Article

Oil Price Forecasting Using a Time-Varying Approach

Lu-Tao Zhao,1,2, Shun-Gang Wang1* and Zhi-Gang Zhang1,*

1 School of Mathematics and Physics, University of Science and Technology Beijing, Beijing 100083, China; ltzhao@ustb.edu.cn (L.-T.Z.); 41721633@xs.ustb.edu.cn (S.-G.W.)
2 Center for Energy and Environmental Policy Research & School of Management and Economics, Beijing Institute of Technology, Beijing 100081, China
* Correspondence: zzgcyf@263.net

Received: 31 January 2020; Accepted: 16 March 2020; Published: 17 March 2020

Abstract: The international crude oil market plays an important role in the global economy. This paper uses a variable time window and the polynomial decomposition method to define the trend term of time series and proposes a crude oil price forecasting method based on time-varying trend decomposition to describe the changes in trends over time and forecast crude oil prices. First, to characterize the time-varying characteristics of crude oil price trends, the basic concepts of post-position intervals, pre-position intervals and time-varying windows are defined. Second, a crude oil price series is decomposed with a time-varying window to determine the best fitting results. The parameter vector is used as a time-varying trend. Then, to quantitatively describe the continuation of the time-varying trend, the concept of the trend threshold is defined, and a corresponding algorithm for selecting the trend threshold is given. Finally, through the predicted trend thresholds, the historical reference data are selected, and the time-varying trend is combined to complete the crude oil price forecast. Through empirical research, it is found that the time-varying trend prediction model proposed in this paper achieves a better prediction than several common models. These results can provide suggestions and references for investors in the international crude oil market to understand the trends of oil prices and improve their investment decisions.

Keywords: Time-varying characteristic; Trend threshold; Time-varying trend decomposition model

1. Introduction

Influenced by many uncertain factors, international crude oil prices are characterized by high randomness. The traditional crude oil price forecasting model, which is based on time series analysis, extracts effective information from historical price series, decomposes those price series into trend, seasonal, periodic and random components, and makes reasonable judgements on the basis of historical trends and future trends. This model plays a key role in judging historical trends reasonably and predicting future trends accurately in investment and decision-making.

Trends in the price sequence of crude oil often depend on a specific time frame that reflects the interrelationships between the data on a time scale. Therefore, predictions of trends tend to occur over specific time intervals. In terms of trend characterization, Noguera (2011) verified the randomness of oil prices by analysing multiple structural changes in actual oil prices [1]. Zhang and Zhang (2014) used the Markov system transformation model with dynamic autoregressive coefficients to discuss the Brent and West Texas Intermediate (WTI) price patterns after the financial crisis and analysed the reasons for the abnormal fluctuations of these two benchmark crude oil prices [2]. Luo et al. (2019) established fuzzy information particles, granulated time series that were used to characterize the time series characteristics, and predicted the trends of time series such as temperature and financial data [3]. Yahyaoui et al. (2018) proposed a new trend-based symbolic approximation (SAX) reduction technique
to classify time series with different trends by using variable segment sizes of time series variation points [4]. Mello et al. (2018) proposed a new method for classifying time series with different trends by using the probability density function (PDF) and K-nearest neighbour (KNN) algorithm, indicating that there are similarities and differences among different time periods of time series [5]. Although the above research introduced a new perspective for describing time series, it remains difficult to grasp the temporal attributes of those time series, and thus the temporal characteristics inherent in the time series (for those time series with nuances) are fundamentally ignored. However, it is difficult to distinguish temporal attributes accurately.

Recognizing that the selection of the time window has an important impact on time series analysis, scholars have proposed a predictive model for variable windows. Gao et al. (2016) examined the evolution of different wave patterns over time from a complex network perspective using a sliding time window to divide the time series of crude oil prices into several parts and defined an autoregressive model based on regression to indicate the fluctuation of each segment. This model proves that there are various autoregressive modes with significantly different statistical characteristics under different periodic time series [6]. An uncertain hidden state can affect the changes in trends [7], creating some difficulty in describing those trends. In addition, there is a popular trend analysis theorem based on time series analysis that is dominated by the moving average method [8]. However, because the values predicted by this method always stay at the previous levels, higher or lower fluctuations cannot be predicted for the future [9,10]. In addition, these methods use a fixed periodic window for each point in the time series, and just because the trend of the time series changes with the length of the window, the same window size is used for all points. This assumption is not reasonable. The trend of the time series is not singular. In contrast, the trend is variable [11–14]. Although the above theories have important application value in finance, biomedicine and environmental science, the selection of the time interval, the sampling frequency of the data and sudden changes in the sample structure will affect the prediction of price trends [15].

Many linear and non-linear models have been widely used in forecasting daily oil prices [16]. For non-stationary data, many scholars often use autoregressive integrated moving average (ARIMA) as a time series prediction model to analyse and predict time series [17]. The autoregressive model (AR) is widely used in economics, informatics, and prediction of natural phenomena and achieves good prediction results [18–20]. With the application of big data technology, many machine learning and deep learning algorithms have been applied to oil price forecasting and have achieved good results. The support vector machine (SVM) is a popular model in the field of machine learning. Its advantage is that it can effectively solve nonlinear problems, and the training samples are very simple. The disadvantage is that the training time for large-scale training data samples is very long, and thus, the SVM is usually used in contrast methods [21–23]. The back propagation neural network (BPNN) is another relatively common model in deep learning. It is widely used in time series prediction. It has the advantages of a strong nonlinear mapping ability, high self-learning and self-adaptive abilities, and a certain fault tolerance capability; however, the downside is that the calculation convergence speed is slow, and it easily falls into local minima. Given the good performance of the BPNN, it is often used in oil price forecasting [24,25]. The above methods can effectively describe the nonlinearity and complexity of crude oil price series, but they still suffer from some problems, such as over-fitting, poor stability and lack of interpretability.

