Energy Strategy Reviews 21 (2018) 121–129

Contents lists available at ScienceDirect

Energy Strategy Reviews

journal homepage: www.elsevier.com/locate/esr

Modeling the price dynamics of three different gas markets-records theory

Hassan Hamie[a,*,1], Anis Hoayek[b,1], Hans Auer[a]

[a] Vienna University of Technology, Institute of Energy Systems and Electrical Drives, Energy Economics Group (EEG), Austria
[b] University of Montpellier, Institute Alexander Grothendieck, France

ARTICLE INFO

Keywords:
Yang model
Random walk model
Natural gas markets
Extremes
Price records

ABSTRACT

Previous research on extreme values has been used in commodity pricing, more specifically in trying to capture the dynamic behavior of the random occurrences of extreme events, such as price spikes/drops. This research uses the records theory to study the effect of extreme gas prices and the probability of future records.

To our knowledge, records theory was not previously applied to gas markets. The aim of this study is to test the stability of three different regional gas markets, each having its own supply and demand characteristics.

Records theory, studies observations that are higher than all previous ones, which is equivalent to say the maximum/minimum observation up to present time, and those records, are concentrated in the tail of a given distribution.

In this study, several models are developed to test and analyze the stability of three main regional gas markets (U.S, Europe and Asia). The *classical* model is used for the case where gas prices are independent and identically distributed (*i. i. d case*). Alternative models, such as Yang model and the discrete-time random walk, are used, where the number of records grows faster than in the *i. i. d* case and where records are not only concentrated among the first observations.

In spite of the non-independent and non-identically distributed properties of the models, the results are distribution free. Consequently, the applicant will not be concerned by identifying the distribution type and the complexity of the models is reduced.

1. Introduction

The global natural gas market is comprised of regional markets that are often grouped based on the regions of natural gas trade (i.e., North America, Europe, and Asia). In recent years, roughly 70% of global natural gas trade has been transported to market destinations within the country of production, while the remaining have crossed international borders, either through long-distance pipelines or through liquefied natural gas [1]. The evolution of the global natural gas market is dependent on a number of factors: natural gas reserves, the production in conjunction with demand, and the ability to meet demand with supplies from other regions.

Three regional gas markets exist around the world, and each one is clustered by a regional spot market:

The US market – cleared by Henry Hub spot prices.
The European market – cleared by spot prices at European hubs.
The Asian market – cleared by spot LNG prices.

In Europe and the US's regional market, natural gas is mostly purchased through pipelines due to large domestic resources and strong grids. The lack of such infrastructures in North East Asia prevents the import of natural gas through pipelines. Therefore, natural gas could be only imported in the form of Liquefied Natural Gas (LNG), which is shipped on maritime tankers. The highest demand for natural gas in 2017 was in North America, a value that is closely followed by Europe, then Asia.[2]

Two basic pricing systems are commonly used for international trade of natural gas. The split in price formation varies deeply between regional markets, depending on several structural factors such as regulation, liberalization process, contracting practices, existence of a spot market, liquidity, and share of imports.

- Gas-on-gas pricing, where the price of natural gas is competitively determined based on gas market spot prices. As such, prices vary as a response to natural gas supply and demand.
- Oil-indexation pricing, where the price of natural gas is determined

* Corresponding author.
E-mail addresses: hasan.hamieh@hotmail.com, Hassan.Hamieh@student.tuwien.ac.at (H. Hamie).

[1] These authors contributed equally to this work.

[2] LNG represents approximately 30% of international natural gas trade, and is used to meet both primary and peak natural gas demand. Japan and Korea, the two largest LNG importers jointly comprised around 50% of the global LNG market in 2015–2016.

https://doi.org/10.1016/j.esr.2018.05.003
Received 23 November 2017; Received in revised form 4 May 2018; Accepted 31 May 2018

based on oil market spot prices. As such, prices vary as a response to oil supply and demand.

North America has the most liberalized system, where gas pricing is highly competitive and is based on supply/demand balances (gas-on-gas pricing). On the other side, Asian and European gas prices are highly influenced by oil-indexation. In Northeast Asia, for example, almost all LNG contract volumes are indexed to crude benchmarks (e.g. JCC, Brent), because of their dependence on external imports [2]. However, as long-term contracts expire, the oil-indexation system is eroding.

It is almost challenging to get an accurate assessment of gas indexation levels and growth. Most of the supply contracts are structural long term portfolio contracts which are considered to be highly sensitive confidential information. Typically contract terms and conditions have evolved over time through the processes of renegotiation and price re-openers [3].

Facts have shown that natural gas, which is traded on the wholesale market, exhibits particularly large increases in price volatility. The rise of competition and deregulation leads to relatively free energy markets, which are characterized by high price shifts. Therefore, the market is, "vulnerable" to price spikes/drops. In response to an unpredictable, volatile and risky environment, protection against market risk has become a necessity.

Accordingly, it is important to model the gas price fluctuations and implement an effective tool for energy price risk management. Value at Risk (VaR) has become a popular risk measure in the financial industry. Since VaR estimations are only related to the tails of a probability distribution, techniques from Extreme Value Theory (EVT) may prove particularly effective.

The study of extremes focuses on outliers, a characteristic which enables a better prediction of unexpected extreme changes [4]. The inherent stylized facts exhibited by commodity markets make the direct use of EVT impossible. For this reason, most of the applications in commodity markets involve a conditional approach two steps introduced [5], and is known as the GARCHEVT approach.

The first step captures the stochastic volatility of the time series. The second step consists of applying EVT to the pseudo-independent and identically distributed ($i.\ i.\ d$) innovations obtained in the first step [6–9].

Another line of research includes other stylized facts, such as long-term memory, change of regimen in volatility and asymmetric effects [10]. For instance [11], analyze the regimen changes on volatilities for crude oil markets (Brent and WTI) and stock markets of UK, France and Japan and find two possible volatility regimens [12]. also consider volatility models including long-range memory for estimating risk measures for some major crude oil and gas commodities. The research showed that, models with long-range memory and asymmetry perform best in one-day-ahead forecasting [13]. explores the relevance of asymmetry and long-term memory to model. A forecast of the conditional volatility was applied in four widely traded commodities (crude oil, natural gas, gold, and silver). The findings show that nonlinear GARCH models, capturing these stylized facts perform better in terms of volatility forecasting.

Finally, the study of extremes can also find the pair wise dependence (co-movement) between different markets that can vary from almost independent to strongly dependent in contrast to previous literature [14,15].

The application of records theory to study extreme events instead of the classical EVT generates many advantages. First, almost all the results of the EVT are asymptotic, non-exact, and depend on the choice of the underlying distribution. While in many of the record theory findings, results are exact and non-asymptotic [16,17].

In addition, several properties of record models are distribution free, i.e. (independent of the choice of the underlying distribution. This helps practitioners to overcome the theoretical complexity, which is hidden behind the choice of the right distribution.

