

Article

## Portfolio Value at Risk Estimate for Crude Oil Markets: A Multivariate Wavelet Denoising Approach

Kaijian He <sup>1,2,\*</sup>, Kin Keung Lai <sup>2</sup> and Guocheng Xiang <sup>1</sup>

<sup>1</sup> Business School, Hunan University of Science and Technology, Xiangtan, Hunan 411201, China; E-Mail: xiangguoch@sohu.com

<sup>2</sup> Department of Management Sciences, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong; E-Mail: mskklai@cityu.edu.hk

\* Author to whom correspondence should be addressed; E-Mail: paulhekj@gmail.com; Tel.: +86-731-58291023; Fax: +86-731-58290306.

Received: 13 February 2012; in revised form: 5 April 2012 / Accepted: 12 April 2012 /

Published: 18 April 2012

---

**Abstract:** In the increasingly globalized economy these days, the major crude oil markets worldwide are seeing higher level of integration, which results in higher level of dependency and transmission of risks among different markets. Thus the risk of the typical multi-asset crude oil portfolio is influenced by dynamic correlation among different assets, which has both normal and transient behaviors. This paper proposes a novel multivariate wavelet denoising based approach for estimating Portfolio Value at Risk (PVaR). The multivariate wavelet analysis is introduced to analyze the multi-scale behaviors of the correlation among different markets and the portfolio volatility behavior in the higher dimensional time scale domain. The heterogeneous data and noise behavior are addressed in the proposed multi-scale denoising based PVaR estimation algorithm, which also incorporates the mainstream time series to address other well known data features such as autocorrelation and volatility clustering. Empirical studies suggest that the proposed algorithm outperforms the benchmark Exponential Weighted Moving Average (EWMA) and DCC-GARCH model, in terms of conventional performance evaluation criteria for the model reliability.

**Keywords:** Portfolio Value at Risk; multivariate wavelet analysis; Exponential Weighted Moving Average (EWMA) model; DCC-GARCH model; multivariate time series model; heterogeneous market hypothesis

---

## 1. Introduction

With the increasing level of global economic integration and the development of financial market, in the crude oil markets we have seen a more volatile as well as more integrated environment [1,2]. On the one hand, the market is influenced by increasingly complicated factors, aside from the fundamental causes such as the basic supply and demand relationship. This is exemplified by the recent intensifying financial crisis, which brought higher level of fluctuations as observed in the crude oil markets. The traditional risk measurement tools such as standard deviations and equilibrium analysis can no longer provide sufficient level of risk coverage when potential risk sources are increasingly differentiated and transient data behaviors are more prevalent. On the other hand, markets across the world are increasingly integrated with the development in both its technological infrastructure and deregulation wave. This is also evidenced from the high level of correlations between both WTI and Brent markets crude oil portfolio forms indispensable part of investment decision making process in the crude oil markets. When the multi-asset portfolio is constructed to reduce the risk level through diversification during portfolio allocation, the reliable and accurate risk measurement for the portfolio is essential for the portfolio management and risk management and affects critically the portfolio performance.

Although the risk measurement research in the crude oil markets has drawn significant research attempts over years, majority of innovative research methodologies are developed for the risk measurement for single asset portfolio. Till now there is only limited attempts to explore various theoretical and empirical issues for portfolio VaR estimation in the literature. One reason is that compared to the numerous attempts to model univariate time series, researches directed toward the understanding and modeling of multivariate time series are much more restricted due to the level of technical difficulties involved. Another reason is that portfolio VaR needs to deal with further complication of modeling second moments including correlation and covariance matrix among multiple assets than the single asset VaR, which involves exponentially increasing level of computational complexity. Typical models used in this field include the much restricted Exponential Weighted Moving Average (EWMA) and the multivariate VARMA-GARCH model for portfolio risk measurement.

Meanwhile, the wavelet based approach is also gaining momentum in the economics and finance literature [3–5]. Firstly, wavelet analysis has been used to identify time varying behaviours of market behaviours, key economic variables and their correlation evolutions. A typical example would be the studies on multi-scale price movements and lead lag relationships in various financial markets [3,6–10]. Empirical studies in these markets reveal the multi-scale structure of prices and evolution of their correlation. Secondly, wavelet analysis is used to smooth data at finer scales. As the magnitude of noise is time varying and the signal to noise ratio (*i.e.*, the threshold level set for the distinction) may be subject to investors' preferences, data smoothness would improve if smoothing is conducted at a finer time scale domain, and thus improved model fitness and performance would result. Empirical studies based on wavelet shrinkage techniques in both electricity and foreign exchange markets have confirmed the potential performance improvement. For instance, Meng *et al.* [11], Amjady and Keynia [12] and Aggarwal *et al.* [13] based their neural network forecasting model on wavelet pre-processed data series and obtained positive performance improvement [11–13]. Thirdly, wavelet analysis has been used to decompose data series for a finer level of modelling accuracy.

Decomposed data series at lower scales are considered dominated by trends and are modelled using econometric approaches. Decomposed data series at higher scales are considered mostly dominated by price volatilities and are modelled using time series approaches. Results of experiments in several energy markets have confirmed the performance improvement [14,15]. For instance, Xu *et al.* [16] attempted to combine wavelet analysis with support vector machine and obtained positive performance improvement in empirical studies on the Australian electricity market [16]. Conejo *et al.* [17] combined wavelet analysis and ARMA model to analyze the Spanish electricity market and obtained positive results [17].

In the crude oil forecasting and risk measurement literature, we have only witnessed limited efforts. Most of them are accumulating quite recently. These attempts mostly use the wavelet analysis to pre-process the data before time series techniques and machine learning techniques are used to analyze the data and make forecasts. For instance, Yousefi *et al.* [15] used wavelet analysis to decompose crude oil price and extended them directly to make forecasts [15]. de Souza e Silva *et al.* [18] uses wavelet analysis to remove high frequency noises, then uses Hidden Markov Model (HMM) to predict future price movement [18]. Naccache [19] uses wavelet analysis analyze the relationship between macroeconomic conditions proxied by Morgan Stanley Capital International (MSCI) and the crude oil price at finer scales [19]. We have conducted some exploratory studies on multi-scale based risk measure estimates. He *et al.* [20] used the Morphological Component Analysis with wavelets as bases to investigate the heterogeneous market microstructure [20]. But the literature on the forecasting aspects of crude oil price is limited to this extent so far.

Meanwhile, risk measurement in the crude oil markets remains another important research topic, but has received much less attention compared to the progress made in the forecasting field. Most of current approaches use the multi-resolution analysis capability of wavelet analysis to look into the past price or correlation behaviours in the capital markets [21,22], e.g., it has been used to analyze the risk distribution across differences the scales [23–30]. Some researches such as that by Karandikar *et al.* [31] has gone a step further to estimate VaR with this information taken into account. VaR estimated following the traditional approaches is adjusted by the proportion of risks that corresponds to the investment time horizon [31]. These approaches are unique in that VaRs are tailored to investor's investment profile [24]. However, backtesting these VaR estimates may involve additional complexities. Unless risks at all levels are taken into account, using these adjusted VaRs to calculate VaR exceedances may result in the underestimated exceedance levels. The reliability may also be statistically overestimated as a result. Aside from the backtesting issue, the performance of these approaches is also constrained by the accuracy of the traditional methodology used. Thus, when it comes to models' reliabilities and accuracies, these approaches are more useful as risk analysis tool than methodologies that bring further performance improvement. Recently He *et al.* [32] used wavelet analysis to analyze the risk structure and construct the VaR estimation algorithm based on it [32]. It is univariate in nature and not directly applicable to the much more complicated multivariate environment. The risk measurement literature is limited to this extent so far concerning the portfolio risk measurement in the multi-scale domain.

Therefore, this paper proposes a novel multivariate wavelet denoising approach for estimating Value at Risk for portfolio in the crude oil markets. Following Heterogeneous Market Hypothesis (HMH), it is assumed that the underlying DGPs vary in their scales, corresponding to different scales of investment. Thus taking different investment scales (*i.e.*, portfolio sizes) as the discriminatory factor, wavelet based

denoising algorithm is used to gain insights into heterogeneous market structures categorized roughly on scales of investment. In this paper, we introduce the multivariate wavelet analysis in higher dimensional space to analyze the second moments evolution of crude oil portfolio data. The introduced multivariate wavelet denoising methodology is then used to extract data of distinct characteristics, which are modeled further incorporating the Multivariate GARCH (MGARCH) framework.

The theoretical motivation for the present research is that recognizing these complicated and heterogeneous nature of the crude oil markets, there exists diverse patterns of behavior of different participants, which is the target for the separation and extraction process. These participants need to be categorized and separated from the original data based on their unique behavioral patterns, resulting from different investment horizons and strategies. For example, the behavior of noise traders are temporal and nonlinear, usually resulting from highly speculative behavioral patterns, while the behavior of fundamental investors are characterized with features subject to the extensive studies over years, including auto correlation, heteroscedasticity, mean reverting and long memory, *etc.*

The practical and technical motivation for the present research is that the challenge to improve estimation reliability and accuracy has been difficult to improve beyond the traditional EWMA and MGARCH models. Recent empirical researches suggest that this is due to its complicated and nonlinear data nature, compared to the linear assumptions adopted before. Besides, crude oil markets typically represent a more volatile and risky environment than the traditional financial markets and demonstrate unique characteristics due to several reasons. Firstly, crude oil is traded less frequently than equities. There are considerable transaction and opportunity costs involved, resulting in higher levels of fluctuations and relatively lower levels of efficiency. Secondly, price movements in crude oil markets are frequently influenced by numerous factors, such as weather conditions, political stability, economic prospects, consumer expectations and business indicators, *etc.* The complex dynamics involved exhibit multi-scale non-linear characteristics, which are beyond explanatory capabilities of current models and methodologies for forecasting and risk measurement purposes. Despite its significance, forecasting of crude price and the measurement of risk in crude oil markets have not received sufficient attentions, as compared to equities markets. Research in this paper attempts to address the unique characteristics of crude oil markets when constructing useful models for forecasting and risk measurement purposes and to fill the gap in the relevant literature. Meanwhile, wavelet denoising in higher dimension involves directional projection of time series, *i.e.*, multivariate time series can be projected into different directions, normally vertical, horizontal and diagonal. Projections into different directions are associated with different correlations among time series at different time lags.

