Energy 239 (2022) 121962

Contents lists available at ScienceDirect

Energy

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

Energy markets — Who are the influencers?

Paulo Ferreira [a, b, c, *], Dora Almeida [c], Andreia Dionísio [c], Elie Bouri [d], Derick Quintino [e]

[a] VALORIZA - Research Center for Endogenous Resource Valorization, 7300-555, Portalegre, Portugal
[b] Instituto Politécnico de Portalegre, 7300-110, Portalegre, Portugal
[c] CEFAGE-UE, IIFA, Universidade de Évora, Largo Dos Colegiais 2, 7000-809, Évora, Portugal
[d] Adnan Kassar School of Business, Lebanese American University, P.O. Box 13-5053, Chouran Beirut, 1102 2801, Lebanon
[e] Department of Economics, Administration and Sociology, University of São Paulo, 13418-900, Piracicaba, SP, Brazil

ARTICLE INFO

Article history:
Received 12 April 2021
Received in revised form
22 July 2021
Accepted 30 August 2021
Available online 4 September 2021

Keywords:
Energy markets
Clean and dirty energy assets
Granger causality
Transfer entropy
Influencer

ABSTRACT

The energy markets have recently undergone important transformations (e.g. deregulation, technological progress, renewable energy deployment and changing energy consumer behaviour) and witnessed a variety of crisis periods, affecting the relationships among energy commodities and their interactions with clean energy indices. This has implications for price discovery, asset allocation and risk management, which requires in-depth analysis to uncover and identify which energy indices (or forms of energy) lead others or are the most influential, while accounting for asymmetry and non-linearity characteristics. To uncover the complex structure of the relationship across the returns of seven different energy commodities and two clean energy stock indices, we apply Granger causality and transfer entropy in both static and dynamic approaches. The results from the Granger causality analysis identify the influence of the other energy products on natural gas, whereas the transfer entropy analysis reveals the importance of WTI oil and the influence of clean energy indices. Diesel is the most influenced energy commodity. A rolling windows analysis confirms those findings and shows evidence of a time-variation that reflects the impacts of crisis periods, especially the pandemic, on the dynamics of relationships.

1. Introduction

The energy markets have undergone important changes and transformations over the past decades arising from the emergence of challenging factors and market developments such as market deregulation, technological advances, the shale gas revolution [1] and renewable energy deployment. Such transformations influence investors' sentiment, market returns and conditions, especially if they are combined with occasional shocks and crisis periods (e.g., the 2007–2008 global financial crisis; the oil price war between Saudi Arabia and Russia in 2014–2015; and the COVID-19 pandemic[1]), and this can disrupt the dynamics of spillovers in the energy markets and make it more complex. They are important for the nature of the return spillovers as well as the individual role of each energy commodity within the structure of return spillovers [2] and have important implications for price discovery, asset allocation and risk management. Previous studies have indicated that through the channels of the financial system, it is possible for information shocks associated with either the demand or supply side to be transmitted to various energy commodities, leading to a distortion and complexity in the network of the relationships among energy commodities (see, for example, [2–4]. In fact, it is often difficult for one energy commodity to resist the shock spillover faced by another energy commodity; this is accentuated by the fact that market participants consider various energy commodities as alternative investment choices [5]. In this regard, the academic literature points to the competitive substitution relationship between fossil energy and clean energy [6], with the argument that higher prices of fossil energy make investments in clean energy more appealing which leads to an increase in the price of energy stocks [7]. However, a competitive substitution relationship has been refuted, suggesting the significance of a decoupling hypothesis [8,9] based on the rationale that crude oil and clean energy assets compete in different markets (i.e., crude oil is used to produce transportation fuel whereas clean energy is

* Corresponding author. VALORIZA - Research Center for Endogenous Resource Valorization, 7300-555, Portalegre, Portugal.

E-mail addresses: pferreira@ipportalegre.pt (P. Ferreira), dmfa1982@gmail.com (D. Almeida), andreia@uevora.pt (A. Dionísio), elie.elbouri@lau.edu.lb (E. Bouri), derickdq@usp.br (D. Quintino).

[1] During the March 2020 outbreak peak, WTI oil prices declined to negative territories for the first time.

https://doi.org/10.1016/j.energy.2021.121962
0360-5442/

mainly used to produce electricity). Some other studies consider clean energy and differentiate between dirty and clean energy investments, highlighting the diversification benefits of adding clean energy investment to commodities in general and to energy commodities in particular [10,11]. However, these studies provide little evidence about the identity of the key influencers in the energy markets covering both clean and dirty energy assets, which is crucial in light of the emergence of various events and crisis periods and the potential asymmetry and nonlinearity features. Addressing this research gap, while considering stylized facts of energy asset returns such as asymmetry and non-linearity, is important to have a better understanding of the identity of the main players and influencers in the energy markets and make inferences for the sake of investors, portfolio managers and policymakers.

Motivated by the above background, the aim of this research paper is to ascertain the main influencers in the wide universe of energy markets covering both dirty energy commodities and clean energy assets, assuming its potential complex and non-linear relationships. Using daily data from November 25, 2003 to December 30, 2020, we based our analysis on a large set of various energy commodities (seven energy commodities, namely natural gas, WTI crude oil, NY gasoline, Gulf gasoline, diesel, heating oil and propane, and two clean energy indices, namely the S&P Global Clean Energy Index and the Renewable Energy Industrial Index). This includes energy commodities which have been less explored in previous literature such as heating oil and propane, and the appealing and environmentally-friendly clean energy stock indices. Methodologically, we apply complementary econometric approaches involving Granger causality and transfer entropy. The latter allows us to overcome the linearity of the traditional Granger causality approach, which is important to account for the presence of non-linear dynamics in energy assets.

While our paper is related to a strand of literature dealing with the return relationship across energy commodities (e.g. Refs. [11–13] or between clean energy indices and “dirty” energy commodities (e.g. Refs. [8,14–17], it differs in several aspects and makes the following contributions. Firstly, we consider the return flow between clean energy indices and “dirty” energy commodities, while extending the analysis of traditional energy commodities to now include other commodities (i.e. heating oil, propane). Secondly, we apply new analytical methodologies involving Granger causality and transfer entropy. Notably, transfer entropy can capture the non-linear dynamics and possible asymmetric information flow between the energy indices under study. More specifically, it is a directional and dynamic measure of predictive information rather than a measure of the causal information flow from a source variable and to a destination variable [18]. Thirdly, our sample period covers the COVID-19 outbreak, during which the stability of financial markets was disturbed and the universe of energy investments adversely affected by the lockdown and the negative prices of WTI oil.

Our main results show that WTI oil and clean energy are the most influential players, while diesel seems to be the most influenced by others. Further analyses provide evidence of time-varying behaviour in the dynamics of relationships between energy assets and point to the disturbing impacts of the COVID-19 outbreak on the relationships.

The rest of the paper is organized as follows: Section 2 provides a literature review, Section 3 describes the data and methodology, and Section 4 presents and discusses the empirical results. Finally, Section 5 presents the conclusions.

