



Research article

Forecasting hourly WTI oil front monthly price volatility densities

Per B. Solibakke

Faculty of Economics, Norwegian University of Science and Technology, Trondheim, Norway

* **Correspondence:** Email: per.b.solibakke@ntnu.no.

Abstract: Financial commodity markets have an impact on company values and cash flows, where price movements within frequent time intervals can be both significant and random. Understanding highly frequent price movements is both important and difficult. In this paper, I measured and forecasted volatility for high-frequency (mostly twelve hours per day) WTI Oil front month price movements from 2012 to 2024 (about 40,500 observations). I created a new stochastic volatility model for extracting latent volatility using the non-linear Kalman filter. Stable and strictly ergodic hourly price series, along with the BIC optimal non-linear (generic) general method of moments model coefficients, enabled this process. The latent volatility seemed to separate into two volatility factors. One factor that was very persistent suggested slow mean reversion, and one choppy strongly mean-reverting factor. The data dependence found in the factor volatility series suggested forecasting ability. I applied classical static forecasts and three machine learning regression techniques to predict one-step-ahead volatility. The two volatility factors and one summarizing exponential factor were reported, along with several fit measures, including the root mean square error and Theil's measure of covariance.

Keywords: data dependency; high frequency; stochastic volatility; international WTI oil prices

JEL Codes: C61, Q4

1. Introduction

Volatility is a crucial concept in financial markets, and its importance is widely recognized in both academic research and practical finance. Key modern financial theory and practice reasons are risk measurement and management, option pricing, portfolio construction (diversification), and a general

understanding of market dynamics (microstructure). Multifactor volatility models are simple to simulate and estimate, and they appear to perform better in derivative calculations due to their smooth sample paths, compatibility with standard hedging arguments, and compatibility with Itô-calculus. Multifactor models, in contrast to jump diffusion models, are more appealing for implementing machine learning (ML) and artificial intelligence (AI) for volatility predictions. The applications of multifactor models are numerous, for example, options pricing, risk assessment, algorithmic trading, portfolio optimization, and general market surveillance. So, this study creates and tests multifactor scientific stochastic volatility (SV) models from high-frequency WTI oil front month (WTI oil) price changes to predict the latent volatility of the liquid fossil fuel market. Volatility is a measure of how widely apart an asset's price swings are from one another. All volatility models must have the ability to anticipate future price changes to be successful. Internationally, volatility models have successfully predicted the absolute size of movements, quantiles, and full densities. Asset volatility is unique in that it is latent or not immediately visible, making the evaluation of volatility models' predicting abilities challenging. I estimate the most liquid hourly WTI oil price movements (continuous) using a general method of moment (GMM) estimation. The SV models give access to conditional moments and potential forecasts (data dependence). The analysis is univariate (non-synchronous trading is irrelevant) and has mainly three objectives. First, an objective to find a general step ahead in WTI oil densities; second, an objective to identify data dependence for potential predictability; and third, an objective to report systematic market features and potential market features applying machine learning techniques.

Gallant and Tauchen (1987, 1992) introduced semi-nonparametric time series analysis (SNP densities)¹ as the methodology. Unlike traditional estimating methods, nonparametric estimation does not depend on selecting data from a known distribution. Nonparametric models, on the other hand, use the underlying data to establish the model structure. The conditional density of the time series processes (non-normality) is approximated by the method using an expansion of hermite functions. For time series data the hermite expansion is attractive both for modelling and computation. For modelling, the Gaussian components of the hermite expansion make it easy to absorb familiar time series models. For computation, the hermite density is easy to evaluate and differentiate. Furthermore, the SNP implementation is easy to sample from (simulations). The SNP expansion process is therefore well approximated using a well-known parametric model as the leading term in the model expansion process; higher-order terms, such as hermite functions, show how the process deviates from the model (Robinson, 1983). To find the proper order of expansion, I fit the SNP model using traditional maximum likelihood and a model selection approach (BIC) (Schwarz, 1978). To compute the nonlinear functions of densities², the model is well-designed.

The remainder of this paper is organized as follows: In Section 2, I discuss the WTI oil front month contracts and look at some literature on the semiparametric (SNP) time-series model, stochastic volatility, and the use of the efficient method of moments (EMM). For the volatility forecasting classical forecasting (OLS) and machine learning (ML) and the lasso, ridge and Decision Forest regression models are presented. Finally, in section 2, neural networks are described briefly for prediction purposes. In Section 3.1, the semi-nonparametric (SNP) establish a maximum likelihood

¹ SNP/EMM: A Program for Efficient Methods of Moments Estimation, Duke University, 09.08.2022 (<http://econ.duke.edu/webfiles/arg/emm>).

² The computer cluster at NTNU, Faculty of Economics and Management, Trondheim is used for estimation/implementation. A special thanks to Professor Asgeir Thomasgaard at NTNU, for access to the computer cluster.

consistent mean and volatility equation details from an optimal BIC criterion (1978). The model approximation for the conditional density, which summarizes the probability distribution and describes the price movement processes, is expanded by the hermite function expansions (non-normality). The residual characteristics report t-statistics for all moment well below two indication a well approximate model fit. In Section 3.2, I look at the effective method of moments (EMM) for estimating SV specifications and use the non-linear Kalman filter to functionally calibrate the volatility vector that hasn't been observed. The methodology evaluates one or two volatility factors for WTI oil. Potential data dependence evaluates the predictability of the volatility factors. In section 4, I study predictions using methodologies from section 2.4. In section 5, I summarize and conclude the study. Appendix A shows similar results for the daily WTI oil data series.

2 Literature review

2.1. WTI oil front month prices

The fossil oil literature shows a link between the oil and stock markets, which appears not only in return but also in volatility. Clark (1973), Tauchen and Pitts (1983), and Ross (1989) observe that the rate of information flow to a market correlates with the volatility of an asset, not its return. Thus, volatility is a good measure of information flow among markets. Exploring information flow may generate new insights. For instance, the study by Vo (2011) shows a bidirectional dependence in volatility between stock and oil markets. That is, shocks to either market help predict not only volatility in their own market but also that in the other market.

I attempt to extract information from WTI oil prices and their movements. It will not focus on the relationship between stock and oil returns, which many papers cited above have studied, but aim more at volatility models that can extract useful information and information flow with good forecasting power. Modelling and forecasting volatility are very important for at least two reasons. First, volatility is an important variable for pricing derivatives, whose trading volume has quadrupled in recent years. Furthermore, volatility is an important input in risk management. For instance, I use it to construct optimal hedge ratios to mitigate risk and estimate the value at risk, to name just two applications. Second, in order to make efficient econometric inferences about a variable's mean, I need a correct specification of its volatility.

2.2. The Semi-Non-Parametric (SNP) model

Our work will suggest conditional models using non-linear stochastic models. To describe the structure of the conditional volatility, I use the term generalized autoregressive conditional heteroscedasticity (GARCH), and to describe the structure of the conditional mean, I use the word autoregressive and moving average (ARMA). In contrast to the ARCH specifications, which were initially explored by Engle (1982) and advanced by Bollerslev (1986), who described the generalized ARCH or GARCH, ARMA models can be investigated in detail in, for instance, Mills (1990). Initially, the number of delays in the ARCH specification was the reason for the development of GARCH from ARCH. The volatility is defined by ARCH/GARCH as a function of historical price changes and volatility. Quite a few studies have demonstrated how the findings from this work have been utilized in the literature on international finance. See, for instance, Nelson (1991), Bollerslev et al. (1987, 1992),

Engle et al. (1986, 1993), and deLima (1995a, 1995b). Gouriéroux (1997) provides a thorough introduction to ARCH models and their uses in finance. While Glosten et al. (1993) described the truncated GARCH (GJR), Ding et al. (1993) extended the symmetric GARCH model into an asymmetric GARCH.

Gallant et al. (1992, 2010) use the term semi-nonparametric, or SNP³, to refer to a method that lies between parametric and nonparametric approaches. Higher-order terms in the series expansion are deviations from the existing parametric model, which is the leading term and is known to provide a decent approximation of the process. The method's theoretical underpinning is the hermite series expansion, which is particularly appealing for time series data due to modelling and computational considerations. In terms of modelling, the Hermite expansion's Gaussian component makes it simple to incorporate into popular time series models, including VAR, ARCH, and GARCH models (Engle, 1982; Bollerslev, 1986). These models typically provide excellent initial approximations for a wide range of applications. Hermite density is simple to assess and distinguish computationally. Furthermore, because its moments are higher than those of the normal, which can be calculated using conventional recursions, they are simple to assess. Finally, sampling from a hermite density is doable, which makes simulation easier.

2.3. Stochastic volatility⁴

Instead of directly defining the predicted distribution of price returns, the stochastic volatility (SV) method does so indirectly through the model's structure. The suggested one-step-ahead distribution of returns recorded over any arbitrary time interval appropriate for the econometrician is not a concern for the SV model because it has its own stochastic process. I begin with the work of Andersen et al. (1994, 2002), which examines the well-known stochastic volatility diffusion for a given stock price S_t , as indicated by Equation (1).

$$\frac{dS_t}{S_t} = \left(\mu + c(V_{1,t} + V_{2,t}) \right) dt + \sqrt{V_{1,t}} dW_{1,t} + \sqrt{V_{2,t}} dW_{2,t}. \quad (1)$$

There are two types of unobserved volatility processes: log linear and square root (affine), for $W_{1,t}$, $i = 1, 2$. Standard Brownian motions $W_{1,t}$, and $W_{2,t}$ may be correlated by $\text{corr}(dW_{1,t}, dW_{2,t}) = \rho$. The μ is mean drift and c is a parameter of volatility-in-mean. The stochastic volatility model was estimated by Andersen et al. (1994, 2002) using daily S&P 500 stock index data spanning from 1953 to December 31, 1996. They sharply reject both SV model versions. A basic SV model, on the other hand, benefits a lot from having a jump component added because it captures two well-known features: Fat non-Gaussian tails and persistent time-varying volatility. Chernov et al.'s (2003) results for an SV model with two stochastic volatility variables are positive. An affine setup and a logarithmic setup are two major categories of setups that the authors take into consideration for the volatility index functions and factor dynamics. The authors estimate the models using daily data on the Dow Index from January 2, 1953, to July 16, 1999. They discover that models with two volatility variables perform significantly better than

³The code and user guide are available at <http://www.aronaldg.org>. The programme is published under the terms of the GNU General Public Licence, version 2 or, at your option, any later version, which is published by the Free Software Foundation. You are free to redistribute and/or modify the programme.

⁴ See also Solibakke (2020, 2022).

models with just one. Additionally, they discover that logarithmic two-volatility component models outperform affine jump diffusion models and provide a good fit to the data. One of the two sources of volatility is quite persistent, whereas the other is substantially mean reverting.

I apply the logarithmic model with two stochastic volatility variables. I expand the model to facilitate the correlation between the mean and the stochastic volatility variables. I use the Cholesky decomposition in the correlation to ensure consistency. The introduction of asymmetry effects (correlation between return innovations and volatility innovations) is the key justification for correlation modelling. With $W_{1,t}$ $i = 1, 2,$ and 3 being standard Brownian motions (random variables), the generic SV model formulation for price change processes (y_t) is therefore Equation (2):

$$\begin{aligned}
 y_t &= a_0 + a_1(y_{t-1} - a_0) + \exp(V_{1t} + V_{2t}) \cdot u_{1t}, \\
 V_{1t} &= b_0 + b_1(V_{1,t-1} - b_0) + u_{2t}, \\
 V_{2t} &= c_0 + c_1(V_{2,t-1} - c_0) + u_{3t}, \\
 u_{1t} &= dW_{1t}, \\
 u_{2t} &= s_1 \left(r_1 \cdot dW_{1t} + \sqrt{1 - r_1^2} \cdot dW_{2t} \right), \\
 u_{3t} &= s_2 \left(r_2 \cdot dW_{1t} + \left(r_3 - (r_2 \cdot r_1) \right) / \sqrt{1 - r_1^2} \cdot dW_{2t} + \right. \\
 &\quad \left. \sqrt{1 - r_2^2 - \left(r_3 - (r_2 \cdot r_1) \right) / \sqrt{1 - r_1^2}} \cdot dW_{3t} \right).
 \end{aligned} \tag{2}$$

The vector of parameters is $(a_0, a_1, b_0, b_1, c_0, c_1, s_1, s_2, r_1, r_2, r_3)$. An internally consistent variance/covariance matrix is enforced by the correlation coefficients, or the three correlation parameters, r_1, r_2, r_3 , obtained by a Cholesky decomposition. c_0 is set to zero not to have two constants in the volatility equation. Taylor (1986), Clark (1973), Tauchen and Pitts (1983), and Rosenberg (1972) are early references. More recent references include Shephard (2004), Andersen (2002), Durham (2003), Gallant et al. (1993, 2010), Taylor (1982, 2005), and Chernov et al. (2003).

The model above contains three stochastic factors (Solibakke, 2020). It is also possible to apply Poisson distributions to jumps, although this greatly complicates calculations. To do a statistical analysis on a stochastic volatility model produced from a scientific procedure, the research employs a computational methodology suggested by Gallant and Tauchen (2010). I first compute the tractable likelihood function of a reduced-form auxiliary model (generous parameterization), which intuitively explains the strategy. The estimated set of score moment functions encode crucial details about the raw data sample's probabilistic makeup. I then use the continuous-time SV model to simulate a lengthy sample. I adjust the parameters in order to maximize the quasi-score moment functions that are assessed using the Metropolis-Hastings algorithm and parallel computing on the simulated data. Among the useful by-products are an explicit metric for assessing the severity of SV model failure and an extensive set of model diagnostics. The scientific stochastic volatility model is easy to simulate, but it cannot generate likelihoods.

The previous SV model estimation yielded a long-simulated realization of the state vector $\{\hat{V}_{i,t}\}_{t=1}^N, i = 1, 2$ and the accompanying $\{\hat{y}_t\}_{t=1}^N$ for $\theta = \hat{\theta}$. By calibrating the functional form of the

conditional distribution of functions supplied $\{\hat{y}_\tau\}_{\tau=1}^t$, assessing the outcome on observed data $\{\tilde{y}_\tau\}_{\tau=1}^n$, and producing predictions for $V_{i,t}, i = 1,2$ through Kalman filtering on y_i , extremely broad functions of can be employed and a large dataset is made available. An SNP model is estimated using the \hat{y}_t . A conditional variance of provided that is one step forward is represented by the final model. Regressions are performed on these series' $\hat{V}_{i,t}$ generously lengthy lags, $\hat{\sigma}_t^2, \hat{y}_t$ and $\{\tilde{y}_\tau\}_{\tau=1}^t$. Values for the volatility factors $\tilde{V}_{i,t}, i = 1,2$ at the original data points are obtained by evaluating these functions on the observed data series $\{\tilde{y}_\tau\}_{\tau=1}^t$.

2.4. Step ahead forecasting

2.4.1. Classic static regression

Classic Static Regression: Static forecasting repeatedly forecasts the dependent variable one step ahead. That is, compute each observation in the forecast sample, compute the $\hat{y}_{s+k} = \hat{\alpha}_1 + \hat{\alpha}_2 \cdot x_{s+k} + \hat{\alpha}_3 \cdot z_{s+k} + \hat{\alpha}_4 \cdot y_{s+k-1} + \dots + \hat{\alpha}_{N+4} \cdot y_{s+k-N}$, always using the actual value of the lagged endogenous variable. I must observe data for both exogenous and any lagged endogenous variables for observations in the forecast sample. This type of volatility forecasting only uses endogenous variables.