The key findings and contributions include the following aspects. There have been limited attempts in the literature to model the conditional correlations among portfolio using multivariate GARCH model [33,34]. The most related one is Vacha and Barunik [35]. However, it ignores the dependence structure among portfolio assets, which can only be investigated in the higher dimension. Besides, it does not deal with the utilization of these time varying correlation information into the modeling of risk evolutions, which this paper addresses using the DCC model in the multi-scale framework. Meanwhile, most research focuses on the single variate VaR estimation, among which the most related one is Cifter [36] that is mainly based on univariate wavelet analysis. This paper instead introduces multivariate wavelet analysis in the estimation of portfolio Value at Risk. Few studies in the literature have

investigated the the portfolio risk measurement with time varying correlation. Work in this paper represents one of few attempts to investigate the estimation of portfolio VaR using multivariate wavelet based techniques.

The rest of the paper proceeds as follows: Section 2 provides a brief account and also literature review on the relevant theories used to construct the proposed algorithm. This includes the portfolio value at risk, multivariate GARCH models and multivariate wavelet analysis theories. Section 3 proposes the multivariate wavelet denoising approach for estimating portfolio value at risk, with detailed economic rationale explained and numerical procedures. The performance of the proposed algorithm against benchmark alternatives are evaluated in Section 4. The empirical studies are conducted in the benchmark US WTI and Europe Brent markets and experiment results are reported to justify the superiority of the proposed algorithm. Section 5 provides some concluding remarks.

## 2. Relevant Theories

### 2.1. Portfolio Value at Risk Estimates

The Value at Risk (VaR) is the percentile based risk sensitive criteria [37]. It is the single statistics summarizing the maximal downside risk given the confidence level and holding horizon. In the more general case of a portfolio of  $n$  assets with returns  $r_t$ , ( $t = 1, 2, \dots, n$ ) and portfolio weight  $\omega_t$ , ( $t = 1, 2, \dots, n$ ), the VaR is defined for a portfolio worth  $X$  at the confidence level  $cl$  over the holding period  $h$  as in (1).

$$P(\omega'_t r_t < VaR_{cl}) = \alpha \quad (1)$$

where  $\alpha = 1 - cl$  and  $P(\cdot)$  is the probability density function,  $VaR_{cl}$  is the VaR threshold at the confidence level  $cl$ .

Similar to the univariate case, the main approaches to estimate portfolio VaR can also be roughly categorized into three categories, *i.e.*, parametric, non-parametric and semi-parametric, but with much less attention than in the univariate case [38,39].

The parametric approach usually refers to a series of models under the well-known variance-covariance framework. Assuming the multivariate normal distribution, the VaR is defined as in (2) [39].

$$VaR(h, \alpha) = [-h\omega\mu + \sqrt{h}\sqrt{\omega \sum \omega^T z_\alpha}]X \quad (2)$$

where  $z_\alpha$  refers to the standard normal distribution,  $h$  is the holding period,  $\sum$  is the variance-covariance matrix,  $\omega$  is the weight matrix.

The  $\mu$  and  $\sum$  can be estimated in the conditional framework such as the Riskmetrics's multivariate Exponential Weighted Moving Average (EWMA) and Multivariate Generalized Autoregressive Conditional Heteroscedasticity (MGARCH). As is well known, EWMA can be viewed as a special case of MGARCH model and has served as the benchmark model in the literature in the past. However, unlike in the univariate case, multivariate GARCH has not been widely accepted yet as the benchmark model.

The multivariate time series models such as MGARCH are extension of univariate time series models in the higher dimensional space. However, it is much more complex. It suffers from a problem called

“curve of dimensions”, which refers to the fact that number of parameters to estimate and the associated computational complexity increase exponentially with the increasing number of variables.

There are mainly three approaches to model the multivariate conditional mean and conditional standard deviation in the literature. To model multivariate conditional mean, Vector Autoregressive (VAR) model and Vector Autoregressive Moving Average (VARMA) are often utilized. In practice the VARMA is too complicated to estimate. The VAR model is usually resorted to in practice to model the multivariate conditional means. There are many multivariate arch type models developed over years to model the multivariate conditional variance-covariance matrix [40]. Typical examples include BEKK, Dynamic Conditional Correlation (DCC) and Constant Conditional Correlation (CCC) [41–44]. Among them, DCC is the most popular because it relaxed the fixed correlation assumptions in the other two approaches. Both Engle [42] and Tse and Tsui [45] propose slightly different DCC MV-GARCH specifications respectively [42,45]. Between them, The DCC MV-GARCH proposed by Engle [42] is used as it allows more flexible GARCH like specification of dynamic correlations than the simpler weighted sum specification of correlations proposed by Tse and Tsui [45].

The DCC MV-GARCH model is specified as in (3) [42].

$$r_t = \alpha_0 + \alpha_1 r_t + \mu_t + \varepsilon_{t+1} \quad (3)$$

where  $r_t$  refers to the return at time  $t$ ,  $r_t \sim N(0, D_{t+1} R D_{t+1})$ ,  $\mu_t$  is the N-dimensional vector of conditional mean,  $D_t^2$  can be defined as any univariate GARCH model as  $D_t^2 = \text{diag}\{\omega_i\} + \text{diag}\{k_i\} r_{t-1} r'_{t-1} + \text{diag}\{\lambda_i\} D_{t-1}^2$ ,  $\varepsilon_t = D_t^{-1} r_t$ .

The dynamic conditional correlation matrix  $R_t$  of  $\varepsilon_t$  at time  $t$  is defined as in (4).

$$R_t = \text{diag}(Q_t)^{-1} Q_t \text{diag}(Q_t)^{-1} \quad (4)$$

where the  $Q_t$  is  $N \times N$  symmetric positive definite unconditional correlation matrix of  $\varepsilon_t$  as in (5).

$$Q_t = S \circ (\mu - A - B) + A \circ \varepsilon_{t-1} \varepsilon'_{t-1} + B \circ Q_{t-1} \quad (5)$$

Only until recently some limited attempts emerged to investigate the MGARCH based VaR estimates, *i.e.*, Rombouts and Verbeek [46], Ku and Wang [47], Huang *et al.* [48].

The non-parametric approach includes Historical Simulation (HS) and Monte Carlo (MC) simulation, both directly extendable from univariate to multivariate case and serve as the benchmark in the literature [39]. The Monte Carlo simulation relies on the repeated draws from the pre-specified distributional assumptions to simulate the real-life scenarios and estimate the portfolio VaR accordingly. However, MC is unconditional in nature and is biased toward normal distribution. It is also limited to the pre-specified distributional assumptions and has difficulty in replicating the extreme event scenario [38]. Some recent attempts have been made to tackle these issues. For instance, Glasserman *et al.* [49] has relaxed the distributional assumption behind Monte Carlo simulation to multivariate t-distribution to improve the performance [49]. The historical simulation approach uses the past information over the pre-specified window to infer directly the future portfolio VaR. This approach is sensitive to the choice of the window’s timing and length and overestimate VaR because of the uniform treatment of risk level when transient events are rare in nature. When transient events are excluded from the window, it would lead to underestimated PVaR.

The semi-parametric approach includes some newly emerging techniques such as copula theory, fuzzy logic, genetic algorithm, *etc.* [48,50–52]. However, to the best of our knowledge, there has been no attempts in analyzing the multi-scale portfolio risk behaviors and constructing the PVaR estimation algorithm using multivariate wavelet analysis in the higher dimension.

## 2.2. Multivariate Wavelet Denoising Algorithm

In recent years, significant progress has been made in multivariate research in engineering fields, which has resulted in significant performance improvement that provides inspiration to the current research. Of more particular relevance to this research is multivariate denoising techniques in the image processing field, which consist of a series of algorithms including spatial filtering and transform domain filtering [53–55]. There are several detailed review and surveys in the literature including Buades *et al.* [56], Mallat and Peyre [57].

In the literature, the traditional denoising algorithms are known to over-suppress high-frequency details [58]. Thus there are several nonlinear alternatives proposed in the literature. For example, Pang and Yang [59] proposes the projection gradient method to solve the transformed image denoising problem using the duality strategy [59]. Meanwhile, multivariate wavelet analysis has received considerable attentions in image denoising fields [53,55]. Multivariate wavelet based image denoising has been explored extensively and positive performance has been reported in different disciplines. For example, in medical image processing, Chen *et al.* [60] proposes a homogeneity similarity based denoising algorithm, which can retain edges and contrast information while avoiding artifacts [60]. Gorgel *et al.* [61] propose wavelet denoising algorithms with adaptive thresholding functions for enhancement of mammographic images in breast cancer detection [61]. Kim *et al.* [58] proposes the use of filtered directional bases and frequency-adaptive shrinkage to reduce noise in edge regions, which is claimed to retain sharp details [58]. Kumar and Agnihotri [62] have attempted discrete wavelet transform and stationary wavelet transform with different threshold techniques in denoising the ECG and MRI images and have found positive performance [62]. Rabbani and Gazor [63] derive estimators using maximum a posteriori and minimum mean squared error methods in different sparse domains projected with discrete wavelet transform and its variant contourlet and curvelet transforms. It was found that the wavelet based method provides superior performance [63]. Saeedi *et al.* [64] used fuzzy logic to incorporate the inter-relation between different channels during the wavelet denoising process and found improved performance compared with the standard single channel wavelet denoising methods [64]. Tian *et al.* [65] proposed a nonparametric model to formulate the marginal distribution of wavelet coefficients in an adaptive manner and integrated it into a Bayesian inference framework to drive a maximum a posterior estimation based image denoising method with improved performance [65]. Zifan *et al.* [66] used the mutliwavelet denoising methods to denoise microarray images and found that the results significantly outperformed wavelet and Wiener filter based methods [66].

