Energy Systems
https://doi.org/10.1007/s12667-022-00515-6

ORIGINAL PAPER

Deep learning based regime-switching models of energy commodity prices

Carlo Mari[1] · **Emiliano Mari**[2]

Received: 29 September 2021 / Accepted: 23 April 2022
© The Author(s) 2022

Abstract
We discuss a deep learning based approach to model the complex dynamics of commodity prices observed in real markets. A regime-switching model is proposed to describe the time evolution of market prices. In this model, the base regime is described by a mean-reverting diffusion process and the second regime is driven by the predictions of a deep neural network trained on the market log-returns time series. A statistical technique, based on the method of simulated moments, is proposed to estimate the model on market data. We applied this methodology to energy commodity price time series with very different characteristics, namely the US wholesale electricity, natural gas and crude oil price daily time series. The obtained results show a good agreement with empirical data. In particular, the model seems to reproduce in a very interesting way the first four central moments of the empirical distributions of log-returns as well as the shape of the observed price time series.

Keywords Commodity prices · Deep learning · Deep neural networks · Regime-switching · Diffusion dynamics · Mean-reversion

JEL Classification G31 · G32 · G33 · M21 · Q40

1 Introduction

Starting from the seminal paper by Hamilton [1], regime-switching models have become very popular in economics and finance especially for describing the non-linear behavior of commodity prices observed in real markets. Very often market

✉ Carlo Mari
carlo.mari@unich.it

Emiliano Mari
emiliano.mari@sydus.it

[1] Department of Economics, University of Chieti-Pescara, Chieti, Italy

[2] Sydus, Italy

Published online: 10 May 2022

prices experience a highly erratic dynamics in which periods of stable prices and periods of turbulent prices can be observed. Regime-switching models, allowing the model parameters to change over time according to an underlying state process as, for example, a finite-state hidden Markov chain [2], are good candidates for describing such a random behavior. Regime-switching hidden Markov models have been widely used in the financial literature to describe the price dynamics of electricity [3–7] and other energy commodities such as natural gas [8, 9] and crude oil [10, 11], as well as to describe the price dynamics of agricultural commodities [12, 13] or to model financial assets price dynamics [14] and interest rates fluctuations [15–17].

Commodity markets prices show a variable and unpredictable behavior, very often characterized by high volatility, jumps and pronounced spikes caused by shocks in the supply-demand balance [18]. This is particularly true in energy commodity markets such as electricity and natural gas markets [19]. Moreover, commodity market prices show a strong mean-reversion component which is responsible for reducing prices after a jump or a spike has occurred [18, 19]. As a consequence, log-returns are characterized by non-normal empirical distributions with high values of the volatility as well as non-zero skewness and large values of the kurtosis. Accurately modeling commodity prices is a fundamental task for implementing risk hedging strategies [20] as well as for pricing commodity derivatives [18, 21]. From this point of view, it is really important that a given model captures the first four central moments of the empirical distribution of log-returns [18, 22]. In addition to the standard deviation of the empirical distribution of log-returns, the model must be able to reproduce well the skewness and the kurtosis. In fact, the skewness accounts for the properties of upward versus downward moves, and the kurtosis accounts for extreme events that may be particularly relevant in the case of commodity markets [18]. Regime-switching models are flexible models that allow us to combine in one model both the dynamics of stable periods and the dynamics of turbulent periods, introducing various mean-reversion rates and volatilities depending on the state of the system. For these reasons, they can be considered good candidates for describing the complex price dynamics observed in real markets. However, the large number of parameters involved makes it necessary to pay great attention to the calibration of regime-switching models on market data [5].

Starting from these considerations, we proposed a new methodology to investigate the dynamics of commodity prices observed in real markets in a proper and accurate way. The method is based on Artificial Intelligence (AI) techniques, in the specific deep learning (DL) techniques, that are used to build a mean-reverting regime-switching model for describing the price dynamics observed in real markets. The proposed methodology is articulated in two steps. In the first step, after removing possible trends and seasonalities from the observed time series of log-prices, a deep neural network (DNN) is trained on the detrended log-return time series. Then, in the second step the DNN predictions of log-returns are used to model the commodity price dynamics. Indeed, we propose a regime-switching AI based hybrid model, hereinafter hybrid model, in which the base regime is described by a mean-reverting diffusion process and the second regime is driven by the predictions of the DNN. The aim is to include in the dynamics a pattern recognition mechanism for log-returns. In our hybrid model, the switching between regimes is not governed by

an underlying hidden Markov process but it is driven by the DNN prediction probability: whenever the DNN prediction probability of a log-return for the next time step is lower than a given probability threshold, the dynamic is described by the base regime otherwise the dynamic is prediction driven.

The use of a DNN in our model is aimed at improving the reliability of the log-return pattern recognition. We employed a DNN with a multi-layer structure. DNNs with a single hidden layer are widely used for modeling and forecasting time series [23], however many hidden layers are required in order to capture non-linear relationships existing between variables [24]. In this regard, Recurrent Neural Networks (RNNs), such as Long Short-Term Memory (LSTM) [25], are particularly suitable to capture long-term temporal dependencies and variable-length observations [26] thus showing strong prediction performance. RNNs are used in in many applications with time series forecasting [26], also in the case of economic and financial time series [27, 28]. In our model, we employed a DNN with a RNN multi-layer architecture in which three stacked hidden LSTM layers were included. LSTM based architectures are specially designed to account for both spatial and temporal variations in order to capture nonlinear dynamics observed in time series [29]. The presence of three stacked LSTM layers allows us to better describe the complex data correlations with the aim to improve the DNN performance [29]. In our specific, the DNN architecture is composed of eight layers, arranged in the following sequence: an input layer, a layer of initial embedding, a one-dimensional convolutional layer, one max pooling layer, three LSTM layers and, finally, one dense layer as output layer. In the following section, the choice of this DNN architecture, the reasoning behind it and the specific task of each layer will be discussed in depth.

The hybrid model proposed in this paper is a parsimonious model characterized by two parameters, namely the mean-reversion coefficient and the volatility parameter of the base regime. Such parameters can be easily calibrated on market data. In this paper, we also provided a statistical method for estimating the model on market data. We showed that the calibration can be performed through a two-step procedure. In the first step, since the mean-reversion component must force back prices to fluctuate around a long-run mean after a jump or a spike has occurred, the mean-reversion parameter is estimated by linear regression on the time series of detrended log-prices. In the second step, the estimation of the volatility parameter is performed by using the method of simulated moments [30, 31] with Monte Carlo techniques [30, 32]. The aim is to obtain a good fit of the first four moments of the empirical distributions of log-returns. In this way, the hybrid model we proposed can be easily calibrated on market data.