2.4.2. Machine learning regression models

In this section, I will define various machine learning regression techniques. I primarily train the models using lagged volatility information. I do not use the potential of future market data for training and prediction. That is, the three models for predicting continuous data are: (1) Lasso regression; (2) Ridge regression; and (3) Decision Forest regression. The Lasso (Least Absolute Shrinkage and Selection Operator) is a linear regression technique that helps to prevent overfitting by shrinking the coefficients of less important features to zero. The method introduces a tuning parameter called lambda (λ) that controls the amount of shrinkage applied to the coefficients. The Lasso model minimises the sum of the squared residuals (difference between predicted and actual values) and the penalty term, which is the absolute value of the sum of the coefficients multiplied by lambda. Increasing the lambda value will shrink more coefficients to zero, resulting in a simpler model with fewer features.

Ridge regression is a linear regression technique that helps to prevent overfitting by adding a penalty term to the sum of squared residuals. This penalty term is the sum of the squared coefficients multiplied by a tuning parameter called lambda (λ), also known as the regularization strength. The Ridge model minimises the sum of the squared residuals and the penalty term, where λ controls the trade-off between the model's fit to the data and the coefficients' magnitude. By increasing the value of lambda (λ), the Ridge model reduces the magnitude of the coefficients towards zero, resulting in a simpler model with smaller coefficients. Cross-validation, like Lasso, also determines the value of lambda in Ridge regression. I chose the best model with the lowest mean squared error on the validation set.

Decision forests, also known as random forests, are an ensemble learning method that combines multiple decision trees to make predictions. Hyperparameters are adjustable parameters that control the behavior of the decision forest model. Techniques such as grid search, random search, or Bayesian optimization, which evaluate different combinations of hyperparameters on a validation set to find the best-performing combination, can optimize hyperparameters for decision forests. Decision trees use

data-based splitting rules to segment the data into subsets. Assign the average of the target variable within a subset as the prediction for all observations that fall inside that subset. Recursive binary splitting splits a sample into segments to implement a single decision tree. This iterative approach determines where and how to split the data based on what leads to the lowest residual sum of squares (RSS). Single-decision trees can have low, non-robust predictive power and suffer from high variance. Random decision forests, which offer performance improvements by combining results from groups, or “forests”, of trees, can overcome this. In summary, the random decision forest algorithm randomly selects predictors as potential candidates for data splitting. Constructs a decision tree from a bootstrapped training set. Repeats the decision tree formation for a specified number of iterations. Averages the results from all trees to make a final prediction.

Lasso and Ridge regression aim to reduce prediction variances using a modified least squares approach. Let us look a little more closely at how this works. Recall that ordinary least squares estimate coefficients by minimizing the residual sum of squares (RSS):

$$RSS = \left[\sum_{i=1}^n \left(y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij} \right) \right]^2. \quad (3)$$

Penalized least squares estimates coefficients using a modified function:

$$S_\lambda = \left[\sum_{i=1}^n \left(y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij} \right) \right]^2 + \lambda J_2. \quad (4)$$

where λ is the tuning parameter and λJ_2 is the penalty term. The penalty term for the Lasso regression is (L1): $\lambda \sum_{j=1}^p |\beta_j|$, and for the ridge regression (L2): $\lambda \sum_{j=1}^p \beta_j^2$.

2.4.3. Artificial intelligence-neural networks

An artificial neural network fits a non-linear model to available data. In a timeseries model, the input lags relate to the first (second) hidden layers, or neurons. Note the lags, avoiding look-ahead bias. Neurons in the final hidden layer calculate the output dimensions and forecasts. Additionally, I use the RELU activation function in both C(R)NN (convolutional/recurrent neural network) and LSTM (long short-term memory) neural networks in this paper⁵. Figure 1 illustrates a simple one-layer neural network with three neurons that predicts step-ahead volatility: Features (lags) Neurons Step-ahead prediction Simple one-layer neural network with S neurons and biases (b) and an activation function, i.e., the softmax function.

⁵ See Hull (2021).

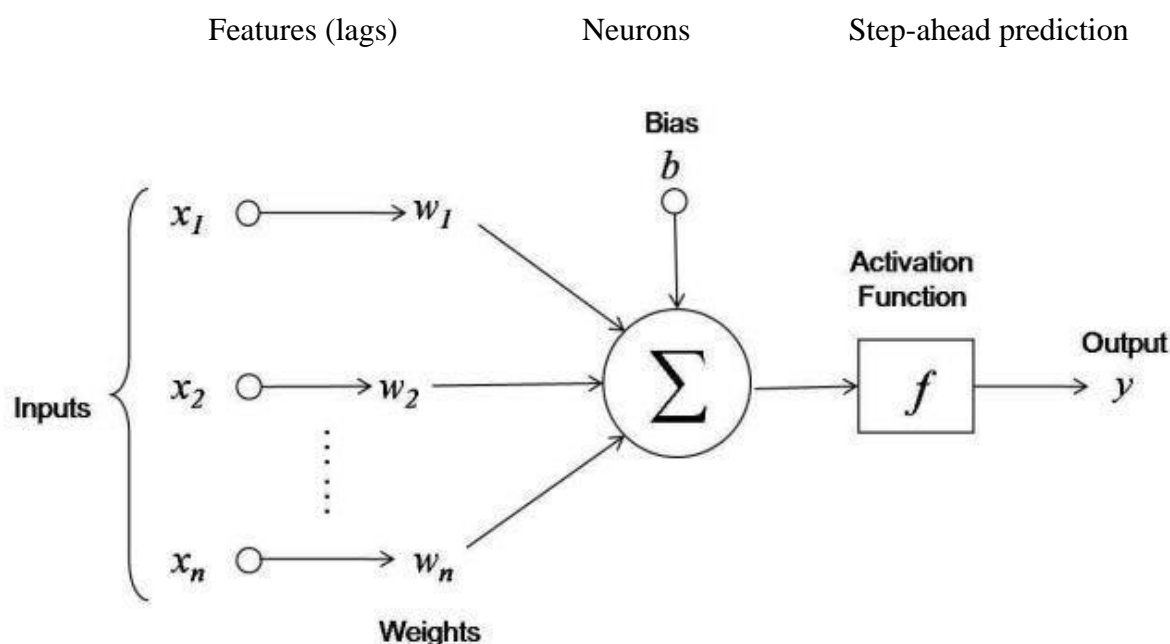


Figure 1. Simple one-layer Neural Network with Σ neurons and biases (b) and an activation function (i.e. softmax).

3. WTI oil front month prices and price movements 2010–2024 (3)

3.1. Data and stationarity

We ensure that the means, variances, and covariances (rather than the entire distribution) are time independent by enforcing weak stationarity. In other words, if that holds for all t , then a process $\{y_t\}$ is weakly stationary $E\{y_t\} = \mu < \infty, V\{y_t\} = E\{(y_t - \mu)^2\} = \gamma_0 < \infty$ as well $cov\{y_t, y_{t-k}\} = E\{(y_t - \mu)(y_{t-k} - \mu)\} = \gamma_k, k = 1, 2, 3, \dots$. Future observations are impacted by a shock to a stationary autoregressive process of order 1 (AR(1)) in a decreasing amount. Table 1 summarizes the characteristics of the price movement series. The data show a negative drift, suggesting a non-increasing WTI oil price level over a twelve years period (2014–2024(3) and 40,600 observations). The standard deviation on an hourly basis is 0.56, suggesting a yearly volatility of approximately 25%. The maximum (minimum) mean is 17 (–15). The kurtosis is high (101.7), indicating many observations around zero, followed by a few outliers. The skewness is positive (1.4), showing an asymmetric distribution for the movement densities. The high kurtosis likely leads to the strong rejection of the Cramer-Mises normality test (117.3). Serial correlation ($Q(12)$) reports serial correlation, while the Breusch-Godfrey LM statistic does not. Heteroscedasticity is present, as evidenced by the $Q^2(12)$ (1,327) and ARCH (12) (1,107) test statistics. These series are stationary (ADF (–184) and Phillips-Perron (–184)). However, the series reports some form of data dependence (BDS Z statistic > 30). The RESET test statistic (Ramsey, 1969) reports minor coefficient instability (14.9). Figure 2 shows the series of continuous movements. Overall, the series seems like usual financial market data. In the movement equations, I also tested with breaking trends, but my findings did not support trend breaks. A popular notion for assessing risk is value at risk (VaR), and Table 1 shows the 2.5% and 1% VaR numbers for market participants. Figure 3 displays the density important for non-normality.

Figure 4 reports correlogram for ordinary returns (top) and absolute returns (bottom) series. Finally, in Figure 5, a simple forecasting of normal, equal weighted, and GARCH(1,1) volatility models are forecasted. The volatility seems well described using a GARCH(1,1) model, showing heterogeneous variables/volatility.

Table 1. WTI oil high frequency (12 hours per day) price movements characteristics from 2014 to 2024(3) giving approx. 40,600 observations.

Mean (all)/ M (-drop)	Median Std.dev.	Maximum/ Minimum	Moment Kurt/Skew	Quantile Kurt/Skew	Quantile Normal	Cramer- von-Mises	Serial dependence Q(12)	Q ² (12)	VaR (1%; 2.5%)
-0.00423	0.00000	16.9792	101.7092	0.48745	403.1285	117.282	28.884	1327.70	-1.585%
0.00000	0.55892	-15.0454	1.36247	-0.01518	{0.0000}	{0.0000}	{0.0000}	{0.0000}	-1.055%
BDS-Z-stat. ($e = 1$) m=2	m=3	m=4	m=5	Phillips- Perron	Augmented DF-test	ARCH (12)	RESET (12;6)	Breusch- Godfrey	CVaR (1%; 2.5%)
33.1594	35.5297	35.3718	36.2117	-184.003	-183.998	1106.922	14.87576	1.56234	-2.435%
{0.0000}	{0.0000}	{0.0000}	{0.0000}	{0.0001}	{0.0001}	{0.0000}	{0.0212}	{0.9552}	-1.734%

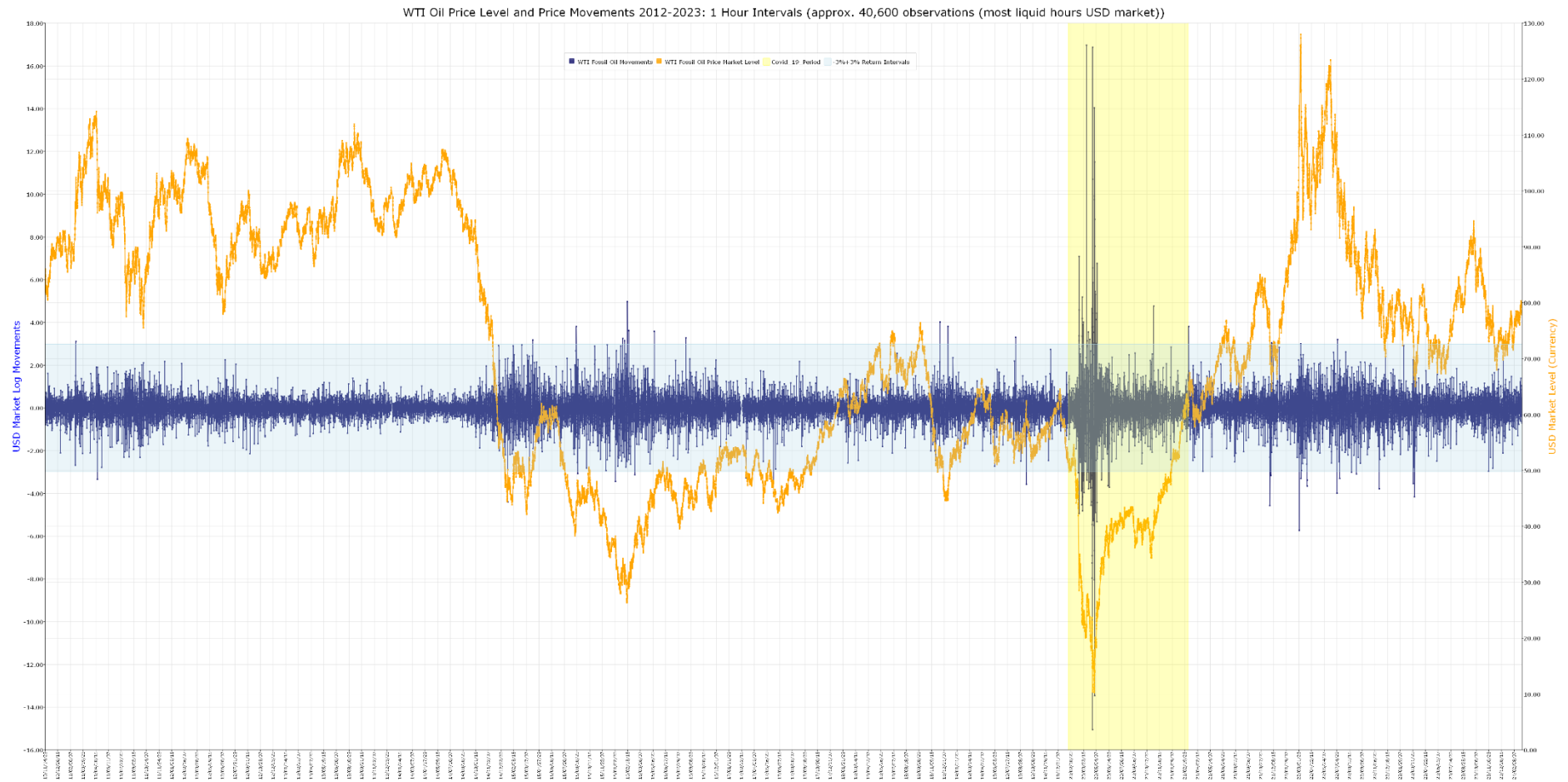


Figure 2. WTI oil High Frequency Price Movements characteristics for 2011–2024(3). The figure reports hourly price process (orange reported on the right axis) and price movements (blue on the left axis) for hourly WTI oil for the period 2016–2024(3). The vertical yellow area is the covid-19 period 2020 and the horizontal blue area is $\pm 3\%$ price movement bands.

