypothèse d'indépendance semble être en désaccord avec les caractéristiques des données considérées. Néanmoins, cette hypothèse constitue la base de la TVE. Le processus de filtrage GARCH a permis de contourner ce problème. Ainsi, nous concluons que l'hypothèse i.i.d est plus plausible pour la chronique des innovations que pour celle des rendements. Toutefois, l'hypothèse de normalité ne peut être retenue pour les deux chroniques en question.

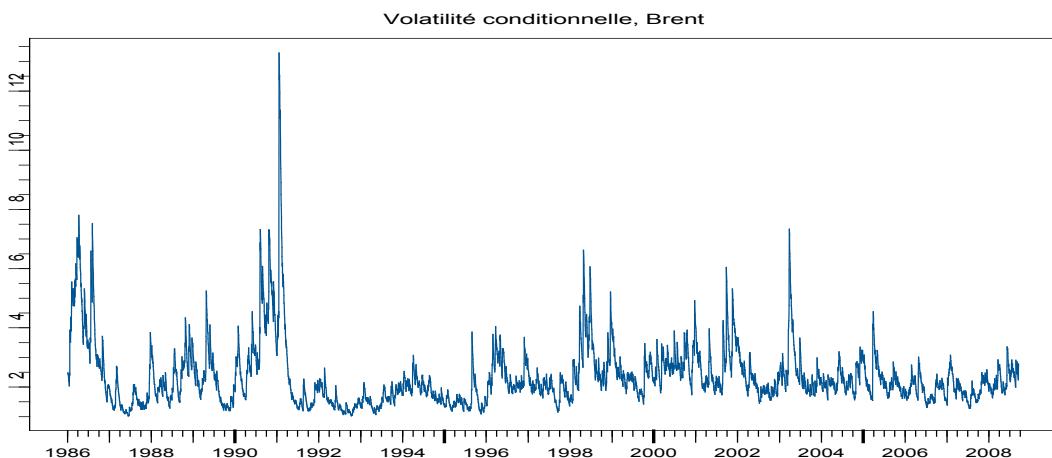


FIGURE 6. Volatilité conditionnelle estimée pour la chronique Brent :  
modèle AR(1)-GARCH(1,1)

## 4.2 PROCÉDURES D'ESTIMATION ET D'ÉVALUATION DES MODÈLES DE LA VAR

Afin d'étudier la performance de l'approche basée sur la TVE ainsi que les autres modèles utilisés pour l'estimation de la  $VaR_p$  pour  $p \in \{0,95; 0,99; 0,995\}$ , nous procérons à un exercice de backtesting<sup>20</sup> basé sur les deux chroniques du brut Brent et WTI. Soient  $r_1, r_2, \dots, r_m$  la chronique des rendements historiques.

La procédure d'estimation est basée sur la technique du rolling qui consiste à estimer la VaR selon une fenêtre glissante « rolling window » de taille  $w$ . Cette procédure s'explique comme suit : on a pour chaque modèle  $T$  prévisions VaR. La première prévision est basée sur le modèle dont les paramètres ont été estimés en utilisant les observations  $r_1, r_2, \dots, r_w$ . La seconde est basée sur les observations  $r_2, r_3, \dots, r_{w+1}$ . Cette démarche se poursuit jusqu'à la dernière prévision qui est basée sur les observations  $r_{T-1}, r_{T-2}, \dots, r_{w+T-1}$ . À chaque fois donc, on ajoute une nouvelle observation et on abandonne la plus ancienne. La VaR re-estimée sera ensuite confrontée à la perte éventuelle du jour. L'avantage principal de cette technique basée sur le rolling est qu'elle permet de capturer les caractéristiques de la dynamique temporelle des données à différentes périodes de temps.

Dans ce travail, nous considérons trois valeurs de la fenêtre de temps  $w$ ,  $w \in \{250; 500; 1000\}$  jours.

Comme il a été expliqué dans la Section 2, le choix de seuil est capital pour l'approche GPD. En effet, un bon estimateur des quantiles issus d'un modèle GPD dépend énormément du choix du seuil. Le seuil optimal est le résultat d'un arbitrage entre le biais et l'inefficacité. Dans ce papier, nous retenons le résultat de la simulation Monte-Carlo faite par Raggad [2007] et Marimoutou *et al.* [2009]. Ces auteurs ont décidé de retenir<sup>21</sup> toujours  $k = 100$  (soit 10 % de 1000 observations utilisées comme fenêtre d'estimation). En conséquence, le seuil sera la valeur tronquant les 10 % des observations les plus petites (les plus grandes) pour la queue gauche (la queue droite).

Il est à noter également qu'il est pratiquement impossible de procéder soigneusement à la vérification et au choix de la meilleure paramétrisation, tout au long de cette période du backtesting. Pour cela nous supposons<sup>22</sup> que les processus AR(1)-GARCH(1,1) et GARCH(1,2) sont respectivement adéquats pour les chroniques Brent et WTI.

Pour chaque  $t \in T$ , un nouveau modèle GARCH (AR(1)-GARCH(1,1) pour le Brent et GARCH(1,2) pour le WTI) est estimé et un nouveau modèle GPD est ajusté pour les résidus standardisés. Cette approche est la TVE conditionnelle, comme il a été mentionné dans la Section 3.2.

---

<sup>20</sup> Le backtesting consiste à confronter la VaR calculée avec les pertes et profits effectivement réalisés. Ainsi, pour un niveau de confiance de 99%, les pertes effectives ne devraient dépasser les prévisions VaR que dans 1 % des cas environ. Sinon, nous serions obligés de remettre en question la validité de cette méthode.

<sup>21</sup> En raison des difficultés pratiques liées au choix de la valeur optimale  $k$  pour chaque fenêtre glissante, nous supposons que la valeur optimale de  $k$  obtenue par simulation reste valable tout au long de la période du backtesting.

<sup>22</sup> Cette hypothèse a été également adoptée par McNeil et Frey [2000], Gençay *et al.* [2003], Raggad [2007] et Marimoutou *et al.* [2009].

