

Wykrywanie anomalii cenowych z zastosowaniem algorytmu isolation forrest na przykładzie notowań Day Ahead Market rynku hiszpańskiego, polskiego i ukraińskiego

Streszczenie

Celem niniejszej pracy jest ocena przydatności algorytmu Isolation Forest do identyfikacji anomalii cenowych na rynkach dnia następnego energii elektrycznej o zróżnicowanej strukturze wytwórczej i regulacyjnej. Analizie poddano dane cenowe z rynków dnia następnego Hiszpanii (ES), Polski (PL) oraz Ukrainy (UA) za rok 2023. W badaniu wykorzystano rzeczywiste dane rynkowe, a algorytm Isolation Forest zastosowano do wykrywania epizodów ekstremalnych zmian cenowych oraz nietypowych rozkładów cen godzinowych. Wyniki analizy wskazują istotne różnice w częstości oraz charakterze anomalii pomiędzy badanymi rynkami, wynikające m.in. z poziomu penetracji odnawialnych źródeł energii, głębokości rynku oraz uwarunkowań regulacyjnych. Szczególną uwagę zwrócono na rynek ukraiński, którego zachowanie cenowe odbiega istotnie od rynków UE w związku z nadzwyczajnymi warunkami funkcjonowania systemu elektroenergetycznego. Przeprowadzone badanie potwierdza, że metody uczenia maszynowego, w tym Isolation Forest, mogą stanowić użyteczne narzędzie wspomagające analizę stabilności i odporności rynków energii elektrycznej.

Słowa kluczowe

rynek energii elektrycznej, rynek dnia następnego, anomalie cenowe, Isolation Forest, uczenie maszynowe, integracja rynków energii

WPROWADZENIE

Rynki energii elektrycznej w Europie charakteryzują się rosnącą zmiennością cen, wynikającą z postępującej transformacji miksu wytwórczego, integracji odnawialnych źródeł energii oraz zmian w mechanizmach rynkowych. W szczególności rynki dnia następnego (Day-Ahead Market, DAM) stanowią kluczowy segment handlu energią, odzwierciedlający bieżącą równowagę podaży i popytu oraz krótkoterminowe ograniczenia systemowe.

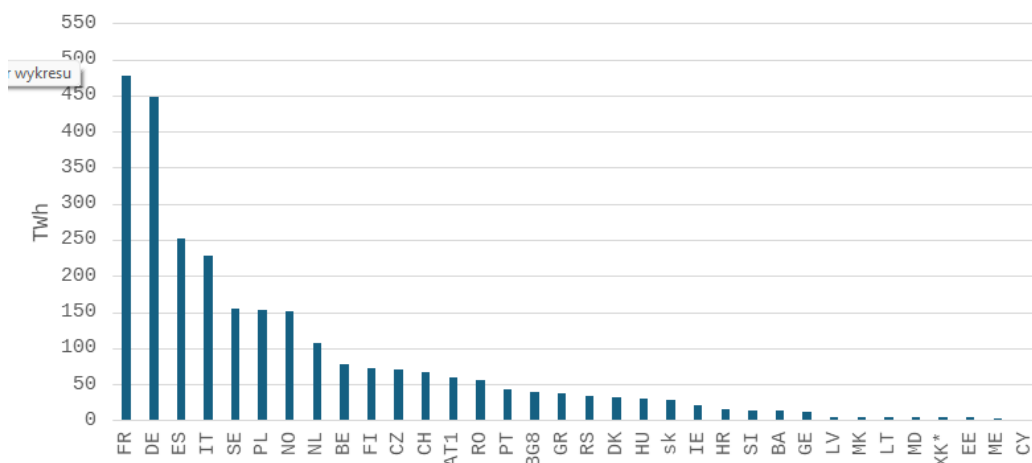
W ostatnich latach obserwuje się wzrost liczby epizodów ekstremalnych cen, takich jak ceny ujemne, nagłe skoki cenowe czy nietypowe rozkłady godzinowe. Zjawiska te są coraz trudniejsze do analizy przy użyciu klasycznych metod statystycznych, co uzasadnia zastosowanie narzędzi opartych na uczeniu maszynowym do identyfikacji i klasyfikacji anomalii cenowych.

Celem niniejszego badania jest ocena możliwości zastosowania algorytmu Isolation Forest do wykrywania anomalii cenowych na rynkach dnia następnego energii elektrycznej o zróżnicowanej strukturze rynkowej. Analiza obejmuje porównanie rynków Hiszpanii, Polski oraz Ukrainy, które różnią się poziomem integracji z rynkiem europejskim, strukturą wytwarzania oraz uwarunkowaniami regulacyjnymi.

W roku 2023 w Unii Europejskiej skonsumowano 2 440 TWh¹. Na wykresie 1 przedstawiono generację energii elektrycznej w poszczególnych krajach.

Wykr. 1. Generacja energii elektrycznej w krajach europejskich w 2023 r.

Energy generation in 2023



Źródło: ENTSE-O, „Statistical Factsheet 2023”, <https://www.entsoe.eu/data/power-stats/>, [data dostępu: 23/05/2024].

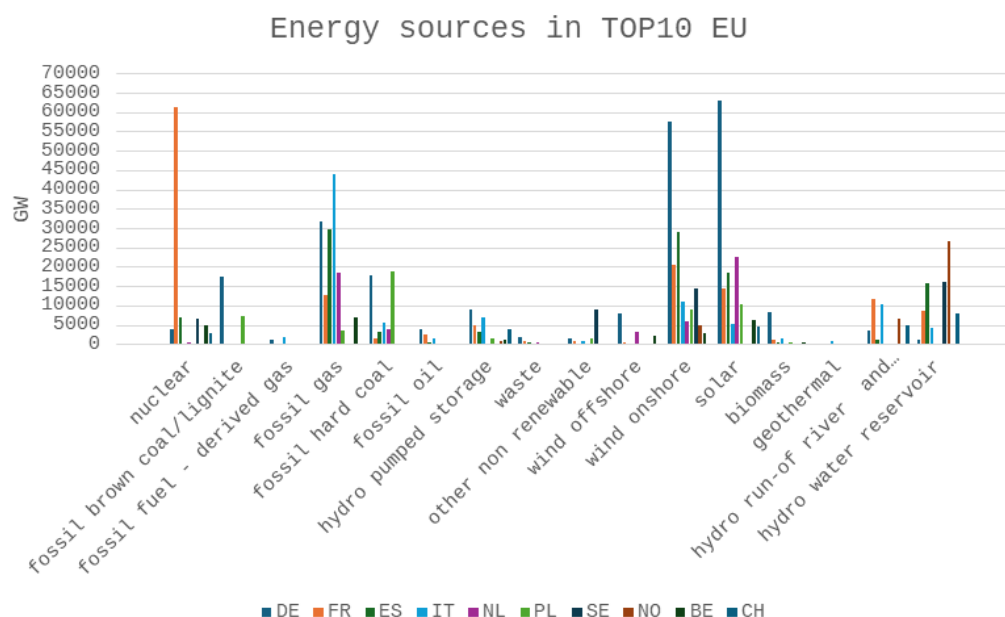
Na wykresie 1 można zaobserwować wielkość generacji energii elektrycznej w poszczególnych krajach. Ma to istotne znaczenie z perspektywy funkcjonowania zintegrowanego europejskiego rynku energii, w szczególności w kontekście roli poszczególnych krajów jako eksporterów lub importerów netto energii elektrycznej.² Dziesięć krajów o największej generacji energii odpowiada za blisko 80% energii elektrycznej wygenerowanej w 2023 roku w Europie. Na wykresie 2 przedstawiono parytet źródeł wytwórczych w dziesięciu krajach Unii Europejskiej o największej zainstalowanej mocy. Największą moc źródeł wytwórczych w Europie posiadają Niemcy o łącznej wartości ok. 233 GW, w tym 4GW elektrowni atomowej, 73 GW mocy jednostek konwencjonalnych, oraz 143 GW mocy jednostek niekonwencjonalnych. Dla odmiany Francja posiada 143 GW mocy, w tym aż 61 GW elektrowni atomowych, 17 GW elektrowni konwencjonalnych i 58 GW źródeł niekonwencjonalnych. Są to zaledwie dwa przykłady wskazujące na różnorodność parytetu energetycznego w poszczególnych krajach.