Second, the EVT approach is generally applied in a context where the observations are independent and identically distributed ($i.\ i.\ d$), which is not always a good hypothesis to be considered. Moreover, going beyond the $i.\ i.\ d$ case in EVT makes the work even more complicated. However, record models beyond $i.\ i.\ d$ context are easily manipulated. Worthy to note, several properties retain their distribution-free nature, which is a big advantage in practical problems.

Finally, in EVT all the results concentrate on the value of the extreme events. This is done by studying the standardized maxima of the observations, without taking into consideration the time these extreme events took place. However, records theory focuses on the values and times of extreme events, a feature making the analysis of the potential results richer. The study of time in record models is accounted for through particular random variables called Record Indicators [18].

Classical econometric models, usually focusing on the whole distribution, have been widely used in literature. Models such as univariate and multivariate GARCH are the most popular methods used for analyzing high-frequency time series data. Authors such as [19–21] have used this classical approach to model natural gas/oil volatility.

As widely agreed in the literature, inferences that do not take into consideration the regime switching phenomenon may lead to unreliable results for much high-frequency time series. In the case of oil and gas markets, sudden short period shocks will not be accounted for.[3] Authors such as [22] and [23], have shown that evidence of regime-switching shall not be ignored in the behavior of natural gas prices, and that the regime switching model performs noticeably better than non-switching models in these cases.

To find evidence of causality between supply and demand variables that affects the natural gas prices, stochastic models such as, multivariate vector autoregressive (VAR) and vector error correction (VECM) models were used by Refs. [25] and [26]. Others, such as [27], study the relationship between international gas market prices and their relation to the oil price through principal components analysis and Johansen likelihood-based co-integration procedure.

In contrast to VAR and VECM models, that assume a stable relationship, the relationship between the variables could be different in the separate regimes. Therefore, authors such as [24] have used the Markov-switching vector autoregressive (MS-VAR) models.

However, all the classical models cited above contain a large number of parameters, a fact that poses estimation challenges, and over-parameterization concerns [28,29]. Such difficulties do not concern the records theory, which does not impose constraints on the quality and distribution of residuals. This certainly alleviates the use of multiple statistical tests that make the classical econometric approaches defined on hypothesis quasi-impossible to be entirely verified.

Non-parametric and non-linear models are also used in literature. The machine learning is essential and can be used to model the complex non-linear relationship between different variables because they are "constraint-free". Therefore, there is no need for additional tests (i.e. normality test for residuals, autocorrelation, etc.). [30], uses machine learning techniques to forecast the movement of the day ahead natural gas spot prices. A second category, such as [31], use the neural network to predict the daily natural gas consumption needed by gas utilities. A third category, like [32], use Gamma test, a mathematically proven smooth test with a wide variety of applications that helps machine learning modelers choose the best input combination before calibrating and testing models, a characteristic that reduces the inputs selection uncertainty.

However, the records theory deals with extreme value of extreme values. As a consequence, the number of available observations is generally small [16]. Thus, the use machine learning models is not

[3] For example, a major event-causing shock will lead to an immediate increase in volatility in natural gas returns.

Table 1
Data collection.

Variable	Frequency	Number of observations	Description	Unit	Source
JCC	Monthly	237	The Japan Customs-cleared Crude prices (1997–2016)	*USD per MMBtu*	PAJ[a]
HH	Monthly	237	The Henry Hub prices (1997–2016)	*USD per MMBtu*	U.S. Energy Information Administration, EIA[b]
GBP	Monthly	237	The German Border Prices (1997–2016)	*USD per MMBtu*	European Energy Exchange, EEX[c]

[a] Available at the Petroleum Association of Japan, http://www.paj.gr.jp/statis/statis/data/07/paj-7_201701.xls
[b] Available at the Energy Information Administration, https://www.eia.gov/dnav/ng/hist/rngwhhdm.htm
[c] Available upon request from the European Energy Exchange AG, marketdata@eex.com

reasonable because of their need to have lots of training data in order to perform well.

The objectives of this research is, first to study the stability of the three regional gas markets by using the records theory, and second, to assess the probability of witnessing a spike/drop in the short term gas prices in three different markets. As previously explained, there have been attempts to forecast the probability of extremes in commodity prices. Yet to the author's knowledge, this is the first time in which the records theory is applied.

Since the time series of the markets are mostly $non-i.\ i.\ d$, and the assumption of the type of distribution is complex, the best approach to find a feasible solution for the gas markets is to use the records theory. Finally the models will be tested to check the reliability of the results. Testing is done by comparing the theoretically expected number of records to the real ones. This will help relevant stakeholders to better estimate their risk portfolio in the short term.

In section 2 we present the most popular record models in $i.\ i.\ d$ case and beyond. In section 3, we explore the significance of the results and highlight its impact. In the final section, we indicate the uses and extension of our work.

2. Material and methods

2.1. Study area and data used

The German Border Price (GBP) is an average of the oil-indexed contracts, and is comprised of a big percentage of Russian gas supplies and spot supplies. The latter is increasingly available at the Dutch-German border and Norwegian pipeline terminals. The GBP became a useful reference in Europe as most of the producer/wholesaler gas imports used the German border as the transfer of ownership point.

Japanese Crude Cocktail (JCC) is the average price of customs-cleared crude oil imports into Japan (formerly the average of the top twenty crude oils by volume) as reported in customs statistics. JCC is currently used as the index of gas prices in the Asia Pacific, of which Japan alone share 37.2% of the total imports in 2013 [33]. The Asian LNG value of imports, in particular, depends the demand for gas, and the long-term contract price (a value that is based on the JCC oil-linked pricing formation) [34].

The Henry Hub, lends its name to the pricing point for natural gas futures contracts traded on the New York Mercantile Exchange (NYMEX) and the OTC swaps traded on Intercontinental Exchange (ICE). Spot and future natural gas prices set at Henry Hub are the primary price set for the North American natural gas market [35].

All prices are denominated in USD per MMBtu (millions of British thermal units). The data set consists of monthly values recorded between March 1997 to November 2016. Table 1 summarizes the definition of the variables used in this study.

Fig. 1 shows the time series of the natural gas prices of the three different markets. The graph is rich in statistical parameters that could be interpreted using statistics (such as correlation, trends, stationary, etc.) and economics tools. We will first divide the time frame into five main periods and explain, in short, the fundamental supply and demand factors that led to the price behavior.

2.1.1. 1997–2006

Gas prices in Asia and Europe moved most closely with crude oil prices, as these were formally indexed to crude/fuel products. Wholesale gas prices in the United States were more volatile, but they, similarly, tended to co-move broadly with oil prices on account of some switching between energy types (natural gas and fuel oil).

During this period, the U.S. natural gas market was the world's largest, with 75% of the supply produced domestically, with the balance between storage withdrawals and imports from Canada (NR). Tight supply-demand balance characterized U.S natural gas markets, and has led to natural gas spikes in several periods during 2000–2008. As shown in the graph, this period witnessed an upward thrust in US gas prices.