Developed in the 1980s, wavelet based transform and its various extension functions have been proven useful for many different disciplines such as signal analysis, numerical analysis, image compression and other applied mathematics and engineering branches [67]. The unique feature of wavelet analysis is its localization capability in both time and scale domain during the analysis. This is in contrast to the traditional spectrum analysis tool that signifies the patterns at different frequency at the cost of ignoring

details at time scales. As financial data usually exhibit nonstationary time varying characteristics, wavelet analysis is appealing when it comes to financial data analysis.

To define wavelet analysis and multi-resolution analysis, in the one dimensional case, firstly a refinable (scale) function  $\phi : \mathbb{R} \rightarrow \mathbb{C}$  is defined as the solution to the two scale recursion difference equation in (6).

$$\phi(x) = \sqrt{2} \sum_{k \in \mathbb{Z}} h_k \phi(2x - k) \tag{6}$$

where  $h_{k,k \in \mathbb{Z}} \in \mathbb{L}^2(\mathbb{Z})$  are recursion coefficients. Then given some additional constraints, (6) gives rise to the wavelet functions  $\varphi$  which forms the orthonormal basis for  $\mathbb{L}^2(\mathbb{R})$  as in (7).

$$\varphi_{j,k}(x) = 2^{-j} \varphi(2^j x - k) \tag{7}$$

The constraints may include orthogonality, compactness of support, vanishing moments, symmetry, regularity, smoothness, time frequency window and decay, etc. [68]. The typical wavelets used include Haar, Daubechies, Coiflets, discrete Meyers, biorthogonal and reverse biorthogonal, etc.

Utilizing the wavelets and their appealing time scale localization characteristics, the original data are projected into the multi-scale domain using wavelet transform as in (8).

$$W(u, s) = \int_{-\infty}^{\infty} x(t) \frac{1}{\sqrt{s}} \psi\left(\frac{t-u}{s}\right) dt x(t) = \frac{1}{C_\psi} \int_0^{\infty} \int_{-\infty}^{\infty} W(u, s) \psi_{u,s}(t) du \frac{ds}{s^2} \tag{8}$$

where  $\psi\left(\frac{t-u}{s}\right)$  refers to the wavelet families translated by location parameter  $u$ , which corresponds to time scale localization, and dilated by scale parameter  $s$ , which corresponds to the frequency scale localization. In practice, wavelet analysis including wavelet transform and synthesis is performed at discrete time interval due to the data availability issue and to avoid the exponential increasing level of information redundancy.

Wavelet analysis in higher dimension are extensions to the wavelet analysis in one dimension, with additional complications. There are mainly two approaches to extend and define wavelet analysis and multi-resolution analysis in the higher dimensional case, i.e., tensor product multi-resolution analysis or the nonseparable scheme [67]. Between them, tensor product based approach has both computational and intuitive appeals. In the tensor product scheme, the 2-dimensional wavelet is constructed by taking the tensor product functions generated by two one-dimensional bases as in (9).

$$\Psi_{j_1,k_1;j_2,k_2}(x_1, x_2) = \psi_{j_1,k_1}(x_1) \psi_{j_2,k_2}(x_2) \tag{9}$$

where both the resulting basis  $\Psi_{j_1,k_1;j_2,k_2}(x_1, x_2)$  and the original bases  $\psi_{j_1,k_1}(x_1)$  and  $\psi_{j_2,k_2}(x_2)$  are orthonormal wavelet bases for  $\mathbb{L}^2(\mathbb{R}^2)$ .

For all  $f, g \in \mathbb{L}^2(\mathbb{R}^2)$ , the decomposition and reconstruction in the two dimensional case is defined as in (10) and (11) respectively.

$$\int_0^{\infty} \frac{da}{a^{n+1}} \int_{-\infty}^{\infty} db (T^{wav} f)(a, b) \overline{(T^{wav} g)(a, b)} = C_\psi \langle f, g \rangle \tag{10}$$

$$f = C_\psi^{-1} \int_0^{\infty} \frac{da}{a^{n+1}} \int_{\mathbb{R}^n} db (T^{wav} f)(a, b) \psi^{a,b} \tag{11}$$

where  $(T^{wav} f)(a, b) = \langle f, \psi^{a,b} \rangle$ ,  $\psi^{a,b}(x) = a^{-n/2} \psi(\frac{x-b}{a})$ ,  $a \in \mathbb{R}_+$ ,  $a \neq 0$ ,  $b \in \mathbb{R}^n$ .

With wavelet analysis, Multi-Resolution Analysis (MRA) can be conducted to analyze the time and frequency characteristics of the data. In the more general case in higher dimension, the MRA on  $\mathbb{L}_2(\mathbb{R}^d)$  is defined as a doubly infinite nested sequence of subspaces  $V_j$  of  $\mathbb{L}_2(\mathbb{R}^d)$  as in (12) [69].

$$\dots \subset \mathbf{V}_2 \subset \mathbf{V}_1 \subset \mathbf{V}_0 \subset \mathbf{V}_{-1} \subset \mathbf{V}_{-2} \subset \dots \subset \mathbb{L}^2(\mathbb{R}^d) \cap_{j \in \mathbb{Z}} \mathbf{V}_j = \{0\} \quad \text{and} \quad \overline{\cup_{j \in \mathbb{Z}} \mathbf{V}_j} = \mathbb{L}^2(\mathbb{R}^d)$$

in which

$$\mathbf{V}_j = \bigotimes_{i=1}^d V_{j,(i)} \subset \mathbb{L}^2(\mathbb{R}^d) \tag{12}$$

where  $\otimes$  is the tensor product operator.

Then for each  $V_{j-1}$  in  $V_j$ , the orthogonal complement space  $W_j$  is defined as in (13).

$$\mathbf{V}_{j-1} = \mathbf{V}_j \oplus \mathbf{W}_j \tag{13}$$

where the orthogonal complement space  $\mathbf{W}_j$  consists of several pieces, *i.e.*, one d-variate scaling function as in (14) and  $2^d - 1$  d-variate wavelet functions as in (15).

$$\phi(x_1, \dots, x_d) = \prod_{i=1}^d \phi(i)(x_i) \tag{14}$$

$$\psi^{(l)}(x_1, \dots, x_d) = \prod_{i=1}^d \xi_{(i)}(x_i), \text{ with } \xi = \phi \text{ or } \psi, \text{ but not all } \xi = \phi \tag{15}$$

Here we take the two dimensional wavelet as a special case for illustration purpose. In the 2-dimensional case, three subspaces are generated with three orthonormal wavelet bases, *i.e.*, horizontal wavelet  $\Psi^{horizontal}(x, y)$ , vertical wavelet  $\Psi^{vertical}(x, y)$  and diagonal wavelet  $\Psi^{diagonal}(x, y)$  [67].

Over years, significant research interests have focused on using the wavelet denoising techniques in the univariate case, but much less attentions on the denoising of multivariate time series data. The basic procedure of multivariate wavelet denoising technique is as follows:

1. It firstly projects the original data series into different scales using wavelet transform as in (10). The resulting coefficients would include approximation coefficients as well as horizontal, vertical and diagonal directions.
2. For approximation and direction coefficients at each direction, the threshold is chosen specifically at different scales for different directions and the wavelet coefficients are processed by either suppression or shrinkage.
3. Using the denoised wavelet coefficients and the scale chosen, the processed wavelet coefficients are reconstructed into the unified denoised data series using wavelet synthesis as in (11).

In the multivariate setting, given the multivariate time series  $X$ , the denoising algorithm assumes that it consists of both deterministic data  $D$  and undesirable stochastic noises  $\epsilon$  in a linearly fashion. Applying the multivariate wavelet analysis, this relationship is defined as in (16).

$$O \equiv vX = vD + v\epsilon \tag{16}$$

where  $v$  is an  $N \times N$  orthonormal matrix,  $O$  is the  $N$  dimensional vector of wavelet transform coefficients  $O_l : l = 0, \dots, N - 1$ .

The mainstream threshold selection rules include Universal and Steins Unbiased Risk Estimate (SURE) [70]. Different threshold selection rules have different target and considerations when setting the noise reduction target. For instance, the universal threshold selection rule aims at reducing the noises at maximum level possible, so the threshold is selected as  $\eta^U = \hat{\sigma}\sqrt{2\log N}$ , where  $N$  is the number of wavelet coefficients and  $\hat{\sigma}$  is the estimate of the volatility level. Statistically this gives the upper bound to noises signals given significant sample size  $N$  and normally distributed noise series. However, this method achieves the maximal level of smoothness at the cost of lower goodness-of-fit. The denoised data risk losing some important data features.

The mainstream thresholding functions are hard and soft thresholding [70,71]. The hard thresholding is the high pass filter, which suppresses the wavelet coefficients below the chosen threshold values and leaves the remaining coefficients intact as in (17).

$$O_l^{(ht)} = \begin{cases} 0 & \text{if } |O_l| \leq \delta \\ O_l & \text{otherwise} \end{cases} \quad (17)$$

where  $O_t$  refers to the wavelet coefficients,  $\eta$  is the set threshold value.

The soft thresholding focus on the signal smoothing. It suppresses the wavelet coefficients below the set threshold value and subtracts the threshold value from the remaining wavelet coefficients. Compared with the hard thresholding, the data processing following soft threshold selection rules are smoother but lost the abrupt changes in the original data. It filters the signal as in (18):

$$O_l^{(st)} = \text{sign}(O_l)(|O_l| - \delta)_+ \quad (18)$$

where

$$\text{sign}(O_l) = \begin{cases} +1 & \text{if } O_l > 0 \\ 0 & \text{if } O_l = 0 \\ -1 & \text{if } O_l < 0 \end{cases}$$

and

$$(x)_+ = \begin{cases} x & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

### 3. Multivariate Wavelet Denoising Model for Portfolio Value at Risk Estimates

In the mainstream literature the Efficient Market Hypothesis (EMH) underpins most econometric models in the crude oil markets, which depicts the market structure and price behavior based on the assumption of homogeneous agents with rational expectations and single time horizon. This approximates the market well when there is relatively stronger control and much regulations in place. However, with the deregulation waves, more empirical researches suggest that the market exhibits diverse

individual characteristics not just in the frequency domain, but over multi-scale time horizons. Thus the Heterogeneous Market Hypothesis (HMH), initially proposed in the exchange rate and equity market, is proposed in this paper to complement the traditional EMH and recognize these new empirical stylized facts [72–80].