Różnorodność parytetu generacyjnego poszczególnych państw ma kluczowy wpływ na kształtowanie się cen energii elektrycznych w danych krajach. Ceny energii na rynkach dnia następnego wyznaczane są zgodnie z zasadą merit order, w której cena rozliczeniowa danej godziny odpowiada kosztowi krańcowemu ostatniej jednostki wytwórczej niezbędnej do zbilansowania zapotrzebowania. W każdej godzinie pracy systemu, dążeniem jest zapewnienie energii elektrycznej na poziomie zapotrzebowania, co wykonuje się korzystając z generacji różnych źródeł, a cena przypadająca na daną godzinę dostawy określana jest na podstawie ostatniego najtańszego źródła generacji, niezbędnego do zakontraktowania, aby zapewnić zapotrzebowanie na energię. W takiej sytuacji bilansowanie systemu polega na pobieraniu ilości energii elektrycznej wymaganej w bieżącej chwili, mając do dyspozycji większą dostępną moc³.

² <https://www.teraz-srodowisko.pl/aktualnosci/ue-zatwierdzila-reforme-rynku-energii-ceny-pradu-maja-byc-bardziej-stabilne-15087.html>, [data dostępu: 27/05/2024].

³ PSE S.A. stosuje praktykę, gdzie w każdej jednostce czasu w sieci powinno być ok. 8% więcej energii, oraz o 20% więcej mocy w dyspozycji aniżeli jest chwilowego zapotrzebowanie.

Wykr. 2. Parytet energetyczny w dziesięciu krajach o największej całkowitej mocy wytwórczej

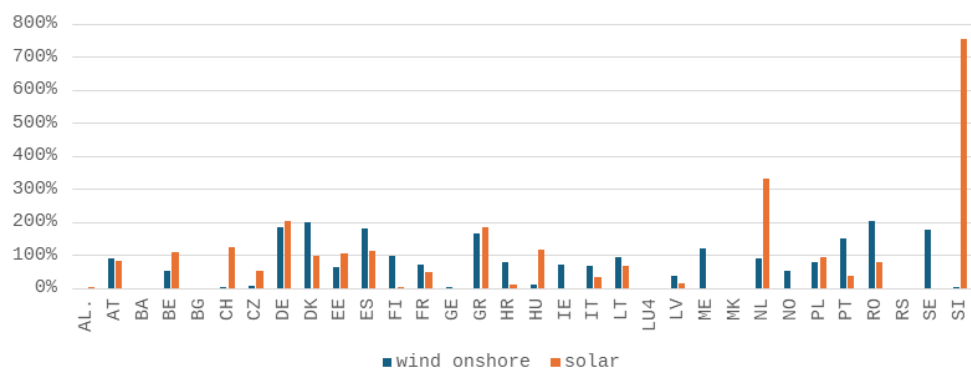


Źródło: ENTSE-O, „Statistical Factsheet 2023”, <https://www.entsoe.eu/data/power-stats/>, [data dostępu: 23/05/2024].

Przez wiele lat sytuacje na rynkach energii można było uznać za stosunkowo stabilne, a ceny w głównej mierze uzależnione były od cen paliw kopalnych. Jednak od kilku lat coraz mocniej zauważalna jest energia pochodząca ze źródeł pogodozależnych. Co więcej, ponieważ źródła te nie są dyspozycyjne, coraz częściej zachodzi zjawisko nadprodukcji energii wobec chwilowego zapotrzebowania. W takich sytuacjach nadmiar energii, prowadzi do typowej sytuacji nadpodaży, wynikiem której są ceny równe zero, lub nawet ujemne ceny. Na wykresie 3 przedstawiono stosunek całkowitej mocy zainstalowanej w lądowych elektrowniach wiatrowych i fotowoltaice do minimalnego zapotrzebowania na energię elektryczną w roku 2023. Na wykresie nie zaprezentowano wyników dla Bułgarii, która osiągnęła chwilowe minimalne zapotrzebowanie na poziomie 31 MWh/h w odniesieniu do mocy lądowej energetyki wiatrowej na poziomie 711 MW i mocy fotowoltaiki na poziomie 2879 MW, co dało wynik odpowiednio 2274% i 9287%. Jak można zauważyć w 12 krajach moc zainstalowana w przynajmniej jednej technologii przekracza 100% minimalnego zapotrzebowania na energię elektryczną.

Wykr. 3. Stosunek zainstalowanej mocy lądowych elektrowni wiatrowych i fotowoltaiki do minimalnego zapotrzebo-

Całkowita moc zainstalowana lądowej energetyki wiatrowej i fotowoltaiki w odniesieniu do minimalnego zapotrzebowania na energię



Źródło: ENTSE-O, „Statistical Factsheet 2023”, <https://www.entsoe.eu/data/power-stats/> [data dostępu: 23/05/2024].

W konsekwencji obserwowanych zmian strukturalnych bilansowanie systemu elektroenergetycznego coraz częściej zachodzi pomiędzy epizodami nadprodukcji źródeł pogodozależnych a okresami dominacji generacji konwencjonalnej, wynikającymi z niekorzystnych warunków meteorologicznych. Należy również podkreślić, że ceny bliskie zeru, lub ujemne ceny nie świadczą o taniości energii elektrycznej pochodzącej ze źródeł pogodozależnych, tylko o ich chwilowej nadprodukcji. Takie zjawiska nie sprzyjają rentowności źródeł pogodozależnych.

