Forecasting oil commodity spot price in a data-rich environment

Sabri Boubaker1,2,3 · Zhenya Liu4,5,6 · Yifan Zhang4

Accepted: 19 September 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract
Statistical properties that vary with time represent a challenge for time series forecasting. This paper proposes a change point-adaptive-RNN (CP-ADARNN) framework to predict crude oil prices with high-dimensional monthly variables. We first detect the structural breaks in predictors using the change point technique, and subsequently train a prediction model based on ADARNN. Using 310 economic series as exogenous factors from 1993 to 2021 to predict the monthly return on the WTI crude oil real price, CP-ADARNN outperforms competing benchmarks by 12.5% in terms of the root mean square error and achieves a correlation of 0.706 between predicted and actual returns. Furthermore, the superiority of CP-ADARNN is robust for Brent oil price as well as during the COVID-19 pandemic. The findings of this paper provide new insights for investors and researchers in the oil market.

Keywords Change point detection · Recursive neural network · Oil price prediction · COVID-19

JEL Classification C22 · C45 · C53 · Q43

1 Introduction
Crude oil is the world’s most actively traded commodity for its strategic importance to national economies. Fluctuations in crude prices have been shown to significantly impact the global
economy in various respects, such as currencies of crude-related countries (Kayalar et al. 2017) and stock return dynamics (Chang and Yu 2013). Similar to other commodities, the spot market price of crude oil is determined, to some extent, by supply and demand. However, additional factors such as financial markets, economic growth, geopolitical uncertainty, and occasional incidents also impact crude oil price formation (Conrad et al. 2014; Baumeister and Kilian 2016; Miao et al. 2017).

Since the oil crisis initiated by the OPEC embargo in 1973, several oil shocks have resulted in the rapid rise in the price of crude oil. For example, the crude oil price rose drastically from $13 per barrel to $34 between 1979 and 1980 due to geopolitical conflict in the Middle East. The oil commodity price is sensitive to the production and supply, as well macroeconomic and financial factors (Ameur et al. 2021). However, the dynamics of time series of these factors are not easily predictable. They are typically considered to be non-stationary, which makes accurately predicting crude oil prices a challenging task.

To address this problem, we propose a deep framework of change point-adaptive recursive neural network (CP-ADARNN) to build a comprehensive forecasting model. The non-stationarity of time series dictates that the data distribution may shift over time, which could explain the inferior performance of conventional prediction models. For example, in the case of crude price, the influential determinants may change with the movement of economic conditions while the “price formulation rules” remain relatively stable. Our objective is to capture the variations within exogenous variables while keeping the conditional distribution of the dependent variable constant. Du et al. (2021) propose the ADARNN model to address this temporal covariate shift problem. To expand upon that model, we introduce the change point detection method as the first module to split the original data into periods with different distributions.

In this paper, we collect 310 monthly variables, including the macroeconomic factors of major economies, demand and supply factors of crude oil, proxy variables measuring economic uncertainty, and financial factors. The sample period covers more than 25 years, from March 1993 to December 2021. We then construct a balanced panel using these selected variables as predictors of crude oil prices. Given the high-dimensional data, we split the first 80% sample as a training set and detect three structure breaks located in April 2000, October 2007, and December 2009. Using the second module of CP-ADARNN, we train a model for one-step-ahead forecasting and the out-performance of proposed model provides new insights for participators in the oil market.

Our empirical results show the superiority of CP-ADARNN for the West Texas Intermediate (WTI) oil price prediction in terms of evaluation metrics including the mean absolute error (MAE), the root mean square error (RMSE), the out-of-sample R squared, and the information coefficient (IC). Specifically, CP-ADARNN outperforms the second-best model by 12.5% in terms of RMSE and achieves an out-of-sample R squared of 0.496 and an IC of 0.706. To check the robustness of the proposed model, we conduct the investigation similarly with the Brent spot crude oil price and design an empirical study to take account of the impact of COVID-19 pandemic. All the results confirm the better performance of CP-ADARNN compared with candidate models.

Our study contributes to the existing literature in several ways. First, we introduce the ADARNN approach in oil price prediction. This allows us to account for the non-stationarity of time series, which dictates that the distribution of data may change over time. Second, building on the original research of Du et al. (2021), we embed the change point detection into the framework of ADARNN to address the problem of distribution distance measuring. The technique is used to detect the changes in means for the high-dimensional panel with a structure of relatively strong cross-sectional dependence. Third, the CP-ADARNN outper-
forms the conventional time series and deep learning benchmarks on the prediction. Fourth, the proposed model performs well when taking into account the impacts of the COVID-19 pandemic.

The remainder of this paper is organized as follows. In Sect. 2, we review the literature for oil price forecasts and time-variant models. In Sect. 3, we propose the CP-ADARNN framework. In Sect. 4, we present the prediction performance on the WTI oil price against other models. In Sect. 5, we check the robustness of the framework for the Brent oil price and considering the impacts of COVID-19. The conclusion of this paper is provided in Sect. 6.

2 Literature review

Crude oil price prediction has drawn significant interest for decades. The first strand of literature on the topic is based on the traditional time series models, such as linear regression, cointegration analysis, and GARCH models (e.g., Ye et al. 2006; He et al. 2010; Mohammadi and Su 2010). Coppola (2018) employs a vector error correction model (VECM) to forecast oil price movements, defining the equilibrium error as the long-run spread between spot price and future price. Kilian and Murphy (2014) propose a structural vector autoregressive (SVAR) model of the global market for crude oil price forecast which includes shocks from speculative demand and flows of supply and demand. The second strand of literature is based on machine learning and continues to grow; it includes, for example, support vector machine (SVM, Xie et al. 2006), tree-based methods (Gumus and Kiran 2017), and ensemble learning (Dbouk and Jamali 2018). Neural networks (NN) are particularly popular for their ability to model complex features. Tang and Zhang (2012) build a multiple wavelet recurrent neural network (MWRNN) model to consider the effects of both trend and random components of the crude oil price. Li et al. (2019) extract hidden characteristics in news media content using a convolutional neural network (CNN) to obtain more accurate forecasts of oil prices.