HMH can be viewed as a more restricted form of FMH, focusing on heterogeneous features of market microstructure, which comes natural when the market has fractal characteristics. It also complements the traditional EMH [74–80]. It assumed that different actors in the heterogeneous market have different time horizons and dealing frequencies. The market is heterogeneous with a fractal structure of time horizons as it consists of short-term, medium-term and long-term components. Different actors are likely to settle for different prices and decide to execute their transactions in different market situations. In other words, they create volatility. The market is also heterogeneous in terms of geographic locations of the participants. Based on these assumptions, the HMH postulates that the market consists of heterogeneous agents with heterogeneous investment strategies and investment time horizons. Compared with the homogeneous reaction to the news shocks in the EMH, HMH states that these agents or investors react to news shocks differently based on their own characteristics. On one hand, their investment time horizon and dealing frequency are diversely different. Their time scale behaviors could range from short and temporal to long and stable characteristics. For instance, pensions funds and central banks tend to have low dealing frequency focusing on long time horizon while the market traders tend to have high dealing frequency with short time horizon. On the other hand, these agents or investors employ diversely different investment strategies or measures based on their own characteristics and focus. Their frequency behaviors could range from stationary to transient characteristics.

In the literature, these micro market structures were studied incorporating information at the operational and transaction level. However, results from these studies are constrained by the restricted access to the concerned data, some insider in nature. The data obtained are also significantly lagged in time due to the regulations on protection of sensitive information. Thus it would be difficult to test the theory and implement it for real time trading environment.

Meanwhile, a different alternative is to use the time series itself without dependence on unreliable external information to extract and separate the underlying components from the original data directly. Certainly it is a difficult problem as the underlying components are unknown latent variables, whose structures are vulnerable to inappropriate separation processes. The multivariate case is made even more complicated by the time varying correlations. The wavelet based approach is promising as it is supported by positive results for decomposing and denoising of univariate time series. However, using wavelet to denoise and separate data features in multivariate time series involves further complications that deserve further investigations.

Therefore, in this paper we propose a different approach to extract and separate the underlying components from the original data directly.

Since the accuracy of the noise and denoised data separation and extraction process affects critically the generalizability of the models fitted, appropriate tools need to be selected. Traditional data denoising techniques include spectrum analysis methods such as simple averages, Kalman filtering and Fourier transforms, *etc.* However, these methods did not receive significant attentions in the mainstream empirical research literature. This is due to the insignificant performance improvement during the

application of these approaches, as they are mostly based on spectrum and frequency scale domain while ignoring the multi-scale heterogeneous structure in the data. More specifically these approaches would result in two critical issues: Firstly, the data may be distorted during the denoising process due to the single time horizon approach, and there are patterns left in the noises removed, which would reduce the forecasting accuracy if ignored. The denoising algorithm with different parameters further complicate this situation. Secondly, following HMH, the noises may vary in level in the multi-scale domain for multi-scale data structure, corresponding to different investment strategies and time horizon. To tackle these issues, the analysis and modeling of the individual components behaviors, including both the separated denoised data and noises, in the multi-scale domain need to be conducted during the denoising process, as it would reduce the risk of using biased estimates with lower goodness of fit and contributes to the overall performance improvement. For example, the behavior of transient data component is significantly different from the behavior of data component driven by normal market activities. Each of them need to be modeled separately to avoid the biased estimate due to violations of assumptions of the unsuitable models.

Thus the wavelet based denoising algorithm is proposed to serve as a promising alternative modeling tool to extract heterogeneous data features of distinct characteristics, in a real time manner without the need of insider information. This is a more refined approach compared with the normal spectrum or frequency domain approach taken by traditional data denoising techniques such as spectrum analysis methods, e.g., simple averages, Kalman filtering and Fourier transforms, *etc.*

Applying the wavelet based time frequency analysis in the higher dimensions, a two stage Multivariate Wavelet Denoising Portfolio Value at Risk Model (MWDNPVaR) is proposed to track portfolio risk evolution in the higher dimensions space and estimate portfolio VaR.

Accordingly the numerical procedure for the two stage hybrid estimation process is listed as follows.

In the first stage, the multivariate wavelet analysis based denoising algorithm is used to separate the noises matrix from the original data matrix, leaving the rest data matrix with more distinct characteristics.

Then the multivariate time series models are used to extract the denoised and noise matrix respectively, resulting in different model parameters tuned for distinct data characteristics and dependency structure.

In the second stage, since we assume that the denoised and noise data are independently distributed, the conditional means are aggregated from the individual conditional mean forecast matrix forecasted for both denoised and noise data components as in (19).

$$R_t = R_{D,t} + R_{N,t} \quad (19)$$

where  $R_t$  is the aggregated conditional mean forecast at time  $t$ ,  $R_{D,t}$  and  $R_{N,t}$  are conditional mean forecasts for both denoised and noise data respectively, at time  $t$ .

As the wavelet decorrelates the data series, suppose data series across different scales are independent. Thus the variance-covariance matrix can be aggregated from the individual variance-covariance matrix forecast at different scales as in (20).

$$\sum_t = \sum_{D,t} + \sum_{N,t} \quad (20)$$

where  $\sum_t$  is the aggregated conditional variance-covariance forecast matrix at time  $t$ ,  $\sum_{D,t}$  and  $\sum_{N,t}$  are conditional variance-covariance forecast matrix for both denoised and noise data respectively, at time  $t$ .  $\sum_t = Cov(X_t, Y_t)$  is the covariance matrix at time  $t$ ,  $Cov(X_{D,t}, Y_{N,t}) = 0$  and  $Cov(X_{N,t}, Y_{D,t}) = 0$ .

Then, suppose a unit worth portfolio investment with  $\omega$  weights and one day holding period with the forecasted conditional mean and covariance matrixes, we follow the traditional variance-covariance approach for estimating portfolio *Var* as in (21).

$$VaR(a, t) = -wR_t + \sqrt{\omega_t \sum_t \omega_t^T z_\alpha} \quad (21)$$

## 4. Empirical Studies

### 4.1. Data and Descriptive Statistics

Empirical studies are conducted in the US West Texas Intermediate (WTI) and Europe Brent markets, which are the marker markets where most trading and transactions are conducted. Together they form essential part of the crude oil portfolio. The data source is the Energy Information Administration, Department of Energy, US. The data set covers the same date range from 2 January, 2002 to 1 October, 2009 for both markets. The starting date for the data set is chosen since the equilibrium price level has shifted to a new regime since 2002. The potential cause is that the balance between the demand and supply of crude oil is increasingly constrained by its limited and scarce nature in recent years, due to the increasing demand and steady level of supply, as well as the wave of liberalization and globalization. Meanwhile, this time period would reduce the impact of previous direct systematic market disruptions such as the Gulf war, the minimum. The end date for the data set is chosen based on data availability. However, this results in different number of daily observations for each market respectively, *i.e.*, 1939 for WTI market and 1996 for Brent market, as normal trading are interrupted at different time points due to different trading regulations and holidays. For the purpose of constructing a valid portfolio in both market, only observations for the valid trading days in both markets are retained in the data set, *i.e.*, if trading in either one market is disrupted with holiday or regulations, observations for this day are excluded in both markets as no valid portfolio can be constructed. This results in 1939 observations. We assume a one dollar investment portfolio with equal holding positions, *i.e.*, 50% in the WTI market and another 50% in the Brent market. The data set is divided into three subsets, *i.e.*, the training set for determination of model parameters (40%), the model tuning set to select and determine the empirically optimal model specifications (20%) and the test set for the out-of-sample test to evaluate the performance of different models (40%). This results in 776 daily observations for out-of-sample test, yielding statistically significant results. The one day ahead forecast is performed using rolling-window method. The size of the window is 775 to cover the relevant information available. As autocorrelation and partial autocorrelation function indicates the trend factors, the daily prices are log differenced at the first order to remove them as in  $y_t = \ln \frac{p_t}{p_{t-1}}$ . We assume one day holding period to conduct one step ahead PVaR estimations.

The descriptive statistics for both markets are listed in Table 1.

**Table 1.** Descriptive statistics and statistical tests.

<b>Data</b>	$r_{WTI}$	$r_{DN,WTI}$	$r_{N,WTI}$
<b>Mean</b>	0.0011	0.0011	0
<b>SD</b>	0.0227	0.0160	0.0160
<b>Skewness</b>	−0.4177	−0.4975	0.0288
<b>Kurtosis</b>	4.8569	6.4738	3.2942
<b>JB Test</b>	0	0	0.1064
<b>BDS Test</b>	0.0037	0	0
<b>Data</b>	$r_{Brent}$	$r_{DN,Brent}$	$r_{N,Brent}$
<b>Mean</b>	0.0011	0.0011	0
<b>SD</b>	0.0219	0.0165	0.0144
<b>Skewness</b>	−0.1009	−0.4202	0.0069
<b>Kurtosis</b>	4.3219	6.2835	2.8970
<b>JB Test</b>	0	0	0.5000
<b>BDS Test</b>	0.6313	0	0

Descriptive statistics in Table 1 show some interesting stylized facts. On the aggregated level, the distribution of the original data approximates the normal distribution, as indicated by four moments. However, the rejection of the null hypothesis for Jarque–Bera (JB) test of normality suggests the deviation of return distribution from the normal distribution [81]. The rejection of the null hypothesis for BDS test suggests the existence of unknown correlations, potentially nonlinear dynamics, in the data [82]. However, compared to descriptive statistics of individually denoised data series in Table 1, the multivariate denoising algorithms extract correlated denoising data and noises data. The paired data in both markets exhibit similar descriptive statistics, suggesting that they are governed by correlated or the same underlying DGP.