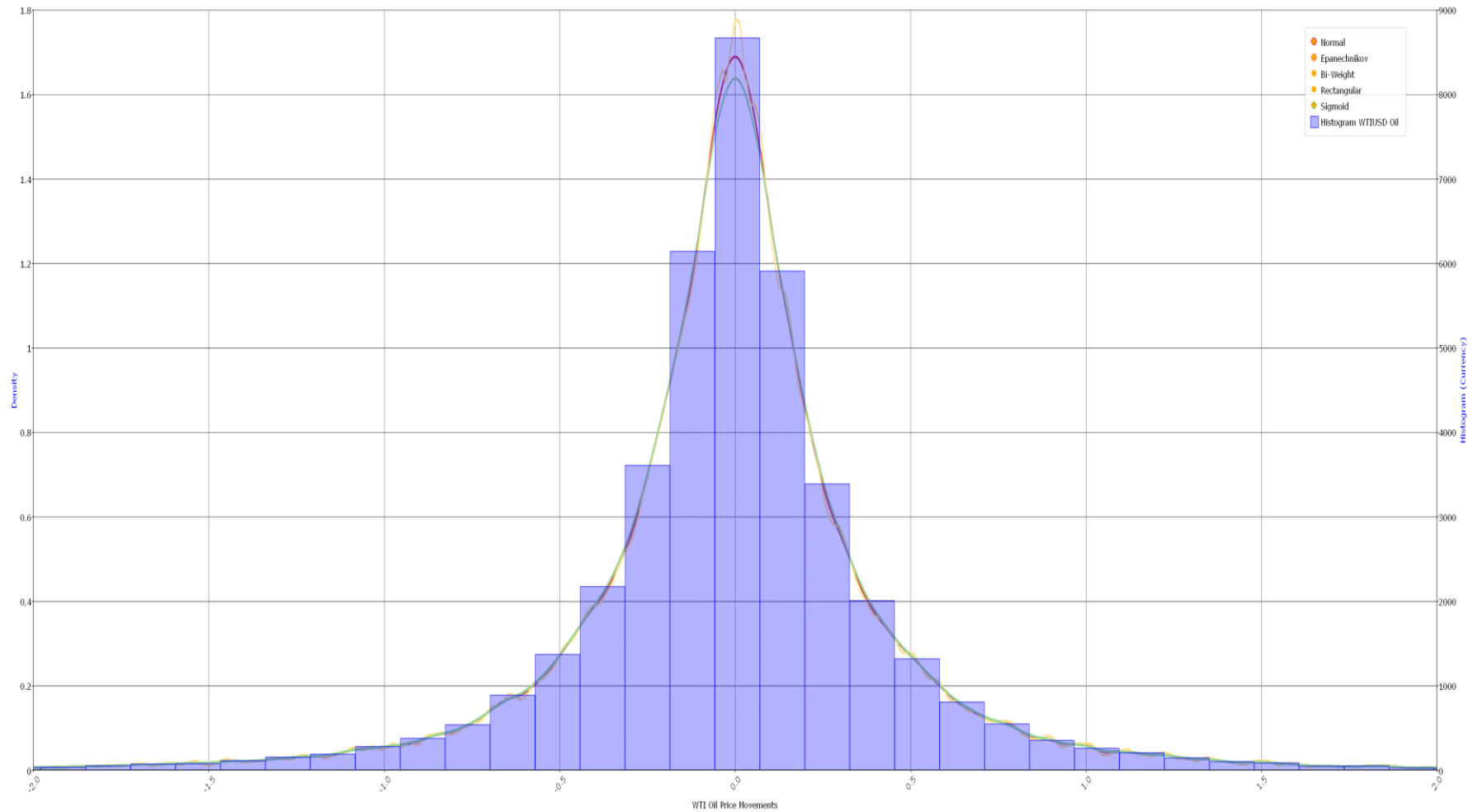


Figure 3. WTI oil price movements density characteristics for 2014–2024(3). The Figure reports kernel WTI oil price movement densities (normal, epanechnikov, bi-weight, rectangular, and sigmoid) with a histogram in the background.

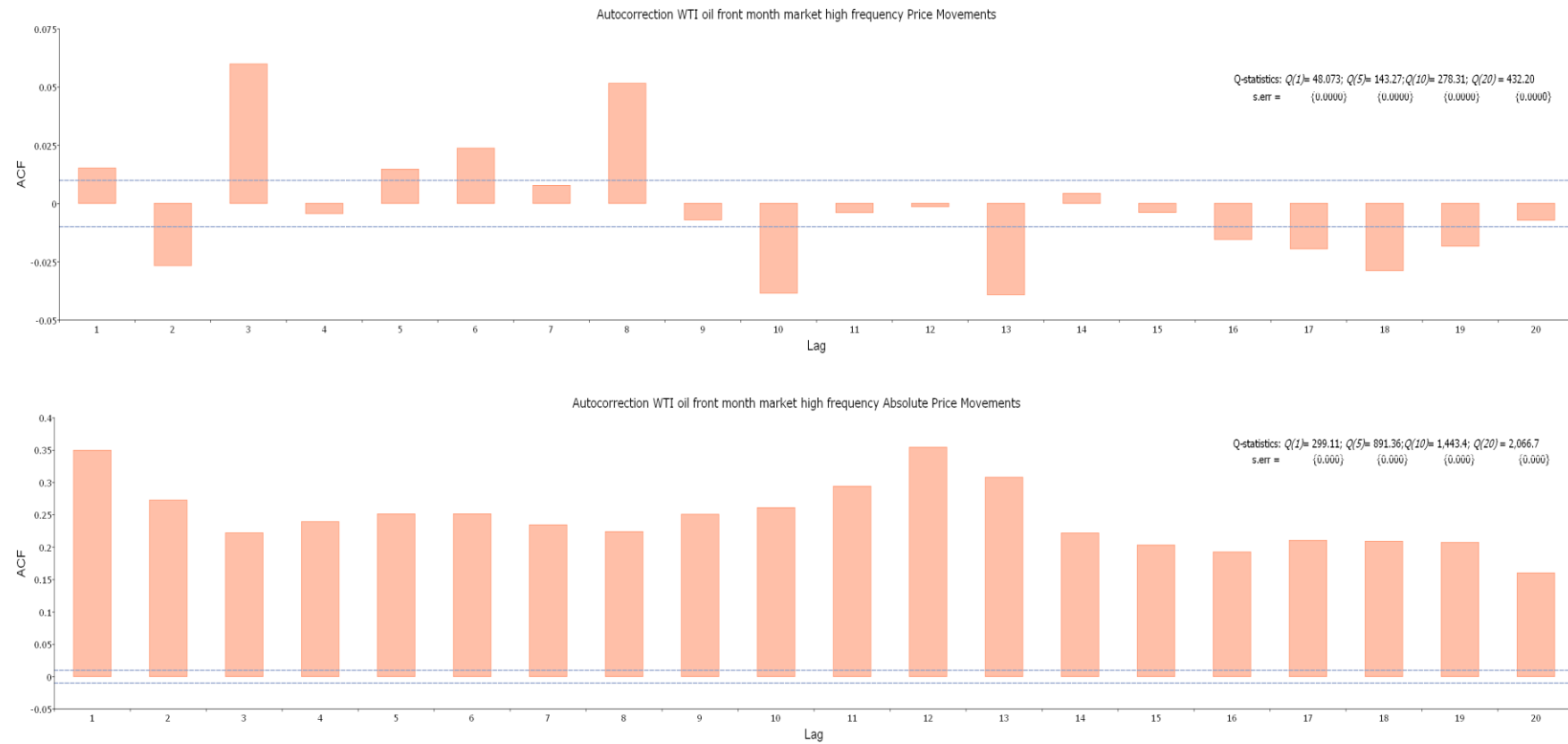


Figure 4. WTI Oil Price Movements autocorrelation characteristics

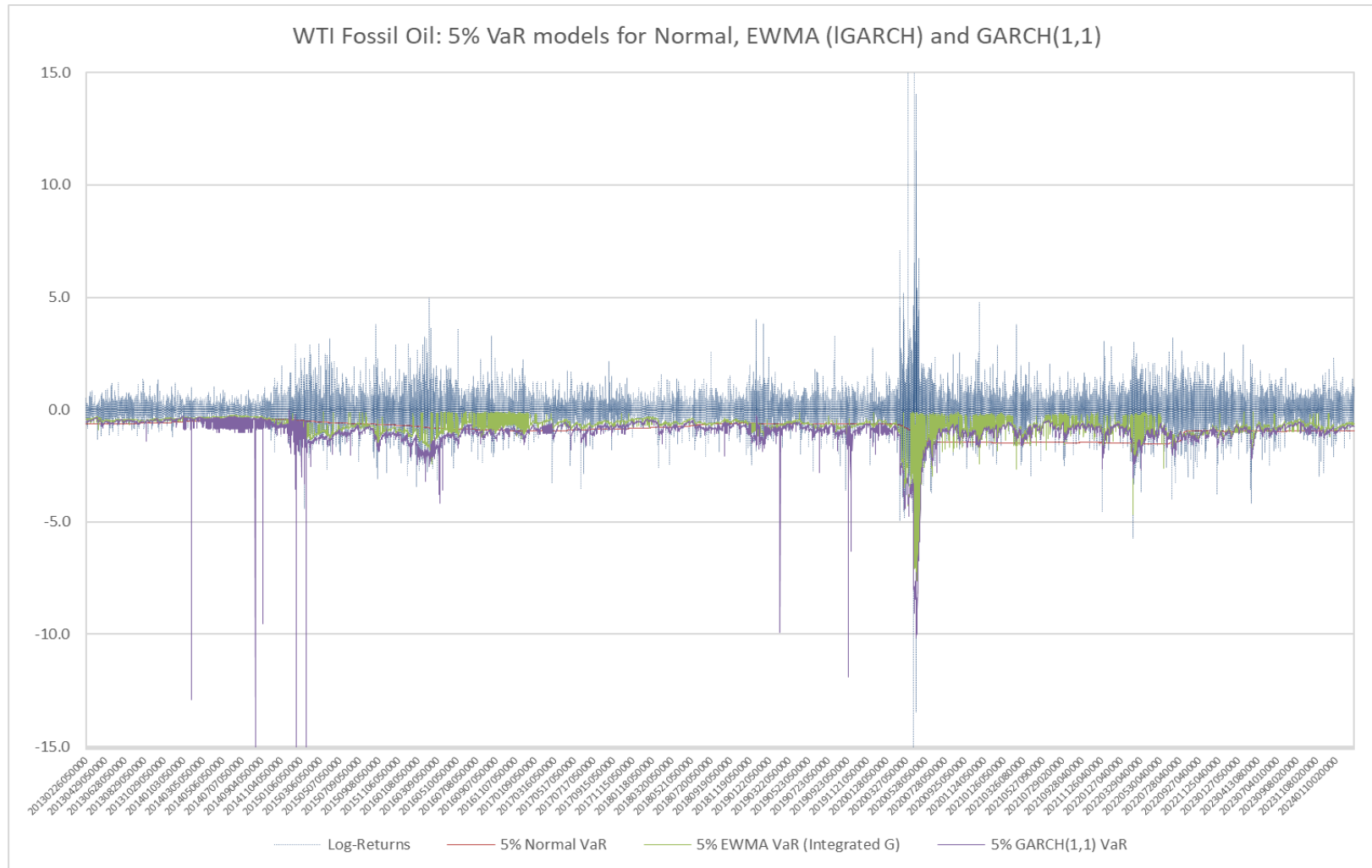


Figure 5. WTI oil Value at Risk (VaR) high frequency models (forecasts) 2014–2024(3). The figure report hourly price changes (blue), and 5% Normal (red), EWMA(1,1) (green), and GARCH(1,1) (purple) Value at Risk (VaR) for the period 2013–2024(3). The plot is computational intensive.

3.2. The SNP density projection

The conditional density, which completely describes how prices move, is, of course, the most important statistical variable. I fit the semi-nonparametric (SNP) model using conventional maximum likelihood and a model selection strategy that determines the appropriate order of expansion (BIC). I compute the Schwarz Bayes information criterion $BIC = s_n(\hat{\theta}) + \left(\frac{1}{2}\right) \left(\frac{p_p}{n}\right) \log(n)$ (Schwarz, 1978), preferring small criterion values.

Table 2 reports the maximum likelihood (ML) estimates of the parameters for the BIC-optimal SNP density models. First, for the mean, the intercept (η_7) is insignificant and the serial correlation (η_8) is negative (significant) (negative dependence). Second, the conditional variance coefficients ($\eta_9 - \eta_{13}$) are all strongly significant. The variance parameters' significance suggests conditional heteroscedasticity for the series. The P & Q companion matrix's greatest eigenvalue for the conditional variance function is 1.006. The spline specification reports mean reversion of the conditional variance even though the P and Q companion matrix is 1.006. Last, the BIC favors the hermite function coefficients ($\eta_1 - \eta_6$) up to the sixth polynomial lag expansion, which indicate modifications in the parametric model. Therefore, it is evident from the hermite result that there may be deviations from the traditional normally distributed and parametric conditional models. Asymmetry between positive and negative movements (η_{12}) seems not to be present. The quadratic and leverage reports corroborate the previously unreported finding (not reported). The high Q[1,1] (η_{11}) suggests volatility persistence.

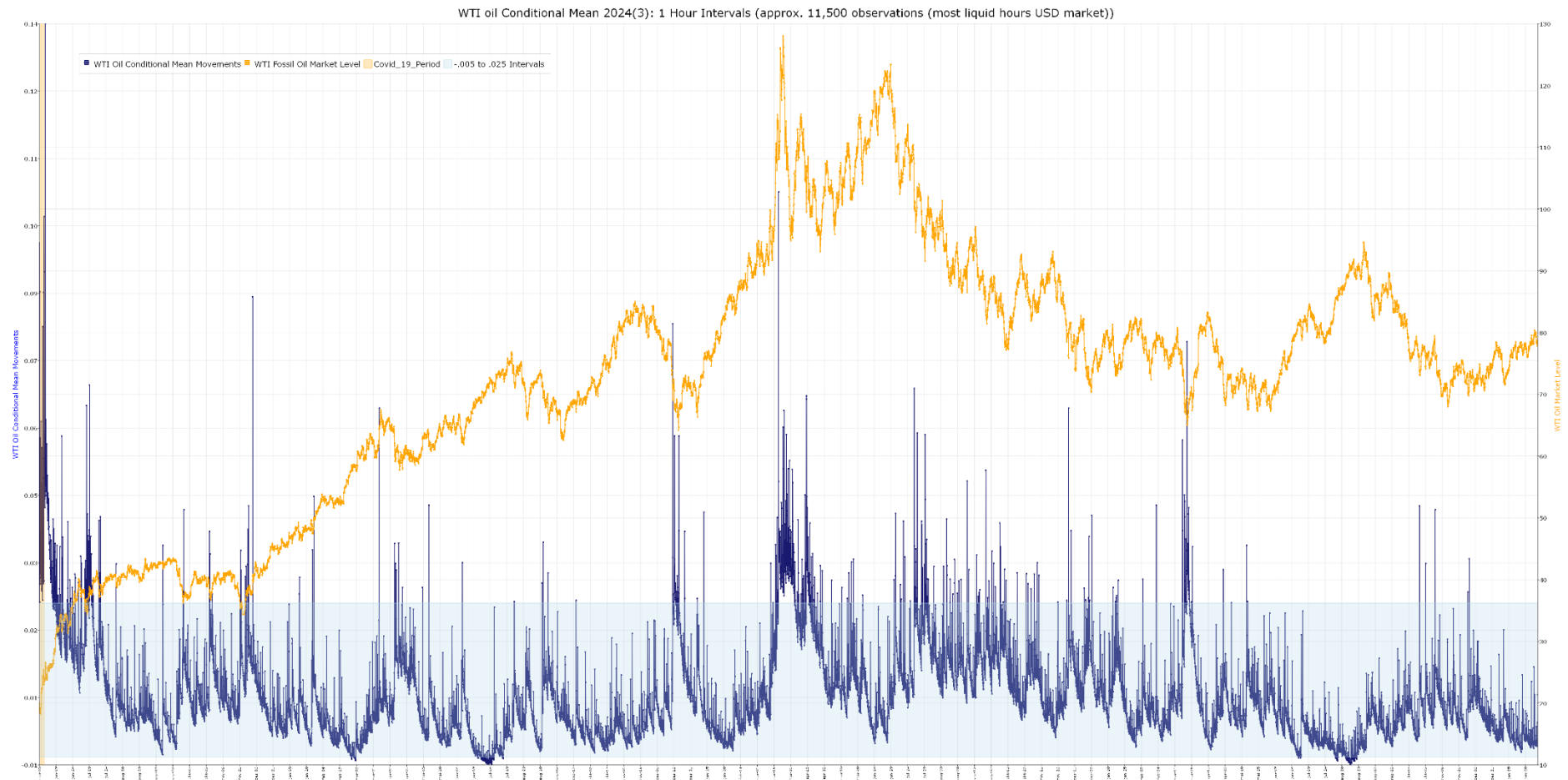
Table 2. SNP-Model Projection Specification for WTI oil price changes.