So far, there has been limited research on oil price forecasting using a time-varying method, and there are only some early attempts by some people [26,27]. Early researchers provided theoretical ground for our research and proved that the method of characterizing oil prices with polynomial function parameters is feasible [28]. The above studies have effectively resolved the description of price trends and the prediction of price fluctuations. However, these methods often describe the trends of time series in a single category or a few categories without taking into account the nuances of different oil price trends over time, which makes it difficult to distinguish between subtle changes in those trends. Therefore, there are two main research problems in the paper: “how can we describe oil
price trends with time-varying characteristics?” and “what are the historical price data we should choose to use?” In response to these questions, this paper proposes a time-varying trend method using variable time windows and parameter vectors in the window based on the function approximation theory [29]. The main methodologies are as follow. Firstly, the oil prices are converted into time-varying trends, and then the time-varying trends are used to find the trend threshold corresponding to each point. Secondly, due to the superiority of machine learning models, we use support vector regression (SVR) to predict the trend threshold. Thirdly, the frequency division regression prediction method is used to predict the time-varying trend of oil price based on the obtained trend threshold and restore it to the predicted value of oil price. Finally, the obtained results are compared with ARIMA and AR models in different hyper parameters, and the significant difference tests are performed on the obtained indicators. Generally, controlling oil prices in the future could provide effective help for policy makers and investors [30]. One of the contributions of this article is that the lag relationship between the time-varying trend components between Brent and WTI. By converting the prices of two oil products into a time-varying trend series, and then by referring to the time-varying trend of WTI, the time-varying trend of Brent after two months can be predicted. It is a new method for studying oil price trends, which verifies the feasibility of combining time-varying windows with parameter vectors to characterize oil prices.

This paper is divided into four parts. The first part introduces the main work of current time series trend forecasting. In the second part, the concept and algorithm of the time-varying trend decomposition model (TV-TD model) are proposed. In the third part, the time-varying trend model is used to construct the time-varying trend and perform a short-term forecast of the Brent and WTI crude oil spot prices. The TV-TD model has been tested by the data from different time periods and been evaluated by the mean absolute percentage error ratio (MAPE-ratio), mean squared prediction error ratio (MSPE-ratio) and success ratio as evaluation indices. Finally, the Diebold–Mariano test and the Pesaran–Timmermann test are performed on the predicted values to determine whether the predicted results have significant differences. In the fourth part, we draw conclusions on our model and provide some suggestions for related scholars and investors about discussing the oil prices with time-varying characteristic to help them make properly decisions.

2. Materials and Methods

Based on the above analysis, this paper uses a variable time window and the polynomial decomposition method to define the trend term and proposes a crude oil price prediction method based on a time-varying trend decomposition model (TV-TD model) to describe the change in the trend over time, and then the crude oil price is predicted. The TV-TD model is divided into three parts, as shown in Figure 1. The first part constructs a time-varying trend to quantitatively describe the trend of time series sample points. The second part constructs a trend threshold. On the one hand, the trend threshold is filtered by overshoot points, and on the other hand, support vector regression is used to perform the rolling prediction according to the obtained trend threshold sequence; then, the trend threshold of the sample point to be predicted, which can be used for the frequency division regression prediction method, is obtained. Finally, the obtained time-varying trend and trend threshold are used to perform the short-term prediction of time series sample points.

2.1. Construction of Time-Varying Trends

The rising and falling state of a time series sample point is determined by the relationship between the selected sample point and its context data. For the same sample point, selecting different time windows will show different trends. Therefore, to characterize the different trends exhibited by the differences in the selected time windows, we first give the following definitions.
2.1.1. Related Definitions of Time-Varying Trends

**Definition 1.** For a given time series \( \{Y_1, Y_2, \ldots, Y_m\} \) \( m \in \mathbb{N}^+ \), for any of the sample points \( Y_t \) \( t = 1, 2, \ldots, m \) and for particular numbers \( t_a, t_b \), the interval \( [t_a, t_b] \) that satisfies \( t_a \leq t \leq t_b \) is called the time-varying window of the sample point \( Y_t \) (Figure 2).

**Definition 2.** For the time-varying window \( [t_a, t_b] \) of the sample point \( Y_t \), let \( a_{Y_t}(t_a) = -t_a + t \); then, call \( a_{Y_t}(t_a) \) the post-position of the sample point \( Y_t \). In addition, let \( b_{Y_t}(t_b) = t_b - t \); then, call \( b_{Y_t}(t_b) \) the pre-position of the sample point \( Y_t \).

![Figure 1. Time-varying trend decomposition model framework.](image1)

![Figure 2. Time-varying window representation of time series sample points.](image2)

Correspondingly, we call the interval \( [t_a, t] \) the post-position interval of the time series sample point \( Y_t \), denoted by the symbol \( LI_t \), we call the interval \( [t, t_b] \) the pre-position interval of the time series sample point \( Y_t \), denoted by the symbol \( PI_t \), and we call \( t_b - t_a \) the length of the time-varying...
where \( t \) windows for each point, the way in which they can be described is not unique, and thus, we have a limited time variation from this point. In the time window, we find the optimal fitting function for the sub sequences containing this point and the time window corresponding to the best fit. At this time, different sets of coefficients and time window of the optimal fit function are referred to as the time-varying trend of the sample point.