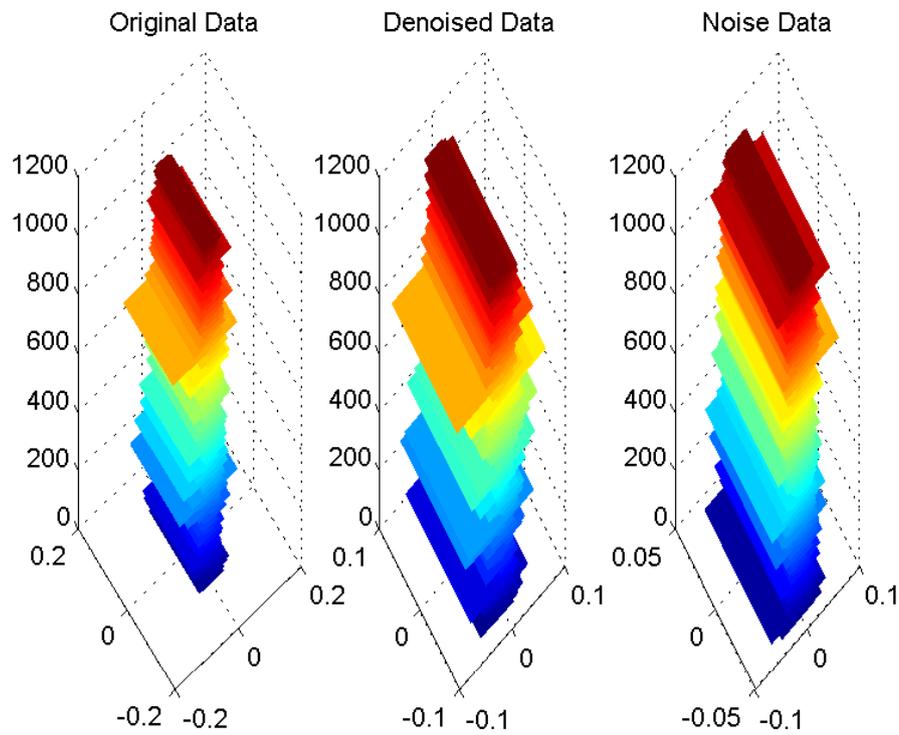
Figure 1 shows the surf plot of return series in both markets.

Results in Figure 1 show that multivariate wavelet analysis effectively separates data of distinct characteristics and shapes. The middle figure for denoised data show distinctively different shapes and behaviors than the right figure for noise data. Behaviors of both of them are different than the left figure for the original data. Interestingly, these figures show that some of the data features in the left figure are signified in the middle figure while others are signified in the right figure. For instance, the abrupt changes in the left figure are mostly captured in the middle figure.

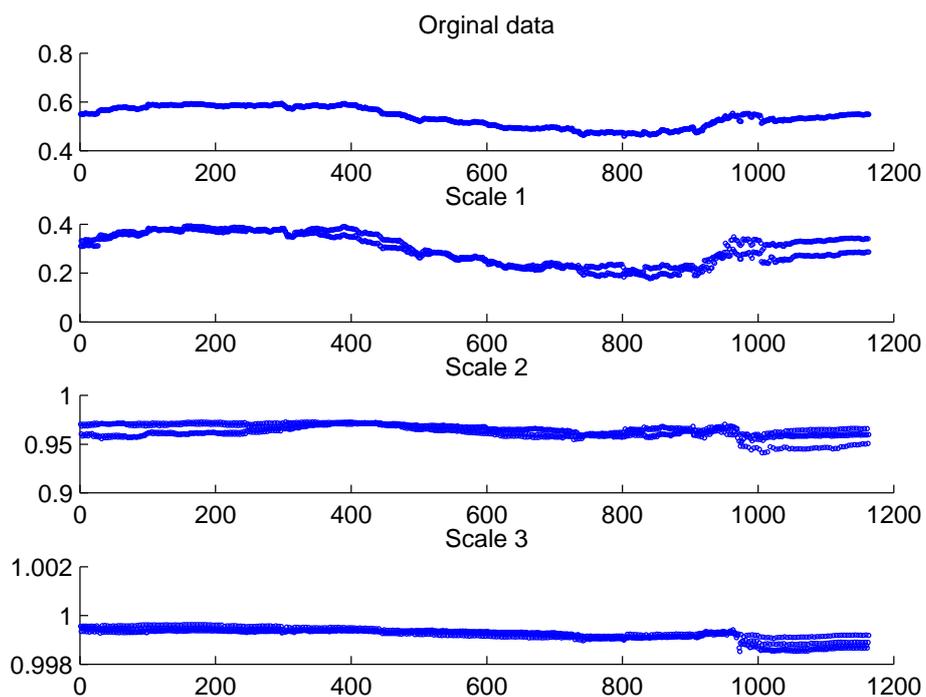
#### 4.2. Multi-scale Analysis of Correlation of Portfolio

As we argue that the correlation between different markets consists of varying strength of correlations at different scales, the cross correlation wavelet analysis is used to analyze and illustrate the multi-scale structure of correlation movements. The rolling windows over which the correlation is calculated is 775. For the portfolio data series decomposed at scale 8 with Daubechies 8 wavelet family, correlation for data series at the first three levels along with the correlation for the original portfolio are shown in the scatter plot in Figure 2.

**Figure 1.** The surf plot of the original, denoised and noise data.



**Figure 2.** Dynamic correlation between WTI and Brent markets at different scales.

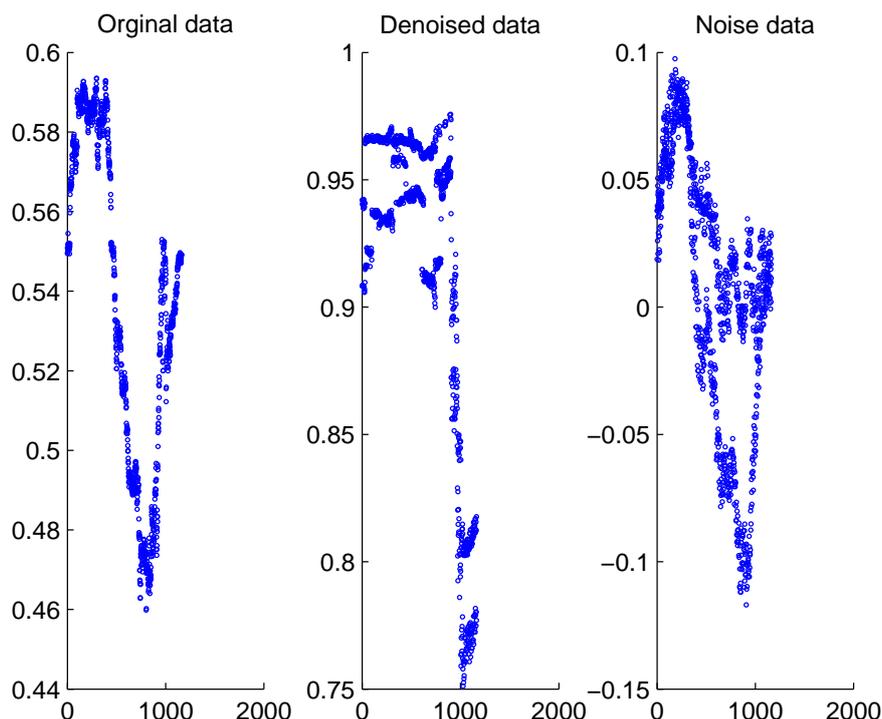


Experiment results in Figure 2 show that the correlation between WTI and Brent market is dynamically changing over the entire investment time horizon, between 0.44 and 0.6. This justifies the use of DCC-GARCH model, in which the dynamic correlation assumptions better conform to the real data.

Using wavelet analysis, the underlying structure of correlation is recovered when relating the strength of correlation to the investment time horizon. For instance, over the short term investment time horizon, the correlation is relatively low for these particular portion of markets. We also observe strong correlation between particular portion of both markets over the medium to long term time horizon.

In addition to the multi-scale correlation structure, there are also heterogeneous structure for the correlation evolution, *i.e.*, the correlation between different markets may be influenced by underlying correlations of different strength and characteristics. Therefore, we further use the cross correlation wavelet analysis to analyze and illustrate the heterogeneous correlation movements. The original, denoised and separated noise correlation between WTI and Brent market are shown in the scatter plot in Figure 3.

**Figure 3.** Dynamic correlation between WTI and Brent markets.



Experimental results in Figure 3 show that correlation for the extracted denoised data series exhibits time varying behavior at higher range between 0.75 and 0.95, while correlation for the noise data series also exhibits time varying behavior at lower range between  $-0.15$  and  $0.1$ . This shows that under the HMH framework, the multi-scale heterogeneous structure of correlations can be recovered using wavelet analysis to extract dependency structures of distinct nature, which drives the dynamic correlation behaviors. In this case, they are roughly categorized into denoised and noise parts to signify their distinct characteristics. Therefore they need to be modeled using different model specifications and parameters.

These experiments signify the multi-scale and heterogeneous correlation structure as important stylized facts that need to be incorporated during the modeling process. Wavelet analysis based denoising algorithm provides an interesting alternative to uncover the multi-scale heterogeneous correlation structure, by projecting the original data into different time scales and extracting different denoised and noise components at these scales.

#### 4.3. Performance Evaluation

The performance of the proposed MWDNPVaR model is evaluated in the empirical studies in the marker WTI and Brent market, against benchmark models. In this paper, we use two models as benchmark models, *i.e.*, the well accepted EWMA as the de facto industry standard and DCC-GARCH due to its theoretical superiority. The generalizability of the proposed model models is evaluated against benchmark models in out-of-sample forecast tests, using the Kupiec backtesting procedure for unconditional risk coverage for reliability of estimates as well as MSE and Clark–West test of predictive accuracy for the accuracy of estimates [83,84].

The frequency based approach, or the so-called unconditional coverage approach, as suggested by Kupiec, tests the statistical significance of the patterns of VaR exceedance, *i.e.*, whether the size of the loss is greater than the estimated VaR values. The null hypothesis for the Kupiec test is that the proposed model is the correct one and corresponds to the theoretical risk coverage. If the model is acceptable, the number of tail losses  $x$  should follow binomial distribution. Thus, at the given total number of observations  $n$  and confidence level  $p$ , the likelihood ratio statistics is calculated as in (22):

$$LR_{UC} = -2\ln[(1-p)^{n-x}p^x] + 2\ln[(1-\frac{x}{n})^{n-x}(\frac{x}{n})^x] \quad (22)$$

The Kupiec likelihood ratio test statistics conforms to the  $\chi^2(1)$  distribution.