Statistical GMM Model: SNP-11116000			
Var	SNP Coeff.	WTI Oil	Std.error
<i>Hermite Polynoms</i>			
h_1	a0[1]	0.03766	{0.01086}
h_2	a0[2]	-0.11787	{0.00680}
h_3	a0[3]	-0.01124	{0.00566}
h_4	a0[4]	0.10584	{0.00437}
h_5	a0[5]	0.03939	{0.00616}
h_6	a0[6]	-0.13513	{0.00505}
<i>Mean Equation (Correlation)</i>			
h_7	b0[1]	-0.05076	{0.01680}
h_8	B(1,1)	0.00069	{0.00855}
<i>Variance Equation (Correlation)</i>			
h_9	R0[1]	0.08822	{0.00244}
h_{10}	P[1,1]	0.08720	{0.00840}
h_{11}	Q[1,1]	0.98646	{0.00057}
h_{12}	V[1,1]	-0.17172	{00000.0}
h_{13}	W[1,1]	0.39511	{0.01235}
<i>Log-Likelihood</i>		-14253.6	
<i>Model</i>	s_n	1.234503	
<i>Selection</i>	aic	1.235629	
<i>Criteras</i>	bic	1.239769	
<i>Largest eigen value mean</i>			0.00069
<i>Largest eigen value variance</i>			0.98072

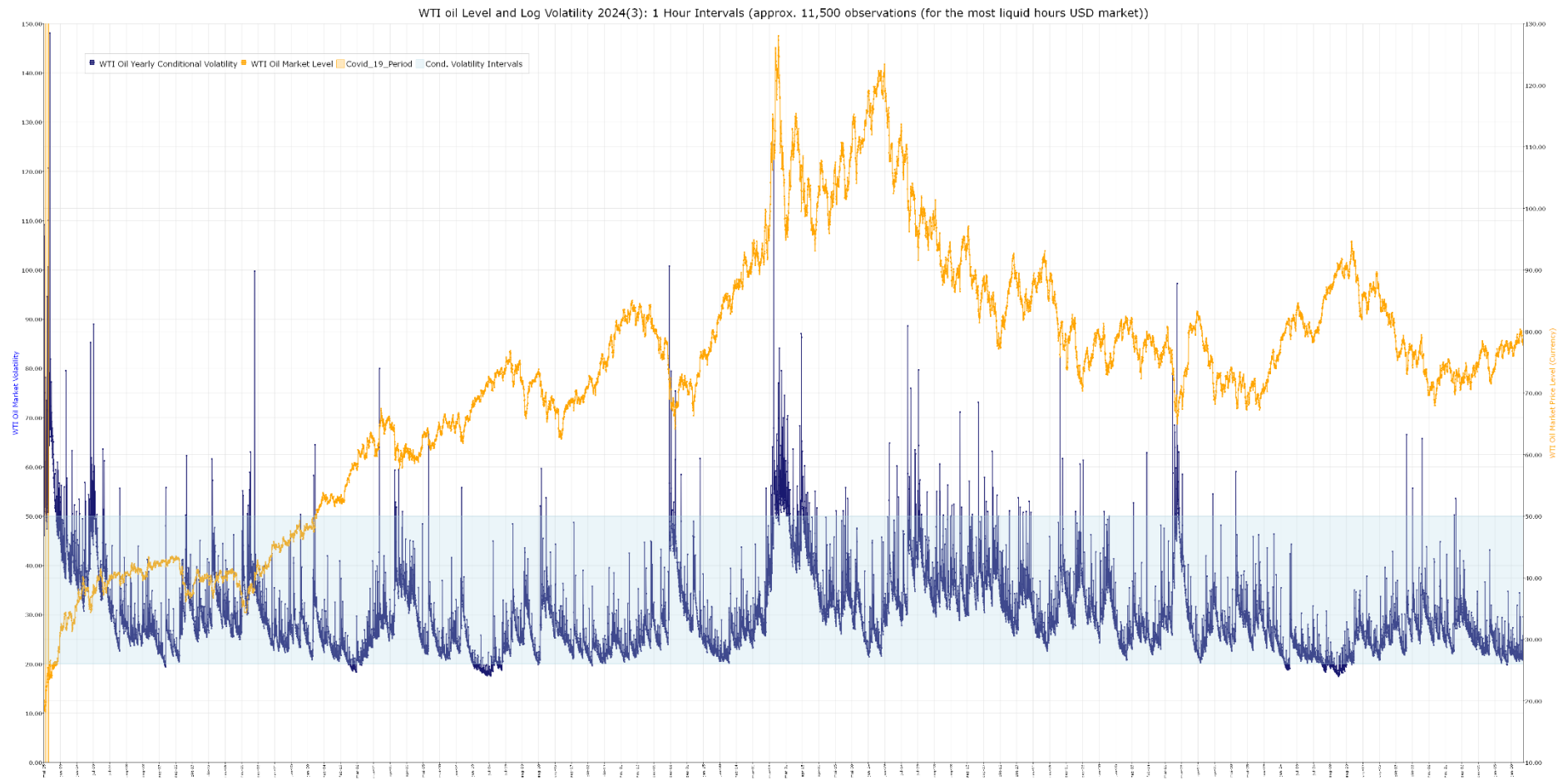
Table 3. SNP-Model Projection WTI oil SNP model residuals.

Mean/ Mode	Median/ Std.dev.	Maximum/ Minimum	Moment Kurt/Skew	Quantile Kurt/Skew	Quantile Normal	Cramer- von-Mises	Serial dependence Q(12)	Q ² (12)
-0.00429	0.00601	7.55081	4.89239	0.28977	40.43639	20.09614	13.4086	1301.9
	1.00018	-7.01311	-0.32293	0.00611	{0.0000}	{0.0000}	{0.3400}	{0.0000}
BDS-stat. ($\varepsilon=1$) m=2	m=3	m=4	m=5	ARCH (12)	RESET (12; 6)	Breusch- Godfrey	VaR 5%/1%	CVaR 5%/1%
3.37240 {0.0007}	3.47479 {0.0005}	1.58471 {0.1130}	0.02899 {0.9769}	76.0590 {0.0000}	8.02566 {0.2362}	1.4894 {0.2256}	-1.5935 % -2.9996 %	-2.457 % -4.036 %

The SNP projection densities give access to the conditional mean and volatility densities reported in Figure 6. Figure 7 shows densities for the same conditional mean and volatility. The figure reports a right skewed conditional volatility density (log-normal), while the mean displays leptokurtosis. Furthermore, Figure 8 allows for the generation and reporting of one-step-ahead mean densities based on lagged values. Any length of simulation path (bootstrapping) is available. All these results are in full compliance with the statistics from Table 1. Finally, for the GMM moment estimations, Table 3 reports residual statistics. The residual test statistics are closer to a normal distribution. However, the Cramer-von-Mises test for normality is significant. Furthermore, the BDS Z-statistics for $m = 2$ and 3 , as well as the ARCH (12) test statistics, indicate data dependence. Most likely, the BDS and ARCH signals remain heteroscedastic.



The figure reports the conditional mean (blue, left axis) for the WTI oil price changes together with WTI oil prices (right axis, yellow). The yellow horizontal bar indicates covid-19 period and the blue vertical band indicate -0.005 – 0.025% volatility band.



The figure reports the conditional volatility (blue, left axis) for the WTI oil price changes together with WTI oil prices (right axis, yellow). The yellow horizontal bar indicates covid-19 period and the blue vertical band indicate 20–50% volatility band.

Figure 6. Conditional mean and volatility for the WTI oil high frequency prices period 2020–2024(3) (hourly).

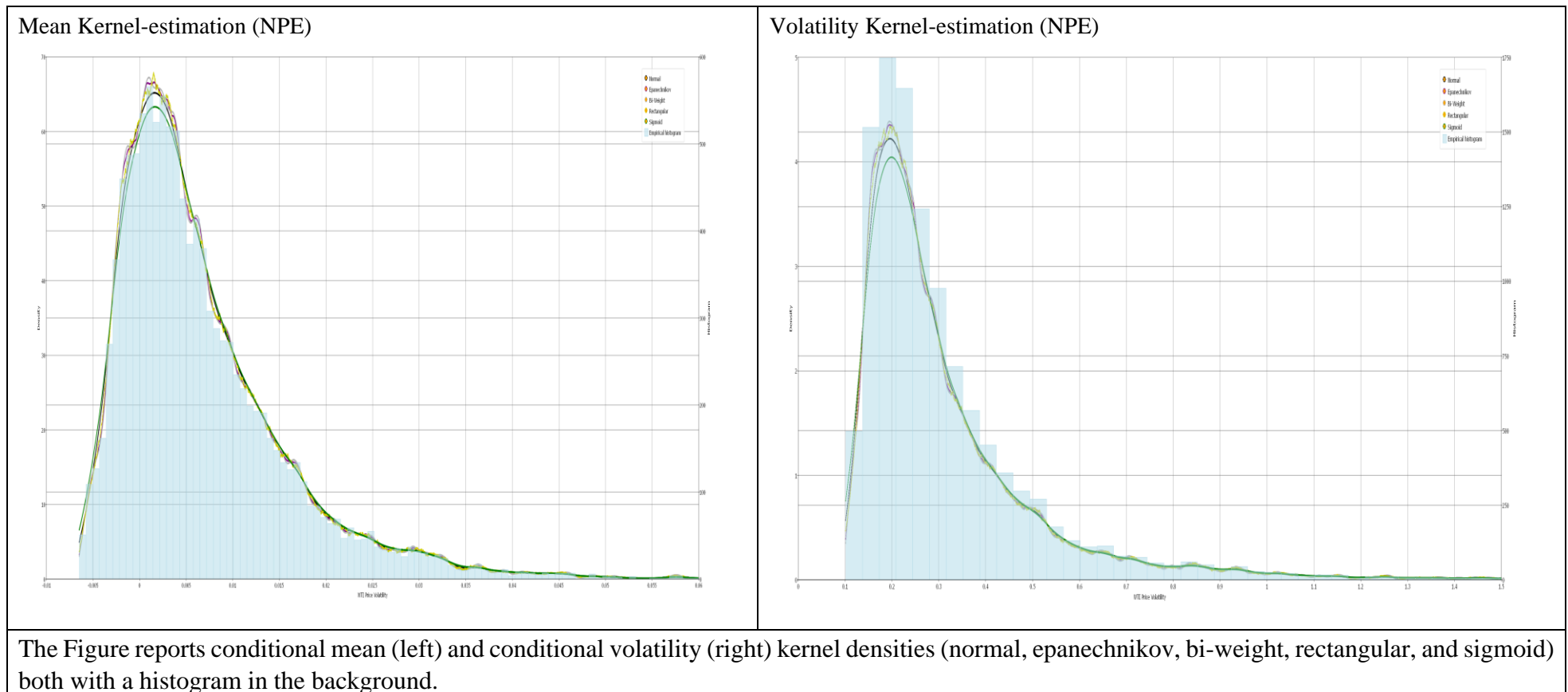


Figure 7. Conditional mean and volatility densities from WTI oil SNP model.

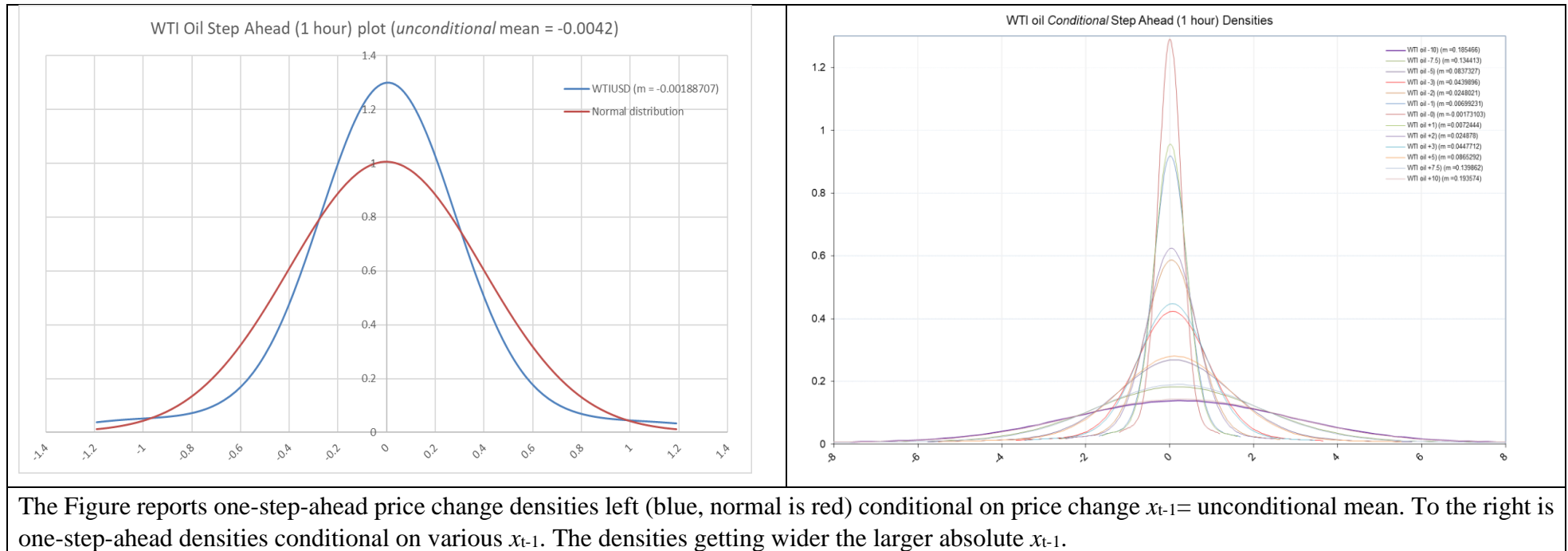
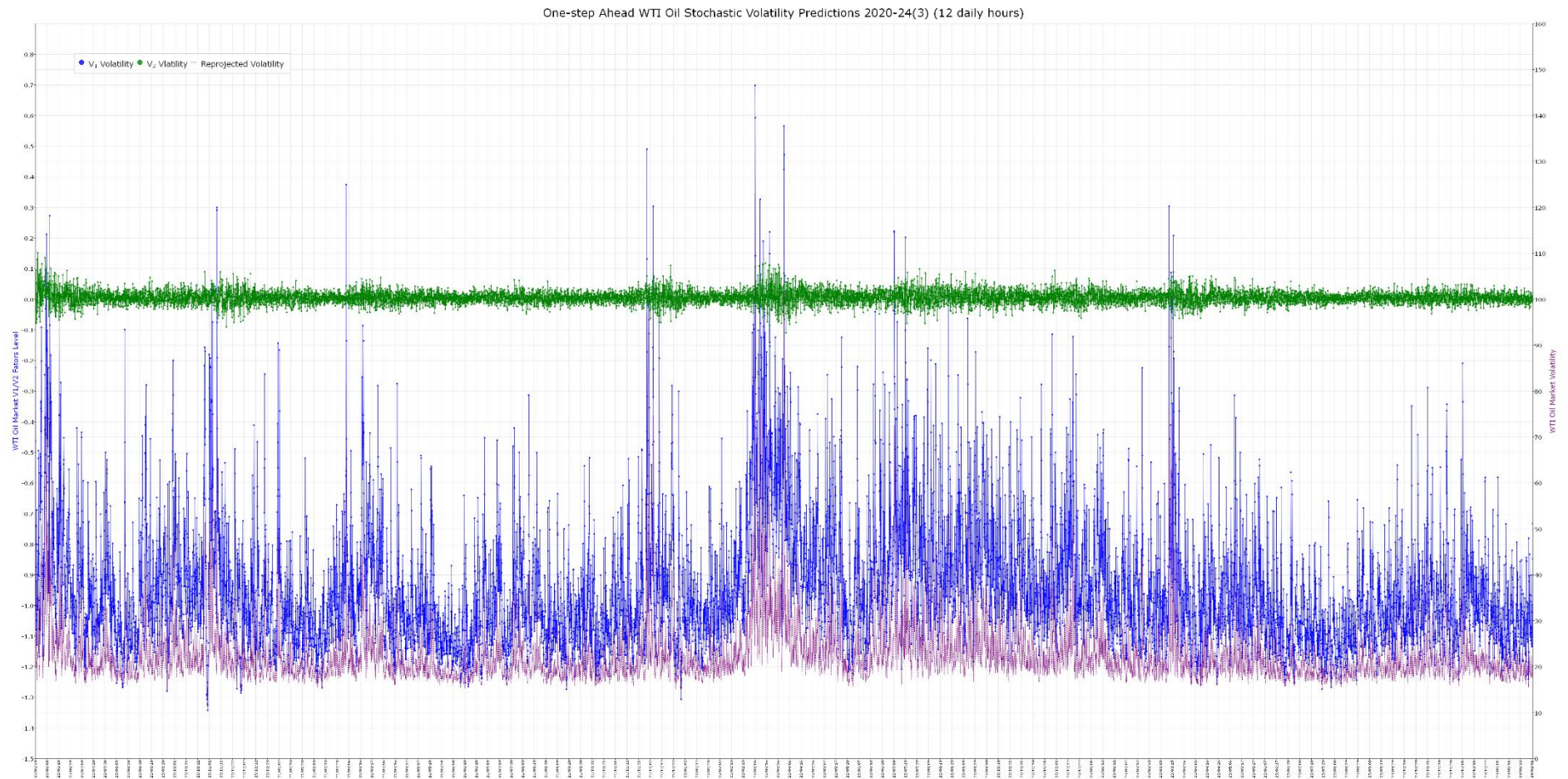


Figure 8. One-step ahead conditional densities (unconditional mean (left), and $-5\%, \dots, 5\%$ (right)).



