UNIWERSYTET OPOLSKI

Instytut Ekonomii i Finansów

PRACA DOKTORSKA

Łukasz Łabanowski

MODELOWANIE PREDYKTYWNE WARTOŚCI ZYSKU OPERACYJNEGO BRUTTO DETERMINANTĄ OCENY EFEKTYWNOŚCI OPERACYJNEJ WYBRANYCH HOTELI. ZASTOSOWANIE LINIOWYCH I NIELINIOWYCH ALGORYTMÓW REGRESYJNYCH

PREDICTIVE MODELLING OF GROSS OPERATING PROFIT AS A DETERMINANT OF OPERATIONAL EFFICIENCY ASSESSMENT IN SELECTED HOTELS. APPLICATION OF LINEAR AND NONLINEAR REGRESSION ALGORITHMS

Praca napisana pod kierunkiem
dr hab. Łukasz Mach, prof. UO

Promotor pomocniczy
Dr Robert Poskart

Opole 2025 r.

Spis treści

1. WSTĘP ... 3

1.1. WPROWADZENIE ... 3
1.2. ZNACZENIE ROZPRAWY DOKTORSKIEJ I JEJ NOWATORSKI CHARAKTER ... 6
1.3. METODYKA BADAŃ ... 7
1.4. PREZENTACJA POSZCZEGÓLNYCH ROZDZIAŁÓW PRACY ... 9

2. KONCEPCJA ZYSKU OPERACYJNEGO BRUTTO JAKO DETERMINANTA OCENY SPRAWNOŚCI OPERACYJNEJ W BRANŻY HOTELARSKIEJ ... 12

2.1. EFEKTYWNOŚĆ FINANSOWA BRANŻY HOTELARSKIEJ I JEJ KLUCZOWE WSKAŹNIKI ... 12
2.2. STANDARDY RACHUNKOWOŚCI STOSOWANE W HOTELARSTWIE ... 15
2.3. ZYSK OPERACYJNY BRUTTO JAKO KLUCZOWA SKŁADOWA OCENY DZIAŁALNOŚCI OPERACYJNEJ W BRANŻY HOTELARSKIEJ ... 21

3. PRZEGLĄD LITERATUROWY STOSOWANIA METOD ANALITYCZNYCH W BRANŻY HOTELARSKIEJ ... 23

3.1. PRZEGLĄD WYBRANYCH METOD ANALIZY I ICH WPŁYW NA BRANŻE HOTELARSKĄ ... 23
3.2. METODY PREDYKCJI STOSOWANE W BRANŻY TURYSTYCZNEJ – NA PODSTAWIE LITERATURY NAUKOWEJ ... 26
3.3. ZAGADNIENIA TEORETYCZNE DOT. WYBRANYCH METOD PREDYKCJI SZEREGÓW CZASOWYCH ... 39

4. ANALIZA ZMIENNOŚCI ZYSKU OPERACYJNEGO BRUTTO DLA WYBRANYCH POLSKICH HOTELI W DZIEDZINIE CZASU I CZĘSTOTLIWOŚCI ... 44

4.1. ZMIENNOŚĆ CZASOWA ZYSKU OPERACYJNEGO BRUTTO ... 45
4.2. ANALIZA ROZKŁADÓW EMPIRYCZNYCH ZYSKU OPERACYJNEGO BRUTTO ... 48
4.3. ANALIZA ZMIENNOŚCI PERIODYCZNEJ ... 55
4.4. ANALIZA ZMIENNOŚCI ZYSKU OPERACYJNEGO BRUTTO W DZIEDZINIE CZĘSTOTLIWOŚCI ... 59

5. OCENA WSPÓŁLINIOWOŚCI ZYSKU OPERACYJNEGO BRUTTO PODSTAWĄ WYBORU REGRESORÓW DO MODELU PREDYKCYJNEGO ... 62

5.1. ANALIZA KORELACJI LINIOWEJ ... 62
5.2. WYZNACZENIE ZMIENNYCH USALI Z UWZGLĘDNIENIEM WŁASNOŚCI WSPÓŁLINIOWOŚCI ... 68
5.3. ANALIZA KOHERENCJI FALKOWEJ ... 75

6. ANALIZA ZMIENNOŚCI PREDYKTORÓW OBJAŚNIAJĄCYCH ZYSK OPERACYJNY BRUTTO ... 81

6.1. ANALIZA ZMIENNOŚCI SZEREGÓW CZASOWYCH PREDYKTORÓW ... 81
6.2. ANALIZA PERIODYCZNOŚCI PREDYKTORÓW ... 82
6.3. ANALIZA PREDYKTORÓW W DZIEDZINIE CZĘSTOTLIWOŚCI ... 84

7. OPRACOWANIE MODELU PREDYKCYJNEGO ZYSKU OPERACYJNEGO BRUTTO ... 87

7.1. MODELOWANIE ZYSKU OPERACYJNEGO BRUTTO Z WYKORZYSTANIEM MODELU EKONOMETRYCZNEGO ... 87
7.1.1. PROCES BUDOWY MODELU PREDYKCYJNEGO Z WYKORZYSTANIEM NIELINIOWYCH ALGORYTMÓW REGRESYJNYCH I UCZENIA MASZYNOWEGO ... 102
7.1.2. ANALIZA ZMIENNYCH USALI JAKO PREDYKTORÓW W MODELACH NIELINIOWYCH ... 104
7.1.3. MODELOWANIE ZYSKU OPERACYJNEGO BRUTTO Z WYKORZYSTANIEM NIELINIOWYCH ALGORYTMÓW REGRESYJNYCH ... 107

8. PODSUMOWANIE I WNIOSKI ... 121

9. LITERATURA ... 125

10. SPIS TABEL ... 135

11. SPIS RYSUNKÓW ... 136

ZAŁĄCZNIKI ... 138

1. Wstęp

1.1. Wprowadzenie

W dynamicznie zmieniającym się środowisku branży hotelarskiej analiza i ocena efektywności operacyjnej przedsiębiorstw nabierają szczególnego znaczenia. Wysoka konkurencja, sezonowość oraz wpływ czynników zewnętrznych, takich jak sytuacja gospodarcza czy preferencje konsumentów, stawiają przed menedżerami hoteli wyzwania związane z optymalizacją kosztów oraz maksymalizacją zysków. W tym kontekście zysk operacyjny brutto jako kluczowy wskaźnik oceny efektywności operacyjnej, staje się istotnym przedmiotem badań, mającym wpływ na podejmowanie strategicznych decyzji biznesowych.

Niniejsza dysertacja szczegółowo analizuje zagadnienia z zakresu analizy danych ekonomicznych, skupiając się na wykorzystaniu nowoczesnych metod i technik analitycznych należących do dziedziny data science. W ramach przeprowadzonych badań zastosowano zaawansowany aparat matematyczny, co pozwoliło na kompleksową interpretację wyników analiz w szerokim kontekście. Wykorzystując metody analizy czasowej, czasowo-częstotliwościowej oraz analizy korelacyjnej, a także korzystając z zaawansowanych algorytmów uczenia maszynowego, opracowano model predykcyjny. Model ten jest zdolny do przewidywania wartości zysku operacyjnego brutto GOP (*ang. Gross Operating Profit*) przedsiębiorstwa hotelowego, wykorzystując do tego celu wybrane wskaźniki USALI (*Uniform System of Accounts for the Lodging Industry*).