2. Literature review

The study of interdependent relationships across different energy commodities has been an appealing area of research for scholars and market participants, and stems from the important role played by energy commodities for international trade, economic activities, accumulation of wealth, and portfolio diversification. The supply and demand theory has been used to explain energy commodity prices. It implies that demand-side factors such as global economic conditions and global real activity are major determinants of energy commodity demand and thereby of energy commodities prices. In this regard, the rapid economic growth in emerging economies such as China and India plays a key role. Furthermore, economic and financial crisis periods such as the 2008 global financial crisis can disrupt economic activity and lead to a decrease in the demand for energy committees. Regarding supply-side factors, production and weather can shape energy commodity supply and thereby energy commodity prices. In the energy markets, US crude oil production and inventory accumulation influence energy prices, especially crude oil prices. However, the financialization phenomenon has facilitated the investability of commodities, including energy commodities, making them accessible to speculators and easily added to multi-asset portfolios [19]. Notably, the associated speculation activity resulting from that phenomenon, as reflected in the index positions and managed-money spread positions, has contributed to energy prices drifting away from fundamental values and intensifying booms and busts and the degree of integration across various energy markets.

The energy sector contains a rich set of energy commodities that do not necessarily exhibit homogenous relationships between each other but rather heterogeneous relationships that can be shaped by crisis periods. The central emphasis of previous studies is the crude oil market and its relationship with the markets of natural gas and clean energy. Much less evidence exists on the relationship between each of those three markets and the rest of energy commodities such as heating oil, gasoline and propane, a relationship which is extended in this paper.

Earlier studies consider the crude oil–natural gas nexus, indicating that these two energy commodities are substitutes in the industrial and power generation sectors [20] as well as in the sectors of transportation and energy consumption, such as heating [21]. Notably, the causal relationship between oil and gas markets is shown to be unstable [12]. In this regard, the empirical results are often mixed, showing a weak cointegration relationship or instability in the price relationship [22] and bidirectional spillover effects between crude oil and natural gas in North America in the medium term ([13]. The shale gas revolution seems to disrupt the long-term oil and gas relationship in the US and there is evidence that the crude oil market dominates natural gas in terms of information transmission (e.g., Ref. [13]. The rationale that crude oil prices are exogenous to natural gas prices resides on the fact that crude oil prices are determined by global supply and demand factors whereas prices of natural gas are more subject to regional and local market conditions (e.g., Ref. [23]. For example, Jadidzadeh and Serletis [1] apply a structural vector autoregressive (VAR) model and find that more than two-thirds of price changes in the US natural gas market can be explained by the structural supply-and-demand influence of international oil prices.

Other studies extend the literature on the crude oil–natural gas nexus by considering other energy commodities while applying multivariate GARCH models. Ahmad and Rais [10] examine spillover effects and portfolio implications for clean energy stocks, technology stocks and four energy commodities: crude oil, Brent crude, gasoline and heating oil. They show evidence of a weak relationship between clean energy stocks and energy commodities, while highlighting the diversification and hedging possibilities. More recently, Asl et al. [11] show weak return linkages between clean energy stocks energy commodities and a stronger return between energy stocks and heating oil, gasoline, crude oil and propane. Natural gas prices are found to be segmented from the other markets under study and results from a portfolio analysis show potential diversification benefits of heating oil for energy stocks and evidence that clean energy and non-clean energy stocks are effective hedges for each other. Saeed et al. [16] assess whether green bonds and clean energy stocks can play as hedging instruments against the downside risk of dirty energy assets. They imply that the degree of conditional correlation varies over time, in the sense that hedgers must act dynamically and that clean energy assets are more effective hedges compared to green bonds, especially with regard to crude oil.

Given the importance of the transition from fossil to clean energy for the sake of energy sustainability, there has been a shift in the academic literature towards clean energy investments and the relationship with crude oil and other dirty energy investments (e.g., Refs. [14–17,24]. Other analyses include the stock prices of technology firms [14,25]. Furthermore, Ahmad [8] considers the price spillover relationship between oil prices and stock prices of technology and clean energy companies. The author argues that technology stocks are important to spillover on returns and volatilities from renewable energy companies and oil prices, and that technology and renewable energy companies are net transmitters, while oil prices are net receivers.

While various methodological frameworks have been applied to the universe of energy markets, covering VAR [1,26], GARCH [10,16], linear and nonlinear Granger causality [13], the decomposition of forecast error variance [9,27], copulas [6,24], wavelets [28], and quantile based models [15,29], some studies cover the possibility of asymmetric effects. For example, Kocaarslan and Soytas [30] use non-linear models and indicate that the impacts of the oil price are asymmetrical, with the increase in oil prices having negative impacts on the prices of clean energy. Uddin et al. [29] show that clean energy has a strong positive dependence on crude oil and that the relationship between oil and clean energy markets is asymmetrical over the quantiles. Xia et al. [7] indicate that the relationship between energy prices and renewable energy stocks is time-varying and exhibits high volatility. Dutta et al. [31] examine the impact of the volatility of the energy sector on clean assets and find a negative relationship, namely, that higher levels of volatility are associated with lower returns on clean assets, and that the impact is asymmetric: when levels of volatility are higher, the relationship is stronger in face of lower levels of volatility. Saeed et al. [15] use quantile-based estimators in the right and left tails and conclude that the relationship of returns between clean and dirty energy assets is larger in the tails compared to mean-based measures, and that the degree of relationship is time-varying and asymmetric with regard to extreme positive and negative returns. Dawar et al. [17] apply quantile regression and show that the relationship between clean energy and oil prices varies in relation to the various economic conditions and exhibits evidence of asymmetry, given that the relationship is strongly negative during bearish periods and non-significant during bullish periods.

In the context of the COVID-19 outbreak, WTI oil prices declined unprecedentedly to negative levels around April 2020 and Foglia and Angelini [32] indicate evidence of a heightened degree of connection between these markets. They show that WTI oil switched from being a transmitter of volatility to becoming a receiver of volatility with the pandemic crisis. Polat [33] studies the spillover effects among energy commodities (crude oil, natural gas, diesel, and gasoline) over the period June 2006–April 2020, showing evidence of instability in the dynamics of spillovers due to the COVID-19 outbreak. Yahya et al. [34] use a regime-switching model and conclude the clean energy index is dominant over oil in the most recent post-coronavirus crisis.

Considering the above literature, this paper extends our understanding on the relationship between dirty (less dirty) energy commodities and clean stock indices within a non-linear bidirectional analysis that involves Granger causality and transfer entropy in a time-varying setting. Interestingly, our dataset covers nine indices covering natural gas, crude oil, two different gasoline prices, diesel, heating oil, propane and two clean energy stock indices (S&P Global Clean Energy Index and RENIXX) for the period November 25, 2003 to December 30, 2020. This allows us to complement the related literature by providing a rich and comprehensive analysis to identify the main influencers in the energy market, which is relevant considering the potential impact of various extreme events and crisis periods.

3. Methodology and data

3.1. Methodology

The analysis of the dependence between financial assets could be made with several approaches. A general approach involves the use of correlation coefficients, which allows the way in which different variables are connected to be quantified and assessed. Many studies have employed several methods to compute correlation coefficients, both linear and non-linear [35–37]. Furthermore, some studies use the correlation coefficient from detrended cross-correlation analysis [38,39].