However, few models above considered the time-variant dynamics in time series, particularly the distributional features. Structural breaks and regime switching approaches have received attention in studies on crude oil over the last decade (e.g., Zhang and Zhang 2015; Kuck and Schweikert 2017; Liu et al. 2020). In this paper, we introduce a change point detection technique to comprehensively capture the distribution changes in economic variables. Change point approaches have been widely used in economic fields, including the stock market (Horváth et al. 2021a), real estate (Horváth et al. 2021b), and macroeconomics (Barigozzi et al. 2018). To forecast oil prices, panel data, particularly high-dimensional panel data, are typically utilized. Common shocks like policy changes, market crashes, and other systemic events have the potential to affect all cross-sections simultaneously. Bai (2010) develops a least-squares type change point test for high-dimensional panel data while assuming the independence of time series observation on cross-section. Horváth and Hušková (2012) propose CUSUM-type tests to detect breaks in high dimensional factor models in terms of means. However, the strong dependence may negate the effectiveness of their methods. We follow Horváth et al. (2021c) in which they provide a valid finite-sample method for high-dimensional panel data with potentially strong cross-sectional dependence.
3 Methodology

In this section, we construct a time series (TS) prediction model based on the change point (CP) detection technique and adaptive recursive neural network (ADARNN) approach. Our proposed method CP-ADARNN depicts a distribution-based forecasting framework with labeled TS data. We begin with a formal formulation of the problem we aim to solve.

3.1 Time series prediction under distribution shift

Given a labeled time series of $T$ cross-sectional observations $\mathcal{D} = \{\mathbf{x}_i, y_i\}_{i=1}^{T}$, where $y_i$ is the scalar label of the $i$th cross-sectional observation $\mathbf{x}_i$, and $\mathbf{x}_i \in \mathbb{R}^p$ is a $p$-dimensional feature. $\mathcal{D}$ is supposed to be split into $K$ intervals, that is, $\mathcal{D} = \bigcup_{k=1}^{K} \mathcal{D}_k$, where $\mathcal{D}_k = \{\mathbf{x}_i, y_i\}_{i=n_k+1}^{n_k+1}$, $n_1 = 0$ and $n_{K+1} = T$. Such partition satisfies the TCS if observations in the same interval follow the identical distribution $P_{D_k}(\mathbf{x}, y)$, while for different intervals that $1 \leq i \neq j \leq K$, $P_{D_i}(\mathbf{x}) \neq P_{D_j}(\mathbf{x})$ and $P_{D_i}(y \mid \mathbf{x}) = P_{D_j}(y \mid \mathbf{x})$.

Given the labeled training set, our aim is to learn an $r$-step time series prediction model $\mathcal{M} : \mathbf{x}_i \rightarrow y_i$ to forecast with the future $r$ observations, $\mathcal{D}_{\text{test}} = \{\mathbf{x}_i\}_{i=T+1}^{T+r}$. We assume a shift in distribution occurs between training and testing sets, i.e., $P_{D_k}(\mathbf{x}) \neq P_{D_k}(\mathbf{x})$ and $P_{D_k}(y \mid \mathbf{x}) = P_{D_k}(y \mid \mathbf{x})$ for any $1 \leq k \leq K$.

A prediction framework under TCS is intuitive in finance. It is commonly accepted that the financial factors could vary from time to time with the fluctuation of the market, which means the data distributions of $\mathbf{x}$ ($P(\mathbf{x})$) are different across periods. While the economic laws or patterns $P(y \mid \mathbf{x})$ may maintain consistent. One problem remaining is the determination of the unknown number of periods $K$. In the original design of ADARNN, the authors use the method known as Temporal Distribution Characterization (TDC) to capture the diversity among the $K$ periods by maximizing the average period-wise distributional distances. However, in such way, the important parameter $K$ is required to be given exogenously.
3.2 Change point detection

We embed the CP detection technique into the framework of ADARNN to automatically determine the number of periods with different distributions. The predictor \( \{x_t\}_{t=1}^T \) under a single-valued prediction model could be considered as a balanced panel, \( x_{i,t}, 1 \leq i \leq N, 1 \leq t \leq T \) which may have the certain breaks in distribution across time. We define the break as the change in mean of the panel following Horváth et al. (2021c) where the cross-sectional dependence is allowed.

**Definition 2** (Change in mean of cross-dependent panel) The Temporal Covariate Shift (TCS) panel \( x_{i,t} \) follows the basic factor model,

\[
x_{i,t} = \mu_{i,k} + \lambda_i^T f_t + e_{i,t}, \quad c_{k-1} < t \leq c_k, 1 \leq k \leq K,
\]

where \( 1 \leq i \leq N \), the panel \( x_{i,t} \) is split into \( K \) sub-periods through \( K-1 \) change points, \( 1 < c_1 < \cdots < c_{K-1} < T \). \( \mu_{i,1}, \mu_{i,2}, \ldots, \mu_{i,K-1} \) are the means of the observations within each sub-period. Given \( d \) common factors \( f_t \in \mathbb{R}^d, 1 \leq t \leq T \), and the corresponding loadings \( \lambda_i \in \mathbb{R}^d \), the \( \lambda_i^T f_t \) presents the cross-sectional dependence. \( \{e_{i,t}, 1 \leq t \leq T\} \) are the error terms.

We are interested in detecting the change points described above, equivalently, testing the null hypothesis, there is no change in means:

\[
H_0: \mu_{i,1} = \mu_{i,2} = \cdots = \mu_{i,K-1}, \quad 1 \leq i \leq N
\]

against the alternative,

\[
H_A: \mu_{i,j} \neq \mu_{i,j+1}, \quad j \in \{1, \ldots, K-1\}, \text{ for some } i \in \{1, \ldots, N\}.
\]

We apply the well-established statistic in Horváth et al. (2021c) to detect the potential changes in means of dependent panel. Modifying the aggregated CUSUM process from Horváth and Hušková (2012), the authors propose a less restrictive form of underlying functional of

\[
V_{N,T}(u) = \sum_{i=1}^N \left( S_i^2(u) - \frac{[uT](T-[uT])}{[\tau T](T-[\tau T])} \right) \mathbb{I}[uT \geq 1]
\]

where \( S_i(u) = \sum_{t=1}^{[uT]} (x_{i,t} - T^{-1} \sum_{t=1}^T x_{i,t}) \) is the CUSUM process of the cross-sectional unit.

Based on their asymptotic results, we conclude the following theorem,

**Theorem 1** If some necessary assumptions\(^1\) hold, then we have that

\[
\frac{1}{(\bar{T} \omega_{N,T})^{1/2}} V_{N,T}(u) \xrightarrow{\mathcal{D}([0,1])} \Delta(u)
\]

under \( H_0 \) and

\[
\frac{1}{(\bar{T} \omega_{N,T})^{1/2}} \sup_{0 \leq u \leq 1} |V_{N,T}(u)| \xrightarrow{P} \infty
\]

under \( H_A \). The detector statistic is hence considered as

\[
v_{N,T} = \frac{1}{(\bar{T} \omega_{N,T})^{1/2}} \sup_{0 \leq u \leq 1} |V_{N,T}(u)|
\]

\(^1\) We refer to Horváth et al. (2021c) for the assumptions 2.1–2.8.
consequently the change point location is given as
\[ c^* = \lfloor T \hat{\theta}_{N,T} \rfloor, \quad \text{and } \hat{\theta}_{N,T} = \min\{u : |V_{N,T}(u)| = \sup_{0 \leq v \leq 1} |V_{N,T}(v)|\} \] (5)
where \( \hat{\omega}_{N,T} \) is the adjusted estimator of the normalizing sequence.\(^2\)

Based on the technique, we could detect the multiple panel change points in the predictors \( \{x_t\}_{t=1}^T \) through the standard binary segmentation (see Csörgö and Horváth 1997).