Zalety zastosowania modelu predykcyjnego w celu zwiększenia efektywności finansowej przedsiębiorstwa hotelowego są wielowymiarowe i mają zasadnicze znaczenie dla strategicznego planowania oraz operacyjnego zarządzania. Po pierwsze, model ten umożliwia dokładne prognozowanie przyszłych przychodów i kosztów, co jest kluczowe dla efektywnego zarządzania cash flow i maksymalizacji rentowności. Dzięki temu menedżerowie mogą lepiej planować inwestycje, optymalizować koszty operacyjne i dostosowywać strategie cenowe w odpowiedzi na przewidywane zmiany w dynamice rynku. Po drugie, model predykcyjny, wykorzystując modele ekonometryczne oraz uczenie maszynowe i zaawansowane algorytmy analityczne, jest w stanie identyfikować ukryte wzorce i zależności w danych historycznych, co otwiera nowe możliwości dla odkrywania potencjalnych źródeł przychodów oraz minimalizowania ryzyka finansowego. Precyzyjne prognozy umożliwiają również efektywne zarządzanie zasobami hotelu, w tym optymalizację zatrudnienia, zarządzanie dostępnością pokoi i usług dodatkowych, co bezpośrednio przekłada się na zwiększenie satysfakcji klienta i lojalności. Po trzecie, wdrożenie modelu predykcyjnego sprzyja zwiększeniu

konkurencyjności przedsiębiorstwa hotelowego na rynku poprzez umożliwienie szybkiego reagowania na zmieniające się warunki rynkowe i preferencje konsumentów. Analiza predykcyjna pozwala na identyfikację trendów rynkowych i dostosowanie oferty hotelu w taki sposób, aby maksymalizować dochody i udział w rynku. Wreszcie, zastosowanie modelu predykcyjnego w przedsiębiorstwie hotelowym znacząco przyczynia się do poprawy kultury decyzyjnej opartej na danych (*ang*. *data-driven decision making*), promując podejmowanie decyzji na podstawie analiz i empirycznych dowodów. To z kolei prowadzi do zwiększenia efektywności operacyjnej, lepszego zarządzania ryzykiem i optymalizacji procesów decyzyjnych na wszystkich poziomach organizacji. Podsumowując, integracja modelu predykcyjnego w strategię finansową przedsiębiorstwa hotelowego oferuje znaczące korzyści, w tym lepsze prognozowanie finansowe, optymalizację operacyjną, zwiększenie konkurencyjności oraz promowanie kultury decyzyjnej opartej na analizie danych. Taka holistyczna aplikacja metod ekonometrycznych oraz *data science* w sektorze hotelarskim nie tylko zwiększa efektywność finansową, ale również stanowi podstawę do zrównoważonego wzrostu i innowacji w branży.

Na potrzeby pracy badawczej sformułowano następujące hipotezy badawcze:

Hipoteza badawcza 1: *Istnieje istotna zależność między wybranymi wskaźnikami finansowymi a zyskiem operacyjnym brutto (GOP) w branży hotelarskiej w Polsce.*

Hipoteza badawcza 2: *Wyselekcjonowany zestaw wskaźników finansowych, skorelowanych ze zmienną GOP, może być użyty do skutecznej predykcji zysku operacyjnego brutto (GOP) w hotelach.*

Hipoteza badawcza 3: *Istnieją sezonowe/cykliczne wzorce w wynikach operacyjnych hoteli w Polsce.*

Hipoteza badawcza 4: *Nowoczesne technologie sztucznej inteligencji, takie jak uczenie maszynowe i sieci neuronowe, mogą skutecznie zastąpić klasyczne metody analityczne w modelowaniu zysku operacyjnego brutto, eliminując konieczność spełnienia założeń wymaganych przez klasyczne metody statystyczne, przy jednoczesnym zapewnieniu porównywalnej lub wyższej dokładności wyników.*

Wspomniane hipotezy stanowią podstawę do badań i analiz, które mogą przyczynić się do lepszego zrozumienia skomplikowanej relacji między wskaźnikami finansowymi a wynikami operacyjnymi w branży hotelarskiej. Celem weryfikacji postawionych hipotez badawczych, w ramach prac zrealizowano następujące etapy:

- Wykonano przegląd piśmiennictwa na temat stosowanych w branży hotelarskiej standardów rachunkowości, wskaźników i statystyk ekonomicznych oraz możliwości ich przewidywania z wykorzystaniem nowoczesnych technologii sztucznej inteligencji, na podstawie którego stwierdzono, że mimo licznie publikowanych prac, nadal istnieje potrzeba doskonalenia metod i narzędzi pozwalających na optymalizacje zarządzania hotelem, celem maksymalizacji zysków.
- Zgromadzono i poddano analizie historyczne dane finansowe dotyczące siedmiu hoteli w Polsce. W szczególności wskaźniki finansowe bazy USALI przebadano następującymi metodami:
  - analizy przebiegów czasowych,
  - analizy periodyczności,
  - analizy rozkładów funkcji gęstości prawdopodobieństwa,
  - analizy sezonowości/cykliczności w domenie częstotliwościowej,
  - analizy współzależności zmiennej zysk operacyjny brutto (GOP) względem pozostałych wskaźników USALI,
  - analizy rozrzutu i współczynnika korelacji,
  - analizy koherencji falkowej w domenie czasowo-częstotliwościowej,
  - analizy istotności wskaźników na możliwości predykcyjne zmiennej GOP za pomocą metod MRMR i statystyki Fishera-Snedecora F.
- Na podstawie uzyskanych wyników analiz wyselekcjonowano zestaw wskaźników skorelowanych ze zmienną GOP.
- Wyselekcjonowane cechy wykorzystano do opracowania modelu ekonometrycznego oraz modeli sztucznej inteligencji do predykcji GOP.
- Wskazano algorytm, który pozwala na skuteczną predykcję zmiennej GOP.

Podsumowując powyższe:

Celem pracy jest zbadanie i ocena wpływu wybranych wskaźników finansowych na zysk operacyjny brutto w branży hotelarskiej w Polsce oraz opracowanie modelu predykcyjnego, z uwzględnieniem możliwości zastosowania nowoczesnych technologii sztucznej inteligencji jako alternatywy dla klasycznych metod analitycznych w modelowaniu i prognozowaniu wyników finansowych.

W procesie realizacji celu pracy, skupiono się na identyfikacji kluczowych determinant efektywności operacyjnej hoteli, uwzględniając zarówno aspekty czasowe, jak i częstotliwościowe zmienności zysku. Osiągnięcie tego celu wymaga zastosowana zaawansowanych metod statystycznych i ekonometrycznych, które pozwolą na kompleksowe zrozumienie dynamiki zmian zysku operacyjnego brutto oraz umożliwią stworzenie narzędzi

wspierających zarządzanie w hotelarstwie. Dzięki realizacji tego celu praca ma przyczynić się do wypełnienia luki badawczej w zakresie analizy kształtowania się wielkości zysku operacyjnego brutto w hotelarstwie, dostarczając zarówno wiedzy teoretycznej, jak i praktycznych narzędzi analitycznych dla menedżerów i analityków finansowych.

Niniejsza dysertacja dotyczy wrażliwych danych finansowych wybranych hoteli w Polsce, dlatego przeprowadzono proces standaryzacji danych.

1.2. Znaczenie rozprawy doktorskiej i jej nowatorski charakter

Znaczenie rozprawy doktorskiej jest ukierunkowane na zrozumienie i ocenę procesów kształtowania się wielkości zysku operacyjnego brutto jako kluczowego wskaźnika efektywności operacyjnej w branży hotelarskiej. W sektorze hotelarskim, który charakteryzuje się dużą zmiennością i zależnością od wielu czynników zewnętrznych, analiza efektywności operacyjnej odgrywa istotną rolę w działalności ekonomiczno-finansowej branży hotelarskiej.

Rozprawa koncentruje się na kilku kluczowych aspektach. Przede wszystkim, praca ta podkreśla znaczenie zysku operacyjnego brutto jako centralnego wskaźnika, który pozwala na obiektywną ocenę efektywności operacyjnej hoteli. W dotychczasowej literaturze przedmiotu, choć istnieje wiele wskaźników finansowych, to zysk operacyjny brutto jest często pomijany w kontekście dogłębnej analizy w branży hotelarskiej. Praca ta wypełnia tę lukę, dostarczając szczegółowej analizy procesu (lub parametrów wpływających na) kształtowania się wielkości tego wskaźnika, biorąc pod uwagę specyfikę i złożoność rynku hotelarskiego.

Kolejnym znaczącym i nowatorskim elementem rozprawy jest przegląd i zastosowanie różnorodnych metod analitycznych i predykcyjnych w badaniu efektywności operacyjnej hoteli. Branża hotelarska jest narażona na szereg czynników sezonowych, społecznych i ekonomicznych, które wpływają na jej kondycję finansową. W związku z tym, analiza zmienności zysku operacyjnego brutto w czasie i częstotliwości, jak również ocena współliniowości predyktorów, stanowi ważne narzędzie w zarządzaniu finansami w sektorze hotelarskim. Rozprawa doktorska dostarcza naukowej podstawy dla praktyków w hotelarstwie, umożliwiając podejmowanie trafnych i lepiej uzasadnionych decyzji finansowych i operacyjnych.

Dodatkowo, innowacyjność rozprawy przejawia się w ocenie współliniowości różnych czynników wpływających na zysk operacyjny brutto. Zastosowanie narzędzi analizy współliniowości oraz identyfikacja kluczowych predyktorów do modelu predykcyjnego

pozwala na bardziej precyzyjne prognozowanie przyszłych wyników finansowych. Wprowadzenie tego rodzaju analizy do badania efektywności operacyjnej hoteli stanowi nowość, która ma potencjał praktycznego zastosowania w branży.