De plus, nous estimons également la GPD inconditionnelle qui correspond à l'expression (17).

Nous introduisons également le modèle GARCH à erreurs normales. Toutefois, l'hypothèse de normalité des résidus standardisés est difficile à justifier du point de vue empirique. En effet, les queues des lois empiriques des rendements sont généralement plus épaisses que celles d'une loi gaussienne. L'utilisation d'une loi Student-*t* ayant des queues de loi plus épaisses que la loi gaussienne peut résoudre potentiellement ce problème. Dans le premier cas, le quantile normal conditionnel des résidus standardisés est tout simplement donné par  $z_p = \pi_p^{-1}(z)$  avec  $\pi_p(\cdot)$  est la fonction de répartition de la loi normale standard. Dans le deuxième, nous pouvons définir le quantile d'une loi Student-*t* par  $t_p = \sqrt{\frac{\hat{v} - 2}{\hat{v}}} F_p^{-1}(z, v)$  avec  $F_p(\cdot)$  suit la loi de Student à  $v$  degrés de liberté, pour  $v > 2$ . À chaque jour  $t$ , on estime un modèle GARCH erreurs Student, un nouveau  $v$  est estimé et un nouveau quantile. La VaR est calculée via la formule (21) pour les deux lois en question : la normale conditionnelle et la Student conditionnelle, en remplaçant  $\overline{\text{VaR}}_p(Z)$  respectivement par  $z_p$  et  $t_p$ .

Pour valider les modèles en question, on procède à un exercice du backtesting. Le backtesting est une confrontation entre la VaR calculée et les pertes et profits effectivement réalisés. Plusieurs tests<sup>23</sup> fondés sur les occurrences de violations ont été proposés afin d'évaluer la validité des mesures et des prévisions de VaR.

Soit la série des violations notée  $I_t$  est définie par une variable dichotomique ayant l'expression suivante :

$$\begin{aligned} \text{Queue gauche }^{24}; \quad I_t &= \begin{cases} 1 & \text{si } r_{t+1} < \text{VaR}_{t+1} \\ 0 & \text{si } r_{t+1} \geq \text{VaR}_{t+1} \end{cases} \\ \text{Queue droite }^{25}; \quad I_t &= \begin{cases} 1 & \text{si } r_{t+1} > \text{VaR}_{t+1} \\ 0 & \text{si } r_{t+1} \leq \text{VaR}_{t+1} \end{cases} \end{aligned}$$

où  $r_t$  est le rendement de l'actif en  $t$ , avec  $t$  compris entre 1 et  $T$ .  $T$  est la période du backtesting. Christoffersen [1998] montre qu'un modèle de VaR est fiable lorsque la suite des violations  $I_t$  satisfait les propriétés de « couverture inconditionnelle » (Unconditional Coverage Property) et « d'indépendance » (Independence Property) des exceptions.

Le test de Couverture inconditionnelle, Unconditional Coverage (UC), développé par Kupiec [1995] permet d'évaluer la pertinence d'un modèle de risque en comparant la proportion  $p$  de violations de la VaR au niveau de risque retenu pour le calcul de cette dernière, c'est-à-dire  $p$ . La fiabilité d'un modèle de VaR sera remise en question lorsque la proportion observée diffère considérablement de  $p$ . Il s'agit de tester les hypothèses suivantes :

---

<sup>23</sup> Pour plus de détails sur ces tests, cf. Campbell S. D., 2005

<sup>24</sup> Pour la queue gauche, la chronique des violations permet d'estimer le nombre de fois où la perte enregistrée dépasse la perte maximale estimée par le modèle d'évaluation.

<sup>25</sup> Pour la queue droite, la chronique des violations permet d'estimer le nombre de fois où le profit enregistré est supérieur au profit maximal estimé par le modèle d'évaluation.

$$\begin{cases} H_0 : \frac{N}{T} = p \\ H_1 : \frac{N}{T} \neq p \end{cases}$$

où  $N = \sum_t^T I_t$  est le nombre des violations des VaR estimées sur  $T$ , fenêtre temporelle utilisée pour le backtesting.

Le ratio de vraisemblance du test de couverture inconditionnelle est donné par :

$$LR_{uc} = 2 \left[ \log \left( \left( \frac{N}{T} \right)^N \left( 1 - \left( \frac{N}{T} \right) \right)^{T-N} \right) - \log \left( P^N (1-p)^{T-N} \right) \right] \quad (26)$$

sous l'hypothèse nulle,  $LR_{uc} \sim \chi^2(1)$ . L'hypothèse  $H_0$  est rejetée lorsque la statistique du test est supérieure à la valeur critique du Khi-deux à un degré de liberté. Il est à noter que ce test est un test bilatéral. De cette manière, il permet de rejeter un modèle s'il génère un nombre élevé ou faible d'exceptions. En effet, un bon modèle ne doit ni sous-estimer, ni sur-estimer le risque.

Christoffersen [1998] a proposé un test plus élaboré qui permet de tester à la fois : (i) si le nombre des exceptions réelles est égal au nombre d'exceptions théoriques et (ii) si le processus des violations VaR est indépendamment<sup>26</sup> distribué.

Le ratio de vraisemblance du test de couverture conditionnelle est donné par :

$$LR_{cc} = -2 \log \left[ (1-p)^{T-N} p^N \right] + 2 \ln \left[ (1-\pi_{01})^{n_{00}} \pi_{01}^{n_{01}} (1-\pi_{11})^{n_{01}} \pi_{11}^{n_{11}} \right] \quad (27)$$

Sous l'hypothèse nulle que le processus des violations VaR est indépendamment distribué et la proportion des violations réelles est égale à la proportion d'exceptions théoriques,  $LR_{cc} \sim \chi^2(2)$ .

Notons que  $n_{01}$  : nombre de jours sans violations suivi d'une violation,  $n_{00}$  : nombre de jours successifs sans violations,  $n_{11}$  : nombre de jours successifs avec violations et  $n_{10}$  : nombre de jours sans violations précédé par une violation. La probabilité d'observer une violation (resp. de ne pas en observer) pour la période courante peut être liée (ou non) à l'occurrence d'une violation à la période précédente,  $\pi_{i,j} = \Pr(I_t = j / I_{t-1} = i)$ . Ainsi  $\pi_{0,1} = \frac{n_{0,1}}{n_{0,0} + n_{0,1}}$  et  $\pi_{1,1} = \frac{n_{1,1}}{n_{1,0} + n_{1,1}}$ .