### 3.3 Time series prediction with CP-ADARNN

Given the sub-period partition using the CP technique, we cooperate with the second module of ADARNN, Temporal Distribution Matching (TDM, Du et al. 2021) to learn the common mechanism within different periods by matching the distributions. Ideally, the learned prediction model \( \mathcal{M} : x_t \rightarrow y_t \) should be more adaptive for the unobserved test data than alternative approaches that take only statistical information into consideration. We begin with formulating the loss of prediction for the parameterized model, \( \mathcal{L}_{\text{pre}} \), in a conventional way,
\[ \mathcal{L}_{\text{pre}}(\theta) = \frac{1}{K} \sum_{k=1}^{K} \frac{1}{|D_k|} \sum_{t \in D_k} \text{loss}(y_t, \mathcal{M}(x_t; \theta)) \] (6)
where \(|D_k|\) is the number of observations in the \( k \) th split period, \( \text{loss}(\cdot, \cdot) \) is a loss function, for example, the MSE loss. \( \mathcal{M}(x_t; \theta) \) denotes the parameterized prediction model.

However, such loss fails to assess the common predictive mechanism shared by different periods. We follow the framework of ADARNN in which the distributions of each pair of periods (e.g., \( D_i \) and \( D_j \)) are matching with methods commonly performed in researches on domain adaptation. We denote the \( V \) hidden states RNN with \( q \)-dimensional features with \( \mathbf{H} = \{h^V_t\}_{t=1}^V \in \mathbb{R}^{V \times q} \). To adaptively match the distributions between cells of RNN, the TDM is formally introduced as,

**Definition 3** (Temporal distribution matching) For a pair of periods \((D_i, D_j)\), the loss function of TDM is given as:
\[ \mathcal{L}_{\text{TDM}}(D_i, D_j; \theta) = \sum_{t=1}^{V} \alpha_{i,j}^t d(h^i_t, h^j_t; \theta) \] (7)
where \( \alpha_{i,j} \) represents the relative importance of the \( V \) hidden states of RNN between the periods of \( D_i \) and \( D_j \). By combining Eqs. (6) and (7), the terminal objective function of TDM with one-layer RNN is:
\[ \mathcal{L}(\theta, \alpha) = \mathcal{L}_{\text{pre}}(\theta) + \lambda \frac{2}{K(K-1)} \sum_{i \neq j} \mathcal{L}_{\text{TDM}}(D_i, D_j; \theta, \alpha) \] (8)
where \( \lambda \) is the parameter to trade off. \( \alpha \) is learned using the boosting-based information evaluation (see Du et al. 2021).

We propose the change point-ADARNN (CP-ADARNN) framework to learn a comprehensive TS prediction model. Figure 1 presents the overview of CP-ADARNN.

\(^2\) For the purpose of saving space, we leave the details of proof, see Horváth et al. (2021c).
To sum up, the proposed framework consists of two main parts. Given the training data, CP-ADARNN first uses the CP technique to split the sample into periods with different distributions. One may note that the original ADARNN applies a TDC module as the first step to capture the distribution switching within time series, but suffering the exogenous choice of one vital parameter, $K$, which can be automatically determined in CP. Then the TDM module performs the distribution matching to establish a prediction model $M$ with which we can forecast the time series with new data.

4 Empirical results

In this section, we investigate the prediction performance of the CP-ADARNN on crude oil real price and make a comparison to alternative candidate models. Utilizing real world data, we forecast the WTI crude oil monthly return on real spot price using numerous influential macroeconomic and financial factors. Using our proposed method, we first detect three structural changes in the cross-sectional means of predictors. We then conduct a forecast embedding our breaks with ADARNN. Overall, we find that our CP-ADARNN outperforms benchmark models in crude oil price prediction.

4.1 Data description

It is well-known that there are many assets concerning crude oil around the world. In this empirical study, we choose the WTI spot price as a global benchmark for crude oil. As one of the largest trading crude oil commodities, and with numerous derivatives targeting, the WTI is considered to be a well-priced and widely-investigated financial instrument (e.g., Chai et al. 2018; Ewald et al. 2019; Boubaker et al. 2021). The monthly WTI spot crude oil price sequence is obtained from the U.S. Energy Information Administration (EIA). The nominal price data are deflated using the U.S. Producer Price Index (PPI) for all commodities, which is obtained from the FRED database of the St. Louis FED.
A diverse set of macroeconomic and financial variables is used in the oil price prediction. To construct the data set of comprehensive exogenous factors, we start with the high-dimensional macroeconomic panel data, FRED-MD (McCracken 2021), which consists of 127 time series. The FRED-MD data set covers eight categories in the US, including output and income, labor market, housing, consumption, orders and inventories, money and credit, interest and exchange rates, prices, and stock market. We select 93 stock and flow time series based on previous studies (Zagaglia 2010; Naser 2016; Zhao et al. 2017). They consist of the crude oil production and stocks, oil products consumption, imports/exports, and other key information on rigs for major countries provided by the EIA. To account for global and regional risks, we include the Economic Policy Uncertainty (EPU, Baker et al. 2016) indexes for 8 countries, Geopolitical Risk (GPR, Caldara and Iacoviello 2021) indexes for 39 countries, and 2 aggregated indexes for geopolitical threats and acts. Forty-one series of financial variables containing market and commodities indexes and crude oil future prices are also included. Finally, our pool of predictors contains 310 time series that include variables concerning fundamentals, demand and supply of crude oil, macroeconomics, and financial markets. We process the 310 series based on the nature of data as input variables.3 Details of the selected variables in the empirical research are provided in the online materials including the data sources and results of the ADF tests.

There are two reasons for choosing these variables to forecast the WTI spot crude oil price. First, they are either directly related to the mechanism of price discovery for crude oil, or they have indirect impacts on the market conditions. Second, the proposed CP-ADARNN framework as a deep learning approach, inherits the powerful ability to model high-dimensional data. The various factors, despite the presence of noise, contain relatively complete information on oil price dynamics.

Our final sample period spans more than 25 years, from March 1993 to December 2021 ($T = 346$ monthly observations in total). For the purpose of modeling and forecasting, we split the data into two parts. The training samples consist of the first 80% of observations (from March 1993 to February 2016 with $T_1 = 276$ months) of predictors and WTI oil price and the remaining ($T_2 = 70$ months) data are used for testing based on Yu et al. (2014). Our proposed CP-ADARNN model as well as competing models are recursively estimated, that is, after forecasting we add an observation at the end of the training sample. To distinguish the differences between the rolling window and anchored window, we refer to Morales-Arias and Moura (2013).