Opracowanie modelu predykcyjnego, bazującego na zaawansowanych metodach analitycznych, stanowi kolejny element nowatorski. Model ten może być wykorzystywany przez menedżerów hoteli do podejmowania lepszych decyzji strategicznych, optymalizacji kosztów oraz maksymalizacji zysków. Tym samym, praca dostarcza narzędzi, które wykraczają poza tradycyjne metody analizy finansowej, oferując bardziej zaawansowane i precyzyjne rozwiązania dostosowane do specyfiki branży hotelarskiej.

Podsumowując, znaczenie rozprawy doktorskiej oraz jej nowatorski charakter polega na wprowadzeniu nowego podejścia do analizy i oceny zysku operacyjnego brutto w hotelarstwie poprzez zastosowanie zaawansowanych technik analitycznych i modelowania. Praca łączy teorię z praktycznymi narzędziami, co czyni ją cenną zarówno z perspektywy naukowej, jak i dla praktyków biznesu w branży hotelarskiej. Dzięki temu rozprawa wnosi istotny wkład w rozwój wiedzy i narzędzi analizy finansowej w sektorze hotelarskim, otwierając nowe możliwości w zakresie prognozowania i zarządzania operacyjnego.

1.3. Metodyka badań

Metodyka badań wyjaśnia podejście oraz narzędzia analityczne, które zostały wykorzystane do przeprowadzenia analizy zysku operacyjnego brutto w branży hotelarskiej. Niniejsza dysertacja charakteryzuje się kompleksowym podejściem badawczym, łącząc różnorodne metody analityczne, predykcyjne oraz statystyczne, aby uzyskać jak najpełniejszy obraz funkcjonowania i efektywności operacyjnej w branży hotelowej w Polsce.

Metodyka badań obejmuje następujące kluczowe elementy:

1. Analiza literatury przedmiotu: W ramach rozdziału trzeciego, który stanowi przegląd literatury dotyczącej metod analitycznych stosowanych w hotelarstwie, dokonano przeglądu istniejących badań naukowych i literatury branżowej. Przegląd ten ma na celu zrozumienie kluczowych wskaźników finansowych, metod predykcyjnych i narzędzi analitycznych, które są istotne w analizie zysku operacyjnego brutto. Ta część stanowi teoretyczną podstawę dla późniejszej analizy empirycznej, a także pomaga w wyborze odpowiednich metod badawczych.
2. Analiza efektywności operacyjnej i standardów rachunkowości: Drugi rozdział skupia się na analizie kluczowych wskaźników efektywności finansowej w branży hotelarskiej oraz standardów rachunkowości. W tym kontekście praca wykorzystuje narzędzia

finansowe i standardy rachunkowości do zrozumienia, jak w branży hotelarskiej określa się i ocenia zysk operacyjny brutto. Ta część analizy pozwala na identyfikację kluczowych elementów wpływających na wynik finansowy hoteli.

3. Analiza empiryczna zysku operacyjnego brutto: Głównym elementem metodyki jest analiza empiryczna zysku operacyjnego brutto w hotelach, która rozciąga się na kilka rozdziałów pracy. W rozdziale czwartym, praca skupia się na badaniu zmienności zysku operacyjnego w różnych wymiarach – zarówno czasowych, jak i częstotliwościowych. W tej części stosowane są metody statystyczne, takie jak analiza rozkładów empirycznych oraz periodyczności. Rozdział ten wykorzystuje techniki analizy szeregów czasowych, pozwalając na zbadanie dynamiki zmian zysku operacyjnego w wybranych hotelach. Analiza w dziedzinie częstotliwości umożliwia również identyfikację okresowych wzorców w danych finansowych.
4. Ocena współliniowości i korelacji: W rozdziale piątym praca koncentruje się na ocenie współliniowości zysku operacyjnego brutto, co jest kluczowe dla wyboru właściwych zmiennych (regresorów) w modelach predykcyjnych. Wykorzystuje się tu analizę korelacji liniowej oraz analizę koherencji falkowej. Te narzędzia statystyczne pozwalają na zidentyfikowanie, które czynniki są ze sobą skorelowane oraz w jakim stopniu wpływają na zmienność zysku operacyjnego, co jest niezbędne dla stworzenia dokładnych modeli predykcyjnych.
5. Analiza zmienności predyktorów: Rozdział szósty opisuje szczegółową analizę zmienności predyktorów objaśniających zysk operacyjny brutto. Stosowane są tutaj zaawansowane techniki analizy szeregów czasowych, analizy rozkładów oraz periodyczności. Wykorzystanie tych metod umożliwia zrozumienie, w jaki sposób zmieniają się różne predyktory (zmienne objaśniające) i jak wpływają na poziom zysku operacyjnego brutto. Analiza ta jest kluczowa dla opracowania późniejszego modelu predykcyjnego.
6. Budowa modelu predykcyjnego: Najważniejsza część metodyki opisana jest w rozdziale siódmym, gdzie praca skupia się na opracowaniu modelu predykcyjnego zysku operacyjnego brutto. W pierwszej kolejności analizowane są różne metody predykcji stosowane w branży turystycznej, co pozwala na wybór najodpowiedniejszych technik. Następnie opisany jest proces budowy modelu predykcyjnego, w którym zastosowane są zarówno modele ekonometryczne, jak i nieliniowe algorytmy regresyjne. W tej części praca wykorzystuje metody statystyczne oraz uczenia maszynowego, aby stworzyć modele zdolne do prognozowania zysku

operacyjnego brutto w hotelach na podstawie zidentyfikowanych wcześniej predyktorów.

7. Analiza wyników i weryfikacja modeli: Chociaż nie jest to bezpośrednio wspomniane w spisie treści a wynika z kontekstu pracy, że zastosowane metody analizy i modelowania muszą zostać poddane ocenie i weryfikacji. W tym celu wykorzystuje się różne miary błędu prognoz, a także analizę wyników modelowania, aby ocenić trafność i skuteczność zaproponowanych modeli predykcyjnych.

Podsumowując, metodyka badań zastosowana w pracy łączy przegląd literatury naukowej, analizę finansową, zaawansowane techniki analizy szeregów czasowych i częstotliwościowych, ocenę współliniowości oraz budowę modeli predykcyjnych przy użyciu metod ekonometrycznych i algorytmów nieliniowych. Takie podejście pozwala na uzyskanie kompleksowego obrazu wpływu różnych czynników na zysk operacyjny brutto w branży hotelarskiej i umożliwia tworzenie skutecznych narzędzi predykcyjnych w tej dziedzinie.

1.4. Prezentacja poszczególnych rozdziałów pracy

Rozdział ten stanowi szczegółową prezentację struktury pracy, która koncentruje się na analizie zysku operacyjnego brutto jako kluczowego elementu oceny efektywności operacyjnej w branży hotelarskiej. Wstępna część pracy składa się z rozdziału pierwszego, który pełni funkcję wprowadzenia do tematyki badawczej. Omawia ono główne założenia pracy, jej znaczenie dla dziedziny naukowej oraz wyjaśnia cel badań. W rozdziale tym znajdziemy również opis koncepcji i planu pracy oraz zastosowanej metodyki badawczej. Jest to fundament całej rozprawy, uzasadniający, dlaczego zysk operacyjny brutto w hotelarstwie został wybrany jako przedmiot analizy i jakie metody badawcze będą zastosowane w dalszych częściach pracy. Drugi rozdział skupia się na koncepcji zysku operacyjnego brutto jako determinancie oceny efektywności operacyjnej w hotelarstwie. Rozpoczyna się od analizy efektywności operacyjnej branży hotelarskiej, w tym omówienia kluczowych wskaźników finansowych. Następnie omawia standardy rachunkowości stosowane w hotelarstwie, które są istotne dla zrozumienia specyfiki funkcjonowania tego sektora. W tej części pracy szczególny nacisk położony jest na zysk operacyjny brutto, ukazując go jako jeden z najważniejszych wskaźników do oceny działalności operacyjnej hoteli. Dzięki takiemu podejściu czytelnik może zrozumieć, jak ważną rolę odgrywa ten wskaźnik w ocenie efektywności finansowej w branży.

Kolejny, trzeci rozdział stanowi przegląd literatury dotyczącej stosowania metod analitycznych w branży hotelarskiej. Rozdział ten prezentuje różne metody analizy oraz ich

wpływ na branżę hotelarską, co pozwala na zrozumienie obecnych trendów i praktyk w tej dziedzinie. W szczególności opisane są tu metody predykcji stosowane w branży turystycznej oraz zagadnienia teoretyczne związane z wybranymi metodami predykcji szeregów czasowych. Takie podejście umożliwia czytelnikowi zapoznanie się z naukowym kontekstem pracy oraz wskazuje na różnorodność metod analitycznych dostępnych dla branży hotelarskiej.