On the time-varying window \([t_a, t_b]\) of the time series sample point \( Y_t \), there is a fitting function \( f(x) x \in [0, t_a + t_b] \) such that \( f(x) \) has the best fitting effect on the time-varying window. Then, the vector formed by the coefficient vector \( \hat{\theta}_t = (\hat{\theta}_{t,0}, \hat{\theta}_{t,1}, \cdots )^T \) of the coefficients of the optimal fitting function \( f(x) \) and the corresponding time window \( W_t \) is the time-varying trend of the time series sample point \( Y_t \), represented by the symbol \( T_t \). It is easy to find that \( T_t = (\hat{\theta}_t, W_t)^T = (\hat{\theta}_{t,0}, \hat{\theta}_{t,1}, \cdots, a_{Y_t}(t_{a'}), b_{Y_t}(t_{b'}))^T \), where \( t_{a'}, t_{b'} \) are the time-varying window endpoints corresponding to the optimal fit.

### 2.1.2. Time-Varying Trend Construction

As seen from the above definition, the use of different fitting functions produces different forms of time-varying trends, which forces us to search for a reliable and convenient fitting function to solve the time-varying trend. For vectors, polynomial functions have good representation capabilities [28]. So, an Algorithm 1 is proposed to solve the time-varying trend for polynomial functions.

#### Algorithm 1. Time-varying trend construction.

1. Input: Time series, search space for fitted polynomials \( S = \{f_1, f_2, \cdots, f_n\}, n \in N \), search space for time-varying window.
2. Output: Time-varying trend of time series sample point \( Y_t \).
3. Step 1: Use different degree polynomials in \( S \) to fit the sample points on the selected window \([t_a, t_b]\) to get different sets of coefficients and errors, and choose a set of coefficients with the smallest error.
4. Step 2: Traverse all windows in the search space and restore the resulting polynomial fitting parameters with time-varying windows. According to the formula \( \hat{Y}_t = \sum_{i=0}^{n} \hat{\theta}_{i,t}x_i^t \) obtain the fitting value \( \hat{Y}_t \). Select the fitting polynomial coefficients with the smallest fitting errors with the corresponding time-varying window to form the time-varying trend of sample point \( Y_t \).

### 2.2. Construction of The Trend Threshold

In the field of machine learning, the SVR is a commonly used predictive model. The SVR model has reliable theoretical support and has a high model prediction accuracy [32]. The SVR model introduces kernel techniques and can fit various forms of functions. SVR, which is mainly applied to
the regression of multivariate predictive variables, estimates the function by solving an optimization problem. It has strong advantages in numerical fitting and parameter optimization and is suitable for price time series with randomness and seasonal characteristics [22]. In this section, this paper first introduces the basic mathematical theory of SVR and then uses SVR to predict the trend threshold.

2.2.1. The Concept of SVR

Support vector regression (SVR) is a machine learning algorithm based on statistical learning theory, with utilization of the structural risk minimization (SRM) principle [33]. Given a training set \( \{(x_1, y_1), \ldots, (x_m, y_m)\} \), where \( x_i \in \mathbb{R}^n \) and \( y_i \in \mathbb{R} \) (\( i = 1, m \) is the number of data pairs). SVR maps the training data into a higher dimensional feature space then builds a linear model \( f(\mathbf{w}, \mathbf{x}) \) to predict the target vector as shown in Equation (1):

\[
f(\mathbf{w}, x_i) = \mathbf{w} \times \phi(x_i) + b
\]

where \( \mathbf{w} \) is the weights vector, \( \phi(x) \) is the non-linear higher dimension mapping of \( x \), and \( b \) is the bias term. The aim of SVR is to find \( \mathbf{w} \) and \( b \) to predict \( y \) using \( x \) as input, with \( f \) being as flat as possible. To achieve a flat function \( f \), the weight vector \( \mathbf{w} \) must be as small as possible [34].

2.2.2. Using SVR to Predict Trend Thresholds

Due to the overshoot phenomenon of a polynomial function, the blind use of a large amount of data will sharply modify the prediction accuracy. Therefore, it is necessary to determine the number of data points involved in the prediction. To determine how much historical data are needed for short-term predictions, we present a method intended to determine the order while using historical data points reasonably and efficiently and thus define the trend thresholds.

First, for the time-varying trend of all sample points, the error between the predicted and true values is calculated by traversing the lag order of the sample training set to be predicted. After the error is obtained, the starting point of the overshoot phenomenon of the error is found in all the results for each point, after which the error suddenly increases, and this point is defined as the trend threshold.

Second, after obtaining the trend threshold of the sample points, the specific regression algorithm described above is used to perform a rolling prediction on the trend threshold. The predicted trend threshold value is then used as the lag order for predicting the data points outside the sample and reducing it to the predicted oil price. For unreasonable predicted values, their trend thresholds are gradually reduced and re-screened until the predicted results fall within the pre-determined allowable range.

In this paper, the SVR model is used to predict the trend threshold sequence. The trend threshold of the previous day is used as the input, and the SVR is used to predict the trend threshold of the next day.

2.3. Oil Price Forecast

Based on the forecasting results of the time-varying trends and trend thresholds, a new method for forecasting oil prices is presented.

By using Algorithm 2, the predicted oil price can be obtained by inputting the time-varying trend and the trend threshold of the sample point to be predicted. At the same time, different models need to be used for comparison to verify the superiority of the proposed model.

To test the prediction accuracy of the TV-TD model outside the sample, we use the mean absolute percentage error ratio (MAPE-ratio), mean squared prediction error ratio (MSPE ratio) and success ratio as evaluation indices. It is standard in the literature to include measures of directional accuracy using the success ratio [18,19]. The MAPE ratio is defined as the MAPE of the model over the MAPE of the no-change forecast. The MSPE ratio is defined as the MSPE of the model over the MSPE of the no-change forecast. The success ratio is defined as the fraction of forecasts that correctly predict the
sign of the change in the price of oil. Then the Diebold–Mariano test [35] and Pesaran–Timmermann test [36] are used to determine whether forecasts are significantly different.