Despite the important information given by correlation coefficients, namely the interconnections between markets, depending on the objectives, directional measures could be desirable as they can distinguish between the intensity of the relationship among variables. One of the most usual ways to make such analysis is using Granger causality (GC). GC is based not only on the premise that the cause precedes the effect but also on the fact that the cause contains exclusive information about the effect. Simultaneously, the effect is unique and is not attributable to any another variable (i.e. in any other cause). Then, it can be said that GC captures causality as predictive of effect. Proposed by Granger [40] in a linear causality structure, GC assesses how past values of a given X variable cause another Y variable, considering present and past values of both in a VAR model. Although there are several areas of knowledge where the GC concept has been applied, from neurophysiology [41,42]; among others) to financial time series analysis [43], linear regression analysis made using VAR models is limited to linear association among variables. In order to better explain this approach, assume two joint stationary processes, X_t and Y_t, according to Granger [40]; a variable Y Granger-cause a variable X (with lags k, l) if and only if:

$$F\left(x_t\middle|x_{t-1}^{(k)},y_{t-1}^{(l)}\right) \neq F\left(x_t|x_{t-1}^{(k)}\right), \tag{1}$$

with $F(x_t|x_{t-1}^{(k)},y_{t-1}^{(l)})$ being the distribution function of the X variable conditional on the joint (k,l)-history $(X_{t-1}^{(k)}|Y_{t-1}^{(l)})$ of itself and the Y variable, and $F(x_t|x_{t-1}^{(k)})$ being the distribution function of X_t conditional exclusively on its own (k)-history.

According to Equation (1), a variable Y Granger-causes a variable X, if past values of Y help to predict the actual value of X, beyond the degree to which X helps to predict itself.

Although GC has been introduced to quantify the coupling direction between variables, it is a measure based on second-order, correlation-centred statistics, limiting its relevance to linear systems [44]. Since an important challenge for an energy researcher is to quantify the informational flow, and because the VAR approach cannot estimate this relationship, it is important to use measures sensitive to non-linear interactions and relations.

According to Dimpfl and Peter [45]; in the financial context, quantifying the flow of information requires the use of time-series properties and an asymmetric measure [46]. Considering transition probabilities (instead of static probabilities), Schreiber [47] introduced a dynamic structure for mutual information (MI) which measures the relationship between two variables. Thus, with the aim of measuring the information flow between two different time series, Schreiber [47] coupled the Shannon [48] entropy and Kullback and Leibler [49] distance concepts, considering that the involved processes are stationary Markov processes. In the bivariate case, the information flow from Y (order process l) to X (order process k) could be measured by quantifying the deviation from the generalized Markov properties, where the probability of observing X at time $t+1$ at the state x, conditional to the previous k observations is given by $p(x_{t+1}\left|x_t^{(k)}\right.) = p(x_{t+1}|x_t^{(k)}, y_t^{(l)})$, with x and y corresponding to the process state with $x_t^{(k)} = (x_t, \ldots, x_{t-k+1})$ and $y_t^{(l)} = (y_t, \ldots, y_{t-l+1})$, defining transfer entropy (TE).

Introduced by Schreiber [47]; TE comes from information theory and is based on Shannon's entropy concept, as a measure of uncertainty. It is a measure of information transfer between time series, regardless of the model used [50].

If variables have Gaussian behaviour, GC and TE are coincident [51]. However, in the presence of non-linear dynamics, which is common in financial markets, TE could be a better measure for analysing the dynamics and the possible asymmetric information flow between assets. In fact, TE is a directional and dynamic measure of predictive information rather than a measure of the causal information flow from a source and to a destination [18]. In other words, while GC is a predictive measure in the statistical sense, TE is framed in terms of uncertainty reduction.

Given the complex behaviour of indexes in the energy markets, for an in-depth evaluation of these relations and their respective magnitude, it is necessary to explore them assuming the presence of non-linear influences between them, justifying the use of TE. According to Schreiber [47]; the equation for TE, calculated based on Shannon's entropy, which defines the information flow from Y to X, is given by:

$$T_{Y\to X}(k,l) = \sum_{x,y} p\left(x_{t+1}, x_t^{(k)}, y_t^{(l)}\right) \log \frac{p\left(x_{t+1}\left|x_t^{(k)}, y_t^{(l)}\right.\right)}{(x_{t+1}|x_t^{(k)})} \tag{2}$$

Considering Eq. (2), it can be said that the TE from Y to X corresponds to the degree in which Y helps to reduce the ambiguity about the future of X beyond the degree on which X helps to reduce the ambiguity about its own future.

The definition of TE is clearly directional, as it considers only the dependency whose origin is in source series Y, while it does not consider dependences caused by common external sources. Moreover, we can also see that TE is an asymmetric measure, as it measures the dependence degree of one given variable from another. Finally, as it evaluates the dependence of Y on X, in an analogous way we can define also the TE from X to Y. The difference between both TEs allows us to obtain the dominant direction of the information flow [52], corresponding to the net transfer entropy (NET TE), which is given by:

$$NET\ TE_{Y\to X}(k,l) = T_{Y\to X}(k,l) - T_{X\to Y}(k,l) \tag{3}$$

If we look at Eqs. (1) and (2), the connection between GC and TE is clear. When $T_{X\to Y}(k,l) \neq 0$, Eq. (2) holds, then TE can be interpreted as a non-parametric test statistic for GC [53]. Thus, TE is a general measure of GC, under the stationary assumption. According to Lizier et al. [54]; TE not only allows information flow estimation between two-time series, it also allows information flow to be measured in a model-free approach; it does not depend on data structure or linearity and is robust to spurious "couplings". In this sense, TE is a non-parametric method. The TE measured is derived for discrete data. However, in any economic application, the obtained time series are continuous, with data partition being necessary to make them discrete. Using a finite number of partitions and a symbolic coding, data discretization could be done (see Behrendt et al. [52] for more details).

In the asset returns case, observations located in the distribution tails are particularly relevant. Then, a data partition based on empirical quantiles is usual, so that the left and the right tail observations, the negative and positive extreme respectively, fall into different categories. The results depend on the number of bins chosen (called alphabet length in Ref. [55]; among others). Even for large data sets, more than a few bins is not compatible with the available data. So, frequently three bins are used [55]. According to this, our time series were divided into three bins, along the quantiles 5% and 95% (represented by $q^r_{[0.05]}$ and $q^r_{[0.95]}$, respectively), as it seems to be consensual in the literature (see also [56]. Thus, symbolic coding will replace each of the values in the series under analysis, $y(t)$, by the corresponding symbol, i.e.,

$$S(t) = \begin{cases} 1 & for\ y(t) \leq q_1 \\ 2 & for\ q_1 < y(t) < q_2 \\ 3 & for\ y(t) \geq q_2 \end{cases} \tag{4}$$

In order to verify the robustness of the TE measurement, the quantile can be changed.