Three are, therefore, the main differences between the AI based dynamic model we proposed in this paper and previous research: (i) the structure of the hybrid model in which the dynamics of the second regime is driven by the prediction of a well defined DNN; (ii) the switching mechanism between regimes that is not governed by an underlying hidden Markov process but it is driven by the DNN probability of forecasting future log-returns; (iii) the reduced number of parameters to be calibrated on market data.

We performed the empirical analysis on energy commodity price time series with very different characteristics, namely on US electricity price time series observed at

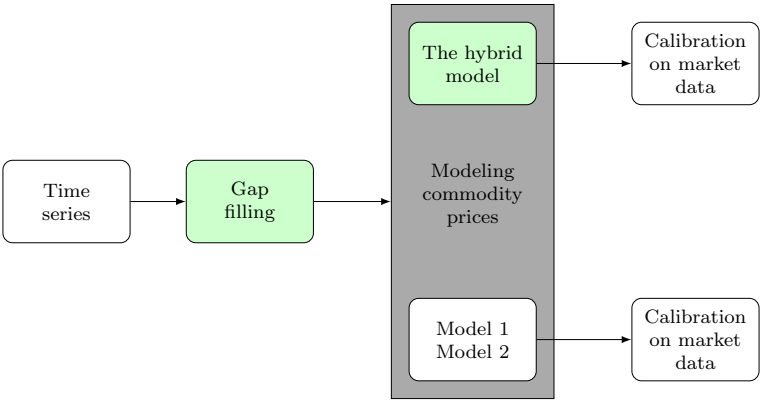

Fig. 1 Block diagram of the whole methodology workflow. The green blocks show the tasks in which the DNN is involved

PJM, on US natural gas spot price time series observed at Henry Hub, and on WTI (West Texas Intermediate) crude oil price time series. Our data set consists of time series at a daily frequency since 1 January 2015 until 31 December 2019. All the observed time series show irregular sampling (lack of daily data points) as a result of weekends, holidays and other missing data, and we decided to perform, first, a data preprocessing by using the prediction of the DNN in order to fill in the gaps in the time series. Although the observed time series exhibit very different behaviors, we will show that the proposed model is flexible enough to reproduce in a very interesting way the first four central moments of the empirical distributions of log-returns as well as the shape of the observed market price time series as a consequence of the pattern recognition mechanism included into the system dynamics.

An empirical comparison with more traditional regime-switching hidden Markov models is provided in this paper. Two alternative models are considered. In both models, the base regime is described by a mean-reverting diffusion process, and the second regime is described by a mean-reverting diffusion process in the first model, called Model 1, and by a mean-reverting jump-diffusion process in the second model, called Model 2. The switching between regime is governed by a hidden Markov process. In both cases, the dynamics is characterized by several parameters that are to be calibrated on market data. We performed the model estimation by maximum likelihood using the Hamilton filtering technique [1] and we showed that our model outperforms both these models in reproducing the first four central moments of the empirical distributions of log-returns in each market under consideration.

The whole methodology workflow is depicted in Fig. 1.

The paper is organized as follows. Section 2 illustrates in detail the proposed methodology. Section 3 is devoted to the empirical analysis. In the same section, estimation results obtained in the commodity markets under investigation are illustrated and discussed. The empirical comparison of our hybrid model with more traditional regime-switching hidden Markov models is provided and discussed in Sect. 4. Section 5 concludes.

2 The model

The main purpose of this study is to discuss a DL based methodology to model the dynamics of commodity market prices starting from the observation of historical prices. To perform this task, a multi-layer DNN with a specific architecture was created with the aim of including DNN predictions in a well defined regime-switching model of market prices. The methodology we proposed can be used also for time series with missing data. In such a case, a filling gap procedure must be used. However, the DNN we built can be employed to perform also this task. We will show that the proposed regime-switching model is flexible enough to reproduce in a very interesting way the first four central moments of the empirical distributions of log-returns as well as the shape of the observed time series of market prices. For these reasons, it can be used for numerous financial applications ranging from hedging financial risk, also through the use of commodity derivatives [18, 20], to the evaluation of long-term investments in energy generation technologies [33].

2.1 The DNN architecture

In this study, a DNN based discrete classification approach was adopted. The reason is that market prices and log-returns are expressed with a finite number of decimal places, and we required that DNN predictions must be characterized by the same number format used, respectively, to represent market prices and log-returns. To implement this approach, we must first define the appropriate range for the classification. In our specific, all the values of the time series were used to define the classification range, from the minimum value to the maximum value. Then, a different class is associated with each value belonging to the classification range rounded to a well defined number of decimal places (two decimal places for prices and four decimal places for log-returns). The DNN architecture we used is very similar to that proposed by Mari and Mari [34] and is showed in Fig. 2. Let us, therefore, briefly summarize its main characteristics. Since the DNN must be capable of making predictions on a given set of observations taking into account the ordering and the nonlinear relationships between them, a multi-layer structure is employed. More specifically, The DNN architecture is composed of eight layers arranged in the following sequence: an input layer, a layer of initial embedding, a one-dimensional convolutional layer, one max pooling layer, three LSTM layers and one dense layer as output layer. After the first two layers, namely the input and the embedding layers, the convolutional layer and the max pooling layer are used to perform data smoothing. First, the convolutional layer allows the neural network to learn smoothing parameters, then the max pooling layer performs the downsampling of the representation generated by the convolutional layer. The subsequent three LSTM layers, stacked on top of each other, use the smoothed raw data and handle the main part of the prediction problem. The presence of three stacked LSTM layers allows us to better capture the nonlinear relationships observed in time series thus improving the DNN prediction performance [29]. Finally, the dense layer complete the prediction task. In this

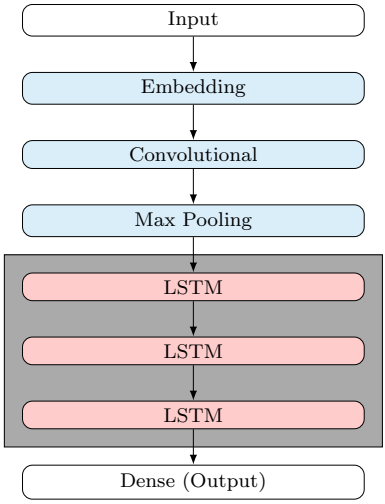

Fig. 2 The DNN architecture

Table 1 DNN main parameters

Layer	Main parameters
Input	shape=(x_train.shape[1])
Embedding	input_dim=15000, output_dim=100, input_length=N
Convolutional	filters=32, kernel_size=5, activation="relu"
MaxPooling	pool_size=2
LSTM	units=256, dropout=0.1
LSTM	units=128, dropout=0.1
LSTM	units=64, dropout=0.1
Dense	units=max(y_train) + 1, activation="softmax"

Training sets are denoted by x_train and y_train

last layer, the Softmax function is used as activation function that returns a probability distribution over the classes belonging to the classification range. Each element of the DNN training dataset is composed of N consecutive observations as input and the next observation as output. Once the DNN is trained, the prediction, i.e., the DNN output (given a set of N consecutive observations as input), is provided by the class value that maximizes the probability distribution returned by the Softmax function. The metric for the loss is the 'sparse categorical crossentropy' and 'Adam' was used as an optimizer.[1] Table 1 provides the main DNN parameters.