The Kupiec backtesting procedure is intuitively and computationally appealing. However, it relies entirely on frequency of losses information and ignores temporal pattern of exceedances and the size of losses. Thus, it suffers from lower level of discriminatory power for small sample sizes. Therefore backtesting procedures addressing other loss exceedances patterns are developed such as the one proposed by Christoffersen to test for the temporal behavior of losses [85]. However, when the sample size is sufficiently large, the Kupiec backtesting procedures would offer an acceptable level of discriminatory power between good and bad models. Therefore, it is adopted in this paper to backtest and evaluate different models.

The MSE is the simple statistics measuring the deviation of forecasts from actual observations, as  $T^{-1} \sum_{i=1}^T \varepsilon_i^2$ , where  $\varepsilon_i = \hat{y}_i - y_i$ .

Diebold Mariano test of equal predictive accuracy is conventionally used to test the statistical significance of the differences of empirical losses between competing models and benchmark models, based on Mean Squared Prediction Error (MSPE). However, it is found that the originally proposed test statistics is upward biased heavily if models tested are nested. Thus the Clark–West test of

equal predictive accuracy is proposed to incorporate this important nested model structure in forecast evaluation problems and adjust for the DM statistics as in (23).

$$\begin{aligned}
 CW &= \frac{\bar{d}}{P\sqrt{\widehat{V}}} & (23) \\
 \bar{d} &= \frac{1}{P} \sum_{t=R+h}^T d_t^*, d_t^* = e_{1t}^2 - e_{2t}^2 + (\widehat{y}_{1t} - \widehat{y}_{2t})^2 \\
 \widehat{V} &= \frac{1}{P} \sum_{t=R+h}^T (d_t - \bar{d})^2 + \frac{2}{P} \sum_{j=1}^m \omega(j, m) \sum_{t=j+r+h}^T (d_t - \bar{d})(d_{t-j} - \bar{d})
 \end{aligned}$$

$\bar{d}$  is sample mean while  $P$  is the sample size.

The null hypothesis for the test is:  $H_0 : E(d_t^*) = 0$ , i.e., equal predictive accuracy or MSPE, for the proposed model against the benchmark models nested. The test statistics has a asymptotically normal distribution.

#### 4.4. Experiment Results

In this study, all models are implemented via the Matlab software package from the Mathworks corporation. As confirmed by numerous empirical studies, the lag order for DCC-GARCH model is set to be 1 and 1 based on the parsimony principle. Since we assume no particular distribution for data and there lacks formal methodology to test the goodness of fit of different wavelet family for the data, we test the in-sample performance of the models with different specifications with the reserved model tuning set to determine model specification. The size of in-sample model tuning set for the selection of model specification is 387. Table 2 lists the in-sample experiment results for the number of exceedances for MWDNPVaR model with different specifications including scales, wavelet basis and shrinkage algorithm. The number of exceedances for different confidence levels  $N_{cl}$ , ( $cl = 95\%, 97.5\%, 99\%$ ) are listed in the parenthesis as  $(N_{99\%}, N_{97.5\%}, N_{95\%})$ . The number of exceedances for EWMA and DCC-GARCH model are (1,4,10) and (3,12,18) respectively. Table 3 lists the  $p$  value of Kupiec backtesting procedure for exceedances observed.

**Table 2.** Exceedances for different model specifications.

Scale	DB2-H	DB2-S	SYM2-H	SYM2-S	COIF2-H	COIF2-S	DMEY-H	DMEY-S
1	(4,12,21)	(3,9,18)	(3,8,20)	(1,8,20)	(3,8,16)	(4,9,14)	(3,13,19)	(2,5,18)
2	(2,12,29)	(4,12,17)	(3,11,23)	(4,9,22)	(5,11,15)	(4,7,14)	(5,8,20)	(4,6,17)
3	(5,13,23)	(5,9,14)	(4,10,20)	(4,12,24)	(6,10,20)	(1,2,12)	(2,12,23)	(2,9,20)
4	(5,8,22)	(5,9,24)	(2,7,17)	(5,7,13)	(9,22,33)	(5,11,25)	(6,10,15)	(1,5,22)
5	(5,9,16)	(5,14,26)	(1,5,19)	(4,8,18)	(4,11,24)	(4,8,18)	(4,8,18)	(6,12,20)
6	(2,7,17)	(4,9,19)	(3,6,19)	(2,12,23)	(1,6,16)	(4,11,20)	(3,8,17)	(2,3,13)
7	(4,13,22)	(3,7,15)	(2,7,20)	(2,4,16)	(0,4,9)	(5,7,16)	(4,11,23)	(5,11,19)
8	(3,9,14)	(4,10,24)	(4,10,18)	(3,9,14)	(4,6,19)	(3,8,19)	(3,7,24)	(5,10,19)
9	(5,10,17)	(3,11,20)	(2,6,19)	(5,10,21)	(5,8,19)	(4,13,22)	(1,6,14)	(4,9,27)

**Table 3.** *P* value of Kupiec backtesting procedure for the range of exceedance values.

$N_{99\%}$	$P_{99\%}$	$N_{97.5\%}$	$P_{97.5\%}$	$N_{97.5\%}$	$P_{97.5\%}$	$N_{95\%}$	$P_{95\%}$	$N_{95\%}$	$P_{95\%}$	$N_{95\%}$	$P_{95\%}$
0	0.0053	4	0.0366	11	0.6728	9	0.0072	17	0.5760	23	0.4078
1	0.0805	6	0.1987	12	0.4653	10	0.0166	18	0.7501	24	0.2951
2	0.2924	7	0.3600	13	0.3032	13	0.1162	19	0.9348	25	0.2064
3	0.6436	8	0.5742	14	0.1864	14	0.1902	20	0.8801	26	0.1396
4	0.9473	9	0.8240			15	0.2916	21	0.7041	27	0.0914
5	0.5807	10	0.9162			16	0.4211	22	0.5449	29	0.0355

In Table 3  $N_{cl}$ , ( $cl = 95\%, 97.5\%, 99\%$ ) refers to the number of exceedances at different confidence levels,  $P_{cl}$ , ( $cl = 95\%, 97.5\%, 99\%$ ) refers to  $p$  value of Kupiec backtesting procedure at different confidence levels. Experiment results from Table 2 show the different in-sample performance for the proposed MWDNPVaR, using different model parameters. The three numbers in the parentheses correspond to exceedances at three different confidence levels, *i.e.*, 99%, 97.5% and 95%. The attempted model parameters are as follows: 4 different wavelet families including Daubechies 2 (DB2), Symlet 2 (SYM2), Coiflet 2 (COIF2) and DMEY 2. Hard and soft thresholding algorithm. Decomposition levels from 1 to 9. The  $p$  value of backtesting procedure is (0.0805, 0.0366, 0.0166) for the EWMA model and (0.6436, 0.4653, 0.7501) for the MGARCH model. Using exceedances and their associated Kupiec backtesting statistics as performance measure as in Tables 2 and 3, we determine the optimal model specification demonstrating the highest level of reliability, as indicated by the higher  $p$  value for the out-of-sample test, *i.e.*, Symlet 2 using hard thresholding algorithm at decomposition level 3. The shrinkage algorithm is minimax threshold. The thresholding is conducted in a level dependent manner. The  $p$  value of backtesting procedure for the chosen model is (0.9473, 0.9162, 0.8801).

The reliability performance measure of the proposed algorithm against alternative benchmark models are listed in Table 4.

**Table 4.** Reliability comparison of different models.

Models	CL	Exceedance	KT <sub><i>pvalue</i></sub>
EWMA	99%	32	0
	97.5%	52	0
	95%	67	0
DCC-GARCH	99%	8	0.9313
	97.5%	29	0.0395
	95%	52	0.0383
MWDNPVaR	99%	7	0.7804
	97.5%	20	0.8908
	95%	39	0.9737

Experiment results in Table 4 show that the proposed MWDNPVaR algorithm improves the reliability significantly compared with the traditional EWMA and DCC-GARCH model. The exceedances of the

MWDNPVaR algorithm decreases substantially across all confidence levels. This result is statistically significant as confirmed by the significant  $p$  value. We also observed significantly improved  $p$  value at both 97.5% and 95% confidence levels from the Kupiec test statistics.

The performance improvement is attributed to the introduction of multivariate wavelet denoising algorithm as the data separation tool. More specifically, this is mainly attributed to the use of multivariate wavelet denoising algorithm to separate data and noises, taking into account their distinct behavior patterns. It is also attributed to the individual modeling of the separated components in the multi-scale time domain, taking into account the heterogeneous data structure.

We further conduct experiments to investigate the predictive accuracy of the proposed algorithm against benchmark alternatives. Results from experiments are listed in Table 5.

**Table 5.** Predictive accuracy comparison of different models.

		MSE	CW <sub><math>p</math>-value,EWMA</sub>	CW <sub><math>p</math>-value,DCC-GARCH</sub>
<b>EWMA</b>	99%	0.0028	N/A	1
	97.5%	0.0022	N/A	1
	95%	0.0018	N/A	1
<b>DCC-GARCH</b>	99%	1.1391	1	N/A
	97.5%	0.8594	1	N/A
	95%	0.6561	1	N/A
<b>MWDNPVaR</b>	99%	18.7991	1	1
	97.5%	14.1771	1	1
	95%	10.8310	1	1

Experiment results in Table 5 show that the proposed MWDNPVaR achieves inferior predictive accuracy than the benchmark EWMA and DCC-GARCH based models, as measured by larger MSE. However, the performance gap is not statistically significant, as shown by the failure to reject the null hypothesis for Clark–West test of predictive accuracy. This implies that the predictive accuracy of the proposed MWDNPVaR model is at the comparable level as the traditional MGARCH based PVaR models in asset allocations and utilization. The main purpose of the downside risk estimate as those presented in this manuscript is to improve the reliability to allow sufficient capital allocation to cover the potential significant capital loss due to adverse market price movement. The proposed algorithm in this manuscript can help to estimate risks more reliably under the turbulent period with the prevalence of transient or extreme events, at the cost of predictive accuracy, if investors are risk averse.