Celem niniejszego artykułu jest ilościowa identyfikacja i porównanie anomalii cenowych na wybranych europejskich rynkach dnia następnego z wykorzystaniem metod uczenia maszynowego, ze szczególnym uwzględnieniem różnic wynikających ze struktury miksu wytwórczego i głębokości rynku.

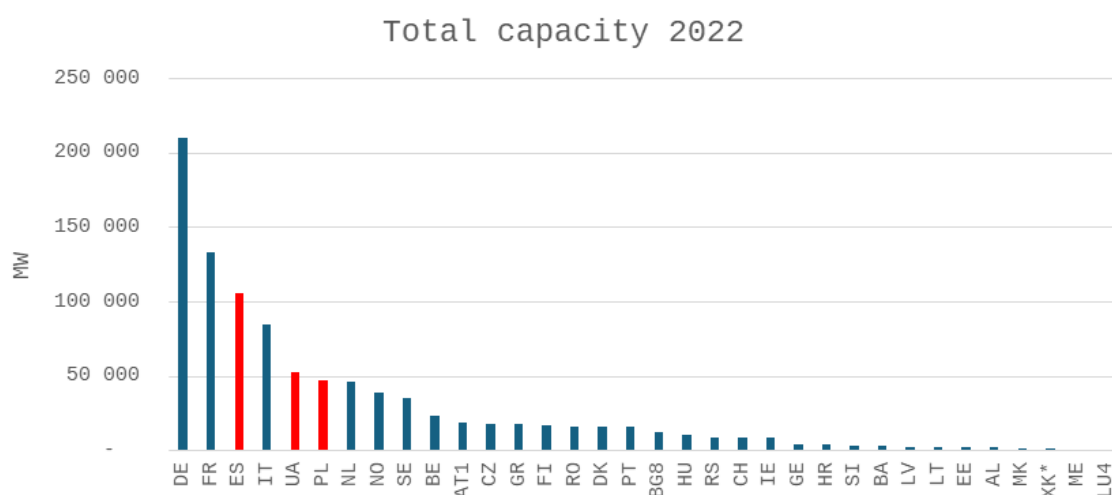
Wkład niniejszej pracy polega na porównawczej analizie anomalii cenowych na rynkach dnia następnego o istotnie odmiennych uwarunkowaniach regulacyjnych i strukturalnych, z wykorzystaniem jednolitej metodologii uczenia nienadzorowanego. W odróżnieniu od dotychczasowych badań koncentrujących się na pojedynczych rynkach lub symulacjach modelowych, analiza oparta jest na rzeczywistych danych rynkowych i uwzględnia kontekst integracji rynków europejskich oraz ograniczeń systemowych.

PRZEDMIOT BADANIA

W ramach niniejszej publikacji prezentowane są wyniki analizy anomalii cenowych z wykorzystaniem algorytmu „isolation forrest” na przykładzie cen rynków dnia następnego Hiszpanii, Polski i Ukrainy w roku 2023. Kraje te zostały wybrane jako zaliczające się do grona dziesięciu krajów o największej zainstalowanej mocy wytwórczej oraz największego zapotrzebowania na energię w Europie⁴.

Na wykresie 4 zaprezentowano całkowitą moc zainstalowaną w krajach europejskiej, wyróżniając systemy hiszpański, ukraiński i polski.

Wykr. 4. Całkowita moc generacji zainstalowana w poszczególnych krajach w roku 2022

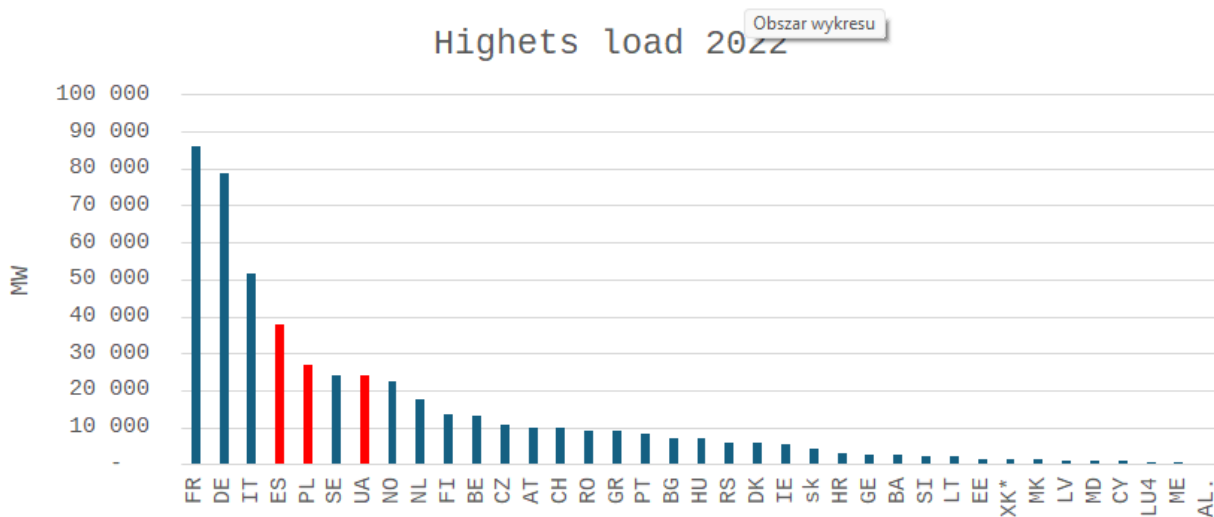


Źródło: ENSTE-O, „Statistical Factsheet 2022”, <https://www.entsoe.eu/data/power-stats/> [data dostępu: 23/01/2024].

Na wykresie 5 przedstawiono szczytowe zapotrzebowanie na energię elektryczną w krajach europejskich. Kolorem czerwonym oznaczono wyniki dla Hiszpanii, Ukrainy i Polski.

⁴ Zainstalowana moc w 2022 roku – Hiszpania miejsce trzecie, Ukraina miejsce piąte, Polska miejsce szóste. Szczytowe zapotrzebowanie na energię – Hiszpania miejsce 4, Polska miejsce piąte, Ukraina miejsce siódme.