[1] Tensorflow 2.2.0 library for Python 3.7 was used to build the artificial intelligence architecture.

2.2 The gap filling procedure

We performed the empirical analysis on US electricity price time series observed at PJM, on US natural gas spot price time series observed at Henry Hub, and on WTI crude oil price time series. All time series are at a daily frequency and show irregular sampling (lack of daily data points) as a result of weekends, holidays and other missing data (hereinafter, market-closure days). A missing values imputation (or gap filling) strategy can be really informative and provide fundamental knowledge [35], a phenomenon also known as *informative missingness* [26]. This is especially true for log-returns, being calculated as the difference in log-prices between two subsequent observations. Since relevant information affecting prices can also be released when the market is closed [36–38], log-returns computed on different length time intervals may have a different informative content. Hence, the DNN is first employed to fill the gaps in the original time series.

Let us denote by $\mathbb{T}_{obs}$ the original time grid composed of M ordered time positions corresponding to trading days in which market prices are available, namely

$$\mathbb{T}_{obs} = \{\bar{t}_1, \bar{t}_2, \cdots, \bar{t}_M\}, \tag{1}$$

and by $p_{\bar{t}_j}$ the market price at time $\bar{t}_j$, with $1 \leq j \leq M$. The information sequence is, therefore, composed of M market price observations. To fill the gaps in the original time series, we generated a set of predicted prices corresponding to market-closure days. In this task the DNN is trained on the set of market prices $p_{\bar{t}_j}$ with $\bar{t}_j \in \mathbb{T}_{obs}$. Each element of the DNN training dataset is composed of N consecutive market price observations as input and the next market price observation as output. Once the DNN is trained, the prediction is provided on the basis of the most probable price value in the DNN classification, given a set of N consecutive market price observations as input. Missing values are then imputed using DNN predictions, thus determining the whole filled in market price sequence p_t with $t \in \mathbb{T}$, where $\mathbb{T}$ is the complete set of daily ordered time grid positions that includes market-closure days as well as business days, namely

$$\mathbb{T} = \{1, 2, \ldots, m\}, \tag{2}$$

where '1' denotes the first day and 'm' denotes the last day of the ordered and filled time interval under investigation. In this way, we expand the cardinality of the set from M to $m = M + C$ by introducing C additional elements corresponding to the number of market-closure days.

2.3 The regime-switching dynamics

The DNN was then employed to build a hybrid regime-switching model of market prices. Let us denote by p_t, with $t \in \mathbb{T}$, the market price at time t and by

$$s_t = \ln p_t, \tag{3}$$

its natural logarithm. We assumed that s_t is a linear superposition of a deterministic component, f_t, accounting for trend and seasonality, and a random component, x_t, namely

$$s_t = f_t + x_t. \tag{4}$$

Log-returns are defined as changes in the stochastic component of log-prices, x_t, namely

$$r_t = x_t - x_{t-1}. \tag{5}$$

The DNN is then trained on the complete time series of log-returns, r_t with $t \in \{2, 3, \dots, m\}$. The aim is to account for correlations over time in a proper way to capture nonlinear dynamics present in the time series of log-returns. As in the gap filling procedure, each element of the DNN training dataset is composed of N consecutive log-return observations as input and the next log-return observation as output. Once the DNN is trained, the prediction is provided on the basis of the most probable value in the DNN classification, given a set of N consecutive log-return observations as input. Let us pose, therefore,

$$\theta_t = \theta(\boldsymbol{r}_{t-1,t-N}), \tag{6}$$

to denote the nonlinear function describing the prediction provided by the DNN on the basis of the observation of the array of N consecutive log-return values

$$\boldsymbol{r}_{t-1,t-N} = \{r_{t-1}, r_{t-2}, \dots, r_{t-N}\}, \tag{7}$$

and P_t to denote the probability of the network prediction, θ_t, once the realized sequence $\boldsymbol{r}_{t-1,t-N}$ is observed.

The dynamic model we propose to describe the complex dynamics of commodity prices observed in real market is a hybrid mean-reverting regime-switching model described by the following stochastic process,

$$x_t = \begin{cases} x_{t-1} - \alpha_0 x_{t-1} + \sigma_0 \epsilon_{0,t} & P_t < P_{\text{th}}, \\ x_{t-1} + \theta_t & P_t \geq P_{\text{th}}, \end{cases} \tag{8}$$

where $\epsilon_{0,t}$ are i.i.d. standard normal random variables and P_{th} is a given probability threshold. The base regime is described by a mean-reverting diffusion process with mean-reversion parameter α_0, and volatility σ_0. The second regime is driven by the predictions of the DNN. Whenever the DNN prediction probability for the next time step is lower than the given probability threshold, P_{th}, the system evolves in time according to the mean-reverting diffusion process of the base regime dynamics, otherwise if the DNN prediction probability for the next time step is greater than the probability threshold, the dynamics is prediction driven. In this model, therefore, the switching mechanism between regimes is not governed by an underlying hidden Markov process but it is driven by the DNN probability of forecasting future log-returns. Since the DNN is trained on observed data, we remark that the second regime too incorporates a mean-reverting behavior. In this way, the model allows us

to introduce different mean-reversion rates and volatilities depending on the state of the system thus providing a realistic description of the price dynamics.

Stochastic paths can be generated by Monte Carlo techniques in the following way. A random seed composed of a sequence of N log-returns generated by the dynamics of the base regime is used to initialize the process. Then, the regime-switching model is used to simulate a whole path on a given time interval. By repeating this procedure, Monte Carlo path samples can be obtained. This is a crucial step because stochastic paths allows us to investigate the characteristics of the hybrid dynamics and to estimate the model on market data. The pseudocode describing the main steps of the stochastic path generation process is reported below.

Algorithm Stochastic path generation

Generating the input seed for the DNN
for iteration $= 1, \cdots, N$ **do**
    Generate the input seed for the DNN by using the base regime of the dynamics (see Equation (8))
end for
Simulating stochastic paths
for iteration $t = 1, \cdots, m$ **do**
    if $P_t > P_{th}$ **then**
        Take the log-return value predicted by the DNN
    else
        Compute the log-return value by using the base regime of the dynamics (see Equation (8))
    end if
    Add to the path the so obtained log-return value
end for

3 The empirical analysis

We performed the empirical analysis on US electricity price time series observed at PJM, on US natural gas spot price time series observed at Henry Hub, and on WTI crude oil price time series. Our data set consists of daily prices since 1 January 2015 until 31 December 2019. Data are freely downloadable from www.eia.gov. Market prices are expressed, respectively, in dollars per megawatthour (electricity), dollars per million btu (natural gas), and dollars per barrel (crude oil). All time series are characterized by missing data corresponding to market-closure days.