The upward of price trend was also seen in Asia and Europe. However, no shocks have been observed in the U.S. The difference of prices between JCC and GBP is mainly due to two reasons: first, Liquefied Natural Gas import prices in the Pacific Basin are more expensive by roughly 1$ per MMBtu as compared with the Atlantic Basin.[4] Second, the gas market in continental Europe does not heavily rely less on gas imports via LNG.

2.1.2. 2006–2008

Over 2007 and into the first half of 2008, natural gas prices maintained their rising behavior in all markets. This is mainly due to tight supplies and unprecedented growth in oil demand. As mentioned previously, oil-indexation in Europe was dominant in international trade; however since 2006, the use of gas- on-gas pricing steadily increased to reach a significant percentage. In Asia, oil-indexation was also dominant, but unlike Europe, has not shown steady decline. Therefore, oil-indexation still contributed a big percentage of the pricing system [36].

2.1.3. 2008–2010

The year 2008, saw a dramatic drop in prices as the depth and spread of the global recession became apparent. Gas prices in key liberalized markets in the US fell from USD 13–14 per MMBtu in mid-2008, to a minimum of USD 4 per MMBtu in April 2009. Starting the second half of 2009, natural gas prices and crude oil prices in the US have stopped co-moving with other markets. This could be interpreted as a result of the growing production of shale gas. As an oversupply of natural gas in the Atlantic basin was created, the gas-on-gas prices dropped regardless of the global economic recession.

Prices in markets linked to oil, such as Japan and Europe, were slowed to drop from their peaks as these prices have lags of 3–6 months [37]. Greater regulatory activity was critical in opening up cross-border activity. Improvements in hub trading were important because they refined price discovery, which ensured competitively priced gas and enhanced energy security.

2.1.4. 2010–2014

Oil linked gas prices in Asia increased in 2010 onwards to above 15 USD per MMBtu, bearing the legacy of the above 100 USD oil prices. In

[4] This premium (the "Asian premium") is due to long-haul shipping of gas, high charges applied to the use of LNG terminals and lastly the absence of competition from piped gas.

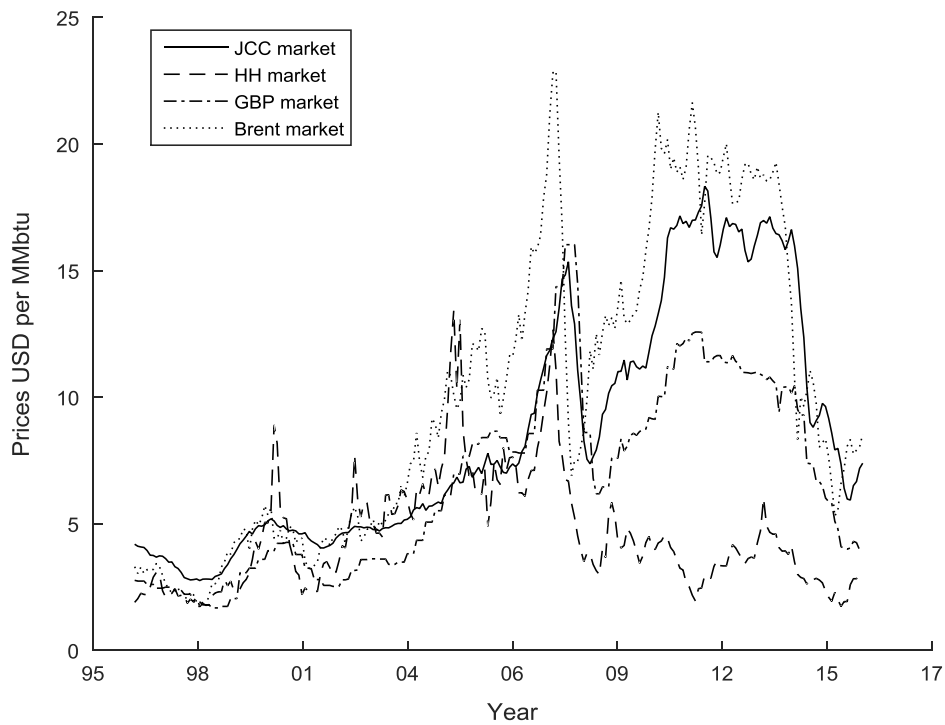

Fig. 1. Monthly Gas Prices for the three Natural Gas markets, USD/MMBtu.

Japan for example, wholesale gas prices continue to co-move with oil prices, although the indexation with oil prices may come under pressure as spot markets in Asia start to develop.

Spot markets for natural gas have grown steadily in importance in Europe. This evolution, combined with high oil prices and low demand for natural gas following the economic crisis, has led to a renegotiation of many indexed contracts, linking new contracts to spot markets (so-called "gas-to-gas pricing"). As shown in Fig. 1, gas prices in Europe have increasingly decoupled from oil prices and did not increase as strongly as oil prices between 2009 and 2011.

2.1.5. 2014-onwards

Spot LNG prices decreased significantly in 2014 and early 2015 in both Asia and Europe. The decrease is driven by weak demand in Asia, increase global supplies, and the drop in oil prices. In recent years, the average German border price followed predominately the gas to gas pricing mechanism [38].

Estimated border prices showed a clear declining trend over 2015 and early 2016. Driven by the oil price drop observed in the second half of 2014, oil-indexed prices fell faster than hub-based prices. Consequently, a significant price convergence in the third quarter of 2015 between the Asian and European gas prices was observed.

In this study, the absolute value of the difference between two consecutive observations for each market is computed, and the new vector is named X_t. Both upward and downward variations are highlighted through the study of upper records.

As we show in the results section, the gas price of the different markets during the period of 1997 and 2016, is vulnerable to price spikes/drops. Given that these sudden movements (records) are rare events, the prediction of spikes/drops is difficult. Accordingly, the stakeholders are put in a sensitive position, especially when they estimate their risk portfolio.

After describing the general trends exhibited by the three gas markets in the past couple of decades, we will explain three distinctive mathematical models which will be used to test the stability of the markets. We start with the classical model, mainly used for the American market where the observations are $i.\,i.\,d$, and then we proceed with the discrete-time random walk and the Yang models which are used for the European and Asian markets respectively. For further information regarding our work, namely the maximum likelihood estimation method and the goodness-of-fit test, the readers are invited to review annexes A and B.

2.2. Records theory and notations

In this paper $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space. X a real random variable defined on Ω with a cumulative distribution function (*CDF*) $F(\cdot)$ and a density function $f(\cdot)$. It is supposed that $(\Omega, \mathcal{F}, \mathbb{P})$ is sufficiently rich to support an infinite sequence $\{X_t,\ t \geq 1\}$ of independent copies of X; they are therefore independent and identically distributed ($i.\,i.\,d$) random variables. An observation X_t is called an upper record if it is higher than all previous observations.