The statistical significance of TE, and consequent statistical inference, is based on the bootstrap method proposed by Dimpfl and Peter [57] with 300 bootstrap replications and not considering 50 observations from the beginning of the bootstrapped Markov chain. Repeating TE estimate calculation, it is possible to obtain the estimates' distribution under the null hypothesis of absence of information flow. Thus, the p-value is given by $1 - \widehat{q}_T$, where $\widehat{q}_T$ denotes the quantile of the simulated distribution, which is determined by the respective transfer entropy estimate [52].

We also calculated the net transfer entropy as follows, given by $NET\ TE_{YX} = TE_{YX} - TE_{XY}$, which tells us if a given asset is a net influencer or influenced by another one, which could give a better idea about the level of influence between the energy markets under study.

For our analysis, we made the estimations firstly considering the whole sample and then, in order to obtain time-varying information, we used consecutive sliding windows, considering windows of 1000 observations (about four years) as a basis. All the estimates used in this paper were made using the package RTranferEntropy, from R software.

3.2. Data

We used data for nine various indices: seven from energy products (natural gas, WTI crude oil, two different gasolines (NY gasoline and Gulf gasoline), diesel, heating oil and propane) and two indices covering clean energy products (S&P Global Clean Energy Index and RENIXX). A detailed description of the data is provided in Table 1. Prices for the seven energy products were retrieved in https://www.eia.gov, while the two indices covering clean energy stocks were retrieved from DataStream. The sample period covers the period from November 25, 2003 to December 30, 2020, where the beginning data was dictated by the data availability of some indices such as the S&P Global Clean Energy Index. The levels of each of the nine indices are transformed into the

Table 1
Information about the prices used in this analysis.

Asset	ID	Description
Natural gas	natgas	Henry Hub spot price, in dollars per million BTU
WTI oil	wti	West Texas Intermediate, FOB prices in dollars per barrel
NY gasoline	gasny	NY Harbor gasoline regular spot prices, dollars per gallon
Gulf Gasoline	gasgulf	Gulf Coast conventional gasoline regular spot price, dollars per gallon
Diesel	diesel	Los Angeles ultra-low sulfur CARB diesel spot price, dollars per gallon
Heating oil	heating	New York Harbor number 2 heating oil spot price, FOB, dollars per gallon
Propane	propane	Mont Belvieu (Texas) price, dollars per gallon
RENIXX	renixx	Renewable Energy Industrial Index
S&P Clean Energy	spclean	S&P Global Clean Energy Index

traditional log returns, resulting in a total of 4188 observations per index. The energy indices and commodities were chosen given that there are some which are widely explored (e.g. gasgulf or gasny), but some others are less explored (e.g. heating oil and propane). At the same time, as we aim to understand the role of clean energy in energy markets, we include the above-mentioned clean energy indices.

4. Empirical findings

As previously explained, we applied Granger causality (GC) and transfer entropy (TE) both for the sample as a whole (static analysis) and considering sliding windows for a time-varying approach. The results of GC for the whole sample are presented in the heatmap in Fig. 1 and those of TE are presented in the heatmap in Fig. 2. In both cases, heatmaps should be read as the variable on the row as the influencer of the value on the column. The lighter shaded pixels are associated with lower CG and TE values and the darker ones with CG and TE's higher values.

Considering the GC results, we conclude that natural gas prices are strongly Granger-caused by the remaining energy prices under study, but not by the clean energy indices. Conversely, natural gas shows some evidence of causality in heating oil and gasoline but does not influence the remaining assets under analysis.

For the remaining energy prices, WTI oil has an influence on gasoline and is influenced by heating oil, diesel and propane. Gasoline prices also show some influence on diesel prices (without the contrary evidence) with NY gasoline also influencing propane prices. Both gasoline prices influence themselves, although Gulf gasoline has a higher influence on NY gasoline than the contrary. For the rest of the relationships, we can see the influence of diesel on heating oil and propane, with diesel also being influenced by heating oil.

Regarding the clean energy indices, linearly they do not show any influence on energy prices, although the S&P clean energy index is influenced by Gulf gasoline and propane. In addition to this, a relationship exists between both indices, with S&P clean energy index strongly influencing the RENIXX, without the contrary influence existing. In general, considering the linear connections, the main influencers appear to be NY gasoline prices and diesel, while natural gas and RENIXX are the most influenced ones.

Regarding TE, with results presented in Fig. 2, we can see many more significant directional relationships, which are evident by the existence of more red cells. Natural gas continues to be more influenced than an influencer. However, contrarily to the GC, WTI oil is a strong influencer, with significant influence on all the remaining assets. In general, the pattern of relationships shows more significance, evidencing the non-linear complexity of energy

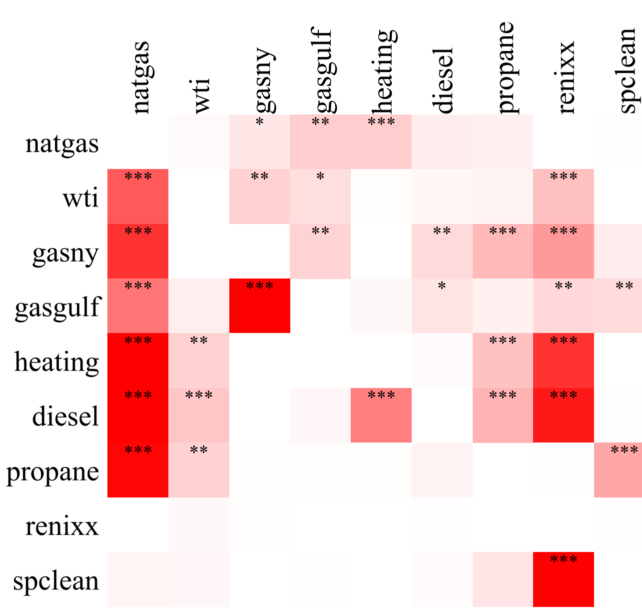

Fig. 1. Heatmap for Granger causality, considering the F statistical test. Notes: Lighter red corresponds to lower causality levels and darker red to higher levels. The number of lags was selected based upon AIC and BIC criteria. ***, ** and * depict significance of the causality at 1%, 5% and 10% levels respectively.

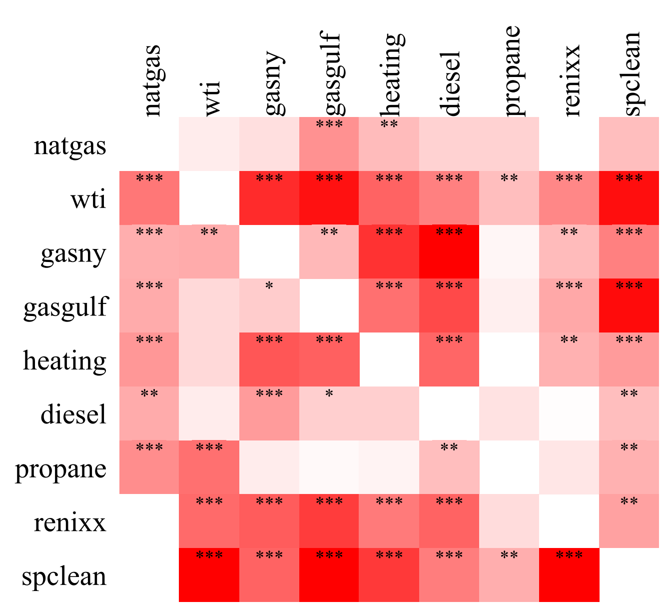

Fig. 2. Heatmap for transfer entropy. Notes: Lighter red corresponds to lower transfer entropy levels and darker red to higher levels. The number of lags was the same as used in the CG analysis. ***, ** and * depict significance of the causality at 1%, 5% and 10% levels respectively.

markets and asymmetric relationships between the energy products and clean energy indexes analysed. In this context, propane seems to be the most segmented asset, influencing natural gas, WTI oil and diesel but only being influenced by WTI oil. Considering the clean energy indices, when considering TE, they show evidence of being influencers of almost all energy prices, although RENIXX is less influenced by them when compared with the S&P clean energy index.