**Algorithm 2.** Frequency division regression prediction method.

1. Input: Time-varying trend \( T_{t-p}, T_{t-p+1}, \ldots, T_t \) of time series sample points \( Y_{t-p}, Y_{t-p+1}, \ldots, Y_t \), where \( t \geq 1 \).
2. Output: The predicted value of the time series sample point \( Y_{t+1} \).
3. Step 1: Extract each component from the input time-varying trend.

\[
\begin{bmatrix}
T_{t-p}, T_{t-p+1}, \ldots, T_t \\
W_{t-p}, W_{t-p+1}, \ldots, W_T
\end{bmatrix}
\]

after which the least squares estimate of the sequence \( \{\theta_{t-p,0}, \theta_{t-p+1,0}, \ldots, \theta_{t,0}\}, \{\theta_{t-p,1}, \theta_{t-p+1,1}, \ldots, \theta_{t,1}\}, \ldots, \{\theta_{t-p,n}, \theta_{t-p+1,n}, \ldots, \theta_{t,n}\} \) and the sequence \( \{\tilde{Y}_{t-p} + b_{Y_{t-p}}, \tilde{Y}_{t-p+1} + b_{Y_{t-p+1}}, \ldots, \tilde{Y}_{t} + b_{Y_{t}}\} \) are obtained, and the regression vector \( \{\hat{\theta}_{t+1,0}, \hat{\theta}_{t+1,1}, \ldots, \hat{\theta}_{t+1,n}\} \) is obtained as well.
4. Step 2: Use the obtained vector \( \{\hat{\theta}_{t+1,0}, \hat{\theta}_{t+1,1}, \ldots, \hat{\theta}_{t+1,n}\} \), use \( x_i = Y_{t-i} + b_{Y_{t-i}} \) to obtain the predicted value \( \hat{Y}_{t+1} \) of the sample point \( Y_{t+1} \) according to the formula \( \hat{Y}_{t+1} = \sum_{i=0}^{n} \hat{\theta}_{t+1,i} x_i \).

### 3. Empirical Analysis

#### 3.1. Data Description

The selected financial time series research subjects in this paper are Brent and WTI crude oil spot prices (daily data from 2 January 1997 to 24 February 2020). Both data samples are from the official website of the US Energy Information Administration (https://www.eia.gov/). For the case of missing data due to the lack of holiday transactions, the missing spot price data are eliminated, and the final sample size are 5869 and 5809.

The basic statistical characteristics of the sample is shown in Table 1. The mean oil prices for Brent and WTI are 58.21 and 56.17, respectively. The ranges for Brent and WTI are 134.85 and 134.49. Therefore, the average and range of oil prices between Brent and WTI are not much different. Brent has a standard deviation of 32.13 and WTI has a standard deviation of 28.50, indicating that the WTI’s oil price distribution is denser than Brent’s oil price distribution. The skewness of Brent is 0.43 indicating that the time series has a positive skew; the time series distribution is not symmetrical but skewed to the right, and the right tail is longer than the left tail. The kurtosis is 2.14; since the kurtosis of a normal distribution is 3, the distribution curve has a small kurtosis and a flat top. Considering the skewness and kurtosis, the crude oil price time series does not obey a normal distribution. It can be seen from the augmented dickey-fuller (ADF) test value that both Brent and WTI oil price time series are non-stationary sequences and generally cannot be used for direct regression. Therefore, it is necessary to find a suitable method to predict the price of crude oil. In this paper, a prediction method based on time-varying trends is proposed for crude oil price time series that can be used to predict crude oil prices.

**Table 1.** Statistical characteristics of the time series of crude oil prices.

| Oil Market | Mean   | Standard Deviation | Range   | Skewness | Kurtosis | ADF  |
|------------|--------|--------------------|---------|----------|----------|------|
| Brent      | 58.21  | 32.13              | 134.85  | 0.43     | 2.14     | -1.70|
| WTI        | 56.17  | 28.50              | 134.49  | 0.40     | 2.25     | -1.88|
3.2. Model Prediction and Inspection

3.2.1. Construction of Time-Varying Trends

The search spaces for the pre-position and post-position sets in this experiment are all consecutive positive integers from 2 to 20. For each crude oil price time series sample point, the fitted polynomial search space $S = \{f_1, f_2, f_3, f_4\}$ following Equation (2).

\[
\begin{align*}
    f_1 &= \theta_1 t + \theta_0 \\
    f_2 &= \theta_2 t^2 + \theta_1 t + \theta_0 \\
    f_3 &= \theta_3 t^3 + \theta_2 t^2 + \theta_1 t + \theta_0 \\
    f_4 &= \theta_4 t^4 + \theta_3 t^3 + \theta_2 t^2 + \theta_1 t + \theta_0
\end{align*}
\] (2)

Using Algorithm 1, the polynomial used in the fitting is first determined; then, a polynomial is selected for each of the 361 time-varying window combinations of each sample point, one of those windows with selected polynomial is selected as the time-varying trend of the sample point, and each of the trends is drawn. Line charts of the components are shown in Figure 3, and the pre-position and post-position charts are shown in Figure 4.

![Figure 3](https://example.com/figure3.png)  
(a)  

![Figure 3](https://example.com/figure3.png)  
(b)  

**Figure 3. Cont.**
Figure 3. Components of the fitted time-varying trends of Brent; (a–e) represent the first to fifth component of the time-varying trend respectively.