4.2 Performance evaluation criterion and benchmarks

For clarification purposes, the forecasting variable in this paper is the log-return ($y_t = \log(Y_t/Y_{t-1})$) of the real oil price ($Y_t$). The price series is often observed with the properties of non-stationarity and persistence; hence, directly predicting prices is much easier than the returns, which may distort the results of the comparison. We focus on the short-term accuracy of prediction; therefore, the lag order of the predictors $r$ is set to one.

3 More specifically, the FRED-MD series are processed following McCracken (2021) in which the author provides corresponding Tcodes, and the other predictors are processed based on the nature of data. We provide the Tcodes for the entire pool of predictors in online materials.
4.2.1 Performance evaluation metrics

In order to evaluate the one-step prediction performance of the models from comprehensive respects, we select two conventional performance metrics, the mean absolute error (MAE) and the root mean square error (RMSE) which have been frequently used in previous studies (e.g., Tang et al. 2015; Lerner and Seru 2021), formulated as:

\[
\text{MAE} = \frac{1}{T_2 - 1} \sum_{t=T_1+2}^{T_1+T_2} |y_t - \hat{y}_t|
\]

\[
\text{RMSE} = \sqrt{\frac{1}{T_2 - 1} \sum_{t=T_1+2}^{T_1+T_2} (y_t - \hat{y}_t)^2}
\]

where \(T_1\) and \(T_2\) are the length of the training and testing set respectively, and \(y_t\) and \(\hat{y}_t\) are the actual and predicted returns on real oil price. Obviously, smaller MAE and RMSE indicate a better prediction model.

In addition, we use another popular evaluation metric, the out-of-sample R-squared, i.e., \(R^2_{OOS}\), which compares different forecast approaches with a benchmark model. In this paper, we use the random walk model as benchmark, which is solidly based on the Efficiency Market Hypothesis (EMH), and consequently, the \(R^2_{OOS}\) statistic is defined as:

\[
R^2_{OOS} = 100 \times \left[ 1 - \frac{\sum_{t=T_1+2}^{T_1+T_2} (y_t - \hat{y}_t)^2}{\sum_{t=T_1+2}^{T_1+T_2} (y_t - \hat{y}_{RW})^2} \right]
\]

where the \(\hat{y}_{RW}\) is the random walk predictions. The \(R^2_{OOS}\) measures the performance of candidate prediction approach relative to the trivial prediction, hence a higher and positive \(R^2_{OOS}\) indicates a better accuracy of forecasting, compared with the benchmark as \(R^2_{OOS} = 0\). For the evaluation of return prediction, the information coefficient (IC) has been widely adopted in prior studies, see for example Guerard et al. (2021). The IC describes the correlation between the predicted and realized asset returns as:

\[
\text{IC} = \text{corr}(y_t, \hat{y}_t)
\]

where \(\hat{y}_t\) is the return forecast and \(y_t\) is the actual return. The IC is calculated on the testing set. Undoubtedly, higher ICs indicate a better prediction.

4.2.2 Candidate models

We use some popular and comparative time series forecasting models in our empirical study to compare with the prediction performance of the proposed framework. They are namely (1) univariate time series models, including the random walk (RW) along with random walk with drift (RWd), and the Autoregressive Integrated Moving Average (ARIMA); (2) regularization models, including the LASSO, the Elastic net (ENet, Zou and Hastie 2005), and the Ridge regression; (3) regression tree model as Random Forest (RF); (4) neural network models, including RNN variants like the long short-term memory (LSTM), the gated recurrent unit (GRU).

Specifically, we clarify the setting of candidate models in WTI crude oil return forecasting as follows.
**Model 1–2:** For the RW models, we assume that log real price of oil follows a random walk process (or with drift), that is, the one-step ahead oil real price log-change is a martingale difference sequence with $E(y_t|\mathcal{F}_t) = 0$ (or $E(y_t|\mathcal{F}_t) = \text{drift} \times 1$).

**Model 3:** For the univariate ARIMA model, we assume the log returns follows an ARIMA($p$, $d$, $q$) process and determine the optimal lag orders $p$, $d$, $q$ according to the Bayesian Information Criterion (BIC). For all cases of modeling log returns, the order of integrated term $d$ equals zero which admits an ARMA($p$, $q$) model.

**Model 4–6:** For the regularization models, the LASSO, ENet and the Ridge regression use parameter shrinkage and variable selection to limit the dimension of regressors which solves the OLS inefficiency problem. For a shrinkage parameter $\lambda \geq 0$ and a positive combination parameter $\alpha$ between 0 (Ridge) and 1 (LASSO), the optimization problem for the regularization model is given as:

$$
\hat{\beta} = \arg \min_{\{\beta_j\}_{i=1}^{k}} \left( \frac{1}{T} \sum_{t=1}^{T} \left( y_{t+1} - \sum_{j=1}^{k} x_{j,t} \beta_j \right)^2 + \lambda P_\alpha(\beta) \right)
$$

and the penalty function $P_\alpha(\beta)$ is:

$$
P_\alpha(\beta) = \sum_{j=1}^{k} \left[ \alpha|\beta_j| + \frac{(1-\alpha)}{2} \beta_j^2 \right]
$$

where $\beta_j$ is the $j$th linear coefficients for the predictor $x_{j,t}$. The optimal shrinkage parameter $\lambda$ can be determined by cross-validation (CV). In this study, we use the 10-fold CV to compute the prediction error and compare different models using training set data. Note that higher $\lambda$ shrinks the coefficients more intensely and $\lambda = 0$ results in an OLS regression. For the weighting parameter $\alpha$, we choose 0, 0.25 and 1, representing the Ridge, ENet and LASSO regressions, respectively.

**Model 7:** For the regression tree model, we consider the RF model that constructs multiple decision trees within training set with bootstrap aggregation (bagging). We consider the number of leaves as $\{5, 10, 20, 50, 100\}$ and choose the optimal one based on the mean squared error (for regression trees) for out-of-bag observations. The number of trees is set to be 100.

**Model 8–10:** For the neural network (NN) models, we consider two widespread used RNN variants and one state-of-the-art ADARNN model. The standard GRU/LSTM is specified with two GRU/LSTM layers and two fully connected (FC) layers and with the dimension of the hidden state to be 64 for each. ADARNN could involve with RNN structure like GRU. We follow Du et al. (2021) and set the network consisting of two-layer GRUs and the dimension of the hidden state to be 64. The GRU layers are followed by two FC layers with output to be $r = 1$. We use the Adam optimizer with learning rate of 0.005 and set the batch size to be 24. To measure the difference of distributions, we exogenously set the number of domain in training set to be 4 and use the cosine distance in Du et al. (2021).