Czwarty rozdział skupia się na analizie zmienności zysku operacyjnego brutto dla wybranych polskich hoteli, zarówno w dziedzinie czasu, jak i częstotliwości. W pierwszej kolejności analizowana jest zmienność czasowa tego wskaźnika, co pozwala na zrozumienie, jak zysk operacyjny zmienia się w określonych ramach czasowych. Następnie przeprowadzona jest analiza rozkładów empirycznych zysku operacyjnego brutto, dzięki której można zbadać różnorodność i specyfikę zmian w zysku w wybranych hotelach. Kolejną częścią jest analiza zmienności periodycznej oraz analiza w dziedzinie częstotliwości, które pomagają zidentyfikować wzorce i trendy w danych finansowych.

Rozdział piąty jest poświęcony ocenie współliniowości zysku operacyjnego brutto jako podstawy wyboru regresorów do modelu predykcyjnego. W ramach tego rozdziału opisane są różne metody oceny współliniowości, a także przeprowadzona zostaje analiza korelacji liniowej, która jest uzupełniona analizą koherencji falkowej. Ta część pracy ma kluczowe znaczenie, ponieważ pozwala na identyfikację zmiennych mających istotny wpływ na zysk operacyjny brutto, co umożliwia budowę dokładniejszego modelu predykcyjnego.

Szósty rozdział koncentruje się na analizie zmienności predyktorów objaśniających zysk operacyjny brutto. Przedstawione są tutaj różne aspekty zmienności szeregów czasowych predyktorów, ich rozkładów oraz periodyczności. Rozdział ten umożliwia zidentyfikowanie czynników wpływających na kształtowanie się zysku operacyjnego i pomaga w opracowaniu bardziej precyzyjnych modeli predykcyjnych. Analiza predyktorów w dziedzinie częstotliwości stanowi kluczowy element tej części pracy, gdyż pozwala na zrozumienie, jakie czynniki mają największe znaczenie dla zmienności zysku operacyjnego.

Siódmy rozdział jest jednym z najważniejszych w całej pracy, ponieważ skupia się na opracowaniu modelu predykcyjnego zysku operacyjnego brutto. Na początku omawiane są metody predykcji stosowane w branży turystycznej, bazując na literaturze naukowej. Następnie przeprowadzana jest szczegółowa analiza teoretyczna wybranych metod predykcji szeregów czasowych. Rozdział ten zawiera także szczegółowy opis procesu budowy modelu predykcyjnego, analizę cech jako predyktorów oraz praktyczne modelowanie zysku operacyjnego przy użyciu modeli ekonometrycznych oraz nieliniowych algorytmów

regresyjnych. Dzięki temu rozdziałowi czytelnik może zrozumieć proces tworzenia modelu predykcyjnego oraz jego praktyczne zastosowanie w analizie operacyjnej hoteli.

Ostatni, ósmy rozdział stanowi podsumowanie i prezentuje kluczowe wnioski płynące z przeprowadzonych badań. Jest to zwieńczenie całej pracy, które pozwala na refleksję nad osiągnięciami badawczymi oraz wkładem w dziedzinę analizy finansowej w branży hotelarskiej. Podsumowanie wskazuje, jakie były główne ustalenia pracy oraz jakie ma ona implikacje dla praktyki zarządzania i finansów w hotelarstwie.

Całość pracy stanowi próbą kompleksowej analizy zysku operacyjnego brutto w branży hotelarskiej, wykorzystując różne metody analityczne i predykcyjne, co przyczynia się do pogłębienia wiedzy na temat efektywności operacyjnej w tym sektorze.

2. Koncepcja zysku operacyjnego brutto jako determinanta oceny efektywności operacyjnej w branży hotelarskiej

2.1. Efektywność finansowa branży hotelarskiej i jej kluczowe wskaźniki

Efektywność finansowa (EF) stanowi kluczowy element analizy ekonomicznej, oceniający zdolność jednostki, przedsiębiorstwa lub organizacji do osiągania maksymalnych rezultatów przy minimalnym zużyciu zasobów (Barburski, 2010; R. Dyduch, 2017; Farrell, 1957; Filip, 2011; Jajuga & Jajuga, 2006; Johnes, 2004; Le, Chuc, & Taghizadeh-Hesary, 2019; Nam, Bon Sesay, Wynne, & Zhang, 2020; Wang & Sarkis, 2017; Yuan, Wu, & Liu, 2022). W zakresie analizy finansowej przedsiębiorstwa, efektywność finansowa odnosi się do optymalnego wykorzystania kapitału i zasobów w celu generowania wartość m.in. w obszarach:

- analizy rentowności zyskowności,
- zarządzania przepływem gotówkowym,
- analizy wskaźników rentowności inwestycji,
- zarządzania ryzykiem finansowym,
- efektywności operacyjnej,
- elastyczności finansowej.

Odnosi się ona również do zdolności organizacji do generowania zysków w stosunku do poniesionych kosztów i zaangażowanych zasobów. Przedsiębiorstwa dążą do maksymalizacji efektywności finansowej, co pozwala na lepsze wykorzystanie zasobów, optymalizację kosztów oraz zwiększenie wartości dla akcjonariuszy. Wysoka efektywność finansowa świadczy o zdrowiu finansowym firmy i jej zdolności do realizacji celów strategicznych w długim okresie.

W kontekście branży hotelarskiej, efektywność finansowa ma kluczowe znaczenie, ponieważ hotele są zmuszone do działalności w sektorze o wysokiej konkurencyjności, gdzie umiejętność optymalizacji kosztów przy jednoczesnym maksymalizowaniu przychodów jest niezbędna do utrzymania rentowności i przewagi konkurencyjnej (Gursoy, Buttle, & Bowie, 2022a).

Należy zaznaczyć, że do oceny efektywności finansowej w branży hotelarskiej wykorzystywane są różnorodne wskaźniki finansowe. Najważniejszymi z nich są zysk operacyjny brutto (GOP), przychód na dostępny pokój (RevPAR), średnia dzienna stawka (ADR) oraz wskaźnik obłożenia. Należy również przy analizie efektywności finansowej pamiętać o pozostałych wskaźnikach tj. ADRRS, TREVPAR, OCC, LC, APPAR, EBIDTAPAR, GOPPAR (Peter Harris & Mongiello, 2012). Szerzej o standardach

rachunkowości stosowanych w hotelarstwie, a więc o standardzie USALI napisano w rozdziale 2.2. niniejszej monografii pt. Standardy rachunkowości stosowane w hotelarstwie.

GOP (*ang. Gross Operating Profit*) jest kluczowym wskaźnikiem rentowności operacyjnej, który mierzy zdolność hotelu do generowania zysków z działalności operacyjnej po odliczeniu kosztów operacyjnych, ale przed uwzględnieniem kosztów amortyzacji, podatków, odsetek i innych kosztów finansowych. GOP odzwierciedla efektywność zarządzania codziennymi operacjami hotelu i jest jednym z głównych wskaźników oceny jego kondycji finansowej (Jagels & Ralston, 2007). Wysoka wartość GOP wskazuje na efektywne zarządzanie kosztami operacyjnymi oraz skuteczną strategię maksymalizacji przychodów, co jest niezbędne dla długoterminowego sukcesu finansowego hotelu (P. Harris & Hazzard, 1992). W praktyce GOP jest często używany przez menedżerów hoteli jako główny wskaźnik przy podejmowaniu decyzji operacyjnych. Analiza tego wskaźnika pozwala na identyfikację obszarów wymagających optymalizacji, takich jak kontrola kosztów operacyjnych, efektywność zarządzania zasobami, czy poprawa polityki cenowej. GOP dostarcza również wartościowych informacji dla właścicieli i inwestorów, umożliwiając im ocenę zwrotu z inwestycji oraz efektywności zarządzania obiektem (Bragg, 2011). Podsumowując GOP to marża zysku brutto, wyrażona w procentach. Jest to wskaźnik finansowy, który określa całkowity zysk generowany przez przedsiębiorstwo hotelarskie na podstawie jego operacji głównych. Obejmuje przychody związane z działalnością operacyjną, takie jak dochody z wynajmu pokoi, usługi gastronomiczne, usługi spa, sala konferencyjna itp., pomniejszone o koszty związane z tymi działaniami, takie jak koszty pracy, koszty zakupu materiałów, koszty utrzymania i eksploatacji, podatki, opłaty itp. Ten wskaźnik, a zwłaszcza jego wartość relatywna (wartość GOP/przychód netto x 100%), stanowi podstawowy element oceny efektywności operacyjnej hotelu jako przedsiębiorstwa.