3.1 Filling gaps in time series data

The first task of the empirical analysis was filling in the gaps in the original time series. To do this, we followed the gap filling procedure proposed by Mari and Mari [34] and briefly summarized below. In the reference dataset we used, market prices are expressed with two decimal places, and we required that the DNN price predictions must be characterized by the same number format. We defined,

therefore, a different class associated with each value belonging to the classification range, i.e., from the minimum to the maximum value of the observed price time series, rounded to two decimal places. For this filling in task, we trained the DNN on a dataset composed of sequences of $N = 5, 10, 20$ consecutive observations as input, and the DNN was trained for 500 epochs. As in our previous study, we observed that the learning process improved significantly from $N = 5$ to $N = 10$ but not as significantly from $N = 10$ to $N = 20$. We supposed that the improvement in the learning process from $m = 5$ to $m = 10$ was certainly due to the lengthening of training sequences but also to the composition of training sequences. In fact, for $N = 5$ some of the training sequences are composed of market prices related to consecutive calendar days and other training sequences are composed of market prices related to five subsequent observations that are not consecutive calendar days. The DNN learning process could be limited by this inhomogeneous composition of the training set. In the empirical analysis, we used the value $N = 10$. Two main reasons guided our choice. The first is that, in such a case, the training set is more homogeneous, being composed of sequences in which both missing data and subsequences of market prices related to consecutive calendar days are certainly present. The DNN could, therefore, learn also by observing how markets react to the information released during closure days. The second is that the learning process did not significantly improve when the number of consecutive observations is increased from $N = 10$ to $N = 20$. The so filled time series are depicted in the top panels of Fig. 3.

Typically, electricity prices may be higher in winter time and in summer time, so that a seasonal component must be included in the deterministic component of the dynamics, f_t, to account for this semiannual periodicity. Moreover, for all the time series a trend component must be also considered to account for expected inflation and possibly for a real escalation rate of the commodity prices (positive or negative). We used a LOESS based decomposition technique in order to detect the deterministic component of the dynamics f_t [39]. LOESS stands for 'LOcally Estimated Scatterplot Smoothing' and it is a flexible seasonal-trend decomposition technique that allows us to control the smoothness of the trend as well as the rate of change of the seasonal component [40]. Figure 3 depicts, for the period 1 January 2015 to 31 December 2019, the filled time series of electricity, gas and oil prices in the top panels; the deterministic components, f_t (in red), and the residual stochastic component of daily log-prices, x_t (in blue), in the middle panels; the log-return time series, r_t, in the bottom panels. The descriptive statistics of log-returns is displayed in Table 2.

The observed time series show very different characteristics. In particular, the standard deviation of empirical log-returns is very high in the case of PJM data with respect to the values observed in the other two markets. Moreover, in both electricity and gas markets, log-returns show large fluctuations with jumps and spikes, and non-normal, leptokurtic empirical distributions. On the other hand, the time series of WTI crude oil does not show such extreme behavior. We notice that the presence of jumps of high magnitude and spikes is revealed by the high values of the kurtosis observed in both electricity and gas markets. These values are significantly greater than the kurtosis value of the WTI crude oil log-return time series.

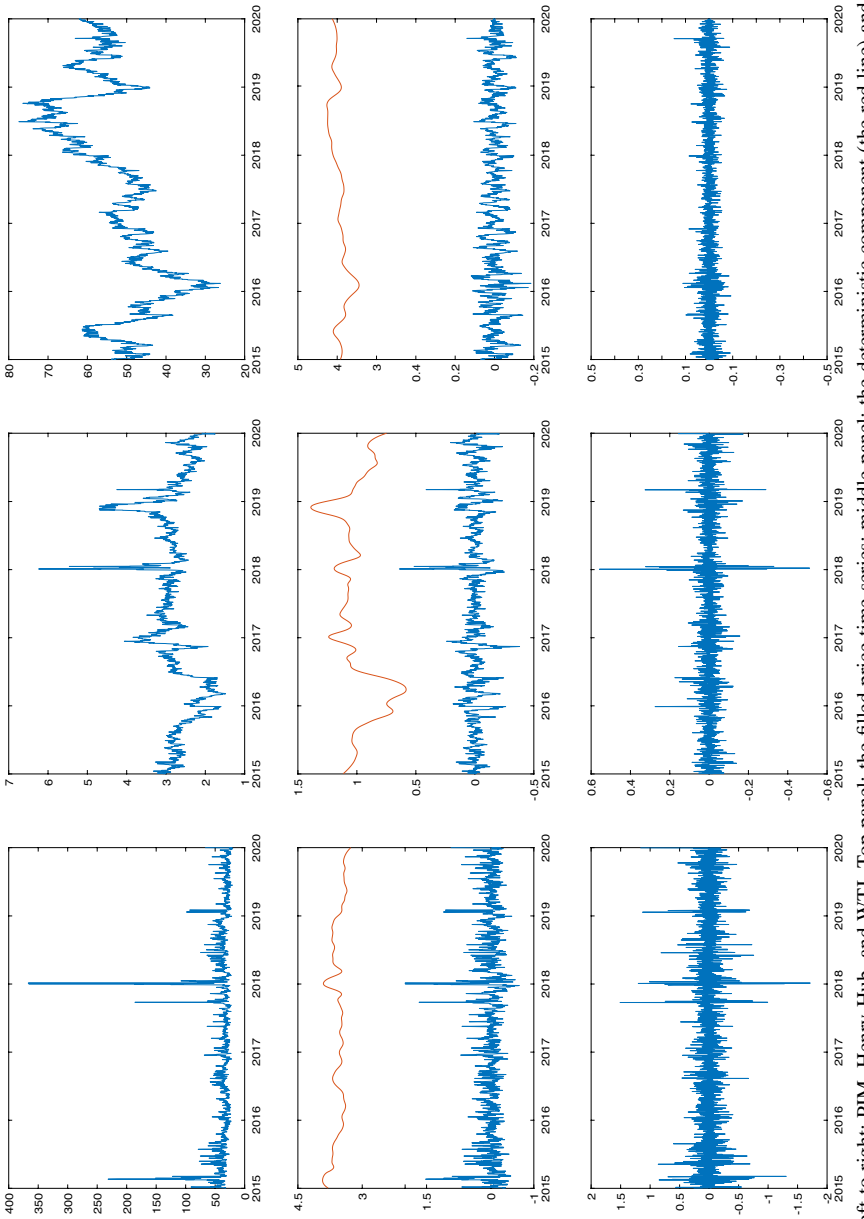

Fig. 3 From left to right: PJM, Henry Hub, and WTI. Top panel: the filled price time series; middle panel: the deterministic component (the red line) and the stochastic component (the blue line) of the price time series; bottom panel: log-return time series

Table 2 Descriptive statistics of log-returns

	Mean	Std	Skewness	Kurtosis
PJM	0.0002	0.1947	−0.0540	14.3641
Henry Hub	0.0000	0.0444	0.3453	34.2795
WTI	0.0000	0.0236	0.2454	5.5603

3.2 Training the DNN on log-return time series

The second task of the empirical analysis was training the DNN on the filled time series of log-returns, r_t. In this task, we expressed log-returns with four decimal places and we required that the DNN price predictions must be characterized by the same number format. We define, therefore, a different class associated with each value belonging to the classification range, i.e., from the minimum to the maximum value of the filled log-return time series, rounded to four decimal places. We trained the DNN using a training dataset composed of sequences of $N = 10, 20, 30$ consecutive observations as input, and the DNN was trained for 250 epochs. Figure 4 shows the behavior of loss functions. The fast decreasing behavior of the loss function in each market demonstrates that the DNN architecture is performing well in all the markets under investigation.