Figure 4. Time-varying window component statistics of time-varying trends.
Figure 3 shows that coefficients of different powers can be used to extract different degrees of oil price change trends. The higher-order terms \( \theta_3 \) and \( \theta_4 \) of the fitting polynomial in the time-varying trend describe the third-order and fourth-order terms, respectively, in the polynomial. Regarding the fluctuations in the trends, it can be considered that they are very small in most of the time period, but in the vicinity of 2009, 2010 and 2015, the value changes greatly, and over the past few years, the oil price has fallen sharply. It can be seen that our model can indeed capture the historical fluctuations in the Brent crude oil spot prices. The coefficients \( \theta_1 \) and \( \theta_2 \) describe the fluctuations in the primary and secondary terms, respectively. The values are small and relatively stable before 2004, but, after 2004, the values change to varying degrees. This shows that our model has a strong ability to capture the historical evolution of trends. It can also be seen that the constant term \( \theta_0 \) that filters out different degrees of fluctuations captures the overall trend of Brent crude oil spot prices. In Figure 4, it is found that the number of pre-position intervals of length 3 is much higher than the number of intervals of other lengths. It is assumed that the trends of most points can be extended to the next three days, and some points can continue to the next 20 days.

The fifth component of the Brent and WTI time-varying trends is displayed in Figure 5. It is interesting that there is a significant lag in the fifth component of the time-varying trend between WTI and Brent. WTI will change the trend before Brent and this time difference is about two months. If the Brent time-varying trend sequence is shifted to the left by a unit length of two months, it will almost coincide with the inflection point of the WTI time-varying trend.

![Figure 5. Lag relationship between the fifth component of the WTI and Brent time-varying trends.](image)

3.2.2. Construction and Prediction of Trend Thresholds

First, the time-varying trend of the predicted point lag from two to 50 steps is selected in turn; then, Algorithm 2 is used to predict the next step, and a series of predicted values are obtained. By comparing the predicted values with the true values, the mean square error (MSE) indicators are shown in Figure 6. According to the fixed-order method proposed in the second part, the overshoot point is selected as the trend threshold (denoted as H); the trend threshold is \( H = 8 \) in Figure 6a, and the trend threshold is \( H = 18 \) in Figure 6b. After confirming the trend thresholds of all sample point data, they are drawn with the corresponding Brent crude oil prices, as shown in Figure 7. It can be
seen from Figure 7 that at the points where the trend threshold is relatively large, the corresponding crude oil prices show little short-term fluctuation in the future and are close to a straight rise, a straight fall or a horizontal trend. In Figure 7, we draw a line which means $H = 20$ and find that the point with a higher $H$ than 20 can be considered as a turning point of oil price fluctuation, that means oil prices will move in the opposite direction after the turning point. The WTI oil prices decreased sharply in the interval from August 2018 to November 2018 with large trend threshold $H = 50$ at start point and continuous small thresholds. However, for the points where the trend threshold is relatively small, the oil price fluctuation changes very frequently.

**Figure 6.** MSEs with different sample points (a) MSE using $H = 8$; (b) MSE using $H = 18$.

**Figure 7.** Crude oil prices and trend thresholds from November 2017 to November 2019.

After determining the trend threshold sequence, the SVR model is used for a rolling prediction, and the scroll window length is set to the length of training set. Because of the existence of the
overshoot phenomenon, the predicted results often do not satisfy the rationality of the financial market; as a consequence, we set the predicted value at the next step to not exceed 10% of the current value: if the predicted value exceeds 10% of the current value, the trend threshold prediction value is gradually reduced until this condition is met. The changes in Brent and WTI’s oil prices have been counted from 2 January 1997 to 24 February 2020. The number of changes in Brent that exceed 10% is 18 (0.31% of the total) and the point exceed doesn’t occur in the past three years. WTI has 33 days (0.57% of the total) and these points only appeared twice in the past five years, so the 10% setting is reasonable.

3.2.3. Model Comparison and Evaluation

This section uses the ARIMA and AR models for a comparison with our proposed TV-TD model. The ARIMA form was determined by ADF test, and autocorrelation/partial autocorrelation function, as ARIMA. We set the rolling window length to the length of training set and use sample as shown in Table 2. The Brent and WTI oil prices are used to train and test models in different parts from January 1998 to January 2020. According to the field of machine learning, we divide the sample into a training set and a test set in a 7:3 ratio.

Table 2. Model evaluation data set information.

| Oil   | Dataset ID | Training Set          | Test Set           | Sample Capacity |
|-------|------------|-----------------------|--------------------|-----------------|
| Brent (WTI) | 1         | 1998.1–2004.12        | 2005.1–2008.1     | 2548 (2503)     |
|       | 2         | 2002.1–2009.12        | 2010.1–2013.1     | 2790 (2762)     |
|       | 3         | 2009.1–2016.12        | 2017.1–2020.1     | 2782 (2765)     |

For all inspections, the results were compared using three common indicators: MAPE ratio, MSPE ratio and success ratio. The oil price prediction of the benchmark no-change forecast model is set to the previous day’s oil price. The rolling window estimation and out-of-sample test with one-step-ahead forecasting are used for all of models. Then the Diebold–Mariano test on the first two indicators and the Pesaran–Timmermann test on the success ratio are performed. The results are conditional on above and shown in Tables 3–5.