**Model 11:** Our proposed CP-ADARNN splits the training set into different sub-periods based on the changes in cross-sectional means of predictors, which is the first step of ADARNN. To demonstrate the superiority of the framework, we keep all the settings same with ADARNN but the way to identify the different distributional periods.

The CP-ADARNN allows the distribution shifting between periods split by the change points detected within predictor panel. To extend this assumption to benchmarks, we also consider the CP- versions of those non-adaptive models by replacing the entire training
Table 1  Tests for common breaks in predictors of the training sample (1993/03–2016/02)

| Periods          | Location of breaks | Relevant events                  |
|------------------|--------------------|----------------------------------|
| Mar 1993–Feb 2016| Oct 2007 (0.00)    | Early phase of financial crisis  |
| Mar 1993–Oct 2007| Apr 2000 (0.00)    | Beginning of oil price rising    |
| Nov 2007–Feb 2016| Dec 2009 (0.00)    | Ending phase of financial crisis |

Note: This table presents the detected change points in a pool of predictors. The critical values of each test are calculated following the bootstrap procedures provided by Horváth et al. (2021c). The corresponding $p$-values are shown in parentheses.

We standardize the predictors using the min-max normalization. All models are performed using MATLAB R2020a software and the ADA models are trained using Python 3.8.5, torch 1.11.0 and CUDA 11.3 on a server with a Ryzen5 2600X CPU, an NVIDIA 1080 GPU and RAM size of 16 GB.

4.3 Change points detected within the predictor panel

We perform the first module of the proposed CP-AdaRNN framework to detect the common breaks in the mean of the predictors (non-normalized) following Horváth et al. (2021c). As discussed in Subsect. 3.2, we apply the change-in-mean test based on the detection statistic $v_{N,T}$ for the high-dimensional cross-dependent panel. To accommodate potential multiple changes, a binary segmentation method is used in the fixed sample of the training set. Accordingly, we find the first and largest change in the predictors structure as October 2007 which divides the whole training sample into two parts—from March 1993 to October 2007 (176 monthly observations) and from November 2007 to February 2016 (100 monthly observations). In addition, two sub-period changes are detected in April 2000 and December 2009.

Our predictor panel covers fundamentals for oil price determination. Within the context of temporal covariate shift, structure breaks could occur in the common factors. With the relatively stable economic mechanism, measured by the conditional distribution $P(y \mid x)$, the marginal distribution of oil price may change before and after the breaks. As shown in Table 1, the first detected change point is October 2007. While it is beyond the scope of this paper to investigate the cause of breaks, that point marks the early phase of 2008 financial crisis during which the values of the macroeconomic variables vary and even reverse. Not surprisingly, the second break point found in the predictor panel is December 2009. These two change points cut the process of crisis out of the whole training sample. The third detected break locates in April 2000 which coincides with the beginning of a prolonged period of oil price increases in the early 2000s.

4.4 The superiority of CP-AdaRNN

In this section, we perform the one-step out-of-sample predictions for the proposed model as well as ten candidate models. For each prediction, we calculate four evaluation metrics

---

4 For example, the RF model is trained with observations indexed with $t = 1, \ldots, T_1$ and forecasts on $t = T_1 + 1, \ldots, T$. Suppose the last change point is detected as $t = T_c$, then the CP-RF model refers to the RF model trained with $t = T_c + 1, \ldots, T_1$ and forecasts on $t = T_1 + 1, \ldots, T$. 

---
Fig. 2 The WTI spot crude oil real price and one-step-ahead forecasts. The figure plots the monthly forecasts of CP-ADARNN and the alternative models. The shadowed areas are sub-periods split by the detected change points.

for the purpose of comparison. The actual value of the WTI spot crude oil real price and the one-step forecasts of each model are shown in Fig. 2. We compute the real price forecast for the month \( t + 1 \) as the observed price in month \( t \) multiplied by the predicted return. There are two results to be noted in Fig. 2: first, most prediction models provide reasonable forecasts in the testing sample; second, the forecasts are more accurate for milder market periods than during intense rises and falls.

Table 2 presents the prediction accuracy metrics of the different models for one-step ahead forecasting. The full-training panel contains models estimated on the full training set. For all but the CP-ADARNN model, the CP- version panel contains models estimated on the last sub-period segmented by change points.

Firstly, comparing four types of competing models—univariate time series models, regularization models, regression tree, and neural networks—it is evident that the regularization models like LASSO, ENet, and Ridge regressions are capable of forecasting individual time series in a data-rich environment. Secondly, we split the training set into sub-periods by applying the change point detection technique, assuming that the factors of oil prices are with distribution shifting. However, the time-variant distributions do not enhance the prediction performance of the models, even for the regularization models, and the CP- versions fail due to fewer observations. Thirdly, the proposed CP-ADARNN model outperforms candidate models and achieves the lowest RMSE, highest \( R^2_{OOS} \) and IC. The possible reason for this is that the CP-ADARNN model captures the stable mechanism of crude oil price formation as the determinants change. Unlike other CP- models based on consistent information, we assume that the distributions of factors could be different from all the periods in the CP-ADARNN framework. As shown in Table 2, the original ADARNN model exhibits lower prediction accuracy, which indicates the change points detected in cross-sectional means provide more appropriate partitions than the distribution distance measures in ADARNN.

Additionally, the RW model is typically used as a benchmark to be beaten in price forecasting. The EMH argues that the markets are efficient, and that asset prices reflect all available
Table 2  Prediction accuracy of oil real return forecasting models

| Models  | Full-training | CP- version |
|---------|---------------|-------------|
|         | MAE  | RMSE  | $R^2_{OOS}$ | IC  | MAE  | RMSE  | $R^2_{OOS}$ | IC  |
| RW      | 0.092 | 0.149 | –          | –   | 0.092 | 0.149 | –          | –   |
| RWd     | 0.106 | 0.182 | –0.496    | 0.253 | 0.106 | 0.149 | –0.005    | 0.273 |
| ARIMA   | 0.085 | 0.145 | 0.043     | 0.218 | 0.088 | 0.145 | 0.044     | 0.245 |
| LASSO   | 0.075 | 0.132 | 0.211     | 0.555 | 0.091 | 0.148 | 0.005     | 0.031 |
| ENet    | 0.074 | 0.126 | 0.278     | 0.574 | 0.091 | 0.148 | 0.005     | 0.016 |
| Ridge   | 0.070 | 0.120 | 0.345     | 0.597 | 0.091 | 0.148 | 0.005     | 0.029 |
| RF      | 0.075 | 0.126 | 0.280     | 0.543 | 0.090 | 0.148 | 0.008     | 0.052 |
| LSTM    | 0.108 | 0.188 | –0.607    | 0.327 | 0.132 | 0.192 | –0.664    | –0.063 |
| GRU     | 0.118 | 0.202 | –0.856    | 0.300 | 0.113 | 0.165 | –0.235    | 0.042 |
| ADARNN  | 0.087 | 0.155 | –0.085    | 0.458 | 0.072 | 0.105 | 0.496     | 0.706 |