### 4.3 RÉSULTATS EMPIRIQUES

#### 4.3.1 Fenêtre glissante de 1000 observations

Après 4477 estimations<sup>27</sup> VaR pour la chronique Brent, et 4826 pour le WTI, on calcule le nombre des violations, c'est-à-dire nombre de fois où les pertes ont dépassé la VaR,

<sup>26</sup> L'historique des violations de la VaR ne doit donner aucune information sur la réalisation ou non d'une violation à la période suivante  $t$ . La présence de regroupement des violations suggère que le modèle de VaR ne réagit pas assez vite aux changements des conditions de marché.

<sup>27</sup> On a 5577 observations pour la chronique des rendements Brent, en utilisant une fenêtre de 1000 observations lors du rolling, on génère 4477 prévisions VaR.

durant toute la période du test. Ce nombre des violations révèle le degré de pertinence du modèle utilisé. En effet, un nombre des violations acceptables doit être en adéquation avec le niveau de confiance auquel la VaR a été calculée. Ainsi, pour un niveau de confiance de 99 %, les pertes effectives ne devraient dépasser les prévisions VaR que dans 1 % des cas environ. Sinon, nous serions obligés de remettre en question la validité de cette méthode.

La performance relative de chaque modèle est tout d'abord évaluée en termes d'un ratio des violations. Le ratio des violations est défini comme étant le rapport du nombre de violations par le nombre total des prévisions VaR d'une période donnée. Si le ratio des violations du quantile d'ordre  $p$  est plus grand que  $p \%$ , ceci nous amène à conclure qu'il y a une sous-estimation excessive du rendement réalisé. Au contraire, si le ratio de violations du quantile d'ordre  $p$  est inférieur à  $p \%$ , alors, on dit qu'il y a surestimation excessive du rendement réalisé par le modèle considéré. Par exemple, si le ratio des violations est 3 % pour une VaR à un niveau de confiance de 95 %, le rendement réalisé est seulement 3 % du temps plus grand que ce que le modèle prévoit.

L'évaluation des modèles VaR se fait donc en comparant les ratios des violations théoriques à ceux réellement observés. Statistiquement, nous utilisons d'abord les deux tests de backtesting proposés par Christoffersen, à savoir le test de couverture inconditionnelle et le test de couverture conditionnelle dans le but de valider les modèles de gestion des risques. Ainsi, un modèle est considéré comme performant s'il est validé à la fois par les deux tests de couvertures inconditionnelle et conditionnelle.

Les Tableaux 4 et 5, présentés en Annexe, montrent les ratios de violation pour les différents niveaux de confiance et les  $p$ -values correspondant aux tests statistiques. Les résultats peuvent être récapitulés comme suit :

- La loi normale inconditionnelle est rejetée par presque tous les niveaux de confiance considérés et pour les deux queues de la loi. En général, les ratios de violations sont supérieurs par rapport à la probabilité théorique. Ce modèle se révèle inadapté pour estimer le risque réel. Statistiquement le test de couverture inconditionnelle, conclut au rejet (sauf, VaR.95 Brent-queue gauche, VaR.99 Brent-queue droite) de l'égalité de ces ratios aux probabilités théoriques d'échec. En couverture conditionnelle le modèle n'est accepté qu'une seule fois. Cet échec est attendu dans la mesure où ce modèle ne permet pas de prendre en compte le caractère leptokurtique des lois des rendements. Si on considère toutes les combinaisons possibles, c'est à dire en tenant compte de 2 chroniques du brut, les 3 valeurs théoriques  $\alpha$  et les 2 queues gauche et droite, le taux de succès du modèle est égal à  $\frac{1}{12}$  soit 8.33 %.
- L'approche normale conditionnelle ou GARCH-N donne des résultats acceptables pour quelques niveaux de confiance mais sa performance se détériore surtout pour des quantiles plus élevés excepté un (VaR995 WTI queue droite). L'avantage de cette approche est qu'elle permet de capturer assez bien les propriétés dynamiques de la loi conditionnelle des rendements. En effet, elle produit une VaR qui reflète la volatilité courante. Néanmoins, ce modèle est incapable de rendre compte de l'ensemble de l'excès de kurtosis et il tend à sous-estimer la vraie exposition aux risques. En effet,

les lois sont leptokurtiques, ce qui implique des événements extrêmes plus fréquents que pour la loi normale. Son taux de succès est égal à  $\frac{4}{12}$  soit 33.33 %.

- Le modèle Student conditionnel ou GARCH-T fournit des résultats nettement plus satisfaisants que le modèle normal conditionnel. Cependant, il tend à produire de meilleures mesures de la VaR pour la queue gauche que pour la queue droite. Le modèle tend à surestimer la queue droite. En effet, le plus souvent, les ratios des violations du côté droit sont inférieurs aux probabilités théoriques. En général, le modèle est très satisfaisant pour l'estimation de la VaR de la queue droite à différents niveaux de confiance mais il échoue à produire le même résultat pour la queue gauche. Son taux total de succès est égal à  $\frac{9}{12}$  soit 75 %.
- Le modèle GPD, bien qu'il satisfasse le test de couverture inconditionnelle, pour presque tous les niveaux de confiance, n'arrive pas à atteindre le critère de couverture conditionnelle dans divers cas. Le non succès relatif de ce modèle peut s'expliquer par l'invalidité de l'hypothèse d'indépendance. En effet, le modèle a été appliqué directement sur les rendements pris sous cette hypothèse. Son taux total de succès est égal à  $\frac{4}{12}$  soit 33.33 %.
- Le modèle GPD conditionnel est très performant en termes de prévisions VaR. Les ratios des violations sont statistiquement égaux aux valeurs théoriques, excepté dans un seul cas (VaR 0,5 % de la queue gauche : Brent). D'ailleurs, ce modèle donne des résultats très satisfaisants pour les deux chroniques étudiées et à travers les deux queues gauche et droite. Les résultats prouvent également que le modèle GPD conditionnel est particulièrement satisfaisant dans le cas où les lois des rendements exhibent des queues extrêmes. Son taux total de succès est égal à  $\frac{11}{12}$  soit 91.66 %.