Table 3. Model prediction results for dataset 1.

| Model/Criterion | MAPE Ratio (p-Value) | MSPE Ratio (p-Value) | Success Ratio (p-Value) |
|-----------------|----------------------|----------------------|-------------------------|
| Brent           |                      |                      |                         |
| TV-TD           | 0.9768 (0.3114)      | 1.1548 (0.0000)      | 0.5237 (0.1209)         |
| ARIMA (1,1,0)   | 0.9991 (0.9993)      | 0.9990 (0.9812)      | 0.4579 (0.9876)         |
| ARIMA (1,2,1)   | 1.1511 (0.0000)      | 1.3255 (0.0000)      | 0.4684 (0.9478)         |
| ARIMA (2,1,0)   | 0.9993 (0.9958)      | 0.9999 (0.9661)      | 0.4605 (0.9823)         |
| AR (25)         | 0.9975 (0.7254)      | 0.9975 (0.5705)      | 0.4618 (0.9762)         |
| WTI             |                      |                      |                         |
| TV-TD           | 1.0615 (0.3393)      | 1.1190 (0.0000)      | 0.5640 (0.0004)         |
| ARIMA (1,1,0)   | 0.9394 (0.0154)      | 0.8179 (0.0000)      | 0.5040 (0.3893)         |
| ARIMA (1,2,1)   | 0.9386 (0.0277)      | 0.8157 (0.0000)      | 0.5027 (0.4224)         |
| ARIMA (2,1,0)   | 0.9406 (0.0177)      | 0.8176 (0.0000)      | 0.5067 (0.3374)         |
| AR (25)         | 0.9449 (0.0295)      | 0.8154 (0.0000)      | 0.4973 (0.5192)         |

Note: All mean absolute percentage error (MAPE) and mean squared prediction error (MSPE) ratio results are relative to the benchmark no-change forecast model. The p-value of the MAPE ratio and MSPE ratio represent significant difference between the model with no-change forecast model using the Diebold–Mariano test with maximum autocovariance lag order h = 11. And we use absolute loss function for MAPE ratio and square loss function for MSPE ratio. The p-value of the success ratio represents the significance of predicting the change in direction of a time series using the Pesaran–Timmermann test.
Table 4. Model prediction results for dataset 2.

| Model/Criterion | MAPE Ratio (p-Value) | MSPE Ratio (p-Value) | Success Ratio (p-Value) |
|-----------------|----------------------|----------------------|-------------------------|
| Brent TV-TD     | 0.9057 (0.0000)      | 0.8099 (0.0000)      | 0.4806 (0.0743)         |
| ARIMA (1,1,0)   | 0.9992 (0.9866)      | 0.9998 (0.9179)      | 0.5234 (0.0909)         |
| ARIMA (1,2,1)   | 1.0004 (0.9139)      | 1.0037 (0.5964)      | 0.5154 (0.1903)         |
| ARIMA (2,1,0)   | 0.9990 (0.9394)      | 1.0006 (0.8013)      | 0.5208 (0.1156)         |
| AR (26)         | 1.0096 (0.3769)      | 1.0228 (0.0077)      | 0.5234 (0.0780)         |
| WTI TV-TD       | 0.9541 (0.0279)      | 0.8886 (0.0077)      | 0.5279 (0.0678)         |
| ARIMA (1,1,0)   | 1.0010 (0.9883)      | 1.0016 (0.8159)      | 0.4974 (0.5546)         |
| ARIMA (1,2,1)   | 1.0021 (0.9568)      | 1.0033 (0.7222)      | 0.4987 (0.5221)         |
| ARIMA (2,1,0)   | 1.0011 (0.9614)      | 1.0024 (0.6492)      | 0.4920 (0.6534)         |
| AR (25)         | 1.0049 (0.4342)      | 1.0079 (0.0048)      | 0.5013 (0.4409)         |

For the MAPE ratio, the TV-TD model has lower MAPE ratio (0.9057 in Brent and 0.9541 in WTI) for prediction than others and has significant difference (p-value less than 10%) with the no-change forecast model about Brent and WTI oil price forecasting in dataset 2. In the above three data sets, the prediction effect of the TV-TD model in the second data set is better. For the MSPE ratio, the TV-TD model has no significant difference about Brent oil price forecasting in dataset 3 and is only better than ARIMA(1,2,1) in dataset 1. But the TV-TD model has a lower MSPE ratio (0.8099 in Brent dataset 2, 0.8886 in WTI dataset 2 and 0.8154 WTI in dataset 3) and a small p-value on these sample sets. So the TV-TD model has better prediction accuracy than other models on Brent and WTI oil prices in datasets 2 and datasets 3. For the success ratio, the TV-TD model has statistical improvements in directional accuracy about WTI price change. Except for the Brent prediction of dataset 2, TV-TD has higher Success ratio than other models. It can be seen that the prediction effect of the TV-TD model for WTI is better than that of Brent. From the time scale, TV-TD has the best prediction ability in the second data set (1998–2008), and the prediction ability in the first data set (2002–2013) performs poorly in a statistical sense.

Analysing the reasons, it can be considered that the amplitude of oil prices change in the training sample of dataset 1 is not large, but the price amplitude is larger in test set. For dataset 2, the data fluctuations on both the training and test sets are significant. Overall, the TV-TD model is better than other models in WTI oil price forecasting and compared with the ARIMA and AR models, the advantages of the TV-TD model are reflected in the range of sharp fluctuations in oil prices.

Table 5. Model prediction results for dataset 3.

| Model/Criterion | MAPE Ratio (p-Value) | MSPE Ratio (p-Value) | Success Ratio (p-Value) |
|-----------------|----------------------|----------------------|-------------------------|
| Brent TV-TD     | 0.9992 (0.8725)      | 0.9427 (0.9205)      | 0.4928 (0.7119)         |
| ARIMA (1,1,0)   | 1.0026 (0.9492)      | 1.0040 (0.8764)      | 0.4928 (0.6230)         |
| ARIMA (1,2,1)   | 1.0050 (0.8783)      | 1.0077 (0.6905)      | 0.4941 (0.4843)         |
| ARIMA (2,1,0)   | 1.0023 (0.9328)      | 1.0039 (0.8433)      | 0.4980 (0.5115)         |
| AR (25)         | 1.0044 (0.8544)      | 1.0126 (0.4021)      | 0.4928 (0.6347)         |
| WTI TV-TD       | 0.8900 (0.1333)      | 0.8154 (0.0000)      | 0.5756 (0.0001)         |
| ARIMA (1,1,0)   | 0.9997 (0.9500)      | 0.9971 (0.8983)      | 0.5074 (0.2277)         |
| ARIMA (1,2,1)   | 1.0010 (0.8385)      | 1.0003 (0.8452)      | 0.5127 (0.1378)         |
| ARIMA (2,1,0)   | 1.0009 (0.9311)      | 0.9986 (0.9397)      | 0.5087 (0.2081)         |
| AR (25)         | 1.0012 (0.8761)      | 1.0061 (0.6809)      | 0.5074 (0.2611)         |
4. Conclusions