The value and the occurrence index of the nth record are respectively given by the following sequences $\{R_n\}_{1\leq n\leq N_T}$ and $\{L_n\}_{1\leq n\leq N_T}$. Such that $R_n = X_{L_n}$ and N_T is the number of records in a time series of length T. In addition, we can define the sequence of record indicators $\{\delta_t\}_{1\leq t\leq T}$ which is equal to one if X_t is a record and zero otherwise. Therefore, it is easy to see that $N_T = \sum_{t=1}^{T} \delta_t$.

2.3. Independent and identically distributed (i. i. d) case

Record's properties in the case where X_t are $i.\,i.\,d$ was well studied by many author [16,17]. It turns out that many of these properties are distribution-free, i.e. they hold for any distribution of X_t. This has led to a huge progress in the overall understanding of the stochastic behavior of records. We present some of these results in order to give an overview of some important facts in the theory of records in the $i.\,i.\,d$ case.

We denote $\mathbb{P}[\delta_t = 1]$, by P_t the record rate at time t. It is the probability that the tth observation X_t , is a record. [17] shows, that for all $T \geq 1$, the random variables $\delta_1, \ldots, \delta_T$ and $M_T = \max(X_1, \ldots, X_T)$ are mutually independent with $\delta_t \sim Bernoulli\left(\frac{1}{t}\right)$, so that

$$P_t = 1/t \tag{1}$$

Note that the record rate goes asymptotically to zero, i.e. records are concentrated among the first observations. Based on the previous reasoning and by writing $N_T = \delta_1 + ... + \delta_T$ we obtain the expected value of N_T:

$$\mathbb{E}[N_T] = \sum_{t=1}^{T} 1/t \tag{2}$$

2.4. Non- independent and identically distributed case

[16] have shown, in the *i. i. d* case, that when t increases, records tend to become more spaced over time. However, in many real data sets, this phenomenon is not true for many reasons [39]. As a result, more comprehensive models were developed, where the number of records grows faster than in the *i. i. d* case, and where records are not only concentrated among the first observations as in the *i. i. d* classical model.

The most popular models beyond the *i. i. d* case are the Linear Drift Model (LDM) and the Yang model where the observations are independent but not identically distributed. On the other hand, another popular model is the Discrete-time random walk, where the observations are dependent and not identically distributed.

Before considering a higher level of complexity and going beyond the *i. i. d* model, one should test if the data fits or not the *i. i. d* case. One option is to test the null hypothesis that the data comes from a sequence of *i. i. d* random variables. This goodness-of-fit test can be based on the result of [16] who showed that under the null hypothesis the statistic

$$\mathcal{N}_T = (N_T - logT)/\sqrt{logT} \tag{3}$$

converges to a standard normal distribution denoted by $N(0,1)$.

Thus, one rejects the *i. i. d* case if $\mathcal{N}_T$ is greater than the theoretical $(1-a)^{th}$ quantile of the standard normal distribution (α is the confidence level of the statistic $\mathcal{N}_T$ generally fixed to 5%). If the null hypothesis is not accepted, then the classical *i. i. d* record model can no longer be considered, and other models that go beyond the classical models should be adopted.

2.4.1. Yang record model

The *Yang − Nevzorov* model was introduced by Ref. [40]; it was shown to fit several sets of data. In the latter model, a non-random integer number ρ_t, of *i. i. d* random variables Y with a cumulative distribution function $F(\cdot)$, is generated and available simultaneously at time t, from which is extracted $X_t = \max(y_1, y_2, ...y_{\rho t})$.

Thus, the sequence $\{X_t,\ t \geq 1\}$ of independent but not identically distributed random variables is considered, with the following cumulative distribution function

$$F_{X_t}(x) = F(x)^{\rho_t},\ \rho_t > 0 \tag{4}$$

Due to the *i. i. d* property of the underlying random variable Y, the probability of a record among the newly generated ρ_t variables is given by:

$$\mathbb{P}[\delta_t = 1] = P_t = \rho_t/S_t\ ,\ t \geq 1, \tag{5}$$

Where

$$S_t = \sum_{k=1}^{t} \rho_k \tag{6}$$

[41] shows that, in general, the expression of the previous probability holds if ρ_t are real and strictly positive. Moreover, he shows that the independence of the record indicators $\{\delta_t,\ t \geq 1\}$ remains valid for any underlying distribution. So, the sequence of δ_t is a Bernoulli process with probability of success P_t. Note that if $\rho_t = 1\ \forall\, t \geq 1$, the Yang-Nevzorov model is simply the classical *i. i. d* model.

We recall that the distribution of the number of records N_T is related to record indicators by the following relationship:

$$N_T = \sum_{t=1}^{T} \delta_t \tag{7}$$

Thus, the expected value and the variance of N_T are:

$$\mathbb{E}[N_T] = \sum_{t=1}^{T} P_t \text{ and } \mathbb{V}[N_T] = \mathbb{E}[N_T] - \sum_{t=1}^{T} P_t^2 \tag{8}$$

Now we assume the following parametric form, originally proposed by Ref. [40], $\rho_t(\gamma) = \gamma^t$ with γ strictly larger than one. Thus, the probability that the X_t, is a record for the Yang model is given by:

$$\mathbb{P}[\delta_t = 1] = P_t(\gamma) = \rho_t(\gamma)/S_t(\gamma) = \gamma^t(\gamma - 1)/\gamma(\gamma^t - 1) \tag{9}$$

Note that $\rho_t(\gamma)$ represents an exponential growth in the number of available random variables at time t. Furthermore, in a Yang model the record rate goes asymptotically to a constant given by:

$$(\gamma - 1)/\gamma \tag{10}$$

This means that the chance of having new record always exists even in the long run forecast. This case is usually encountered when the analyzed variable is more or less unstable (in other economic terms, volatile).

In order to make Yang model applicable in practice, our goal is to estimate the parameter γ, based on the maximum likelihood estimation method[5] and by using the probability distribution of the record indicators. As the record indicators δ_t are independent and follow the Bernoulli distribution of parameter $P_t(\gamma)$, our work consists in finding the γ, which maximizes the Log-Likelihood function[6]:

$$\log L(\gamma) = \log \mathbb{P}[\delta_1, ..., \delta_T; \gamma] \tag{11}$$

Still in the context of a Yang model, we define:

$$\Delta_{L_n} = L_{n+1} - L_n,\ n \geq 1 \tag{12}$$

As the inter-record time, i.e. the time recorded between two consecutive records n^{th} and the $(n+1)^{th}$. It has been shown that the inter-record time Δ_{L_n} are asymptotically geometrically distributed.

To summarize, the Yang model distinguishes itself from other models as, first it can fit observations that are independent but not identically distributed, second, the inter-record time follows a geometric distribution and last, the record rate goes asymptotically to a constant.

For the remaining of this section, the goodness-of-fit for a Yang model is evaluated by checking whether the observed values of the inter-record times, after a warm-up period allowing the asymptotic effect to settle, are in agreement with the geometric distribution.