Comme nous pouvons le constater sur la Figure 4 pour la chronique Brent, l'approche GPD conditionnelle fournit une quantification plus flexible de VaR, qui tient compte des dynamiques de la volatilité. La procédure du backtesting du modèle GPD conditionnel correspondant à la Figure 7 est donnée par la Figure 8. Au-dessous de l'axe des abscisses figurent les rendements négatifs (pertes) superposés aux violations des prévisions VaR 1 % pour la queue gauche. Les violations de ce modèle au niveau de la queue gauche (droite) sont montrées avec des croix. Celles de la queue droite sont représentées avec des triangles.

#### 4.3.2 Fenêtre glissante de tailles inférieures

Comme cela a été constaté pour la première fois par Jorion [1996], la mesure VaR est sujet à des erreurs d'estimations du fait qu'elle est calculée à partir des réalisations passées des variables aléatoires. Ses résultats montrent que la mesure VaR peut présenter un biais lors de son estimation.

Parmi les critères les plus importants lors de l'estimation des paramètres à partir des observations des chroniques il y a la taille de la période d'estimation prise en considération. La majorité des travaux récents, surtout ceux basés sur la théorie des valeurs extrêmes, considèrent une fenêtre de taille égale à 1000 observations (McNeil et Frey [2000], Raggad [2007], Marimoutou *et al.* [2009]), toutefois, il serait peut être

décisif de connaître la pertinence de nos résultats lorsque nous utilisons différentes fenêtres d'estimations, en particulier, des fenêtres de tailles inférieures.

Ainsi, dans le souci d'étudier la sensibilité de nos résultats par rapport au choix de la taille de la fenêtre d'estimation, nous répétons les calculs faits dans la sous-section 6.1 en considérant deux autres fenêtres à savoir  $w = 250$  et  $w = 500$ . Il est à noter que lors de l'estimation des modèles fondés sur la théorie des valeurs extrêmes, nous supposons toujours que le seuil optimal est la valeur tronquant les 10 % de la taille de la fenêtre utilisée.

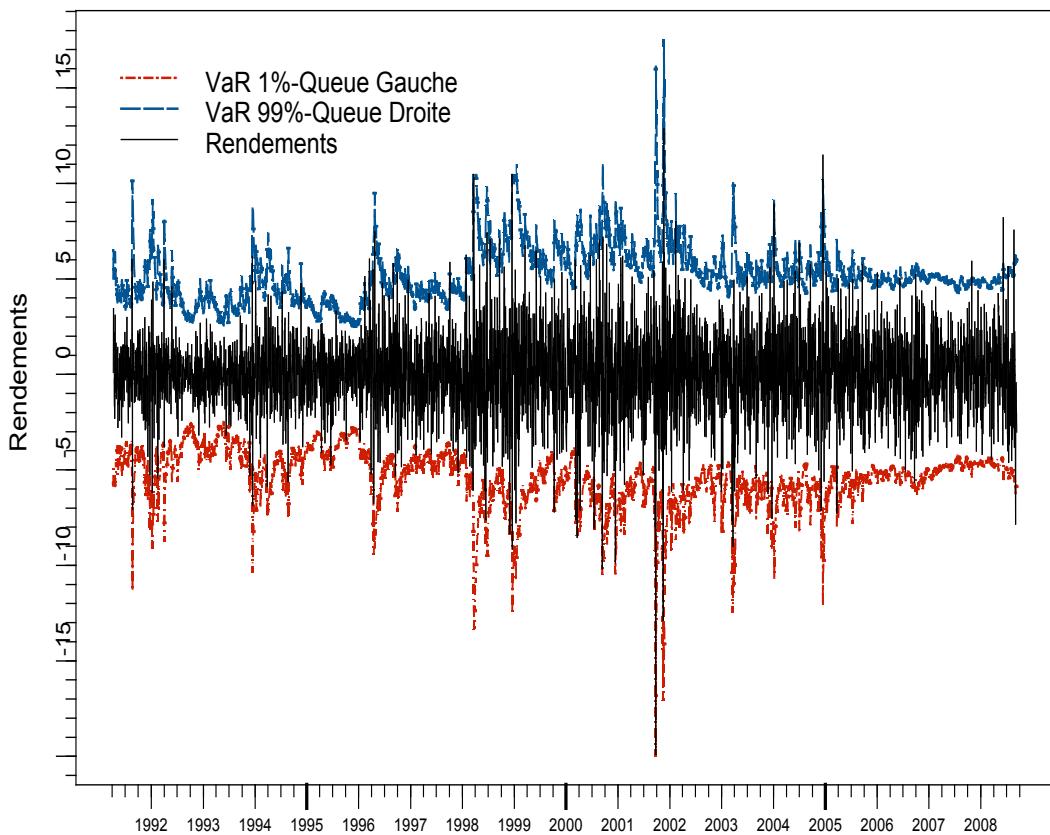


FIGURE 7. Rendements journaliers et prévisions VaR pour la chronique Brent, modèle GPD conditionnel