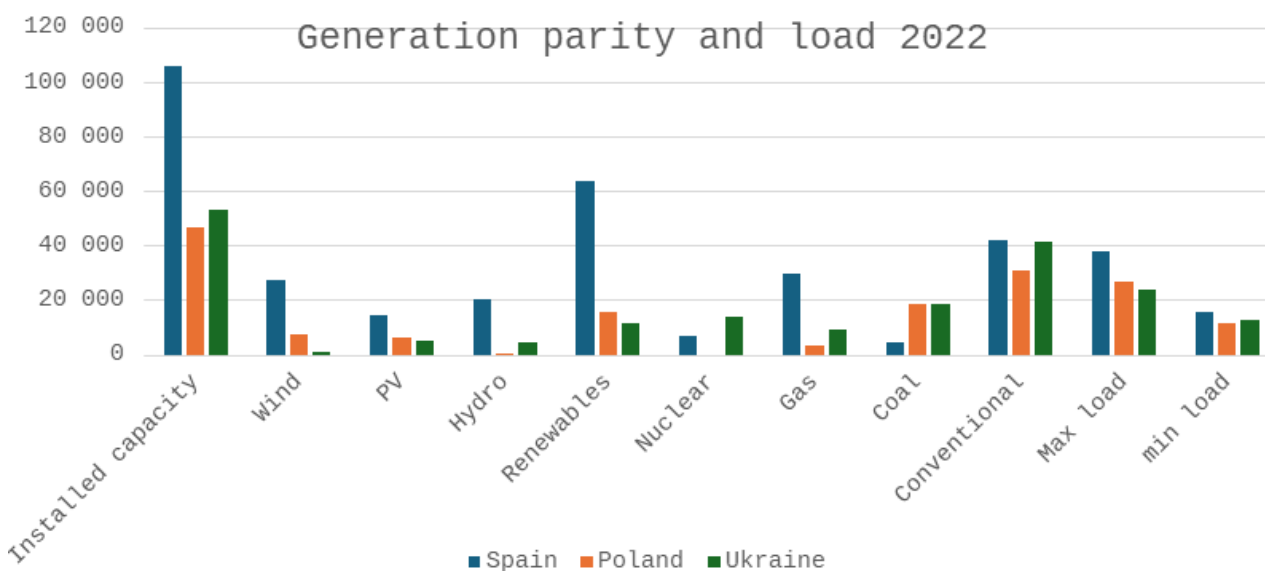
Wykr. 5. Szczytowe zapotrzebowanie na energię elektryczną w poszczególnych krajach w roku 2022



Źródło: ENSTE-O, „Statistical Factsheet 2022”, <https://www.entsoe.eu/data/power-stats/> [data dostępu: 23/01/2024].

Jak można zauważyć na wykresie 4 i 5, wybrane kraje do badania kwalifikują się do grona liderów energetycznych Europy ze względu zarówno na zapotrzebowanie na energię jak i ze względu na zainstalowaną moc wytwórczą. Jednak różnią się one parytetem technologii wytwórczych w swoich systemach energetycznych. Na wykresie 6 zaprezentowano miks energetyczny w Hiszpanii, Ukrainie i Polsce. Można zauważyć dużą rozbieżność w przypadku energetyki wiatrowej, wodnej, atomowej, gazowej czy węglowej. Ma to istotne znaczenie dla sposobu kształtowania się cen energii opartych o metodę merit order, czyli tzw. „ostatniego zaszeregowania”.

Wykr. 6. Miks systemów energetycznych Hiszpanii, Ukrainy i Polski



Źródło: ENSTE-O, „Statistical Factsheet 2022”, <https://www.entsoe.eu/data/power-stats/> [data dostępu: 23/01/2024].

W tabeli 1 przedstawiono wyniki analizy statystycznej notowań cen energii rynku dnia następnego dla okresu od 12.02.2023 do 31.12.2023 roku w Hiszpanii, Ukrainie i Polsce, oraz podstawowe dane charakteryzujące strukturę wytwórczą systemów energetycznych.

Tab. 1. Wyniki analizy statystycznej cen rynku dnia następnego oraz podstawowe dane struktury wytwórczej systemów elektroenergetycznych Hiszpanii, Ukrainy i Polski

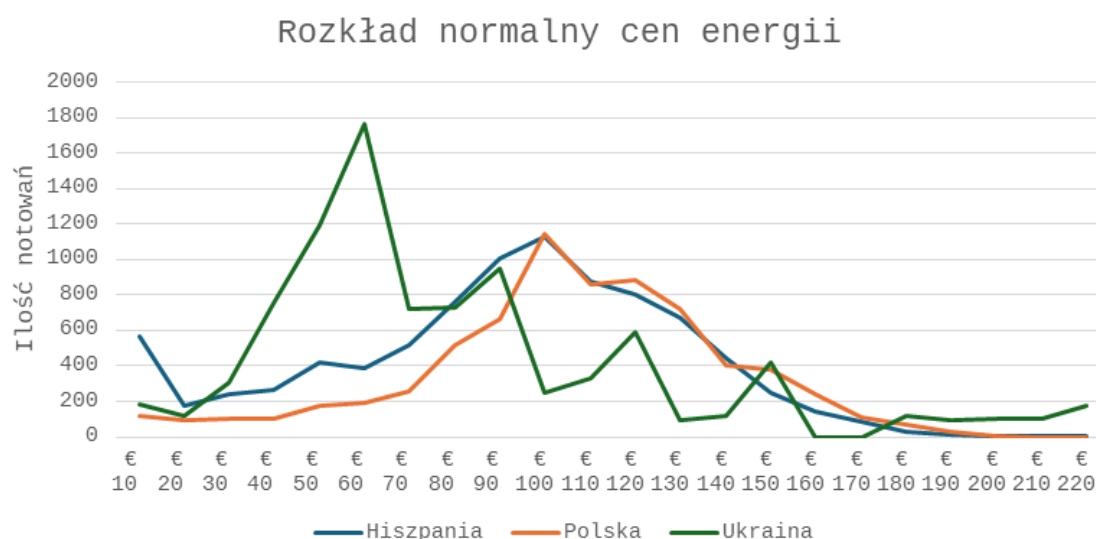
	Spain	Polska	Ukraina
Średnia	€ 82	€ 115	€ 80
Mediana	€ 88	€ 116	€ 70
odch. Stand.	€ 37	€ 40	€ 37
max	€ 220	€ 232	€ 176
min	€ -	-€ 12	€ -
kwartyl 25	€ 60	€ 95	€ 50
kwartyl 75	€ 108	€ 140	€ 94
Skośność	- 0,51	- 0,53	0,71
	Spain	Polska	Ukraina
Zainstalowana moc	106 046	47 163	53 248
Wiatr	27 735	7 950	1 111
PV	14 640	6 664	5 363
El. Wodna	20 342	785	4 829
OZE	63 691	16 046	11 502
Atom	7 117	-	13 835
Gaz	29 927	3 807	9 360
Węgiel	4 642	19 073	18 587
Konwencjonalna	42 335	31 117	41 782
Max load	37 898	27 211	24 171
min load	16 101	11 824	12 899
Max load/Konwencjonalna	90%	87%	58%

Źródło: opracowanie własne.

Zróżnicowanie relacji pomiędzy mocą zainstalowaną w źródłach konwencjonalnych a zapotrzebowaniem szczytowym stanowi istotne tło interpretacyjne dla dalszej analizy rozkładów cenowych oraz identyfikowanych anomalii.