In addition to the fact that heatmaps provide an intuitive image of results, given the high number of analysed pairs of relations, we calculated the net transfer entropy ($NET\ TE_{YX} = TE_{YX} - TE_{XY}$) which shows if a given asset is a net influencer or influenced by another one. The results are presented in Fig. 3. Natural gas is more influenced than an influencer and it is a net influencer of just two other assets (Gulf gas and S&P clean energy index). Considering energy prices, WTI oil shows its important role, influencing all the others, except propane. Moreover, in general, the influence of WTI on other indices is high, as can be seen in the darker red cells. Propane also has an influence on several other energy prices, although with lower intensity. Diesel is clearly the energy price which is more influenced by the others. Regarding clean energy indices, it is possible to see that they have a net influence on most of the energy prices, with natural gas being the only exception. RENIXX and natural gas share the same transfer entropy between them (and thus, the net entropy is null), while S&P clean energy index is influenced by natural gas. Between them, S&P is a net influencer of RENIXX.

We proceed with our analysis and apply the sliding windows approach considering consecutive windows of 1000 observations, i.e., calculating the TE for the window from $t = 1, \ldots, 1000$, then for $t = 2, \ldots, 1001$, and so on, for a total of 3189 estimates for TE for each pair under analysis. We apply this approach in order to identify the time-varying dynamics of the TE as well as of net TE. Fig. 4 depicts the evolution of TE between WTI oil and natural gas, the reverse TE, and net TE for this pair. In this case, we can see that most of the time, WTI oil was responsible for transmitting more information to natural gas than the opposite. The exceptions are a short period in 2013, with a relatively low difference between TE and a larger period, from 2019 and part of 2020, in this case with higher net TE values.

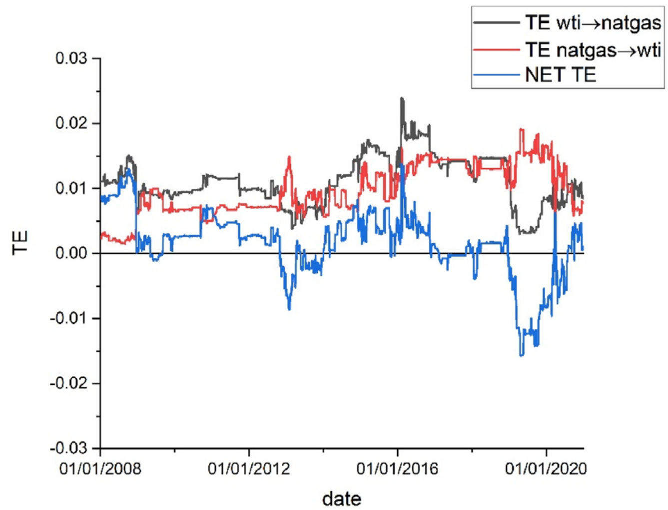

Fig. 4. Time evolution of TE between WTI oil and natural gas and net TE. Note: The analysis involves a sliding windows approach based on a window size of 1000 observations.

To simplify the visualization and interpretation of the information, we transformed the whole set of TE estimates into yearly based heatmaps, calculating the mean net TE (possible due to the additivity property of TE). Note that the heatmaps are built based on the type of information of Fig. 4. However, due to space constraints, we do not show the whole set of paired TE; however, results can be supplied upon request. The different heatmaps are represented in Fig. 5., through which we can have several important insights, both in a general way or in a more particular way, for each of the variables under analysis.

It is possible to identify an interesting time-varying pattern in several cases, identified by some colour changing in the different heatmaps, with no particular pattern. Natural gas was highly influenced by WTI oil, RENIXX and NY gas and in a higher extension, also influenced by propane and S&P clean energy index. Conversely, it exerted a higher influence on heating oil and diesel, mainly between 2016 and 2019 and on Gulf gas, during most of the sample. It is possible to highlight the high influence of natural gas on WTI oil in 2019, contrary to the historical influence. WTI oil plays a fundamental role in the energy market, being the most influential of all. This commodity has an almost constant influence on gasoline and diesel prices, which is not a surprising result, as well as on propane and natural gas. On the contrary, clean energy indices have more influence on oil, although in 2019 and 2020, oil inverted the global trend. In fact, the last year of the sample shows an evident influence of oil on the whole energy market, probably related to the COVID-19 context. NY gasoline seems to be more influenced by other prices, particularly by WTI oil (as previously identified) and by the S&P clean energy index and propane. Regarding the other prices under analysis, the evidence is mixed, although more marked by net negative TE (blue cells). RENIXX, natural gas and diesel are, however, more influenced by NY gasoline. As concerns Gulf gasoline, it has a positive influence on diesel in almost all years of the sample under analysis, propane mainly in the last years, and also a significant influence on the clean energy indices. On the contrary, it is mostly influenced by natural gas and WTI oil. Heating oil shows several mixed results, with time-varying patterns and an uncertain type of influence with several prices, such as NY and Gulf gasoline types, propane and both clean energy indices used in the analysis. Although it shows its higher influence on WTI oil and

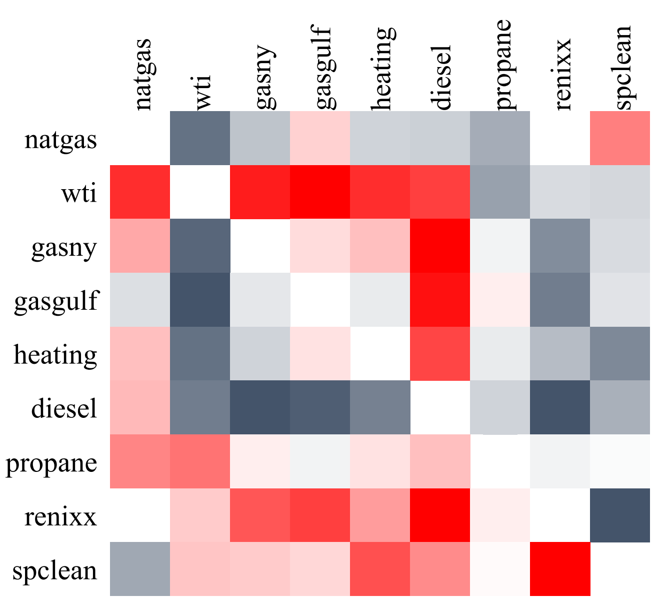

Fig. 3. Heatmap for net transfer entropy. Notes: Negative values (asset is net influenced) are represented in blue and positive values (asset is net influencer) in red.