The most popular approach consists in constructing a goodness-of-fit test based on Pearson's chi-square test [42]. It is assumed that the event $N_T = m$ is not random and we chose a $K > 1$ in order to partition the set $\{1,2, ..., \infty\}$ into K subsets. Concerning the problem of choosing K, this is a question that users have asked statisticians at the very moment when the chi-square test appeared [43–47]. Few of these responses reinforce each other, while others contradict. However none appears to be universally accepted. Recently [48], write that the problem must be solved ad hoc in relation to contextual elements as a serious and rigorous user of this test will not fail to obtain.[7]

If the Pearson chi-square test rejects the geometric distribution hypothesis, this casts doubt on Yang's model. Then, we should consider a more general model where the observations are dependent and not identically distributed.

[5] The maximum likelihood estimation is a method of estimating the parameters of a statistical model.

[6] Refer to annex A, for further information regarding the maximum likelihood method.

[7] Refer to annex B, for further information regarding the goodness-of-fit test of the Yang model.

The simplest way to check if the observations are independent is to use the $Ljung - Box$ test, and if equals to unity, indicates that the observations belong to a dependent set of data.

2.4.2. Discrete–time random walk model (DTRW)

One can think of a more general model, where the observations are dependent and not identically distributed. The most used model in this context is the Discrete-time random walk model introduced by Ref. [49].

In this model, the tth observation of the upper record is given by

$$X_t = X_{t-1} + \eta_t, \tag{13}$$

where η_t i.i.d increments are drawn from a continuous distribution.

The stochastic behavior of this model was carefully studied, and it has been shown that the record rate is given by:

$$\mathbb{P}[\delta_t = 1] = P_t = \binom{2t}{t} 2^{-2t}, \tag{14}$$

In this model, all the above properties are in a distribution-free context (independent from the choice of the distribution of X_t). In addition, it is remarked that for a big value of t, the record rate goes asymptotically to zero. This means that records are concentrated among the first observations.

In the next section, results of the different tests that were used in the models are presented. The analysis and choice of models is explained and validated based on economical and mathematical reasoning.

3. Results and discussion

3.1. Results of the three models

The analysis starts by extracting the record values, indices and indicators, as well as the number of records that each market has witnessed within the period of March 1997 to November 2016.

We start by applying the goodness of fit test, which tests the null hypothesis that the data comes from a sequence of $i.\ i.\ d$ random variables. As shown in Table 3, both the German and Asian market rejects the null hypothesis, which explains empirically that those markets are struggling with sudden price variations going beyond the $i.\ i.\ d$ case.

This result can be expected based on the analysis of Table 2, where the American market has the lowest number of records, while the Asian market has the highest number of records. Worthy to note, the record are concentrated among the first observations in the results of the American market, which is typical for the $i.\ i.\ d$ case.

Table 2
Number of records and record index.

Markets	HH	GBP	JCC
Number of records	8	10	17
Record index	Apr-97	Apr-97	Apr-97
	May-97	Jul-97	Jul-97
	Aug-97	Oct-97	Aug-97
	Sep-97	Jul-98	Apr-98
	Dec-97	Oct-98	Jul-98
	Jun-00	Oct-99	Aug-99
	Dec-00	Jan-00	Oct-99
	Jan-06	Jan-08	Sep-04
		Jul-08	Jul-05
		Mar-09	Feb-06
			Oct-07
			Jan-08
			Aug-08
			Sep-08
			Dec-08
			Feb-09
			Apr-15

Table 3
Goodness of fit test results at a confidence level of 5%.

Markets	HH	GBP	JCC
P-value	0.139	0.026	4e-007
Result	Accept H_0	Reject H_0	Reject H_0

Table 4
Pearson chi-squared test results at a confidence level of 5%.

Markets	GBP	JCC
P-value	0.127	0.358
Result	Accept H_0	Accept H_0

This is in line with mathematical theory, which suggests that the number of records for $non - i.\ i.\ d$ case grows faster than in the $i.\ i.\ d$ case.

We now proceed, to the less stable markets, where the classical models cannot be adopted.

Note that the assumption of a Yang model is increasingly consistent, when the p-value of the $\chi(\tilde{\gamma})$ statistic is increasingly greater than the confidence level α generally fixed at 5%. The results shown in Table 4, suggest that both markets accept the null hypothesis, which means that the Yang model can be applied to both the JCC and GBP markets. However the GBP market has a p-value which is not considerably high. Accordingly, we explored other general models where the observations are dependent and not identically distributed.

In order to test whether the observations in the GBP market come from a dependent set of data, the $Ljung - Box$ test is used, with the aim of knowing if the observations are auto-correlated. The result of the $Ljung - Box$ is equal to unity. This implies that the series of GBP belongs to a dependent set of data, which is common to see in a commodity that exhibits lots of volatility. Consequently the Random walk model, can be applied for the GBP market.

In addition to the mathematical explanation, we can demonstrate that GBP exhibits more volatility in economic reasoning:

First, gas in continental Europe is slowly, but surely, heading towards a gas to gas pricing rule (The latest report of IGU states that over 60% of traded gas in Europe is indexed to the continental gas hubs). Therefore, this could bring more volatility to the market.

Second, the supply portfolio in such markets is an important factor. Thus, all supply sources coming from regions that show sign of volatility or insecurity or instability could increase volatility at times where the system does not work properly (excess of gas supply, implies less volatility and vice versa).

Last but not least, the high level of storage capacity in European gas markets would likely have a strong and important effect on price volatility as it will tend to decrease, mainly because it will increase the strategic seasonal supply of gas. However, with no diversified, secured and long lasting gas supply to Europe, this could make the presence of storage capacity unexploited.

The probability of having a record can be computed for each market and for any time in the near future. Table 5, shows the result of our mathematical models. The probabilities were computed for the observation that coincides with the date of November 2013.

The probability of having a new record is highest in Asian markets. This provide evidence that the theoretical reasoning used in the study, is precise. The following are the main outcomes:

1) The probability of having a record is low for all markets. Unless a major rupture in supply and demand fundamentals occurs the probability of a major spike/drop is not significant.
2) Viewed as increasingly scarce a decade ago, when its price rose above $15 per million Btu (Fig. 1), the substantial shale gas exploitation in the US, led to a resource surplus and substantial low

Table 5
Probability of Records for each market.

Markets	Probability of records	Probability of having a record on t = 200 (November 2013)
HH	$\mathbb{P}[\delta_t = 1] = P_t = \frac{1}{t}$	$\mathbb{P}[\delta_t = 1] = 0.005$
JCC, where γ is equal to 1.058	$\mathbb{P}[\delta_t = 1] = P_t(\gamma) = (\gamma^t(\gamma-1))/(\gamma(\gamma^t - 1))$	$\mathbb{P}[\delta_t = 1] = 0.0555$
GBP	$\mathbb{P}[\delta_t = 1] = P_t = \binom{2t}{t} 2^{-2t}$,	$\mathbb{P}[\delta_t = 1] = 0.0399$

prices. The result of the classical model that is applied to the *i. i. d* case is in line with this economic fact, and the probability of having a record tends to zero in the American Market (HH). This means that there is little chance in encountering a price/drop in gas prices in the short term, a result also endorsed by recent literature [50].