Na wykresie 7 przedstawiono rozkład normalny notowań cen energii na rynkach dnia następnego Hiszpanii, Ukrainy i Polski.

Wykr. 7. Rozkład normalny notowań cen energii na rynkach dnia następnego Hiszpanii, Ukrainy i Polski



Źródło: Opracowanie własne.

Jak można zaobserwować, każdy kraj posiada własną charakterystykę notowań. W przypadku Hiszpanii można wyróżnić dwa maksima lokalne, jeden w okolicach 0 EUR i drugi w okolicy 100 EUR. Rynek ukraiński charakteryzuje się czterema obszarami maksimum lokalnego, na poziomie 60 EUR, 90 EUR, 120 EUR i 150 EUR. Polski rynek dnia następnego charakteryzuje się rozkładem najbardziej zbliżonym do rozkładu normalnego z pojedynczym maksimum w okolicach 100 EUR. Wykres 7 dowodzi różnorodności rynków dnia następnego badanych krajów, co jest konsekwencją różnych struktur wytwórczych.

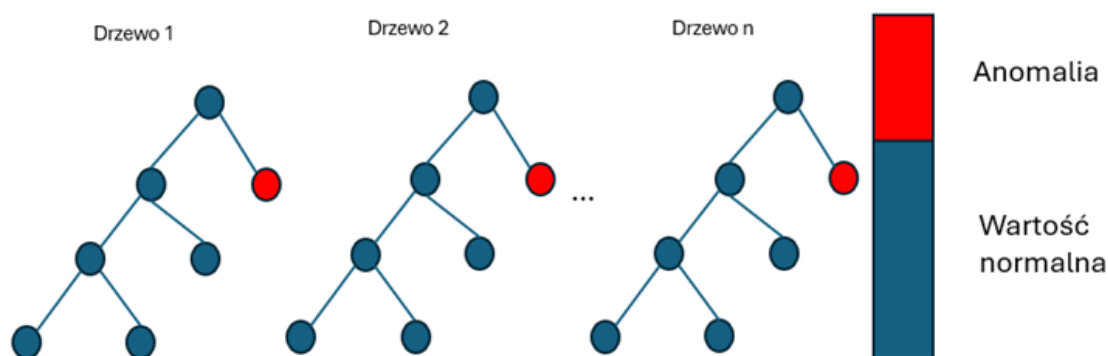
METODYKA BADANIA

Dla potrzeb przeprowadzenia badania zastosowano algorytm „lasu separującego” (ang. Isolation forest). Jest to algorytm, który pozwala w badanej populacji klasyfikować próbki, jako normalne lub odstające anomalie. Algorytm ten został zaproponowany w 2008 roku przez Fei Tony Liu, Kai Ming Ting i Zhi-Hua Zhou. Algorytm ten zalicza się do uczenia nienadzorowanego, z grupy drzew decyzyjnych. Algorytm skupia się na dwóch, zasadniczych cechach anomalii, tj.:

- jest ich stosunkowo niewiele w populacji,
- charakteryzują się wartościami znacząco odbiegającymi od pozostałych próbek populacji.

Na rysunku 1 przedstawiono graficzną reprezentację funkcjonowania algorytmu separującego. Populacja próbek jest poddawana analizie poprzez konstruowanie drzew, w których określone są odległości od korzeni. Próbki o małych odległościach od korzenia najczęściej stanowią anomalie. Jednak z racji stworzenia lasu, niezbędne jest zastosowanie oceny zaufania do wskazanych anomalii (ang. Anomaly score). Jest to funkcja licząca średnią długość ścieżki danej próbki od wierzchołka.

Rys. 1. Graficzna reprezentacja tworzenia drzew separujących, oraz funkcji zaufania



Źródło: opracowanie własne.

Algorytm ma różne implementacje i różną identyfikację próbek, aczkolwiek pierwotna zakładała wartości zaufania w przedziale od 0 do 1, gdzie wartości zbliżone do 1 stanowiły anomalie, natomiast wartości zbliżone do 0 wartości normalne. Wynik anomalii reprezentuje następujący wzór:

$$S(x, n) = 2^{-\frac{E(h(x))}{c(n)}}$$

Gdzie:

$h(x)$ – długość ścieżki punktu danych x w danym drzewie izolacyjnym,

$E(h(x))$ – średnia wartość długości ścieżki we wszystkich drzewach separacji,

$c(n)$ – średnia wartość $h(x)$ przy danej wielkości próbki (w powyższym przypadku n razy).

W analizie wykorzystano algorytm Isolation Forest, będący metodą nienadzorowanego uczenia maszynowego przeznaczoną do wykrywania obserwacji odstających. Model został zaimplementowany w języku Python z wykorzystaniem biblioteki scikit-learn. Parametry algorytmu dobrano empirycznie, koncentrując się na identyfikacji ekstremalnych epizodów cenowych bez nadmiernego dopasowania do szumu danych.

Przyjęta metodyka umożliwia porównywalną identyfikację anomalii pomiędzy rynkami o odmiennych charakterystykach strukturalnych, bez konieczności stosowania założeń dotyczących rozkładu danych wejściowych.

WYNIKI PRZEPROWADZONEJ ANALIZY

W ramach wykonanych badań, przyjęte podejście badawcze pozwalała odgórnie określić poziom skażenia populacji danych, tzn. określić ile próbek z populacji uznajemy za anomalie. W tabeli 2 przedstawiono podstawowe założenia dla przeprowadzonych badań.

Tab. 2. Przyjęte założenia dla wykonanych badań

	Hiszpania	Polska	Ukraina			
Poziom skażenia	5%	10%	5%	10%	5%	10%
Próbki normalne	7394	6987	7367	6978	7372	7050

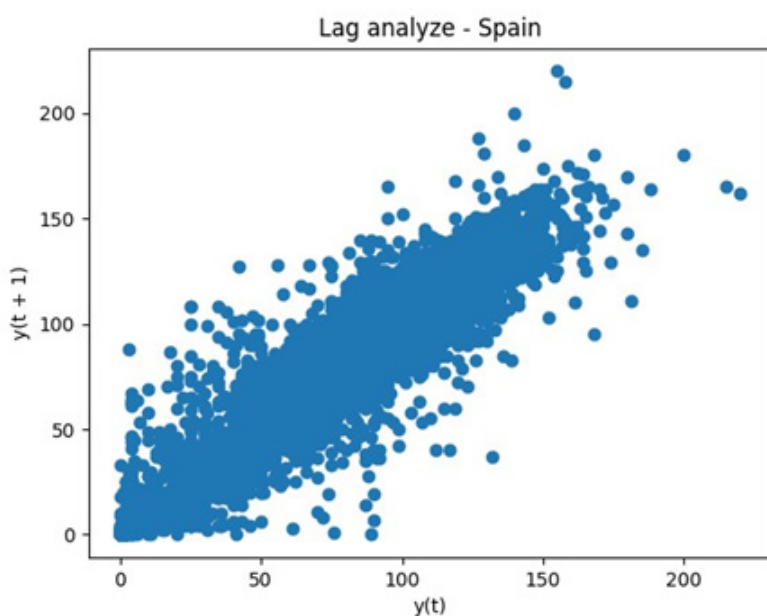
Anomalie	258	765	385	774	380	702
Populacja	7752					

Źródło: Opracowanie własne.