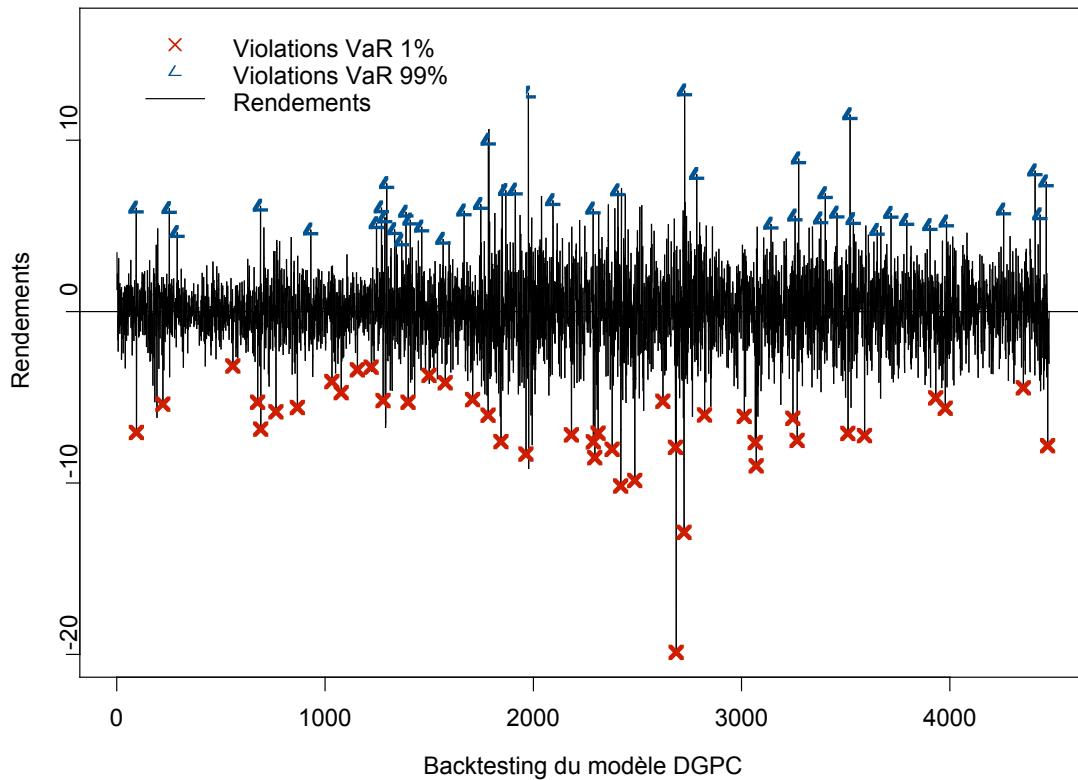


FIGURE 8. Procédure du backtesting du modèle GPD conditionnel, pour la chronique Brent.

Le Tableau 6 résume le taux de succès de l'ensemble des modèles considérés selon les trois fenêtres d'estimation  $w = 250$ ,  $w = 500$  et  $w = 1000$ . Le taux de succès du modèle est égal au rapport du nombre de fois où le modèle est validé par les deux tests de couvertures inconditionnelle et conditionnelle par le nombre des combinaisons possibles, c'est-à-dire en tenant compte de 2 chroniques du brut, les 3 valeurs théoriques  $\alpha$  et le 2 queues gauche et droite, soit 12.

La lecture de ce tableau permet de tirer les constatations suivantes :

- Pour  $w = 500$ , nous aboutissons aux mêmes résultats trouvés précédemment ; les modèles les plus performants en termes de prévisions de la VaR sont successivement le modèle GPDC et le modèle GARCH-T. De même, nous trouvons également que les modèles Normale, GARCH-N, GPD sont les moins performants si on fait varier la taille de la fenêtre d'estimation de 1000 à 500.
- Les modèles conditionnels (GARCH-N, GARCH-T, GPDC) sont les moins affectés, comparés aux modèles inconditionnels (Normal, GPD), par une diminution de la taille de la fenêtre d'estimation. Toutefois, il est à noter que la performance des modèles se dégrade lorsque nous réduisons la fenêtre d'estimation à 500.
- À l'exception du modèle GARCH-N, la performance de l'ensemble des modèles étudiés se détériore pour des fenêtres de taille 250.

Modèles	$W = 1000$	$W = 500$	$W = 250$
Normal	8.33 %	0 %	0 %
GARCH-N	33.33 %	16.66 %	41.66 %
GARCH-T	75 %	66.66 %	50 %
GPD	33.33 %	0 %	0 %
GPDC	91.66 %	91.66 %	8.33 %

TABLEAU 6. Comparaison des taux de succès des modèles,  
selon les trois fenêtres d'estimation

En nous appuyant sur l'ensemble des résultats, nous constatons que pour la majorité des modèles considérés, le choix de taille de la fenêtre utilisée pour l'estimation de VaR est capitale. Généralement, il y a un arbitrage entre la taille de l'échantillon et la robustesse du modèle. Le choix d'une taille inférieure à 1000 observations pourra réduire la performance des modèles en question. La principale conclusion de cette étude est qu'il y a généralement une supériorité des modèles conditionnels (GPDC, GARCH-T) pour l'estimation de la VaR. Ces modèles possèdent des propriétés attrayantes, ils permettent à la fois de capturer la lourdeur des queues et la nature hétéroscléastique des chroniques financières ou macroéconomiques.

## 5. CONCLUSION

Le marché du pétrole est fondamentalement très volatile dans la mesure où il dépend de facteurs très difficiles à prévoir, que ce soient les facteurs géopolitiques, les aléas climatiques ou encore d'événements tels que les perspectives économiques ou encore l'évolution des capacités de production. Ainsi, les investisseurs et les gestionnaires des risques sont particulièrement sensibles à l'occurrence de pertes substantielles. Dans ce contexte, il est primordial de disposer des outils permettant de mieux quantifier les risques liés aux fluctuations des prix du brut.

Nous avons présenté une classe des modèles basée sur la théorie des valeurs extrêmes, qui permet d'évaluer les événements rares et les pertes associées à leur survenance. Ce travail offre des réponses intéressantes à la question de l'évaluation du risque sur le marché pétrolier. Nous avons mis en évidence l'apport distinctif de l'approche conditionnelle de la théorie des valeurs extrêmes, en matière de prévision de la mesure VaR. En particulier, la comparaison du modèle GPD conditionnel à un ensemble de modèles souvent rencontrés dans la littérature VaR a montré la supériorité de ce premier. De plus, le GARCH-T peut aboutir à des résultats également satisfaisants. Ces deux modèles possèdent des propriétés attrayantes, ils produisent une VaR qui reflète l'aspect dynamique de la volatilité et permettent de capturer la lourdeur des queues. La VaR Normale est une mesure qui nous semble inadaptée pour évaluer le risque extrême des rendements du brut dont les queues de loi sont les plus épaisses. Enfin, nous constatons que pour l'ensemble des modèles considérés, le choix de taille de la fenêtre utilisée pour l'estimation de VaR est capital. Généralement il y a un arbitrage entre la taille de l'échantillon et la robustesse du modèle. Le choix d'une taille inférieure à 1000 observations pourra réduire la performance des modèles en question.