Note: Other than the proposed CP-AdARNN framework, models with “CP-” mean that their training samples are replaced with the last sub-period split by the change points. The bold values represent the best performance among the 20 models in terms of MAE, RMSE, $R^2_{OOS}$ and IC.

information. Therefore, the best predictor of future price is the present price, which is equivalent to a zero-return forecast. We compute the $R^2_{OOS}$ based on the RW forecasts. Surprisingly, all the neutral network models have negative $R^2_{OOS}$. We conjecture that the monthly data frequency and the invariant distribution assumption over such a long horizon weaken the predictability of NN models. Specifically, the CP-ADARNN model provides a much more precise forecast which generates the lowest RMSE of 0.105 (12.5% improvement from the second-best model Ridge and 32.3% improvement from ADARNN) the highest $R^2_{OOS}$ of 0.496 and IC of 0.706.

The superiority of the CP-ADARNN model suggests the potential occasional distribution shifting in macroeconomic variables; it also indicates the importance of considering time-variant data properties for prediction. In addition, the previous study of ADARNN raised the interest in prediction under TCS, and the empirical results show that a more appropriate method to distinguish the distributional difference may enhance the capacity of the model. Our proposed framework for forecasting oil commodity price may provide helpful insights to investors and industry stakeholders.

5 Robustness test

In this section, we perform two empirical studies to verify the robustness of the CP-ADARNN model for oil price prediction. The results of the Brent oil forecast and also the COVID sample reaffirm the superiority of the proposed framework.

5.1 Forecasting the Brent spot crude oil price

To determine if the prediction performance of the CP-ADARNN framework presented in the previous section is sensitive to the choice of crude oil underlying, we replicate the procedures with the WTI spot crude oil price with the Brent spot price. As the most prevalent global benchmarks for crude oil with a similar quality, the two spot prices are highly correlated with
a correlation of 0.9919 from March 1993 to December 2021.\textsuperscript{5} However, the WTI and Brent spot oil prices began to decouple in 2011, with the spread between the real prices reaching 30.58 dollar per barrel in September 2011, as shown in Fig. 3.

As argued by Geyer-Klingeberg and Rathgeber (2021), regarding the Brent-WTI spread, the influence of paper market trading on the physical spot market heavily increased after 2010. It is natural to consider the effect of the gap between underlyings. We keep the predictors constant, therefore, the change points detected in the panel remain the same. The CP-ADARNN and alternative forecasting models are performed on the Brent monthly crude oil price. The resulting evaluation metrics are presented in Table 3. It can be seen that the proposed prediction model with the lowest RMSE, highest $R^2_{OOS}$ and IC, exhibits higher forecasting accuracy than the other candidate models.

The results of the Brent oil prediction not only illustrate the robustness of the CP-ADARNN model, they also demonstrate the economic implications of the temporal covariate shift assumption. By controlling for the predictors, the distribution shifting in macroeconomic variables remains the same for the baseline results in Sect. 4. As expected, even when replacing the target crude oil, the economic rule of oil price formation is much more stable despite the variations in market conditions over time.

5.2 Forecasting crude oil prices during COVID-19

In the baseline empirical studies, we exogenously split the whole sample into 80% for training and 20% for testing. Under TCS, it is assumed that the unconditional distributions of predictors differ for the testing set and all the sub-periods of the training set. The COVID-19 pandemic has had a significant impact on the price of oil by affecting importing and exporting (Khalfaoui et al. 2022), risk exposure (Akhtaruzzaman et al. 2021), and other factors.

\textsuperscript{5} The correlation between WTI and Brent spot crude oil real prices is 0.9990 from March 1993 to December 2010.
### Table 3: Prediction accuracy of Brent oil real return forecasting models

| Models  | Full-training | CP- version |
|---------|---------------|-------------|
|         | MAE | RMSE | $R^2_{DOS}$ | IC | MAE | RMSE | $R^2_{DOS}$ | IC |
| RW      | 0.092 | 0.146 | – | – | 0.092 | 0.146 | – | – |
| RWd     | 0.105 | 0.172 | –0.404 | 0.296 | 0.116 | 0.158 | –0.173 | 0.200 |
| ARIMA   | 0.084 | 0.137 | 0.118 | 0.348 | 0.091 | 0.146 | –0.011 | 0.041 |
| LASSO   | 0.071 | 0.113 | 0.394 | 0.648 | 0.090 | 0.145 | 0.009 | 0.061 |
| ENet    | 0.073 | 0.117 | 0.354 | 0.621 | 0.091 | 0.145 | 0.005 | 0.010 |
| Ridge   | 0.069 | 0.111 | 0.415 | 0.659 | 0.091 | 0.145 | 0.009 | 0.080 |
| RF      | 0.078 | 0.125 | 0.268 | 0.526 | 0.090 | 0.144 | 0.019 | 0.147 |
| LSTM    | 0.145 | 0.208 | –1.046 | 0.516 | 0.122 | 0.171 | –0.382 | 0.085 |
| GRU     | 0.102 | 0.169 | –0.353 | 0.382 | 0.104 | 0.148 | –0.027 | 0.301 |
| ADARNN  | 0.090 | 0.123 | 0.288 | 0.676 | 0.072 | 0.106 | 0.472 | 0.698 |

Note: Other than the proposed CP-ADARNN framework, models with “CP-” mean that their training samples are replaced with the last sub-period split by the change points. The bold values represent the best performance among the 20 models in terms of MAE, RMSE, $R^2_{DOS}$ and IC.

### Table 4: Tests for common breaks in predictors of the pre-COVID sample (1993/03—2019/12)

| Periods     | Location of breaks | Relevant events                      |
|-------------|---------------------|--------------------------------------|
| Mar 1993–Dec 2019 | Oct 2007 (0.00) | Early phase of financial crisis      |
| Mar 1993–Oct 2007 | Apr 2000 (0.00) | Beginning of oil price rising        |
| Nov 2007–Dec 2019 | Jun 2017 (0.00) | Rising of geopolitical conflicts     |

Note: This table presents the detected change points in a pool of predictors. The critical values of each test are calculated following the bootstrap procedures provided by Horváth et al. (2021c). The corresponding $p$-values are shown in parentheses.

Accordingly, we observe a bounce in oil prices during the COVID-19 pandemic. Hence, the training set for this robustness test is extended to December 2019. We choose January 2020 as the starting point of the testing set because the first officially confirmed case was reported in the end of 2019 and the Chinese government imposed a lock-down on Wuhan in January 2020.