In this paper, a new perspective on the description of time series trends—time-varying trends—is introduced, and time-varying trends are constructed using coefficients of fitted polynomials and time-varying windows; then, short-term predictions of time series are presented. The TV-TD model proposed in this paper consists of three parts. The first part constructs the time-varying trends. For this part, a new method for decomposing time series is proposed. Time series are decomposed by polynomial fitting into coefficients of the fitting polynomials and the corresponding time-varying window lengths. At the same time, this study also develops a specific algorithm for finding the time-varying trends of time series sample points. The second part constructs and predicts the trend threshold. For this part, we define the trend threshold to select valid historical data points and use the SVR model to predict the trend threshold sequence. The third part predicts the time series sample points using time-varying trends and trend thresholds to predict the oil price on a future day. The proposed TV-TD model has the following contributions:

1. The introduction of time-varying trends provides a new method for researchers to analyse the trends of time series, thereby resolving the limitation of using fixed-length time windows in traditional methods; moreover, we define time-varying trends by defining the time-varying window. The trend of time series sample points is characterized by different optimal time window fitting parameters. In addition, our model has a good ability to express the historical trend of Brent and WTI oil spot prices; hence, the time series trend is more scientifically described and thus more realistic. We recommend that policymakers and investors can use WTI’s time-varying trends to predict future Brent’s time-varying trends according to the significant lag in the fifth component of the time-varying trend.

2. The presentation of the trend thresholds acknowledges that the data used for the prediction are not perfect, but historical data need to be selected according to the trend characteristics of each point. In addition, the trend thresholds are defined according to the overshoot phenomenon of polynomial functions; therefore, the predicted trend thresholds are obtained by a rolling prediction, and the historical data used for the prediction are determined in combination with a preset controlled range. If investors can correctly grasp the value brought by the trend threshold, it will help investors to accurately control the crude oil market and adjust their investment strategies in time. And the TV-TD model has a good prediction ability in a specific time period, and this ability is significantly improved.

In future research, the relationship between time-varying trends and oil prices will be explored so that investors could predict the future changes in Brent oil prices based on changes in WTI’s oil prices. On the other hand, the TV-TD model can significantly improve accuracy in areas with large fluctuation ranges. Therefore, the model could be improved to suitable for small price fluctuation ranges.

Author Contributions: L.-T.Z., S.-G.W. and Z.-G.Z. performed the research; L.-T.Z., S.-G.W. and Z.-G.Z. co-wrote the paper. Conceptualization, L.-T.Z. and S.-G.W.; Data curation, L.-T.Z. and S.-G.W.; Formal analysis, S.-G.W.; Funding acquisition, L.-T.Z.; Methodology, L.-T.Z., S.-G.W. and Z.-G.Z.; Project administration, Z.-G.Z.; Resources, L.-T.Z.; Software, S.-G.W.; Supervision, L.-T.Z. and Z.-G.Z.; Validation, L.-T.Z.; Visualization, Z.-G.Z.; Writing—original draft, L.-T.Z. and S.-G.W. All authors have read and agreed to the published version of the manuscript.

Funding: The authors gratefully acknowledge the financial support from the National Natural Science Foundation of China under Grant Nos. 71871020, 71521002.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Noguera, J. Oil prices: Breaks and trends. Energy Econ. 2013, 37, 60–67. [CrossRef]
2. Zhang, Y.; Zhang, L. Interpreting the crude oil price movements: Evidence from the Markov regime switching model. Appl. Energy 2015, 143, 96–109. [CrossRef]
3. Luo, C.; Tan, C.H.; Zheng, Y.J. Long-term prediction of time series based on stepwise linear division algorithm and time-variant zonary fuzzy information granules. *Int. J. Approx. Reason.* 2019, 108, 38–61. [CrossRef]

4. Hamdi, Y.; Reem, A. A novel trend based SAX reduction technique for time series. *Expert Syst. Appl.* 2019, 130, 113–123.

5. Mello, C.E.; Carvalho, A.S.T.; Lyra, A.; Pedreira, C.E. Time series classification via divergence measures between probability density functions. *Pattern Recogn. Lett.* 2019, 125, 42–48. [CrossRef]

6. Gao, X.; Fang, W.; An, F.; Wang, Y. Detecting method for crude oil price fluctuation mechanism under different periodic time series. *Appl. Energy* 2017, 192, 201–212. [CrossRef]

7. Zhang, M.; Jiang, X.; Fang, Z.; Zeng, Y.; Xu, K. High-order Hidden Markov Model for trend prediction in financial time series. *Physica A* 2019, 517, 1–12. [CrossRef]

8. Huang, P.; Ni, Y. Board structure and stock price informativeness in terms of moving average rules. *Q. Rev. Econ. Financ.* 2017, 63, 161–169. [CrossRef]

9. Brock, W.; Lakonishok, J.; Lebaron, B. Simple technical trading rules and the stochastic properties of stock returns. *J. Financ.* 1992, 47, 1731–1764. [CrossRef]

10. Chiarella, C.; He, X.; Hommes, C. A dynamic analysis of moving average rules. *J. Econ. Dyn. Control* 2006, 30, 1729–1753. [CrossRef]