La fiabilité de l'approche basée sur la TVE en terme de prévisions VaR et sa supériorité conduisent à recommander son implémentation dans le cadre de la gestion

du risque sur le marché pétrolier. L'intérêt pour les différents acteurs de marchés (investisseurs, gérants, gestionnaires des risques, législateurs) est d'obtenir une information fiable sur le risque extrême et d'éviter les surestimations et les sous-estimations qu'impliquent l'emploi des indicateurs de risque traditionnellement utilisés.

Dans ce travail nous avons adopté une approche conditionnelle de la théorie des valeurs extrêmes. Cette approche initialement introduite par McNeil et Frey [2000] consiste à filtrer les chroniques des rendements à travers un modèle de type GARCH et d'appliquer dans une seconde étape les outils de la TVE. L'objectif de ce processus de filtrage est de préserver l'hypothèse i.i.d., hypothèse de base de la TVE. Les méthodes proposées ici donnent des résultats satisfaisants, mais il serait aussi intéressant de tester d'autres processus de filtrage, incluant la classe des modèles à volatilité stochastique.

*Remerciements.* L'auteur remercie le Comité Mixte pour la Coopération Universitaire franco-tunisien (CMCU), sans lequel ce travail n'aurait jamais abouti. Ses remerciements s'adressent également à Claude Deniau pour son aide et ses conseils, à un rapporteur anonyme et au Directeur de la rédaction pour leurs critiques constructives.

## BIBLIOGRAPHIE

- BALI T. G., NEFTCI S. N., "Disturbing extremal behaviour of spot price dynamics", *Journal of Empirical Finance* 10, 2003, p. 455-477.
- BALKEMA A., DE HAAN L., "Residual Life at Great Age", *Annals of Probability* 2, 1974, p. 792-804.
- BARBUT M., "Homme moyen ou homme extrême...", *Journal de la Société Française de la Statistique* 144, 2003, p.1-2.
- BARBUT M., « Une famille de distributions : des paretaines aux conta-paretaines, applications à l'étude de la concentration urbaine et de son évolution », *Mathématiques, Informatique et Sciences humaines* 141, 1998, p. 43-72.
- BEIRLANT J., VYNCKIER P., TEUGELS J., "Tail index estimation, Pareto quantile plots, and regression diagnostics", *Journal of American Statistical Association* 91, 1996, p.1659-1667.
- BOLLERSLEV T., "Generalized autoregressive conditional heteroskedasticity", *Journal of Econometrics* 31, 1986, p. 307-327.
- BOX G. E. P., JENKINS G. M., "Time Series Analysis: Forecasting and Control", Holden-Day, 1976, [revised edition].
- BYSTRÖM H., "Managing extreme risks in tranquil and volatile markets using conditional extreme value theory", *International Review of Financial Analysis* 13, 2004, p.133-152.
- CAMPBELL S. D., "A review of backtesting and backtesting procedures", *Finance and Economics Discussion Series Divisions of Research Statistics and Monetary Affairs Federal Reserve Board*, Washington (D.C), 2005.
- CHRISTOFFERSEN P., "Evaluating Interval Forecasts", *International Economic Review* 39, 1998, p. 841-862.
- CHRISTOPEIT N., "Estimating parameters of an extreme value loi by the method of moments", *Journal of Statistical Planning and Inference* 41, 1994, p. 173-186.
- COLES S.G., WALSHAW, D., "Directional modeling of extreme wind speeds", *Appl. Statist.* 43, 1994, p. 139-157.
- DANIELSSON J., DE HAAN L., DE VRIES, C. G., "Using a bootstrap method to choose the sample fraction in tail index estimation", *Journal of Multivariate Analysis* 76, 2001, p. 226-248.
- DANIELSSON J., DE VRIES C. G., "Tail index estimation with very high frequency data", *Journal of Empirical Finance* 4, 1997, p. 241-257.