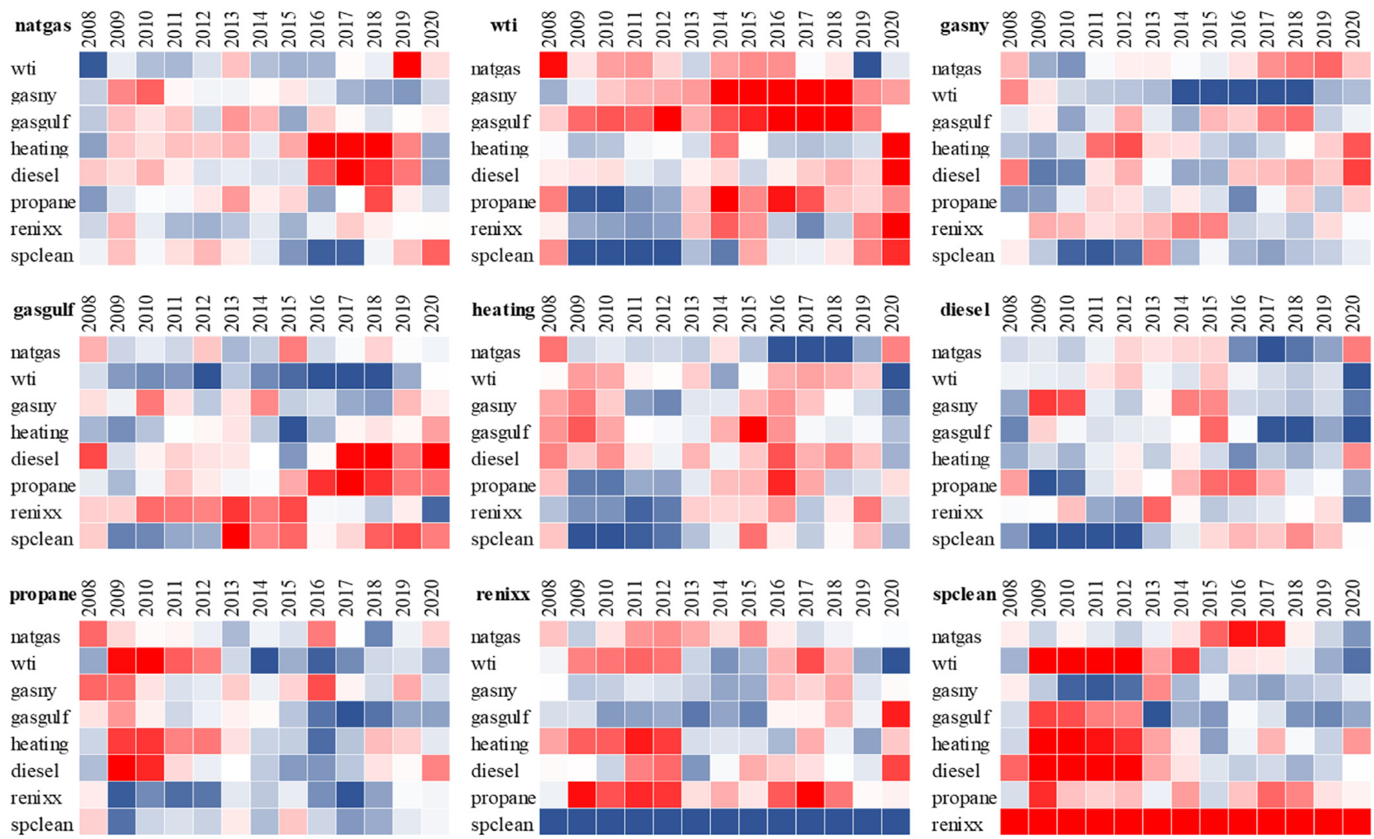

Fig. 5. Yearly evolution of the net TE. Notes: The basis-index is represented in each individual map at the top-left cell. Negative values (asset is net influenced) are represented in blue and positive values (asset is a net influencer) in red.

diesel prices, it is more influenced by natural gas. The diesel price is clearly the most influenced by the others. Despite the fact that most of the cells of its heatmap are blue, i.e., it is more influenced than an influencer, it is possible to conclude that in any case, diesel is more often influential. However, it is possible to highlight the fact that in the last years, diesel had some influence on the S&P clean energy index. Propane is also more influenced by the remaining indices. Firstly, it is clearly more influenced by the clean energy indices. However, in most cases, this type of energy is highly influenced by the others. It should be highlighted that until 2013, propane had the capacity to influence other energy prices (in particular WTI and heating oil and diesel), although since 2014 its influence was reduced.

Finally, both clean energy indices show the capacity of influencing the remaining energy products. In absolute terms, it is more visible in the case of the S&P clean energy index, mainly until 2013, with many dark red cells in its heatmap. After that, this clean energy index continues to have some influence on propane and natural gas (although in this case, in the last two years, the direction of the influence changed), but for the remaining indices, it became more influenced. Although the degree of intensity of the influence may be not so high, RENIXX has more influence on the other prices, evidenced by more red cells than S&P. The exceptions are gasoline prices, although the relationship changed in time (at the beginning, RENIXX was more influenced than an influencer, with the pattern changing since then). Comparing both clean energy indices between themselves, the S&P clean energy index clearly influences RENIXX: this happens in all years under analysis and with relative strength.

5. Discussion and conclusions

The results presented in the previous section give some interesting insights into the non-linear dynamics and return flow between the various energy indices under study. They identified the most influential and influenced in the universe of energy indices, with the possibility of differentiating between clean and dirty energies. Based on the linear Granger causality analysis, we conclude that natural gas is strongly influenced by other energy prices, corroborating the findings of Hartley et al. [58]. At the same time, natural gas prices are not Granger-caused by clean energy indices, which is in line with Geng et al. [13]; who conclude that the pricing mechanism of the North American natural gas market is based on market supply-demand and that crude oil prices were initially used as a pricing benchmark. It is noticeable that the influence of WTI on the market is limited based on the linear analysis. However, the use of transfer entropy complements our analysis and shows that the impact of WTI oil on the market is more relevant, especially influencing gasoline prices. For the remaining indices in general, RENIXX and the S&P clean energy index are strong influencers, corroborating the findings of Uddin et al. [29] who indicated a strong positive dependence between clean energy and oil markets. At the same time, RENIXX and S&P clean energy index are also both influenced, contradicting Ahmad and Rais [10] who found limited dependence between clean energy stocks and energy commodities, especially by WTI oil price (corroborating Kumar et al. [26]; and Ferrer et al. [9] and by the price of gasoline in the Gulf. This mutual information flow “from” and “to” RENIXX and S&P clean energy index can point to the competitive substitution between fossil energy and clean energy, implying that the oil market uncertainty plays a central role in determining “clean” energy returns.