3) There is always a certain probability of having a record in the European market and the Asian market. Again, the Yang and Discrete-time random walk models performed well. This is in line with economic reasoning for both markets.

Research papers published by Refs. [51,52], have increasingly observed that the move from oil-indexed to hub or market pricing is a clear secular trend, strongest in northwest Europe and spreading southwards and eastwards. The strong gas markets integration between the different European hubs and, the diversification in gas supply will keep the gas prices resistant to sudden spikes/drops. This is the result of many years of pro-competition EU regulatory initiatives implied by the several EU directives, namely the legislation concerning security of supply at EU level (994/2010). The annual surveys on pricing of wholesale gas undertaken by the International Gas Union also lend quantitative evidence of these trends [53].

On the other hand, Asian gas prices are still dominated by oil indexation of LNG contracts, which cannot follow gas market fundamentals in a timely manner. This economic reasoning is in line with recent publications, particularly for the Asian market. Research showed, that the Asian Pacific LNG is currently undergoing considerable change and uncertainty, and the risk of unexpected oil price shocks has always been the main factor in analyzing LNG trade and market interactions in the Asian gas market [34,50,54,55].

4) Results suggest that the Yang model fit the JCC market, and that the result of the maximum likelihood principle based on record indicators (described in Section 2.4.1) gives a value for $\hat{\gamma}$ to be equal to 1.058. As $\hat{\gamma}$ exceeds unity, the choice for the Yang model was reasonable.
5) The record rate in some of the models converges to a certain constant in the short run. This is a significant indicator showing that the markets of these models are in an unstable situation and vulnerable to future spikes/drops.

3.2. Difference between empirical and theoretical findings

The analysis carried in this study, started by applying several goodness-of-fit tests in order to assign each set of data to the model that fits observations the most.

In a second step, the probability of the records for each model was computed.

In a final step, in order to test the models, we computed the theoretically expected number of records for each model based on the expected value of N_T, and compared it to the real number of records extracted directly from the series of observations.

By completing all three steps, we evaluate the forecasting performance of each model. The closeness between the theoretical and empirical results is a proof of good performance.

Table 6
Results of empirical and theoretical probabilities.

Markets	HH	GBP	JCC
Number of records (Actual)	8	10	17
Number of records (Theoretical)	6.043	16.36	16.47
Percentage error (%)	24	63	3

As shown in Table 6, the percentage error is low for both the HH and JCC markets. This is perceived as a good indicator regarding the choice of our models: they fit the considered data sets.

The percentage error in the JCC market is minimal, which proves that the Yang model has successfully predicted the number of records in the last twenty years. Since the record rate in a Yang model goes asymptotically to a constant, this is further proof that the Asian market is unstable in the short and long run prediction and will always be vulnerable to spikes and drops in natural gas prices.

The gas regulator in the relevant Asian countries can benefit from our theoretical findings as basis to develop, support and improve business environment for developing a better functioning LNG market. The scientific and political community have shared policy recommendation to improve the situation in the Asian gas market, including:

- Liberalizing domestic gas market and developing adequate and accessible gas infrastructure capacities by promoting third party access[8] [56].
- Moving away from JCC to a price mechanism which reflects anticipated market fundamentals of the Asian buyers' country[9] [57].
- Securing new gas sources, such as East Africa and Australia to diversify the supply portfolio of Asian importers.
- Promoting governmental adequate and accessible infrastructure developments (upstream/downstream) through: equity participation of importers/Asian shippers in LNG projects, investment in storage in order to increase domestic flexibility, and investment in new Asian upstream projects.

Consequently, we have successfully modeled and used the records theory on a data set that goes beyond the *i. i. d* case without going through the conditional approach which is mostly used in EVT. In addition, the results are distribution free, which minimize the complexity of the models.

4. Conclusion

To date the literature studying the tails of the distribution, have used the extreme value theory to study the price behavior of commodity prices. This research; however offers an alternative and concentrates on

[8] The TPA should be complemented by enhancing contractual flexibility. This can done by eliminating destinations restrictions in LNG contracts.

[9] The imperative for buyers will be not to lock themselves into long term inflexible price arrangements during a period when market dynamics will be changing rapidly.

the records theory.

Our findings suggests that out of the three main regional gas markets, the Asian market seems to be less stable than the others, and that the probability of having a record in the coming years is the highest. In addition, there is always a probability of having a record in the long run for the Asian market, which proves that the market is not stable, unlike the American and European markets which seems to be more stable in both the short and long run.

While this study offers binding mathematical models, it does also reveal concrete results that are distribution free. Most importantly the validity of the models/results is validated in the two step analysis we have conducted, as both the empirical and theoretical findings seems to be close for all three models, especially the Asian market.

The analytical framework that is based on distinctive mathematical models described in the paper, and the ability to forecast future spike/drops should be used as an incentive, by Asian gas regulators, where there are signs of price instability.

While the private sector is mainly responsible for commercial deals, the public sector is encouraged to support and improve business environment for developing well-working rules of gas markets conductive to a better functioning LNG market, especially in terms of flexibility, price formation, gas supply security, and securing necessary investment.

Worthy to mention, the record models will work best when used to predict the probability of a record in the short run. As known that in volatile commodity price forecasting, models should be updated frequently, because of the vulnerability of the market to sudden disruptions in supply and demand fundamentals, which needs to be reflected in the data.

The study used three record models to explain the different sets of data. Pursuing further research and considering other record models for future research, such as Linear Drift Model and Increasing Variance Model, is needed. Record values were not computed in this analysis. In the future, researchers could study the joint distribution of record values based on a Markov chain analysis, which can provide more information to our inference procedure. Using this information, it is possible to predict future record values with a given confidence interval.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.esr.2018.05.003.

Annex A. Maximum likelihood estimation method

Maximizing the Log-Likelihood function

$$\log L(\gamma) = \log \mathbb{P}[\delta_1, ..., \delta_T; \gamma]$$
$$= N_T \log(1 - 1/\gamma) + (T - N_T)\log(1/\gamma) - \log(1 - 1/\gamma^T) - \sum_{t=2}^{T} \delta_t \log(1 - 1/\gamma^{t-1})$$

Thus, we must find the value $\hat{\gamma}$ of γ, which denotes our estimator by the maximum likelihood method, such that:

$$(d \log L(\gamma)/d\gamma)_{\gamma=\hat{\gamma}} = 0 \tag{15}$$

The asymptotic behavior of this estimator is:

$$(\hat{\gamma} - \gamma)/\sqrt{I_T^{-1}(\gamma)} \to N(0,1), \tag{16}$$

where $I_T(\gamma)$ denotes Fisher information. Note that the behavior of this estimator is also distribution-free. The importance of calculating the asymptotic behavior is double folded: It shows that the behavior of the estimator is distribution-free, which means independent from the choice of the underlying distribution Y. In addition, it gives the possibility to construct the confidence intervals of the parameter γ for a given confidence level alpha.