RevPAR (*ang. Revenue per Available Room*) to wskaźnik przychodów na dostępny pokój, który stanowi połączenie wskaźnika obłożenia oraz ADR. RevPAR pozwala na ocenę efektywności zarządzania przychodami z pokoi, uwzględniając zarówno ceny, jak i stopień obłożenia (Cote, 2001). Dzięki temu wskaźnikowi, menedżerowie mogą monitorować, jak dobrze hotel radzi sobie w maksymalizowaniu przychodów z dostępnych pokoi. Uszczegóławiając jest to wskaźnik mierzący przychód ze sprzedaży ogółem podzielony przez

liczbę wszystkich pokoi dostępnych w danym okresie. Informuje on o średnim przychodzie na jeden dostępny pokój, biorąc pod uwagę bieżące stawki obłożenia. REVPAR jest istotnym wskaźnikiem w branży hotelarskiej, umożliwiającym monitorowanie wydajności finansowej hotelu. Należy jednak pamiętać, że nie uwzględnia on kosztów.

ADR (*ang. Average Daily Rate*), czyli średnia dzienna stawka, to wskaźnik odzwierciedlający średni dochód uzyskiwany za wynajęty pokój. Jest on istotny dla oceny polityki cenowej hotelu i jego pozycji na rynku. Im wyższy ADR, tym lepsza jest zdolność hotelu do przyciągania klientów gotowych zapłacić wyższe ceny za nocleg (Atkinson & Brander Brown, 2001). Wskaźnik obłożenia mierzy procentowy udział wynajętych pokoi w stosunku do wszystkich dostępnych pokoi w danym okresie. Jest to wskaźnik, który pozwala na ocenę skuteczności działań marketingowych i operacyjnych hotelu, wpływających na przyciąganie klientów. Wysoki wskaźnik obłożenia zazwyczaj prowadzi do wyższych przychodów, jednak musi być on analizowany w połączeniu z ADR, aby w pełni ocenić efektywność finansową (Peter Harris & Mongiello, 2012). Oblicza się go poprzez podzielenie łącznego przychodu ze wszystkich zajętych pokoi w danym dniu przez liczbę tych pokoi. ADR umożliwia zatem monitorowanie efektywności cenowej oraz dochodowości poszczególnych pokoi.

ADRRS (*ang. Average Daily Rate by Rate Segment*), jest to średnia cena pokoju w określonym segmencie sprzedaży, na przykład biznesowym lub weselnym. Segmentacja gości hotelowych jest ważna nie tylko z perspektywy planowania przychodów, ale także pozwala identyfikować grupy gości, które generują najwyższe przychody związane z ceną pokoju.

TREVPAR (*ang. Total Revenue Per Available Room*), jest to wskaźnik mierzący całkowity przychód wygenerowany przez hotel na jeden dostępny pokój. Obejmuje on wszystkie źródła przychodu związane z nieruchomością hotelową, takie jak sprzedaż z barów, obsługa pokoju, śniadania, spa, usługi sportowe itp. TREVPAR dostarcza informacji o całkowitym dochodzie generowanym przez hotel w kontekście dostępnych pokoi, umożliwiając ocenę ogólnej wydajności operacyjnej hotelu. Jest to istotny wskaźnik dla branży hotelarskiej, pozwalający na monitorowanie efektywności dochodowej i zarządzanie przychodami z różnych źródeł wewnątrz hotelu.

OCC (*ang. Occupancy*), oznacza obłożenie pokoi w hotelu i jest wyrażane jako procent zajętych pokoi w stosunku do całkowitej liczby dostępnych pokoi. OCC jest kluczowym wskaźnikiem, który informuje o poziomie wykorzystania potencjału noclegowego hotelu.

LC (*ang. Labor Cost*), oznacza koszty związane z pracą, które oblicza się poprzez podzielenie kosztów osobowych przez przychody ze sprzedaży obiektu lub działu, a następnie pomnożenie wyniku przez 100%. Jest to wskaźnik, który informuje o stosunku kosztów pracy do przychodów, służący do monitorowania efektywności zarządzania kosztami personalnymi.

ARPAR (*ang. Average Revenue Per Available Room*), to średnie koszty zmienne w przychodzie na jeden dostępny pokój. Jest to wskaźnik zaproponowany przez Ira Vouk, doświadczonego konsultanta i doradcę hotelowego, który pozwala na analizę kosztów zmiennych w odniesieniu do dostępnych pokoi.

EBIDTAPAR (*ang. Earnings Before Interest, Taxes, Depreciation, Amortization, Per Available Room*), jest to dochód przed kosztami finansowymi, podatkami, amortyzacją i podzielony przez liczbę dostępnych pokoi. Ten wskaźnik dostarcza informacji o efektywności operacyjnej hotelu, uwzględniając podstawowe czynniki finansowe.

GOPPAR (*ang. Gross Operating Profit Per Available* Room) jest uzyskiwany przez podzielenie zysku operacyjnego brutto przez liczbę dostępnych pokoi. Wskaźnik ten pozwala na ocenę rentowności operacyjnej hotelu i stanowi ważną miarę efektywności działalności.

Powyższe wskaźniki nie są tylko narzędziami oceny efektywności finansowej, ale również umożliwiają menedżerom hotelowym podejmowanie decyzji strategicznych, które mają na celu poprawę rentowności. Przykładowo, analiza GOP może pomóc w identyfikacji nieefektywności w zarządzaniu kosztami, natomiast RevPAR i ADR mogą dostarczyć informacji na temat skuteczności strategii cenowych i marketingowych. Podsumowując, efektywność finansowa w branży hotelarskiej jest mierzona za pomocą szeregu kluczowych wskaźników, które pozwalają na kompleksową ocenę rentowności operacyjnej i skuteczności zarządzania. GOP, ze względu na swoją zdolność do odzwierciedlania rzeczywistej rentowności operacyjnej, odgrywa szczególną rolę w monitorowaniu i poprawie efektywności finansowej hoteli. Dążenie do optymalizacji tych wskaźników jest niezbędne dla utrzymania konkurencyjności i długoterminowego sukcesu na dynamicznym rynku hotelarskim.

2.2. Standardy rachunkowości stosowane w hotelarstwie

Rachunkowość w branży hotelarskiej odgrywa kluczową rolę w zarządzaniu finansami oraz operacjami, umożliwiając właścicielom i menedżerom skuteczne monitorowanie i ocenę

wyników finansowych. Standardy rachunkowości stosowane w hotelarstwie, takie jak USALI (*ang. Uniform System of Accounts for the Lodging Industry*), zostały opracowane w celu ujednolicenia sposobu raportowania finansowego i zapewnienia porównywalności wyników operacyjnych pomiędzy różnymi obiektami hotelarskimi (Atkinson & Brander Brown, 2001). Dzięki tym standardom, branża hotelarska może w sposób spójny i przejrzysty prezentować swoje wyniki finansowe, co ułatwia zarówno wewnętrzne zarządzanie, jak i komunikację zewnętrzną z interesariuszami.

USALI jest najczęściej stosowanym systemem rachunkowości w hotelarstwie, który obejmuje szczegółowe wytyczne dotyczące klasyfikacji i prezentacji przychodów, kosztów oraz zysków. System ten umożliwia hotelom dokładne śledzenie wyników finansowych na poziomie operacyjnym, co jest niezbędne do podejmowania świadomych decyzji biznesowych (Bragg, 2011). Dzięki USALI, hotele mogą lepiej zarządzać swoimi zasobami, optymalizować koszty i zwiększać efektywność operacyjną.

W branży hotelarskiej, rachunkowość odgrywa istotną rolę w zarządzaniu finansami i monitorowaniu efektywności działalności (Bednarska, 2014; Enz, 2010; Kucharczyk & Mitura, 2010; Mauri, 2012; Milewska & Włodarczyk, 2018; Mitura & Koniuszewska, 2006; Sadkowski & Sala, 2022; Szmidt, 2017; Tylińska & Wajgner, 2012). Zasady ewidencji danych w systemach finansowo-księgowych są określane przez rachunkowość, a podstawowym aktem normatywnym w Polsce jest ustawa o rachunkowości, która reguluje ogólne zasady ewidencji dla różnych rodzajów działalności. Jednakże, ustawa o rachunkowości nie uwzględnia specyficznych potrzeb branży hotelarskiej, dlatego w międzynarodowej branży hotelarskiej stosuje się standardowe systemy księgowe dedykowane dla hoteli.