- DAVISON A. C., SMITH R. L., "Models for exceedances over high thresholds", *Journal of Royal Statistical Society Ser. B* 52, 1990, p. 393-442.
- DEKKERS A., EINMAHL J., DE HAAN L., "A moment estimator for the index of an extreme-value loi", *Ann. Statist.* 17, 1989, p. 1833-1855.
- EMBRECHTS P., KLUPPELBERG C., MIKOSCH T., *Modelling Extremal Events for Insurance and Finance*, Berlin, Springer, 1997.
- EMBRECHTS P., RESNICK S., SAMORODNITSKY G., "Extreme value theory as a risk management tool", *North American Actuarial Journal* 26, 1999, p. 30-41.
- ENGLE R. F., "Autoregressive conditional heteroscedasticity with estimates of the variance of U.K. inflation", *Econometrica* 50, 1982, p. 987-1008.
- FELLER W., "An Introduction to Probability Theory and Its Applications", Vol. I. (3rd ed.), New York, Wiley, 1968.
- FISHER R., TIPPETT L., "Limiting forms of the frequency loi of the largest or smallest member of a sample", *Proceedings of the Cambridge Philosophical Society* 24, 1928, p. 180-190.
- GENÇAY R., SELÇUK F., "Extreme value theory and value-at-risk: Relative performance in emerging markets", *International Journal of Forecasting* 20, 2004, p. 287-303.
- GREENWOOD J., LANDWEHER J., MATALAS N., WALLIS J., "Probability weighted moments: Definition and relation to parameters of several lois expressible in inverse form", *Water Resources Research* 15, 1979, p. 1049-1054.
- HALL P., "Using the bootstrap to estimate mean squared error and select smoothing parameter in nonparametric problems", *J. Multivariate Anal.* 32, 1990, p. 177-203.
- HAN Z., "Living to 100 and Beyond: An Extreme Value Study" *Living to 100 and Beyond Symposium Monograph*, 2005. [Available at: <http://www.soa.org/ccm/content/research-publications/library-publications/mono-graphs/life-monographs/living-to-100-and-beyond-monograph/>].
- HILL B. M., "A simple general approach to inference about the tail of a Loi", *Annals of Statistics* 5, 1975, p. 1163-1174.
- HORIUCHI S., WILMOTH J. R., "Deceleration in the age pattern of mortality at older ages", *Demography* 35, 1998, p. 391-412.
- HOSKING J., WALLIS J., "Parameter and quantile estimation for the generalized Pareto loi", *Technometrics* 29, 1987, p. 339-349.
- HULL J., WHITE A., "Incorporating volatility updating for value-at-risk", *Journal of Risk* 1, 1998, p. 5-19.
- JORION P., "Measuring the risk in value at risk", *Risk 2: Financial Analysts*, 1996, p.47-56.
- KATZ R. W., PARLANGE M. B., NAVÉAU P., "Statistics of extremes in hydrology", *Advances in Water Resources* 25, 2002, p. 1287-1304.
- CLAJNMIC H., « Niveaux de retour pour les vitesses extrêmes des vents », *XXXV<sup>es</sup> Journées de Statistique*, Lyon, 2003.
- KREHBIEL T., ADKINS L. C., "Price risk in the NYMEX energy complex: An extreme value approach", *Journal of Futures Markets* 25, 2005, p. 309-337.
- KUPIEC P., "Techniques for Verifying the Accuracy of Risk Management Models", *Journal of Derivatives* 3, 1995, p. 73-84.
- LONGIN F. M., "From value at risk to stress testing: The extreme value approach", *Journal of Banking and Finance* 24, 2000, p. 1097-1130.
- LONGIN F. M., « Value at risk: Une nouvelle approche fondée sur les valeurs extrêmes », *Annales d'économie et de statistique* 52, 1998, p. 23-51.
- LYE L., HAPUARACHCHI K., RYAN S., "Bayes estimation of the extrema-value reliability function", *IEEE Transactions on Reliability* 42, 1993, p. 641-644.
- MANDELBROT B. B., « Sur certains prix spéculatifs : faits empiriques et modèle basé sur les processus stables additifs non gaussiens de Paul Lévy », *Comptes rendus à l'Académie des Sciences* 254, 1962, p. 3968-3970.

- MANDELBROT B. B., "The variation of certain speculative prices", *Journal of Business* 26, 1963, p. 394-419.
- MANGANELLI S., ENGLE R. F., "Value at risk models in finance", *European central bank, Working paper series* 75, 2001.
- MARIMOUTOU V., RAGGAD B., TRABELSI A., "Extreme Value Theory and Value at Risk: Application to Oil Market", *Energy Economics* 31, 2009, p. 519-530.
- MATTHYS G., BEIRLANT J., "Adaptive threshold selection in tail index estimation", *Working paper University Center for Statistics at the Catholic University Leuven*, 2000.
- MCNEIL A. J., *Calculating quantile risk measures for financial time series using extreme value theory*, Department of Mathematics, ETH. Swiss Federal Technical University E-Collection, 1998, [<http://e-collection.ethbib.ethz.ch/>].
- MCNEIL A. J., FREY R., "Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach", *Journal of Empirical Finance* 7, 2000, p. 271-300.
- MCNEIL A. J., SALADIN T., "The peaks over thresholds method for estimating high quantiles of loss lois", Department of Mathematic ETH Zurich, 1997.
- PICKANDS J., *Statistical inference using extreme order statistics*, Annals of Statistics 3, 1975, p. 119-131.
- RAGGAD B., *Gestion des risques : théorie et application au marché pétrolier*, thèse, Institut Supérieur de Gestion de Tunis, Université de Tunis, 2007.
- REISS R., THOMAS M., *Statistical Analysis of Extreme Value with Applications to Assurance, Finance, Hydrology and Other Fields*, Basel, Birkhauser, Verlag, 2001.
- ROOTZEN H., TAJVIDI N., "Extreme value statistics and wind storm losses: a case study", *Scandinavian Actuarial Journal*, 1997, p. 70-94.
- SMITH R. L., "Maximum likelihood estimation in a class of nonregular cases", *Biometrika* 72, 1985, p. 67-90.
- SMITH R. L., "Extreme value theory based on the  $r$  largest annual events", *J. Hydrology* 86, 1986, p. 27-43.
- SMITH R. L., "Estimating tails of probability lois", *The Annals of Statistics* 3, 1987, p. 1174-1207.
- SMITH R. L., *Statistics of extremes, with application in environment, insurance and finance*, 2003, [<http://www.stat.unc.edu/postscript/rs/semstatrls.ps>].
- SMITH R. L., "Extreme value statistics in meteorology and the environment. In environmental Statistics", Chapter 8, *NSF-CBMS conference notes*, 2001, p. 300–357, [available at; [www.unc.edu/depts/statistics/postscript/rs/envnotes.pdf](http://www.unc.edu/depts/statistics/postscript/rs/envnotes.pdf)].
- THATCHER R. A., "The Long-Term Pattern of Adult Mortality and the Highest Attained Age", [with discussion], *Journal of the Royal Statistical Society Series A* 162, 1999, p. 5–43.
- TSAY R. S., *Analysis of Financial Time Series*, New York, Wiley, 2002.
- YAFFEE R. A., McGEE M., *Time Series Analysis and Forecasting with Applications of SAS and SPSS*, Academic Press, INC, 1999.