Looking at Fig. 4, we can observe that after 200 epochs the learning process does not improve significantly from $N = 10$ to $N = 30$. In the empirical analysis we used the value $N = 10$ for log-return predictions in the case of PJM and WTI, and the value $N = 30$ for log-return predictions in the Henry Hub market (in order to test the procedure also in this case).

3.3 Estimating the hybrid regime-switching model

The third task was estimating the regime-switching model on market data. As shown by Eq. 8, the dynamics depends on two parameters, namely the mean reversion parameter, α_0, and the volatility parameter, σ_0. We notice that the threshold probability, P_{th}, could be considered as a further parameter to be estimated on market data thus adding flexibility to the model. However, in our analysis we set $P_{th} = 0.8$.

The mean-reversion and the volatility parameters were estimated by using a two-step procedure. Specifically, the mean-reversion parameter was estimated in the first step, and the volatility parameter in the second step.

In the first step, since the mean-reversion component must force back prices to fluctuate around a long-run mean after a jump or a spike has occurred, the mean-reversion parameter, α_0, was estimated on the filled time series of detrended log-prices, x_t, by linear regression. Estimation results are displayed in Table 3.

In the second step, the volatility parameter, σ_0, was estimated. The approach we followed is based on the method of simulated moments [30, 31] by using Monte Carlo techniques [30, 32]. For each value of the parameter σ_0 belonging to a grid with values varying from 0.1 to 0.2 (in increments of 0.001) for PJM data, and with

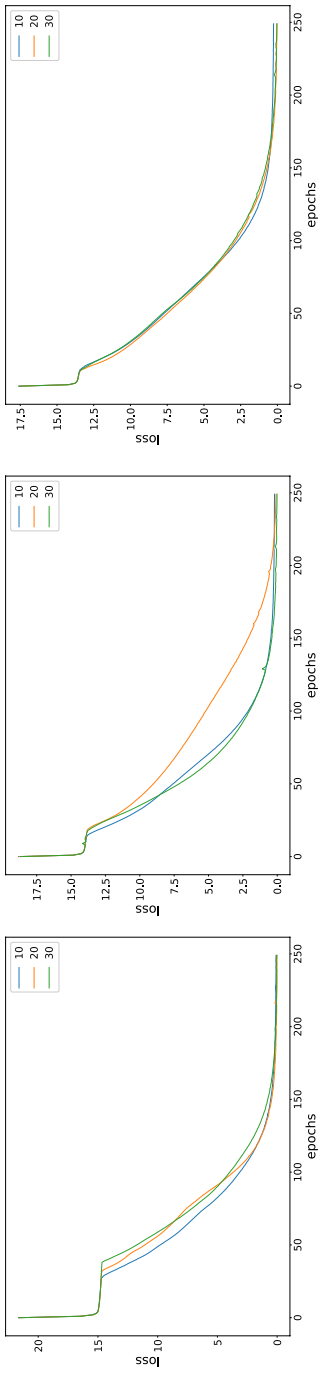

Fig. 4 Loss functions. From left to right: PJM, Henry Hub, and WTI

Table 3 Estimation results: step one. Standard errors (std err) are between parentheses

	α_0	std err
PJM	0.3452	(0.0178)
Henry Hub	0.1554	(0.0125)
WTI	0.1525	(0.0124)

Table 4 Estimation results: step two

	σ_0	Mean	Std	Skewness	Kurtosis
PJM	0.176	0.0002	0.1944	0.0209	13.4547
		(0.0001)	(0.0091)	(0.2009)	(4.1252)
Henry Hub	0.027	0.0000	0.0434	0.5030	36.6399
		(0.0001)	(0.0028)	(0.3787)	(7.4623)
WTI	0.019	0.0000	0.0231	0.1061	5.2611
		(0.0001)	(0.0012)	(0.3604)	(0.8903)

Estimated σ_0 and the first four moments of the simulated log-return distributions. Sample standard deviations are reported between parentheses

Table 5 Some statistical parameters

	n_{av}	p_{av}
PJM	1341	97.9%
Henry Hub	1317	98.4%
WTI	1212	98.5%

values varying from 0 to 0.05 (in increments of 0.001) for Henry Hub and WTI data,[2] a sample of one hundred random paths was generated by using Monte Carlo simulations from Eq. (8) with α_0 given by the estimate obtained in the first step. Each path has a length N equal to the number of calendar days in the interval under investigation, i.e., $m = 1826$ in the time interval 1 January 2015 to 31 December 2019. Along each path, the mean, standard deviation, skewness and kurtosis of log-returns (hereinafter, the first four moments of log-returns) were computed and averaged over the sample. We assumed that a given values of σ_0 offers a good fit if, for each moment, the difference between the sample average value and the observed value reported in Table 2 was less than one half the sample standard deviation for that moment. Estimation results are depicted in Table 4. The fitting value of σ_0 is shown together with the first four moments of the log-return distribution computed on the Monte Carlo sample.

Table 5 shows the average number of the DNN predictions, n_{av}, and the average probability of the DNN prediction, p_{av}, computed in a Monte Carlo sample

[2] Volatility ranges were set consistently with the standard deviation values of log-returns reported in Table 2.