Istnieje długa historia rozwoju tych rozwiązań, sięgająca roku 1926 (Gawlik & Puciato, 2016; Sarnecki, 2023b). Uniform System of Accounts for Hotels (USAH) i Uniform System of Accounts for the Lodging Industry (USALI) to dwa różne standardy rachunkowości stosowane w branży hotelarskiej. USAH został początkowo opracowany dla amerykańskiego sektora hotelarskiego, ale jest również szeroko stosowany w hotelach na całym świecie. Z kolei USALI jest globalnym standardem rachunkowości dla branży hotelowej i również znajduje zastosowanie w hotelach na całym świecie, obejmując różne kraje i regiony. USAH został opracowany przez Hospitality Financial and Technology Professionals (HFTP, 2018) i składa się z wytycznych i zasad dotyczących klasyfikacji, prezentacji i analizy danych finansowych w hotelach. USALI, stworzone przez Hotel Association of New York City, stanowi standardowe

wytyczne dotyczące klasyfikacji, prezentacji i analizy danych finansowych w branży hotelowej. Składa się z różnych rozdziałów, które obejmują różne aspekty działalności hotelowej, takie jak przychody, koszty, zyski, pokoje, gastronomia itp. USAH koncentruje się na rachunkowości hotelowej i uwzględnia specyficzne potrzeby oraz charakterystykę operacji w hotelach, takie jak obszar pokoi, gastronomia czy zarządzanie nieruchomościami. Z kolei USALI zostało stworzone z myślą o całej branży noclegowej, obejmującej hotele, motele, pensjonaty i inne obiekty zakwaterowania. Standard ten uwzględnia różnorodne typy obiektów oraz ich specyficzne wymagania finansowe. Oba modele różnią się zakresem zastosowania, organizacją oraz uwzględnianymi specyficznymi wymaganiami branżowymi. Oba standardy mają na celu zapewnienie spójności w zarządzaniu finansami hotelowymi i ułatwienie analizy danych finansowych w branży hotelarskiej na całym świecie. W tradycyjnym podejściu rachunkowości hotelowej często brakuje jednolitości i spójności w sposób, w jaki dane finansowe są gromadzone, klasyfikowane i raportowane. Brak standardu może prowadzić do trudności w porównywaniu wyników finansowych między różnymi hotelami, a także utrudniać analizę i ocenę efektywności operacyjnej.

W przeciwieństwie do tego, model USALI stanowi szczegółowy i jednolity standard rachunkowości dla branży hotelarskiej. Określa on szczegółowe wytyczne dotyczące klasyfikacji kont, metod rachunkowych i raportowania finansowego. System USALI, wraz z ekspansją dużych amerykańskich sieci hotelowych na terenie Europy i innych kontynentów, zdobył ogólną akceptację jako rozwiązanie branżowe, również w Polsce. Wszystkie międzynarodowe sieci hotelowe działają zgodnie z USALI. W Polsce, ze względu na obowiązujące krajowe przepisy legislacyjne dotyczące rachunkowości finansowej i podatkowej, rachunkowość oparta na standardzie USALI jest uważana za integralną część rachunkowości zarządczej.

Podstawowym raportem hotelowym w standardzie USALI jest Sprawozdanie Operacyjne (*ang. Operating Statement*), patrz tabela 1, które odpowiada polskiemu rachunkowi zysków i strat. Warto zauważyć różnicę filozofii w układzie raportów między USALI a polskim RZiS (*Rachunek Zysków i Strat*) według układu rodzajowego. Rachunek zysków i strat grupuje koszty bez analizowania, gdzie dokładnie powstały. Natomiast USALI grupuje zarówno przychody, jak i koszty w zależności od miejsca ich powstania oraz czy są one kontrolowane w danym departamencie. Dzięki temu możliwe jest precyzyjne określanie i wpływanie na

efektywność operacyjną poszczególnych działów oraz hotelu jako całości. Korzystając wyłącznie z RZiS, nie można osiągnąć tego typu efektu.

Tabela 1. Ogólny format sprawozdania operacyjnego w języku polskim i angielskim.

Wyszczególnienie	**Desription**
PRZYCHODY	REVENUE
Pokoje	Rooms
Gastronomia	Food and Beverage
Inne departamenty operacyjne	Other Operated Departments
Wynajem i inne przychody	Rental and Other Income
Razem przychody	Total Revenue
KOSZTY W DEPARTAMENTACH	DEPARTAMENTAL EXPENCES
Pokoje	Rooms
Gastronomia	Food and BeveraGE
Inne departamenty operacyjne	Other Operated Departments
Razem koszty w departamentach	Total Departamental Expences
DOCHÓD OPERACYJNY	TOTAL DEPARTAMENTAL INCOME
KOSZTY NIEPODZIELONE	UNDISTRIBUTED DEPARTAMENTAL EXPENCES
Administracja i dyrekcja	Administrative and General
Sprzedaż i marketing	Sales and Marketing
Koszty utrzymania budynku	Property Operations and Maintenance
Media	Utilities
Razem koszty niepodzielone	Total Undistributed Expences
ZYSK OPERACYJNY BRUTTO	GROSS OPERATING PROFIT
OPŁATY ZA ZARZĄDZANIE	MANAGEMENT FEES
DOCHÓD PRZED KOSZTAMI STAŁYMI	INCOME BEFORE FIXED CHARGES
KOSZTY STAŁE	FIXED CHARGES
Wynajem	Rent
Podatki i opłaty lokalne	Property and Other Taxes
Ubezpieczenie	Insurance
Razem koszty stałe	Total Fixed Charges
DOCHÓD OPERACYJNY NETTO	NET OPERATING INCOME
ODPISY NA FUNDUSZ REMONTOWY	LESS: REPLACEMENT RESERVES

Wyszczególnienie	Desription
SKORYGOWANY DOCHÓD OPERACYJNY NETTO	ADJUSTED NET OPERATING INCOME

Źródło: Opracowanie własne na podstawie: (Qualpro, 2021; Rachuba, 2020; Sarnecki, 2023a).

Standard USALI dostarcza szczegółowych wytycznych dotyczących klasyfikacji, prezentacji i interpretacji danych finansowych w hotelach. Określa on szczegółowe kategorie i podkategorie rachunkowe, które powinny być używane do zgłaszania przychodów, kosztów i innych aspektów finansowych związanych z operacjami hotelowymi. Dzięki zastosowaniu Standardu USALI, hotele mogą prowadzić spójną i porównywalną rachunkowość, co ułatwia analizę finansową, raportowanie i benchmarking zarówno wewnętrzny, jak i zewnętrzny. Umożliwia również dokładne śledzenie wydatków, zarządzanie kosztami i ocenę rentowności działalności hotelu.

Tradycyjny sposób prowadzenia rachunkowości hotelowej różni się od modelu USALI. W tradycyjnym podejściu rachunkowości hotelowej zastosowane są ogólne zasady rachunkowości, które mogą być stosowane w różnych branżach. W tym przypadku rachunkowość hotelowa może być prowadzona przy użyciu ogólnych kont księgowych i procedur rachunkowych. Standard USALI obejmuje różne obszary działalności hotelowej, takie jak przychody z pokoi, przychody z usług gastronomicznych, koszty operacyjne, przychody i koszty nieoperacyjne oraz analizę wydajności operacyjnej. Dzięki konsekwentnemu stosowaniu tych wytycznych, pomaga hotelarzom w podejmowaniu strategicznych decyzji biznesowych i doskonaleniu zarządzania operacjami hotelowymi. Przejście z tradycyjnego sposobu prowadzenia rachunkowości hotelowej na model USALI może wymagać dostosowania i przystosowania systemów księgowych oraz szkolenia personelu. Jednak korzyści z zastosowania standardu USALI obejmują lepszą kontrolę finansową, większą spójność danych, łatwiejszą analizę wyników oraz możliwość porównywania się z innymi hotelami w branży. Model USALI definiuje szczegółowe konta księgowe, które są dedykowane dla branży hotelarskiej, uwzględniając specyfikę operacji hotelowych. Określa również standardowe raporty finansowe, takie jak Raport P&L (*ang. Profit and Loss*), GOPPAR czy TREVPAR, które pozwalają na dokładną ocenę efektywności operacyjnej hotelu.