Looking simultaneously at the results obtained through GC and TE analysis, we notice that while RENIXX seems to be influenced by other energy products, the S&P clean energy index does not show this pattern. On the contrary, TE analysis reveals that the S&P clean energy index is more influenced by the energy products (assets) than RENIXX. This pattern could seem antagonistic. According to our point of view, this different pattern could be a reflex not only of the composition of the indices (RENIXX is composed by companies of renewable energies located in 15 different countries from three continents, while the S&P clean index is composed by companies from the utilities, information technology, industrial and energy sectors, located in 18 different countries from four continents) but also a more accurate capacity of TE to analyse assets relationship in the presence of non-linear dynamics, which is common in financial markets (the energy market being no exception).

Looking at the net TE, we can see that natural gas is more influenced than an influencer, only registering an influence on Gulf gasoline and with the S&P clean index. Conversely, the natural gas index is clearly influenced by WTI (corroborating [1,21,58] as well as by other energy prices. WTI oil is a stronger influencer in the energy market, which is not entirely a new finding. However, in net terms, it is more influenced by clean energy markets, which could be related to their substitutability [6].

Despite this information for the whole sample, the existence of time-varying patterns could be important in explaining market dynamics, and to this end, we applied a sliding windows approach for the relationships. Based on the net TE, we can find some time-varying behaviours. For example, propane passed from being an influencer to being influenced, with some evidence for the same happening with S&P clean energy index. Natural gas also gained some relevance as an influencer of some other energy prices, such as heating oil, diesel or even WTI oil, which is consistent with the results shown by Lovcha and Perez-Laborda [27]. The sliding window approach also revealed that the S&P clean energy index had been an important influencer during the analysed period, (including on RENIXX). Issues like the liquidity of both indices could be related to these results, which could be the subject of future studies. Concerns about climate change and the need to achieve sustainable development increased the research and the demand for clean energies, in counterpoint to the use of other conventional forms of energy, like fossil fuels [59]. This could lead to a change in some patterns of relationships, and the time-varying patterns seem to confirm some changes. In fact, most of the time, both clean energy indices were more influencers than influenced by WTI oil, which is consistent with Reboredo et al. [28] or Uddin et al. [29]. However, mainly in the last two years the relationship seems to have changed, with a higher effect of WTI on these clean indices in 2020, which could be related to the health crisis caused by COVID-19, in line with Foglia and Angelini [32].

Our main findings evidence a complex system with non-linear characteristics as well as dynamic relationships across the various indices in the energy markets. The information of net TE between the various indices is quite rich and complex and does not point to a particular isolated influencer. However, we can see the growing strength of the clean energy markets compared to diesel, propane, and natural gas. The rise in the level of influence of clean energy leads us to believe that clean energies are increasingly capturing the attention of investors and markets. Is this good news? Probably so, especially if clean, non-polluting energies manage to lead energy markets and motivate respect for life on Earth.

Considering the above findings, some policy implications can be drawn. Firstly, they are important to market participants and decision makers who are concerned with return flow between the various energy indices. According to our results they should pay attention to non-linear dynamics and asymmetric return flow when constructing portfolio strategies and making inferences about risk management. Furthermore, given the evidence of time variation in return flow, especially during crisis periods, portfolio managers can construct less exposed portfolios to preserve the interests of investors. Secondly, given that the two clean energy indices have been gaining influence compared to diesel, propane, and natural gas, their potential diversification benefits become less pronounced, especially during crisis periods. However, their differences might imply that portfolio managers should not treat the S&P clean energy index and RENIXX as a single asset when it comes to their individual relationship with energy commodities. Thirdly, policymakers concerned with the stability and return flow among energy assets should have a closer look into the most influencers, especially during periods of market turmoil.

The use of bivariate techniques employed in this paper, although they complement each other, is not free of possible shortcomings, given the possible interconnection of the energy assets within a multivariate setting. So, multivariate techniques such as Granger and transfer entropy conditional and partial analysis (see, for example [60], may be alternative tools to reveal more notable results. Another possible extension can involve the use of transfer entropy within network-based analyses, while accounting for the potential effect of macro-economic and uncertainty factors.

Credit author statement

Paulo Ferreira: Conceptualization, Methodology, Software, Formal analysis, Writing – original draft, Writing – review & editing, Dora Almeida: Conceptualization, Methodology, Software, Formal analysis, Writing – original draft, Writing – review & editing, Andreia Dionísio: Conceptualization, Methodology, Software, Formal analysis, Writing – original draft, Writing – review & editing, Elie Bouri: Conceptualization, Methodology, Software, Formal analysis, Writing – original draft, Writing – review & editing, Derick Quintino: Conceptualization, Methodology, Software, Formal analysis, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Jadidzadeh A, Serletis A. How does the US natural gas market react to demand and supply shocks in the crude oil market? Energy Econ 2017;63:66–74.
[2] Lin B, Su T. Does COVID-19 open a Pandora's box of changing the connectedness in energy commodities? Res Int Bus Finance 2021;56:101360.
[3] Kilian L. The economic effects of energy price shocks. J Econ Lit 2008;46(4): 871–909.
[4] Umar Z, Gubareva M, Naeem M, Akhter A. Return and volatility transmission between oil price shocks and agricultural commodities. PloS One 2021;16(2): e0246886.
[5] Lin B, Bai R. Oil prices and economic policy uncertainty: evidence from global, oil importers, and exporters' perspective. Res Int Bus Finance 2021;56: 101357.
[6] Reboredo JC. Is there dependence and systemic risk between oil and renewable energy stock prices? Energy Econ 2015;48:32–45.
[7] Xia T, Ji Q, Zhang D, Han J. Asymmetric and extreme influence of energy price changes on renewable energy stock performance. J Clean Prod 2019;241: 118338.
[8] Ahmad W. On the dynamic dependence and investment performance of crude oil and clean energy stocks. Res Int Bus Finance 2017;42:376–89.
[9] Ferrer R, Shahzad SJH, López R, Jareño F. Time and frequency dynamics of connectedness between renewable energy stocks and crude oil prices. Energy Econ 2018;76:1–20.
[10] Ahmad W, Rais S. Time-varying spillover and the portfolio diversification