The Figure reports the reprojected volatility factors, V_1 (blue) and V_2 (green). The reprojected volatility (purple) seems quite similar to V_1 (blue).

Figure 9. Volatility paths for two factors (V_1 and V_2) and re-projected volatility ($\exp(V_1+V_2)$).

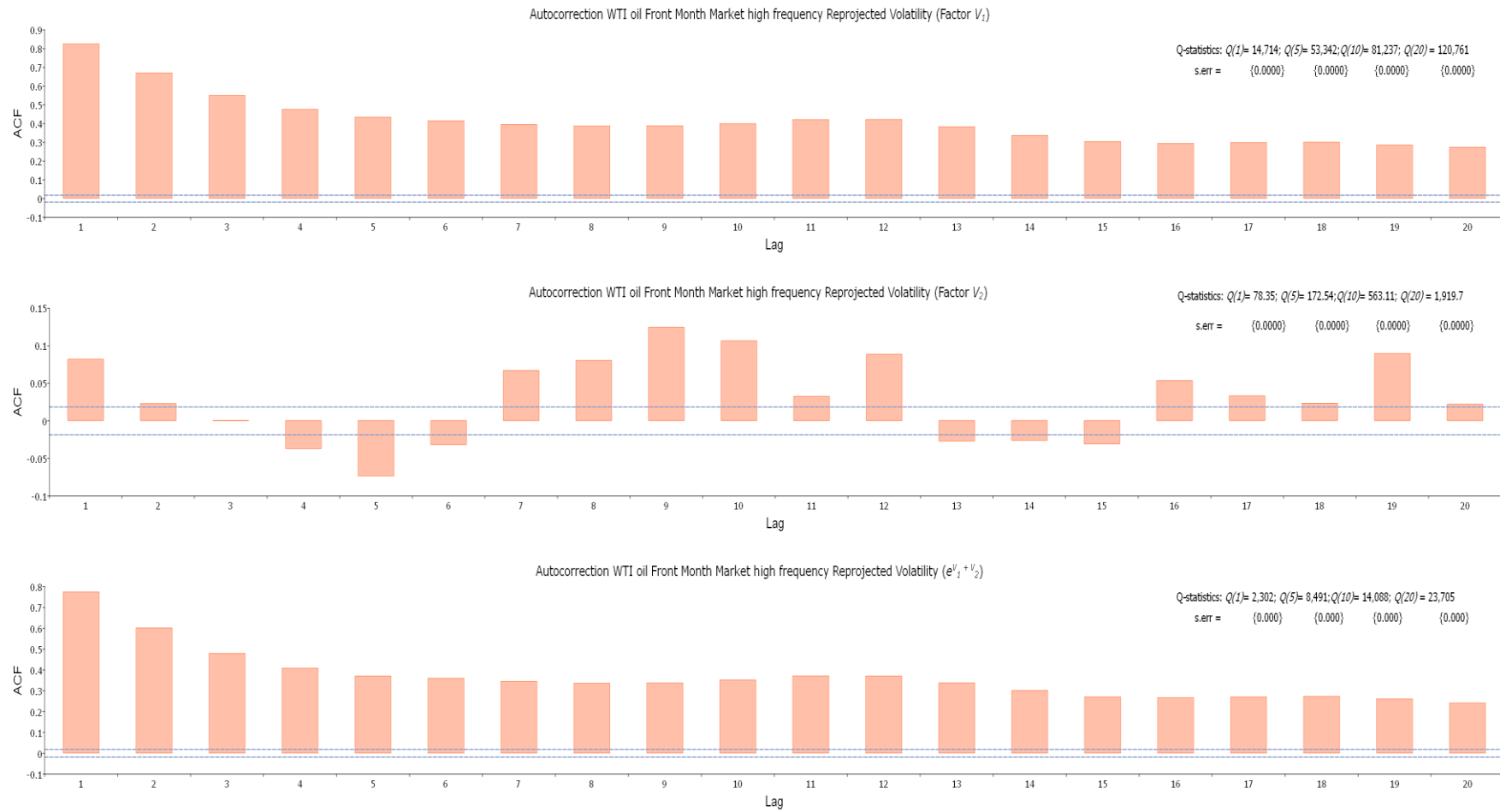


Figure 10. Correlogram for the volatility factors (V_1 and V_2) and $exp(V_1+V_2)$.

3.3. Stochastic volatility (EMM) for WTI oil front month prices

The parameter vector for the EMM model described in section 2.3, is reported in Table 4 below.

Table 4. SV model parameter vector θ for WTI oil high frequency prices.

Parameter values for Scientific model			
The WTI Oil High Frequency Prices model			
θ	Mode	Mean	Standard error (hess)
$a0$	0.013855	0.012895	0.001403
$a1$	0.001221	0.001529	0.001685
$b0$	-0.959530	-0.961890	0.003432
$b1$	0.785280	0.785230	0.000400
$c0/c1$	0	0	0
$s1$	0.305690	0.305910	0.000501
$s2$	0.253880	0.254000	0.000362
$r1$	0.051086	0.051039	0.001003
$r2$	-0.019867	-0.018614	0.001214
$r3$	0	0	0
Distributed Chi-square (freedoms)			$\chi^2(5)$
Posterior at the mode			-4.7650
Chi-square test statistic			{0.4452}

Table 5. Characteristics volatility-factors

Category	Mean (all)/ Mode	Median Std.dev.	Max/ Min	Moment Kurt/Skew	Quantile Kurt/Skew	Quantile Normal	Cramer- Mises	Serial dep. Q(12)	VaR (1%, 2.5%)	
Factor V_{1t}	-0.94070	-0.98748	0.6980	5.2064	0.12966	58.4788	76.844	2404.76	-1.2155	
		0.20338	-1.3432	1.76035	0.16421	{0.0000}		{0.0000}	-1.1896	
		BDS-Z-stat ($e = 1$)			Phillips- Perron	Augm DF-test		Breusch- Godfrey	CVaR (1%, 2.5%)	
		41.303	48.614	56.757	64.268	-66.813	-7.323		218.566	-1.2463
		{0.0000}	{0.0000}		{0.0000}	{0.0001}	{0.0000}		{0.0000}	-1.2187
		Mean	Median Std.dev.	Max/ Min	Moment Kurt/Skew	Quantile Kurt/Skew	Quantile Normal	Cramer- Mises	Serial dep. Q(12)	VaR (1%, 2.5%)
Factor V_{2t}	0.00592	0.00530	0.1529	2.2677	0.08283	3.4142	290.172	2106.38	-0.0471	
		0.02104	-0.1102	0.34166	0.01023	{0.1814}		{0.0000}	-0.0362	
		BDS-Z-stat ($e = 1$)			Phillips- Perron	Augm DF-test		Breusch- Godfrey	CVaR (1%, 2.5%)	
		27.090	33.668	38.903	44.264	-101.116	-8.971		204.872	-0.0604
		{0.0000}	{0.0000}		{0.0000}	{0.0001}	{0.0000}		{0.0000}	-0.0487
		Mean	Median Std.dev.	Max/ Min	Moment Kurt/Skew	Quantile Kurt/Skew	Quantile Normal	Cramer- Mises	Serial dep. Q(12)	VaR (1%, 2.5%)
Volatility $\exp(V_{1t}+V_{2t})$ (yearly)	22.131	20.595	123.06	31.4565	0.21565	107.722	138.291	2318.3	16.253	
		5.817	14.1801	3.96985	0.21397	{0.0000}		{0.0000}	16.762	
		BDS-Z-stat ($e = 1$)			Phillips- Perron	Augm DF-test		Breusch- Godfrey	CVaR (1%, 2.5%)	
		21.416	25.855	29.051	31.818	-117.090	-7.945		187.52	15.726
		{0.0000}	{0.0000}		{0.0000}	{0.0001}	{0.0000}		{0.0000}	16.211
		Mean	Median Std.dev.	Max/ Min	Moment Kurt/Skew	Quantile Kurt/Skew	Quantile Normal	Cramer- Mises	Serial dep. Q(12)	VaR (1%, 2.5%)

I use the efficient methods of moments (GMM-MCMC) to estimate the unobserved state vector, applying the Gallant and Tauchen (2010) method. In short, the EMM is a systematic approach generating moment conditions from the generalized method of the SNP moments estimator. The test-statistic $\chi^2(5) = 4.765$ with associated test statistic 0.45 from Table 4 (bottom) measure the extent of SV model failure⁶. A nonlinear Kalman filtering technique brings the unobserved vector back to reality. Figure 9 (11,500 observations) reports the unobserved vectors for the sub-period 2020–2024(3). The volatility appears to be divided into two factors: One factor that is very persistent (slow mean reversion) and a strongly mean-reverting factor. Table 5 reports the same figures' characteristics in numerical form.

The re-projected volatility, which is skewed to the right and non-normal, reports a yearly average of 22.1%. Both Phillips-Perron and Augmented Dickey-Fueller confirm the stationary nature of the data. Data dependence appears to be present, with a clearly higher value for volatility factor 1 (V_1) than for factor 2 (V_2). The correlogram for the factor volatilities in Figure 10 clearly suggests a strong serial correlation in the two series and in particular for V_1 and $\exp(V_1 + V_2)$. Moreover, the Brock, Decker, and Scheinkman (BDS) test statistic for the correlation integral from 2 to 5 (m) suggests significant data dependence.

Table 6. WTI oil hourly volatility ordinary OLS (qr) step ahead fit measures.

Hourly Estimated Stochastic Volatility Forecast Fit Measures (EViews)					
Contracts	Error Measures	Factor 1 V_{1t}	Factor 2 V_{2t}	Reprojected Volatility	
	Root Mean Square Error (RMSE)	0.00778	0.07580	0.0334	
	Mean absolute Error (MAE)	0.00555	0.05405	0.0223	
WTI Oil	Mean absolute percent error (MAPE)	0.5624	528.168	5.3733	
Front Month	Theil inequality coefficient (U_1)	0.00381	0.37021	0.04165	
Movements	Bias proportion		0.0004	0.0004	0.00187
(Hourly)	Variance Proportion		0.0198	0.2915	0.01952
	Covariance Proportion		0.9798	0.7082	0.97861
	Theil U_2 Coefficient	0.0774	0.72245	0.86221	
	Symmetric MAPE	0.563083	83.4825	5.40013	

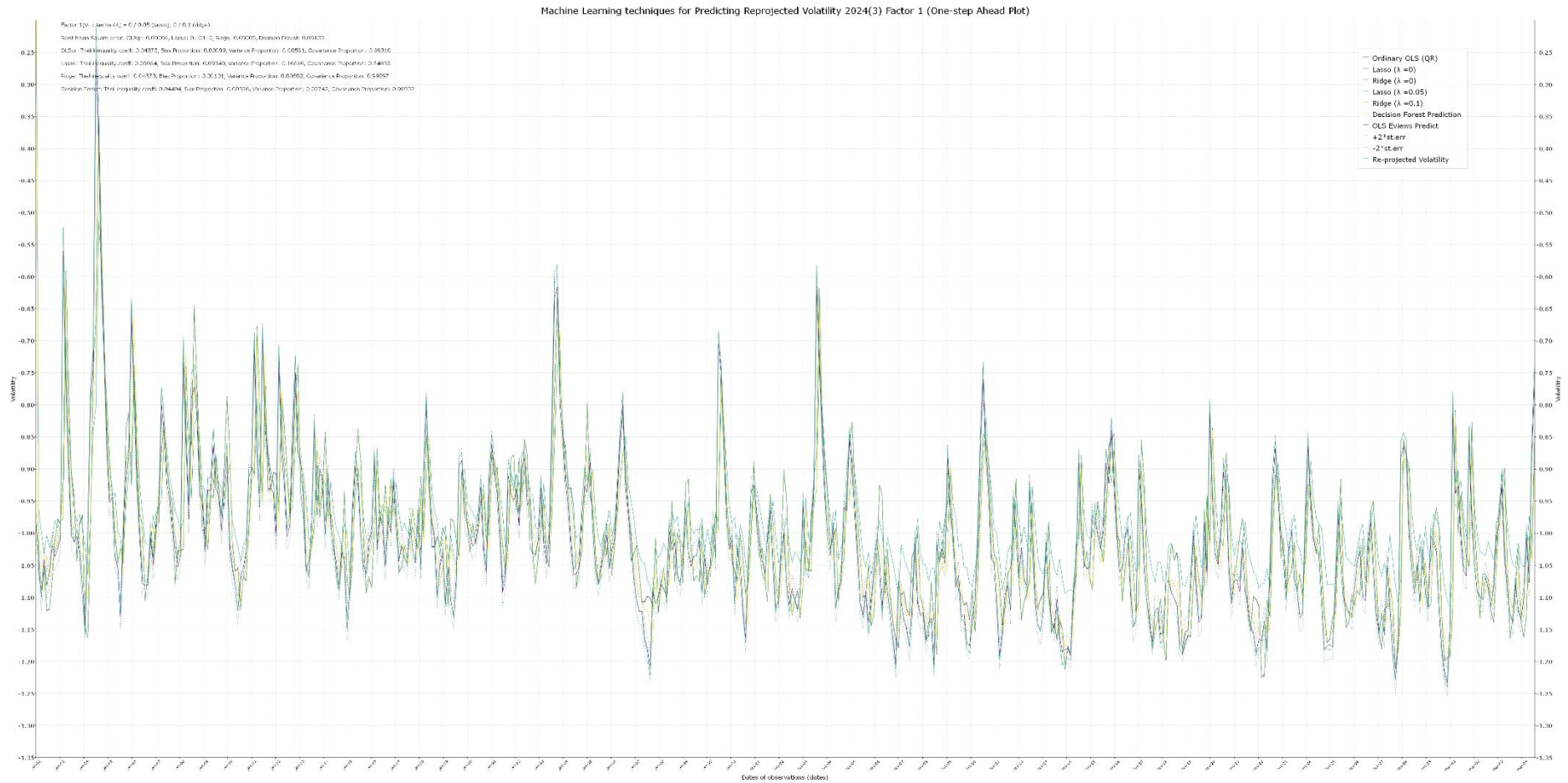
4. Forecasting re-projected volatility factors