W standardzie USALI istnieje wyraźny podział na przychody (*ang. Revenue*) i koszty departamentów (*ang. Department Expenses*). Przychody są podzielone na kilka kluczowych grup, takich jak pokoje, gastronomia, inne departamenty operacyjne, wynajem i inne źródła przychodów. Każda z tych grup jest dokładnie wyodrębniona w celu dokładnego śledzenia i raportowania przychodów z poszczególnych obszarów działalności hotelu. Podobnie jak przychody, koszty departamentów są również szczegółowo rozdzielone zgodnie z

odpowiadającymi im grupami przychodów. Na przykład, koszty związane z pokojami, gastronomią i innymi departamentami operacyjnymi są osobno identyfikowane i monitorowane w celu dokładnego zrozumienia struktury kosztów hotelu. Ponadto, istnieje zbiór kosztów niepodzielnych (ogólnych), które obejmują koszty administracyjne, koszty marketingu, koszty ogólnego utrzymania budynku, koszty mediów i inne. Te koszty sąanalizowane jako elementy ogólnych kosztów operacyjnych hotelu.

Podział na departamenty w modelu USALI:

- Departament pokoi: Jest to najważniejszy departament ze względu na specyfikę działalności hotelarskiej. Obejmuje on koszty związane z wyposażeniem pokoi, wynagrodzeniem personelu sprzątającego oraz usługami takimi jak pranie pościeli. Dzięki analizie wyników finansowych w tym departamencie, możemy ocenić rentowność tego obszaru i dostosować ceny pokoi w zależności od sezonu.
- Departament gastronomii: Jest to drugi co do wielkości departament w większości hoteli. Prowadząc szczegółową rachunkowość w tym obszarze, możemy analizować zyski generowane przez serwowane posiłki oraz tworzyć pakiety, takie jak "B&B" (*ang. Bed & Breakfast*). Informacje dotyczące wyników finansowych w dziale gastronomii pozwalają na efektywne zarządzanie personelem (wynagrodzenia dla kelnerów, barmanów i kucharzy), zarządzanie zaopatrzeniem w artykuły spożywcze oraz określenie rentowności usług gastronomicznych.
- Inne departamenty: Oprócz dwóch najważniejszych departamentów, wyróżnia się również inne, takie jak dochody z parkingu czy ewentualnie z kasyna, jeśli obiekt taki posiada. Ponadto, poza departamentami generującymi przychody i koszty, istnieją również centra kosztów. W modelu USALI wyróżnia się departament kosztów niepodzielnych, stałych, odpisy na remonty, koszty finansowe i amortyzacje. Te departamenty odpowiadają za rejestrowanie i monitorowanie różnych rodzajów kosztów, które nie są bezpośrednio związane z przychodami.

Podsumowując, kluczowe wskaźniki finansowe stosowane w ramach USALI, takie jak GOP, RevPAR oraz ADR, stanowią podstawę oceny efektywności operacyjnej hoteli. Wskaźniki te pozwalają menedżerom na dokładną analizę wyników finansowych oraz identyfikację obszarów wymagających poprawy (Schmidgall & DeFranco, 2023). GOP, w szczególności, jest kluczowym wskaźnikiem rentowności operacyjnej, który odzwierciedla zyskowność działalności operacyjnej hotelu po uwzględnieniu kosztów operacyjnych, ale przed odliczeniem kosztów amortyzacji, podatków i odsetek (Jagels & Ralston, 2007).

Stosowanie standardów rachunkowości, takich jak USALI, jest również istotne z perspektywy zgodności z przepisami prawnymi oraz standardami branżowymi. Ujednolicenie rachunkowości w hotelarstwie pozwala na lepszą kontrolę finansową oraz zgodność z wymogami regulacyjnymi, co jest kluczowe dla utrzymania wiarygodności i zaufania interesariuszy. Ponadto, standardy te ułatwiają hotelom dostosowanie się do zmieniających się warunków rynkowych i gospodarczych, umożliwiając elastyczne reagowanie na nowe wyzwania i możliwości. GOP jako kluczowy wskaźnik rentowności operacyjnej, jest szczególnie przydatny w identyfikacji mocnych i słabych stron działalności operacyjnej hotelu. Przyszłe badania powinny koncentrować się na dalszym rozwijaniu narzędzi analitycznych, które mogą wspierać bardziej precyzyjną ocenę efektywności operacyjnej w branży hotelarskiej.

W międzynarodowym kontekście, standardy rachunkowości w hotelarstwie mogą różnić się w zależności od lokalnych przepisów i norm, co może prowadzić do pewnych wyzwań związanych z porównywalnością wyników finansowych między hotelami z różnych krajów (Guilding & Ji, 2022). Dlatego ważne jest, aby hotele działające na rynku międzynarodowym były świadome tych różnic i odpowiednio dostosowywały swoje praktyki rachunkowe, aby zachować spójność i dokładność w raportowaniu finansowym.

Standardy rachunkowości stosowane w hotelarstwie, a w szczególności USALI, stanowią fundament efektywnego zarządzania finansami w tej branży. Dzięki tym standardom, hotele mogą skutecznie monitorować swoje wyniki operacyjne, zarządzać kosztami i podejmować świadome decyzje strategiczne. Stosowanie ujednoliconych standardów rachunkowości jest nie tylko kluczowe dla wewnętrznej organizacji, ale również dla zachowania zgodności z regulacjami i utrzymania zaufania wśród interesariuszy.

2.3. Zysk operacyjny brutto jako kluczowa składowa oceny działalności operacyjnej w branży hotelarskiej

Jak zaznaczono powyżej, zysk operacyjny brutto (*GOP, ang. Gross Operating Profit*) stanowi fundamentalny wskaźnik efektywności działalności operacyjnej przedsiębiorstw hotelarskich. Wskaźnik ten odgrywa kluczową rolę w ocenie zdolności hotelu do generowania zysków z podstawowej działalności operacyjnej, niezależnie od wyników finansowych związanych z inwestycjami czy kosztami finansowymi. GOP odzwierciedla rentowność operacyjną, wskazując na zdolność hotelu do efektywnego zarządzania kosztami operacyjnymi i maksymalizacji przychodów z działalności podstawowej (Gursoy, Buttle, & Bowie, 2022b).

Z definicji zysk operacyjny brutto to różnica między całkowitymi przychodami z działalności operacyjnej a całkowitymi kosztami operacyjnymi, z wyłączeniem kosztów amortyzacji, podatków, odsetek oraz innych kosztów finansowych. GOP jest wskaźnikiem pozwalającym na ocenę efektywności operacyjnej, skupiając się na zarządzaniu codziennymi operacjami w hotelu (Jones & Lockwood, 1999). Zysk ten odgrywa kluczową rolę w określaniu zdolności przedsiębiorstwa do generowania wartości dodanej dla właścicieli i inwestorów, a także w ocenie efektywności zarządzania kosztami (Schmidgall & DeFranco, 2023). W analizie działalności operacyjnej hotelu, GOP pełni również rolę narzędzia diagnostycznego, które umożliwia identyfikację obszarów wymagających poprawy, takich jak koszty operacyjne, struktura przychodów, czy efektywność zarządzania zasobami. Wskaźnik ten pozwala na porównywanie wyników operacyjnych między różnymi hotelami, a także monitorowanie zmian w czasie w kontekście efektywności operacyjnej. Dzięki temu, zarządzający hotelem mogą podejmować świadome decyzje dotyczące optymalizacji kosztów oraz strategii zwiększania przychodów (Kim & Ham, 2006). Należy zaznaczyć, że zysk operacyjny brutto jest również wrażliwy na wpływ czynników zewnętrznych, takich jak zmiany w popycie na usługi hotelarskie, poziom konkurencji na rynku, polityka cenowa, czy też warunki gospodarcze. Z tego względu, analizując GOP, istotne jest uwzględnienie zarówno czynników wewnętrznych, jak i zewnętrznych, które mogą wpłynąć na wynik finansowy hotelu (Olsen & Szabo, 2000).

GOP jest kluczowym wskaźnikiem oceny działalności operacyjnej w branży hotelarskiej, umożliwiającym analizę efektywności zarządzania kosztami oraz zdolności hotelu do generowania zysków z działalności podstawowej. Przyszłe badania w tym obszarze powinny skupić się na dalszej eksploracji wpływu czynników zewnętrznych na GOP oraz na opracowaniu metodologii pozwalających na bardziej precyzyjną ocenę efektywności operacyjnej hoteli (Slattery, 2002). Wartość GOP, w szczególności, jest regularnie analizowana w celu oceny, jak zmiany w polityce zarządzania, takie jak strategie cenowe czy inicjatywy marketingowe, wpływają na zyskowność operacyjną (Olsen & Szabo, 2000). Wskaźniki finansowe USALI, w tym GOP, są również wrażliwe na wpływ czynników zewnętrznych, takich jak warunki ekonomiczne, konkurencja na rynku, czy zmieniające się potrzeby klientów. Analiza tych wskaźników powinna uwzględniać zarówno czynniki wewnętrzne, jak i zewnętrzne, aby dokładnie ocenić, jakie działania należy podjąć, aby poprawić wyniki finansowe hotelu (Slattery, 2002).