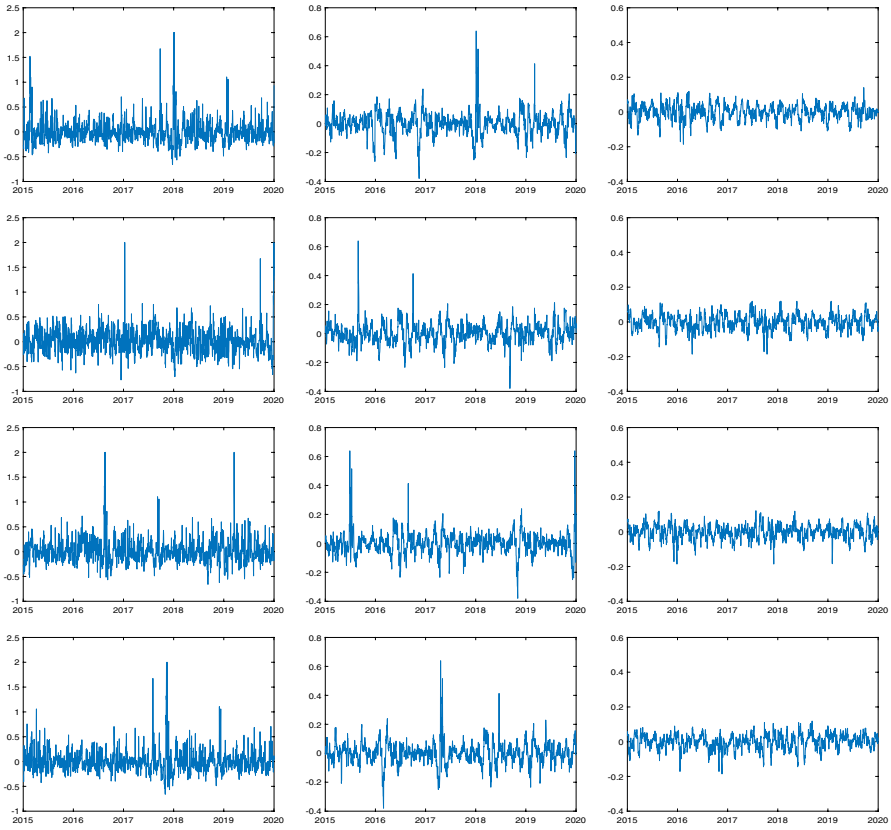

Fig. 5 From left to right: PJM, Henry Hub, and WTI. From top to bottom: the detrended log-price time series and three log-price paths obtained by Monte Carlo model simulations using estimated parameters

of one hundred simulated paths by using estimated parameters. We notice that, on average, the dynamics is driven by the DNN predictions at about seventy percent and that, on average, the prediction probability is very high (about 98%) in all the markets under investigation.

The proposed model provides an interesting description of the price dynamics observed in real markets thus offering a good interpretation of the main stylized facts of commodity price dynamics and a remarkable agreement with experimental data. Moreover, we observe that the model can generate log-price paths with various and interesting shapes and that, due to the pattern recognition mechanism based on the DNN predictions in the second regime, some of them are very similar to those observed in real markets. This can be seen by looking at Fig. 5 that depicts, from top to bottom, the detrended log-price observed time series and three log-price paths obtained by Monte Carlo simulations using estimated parameters.

4 An empirical comparison with regime-switching hidden Markov models

Regime-switching hidden Markov models offer the possibility of combining in one model periods of stable dynamics and periods of turbulent dynamics, depending on the realization of a stochastic latent state variable of the system. Like jump-diffusion processes, regime-switching hidden Markov models provide examples of non-Gaussian models with stochastic volatility. In this section, we perform a comparative analysis between the model proposed in this paper and two alternative regime-switching hidden Markov models properly designed for capturing the main features of log-returns observed in commodity markets. The main features of these models, called respectively Model 1 and Model 2, are described below.

4.1 Model 1

In Model 1, the dynamics of both the base regime and the turbulent regime is described by a mean-reverting diffusion process. Using the Euler discretization with time step equal to one day, Model 1 can be represented by the following process,

$$x_t = \begin{cases} x_{t-1} - \alpha_0 x_{t-1} + \sigma_0 \epsilon_{0,t}, \\ x_{t-1} - \alpha_1 x_{t-1} + \sigma_1 \epsilon_{1,t}, \end{cases} \tag{9}$$

where both $\epsilon_{0,t}$ and $\epsilon_{1,t}$ are i.i.d. standard normal random variables for all t. We assumed that $\epsilon_{0,t}$ and $\epsilon_{1,t}$ are mutually independent random variables for all t. The parameters α_0 and α_1 as well as σ_0 and σ_1 allow us to account for different mean-reversion rates and volatilities depending on the state of the system.

4.2 Model 2

Model 2 adds a further degree of freedom to the description of the dynamics of market prices including Poisson jumps in the second regime. In the specific, the dynamics of the base regime is described by a mean-reverting diffusion process in order to account for the motion during stable periods; the dynamics of the turbulent regime is described by a mean-reverting jump-diffusion process. Using the Euler discretization with time step equal to one day, Model 2 can be represented by the following process,

$$x_t = \begin{cases} x_{t-1} - \alpha_0 x_{t-1} + \sigma_0 \epsilon_{0,t}, \\ x_{t-1} - \alpha_1 x_{t-1} + \sigma_1 \epsilon_{1,t} + J\zeta_t, \end{cases} \tag{10}$$

where, as in the previous model, $\epsilon_{0,t}$ and $\epsilon_{1,t}$ are i.i.d. standard normal random variables for all t. The random jump amplitude J is assumed to be distributed as a Gaussian random variable with zero mean and standard deviation σ_J, i.e., $J \sim N(0, \sigma_J^2)$. The zero-mean assumption for the jump amplitude is due to the low values of the skewness observed in the log-return time series. The random variables ζ_t are i.i.d. binary random variables assuming the value 1 with probability λ and the value 0

Table 6 Estimation results for PJM electricity prices

	Model 1	Model 2
α_0	0.3230	0.3147
	(0.0198)	(0.0287)
σ_0	0.1238	0.1156
	(0.0034)	(0.0037)
α_1	0.3669	0.3729
	(0.0265)	(0.0284)
σ_1	0.4059	0.2328
	(0.0337)	(0.0171)
λ		0.1619
		(0.0314)
σ_J		0.5665
		(0.0590)
$1-\gamma$	0.9827	0.9719
	(0.0053)	(0.0076)
$1-\eta$	0.8557	0.8832
	(0.0398)	(0.0274)
LL	884.2	896.0
SC	−1723.4	−1731.9

Standard errors are between parentheses

with probability $1-\lambda$. We assumed that $\epsilon_{0,t}$, $\epsilon_{1,t}$, ζ_t and the jump amplitude J are mutually independent random variables for all t.

In both models, the switching between regimes is driven by a two-valued hidden Markov process characterized by the following transition probability matrix,

$$\boldsymbol{\pi} = \begin{pmatrix} 1-\xi & \eta \\ \xi & 1-\eta \end{pmatrix}, \tag{11}$$

where ξ denotes the transition probability of a switch from the base regime to the turbulent regime in the time interval $[t-1, t]$, and η is the probability of the opposite transition. Both models were estimated on market data by maximum likelihood using the Hamilton filtering technique [1]. In fact, the six parameters of Model 1 and the eight parameters of Model 2 are too many to be calibrated by the method of simulated moments with Monte Carlo techniques. Estimation results are depicted in Table 6 in the case of PJM, in Table 7 for the Henry Hub market, and in Table 8 for WTI crude oil prices. For each market, the parameters estimates, the log-likelihood (LL), and the value of the Schwartz Criterion (SC) are reported. In the case of the Henry Hub market the empirical analysis reveled that the mean-reversion component of the turbulent dynamics is negligible in both Model 1 and Model 2, and this is the reason why such a parameter does not appear in Table 7. Model 2 seems to outperforms Model 1 in all markets under consideration, as it can been observed by looking at the value of the Schwartz Criterion.