4.1. Classical OLS static forecasting

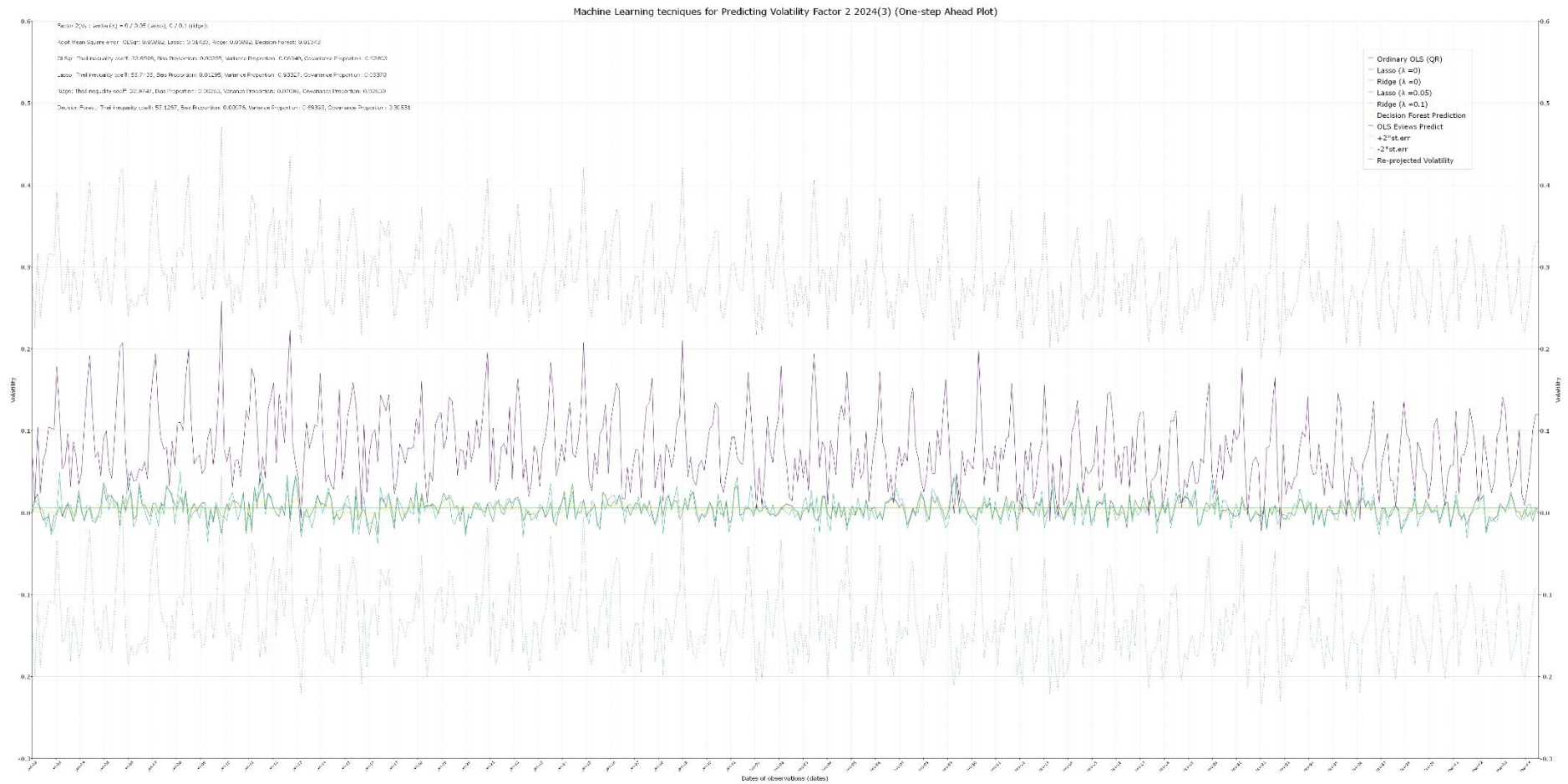
The WTI oil Hourly Volatility Ordinary OLS Step-Ahead Fit Measures demonstrate a data dependence that suggests predictability through serial correlations. Table 6 reports the classical step-ahead forecasts using only endogenous variables (lags), and provide numbers for the volatility fit. The Theil's covariance measure is quite high. Factor 1 (V_1) has a 97.9% Theil covariance fit, indicating that it is clearly predictable. Factor 2 (V_2) report a lower Theil covariance measure of 70.8%. The re-projected volatility ($\exp(V_1 + V_2)$) reports a close to factor 1 (V_1) result of 97.8%. Therefore, it appears that I can project the volatility. However, trading volatility from these factors can become challenging, necessitating a close watch on other factors to prevent sudden losses from

⁶ Degrees of freedom (5) is the number of SNP model parameters minus SV-model parameters –1.

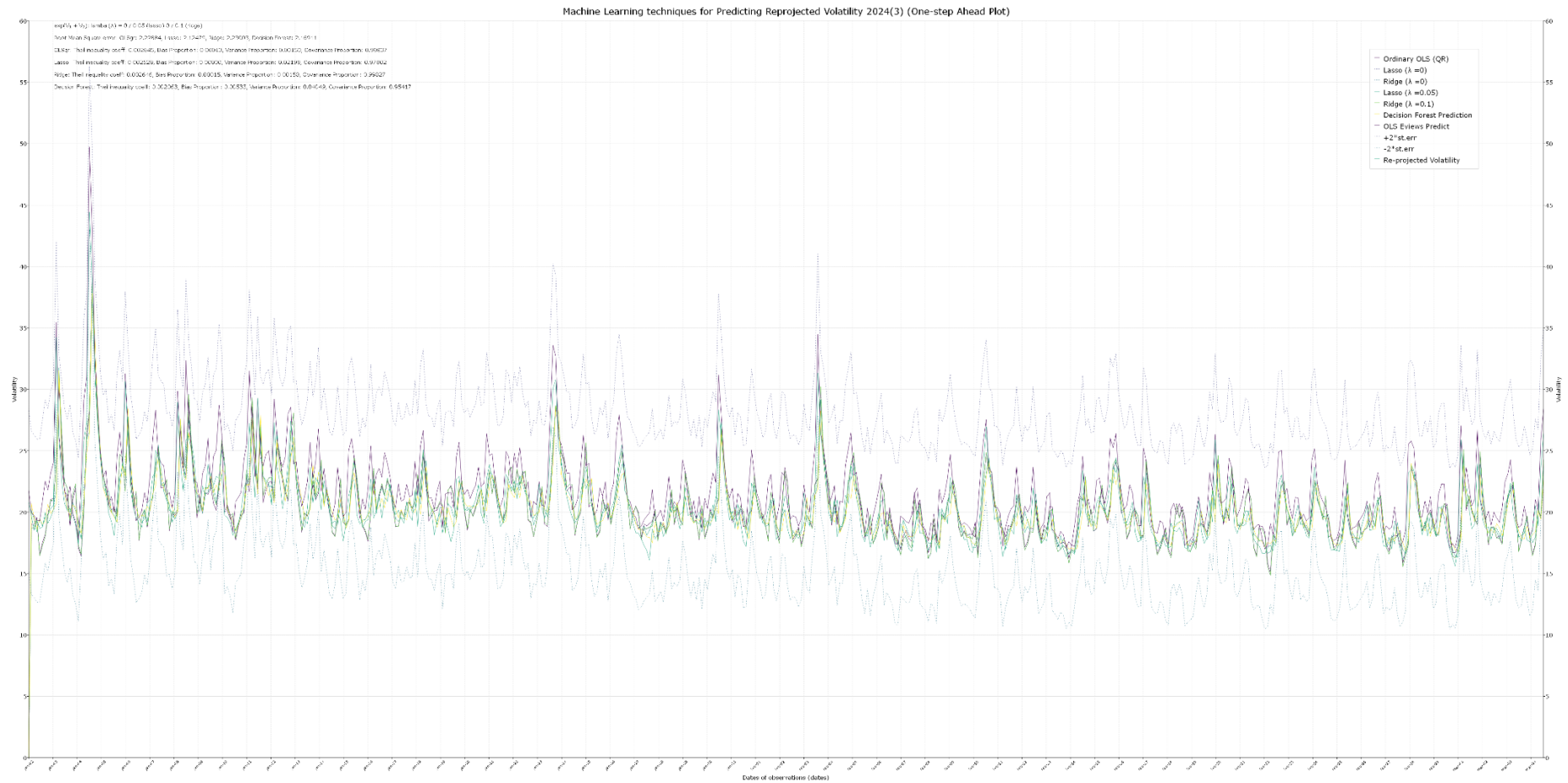
the mean reversion process. Furthermore, when considering portfolio theory and systematic risk compensation, it's crucial to keep in mind the significance of covariance.



The Figure reports V_1 actual volatility together with predicted volatility (ordinary, lasso, ridge, decision forest) and ± 2 ordinary regression robust standard deviations



The Figure reports V_2 actual volatility together with predicted volatility (ordinary, lasso, ridge, decision forest) and ± 2 ordinary regression robust standard deviations



The Figure reports actual Reprojected volatility ($\exp(V_1+V_2)$) together with predicted volatility (ordinary, lasso, ridge, decision forest) and ± 2 ordinary regression robust standard deviations.

Figure 11. Predictions from OLS, ML with 95% robust confidence intervals.

4.2. Machine learning regression techniques and step ahead forecasting

Table 7 reports the prediction results from machine learning (ML) regression techniques for the volatility factors (V_1 (left), V_2 (middle) and $\exp(V_1 + V_2)$ (right column)). Figure 11 contains the prediction plots for the same ML regression models. For the reprojected volatility, the Lasso model with a $\lambda = 0.05$ ($L1$ penalty is marginally positive) seems to report the lowest RMSE, followed by the Decision Forest model. Furthermore, the Theil covariance measure is clearly highest for the Lasso model, with $\lambda = 0$ (99.8%). The Ridge model with $\lambda = 0.1$ follows closely (99.8%). Generally, the Lasso model with lambda regularization serves as a powerful feature selection technique, enabling the construction of accurate and interpretable regression models. When the data contains numerous correlated features, people often use ridge regression. The regularization term shrinks the coefficients towards each other, which can improve the stability and interpretability of the model. Ridge regression with lambda regularization is therefore a useful technique for reducing overfitting and improving the performance of linear regression models. In this paper the λ uses values from $0 \leq \lambda < 1$. I use hyperparameters to tune and improve the decision forest models.

Confidence intervals are not available using machine learning (ML) techniques. However, Table 6 reports fit measures for all regression models, which are all comparable with the static OLS forecasts. For volatility factor 1 (V_1), the regression techniques Lasso ($\lambda = 0$), Lasso ($\lambda = 0.05$), Ridge ($\lambda = 0.1$), and Decision Forest report RMSEs of 0.090, 0.101, 0.090, and 0.091, respectively. Furthermore, for factor V_1 , the Lasso ($\lambda = 0$), Lasso ($\lambda = 0.05$), Ridge ($\lambda = 0.1$), and Decision Forest regressions report a Theil covariance measure of 99.3%, 74.0%, 99.3%, and 96.9%, respectively. Note however, that the classical static forecasts (OLS) in Table 6 report a Theil covariance fit of 99.3%, approximately equal to Ridge regression. The covariance values for volatility factor 2 (V_2) are 92.8% for Lasso ($\lambda = 0$), 3.4% for Lasso ($\lambda = 0.05$), 92.7% for Ridge ($\lambda = 0.1$), and 30.5% for Decision Forest regression. These values are clearly lower than those for factor 1 (V_1). Hence, neither the OLS nor the ML techniques seem to report good fits for factor 2 (V_2). For the re-projected volatility, the Lasso ($\lambda = 0$), Lasso ($\lambda = 0.05$), Ridge ($\lambda = 0.1$), and Decision Forest regression techniques report covariance measures of 99.8%, 97.8%, 99.8%, and 95.4%, respectively. The regression results for the reprojected volatility is clearly closer to factor 1 (V_1) than factor 2 (V_2).

The improvements made using ML are marginal, so it is important to evaluate ML techniques in light of the use of computer resources. Moreover, thorough tuning of the hyperparameters for the ML techniques may improve performance. Consequently, based on the regression results, the machine learning regression models seem to expand the possibilities for volatility predictions. For the re-projected volatility ($\exp(V_1 + V_2)$), the ML Lasso regression techniques ($\lambda = 0$ and 0.05) and ridge ($\lambda = 0.1$) report overall best fits ($\approx 98 - 99\%$). In fact, the Lasso and Ridge regression models report a Theil covariance measure that is 5% better than the Decision Forest regression model.

In ML, over-(under-)fitting is important to avoid, referred to as the bias-variance trade-off. For all methods, hyperparameters are used for tuning and therefore minimize the trade-off. It will therefore be important to use hyperparameters to tune and improve the models. I can categorize the error of a machine learning model into two major categories: Bias and variance (refer to the Theil measure). The error that occurs when I fit a simple model to a more complex data-generating process. A model with high bias will underfit the training data. When I apply our model to a new dataset that it has not seen, the expected prediction error occurs. A model with a high variance will typically overfit the training

data, resulting in lower training set errors but higher errors on any data not used for training. The model validation process in the previous section works when I have large datasets. When data is scarce, I must resort to a technique known as cross-validation. The purpose of cross-validation is to provide a better estimate of a model's ability to perform on unseen data. It provides an unbiased estimate of the generalization error, especially in the case of limited data. There are many reasons I may want to do this: (1) To have a clearer measure of how the model performs; (2) to tune hyperparameters; and (3) to make model selections. The idea behind cross-validation is simple: rather than training our models on one training set, I train them on multiple subsets of data. I follow the basic steps of cross-validation: (1) Divide the data into portions; (2) train our model on a subset of these portions; (3) test our model on the remaining subsets of the data; and (4) repeat steps 2–3 until I have trained and tested the model on the entire dataset. Average the model performance across all iterations of testing to get the total model performance.

4.3. Neural networks (CNN/RNN/LSTM)

Figure 12 reports a neural network (CNN/RNN) for V_{1t} . The fit is satisfactory, and the one-step-ahead predictions have a low MSE (0.09). Figure 13 shows that V_{2t} has a much higher MSE (0.6). Hence, the neural networks report satisfactory forecasts for V_{1t} . However, V_{2t} seems difficult to forecast for neural networks. Note that V_{1t} seems to be the most important factor for the accumulated for the reprojected volatility ($\exp(V_{1t} + V_{2t})$).