Annex B. Goodness-of-fit for a Yang model

In the following section N_T denotes the number of records after the warm-up period and T the present time.

On the basis of a partition (of disjoint sub-sets) $\Pi_1 \cup ... \cup \Pi_K$ of the set $\{1,2, ..., \infty\}$, n_k, $1 \le k \le K$ denotes the number of Δ_{L_n} which fall within Π_k with $n_1 + ... + n_K = m - 1$ (Because for $N_T = m$ records it corresponds $m - 1$ inter-records). In addition, we denote $\pi_k(\gamma) = \sum_{j \in \Pi_k} p_j(\gamma)$, $1 \le k \le K$. The Pearson chi-square statistic is given by:

$$\chi(\gamma) = \sum_{k=1}^{K} (n_k - (m-1)\pi_k(\gamma))^2/((m-1)\pi_k(\gamma)) \tag{17}$$

The value of this statistic is compared to $x^2_{K-1,\,1-\alpha}$ the quantile of order $(1 - \alpha)$ of the chi-square with $K - 1$ degrees of freedom, denoted χ^2_{K-1}. When $\chi(\gamma) > x^2_{K-1,\,1-\alpha}$, the test rejects the hypothesis H_0 at the confidence asymptotic level α. Thus, the observed values of Δ_{L_n} are not in agreement with the geometric distribution and the model does not fits Yang.

However, the statistic $\chi(\gamma)$ is unusable because the parameter γ is unknown. We must therefore estimate it. To do this, we calculate the value $\tilde{\gamma}$ that minimizes $\chi(\gamma)$. Then, the usable statistic is

$$\chi(\tilde{\gamma}) = \underset{\gamma}{argmin}\chi(\gamma) \tag{18}$$

According to a classical result [47], if the data comes from a geometric distribution, the statistic $\chi(\tilde{\gamma}) \to \chi^2_{K-2}$ converges in distribution to a chi-square with $K - 2$ degrees of freedom. If $\chi(\tilde{\gamma})$ exceeds $x^2_{K-2,\,1-\alpha}$ (the quantile of order $1 - \alpha$ of the chi-square with $K - 2$ degrees of freedom), the test rejects, at the confidence level α, the null hypothesis that the Δ_{L_n} follow a Geometric distribution, which casts doubt on Yang's model.