Table 7 Estimation results for Henry Hub gas prices

	Model 1	Model 2
α_0	0.0937	0.1133
	(0.0097)	(0.0126)
σ_0	0.0307	0.0352
	(0.0008)	(0.0008)
σ_1	0.1498	0.0035
	(0.0113)	(0.0005)
λ		0.1274
		(0.0146)
σ_J		0.1976
		(0.0262)
$1-\gamma$	0.9692	0.8406
	(0.0059)	(0.0170)
$1-\eta$	0.3631	0.2771
	(0.0141)	(0.0250)
LL	3500.8	3598.9
SC	−6964.1	−7145.2

Standard errors are between parentheses

Table 8 Estimation results for WTI crude oil prices

	Model 1	Model 2
α_0	0.1299	0.1769
	(0.0136)	(0.0423)
σ_0	0.0182	0.0329
	(0.0006)	(0.0025)
α_1	0.1808	0.1190
	(0.0198)	(0.0116)
σ_1	0.0316	0.0135
	(0.0016)	(0.0026)
λ		0.4958
		(0.0429)
σ_J		0.0192
		(0.0028)
$1-\gamma$	0.9911	0.9830
	(0.0055)	(0.0081)
$1-\eta$	0.9738	0.9963
	(0.0120)	(0.0026)
LL	4407.8	4417.5
SC	−8770.5	−8775.0

Standard errors are between parentheses

Table 9 The first four moments of the simulated log-return distributions for PJM

	Mean	Std	Skewness	Kurtosis
Model 1	0.0002	0.1940	0.0166	9.0265
	(0.0001)	(0.0131)	(0.2271)	(1.3078)
Model 2	0.0002	0.1944	0.0174	12.0881
	(0.0001)	(0.0124)	(0.4182)	(2.7974)

Sample standard deviations are reported between parentheses

Table 10 The first four moments of the simulated log-return distributions for the Henry Hub market

	Mean	Std	Skewness	Kurtosis
Model 1	0.0000	0.0450	0.0203	17.1262
	(0.0001)	(0.0025)	(0.6639)	(3.6851)
Model 2	0.0000	0.0450	0.0109	25.5072
	(0.0001)	(0.0028)	(1.0348)	(7.9593)

Sample standard deviations are reported between parentheses

Table 11 The first four moments of the simulated log-return distributions for WTI

	Mean	Std	Skewness	Kurtosis
Model 1	0.0000	0.0232	0.0041	3.9910
	(0.0001)	(0.0013)	(0.0916)	(0.2976)
Model 2	0.0000	0.0234	−0.0018	4.2271
	(0.0001)	(0.0016)	(0.0918)	(0.3082)

Sample standard deviations are reported between parentheses

A comparative analysis among the hybrid regime-switching model proposed in this paper and the regime-switching Model 1 and Model 2 can be performed on the basis of their ability to reproduce the first four moments of empirical log-return distributions. To accomplish this task, we created, for each market under investigation, a sample of one thousand paths randomly generated by Monte Carlo simulations of the dynamics described by Model 1 and Model 2 by using estimated parameters. Each path in the sample has a length N equal to the number of calendar days in the interval under investigation, i.e., $m = 1826$ in the time interval 1 January 2015 to 31 December 2019. Along each path, the first four moments of log-returns were computed and averaged over the sample. Average values are reported in Table 9 for PJM, in Table 10 for the Henry Hub market, and in Table 11 for WTI.

Looking at Table 2 that depicts the first four moments of the observed log-return distributions, we observe that also in this respect Model 2 outperforms Model 1. The inclusion of a jump component in the turbulent regime of Model 2 improves significantly the agreement with experimental data with respect to Model 1 in both the electricity and the natural gas markets. However, the improvement is not so significant in the case of WTI crude oil prices.

The hybrid regime-switching model proposed in this paper shows a better performance with respect to both Model 1 and Model 2, thus reproducing in a very remarkable way the first four moments of the log-return distributions observed in all the markets under investigation. Two are the main theoretical reasons of such performance improvement. The first is due to the pattern recognition mechanism included in the second regime of the hybrid model based on the DNN predictions. The second is that the switching mechanism between regimes in the hybrid model is not governed by an underlying hidden Markov process but it is driven by the DNN probability of forecasting future log-returns: the dynamics switches from the base regime to the second regime whenever the DNN prediction probability is greater than a given probability threshold and it comes back to the base regime whenever the DNN prediction probability is lower than the probability threshold. As a consequence, the model seems to behave more realistically thus reproducing well the empirical log-return distributions and the shape of the empirical price paths observed in real markets. However, it is worth to remark that, due to the characteristics of the switching mechanism in which the DNN prediction probability must be calculated at each time step, the main disadvantage of the proposed approach could be the computational time needed for estimating the model on experimental data. As it will be discussed in the next, concluding section, this fact could be relevant when considering possible extensions of the hybrid model proposed in this study.

5 Concluding remarks

To investigate the complex dynamics of commodity prices observed in real markets, we presented, in this paper, an AI based hybrid model with regime-switching. In the proposed model, the base regime is described by a mean-reverting diffusion process and the second regime is driven by the predictions of a multi-layer DNN with a well defined architecture. Since very often market price time series exhibit irregular sampling with missing data as a result of weekends, holidays and other market specific reasons, we also used forecasts from the DNN itself to fill gaps in the original time series. Then, we provided an estimation procedure for the hybrid model. Being defined by only two parameters, the model can be easily calibrated on market data. The obtained results showed an interesting agreement with empirical data. Specifically, the model reproduces in a remarkable way the first four moments of the empirical distribution of log-returns as well as the shape of the observed price time series.

The hybrid model proposed in this paper seems to be a powerful tool of analysis for investigating the features of the dynamics of commodity prices. In this regard, we remark that having good models capturing well the main stylized fact of observed market prices is of crucial importance for numerous financial applications ranging from hedging financial risk, also through the use of commodity derivatives, to the evaluation of long-term investments in energy generation technologies.