Dla każdego badanego rynku wielkość populacji wynosiła 7752 próbki, natomiast w badaniu przyjęto dwie wartości funkcji zaufania, tj. 5% i 10% całkowitej populacji, jako wielkości izolacji anomalii od wartości normalnych.

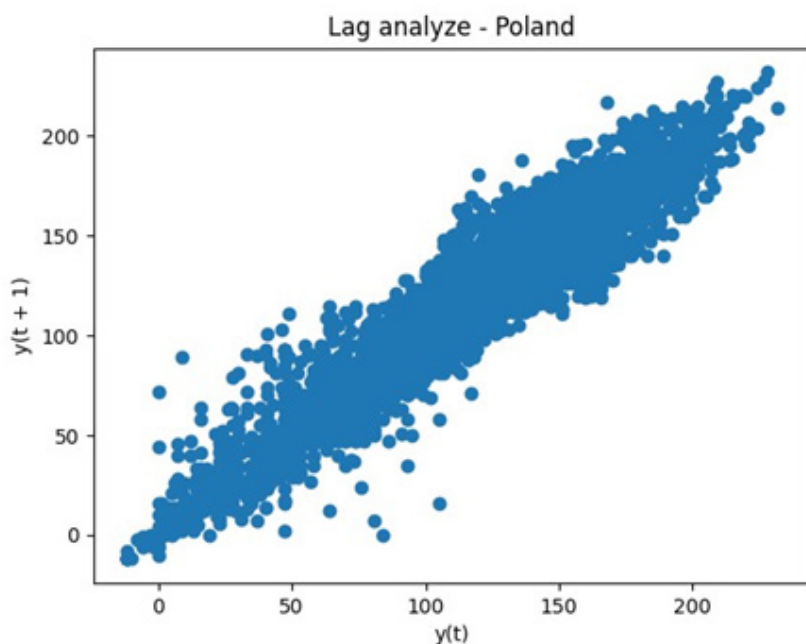
W celu zweryfikowania trafności zastosowania algorytmu lasu separującego, dla badanych zbiorów przeprowadzono analizę korelacji opóźnienia, która bada korelację próbek następujących po sobie. Pozwala ona ocenić specyfikę zbioru danych i ciągłości zmienności próbek po sobie. Na wykresach 8, 9 i 10 przedstawiono analizę korelacji opóźnienia dla danych reprezentujących rynek dnia następnego w Hiszpanii, Ukrainie i Polsce.

Wykr. 8. Analiza korelacji opóźnienia – hiszpański rynek dnia następnego



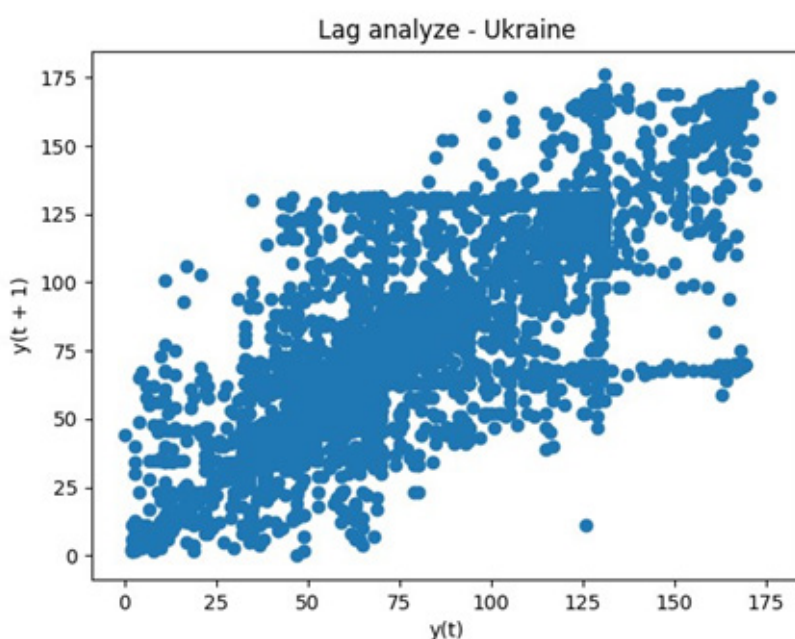
Źródło: Opracowanie własne.

Wykr. 9. Analiza korelacji opóźnienia – polski rynek dnia następnego



Źródło: Opracowanie własne.

Wykr. 10. Analiza korelacji opóźnienia – ukraiński rynek dnia następnego

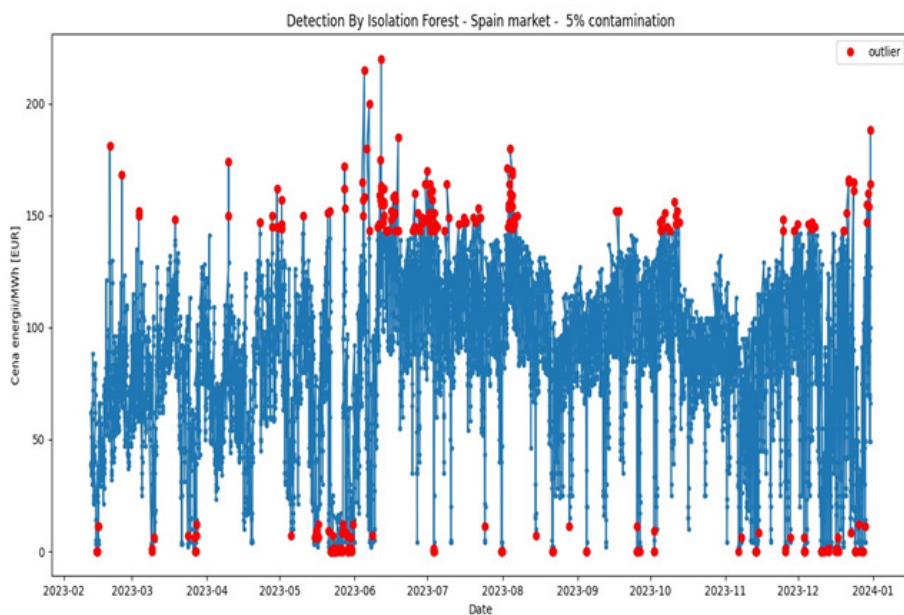


Źródło: Opracowanie własne.