There are some important implications drawn from the empirical investigations in the paper. Firstly, although previous researches show that there are increasingly diversified data features, such as extreme events, that can not be explained by MGARCH models based on normal distribution, our results show that it could be attributed to the heterogeneous structure of the underlying data components, potentially much simpler than previously assumed extreme distribution. Thus based on the parsimony principle, it is inappropriate to disregard MGARCH based models in favor of others, such as the models based on extreme value theories. By deliberately separating and modeling the data and noise, the performance of DCC-GARCH model can be improved further to the competent level. Secondly, the optimal performance

is achieved when the individual components are extracted using Symlet 2 wavelet family and modeled with different parameters. These observations support the argument that the risk in the crude oil prices are complicated processes with a mixture of underlying DGPs. Its modeling and analysis is a tricky issue which requires the detailed look into the underlying structure in the time scale domain. Thus appropriate recovery of the underlying structure is critical to the further performance improvement and more thorough understanding of the DGPs during the modeling process.

Results and conclusions are generalizable and are supposed to be valid across different data sets and time windows, *i.e.*, the proposed algorithm is expected to lead to improved reliability since the disruptive noise parts are separated in the multiscale domain.

## 5. Conclusions

In this paper, we introduce the multivariate wavelet analysis to analyze the multi-scale behaviors of the correlation among different markets. Then we propose a novel multivariate wavelet denoising based approach for estimating PVaR. The superior performance of the proposed algorithm is supported by performance evaluation in the empirical studies in the crude oil markets. This is attributed to the separation of data and noises taking into account their multi-scale dependency structure and the construction of the hybrid methodology incorporating these information.

The contributions of the proposed methodology is threefold. Firstly, the multivariate wavelet analysis is introduced to analyze data and noises components in the portfolio, especially their distinct behavioral and correlation patterns, recognizing the heterogeneous investment strategies over different time horizons and across different markets. Secondly, the risk for the separated data and noises components of the portfolio are modeled using DCC-GARCH model with different parameters to emphasize their different behavioral and correlation characteristics. The incorporation of heterogeneous investment strategies for portfolio across different markets in the time scale domain leads to improved PVaR estimates with higher reliability. Thirdly, the proposed multivariate wavelet denoising GARCH methodology is not sensitive to the underlying distributional assumptions and thus is more robust and effective. Results from this research may be limited in that the level of accuracy for the boundary or the edge set by the denoising algorithm could critically affect the performance. These can be possible directions for future research.

## Acknowledgement

The authors would like to thank the editors and two anonymous referees for their valuable comments and suggestions, which help improve the quality of the paper tremendously. This work is partially supported by grants from the Major Program of the National Social Science Foundation of China (No. 09&ZD041), the National Social Science Foundation of China (No. 11BJY108), the Humanities and Social Science Foundation of Ministry of Education of China (No. 10YJA790064, 12YJAZH039), the Major Program of Humanities and Social Science Foundation of Ministry of Education of China (No. 11JZD018).

## References

1. Yang, C.; Hwang, M.; Huang, B. An Analysis of Factors Affecting Price Volatility of the US Oil Market. *Energy Econ.* **2002**, *24*, 107–119.
2. Plourde, A.; Watkins, G. Crude Oil Prices between 1985 and 1994: How Volatile in Relation to Other Commodities? *Resour. Energy Econ.* **1998**, *20*, 245–262.
3. Ramsey, J.B. The Contribution of Wavelets to the Analysis of Economic and Financial Data. *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **1999**, *357*, 2593–2606.
4. Ramsey, J. Wavelets in Economics and Finance: Past and Future. *Stud. Nonlinear Dyn. Econom.* **2002**, *6*, doi: 10.2202/1558-3708.1090.
5. Struzik, Z.R. Wavelet Methods in (financial) Time-Series Processing. *Phys. A* **2001**, *296*, 307–319.
6. Kim, S.; In, F. The Relationship between Stock Returns and Inflation: New Evidence from Wavelet Analysis. *J. Empir. Financ.* **2005**, *12*, 435–444.
7. In, F.; Kim, S. A Note on the Relationship between Fama-French Risk Factors and Innovations of Icapm State Variables. *Financ. Res. Lett.* **2007**, *4*, 165–171.
8. Gallegati, M. Wavelet Analysis of Stock Returns and Aggregate Economic Activity. *Comput. Stat. Data Anal.* **2008**, *52*, 3061–3074.
9. Mitra, S.; Nandi, B.; Mitra, A. Study of Dynamic Relationships Between Financial and Real Sectors of Economies with Wavelets. *Appl. Math. Comput.* **2007**, *188*, 83–95.
10. Sharkasi, A.; Crane, M.; Ruskin, H.J.; Matos, J.A. The Reaction of Stock Markets to Crashes and Events: A Comparison Study between Emerging and Mature Markets Using Wavelet Transforms. *Phys. A: Stat. Mech. Appl.* **2006**, *368*, 511–521.
11. Meng, K.; Dong, Z.Y.; Wong, K.P. Self-Adaptive Radial Basis Function Neural Network for Short-Term Electricity Price Forecasting. *IET Gener. Transm. Distrib.* **2009**, *3*, 325–335.
12. Amjady, N.; Keynia, F. Day Ahead Price Forecasting of Electricity Markets by a Mixed Data Model and Hybrid Forecast Method. *Int. J. Electr. Power Energy Syst.* **2008**, *30*, 533–546.
13. Aggarwal, S.K.; Saini, L.M.; Kumar, A. Electricity Price Forecasting in Ontario Electricity Market Using Wavelet Transform in Artificial Neural Network Based Model. *Int. J. Control Automat. Syst.* **2008**, *6*, 639–650.
14. Soltani, S. On the Use of the Wavelet Decomposition for Time Series Prediction. *Neurocomputing* **2002**, *48*, 267–277.
15. Yousefi, S.; Weinreich, I.; Reinartz, D. Wavelet-Based Prediction of Oil Prices. *Chaos Solitons Fractals* **2005**, *25*, 265–275.
16. Xu, Z.; Dong, Z.; Liu, W. Short-Term Electricity Price Forecasting Using Wavelet and Svm Techniques. In *Dynamics of Continuous Discrete and Impulsive Systems, Series B, Applications Algorithms*, Proceedings of the Third International DCDIS Conference on Engineering Applications and Computational Algorithms, Guelph, Canada, 15–18 May 2003; pp. 372–377.
17. Conejo, A.; Plazas, M.; Espinola, R.; Molina, A. Day-Ahead Electricity Price Forecasting Using the Wavelet Transform and Arima Models. *IEEE Trans. Power Syst.* **2005**, *20*, 1035–1042.

18. de Souza e Silva, E.G.; Legey, L.F.; de Souza e Silva, E.A. Forecasting oil price trends using wavelets and hidden Markov models. *Energy Econ.* **2010**, *32*, 1507–1519.
19. Naccache, T. Oil price cycles and wavelets. *Energy Econ.* **2011**, *33*, 338–352.
20. He, K.; Lai, K.K.; Yen, J. Value-at-risk estimation of crude oil price using MCA based transient risk modeling approach. *Energy Econ.* **2011**, *33*, 903–911.
21. Fernandez, V. The Capm and Value at Risk at Different Time-Scales. *Int. Rev. Financ. Anal.* **2006**, *15*, 203–219.
22. Fernandez, V.P. The International Capm and a Wavelet-Based Decomposition of Value at Risk. *Stud. Nonlinear Dyn. Econom.* **2005**, *9*, Article 4.
23. Gencay, R.; Gradojevic, N.; Selcuk, F.; Whitcher, B. Asymmetry of Information Flow Between Volatilities Across Time Scales. *Quant. Financ.* **2010**, *10*, 895–915.
24. Fernandez, V.; Lucey, B.M. Portfolio Management Under Sudden Changes in Volatility and Heterogeneous Investment Horizons. *Phys. A: Stat. Mech. Appl.* **2007**, *375*, 612–624.
25. Manchanda, P.; Kumar, J.; Siddiqi, A. Mathematical Methods for Modelling Price Fluctuations of Financial Times Series. *J. Franklin Inst.* **2007**, *344*, 613–636.
26. DiSario, R.; Saraoglu, H.; McCarthy, J.; Li, H. Long Memory in the Volatility of an Emerging Equity Market: The Case of Turkey. *J. Int. Financ. Mark. Inst. Money* **2008**, *18*, 305–312.
27. Gencay, R.; Selck, F.; Whitcher, B. Multiscale Systematic Risk. *J. Int. Money Financ.* **2005**, *24*, 55–70.
28. Gencay, R.; Selcuk, F.; Ulugulyagci, A. High Volatility, Thick Tails and Extreme Value Theory in Value-at-Risk Estimation. *Insur. Math. Econ.* **2003**, *33*, 337–356.
29. Capobianco, E. Empirical Volatility Analysis: Feature Detection and Signal Extraction with Function Dictionaries. *Phys. A: Stat. Mech. Appl.* **2003**, *319*, 495–518.
30. Kyaw, N.A.; Los, C.A.; Zong, S. Persistence Characteristics of Latin American Financial Markets. *J. Multinat. Financ. Manag.* **2006**, *16*, 269–290.
31. Karandikar, R.G.; Deshpande, N.R.; Khaparde, S.A.; Kulkarni, S.V. Modelling Volatility Clustering in Electricity Price Return Series for Forecasting Value at Risk. *Eur. Trans. Electr. Power* **2009**, *19*, 15–38.
32. He, K.; Xie, C.; Chen, S.; Lai, K.K. Estimating Var in crude oil markets: A Novel Multi-Scale Non-Linear Ensemble Approach Incorporating Wavelet Analysis and Neural Network. *Neurocomputing* **2009**, *72*, 3428–3438.
33. Chang, C.L.; McAleer, M.; Tansuchat, R. Analyzing and forecasting volatility spillovers, asymmetries and hedging in major oil markets. *Energy Econ.* **2010**, *32*, 1445–1455.
34. Lanza, A.; Manera, M.; McAleer, M. Modeling Dynamic Conditional Correlations in WTI Oil Forward and Futures Returns. *Financ. Res. Lett.* **2006**, *3*, 114–132.
35. Vacha, L.; Barunik, J. Co-movement of energy commodities revisited: Evidence from wavelet coherence analysis. *Energy Econ.* **2012**, *34*, 241–247.
36. Cifter, A. Value-at-risk estimation with wavelet-based extreme value theory: Evidence from emerging markets. *Phys. A: Stat. Mech. Appl.* **2011**, *390*, 2356–2367.
37. Yoshida, Y. An Estimation Model of Value-at-Risk Portfolio Under Uncertainty. *Fuzzy Sets Syst.* **2009**, *160*, 3250–3262.