11. Ahrens, W.A.; Sharma, V.R. Trends in natural resource commodity prices deterministic or stochastic. *J. Environ. Econ. Manag.* 1997, 33, 59–74. [CrossRef]

12. Huang, S.; An, H.; Huang, X.; Wang, Y. Do all sectors respond to oil price shocks simultaneously. *Appl. Energy* 2018, 227, 393–402. [CrossRef]

13. Wang, M.; Chen, Y.; Tian, L.; Jiang, S.; Tian, Z.; Du, R. Fluctuation behavior analysis of international crude oil and gasoline price based on complex network perspective. *Appl. Energy* 2016, 175, 109–127. [CrossRef]

14. Ghoshay, A.; Johnson, B. Trends in world energy prices. *Energy Econ.* 2010, 32, 1147–1156. [CrossRef]

15. Naser, H. Estimating and forecasting the real prices of crude oil: A data rich model using a dynamic model averaging (DMA) approach. *Energy Econ.* 2016, 56, 75–87. [CrossRef]

16. Dbouk, W.; Jamali, I. Predicting daily oil prices: Linear and non-linear models. *Res. Int. Bus. Financ.* 2018, 46, 149–165. [CrossRef]

17. Reisen, V.A.; Lopes, S. Some simulations and applications of forecasting long-memory time-series models. *J. Stat. Plan. Infer.* 1999, 80, 269–287. [CrossRef]

18. Alquist, R.; Kilian, L.; Vigfusson, R.J. Forecasting the Price of Oil. In *Handbook of Economic Forecasting*; Elsevier: Amsterdam, The Netherlands, 2013; Volume 2, pp. 427–507.

19. Baumeister, C.; Kilian, L. Real-time forecasts of the real price of oil. *J. Bus. Econ. Stat.* 2012, 30, 326–336. [CrossRef]

20. Snudden, S. Targeted growth rates for long-horizon crude oil price forecasts. *Int. J. Forecast.* 2018, 34, 1–16. [CrossRef]

21. Yi, S.; Guo, K.; Chen, Z. Forecasting China’s Service Outsourcing Development with an EMD-VAR-SVR Ensemble Method. *Procedia Comput. Sci.* 2016, 91, 392–401. [CrossRef]

22. Al-Musaylh, M.S.; Deo, R.C.; Adamowski, J.F.; Li, Y. Short-term electricity demand forecasting with MARS, SVR and ARIMA models using aggregated demand data in Queensland, Australia. *Adv. Eng. Inform.* 2018, 35, 1–16. [CrossRef]

23. Fan, L.; Pan, S.; Li, Z.; Li, H. An ICA-based support vector regression scheme for forecasting crude oil prices. *Technol. Forecast. Soc.* 2016, 112, 245–253. [CrossRef]

24. Barunik, J.; Malinská, B. Forecasting the term structure of crude oil futures prices with neural networks. *Appl. Energy* 2016, 164, 366–379. [CrossRef]

25. Chiroma, H.; Abdulkareem, S.; Herawan, T. Evolutionary Neural Network model for West Texas Intermediate crude oil price prediction. *Appl. Energy* 2015, 142, 266–273. [CrossRef]

26. Zhang, Y.J.; Wu, Y.B. The time-varying spillover effect between WTI crude oil futures returns and hedge funds. *Int. Rev. Econ. Financ.* 2019, 61, 156–169. [CrossRef]

27. Shao, Y.H.; Yang, Y.H.; Shao, H.L.; Stanley, H.E. Time-varying lead–lag structure between the crude oil spot and futures markets. *Physica A* 2019, 523, 723–733. [CrossRef]

28. Zhao, L.T.; Wang, Y.; Guo, S.Q.; Zeng, G.R. A novel method based on numerical fitting for oil price trend forecasting. *Appl. Energy* 2018, 220, 154–163. [CrossRef]
29. Ghosh, S.; Deb, A.; Sarkar, G. Taylor series approach for function approximation using ‘estimated’ higher derivatives. *Appl. Math. Comput.* 2016, 284, 89–101. [CrossRef]
30. Baumeister, C.; Kilian, L. What central bankers need to know about forecasting oil prices. *Int. Econ. Rev.* 2014, 55, 869–889. [CrossRef]
31. Zhao, L.T.; Liu, L.N.; Wang, Z.J.; He, L.Y. Forecasting Oil Price Volatility in the Era of Big Data: A Text Mining for VaR Approach. *Sustainability* 2019, 11, 3892. [CrossRef]
32. Drucker, H.; Burges, C.J.; Kaufman, L.; Smola, A.J.; Vapnik, V. Support vector regression machines. *Adv. Neural Inf. Process. Syst.* 1997, 28, 779–784. Available online: https://www.researchgate.net/publication/309185766_Support_vector_regression_machines (accessed on 16 March 2020).
33. Smola, A.J.; Schölkopf, B. A tutorial on support vector regression. *Stat. Comput.* 2004, 14, 199–222. [CrossRef]
34. Mathur, N.; Glesk, I.; Buis, A. Comparison of adaptive neuro-fuzzy inference system (ANFIS) and Gaussian processes for machine learning (GPML) algorithms for the prediction of skin temperature in lower limb prostheses. *Med. Eng. Phys.* 2016, 38, 1083–1089. [CrossRef] [PubMed]
35. Diebold, F.X.; Mariano, R.S. Comparing predictive accuracy. *J. Bus. Econ. Stat.* 1995, 13, 253–263.
36. Pesaran, M.H.; Timmermann, A. A simple nonparametric test of predictive performance. *J. Bus. Econ. Stat.* 1992, 10, 461–465.

© 2020 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 (CC BY) license (http://creativecommons.org/licenses/by/4.0/).