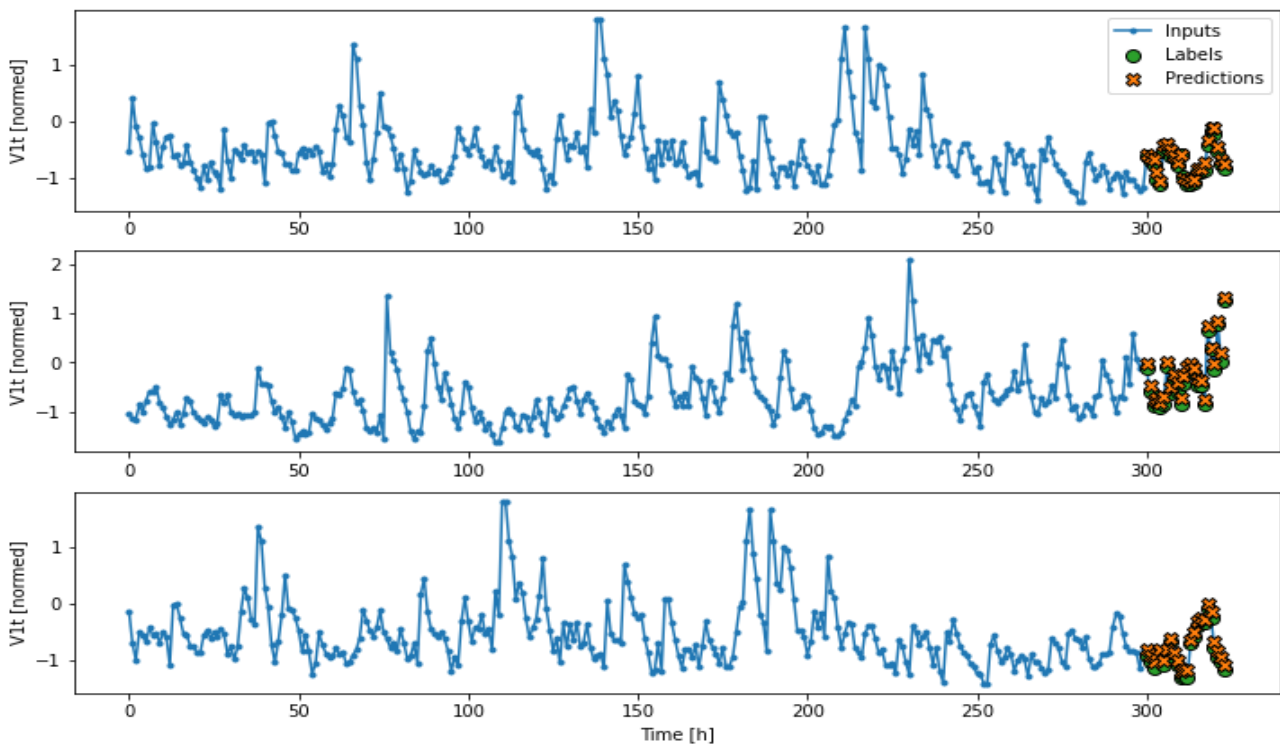
Table 7. WTI oil hourly volatility machine learning step ahead fit measures.

Hourly Estimated Stochastic Volatility Forecasts Fit				
Category	Factor 1 V_{1t}	Factor 2 V_{2t}	Reprojected Volatility ($e^{(V_1+V_2)}$)	
<i>Lasso Regression</i>				
<i>Ridge Regression</i>	RMSE	0.090061	0.008824	2.22884
($l = 0.0$)	MSE	0.059361	0.006768	1.43976
	MAPE	6.616165	190.1440	6.80762
	Theil inequality c 1 (U_1)	0.043735	22.85882	0.00264
	Bias Proportion	0.00099	0.00255	0.00013
	Variance P	0.00591	0.06940	0.00150
	Covariance P	0.99310	0.92805	0.99837
	Theil U_2 Coefficient	0.77736	0.48436	1.04657
	Symmetric MAPE	0.01071	0.11457	0.01288
<i>Lasso Regression</i>				
($l = 0.05$)	RMSE	0.101101	0.014326	2.12479
	MSE	0.074398	0.011201	1.27648
	MAPE	8.125197	248.4738	5.94658
	Theil inequality c 1 (U_1)	0.050636	55.74350	0.00253
	Bias Proportion	0.09349	0.01295	0.00000
	Variance P	0.16616	0.95327	0.02198
	Covariance P	0.74035	0.03378	0.97802
	Theil U_2 Coefficient	1.07112	0.41586	0.93108

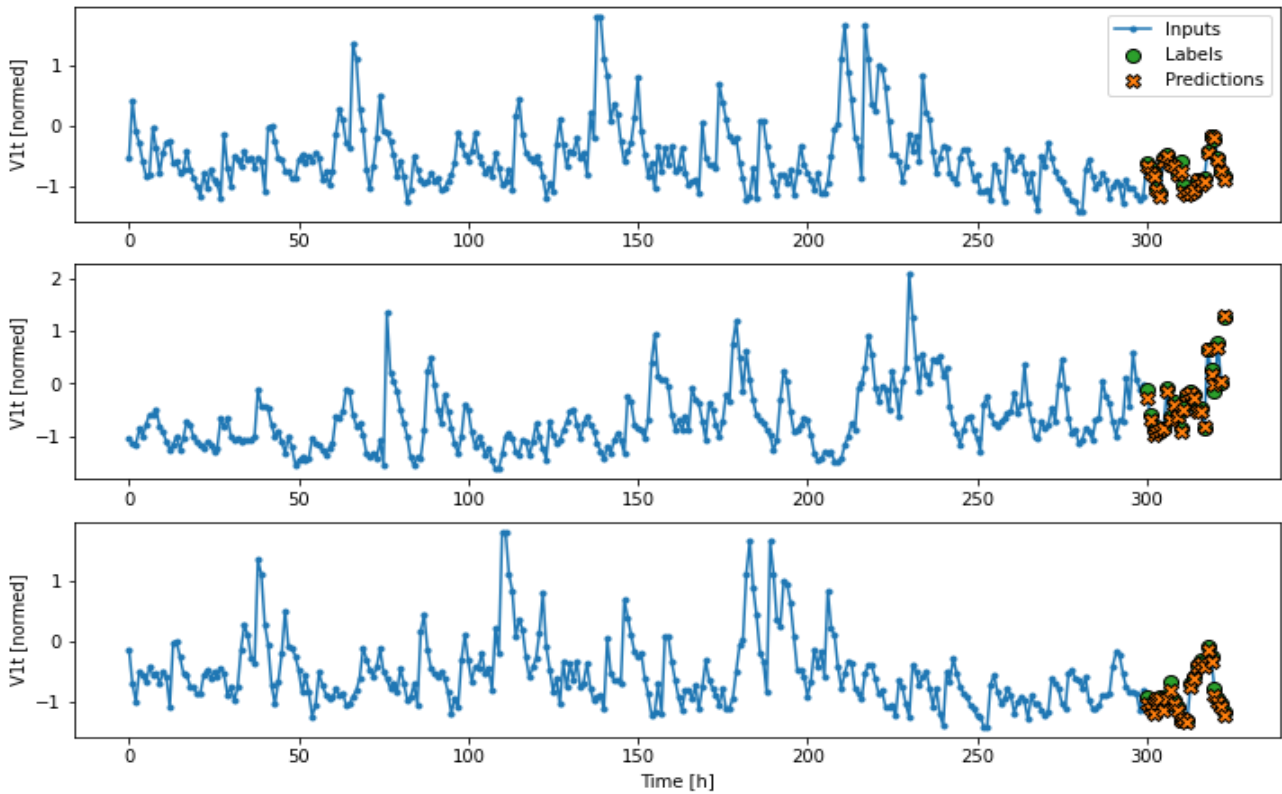
Continued on next page

Hourly Estimated Stochastic Volatility Forecasts Fit

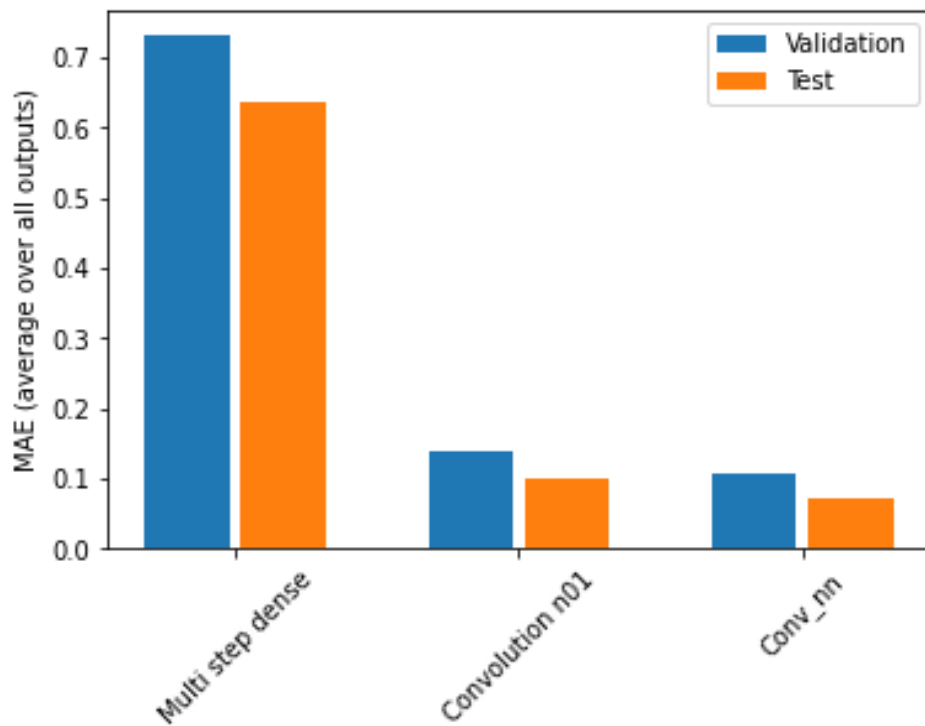
Category		Factor 1 V_{1t}	Factor 2 V_{2t}	Reprojected Volatility ($e^{(V_1+V_2)}$)	
	Symmetric MAPE	0.01362	0.23404	0.01142	
<i>Ridge Regression</i>					
$(l = 0.1)$	RMSE	0.090052	0.008819	2.23003	
	MSE	0.059335	0.006763	1.44031	
	MAPE	6.613866	190.7067	6.81054	
	Theil inequality c 1 (U_1)	0.043732	22.87471	0.00265	
	Bias Proportion		0.00101	0.00265	0.00015
	Variance P		0.00602	0.07086	0.00158
	Covariance P		0.99297	0.92650	0.99827
	Theil U_2 Coefficient	0.77744	0.47533	1.04776	
	Symmetric MAPE	0.01071	0.11452	0.01288	
<i>Decision Forest</i>					
	RMSE	0.091326	0.013431	2.16911	
	MSE	0.057668	0.010480	1.28146	
	MAPE	6.611637	180.5500	5.90789	
	Theil inequality c 1 (U_1)	0.044041	53.12674	0.00260	
	Bias Proportion		0.00326	0.00076	0.00533
	Variance P		0.02742	0.69393	0.04049
	Covariance P		0.96932	0.30531	0.95417
	Theil U_2 Coefficient	0.81876	0.41527	0.97366	
	Symmetric MAPE	0.01037	0.23436	0.01151	



The Figure reports V_1 actual volatility lagged for 25 days (and 12 hours) and a CNN neural network predicted volatility.

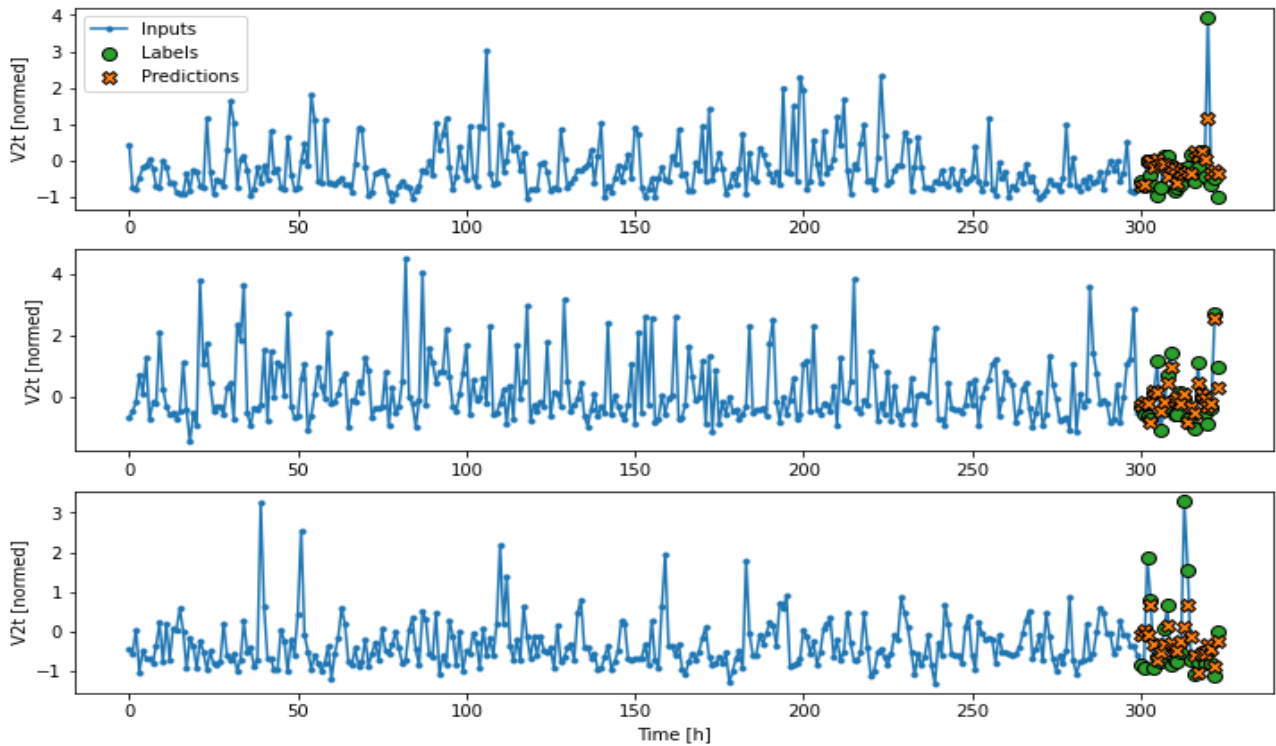


The Figure reports V_{1t} actual volatility lagged for 25 days (and 12 hours) and a RNN/LSTM neural network predicted volatility.

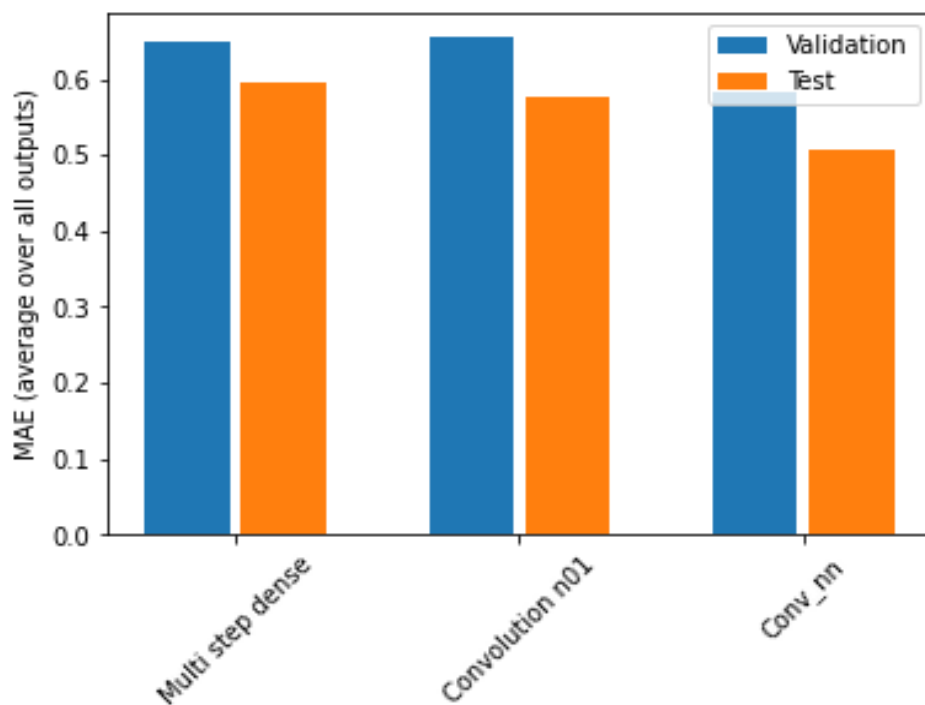


Neural network Mean Absolute Error (MAE) comparisons for V_{1t} .

Figure 12. Hourly neural networks and performance WTI oil V_{1t} .



The Figure reports V_2 actual volatility lagged for 25 days (and 12 hours) and a CNN/RNN/LSTM neural network predicted volatility.