The proposed approach can be extended in three main directions. Two of them concern the possibility of making the model more flexible in order to improve the prediction approach and the agreement with experimental data. Indeed, in the empirical analysis proposed in this study we set the threshold prediction probability to a

well defined value. However, such a parameter can be considered as a further parameter to be estimated on market data thus adding flexibility to the model. The main drawback is that the computational time needed to perform the parameter estimation can increase significantly. To overcome this difficulty, ad-hoc numerical optimization procedures must be investigated. The second direction regards the possibility to extend the model by including more general processes to describe the dynamics of the base regime. For example, the inclusion of a jump component in the base regime can add further flexibility (and complexity) to the proposed hybrid model. In such a case, appropriate estimation techniques must be investigated in order to limit the computational time. Finally, although our analysis focused on commodity market prices, the proposed model is general and could be used to describe phenomena belonging to different context ranging from physical science to social science. These topics will be left for future investigations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. Hamilton, J.D.: A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrica **57**, 357–384 (1989)
2. Elliott, R.J., Siu, T., Lau, J.W.: A hidden Markov regime-switching smooth transition model. Stud. Nonlinear Dyn. Econ. **22**(4), 20160061 (2018). https://doi.org/10.1515/snde-2016-0061
3. Eichler, M., Türk, D.: Fitting semiparametric Markov regime-switching models to electricity spot prices. Energy Econ. **36**, 614–624 (2013)
4. Huisman, R., Mahieu, R.: Regime jumps in electricity prices. Energy Econ. **25**, 423–434 (2003)
5. Weron, R., Bierbrauer, M., Trück, S.: Modeling electricity prices: jump-diffusion and regime switching. Physica A **336**, 39–48 (2004)
6. Mari, C.: Regime-switching characterization of electricity prices dynamics. Physica A **371**, 552–564 (2006)
7. Paraschiv, P., Fleten, S.E., Schürle, M.: A spot-forward model for electricity prices with regime shifts. Energy Econ. **47**, 142–153 (2015)
8. Chen, Z., Forsyth, P.A.: Implications of a regime-switching model on natural gas storage valuation and optimal operation. Quant. Finance **10**(2), 159–176 (2010)
9. Leonhardt, D., Ware, A., Zagst, R.: A cointegrated regime-switching model approach with jumps applied to natural gas futures prices. Risks **5**, 48 (2017). https://doi.org/10.3390/risks5030048
10. Bernabe, A., Martina, E., Alvarez-Ramirez, J., Ibarra-Valdez, C.: A multi-model approach for describing crude oil price dynamics. Physica A **338**, 567–584 (2004)
11. Zhu, D., Ching, W., Elliott, R.J., Siu, T., Zhang, L.: Hidden Markov models with threshold effects and their applications to oil price forecasting. J. Ind. Manag. Optim. **13**(2), 757–773 (2017)
12. De la Torre-Torres, O.V., Aguilasocho-Montoya, D., Álvarez-García, J., Simonetti, B.: Using Markov-switching models with Markov chain Monte Carlo inference methods in agricultural commodities trading. Soft Comput. **24**, 13823–13836 (2020)
13. Chen, S., Insley, M.: Regime switching in stochastic models of commodity prices: an application to an optimal tree harvesting problem. J. Econ. Dyn. Control **36**(2), 201–2019 (2012)

14. Wang, M., Lin, Y., Mikhelson, I.: Regime-switching factor investing with hidden Markov models. J. Risk Financ. Manag. **13**, 311 (2020). https://doi.org/10.3390/jrfm13120311
15. Bansal, R., Zhou, H.: Term structure of interest rates with regime shifts. J Finance **57**(5), 1997–2043 (2002)
16. Ang, A., Bekaert, G.: Regime switches in interest rates. J. Bus. Econ. Stat. **20**, 163–182 (2002)
17. Zhou, N., Mamon, R.: An accessible implementation of interest rate models with Markov-switching. Exp. Syst. Appl. **39**(5), 4679–4689 (2012)
18. Geman, H.: Commodities and Commodity Derivatives. Wiley, Chichester, UK (2005)
19. Weron, R.: Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach. Wiley, Chichester (2013)
20. Eydeland, A., Wolyniec, K.: Energy and Power Risk Management. Wiley, Chichester, UK (2003)
21. Fiorenzani, S.: Quantitative Methods for Electricity Trading and Risk Management. Palgrave Macmillan, London (2006)
22. Geman, H., Roncoroni, A.: Understanding the fine structure of electricity prices. J. Bus. **79**, 1225–1262 (2006)
23. Zhang, P., Qi, G.M.: Neural network forecasting for seasonal trend time series. Eur. J. Oper. Res. **160**, 501–514 (2005)
24. Wong, W.K., Xia, M., Chu, W.C.: Adaptive neural networks models for time-series forecasting. Eur. J. Oper. Res. **207**, 807–816 (2010)
25. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Comput. **9**, 1735–1780 (1997)
26. Che, Z., Purushotham, S., Cho, K., Sontag, D., Liu, Y.: Recurrent neural networks for multivariate time series with missing values. Sci. Reports **8**, 6085 (2018)
27. Xu, Z., Zhang, J., Wang, J., Xu, Z.: Prediction research of financial time series based on deep learning. Soft Comput. **24**, 8295–8312 (2020)
28. Bao, W., Yue, J., Rao, Y.: A deep learning framework for financial time series using stacked autoencoders and long-short term memory. PLoS One (2017). https://doi.org/10.1371/journal.pone.0180944
29. Ren, H., Cromwell, E., Kravitz, B., Chen, X.: Using deep learning to fill spatio-temporal data gaps in hydrological monitoring networks. Hydrol. Earth Syst. Sci. Discuss. (2019). https://doi.org/10.5194/hess-2019-196
30. McFadden, M.: A method of simulated moments for estimation of discrete response models without numerical integration. Econometrica **57**(5), 995–1026 (1989)
31. Duffie, D., Singleton, K.: Simulated moments estimation of Markov models of asset prices. Econometrica **61**, 929–952 (1993)
32. Gelman, A.: Method of moments using Monte Carlo simulation. J. Comput. Graph. Stat. **4**(1), 36–54 (1995)
33. Mari, C.: Short-term movements of electricity prices and long-term investments in power generating technologies. Energy Syst. (2021). https://doi.org/10.1007/s12667-020-00422-8
34. Mari, C., Mari, E.: Gaussian clustering and jump-diffusion models of electricity prices: a deep learning analysis. Dec. Econ. Finance (2021). https://doi.org/10.1007/s10203-021-00332-z
35. Owen, M.: Practical Signal Processing. Cambridge University Press, Cambridge (2007)
36. French, K.R.: Stock returns and the week-end effect. J. Financ. Econ. **8**, 55–69 (1980)
37. Mantegna, R..N., Stanley, H..E.: An Introduction to Econophysics - Correlations and Complexity in Finance. Cambridge University Press, Cambridge (2007)
38. Voit, J.: The Statistical Mechanics of Financial Markets. Springer, Berlin (2005)
39. Cleveland, R.B., Cleveland, W.S., McRae, J.E., Terpenning, I.J.: STL: a seasonal-trend decomposition procedure based on LOESS. J. Off. Stat. **6**(1), 3–33 (1990)
40. Dagum, E..B., Bianconcini, S.: Seasonal Adjustment Methods and Real Time Trend-Cycle Estimation. Springer, Cham (2016)

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