We re-detect the change points in variables with a potential impact on oil prices as the additional observations included. The first two breaks are exactly the same as those shown in Sect. 4.3. The profound impact of the global financial crisis in 2008 results in a considerable structure break in oil price fundamentals from its early phase. The third break located in Jun 2017 marks the rise of geopolitical conflicts, such as the economic conflict between China and the United States. We present the results of the forecast evaluation in Table 5. Regardless of the WTI or Brent crude oil price, the CP-ADARNN model generally outperforms the other models on one-step-ahead return forecasts.

The exogenous shock of the COVID-19 outbreak has taken a toll on the financial markets and on macroeconomic conditions. As a result, such a shock provides a powerful structure break of oil predictors. As shown in Table 5, for regression models like LASSO, ENet, and Ridge to provide accurate forecasts, a necessary assumption is that the distribution of regressors remain constant between the training and testing sets. Hence, from the full-training to the CP- version of the regression models, the differences of predictor distribution are much...
## Table 5 Prediction accuracy of oil price forecasting models during the COVID-19 pandemic

| Models         | Full-training | CP- version |
|----------------|---------------|-------------|
|                | MAE    | RMSE   | $R^2_{DOS}$ | IC      | MAE    | RMSE   | $R^2_{DOS}$ | IC      |
| Panel A: WTI spot crude oil price |          |         |             |         |          |         |             |         |
| RW             | 0.154  | 0.234  | –           | –       | 0.154  | 0.234  | –           | –       |
| RWd            | 0.153  | 0.281  | –0.432      | 0.277   | 0.159  | 0.230  | 0.035       | 0.199   |
| ARIMA          | 0.136  | 0.229  | 0.047       | 0.231   | 0.158  | 0.236  | –0.010      | 0.059   |
| LASSO          | 0.129  | 0.209  | 0.208       | 0.583   | 0.131  | 0.201  | 0.267       | 0.515   |
| ENet           | 0.120  | 0.201  | 0.265       | 0.574   | 0.138  | 0.199  | 0.277       | 0.759   |
| Ridge          | 0.115  | 0.190  | 0.344       | 0.610   | 0.142  | 0.213  | 0.173       | 0.491   |
| RF             | 0.131  | 0.212  | 0.185       | 0.454   | 0.155  | 0.232  | 0.020       | 0.305   |
| LSTM           | 0.148  | 0.222  | 0.105       | 0.523   | 0.179  | 0.249  | –0.128      | 0.011   |
| GRU            | 0.108  | 0.149  | **0.597**   | 0.821   | 0.153  | 0.209  | 0.207       | 0.500   |
| ADARNN         | 0.109  | 0.166  | 0.499       | 0.713   | **0.094** | 0.153  | **0.573**   | **0.765** |
| Panel B: Brent spot crude oil price |          |         |             |         |          |         |             |         |
| RW             | 0.152  | 0.227  | –           | –       | 0.152  | 0.227  | –           | –       |
| RWd            | 0.145  | 0.261  | –0.321      | 0.335   | 0.157  | 0.226  | 0.010       | 0.161   |
| ARIMA          | 0.132  | 0.211  | 0.141       | 0.398   | 0.153  | 0.231  | –0.029      | 0.025   |
| LASSO          | 0.105  | 0.169  | 0.444       | 0.683   | 0.130  | 0.189  | 0.305       | 0.642   |
| ENet           | 0.107  | 0.176  | 0.402       | 0.658   | 0.130  | 0.193  | 0.278       | 0.548   |
| Ridge          | **0.105** | 0.168  | 0.452       | 0.708   | 0.140  | 0.211  | 0.140       | 0.392   |
| RF             | 0.114  | 0.182  | 0.356       | 0.666   | 0.150  | 0.224  | 0.025       | 0.344   |
| LSTM           | 0.146  | 0.228  | –0.006      | 0.519   | 0.156  | 0.197  | 0.249       | 0.675   |
| GRU            | 0.136  | 0.221  | 0.051       | 0.435   | 0.157  | 0.226  | 0.015       | 0.194   |
| ADARNN         | 0.155  | 0.223  | 0.033       | 0.513   | 0.117  | **0.160** | **0.502**   | **0.713** |

Note: Other than the proposed CP-ADARNN framework, models with “CP-” mean that their training samples are replaced with the last sub-period split by the change points. The bold values represent the best performance among the 20 models in terms of MAE, RMSE, $R^2_{DOS}$ and IC.

More significant, resulting in a worse performance. However, the CP-ADARNN framework adaptively learns the distribution shifting and allows the constant conditional distribution of oil returns given predictors, allowing for more precise forecasts on oil real returns.

### 6 Conclusion

This paper proposes a deep learning framework to train an adaptive prediction model for time series forecasting in the crude oil markets. CP-ADARNN consists of two modules: change point detection to characterize the distributional information in TS and temporal distribution matching to build a generalized RNN model. We apply the framework to the WTI spot oil price market from March 1993 to December 2021 and utilize the first 80% of the data as the training set. Our forecasting results on the testing sample show that the proposed CP-ADARNN model outperforms benchmarks including conventional time series models (i.e., RW, RWd, and ARIMA), regularization models (i.e., LASSO, ENet, and Ridge regressions), regression tree model like RF, and neural network models (i.e., GRU, LSTM, ADARNN).
in terms of MAE, RMSE, $R^2_{OOS}$, and IC. We also confirm the prediction accuracy of CP-ADARNN on the Brent oil price. In addition, for the time period inclusive of COVID-19, our model exhibits better performance when accounting for the distribution shifting of factors after the outbreak of the pandemic.

Our finding that distribution shifting should be addressed in time series analysis has important implications for investors, scholars, and other stakeholders. In the future, this research can be extended to include the various definitions of distributional change points. The CP-ADARNN framework provides a novel view of an important modeling problem for financial time series, that is, the consistent assumption on the process. This study can be used for both academic research and market participants for oil price forecasts, and it can be easily extended to other macroeconomic forecasting fields.

**Funding** Yifan Zhang was supported by the Outstanding Innovative Talents Cultivation Funded Programs 2021 of Renmin University of China.

**References**

Akhtaruzzaman, M., Boubaker, S., Chiah, M., & Zhong, A. (2021). Covid-19 and oil price risk exposure. *Finance Research Letters*, 42, 101882.

Ameur, H.B., Fititi, Z., and Louhichi, W., (2021). Revisiting the relationship between spot and futures markets: Evidence from commodity markets and nardl framework. *Annals of Operations Research*: 1–19.

Bai, J. (2010). Common breaks in means and variances for panel data. *Journal of Econometrics*, 157(1), 78–92.

Baker, S. R., Bloom, N., & Davis, S. J. (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics*, 131(4), 1593–1636.