References

[1] International Gas Union, World LNG Report, Int. Gas union world gas LNG Rep 88 (2016) http://www.igu.org/publications.
[2] EIA, Gas Strategies LNG Data Service, (2015).
[3] Reuters, Timera Energy, (2012).
[4] V. Marimoutou, B. Raggad, A. Trabelsi, Extreme value theory and value at risk: application to oil market, Energy Econ. 31 (2009) 519–530, http://dx.doi.org/10.1016/j.eneco.2009.02.005.
[5] A.J. McNeil, R. Frey, Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach, J. Empir. Finance 7 (2000) 271–300, http://dx.doi.org/10.1016/S0927-5398(00)00012-8.
[6] H.N.E. Byström, Managing extreme risks in tranquil and volatile markets using conditional extreme value theory, Int. Rev. Financ. Anal. 13 (2004) 133–152, http://dx.doi.org/10.1016/j.irfa.2004.02.003.
[7] Qinlu Chen, David E. Giles, Risk Analysis for Three Precious Metals: An Application of Extreme Value Theory, Econometrics Working Papers 1704 Department of Economics, University of Victoria, 2017, https://ideas.repec.org/p/vic/vicewp/1704.html.
[8] K. Fong Chan, P. Gray, Using extreme value theory to measure value-at-risk for daily electricity spot prices, Int. J. Forecast. 22 (2006) 283–300, http://dx.doi.org/10.1016/j.ijforecast.2005.10.002.
[9] R. Herrera, A. Rodriguez, G. Pino, Modeling and forecasting extreme commodity prices: a Markov-Switching based extreme value model, Energy Econ. 63 (2017) 129–143, http://dx.doi.org/10.1016/j.eneco.2017.01.012.
[10] Q. Liu, I. Wong, Y. An, J. Zhang, Asymmetric information and volatility forecasting in commodity futures markets, Pac. Basin Finance J. 26 (2014) 79–97, http://dx.doi.org/10.1016/j.pacfin.2013.10.007.
[11] C. Aloui, R. Jammazi, The effects of crude oil shocks on stock market shifts behaviour: a regime switching approach, Energy Econ. 31 (2009) 789–799, http://dx.doi.org/10.1016/j.eneco.2009.03.009.
[12] C. Aloui, S. Mabrouk, Value-at-risk estimations of energy commodities via long-memory, asymmetry and fat-tailed GARCH models, Energy Pol. 38 (2010) 2326–2339, http://dx.doi.org/10.1016/j.enpol.2009.12.020.
[13] W. Chkili, S. Hammoudeh, D.K. Nguyen, Volatility forecasting and risk management for commodity markets in the presence of asymmetry and long memory, Energy Econ. 41 (2014) 1–18, http://dx.doi.org/10.1016/j.eneco.2013.10.011.
[14] R. Aloui, M.S. Ben Aïssa, D.K. Nguyen, Conditional dependence structure between oil prices and exchange rates: a copula-GARCH approach, J. Int. Money Finance 32 (2013) 719–738, http://dx.doi.org/10.1016/j.jimonfin.2012.06.006.
[15] E. Lindström, F. Regland, Modeling extreme dependence between European electricity markets, Energy Econ. 34 (2012) 899–904, http://dx.doi.org/10.1016/j.eneco.2012.04.006.
[16] Barry C. Arnold, N. Balakrishnan, Haikady N. Nagaraja, The Comprehensive Guide to Modern Record Theory, John Wiley & Sons, New York, 1998https://www.wiley.com/en-us/Records-p-9780471081081.
[17] Nevzorov, B. Valery, Records-mathematical Theory, American Mathematical Society, Rhode Island, 2001.
[18] Rocco Ballerini, Sidney I. Resnick, Embedding sequences of successive maxima in extremal processes, with applications, J. Appl. Probab. 24 (1987) 827–837.
[19] O. Efimova, A. Serletis, Energy markets volatility modelling using GARCH, Energy Econ. 43 (2014) 264–273, http://dx.doi.org/10.1016/j.eneco.2014.02.018.
[20] X. Mu, Weather, storage, and natural gas price dynamics: fundamentals and volatility, Energy Econ. 29 (2007) 46–63, http://dx.doi.org/10.1016/j.eneco.2006.04.003.
[21] B. Ewing, F. Malik, O. Ozfidan, Volatility transmission in the oil and natural gas markets, Energy Econ. 24 (2002) 525–538, http://dx.doi.org/10.1016/S0140-9883(02)00060-9.
[22] B. Lin, P.K. Wesseh, What causes price volatility and regime shifts in the natural gas market, Inside Energy 55 (2013) 553–563, http://dx.doi.org/10.1016/j.energy.2013.03.082.
[23] W.M. Fong, K.H. See, A Markov switching model of the conditional volatility of crude oil futures prices, Energy Econ. 24 (2002) 71–95, http://dx.doi.org/10.1016/S0140-9883(01)00087-1.
[24] F. Fallahi, Causal relationship between energy consumption (EC) and GDP: a Markov-switching (MS) causality, Inside Energy 36 (2011) 4165–4170, http://dx.doi.org/10.1016/j.energy.2011.04.027.
[25] S. Nick, S. Thoenes, What drives natural gas prices? - A structural VAR approach, Energy Econ. 45 (2014) 517–527, http://dx.doi.org/10.1016/j.eneco.2014.08.010.
[26] T. Panagiotidis, E. Rutledge, Oil and gas markets in the UK: evidence from a co-integrating approach, Energy Econ. 29 (2007) 329–347, http://dx.doi.org/10.1016/j.eneco.2006.10.013.
[27] B. Siliverstovs, G. L'Hégaret, A. Neumann, C. von Hirschhausen, International market integration for natural gas? A cointegration analysis of prices in Europe, North America and Japan, Energy Econ. 27 (2005) 603–615, http://dx.doi.org/10.1016/j.eneco.2005.03.002.
[28] C. Gerda, H. Nils, Model Selection and Model Averaging, Cambridge Series in Statistical and Probabilistic Mathematics, (2008).
[29] F. Harrell, Regression Modeling Strategies, Springer, 2015, https://www.springer.com/gp/book/9783319194240.
[30] S. Busse, P. Helmholz, M. Weinmann, Forecasting day ahead spot price movements of natural gas - an analysis of potential influence factors on basis of a NARX neural network, Multikonferenz Wirtschaftsinformatik 2012-Tagungsband Der MKWI, 2012, p. 2012.
[31] A. Khotanzad, H. Elragal, T.L. Lu, Combination of artificial neural-network forecasters for prediction of natural gas consumption, IEEE Trans. Neural Network. 11 (2000) 464–473, http://dx.doi.org/10.1109/72.839015.
[32] N. Salehnia, M.A. Falahi, A. Seifi, M.H. Mahdavi Adeli, Forecasting natural gas spot prices with nonlinear modeling using Gamma test analysis, J. Nat. Gas Sci. Eng. 14 (2013) 238–249, http://dx.doi.org/10.1016/j.jngse.2013.07.002.
[33] P. Economist, World LNG Database, (2015).
[34] T.N.C. Tom Kompas, A structural and stochastic optimal model for projections of LNG imports and exports in Asia-Pacific, Heliyon 2 (2016).
[35] Platts, No Title, (2017) https://www.platts.com/commodity/natural-gas.
[36] I.E. Agency, No Title, Paris, (2013).
[37] R. Davoust, Gouvernance europeene et geopolitique de l'energie, Bruxelles, (2008), p. 12 http://www.iaea.org/inis/collection/NCLCollectionStore/_Public/42/050/42050176.pdf.
[38] D. Energy, Quarterly Report on European Gas Markets, Bruxelles, (2016), p. 17 https://ec.europa.eu/energy/sites/ener/files/documents/quarterly_report_on_european_gas_markets_q3_2015.pdf.
[39] A.S. Hoayek, G.R. Ducharme, Z. Khraibani, Distribution-free inference in record series, Extremes 20 (2017) 585–603, http://dx.doi.org/10.1007/s10687-017-0283-7.
[40] M.C.K. Yang, On the distribution of the inter-record times in an increasing population, J. Appl. Probab. 12 (1975) 148–154, http://dx.doi.org/10.2307/3212417.
[41] Nevzorov, Records for nonidentically distributed random variables, Proc. Fifth Vilnius Conf. Probab. Stat, 1990, pp. 227–233.
[42] X. Karl Pearson, On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling, London, Edinburgh, Dublin Philos. Mag. J. Sci 50 (1990).
[43] A.W.H.B. Mann, On the choice of the number of class intervals in the application of the chi square teste, Ann. Math. Stat. (1942) 306–317.
[44] T. Inglot, T. Ledwina, Asymptotic optimality of data-driven Neyman's tests for uniformity, Ann. Stat. 24 (1996) 1982–2019, http://dx.doi.org/10.1214/aos/1069362306.
[45] J.C.W. Rayner, D.J. Best, The choice of class probabilities and number of classes for the simple $\chi ^{2}$ goodness of fit test, Indian J. Stat 44 (1982) 28–38 http://www.jstor.org/stable/25052234.
[46] F.C. Drost, A limit theorem for some modified chi-square statistics when the number of classes increases, Math. Stat. Probab. Theory (1987) 43–58.
[47] Y.M. Bishop, S.E. Fienberg, P.W. Holland, Discrete Multivariate Analysis Theory and Practice, (2013), http://dx.doi.org/10.1017/CBO9781107415324.004.
[48] J.C.W.R. Glen, D. Rayner, Power of the Neyman smooth tests for the uniform distribution, J. Appl. Math. Decis. Sci. 5 (2001) 181–191.
[49] S.N. Majumdar, R.M. Ziff, Universal record statistics of random walks and lévy flights, Phys. Rev. Lett. 101 (2008), http://dx.doi.org/10.1103/PhysRevLett.101.050601.
[50] D. Zhang, Q. Ji, Further evidence on the debate of oil-gas price decoupling: a long memory approach, Energy Pol. 113 (2018) 68–75, http://dx.doi.org/10.1016/j.enpol.2017.10.046.
[51] Heather Patrick, The Evolution of European Traded Gas Hubs, Oxford Inst. Energy Stud., 2015, https://www.oxfordenergy.org/wpcms/wp-content/uploads/2016/02/NG-104.pdf.
[52] P. Beatrice, The Cost of Price De-Linkages between European Gas Hubs, Oxford Inst. Energy Stud., 2015, https://www.oxfordenergy.org/wpcms/wp-content/uploads/2015/09/NG-101.pdf.
[53] International Gas Union, A Global Review of price Formation Mechanisms, Barcelona, 2017.
[54] X. Shi, H.M.P. Variam, East Asia's gas-market failure and distinctive economics—a case study of low oil prices, Appl. Energy 195 (2017) 800–809, http://dx.doi.org/10.1016/j.apenergy.2017.03.091.
[55] T. Kompas, T.N. Che, Economic Analysis of Cost Competitiveness of Liquified Natural Gas Exporters to the Asia-Pacific Region, (2015).
[56] Economic Research Institute for ASEAN and East Asia, Recommendations for a Better Functioning LNG Market in Asia, Inst. Electr. Eng. Japan, 2015, p. 5 https://eneken.ieej.or.jp/data/6302.pdf.
[57] J. Stern, International gas pricing in Europe and Asia: a crisis of fundamentals, Energy Pol. 64 (2014) 43–48, http://dx.doi.org/10.1016/j.enpol.2013.05.127.