Najlepszy wynik analizy korelacji opóźnień przypada Polsce, następnie Hiszpanii i na końcu Ukrainie. Co więcej na wykresie dla rynku ukraińskiego można zaobserwować anomalie, które wynikają z faktu, że z racji stanu wojny, rynek energii był okresowo odgórnie kontrolowany, a ceny narzucane. Stąd charakterystyczne proste pionowe i poziome na wykresie, utworzone przez poszczególne próbki.

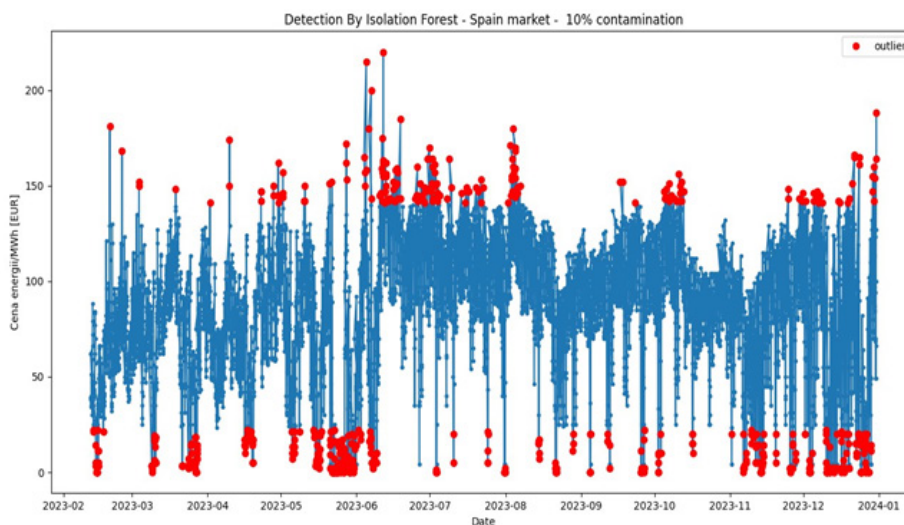
Na wykresach 11, 12, 13, 14, 15 i 16 przedstawiono wykresy reprezentujące identyfikację anomalii cen energii na rynkach dnia następnego Hiszpanii, Polski i Ukrainy.

Wykr. 11. Analiza korelacji opóźnienia – hiszpański rynek dnia następnego (poziom skażenia 5%)



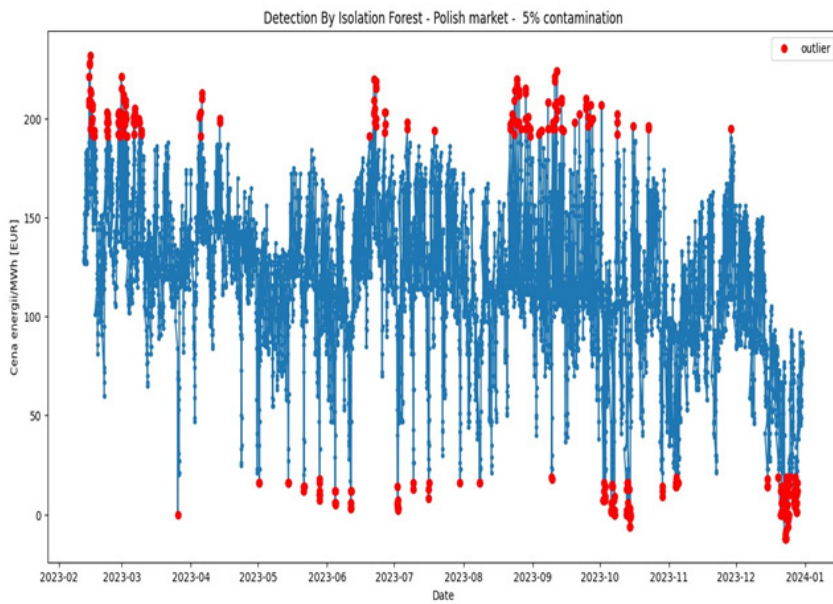
Źródło: Opracowanie własne.

Wykr. 12. Analiza korelacji opóźnienia – hiszpański rynek dnia następnego (poziom skażenia 10%)



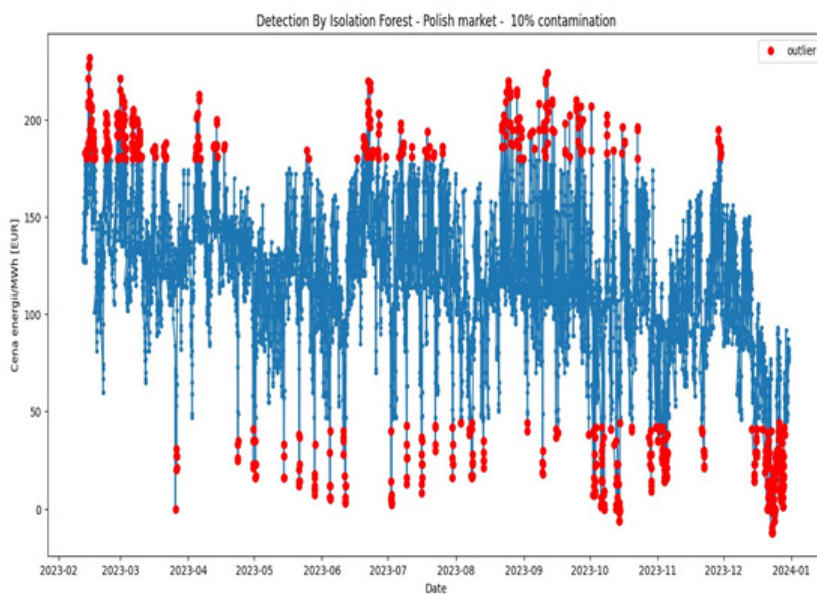
Źródło: Opracowanie własne.

Wykr. 13. Analiza korelacji opóźnienia – polski rynek dnia następnego (poziom skażenia 5%)



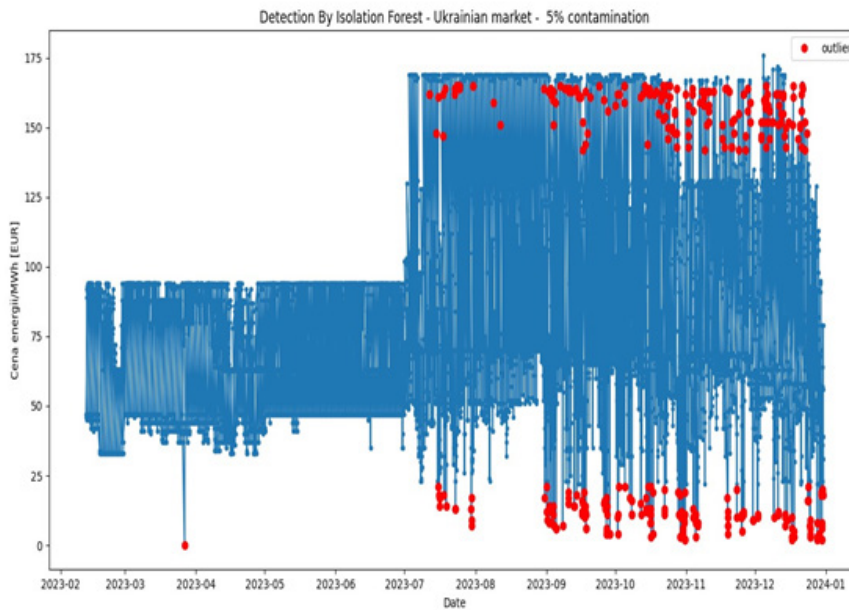
Źródło: Opracowanie własne.