implications of clean energy equity with commodities and financial assets. Emerg Mark Finance Trade 2018;54(8):1837–55.
[11] Asl MG, Canarella G, Miller SM. Dynamic asymmetric optimal portfolio allocation between energy stocks and energy commodities: evidence from clean energy and oil and gas companies. Resour Pol 2021;71:101982.
[12] Batten JA, Ciner C, Lucey BM. The dynamic linkages between crude oil and natural gas markets. Energy Econ 2017;62:155–70.
[13] Geng JB, Ji Q, Fan Y. The relationship between regional natural gas markets and crude oil markets from a multi-scale nonlinear Granger causality perspective. Energy Econ 2017;67:98–110.
[14] Bondia R, Ghosh S, Kanjilal K. International crude oil prices and the stock prices of clean energy and technology companies: evidence from non-linear cointegration tests with unknown structural breaks. Energy 2016;101: 558–65.
[15] Saeed T, Bouri E, Alsulami H. Extreme return connectedness and its determinants between clean/green and dirty energy investments. Energy Econ 2020a;96:105017.
[16] Saeed T, Bouri E, Vo XV. Hedging strategies of green assets against dirty energy assets. Energies 2020b;13(12):3141.
[17] Dawar I, Dutta A, Bouri E, Saeed T. Crude oil prices and clean energy stock indices: lagged and asymmetric effects with quantile regression. Renew Energy 2021;163:288–99.
[18] Moldovan A, Caţaron A, Andonie R. Learning in feedforward neural networks accelerated by transfer entropy. Entropy 2020;22(1):102.
[19] Silvennoinen A, Thorp S. Financialization, crisis and commodity correlation dynamics. J Int Financ Mark Inst Money 2013;24:42–65.
[20] Serletis A, Timilsina G, Vasetsky O. International evidence on aggregate short-run and long-run interfuel substitution. Energy Econ 2011;33(2):209–16.
[21] Brown SP, Yucel MK. What drives natural gas prices? Energy J 2008;29(2): 45–60.
[22] Brigida M. The switching relationship between natural gas and crude oil prices. Energy Econ 2014;43:48–55.
[23] Nick S, Thoenes S. What drives natural gas prices? — a structural VAR approach. Energy Econ 2014;45:517–27.
[24] Bouri E, Naji J, Dutta A, Uddin GS. Gold and crude oil as safe-haven assets for clean energy stock indices: blended copulas approach. Energy 2019;178: 544–53.
[25] Sadorsky P. Correlations and volatility spillovers between oil prices and the stock prices of clean energy and technology companies. Energy Econ 2012;34(1):248–55.
[26] Kumar S, Managi S, Matsuda A. Stock prices of clean energy firms, oil and carbon markets: a vector autoregressive analysis. Energy Econ 2012;34(1): 215–26.
[27] Lovcha Y, Perez-Laborda A. Dynamic frequency connectedness between oil and natural gas volatilities. Econ Modell 2020;84:181–9.
[28] Reboredo JC, Rivera-Castro MA, Ugolini A. Wavelet-based test of co-movement and causality between oil and renewable energy stock prices. Energy Econ 2017;61:241–52.
[29] Uddin GS, Rahman ML, Hedström A, Ahmed A. Cross-quantilogram-based correlation and dependence between renewable energy stock and other asset classes. Energy Econ 2019;80:743–59.
[30] Kocaarslan B, Soytas U. Asymmetric pass-through between oil prices and the stock prices of clean energy firms: new evidence from a nonlinear analysis. Energy Rep 2019;5:117–25.
[31] Dutta A, Bouri E, Saeed T, Vo XV. Impact of energy sector volatility on clean energy assets. Energy 2020;212:118657.
[32] Foglia M, Angelini E. Volatility connectedness between clean energy firms and crude oil in the COVID-19 Era. Sustainability 2020;12(23):9863.
[33] Polat O. Measuring dynamic connectedness networks in energy commodities: evidence from the D-Y and frequency connectedness approaches. OPEC Energy Rev 2020;44(4):404–28.
[34] Yahya M, Kanjilal K, Dutta A, Uddin GS, Ghosh S. Can clean energy stock price rule oil price? New evidences from a regime-switching model at first and second moments. Energy Econ 2021;95:105116.
[35] Niu H. Correlations between crude oil and stocks prices of renewable energy and technology companies: a multiscale time-dependent analysis. Energy 2021;221:119800.
[36] Fiedor P. Networks in financial markets based on the mutual information rate. Phys Rev E - Stat Nonlinear Soft Matter Phys 2014;89(5):1–7.
[37] Benedetto F, Mastroeni L, Quaresima G, Vellucci P. Does OVX affect WTI and Brent oil spot variance? Evidence from an entropy analysis. Energy Econ 2020;89:104815.
[38] Guedes E, Dionísio A, Ferreira PJ, Zebende GF. DCCA cross-correlation in blue-chips companies: a view of the 2008 financial crisis in the Eurozone. Physica A 2017;479:38–47.
[39] Pal D, Mitra SK. Interdependence between crude oil and world food prices: a detrended cross correlation analysis. Physica A 2018;492:1032–44.
[40] Granger CWJ. Investigating causal relations by econometric models and cross-spectral methods. Econometrica 1969;37(3):424–38.
[41] Gourévitch B, Le Bouquin-Jeannès R, Faucon G. Linear and nonlinear causality between signals: methods, examples and neurophysiological applications. Biol Cybern 2006;95(4):349–69.
[42] Hosseini SAH, Sohrabpour A, He B. Electromagnetic source imaging using simultaneous scalp EEG and intracranial EEG: an emerging tool for interacting with pathological brain networks. Clin Neurophysiol 2018;129(1):168–87.
[43] Lai Y, Hu Y. A study of systemic risk of global stock markets under COVID-19 based on complex financial networks. Physica A 2021;566(127):125613.
[44] Gencaga D, Knuth KH, Rossow WB. A recipe for the estimation of information flow in a dynamical system. Entropy 2015;17(1):438–70.
[45] Dimpfl T, Peter FJ. The impact of the financial crisis on transatlantic information flows: an intraday analysis. J Int Financ Mark Inst Money 2014;31(1): 1–13.
[46] Altiparmak F, Dengiz B. A cross entropy approach to design of reliable networks. Eur J Oper Res 2009;199(2):542–52.
[47] Schreiber T. Measuring information transfer. Phys Rev Lett 2000;85:461–4.
[48] Shannon CE. A mathematical theory of communication. Bell Syst. Tech. J. 1948;27:379–423.
[49] Kullback S, Leibler RA. On information and sufficiency. Ann Math Stat 1951;22(1):79–86.
[50] Korbel J, Jiang X, Zheng B. Transfer entropy between communities in complex financial networks. Entropy 2019;21(11):1–13.
[51] Barnett L, Barrett AB, Seth AK. Granger causality and transfer entropy are equivalent for Gaussian variables. Phys Rev Lett 2009;103:238701.
[52] Behrendt S, Dimpfl T, Peter FJ, Zimmermann DJ. RTransferEntropy — quantifying information flow between different time series using effective transfer entropy. Software 2019;10:100265.
[53] García-Medina A, Hernández JBC. Network analysis of multivariate transfer entropy of cryptocurrencies in times of turbulence. Entropy 2020;22(7):1–18.
[54] Lizier JT, Heinzle J, Horstmann A, Haynes JD, Prokopenko M. Multivariate information-theoretic measures reveal directed information structure and task relevant changes in fMRI connectivity. J Comput Neurosci 2011;30(1): 85–107.
[55] Jizba P, Kleinert H, Shefaat M. Rényi's information transfer between financial time series. Physica A 2012;391(10):2971–89.
[56] Bae K-H, Karolyi GA, Stulz RM. A new approach to measuring financial contagion. Rev Financ Stud 2003;16(3):717–63.
[57] Dimpfl T, Peter FJ. Using transfer entropy to measure information flows between financial markets. Stud Nonlinear Dynam Econom 2013;17(1):85–102.
[58] Hartley PR, Medlock III KB, Rosthal JE. The relationship of natural gas to oil prices. Energy J 2008;29(3):47–65.
[59] Managi S, Okimoto T. Does the price of oil interact with clean energy prices in the stock market? Jpn World Econ 2013;27:1–9.
[60] Siggiridou E, Koutlis C, Tsimpiris A, Kugiumtzis D. Evaluation of granger causality measures for constructing networks from multivariate time series. Entropy 2019;21(11):1–24.