Barigozzi, M., Cho, H., & Fryzlewicz, P. (2018). Simultaneous multiple change-point and factor analysis for high-dimensional time series. *Journal of Econometrics*, 206(1), 187–225.

Baumeister, C., & Kilian, L. (2016). Forty years of oil price fluctuations: Why the price of oil may still surprise us. *Journal of Economic Perspectives*, 30(1), 139–60.

Boubaker, S., Liu, Z., and Zhan, Y., (2021). Risk management for crude oil futures: An optimal stopping-timing approach. *Annals of Operations Research*: 1–19.

Caldara, D. and Iacoviello, M., (2021) Measuring geopolitical risk. *Board of Governors of the Federal Reserve*.

Chai, J., Xing, L. M., Zhou, X. Y., Zhang, Z. G., & Li, J. X. (2018). Forecasting the wti crude oil price by a hybrid-refined method. *Energy Economics*, 71, 114–127.

Chang, K. L., & Yu, S. T. (2013). Does crude oil price play an important role in explaining stock return behavior? *Energy Economics*, 39, 159–168.

Conrad, C., Loch, K., & Rittler, D. (2014). On the macroeconomic determinants of long-term volatilities and correlations in us stock and crude oil markets. *Journal of Empirical Finance*, 29, 26–40.

Coppola, A. (2018). Forecasting oil price movements: Exploiting the information in the futures market. *Journal of Futures Markets*, 28(1), 34–56.

Csörgő, M. and Horváth, L., (1997). Limit theorems in change-point analysis.

Dbouk, W., & Jamali, I. (2018). Predicting daily oil prices: Linear and non-linear models. *Research in International Business and Finance*, 46, 149–165.

Du, Y., Wang, J., Feng, W., Pan, S., Qin, T., Xu, R., and Wang, C., (2021). Adarnn: Adaptive learning and forecasting of time series. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*, pp. 402–411.

Ewald, C. O., Zhang, A., & Zong, Z. (2019). On the calibration of the Schwartz two-factor model to wti crude oil options and the extended Kalman filter. *Annals of Operations Research*, 282(1), 119–130.

Geyer-Klingeberg, J., & Rathgeber, A. W. (2021). Determinants of the wti-brent price spread revisited. *Journal of Futures Markets*, 41(5), 736–757.

Guerard, J. B., Xu, G., & Markowitz, H. (2021). A further analysis of robust regression modeling and data mining corrections testing in global stocks. *Annals of Operations Research*, 303(1), 175–195.

Gumus, M. and Kiran, M.S., (2017). Crude oil price forecasting using xgboost. In *2017 International Conference on Computer Science and Engineering (UBMK)*, pp. 1100–1103. IEEE.

He, Y., Wang, S., & Lai, K. K. (2010). Global economic activity and crude oil prices: A cointegration analysis. *Energy Economics*, 32(4), 868–876.
Horváth, L., & Hušková, M. (2012). Change-point detection in panel data. *Journal of Time Series Analysis, 33*(4), 631–648.

Horváth, L., Li, H., and Liu, Z., (2021). How to identify the different phases of stock market bubbles statistically? *Finance Research Letters*, 102366.

Horváth, L., Liu, Z., and Lu, S., (2021). Sequential monitoring of changes in dynamic linear models, applied to the us housing market. *Econometric Theory*, 1–64.

Horváth, L., Liu, Z., Rice, G., and Zhao, Y., (2021). Detecting common breaks in the means of high dimensional cross-dependent panels. *The Econometrics Journal*.

Kayalar, D. E., Küçüközmen, C. C., & Selçuk-Kestel, A. S. (2017). The impact of crude oil prices on financial market indicators: Copula approach. *Energy Economics, 61*, 162–173.

Khalfaoui, R., Solarin, S.A., Al-Qadasi, A., and Ben Jabeur, S., (2022). Dynamic causality interplay from covid-19 pandemic to oil price, stock market, and economic policy uncertainty: Evidence from oil-importing and oil-exporting countries. *Annals of Operations Research*, 1–39.

Kilian, L., & Murphy, D. P. (2014). The role of inventories and speculative trading in the global market for crude oil. *Journal of Applied Econometrics, 29*(3), 454–478.

Kuck, K., & Schweikert, K. (2017). A Markov regime-switching model of crude oil market integration. *Journal of Commodity Markets, 6*, 16–31.

Lerner, J. and Seru, A., (2021). The use and misuse of patent data: Issues for finance and beyond. *The Review of Financial Studies*.

Li, X., Shang, W., & Wang, S. (2019). Text-based crude oil price forecasting: A deep learning approach. *International Journal of Forecasting, 35*(4), 1548–1560.

Liu, Y., Sun, H., Zhang, J., & Taghizadeh-Hesary, F. (2020). Detection of volatility regime-switching for crude oil price modeling and forecasting. *Resources Policy, 69*, 101669.

McCracken, M. (2021). Fred-md and fred-qd: monthly and quarterly databases for macroeconomic research.

Miao, H., Ramchander, S., Wang, T., & Yang, D. (2017). Influential factors in crude oil price forecasting. *Energy Economics, 68*, 77–88.

Mohammadi, H., & Su, L. (2010). International evidence on crude oil price dynamics: Applications of Arima-Garch models. *Energy Economics, 32*(5), 1001–1008.

Morales-Arias, L., & Moura, G. V. (2013). Adaptive forecasting of exchange rates with panel data. *International Journal of Forecasting, 29*(3), 493–509.

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

Tang, M., & Zhang, J (2012). A multiple adaptive wavelet recurrent neural network model to analyze crude oil prices. *Journal of Economics and Business, 64*(4), 275–286.

Tang, L., Wang, S., He, K., & Wang, S. (2015). A novel mode-characteristic-based decomposition ensemble model for nuclear energy consumption forecasting. *Annals of Operations Research, 234*(1), 111–132.

Xie, W., Yu, L., Xu, S., and Wang, S., (2006) A new method for crude oil price forecasting based on support vector machines. In *International Conference on Computational Science*, pp. 444–451. Springer.

Ye, M., Zyren, J., & Shore, J. (2006). Forecasting short-run crude oil price using high-and low-inventory variables. *Energy Policy, 34*(17), 2736–2743.

Yu, L., Zhao, Y., & Tang, L. (2014). A compressed sensing based ai learning paradigm for crude oil price forecasting. *Energy Economics, 46*, 236–245.

Zagaglia, P. (2010). Macroeconomic factors and oil futures prices: A data-rich model. *Energy Economics, 32*(2), 409–417.

Zhang, Y. J., & Zhang, L. (2015). Interpreting the crude oil price movements: Evidence from the Markov regime switching model. *Applied Energy, 143*, 96–109.

Zhao, Y., Li, J., & Yu, L. (2017). A deep learning ensemble approach for crude oil price forecasting. *Energy Economics, 66*, 9–16.

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 67*(2), 301–320.

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.