Wykr. 14. Analiza korelacji opóźnienia – polski rynek dnia następnego (poziom skażenia 10%)



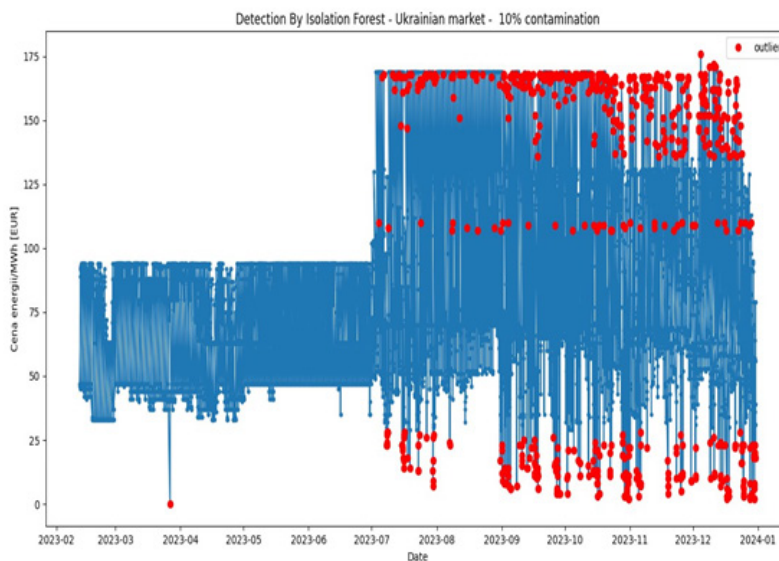
Źródło: Opracowanie własne.

Wykr. 15. Analiza korelacji opóźnienia – ukraiński rynek dnia następnego (poziom skażenia 5%)



Źródło: Opracowanie własne.

Wykr. 16. Analiza korelacji opóźnienia – ukraiński rynek dnia następnego (poziom skażenia 10%)



Źródło: Opracowanie własne.

Uzyskane wyniki wskazują, że liczba oraz intensywność wykrywanych anomalii cenowych istotnie różnią się pomiędzy analizowanymi rynkami. W szczególności rynek hiszpański charakteryzuje się wysoką częstością epizodów cen ujemnych, co koreluje z dużą penetracją odnawialnych źródeł energii. Z kolei rynek polski wykazuje mniejszą liczbę anomalii, lecz o większej amplitudzie cenowej.

WNIOSKI

analiza potwierdza, że algorytm Isolation Forest może być skutecznym narzędziem do identyfikacji anomalii cenowych na rynkach dnia następnego energii elektrycznej. Wyniki badania wskazują na istotne różnice w charakterze i częstotliwości anomalii pomiędzy rynkami Hiszpanii, Polski oraz Ukrainy, wynikające z odmiennych uwarunkowań strukturalnych i regulacyjnych. Narzędzie to może znaleźć praktyczne zastosowanie dla osób funkcjonujących w obszarze obrotu energią na rynkach bieżących jak i dla analityków systemu.

Badanie pokazuje również, że metody uczenia maszynowego mogą stanowić wartościowe uzupełnienie klasycznych analiz rynkowych, zwłaszcza w kontekście rosnącej zmienności cen energii. Dalsze prace badawcze powinny koncentrować się na rozszerzeniu analizy o inne segmenty rynku oraz na integracji wyników z modelami fundamentalnymi systemu elektroenergetycznego.

Uzyskane wyniki wskazują, że dalsza integracja rynków oraz wzrost elastyczności systemów elektroenergetycznych będą miały kluczowe znaczenie dla ograniczenia skali anomalii cenowych w warunkach rosnącej penetracji źródeł odnawialnych.

BIBLIOGRAFIA:

1. Soerio S., Ferreira Dias M., "Renewable energy community and the European energy Market: main motivations", *Heliyon* 6 (2020), e04511
2. Roldan J., Burgos M., Riquelme J., Trigo A., "The merit order effect of energy efficiency", *Energy Procedia* 2016, 175 – 184,
3. Loumakis S., Giannini E., Maroulis Z., "Merit order effect modeling: The case of the Hellenic electricity market", *Energies* 2019,
4. Kusz K., „Bilansowanie handlowe niespokojnych źródeł wytwórczych – ujęcie i wycena”, *Polityka Energetyczna – Energy Policy Journal* 2017,
5. Roldan – Fernandez, Burgos – Payan, Requelme-Santos, Trigo – Garcia, "The merit – Order Effect of Energy Efficiency", *Energy Economics Iberian Conference*, 2016
6. Liu F., Ting K., Zhou Z., "Isolation forest", *Eight IEEE International Conference on Data Mining*, 12.2008
7. Kirschen, D. S., Strbac, G. *Fundamentals of Power System Economics*. Wiley. 2018
8. Clò S., Cataldi A., Zoppoli P. The merit-order effect in the Italian power market. *Energy Policy*, 77, 79–88. 2015

DETECTION OF PRICE ANOMALIES USING THE ISOLATION FOREST ALGORITHM: EVIDENCE FROM DAY-AHEAD MARKET PRICES IN THE SPANISH, POLISH, AND UKRAINIAN ELECTRICITY MARKETS

Abstract

The aim of this study is to assess the applicability of the Isolation Forest algorithm for identifying price anomalies in day-ahead electricity markets characterized by diverse generation structures and regulatory frameworks. The analysis is based on day-ahead price data from the Spanish (ES), Polish (PL), and Ukrainian (UA) electricity markets for the year 2023. Real market data were used, and the Isolation Forest algorithm was applied to detect episodes of extreme price fluctuations as well as atypical hourly price distributions.

The results reveal significant differences in both the frequency and nature of price anomalies across the analyzed markets, stemming, among others, from varying levels of renewable energy penetration, market depth, and regulatory conditions. Particular attention is given to the Ukrainian market, whose price behavior deviates substantially from EU markets due to extraordinary operating conditions of the power system.

The study confirms that machine learning methods, including Isolation Forest, can serve as a useful tool supporting the analysis of stability and resilience in electricity markets.

Keywords

electricity market; day-ahead market; price anomalies; Isolation Forest; machine learning; energy market integration