38. Miller, D.J.; Liu, W.H. Improved Estimation of Portfolio Value-At-Risk Under Copula Models with Mixed Marginals. *J. Futures Mark.* **2006**, *26*, 997–1018.
39. Dowd, K. *Measuring Market Risk*; John Wiley & Sons Inc.: Chichester, UK and Hoboken, NJ, USA, 2005.
40. Bauwens, L.; Laurent, S.; Rombouts, J. Multivariate GARCH models: A survey. *J. Appl. Econom.* **2006**, *21*, 79–109.
41. Engel, R. *Anticipating Correlations: A New Paradigm for Risk Management*; The Econometric and Tinbergen Institutes Lectures, Princeton University Press: New Jersey, NJ, USA, 2009.
42. Engle, R. Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models. *J. Bus. Econ. Stat.* **2002**, *20*, 339–350.
43. Bollerslev, T. Modelling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized Arch Model. *Rev. Econ. Stat.* **1990**, *72*, 498–505.
44. Engle, R.; Kroner, K. Multivariate Simultaneous Generalized Arch. *Econom. Theory* **1995**, *11*, 122–150.
45. Tse, Y.; Tsui, A. A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations. *J. Bus. Econ. Stat.* **2002**, *20*, 351–362.
46. Rombouts, J.V.K.; Verbeek, M. Evaluating Portfolio Value-at-Risk Using Semi-Parametric Garch Models. *Quant. Financ.* **2009**, *9*, 737–745.
47. Ku, Y.H.H.; Wang, J.J. Estimating Portfolio Value-at-Risk Via Dynamic Conditional Correlation Mgarch Model—An Empirical Study on Foreign Exchange Rates. *Appl. Econ. Lett.* **2008**, *15*, 533–538.
48. Huang, J.T.; Lee, K.J.; Liang, H.; Lin, W.F. Estimating Value at Risk of Portfolio by Conditional Copula–Garch Method (Vol 43, Pg 315, 2009). *Insur. Math. Econ.* **2010**, *46*, 436.
49. Glasserman, P.; Heidelberger, P.; Shahabuddin, P. Portfolio Value-at-Risk with Heavy-Tailed Risk Factors. *Math. Financ.* **2002**, *12*, 239–269.
50. Zmeskal, Z. Value at Risk Methodology of International Index Portfolio Under Soft Conditions (fuzzy-Stochastic Approach). *Int. Rev. Financ. Anal.* **2005**, *14*, 263–275.
51. Lin, P.C.; Ko, P.C. Portfolio Value-At-Risk Forecasting with Ga-Based Extreme Value Theory. *Expert Syst. Appl.* **2009**, *36*, 2503–2512.
52. Kozaki, M.; Sato, A.H. Application of the Beck Model to Stock Markets: Value-at-Risk and Portfolio Risk Assessment. *Phys. A: Stat. Mech. Appl.* **2008**, *387*, 1225–1246.
53. Katkovnik, V.; Foi, A.; Egiazarian, K.; Astola, J. From Local Kernel to Nonlocal Multiple-Model Image Denoising. *Int. J. Comput. Vis.* **2010**, *86*, 1–32.
54. Palhano Xavier de Fontes, F.; Andrade Barroso, G.; Coup, P.; Hellier, P. Real Time Ultrasound Image Denoising. *J. Real-Time Image Process.* **2011**, *6*, 15–22.
55. Motwani, M.; Gadiya, M.; Motwani, R.; Harris, F.J. Survey of Image Denoising Techniques. *In Proceedings of Global Signal Processing Expo Conference (GSPx)*, Santa Clara, CA, USA, 2004; pp. 27–30.
56. Buades, A.; Coll, B.; Morel, J. A Review of Image Denoising Algorithms, with a New One. *Multiscale Model. Simul.* **2005**, *4*, 490–530.

57. Mallat, S.; Peyre, G. A Review of Bandlet Methods for Geometrical Image Representation. *Numer. Algorithms* **2007**, *44*, 205–234.
58. Kim, S.; Kang, W.; Lee, E.; Paik, J. Wavelet-Domain Color Image Enhancement Using Filtered Directional Bases and Frequency-Adaptive Shrinkage. *IEEE Trans. Consum. Electron.* **2010**, *56*, 1063–1070.
59. Pang, Z.F.; Yang, Y.F. A Two-Step Model for Image Denoising Using a Duality Strategy and Surface Fitting. *J. Computat. Appl. Math.* **2010**, *235*, 82–90.
60. Chen, Q.; Sun, Q.S.; Xia, D.S. Homogeneity Similarity Based Image Denoising. *Pattern Recognit.* **2010**, *43*, 4089–4100.
61. Gorgel, P.; Sertbas, A.; Ucan, O.N. A Wavelet-Based Mammographic Image Denoising and Enhancement with Homomorphic Filtering. *J. Med. Syst.* **2010**, *34*, 993–1002.
62. Kumar, P.; Agnihotri, D. Biosignal Denoising Via Wavelet Thresholds. *IETE J. Res.* **2010**, *56*, 132–138.
63. Rabbani, H.; Gazor, S. Image Denoising Employing Local Mixture Models in Sparse Domains. *IET Image Process.* **2010**, *4*, 413–428.
64. Saeedi, J.; Moradi, M.H.; Faez, K. A New Wavelet-Based Fuzzy Single and Multi-Channel Image Denoising. *Image Vis. Comput.* **2010**, *28*, 1611–1623.
65. Tian, J.; Chen, L.; Ma, L. A Wavelet-Domain Non-Parametric Statistical Approach for Image Denoising. *IEICE Electron. Exp.* **2010**, *7*, 1409–1415.
66. Zifan, A.; Moradi, M.H.; Gharibzadeh, S. Microarray Image Enhancement by Denoising Using Decimated and Undecimated Multiwavelet Transforms. *Signal Image Video Process.* **2010**, *4*, 177–185.
67. Daubechies, I. *Ten Lectures on Wavelets*; Society for Industrial and Applied Mathematics: Philadelphia, PA, USA, 1992.
68. Percival, D.B.; Walden, A.T. *Wavelet Methods for Time Series Analysis*; Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press: Cambridge, UK, 2000.
69. Vidakovic, B. *Statistical Modeling by Wavelets*; Wiley Series in Probability and Mathematical Statistics Applied Probability and Statistics Section, Wiley: New York, NY, USA, 1999.
70. Donoho, D.L.; Johnstone, I.M. Adapting to Unknown Smoothness via Wavelet Shrinkage. *J. Am. Stat. Assoc.* **1995**, *90*, 1200–1224.
71. Donoho, D.; Johnstone, I. Ideal Denoising in an Orthonormal Basis Chosen from a Library of Bases. *C. R. Acad. Sci. Paris A* **1994**, *319*, 1317–1322.
72. Muller, U.A.; Dacorogna, M.M.; Olsen, R.D.D.R.B.; Pictet, O.V.; von Weizsacker, J.E. Volatilities of Different Time Resolutions—Analyzing the Dynamics of Market Components. *J. Empir. Financ.* **1997**, *4*, 213–239.
73. Lux, T.; Marchesi, M. Scaling and Criticality in a Stochastic Multi-Agent Model of a Financial Market. *Nature* **1999**, *397*, 498–500.
74. Brock, W.A.; Hommes, C.H. A Rational Route to Randomness. *Econometrica* **1997**, *65*, 1059–1095.
75. Brock, W.A.; Kleidon, A.W. Periodic Market Closure and Trading Volume: A Model of Intraday Bids and Asks. *J. Econ. Dyn. Control* **1992**, *16*, 451–489.

76. Brock, W.A.; LeBaron, B.D. A Dynamic Structural Model for Stock Return Volatility and Trading Volume. *Rev. Econ. Stat.* **1996**, *78*, 94–110.
77. Hommes, C.H. Financial Markets as Nonlinear Adaptive Evolutionary Systems. *Quant. Financ.* **2001**, *1*, 149–167.
78. Dacorogna, M.M.; Gencay, R.; Muller, U.; Olsen, R.B.; Pictet, O.V. *An Introduction to High-Frequency Finance*; Academic Press: San Diego, CA, USA, 2001.
79. Goeree J.K., H.C. Heterogeneous Beliefs and the Non-Linear Cobweb Model. *J. Econ. Dyn. Control* **2000**, *24*, 761–798.
80. Farmer, J. Market Force, Ecology and Evolution. *Ind. Corp. Change* **2002**, *11*, 895–953.
81. Jarque, C.M.; Bera, A.K. A Test for Normality of Observations and Regression Residuals. *Int. Stat. Rev.* **1987**, *55*, 163–172.
82. Brock, W.A.; Scheinkman, J.A.; Dechert, W.D.; LeBaron, B. A Test for Independence Based on the Correlation Dimension. *Econom. Rev.* **1996**, *15*, 197–235.
83. Clark, T.E.; West, K.D. Using Out-Of-Sample Mean Squared Prediction Errors to Test the Martingale Difference Hypothesis. *J. Econom.* **2006**, *135*, 155–186.
84. Clark, T.E.; West, K.D. Approximately Normal Tests for Equal Predictive Accuracy in Nested Models. *J. Econom.* **2007**, *138*, 291–311.
85. Christoffersen, P.F. Evaluating Interval Forecasts. *Int. Econ. Rev.* **1998**, *39*, 841–862.

© 2012 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).