Neural network Mean Absolute Error (MAE) comparisons for V_{2t} .

Figure 13. Hourly neural networks and performance WTI oil V_{2t} .

5. Summary and conclusions

In this paper, I have modelled and successfully estimated a multifactor stochastic model for the conditional mean and variance for the WTI oil price movements, period 2014 to 2024(3). This multifactor model aligns with the efficient market hypothesis, which posits that relevant random information, accessible to all traders, drives financial market prices, rendering prediction impossible (akin to a random walk). The applications of option pricing, risk assessment, trading, portfolio optimization, and market surveillance show that volatility is a cornerstone of modern financial theory and practice. The analysis reports success with two volatility factors: one is extremely persistent, and the other strongly implies reverting. Furthermore, the latent volatility's data dependence suggests predictability. The persistent factor is predictable; the mean reversion factor is most likely not predictable. The interesting question now is whether to trade volatility for only the predictable, persistent factor. I leave the market implementation to future research using ML and AI methods.

Even though price processes in energy markets seem stochastic and are unpredictable, I can use observed previous prices and their variations to determine the time-dependent variance of forecast errors. For derivatives time series volatility is important for market participants in actual financial derivatives markets as well as synthetically for index and assets (B&S formula). The static predictions of the projected volatility for the WTI oil contracts show a Theil's inequality coefficient close to zero and a covariance portion of 99.8% using Ridge regression (ML) with a $\lambda^2 = 0.1$ (L2). Furthermore, using continuous prediction updates (i.e., hourly) may further improve these measures of fit. Also, V_1 , the main and long-lasting factor for the expected volatility in WTI oil prices, has a Theil inequality coefficient that is very close to zero and a similar covariance portion of 99.3%. Furthermore, note that the reprojected volatility is strongly influenced of the persistent and predictable volatility factor V_1 . Market participants can apply a continuous multifactor SV model with associated volatility trading strategies to achieve superior profit from the energy markets. Hence, extending our results to AI/ML for program trading and applying volatility forecasts may turn out to be warranted.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Data availability statement

The article uses a freely downloadable dataset (40546 hourly and 2443 daily observations).

The paper encloses the hourly WTI oil datasets for the period 2014–2024(3):

1. 001-High_Frequency_WTI_FM_price_Movements_2014-2024(3).txt.
2. 001-Daily_Frequency_WTI_FM_price_Movements_2014-2024(3).txt.

Conflict of interest

The author declares no conflicts of interest in this paper.

References

- Andersen TG (1994) Stochastic autoregressive volatility: a framework for volatility modelling. *Math Financ* 4: 75–102. <https://doi.org/10.1111/j.1467-9965.1994.tb00063.x>
- Andersen TG, Benzoni L, Lund J (2002) Towards an empirical foundation for continuous-time models. *J Financ* 57: 1239–1284. <https://doi.org/10.1111/1540-6261.00460>

- Brock WA, Dechert WD, Scheinkman JA, et al. (1996) A test for independence based on the correlation dimension. *Econom Rev* 15: 197–235. <https://doi.org/10.1080/07474939608800353>
- Black F (1976) Studies of Stock Market Volatility Changes, Proceedings of the American Statistical Association, Business and Economics Section, 177–181.
- Bollerslev T (1986) Generalised Autoregressive Conditional Heteroscedasticity. *J Econom* 31: 307–27. [https://doi.org/10.1016/0304-4076\(86\)90063-1](https://doi.org/10.1016/0304-4076(86)90063-1)
- Bollerslev T (1987) A Conditionally heteroscedastic Time Series Model for Speculative Prices and Rates of Return. *Rev Econ Stat* 64: 542–547. <https://doi.org/10.2307/1925546>
- Bollerslev T, Chou RY, Kroner KF (1992) ARCH modeling in finance. A review of the theory and empirical evidence. *J Econom* 52: 5–59. [https://doi.org/10.1016/0304-4076\(92\)90064-X](https://doi.org/10.1016/0304-4076(92)90064-X)
- Box GEP, Jenkins GM (1976) *Time Series Analysis: Forecasting and Control*, Revised Edition, San Francisco: Holden Day.
- Brock WA, Deckert WD (1988) Theorems on Distinguishing Deterministic from Random Systems, In: Barnett, W.A., Berndt, E.R., White, H., *Dynamic Econometric Modelling*, Cambridge University Press, 247–268.
- Brock WA, Dechert WD, Scheinkman JA, et al. (1996) A test for independence based on the correlation dimension. *Econom Rev* 15: 197–235. <https://doi.org/10.1080/07474939608800353>
- Campbell JY, Grossman SJ, Wang J (1993) Trading Volume and Serial Correlation in Stock Returns. *Q J Econ* 108: 905–939. <https://doi.org/10.2307/2118454>
- Campbell J, Mankiw NG (1987) Are output fluctuations transitory?. *Q J Econ* 102: 875–880. <https://doi.org/10.2307/1884285>
- Chan KF, Gray P (2006) Using Extreme Value Theory to Measure Value-at-Risk for Daily Electricity Prices. *Int J Forecast* 22. <https://doi.org/10.1016/j.ijforecast.2005.11.005>
- Chernov M, Gallant AR, Ghysel E, et al. (2003) Alternative models for stock price dynamics. *J Econom* 56: 225–257. [https://doi.org/10.1016/S0304-4076\(02\)00202-5](https://doi.org/10.1016/S0304-4076(02)00202-5)
- Christie A (1982) The Stochastic Behavior of Common Stock Variances: Value, Leverage and Interest Rate Effects. *J Financ Econ* 10: 407–432. [https://doi.org/10.1016/0304-405X\(82\)90018-6](https://doi.org/10.1016/0304-405X(82)90018-6)
- Clark PK (1973) A subordinated stochastic Process model with finite variance for speculative prices. *Econometrica* 41: 135–156. <https://doi.org/10.2307/1913889>
- de Lima PJF (1995a) Nonlinearities and Nonstationarities in stock returns. Working Paper in Economics, The Johns Hopkins University, Department of Economics.
- de Lima PJF (1995b) Nuisance parameter free properties of correlation integral based statistics. *Econom Rev* 15: 237–259. <https://doi.org/10.1080/07474939608800354>
- deVany AS, Wall WD (1999) Cointegration analysis of spot electricity prices: insights on transmission efficiency in the western US. *Energy Econ* 21: 435–448. [https://doi.org/10.1016/S0140-9883\(99\)00011-0](https://doi.org/10.1016/S0140-9883(99)00011-0)
- Dickey DA, Fuller WA (1979) Distribution of the estimators for autoregressive time series with a unit root. *J Am Stat Assoc* 74: 427–431. <https://doi.org/10.2307/2286348>
- Ding Z, Engle RF, Granger CWJ (1993) A Long memory Property of Stock Market Returns and a New Model. *J Empir Financ* 1: 83–106. [https://doi.org/10.1016/0927-5398\(93\)90006-D](https://doi.org/10.1016/0927-5398(93)90006-D)
- Doan T, Litterman R, Sims C (1984) Forecasting and Conditional Projection using Realistic Prior Distributions. *Econom Rev* 3: 1–100. <https://doi.org/10.1080/07474938408800053>
- Durham G (2003) Likelihood-based specification analysis of continuous-time models of the short-term interest rate. *J Financ Econ* 70: 463–487. [https://doi.org/10.1016/S0304-405X\(03\)00105-7](https://doi.org/10.1016/S0304-405X(03)00105-7)
- Engle RF (1982) Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation. *Econometrica* 50: 987–1008. <https://doi.org/10.1080/16843703.2004.11673078>

- Engle RF, Bollerslev T (1986) Modelling the Persistence of Conditional Variances. *Econom Rev* 5: 1–50. <https://doi.org/10.1080/07474938608800095>
- Engle RF, Ng VK (1993) Measuring and Testing the Impact of News on Volatility. *J Financ* 48: 1749–1778. <https://doi.org/10.1111/j.1540-6261.1993.tb05127.x>
- Engle RF, Patton AJ (2001) What Good is a Volatility Model. Available from: <http://archive.nyu.edu/bitstream/2451/26881/2/S-DRP-01-03.pdf>.
- Gallant AR, Nychka DW (1987) Semi nonparametric maximum likelihood estimation. *Econometrica* 55: 363–390. <https://doi.org/10.2307/1913241>
- Gallant AR, Rossi PE, Tauchen G (1992) Stock Prices and Volume. *Rev Financ Stud* 5: 199–242. <https://doi.org/10.1093/rfs/5.2.199>
- Gallant AR, Rossi PE, Tauchen G (1993) Nonlinear dynamic structures. *Econometrica* 61: 871–907. <https://doi.org/10.2307/2951766>
- Gallant AR, Rossi PE, Tauchen G (1992) A nonparametric approach nonlinear time series analysis: Estimation and Simulation, In: Parzen, E., Brillinger, D., Rosenblatt, M., et al., *New dimensions in time series analysis*, New York: Springer-Verlag. https://doi.org/10.1007/978-1-4613-9296-5_5
- Gallant AR, Tauchen G (2010) Simulated Score Methods and Indirect Inference for Continuous Time Models, In: Aït-Sahalia, Y., Hansen, L.P., *Handbook of Financial Econometrics*, North Holland, 1: 427–477. <https://doi.org/10.1016/B978-0-444-50897-3.50011-0>
- Glosten L, Jagannathan, Runkle D (1993) Relationship between the expected value and the volatility of the nominal excess return on stocks. *J Financ* 48: 1779–1801. <https://doi.org/10.1111/j.1540-6261.1993.tb05128.x>
- Gouriéroux C (1997) *ARCH Models and Financial Applications*, New York: Springer.
- Hamilton JD (1994) *Time Series Analysis*, Princeton, New Jersey: Princeton University Press. <https://doi.org/10.1515/9780691218632>
- Hull JC (2021) *Machine Learning in Business: An Introduction to the Wprld of Dats Science*, GFS Press.
- Koop G (1994) Parameter uncertainty and impulse response analysis. *J Econom* 72: 135–149. [https://doi.org/10.1016/0304-4076\(94\)01717-4](https://doi.org/10.1016/0304-4076(94)01717-4)
- Koop G, Osiewalski J, Steel MFJ (1994) Bayesian long-run prediction in time series models. *J Econom* 69: 61–80. [https://doi.org/10.1016/0304-4076\(94\)01662-J](https://doi.org/10.1016/0304-4076(94)01662-J)
- Lee K, Pesaran MH (1993) Persistence profiles and business cycle fluctuations in a disaggregated model of UK output growth. *Richerche Economice* 47: 293–322. [https://doi.org/10.1016/0035-5054\(93\)90032-X](https://doi.org/10.1016/0035-5054(93)90032-X)
- Neftci S (1984) Are economic time series asymmetric over the business cycle. *J Polit Econ* 93: 307–328. <https://doi.org/10.1086/261226>
- Pesaran MH, Potter S (1994) A floor and ceiling model of U.S. output. *J Econ Dyn Control* 21: 661–695. [https://doi.org/10.1016/S0165-1889\(96\)00002-4](https://doi.org/10.1016/S0165-1889(96)00002-4)
- Pesaran MH, Shin Y (1995) Cointegration and speed of convergence to equilibrium. *J Econom* 71: 117–143. [https://doi.org/10.1016/0304-4076\(94\)01697-6](https://doi.org/10.1016/0304-4076(94)01697-6)
- Pesaran MH, Pierse R, Lee K (1993) Persistence, cointegration and aggregation: A disaggregated analysis of output fluctuations in the US Economy. *J Econom* 56: 57–88. [https://doi.org/10.1016/0304-4076\(93\)90101-A](https://doi.org/10.1016/0304-4076(93)90101-A)
- Potter S (1995) A nonlinear approach to US GNP. *J Appl Econom* 10: 109–125. <https://doi.org/10.1002/jae.3950100203>
- Potter S (1994) Nonlinear impulse response functions. Department of Economics working paper (University of California, Los Angeles, CA). <https://doi.org/10.2139/ssrn.163169>
- Ripley BD (1987) *Stochastic simulation*, New York: Wiley. <https://doi.org/10.1002/9780470316726>

- Tiao G, Tsay R (1994) Some advances in nonlinear and adaptive modeling in time series analysis. *J Forecast* 13: 109–131. <https://doi.org/10.1002/for.3980130206>
- Kwiatkowski D, Phillips CB, Schmidt P (1992) Testing the Null Hypothesis of Stationary against the Alternative of a Unit Root. *J Econom* 54: 154–179. <https://doi.org/10.1111/1467-9892.00213>
- Ljung GM, Box GEP (1978) On a Measure of Lack of Fit in Time Series Models. *Biometrika* 66: 67–72. <https://doi.org/10.1093/biomet/65.2.297>
- Mills TC (1990) *Time Series Techniques for Economists*, Cambridge University Press.
- Nelson D (1991) Conditional heteroscedasticity in asset returns; A new approach. *Econometrica* 59: 347–370. <https://doi.org/10.2307/2938260>
- Ramsey JB (1969) Tests for specification errors in classical least square regression analysis. *J R Stat Soc Ser B-Stat Methodol* 31: 350–371. <https://doi.org/10.1111/j.2517-6161.1969.tb00796.x>
- Robinson PM (1983) Nonparametric Estimators for Time Series. *J Time Ser Anal* 4: 185–207. <https://doi.org/10.1111/j.1467-9892.1983.tb00368.x>
- Rosenberg B (1972) The behavior of random variables with nonstationary variance and the distribution of security prices, unpublished paper, Research Program in Finance, University of California, Berkeley.
- Ross C (1989) Institutional Markets, Financial Marketing, and Financial Innovation. *J Financ* 44: 541–556. <https://doi.org/10.1111/j.1540-6261.1989.tb04377.x>
- Scheinkman JA (1990) Nonlinearities in Economic Dynamics. *Econ J* 100: 33–48. <https://doi.org/10.2307/2234182>
- Schwarz G (1978) Estimating the dimension of a model. *Ann Stat* 6: 461–464. <https://doi.org/10.1214/aos/1176344136>
- Shepard N (2004) *Stochastic Volatility: Selected Readings*, Oxford University Press, <https://doi.org/10.1093/oso/9780199257195.001.0001>
- Sims C (1980) Macroeconomics and Reality. *Econometrica* 48: 1–48. <https://doi.org/10.2307/1912017>
- Solibakke PB (2020) Stochastic Volatility Models Predictive Relevance for Equity Markets, In: Valenzuela, O., Rojas, F., Herrera, L.J., et al., *Theory and Applications of Time Series Analysis: Selected Contributions from ITISE 2019*, 1st ed., Springer, Cham, 125–143. https://doi.org/10.1007/978-3-030-56219-9_9
- Solibakke Per B (2022) Bootstrapped Nonlinear Impulse-Response Analysis: The FTSE100 (UK) and the NDX100 (US) Indices 2012–2021. *Int J Comput Econ Econom* 12: 197–221. <https://doi.org/10.1504/IJCEE.2021.10043332>
- Tauchen G, Pitts M (1983) The price variability volume relationship on speculative markets. *Econometrica*, 485–505. <https://doi.org/10.2307/1912002>
- Taylor S (1982) Financial returns modelled by the product of two stochastic processes—a study of daily sugar prices 1961–79, In: Anderson, O. D.(ed.), *Time Series Analysis: Theory and Practice*, Amsterdam, North-Holland, 1: 203–226. <https://doi.org/10.1093/oso/9780199257195.003.0003>
- Taylor S (2005) *Asset Price Dynamics, Volatility, and Prediction*, Princeton University Press, <https://doi.org/10.1515/9781400839254>
- Vo M (2011) Oil and Stock market volatility: A multivariate stochastic volatility perspective. *Energy Econ* 33: 956–965. <https://doi.org/10.1016/j.eneco.2011.03.005>



AIMS Press

© 2024 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0>)