## ANNEXE

Panel A : ratios des violations de VaR (en %). Les nombres entre parenthèses représentent le rang du modèle par rapport à l'ensemble des modèles étudiés.						
$\alpha$	5 %		1 %		0.5 %	
	Gauche	Droite	Gauche	Droite	Gauche	Droite
Normale	4.6 % (3)	3.6 % (5)	1.4 % (3)	1.2 % (2)	0.9 % (3)	0.8 % (2)
GARCH-N	5.1 % (1)	4 % (4)	1.4 % (3)	1.4 % (4)	1 % (4)	0.8 % (2)
GARCH-T	5.6 % (4)	4.3 % (3)	1 % (1)	0.8 % (2)	0.5 % (1)	0.5 % (1)
GPD	5.1 % (1)	4.8 % (2)	0.9 % (2)	0.8 % (2)	0.5 % (1)	0.5 % (1)
GPDC	5.2 % (2)	5.1 % (1)	0.9 % (2)	1 % (1)	0.4 % (2)	0.5 % (1)
Panel B : test de couverture inconditionnelle						
Varp	VaR.95		VaR.99		VaR.995	
	Gauche	Droite	Gauche	Droite	Gauche	Droite
Normale	<u>0.215</u>	0	0.021	<u>0.29</u>	0.001	0.005
GARCH-N	<u>0.725</u>	0.001	0.014	0.021	0	0.022
GARCH-T	<u>0.09</u>	0.021	<u>0.972</u>	<u>0.173</u>	<u>0.735</u>	<u>0.935</u>
GPD	<u>0.829</u>	<u>0.637</u>	<u>0.466</u>	<u>0.296</u>	<u>0.735</u>	<u>0.767</u>
GPDC	<u>0.448</u>	<u>0.725</u>	<u>0.615</u>	<u>0.908</u>	<u>0.025</u>	<u>0.935</u>
Panel C : test de couverture conditionnelle						
VaRp	VaR.95		VaR.99		VaR.995	
	Gauche	Droite	Gauche	Droite	Gauche	Droite
Normale	0	0	0	<u>0.511</u>	0	0.011
GARCH-N	0	0.04	0.001	<u>0.364</u>	0.022	<u>0.223</u>
GARCH-T	<u>0.585</u>	0.002	0.028	0.03	0	<u>0.057</u>
GPD	<u>0.09</u>	0.033	<u>0.776</u>	<u>0.295</u>	<u>0.83</u>	<u>0.894</u>
GPDC	<u>0.665</u>	<u>0.151</u>	<u>0.674</u>	<u>0.642</u>	<u>0.079</u>	<u>0.894</u>

Les panels A, B et C rapportent successivement les ratios des violations des VaR journalières, les p-values du test de couverture inconditionnelle et du test de couverture conditionnelle pour l'ensemble des modèles étudiés. Les modèles sont successivement la loi normale inconditionnelle (Normale), le modèle GARCH à innovations normale (GARCH-N), le modèle GARCH à innovations Student (GARCH-T), le modèle généralisé de Pareto (GPD) et le modèle GPD Conditionnel (GPDC). Notons qu'une p-value supérieure à 5 % indique que le modèle en question est adéquat (valeurs soulignées).

TABLEAU 4. Résultats de Backtesting (Brent)

Panel A : ratios des violations de VaR (en %). Les nombres entre parenthèses représentent le rang du modèle par rapport à l'ensemble des modèles étudiés						
$\alpha$	5 %		1 %		0.5 %	
	Gauche	Droite	Gauche	Droite	Gauche	Droite
Normale	4.6 % (2)	3.9 % (5)	1.6 % (2)	1.4 % (4)	1.3 % (4)	1.1 % (4)
GARCH-N	4.7 % (3)	4.2 % (4)	1.6 % (2)	1.2 % (2)	1 % (3)	0.7 % (3)
GARCH-T	5.7 % (4)	4.4 % (2)	1.1 % (1)	0.7 % (3)	0.6 % (2)	0.4 % (2)
GPD	5.4 % (2)	4.2 % (3)	1.1 % (1)	1.2 % (2)	0.6 % (2)	0.7 % (3)
GPDC	5 % (1)	4.9 % (1)	0.9 % (1)	1.1 % (1)	0.5 % (1)	0.5 % (1)
Panel B : test de couverture inconditionnelle						
VaRp	VaR.95		VaR.99		VaR.995	
	Gauche	Droite	Gauche	Droite	Gauche	Droite
Normale	<u>0.174</u>	0	0	0.005	0	0
GARCH-N	<u>0.413</u>	0.006	0	<u>0.275</u>	0	<u>0.126</u>
GARCH-T	<u>0.029</u>	<u>0.057</u>	<u>0.694</u>	<u>0.063</u>	<u>0.249</u>	<u>0.19</u>
GPD	<u>0.157</u>	<u>0.006</u>	<u>0.415</u>	<u>0.275</u>	<u>0.249</u>	<u>0.126</u>
GPDC	<u>0.879</u>	<u>0.776</u>	<u>0.438</u>	<u>0.694</u>	<u>0.979</u>	<u>0.979</u>
Panel C : test de couverture conditionnelle						
VaRp	VaR.95		VaR.99		VaR.995	
	Gauche	Droite	Gauche	Droite	Gauche	Droite
Normale	0	0	0	0	0	0
GARCH-N	<u>0.569</u>	0.023	0.001	<u>0.507</u>	0	<u>0.251</u>
GARCH-T	<u>0.006</u>	<u>0.141</u>	<u>0.788</u>	0.017	<u>0.426</u>	<u>0.396</u>
GPD	0	0.023	0.009	<u>0.507</u>	0.024	<u>0.251</u>
GPDC	<u>0.464</u>	<u>0.941</u>	<u>0.503</u>	<u>0.788</u>	<u>0.887</u>	<u>0.887</u>

Les panels A, B et C rapportent successivement les ratios des violations des VaR journalières, les p-values du test de couverture inconditionnelle et du test de couverture conditionnelle pour l'ensemble des modèles étudiés. Les modèles sont successivement la loi normale inconditionnelle (Normale), le modèle GARCH à innovations normale (GARCH-N), le modèle GARCH à innovations Student (GARCH-T), le modèle généralisé de Pareto (GPD) et le modèle GPD Conditionnel (GPDC). Notons qu'une p-value supérieure à 5% indique que le modèle en question est adéquat (valeurs soulignées).

TABLEAU 5. Résultats de Backtesting (WTI)