3. Metody analityczne w badaniu wyników ekonomicznych sektora hotelarskiego – przegląd badań

3.1. Przegląd wybranych metod analizy i ich wpływ na branże hotelarską

Profesor Peter Drucker, czołowy teoretyk zarządzania XX wieku, zapisał swoją myśl: „Jeśli czegoś nie można zmierzyć, nie można tym zarządzać" (Klimczuk-Kochańska, 2016). To stwierdzenie nadal zachowuje swoją aktualność. Rozwiązania oparte na intuicji, które nie są oparte na wiarygodnych informacjach i analizie danych, są podatne na subiektywną ocenę i mogą prowadzić do decyzji, które nie są obiektywne i racjonalne.

W miarę jak technologia informacyjna coraz bardziej przenika wszystkie aspekty współczesnego społeczeństwa, hotele na całym świecie wydają miliony dolarów na wdrożenie narzędzi handlu elektronicznego (e-commerce) i opracowanie odpowiednich strategii przyciągania i utrzymywania gości (Gallego, González-Rodríguez, & Font, 2023). Niemniej jednak, mimo że literatura odzwierciedla czynniki wpływające na przyjęcie określonych narzędzi e-commerce, nie dostarcza ona jednoznacznej odpowiedzi co do skuteczności wydatków na e-commerce. Konkretnie nie jest jasne czy takie wydatki wpływają na wyniki pod względem kształtowania się wielkości przychodów i zysku brutto, czy te efekty są stałe w czasie i dotyczą hoteli o różnej skali działalności (Pavlatos & Paggios, 2009). Autorzy pracy (Hua, Morosan, & DeFranco, 2015) wykonali badanie na podstawie danych finansowych zgłoszonych przez próbkę 275 hoteli w latach 2007- 2012, które wykazało, że wydatki na e-commerce miały istotny i dodatni wpływ na przychody z pokoi we wszystkich badanych latach próbki, z wyjątkiem 2007 roku. W analizach podgrupowych według skali sieci hoteli stwierdzono, że wydatki na e-commerce miały istotny wpływ na zysk operacyjny brutto w przypadku hoteli klasy średniej i wyższej, ale nie miały tego wpływu w przypadku hoteli luksusowych, bardzo luksusowych oraz hoteli klasy średniej wyższej. Badanie przeprowadzone przez autorów pracy (Assaf, Josiassen, & Cvelbar, 2012) miało na celu zrozumienie związku między raportowaniem zgodnym z podejściem TBL (*ang. Triple Bottom Line*), a wydajnością hoteli. Raportowanie TBL reprezentuje kompleksowe podejście do osiągania zrównoważonego rozwoju, łącząc w sobie aspekty środowiskowe, społeczne i finansowe. Mimo rosnącej presji społeczeństwa w kierunku zwiększonego zaangażowania hoteli w raportowanie TBL, istnieją podziały co do tego, czy ma ono wpływ na wyniki działalności hoteli. W ramach cytowanej pracy badawczej, autorzy zastosowali metodę DEA (*ang. Data Envelopment Analysis*). Wyniki badania, opierające się na próbie słoweńskich hoteli, potwierdziły hipotezy, że bardziej kompleksowe raportowanie dotyczące kwestii środowiskowych, społecznych i finansowych przekłada się na lepsze wyniki działalności hoteli.

Autorzy zauważyli również, że raportowanie w zakresie kwestii środowiskowych wydaje się mieć nieco większy wpływ na wyniki niż raportowanie w zakresie kwestii społecznych i finansowych.

Autorzy pracy (Yadegaridehkordi, Nilashi, Nasir, & Ibrahim, 2018) przeprowadzili badanie mające na celu zidentyfikowanie względnej ważności parametrów CSFs (*ang. Critical Success Factors*), które przyczyniły się do osiągnięcia sukcesu i rozwoju branży hotelowej w Malezji. W badaniu wykorzystali modele TOE (*ang. Technology- Organization-Environment*) oraz HOT (*ang. Human-Organization-Technology*) w celu stworzenia modelu matematycznego, który miał za zadanie predykcję istotnych czynników wpływających na sukces i rozwój sektora hotelowego. Dwuetapowe badanie obejmowało zastosowanie metod SEM (*ang. Structural Equation Modelling*) oraz ANFIS (*ang. Adaptive Neuro-Fuzzy Inference Systems*) w celu przetestowania modelu badawczego i określenia względnej ważności czynników w tym modelu. Wyniki analizy SEM wykazały, że kluczowymi czynnikami wpływającymi na sukces i rozwój hoteli w Malezji były: wykorzystanie technologii informatycznych, wyniki finansowe, benchmarki, standaryzacja usług, wsparcie kierownictwa, satysfakcja klienta, jakość obsługi, projektowanie wnętrza i fasady hotelu, lokalizacja, szkolenia i upoważnienie pracowników. Ponadto, wyniki analizy ANFIS wskazały, że satysfakcja klienta w kontekście ludzkim była najważniejszym predyktorem sukcesu i rozwoju hoteli w Malezji. Te wyniki podkreślają znaczenie zarówno technologicznych aspektów działalności hotelowej, jak i ludzkich czynników, takich jak satysfakcja klienta i jakość obsługi, jako kluczowych elementów sukcesu w branży hotelowej w badanym regionie.

W badaniu przeprowadzonym przez (Tsionas & Assaf, 2021) zaprezentowano elastyczną funkcję przychodów, która jest szczególnie użyteczna dla hoteli działających w konkurencyjnych klastrach lokalizacyjnych. W przypadku tych hoteli zakłada się, że podejmują one swoje decyzje cenowe, kierując się działaniami konkurencji. Badacze wykorzystali podejście dualne, aby oszacować elastyczność popytu oraz zmiany koniunkturalne. Wskazali, że ważność oszacowania tych parametrów nie wymaga bezpośredniego określania funkcji popytu, a także umożliwia obliczenie w relatywnie sposób prosty wskaźnika Lernera (*ang. Lerner Index*, jest to miara makroekonomiczna do oceny stopnia władzy rynkowej lub siły monopolistycznej przedsiębiorstwa (konkurencji) na rynku, wyższa wartość oznacza mniejszą konkurencję, co może wiązać się z wyższymi cenami i niższą efektywnością alokacji zasobów). Zarówno elastyczność popytu, jak i zmiany koniunkturalne, mają istotne implikacje dla zarządzania cenami oraz pomiaru wrażliwości na działania innych hoteli w tych samych klastrach. Badacze skonstruowali elastyczną funkcję przychodów, uwzględniającą specyficzne cechy hoteli działających w konkurencyjnych klastrach i wykazali,

że można było oszacować szereg kluczowych parametrów, w tym elastyczność popytu, zmiany koniunkturalne, efektywność przychodów, zmiany techniczne, zmiany efektywności, wzrost wydajności oraz wskaźnik Lernera, bez konieczności dokładnego określania funkcji popytu. Badanie dotyczyło hoteli działających na terenie Stanów Zjednoczonych za pomocą podejścia bayesowskiego.

W branży hotelarskiej stosowane są różne mierniki operacyjne, czyli narzędzia, które służą do pomiaru, oceny i monitorowania różnych aspektów operacyjnych hotelu. Są to wskaźniki, które pozwalają ocenić wydajność działalności hotelowej, zarówno na poziomie ogólnym, jak i na poziomie poszczególnych działów i procesów. Mierniki operacyjne dostarczają informacji na temat efektywności działania hotelu, zgodności z założonymi celami, wykorzystania zasobów, satysfakcji gości, jakości obsługi, przychodów i kosztów, a także innych kluczowych czynników wpływających na sukces operacyjny hotelu. Dzięki tym miernikom zarządzający hotelami mogą śledzić postępy, identyfikować obszary wymagające poprawy i podejmować odpowiednie działania w celu optymalizacji działalności operacyjnej i osiągnięcia lepszych wyników. Mierniki operacyjne (BnB_Holiday, 2021; HospitalityNet, 2022; Straus, 2020) umożliwiają dokonanie porównań w następujących obszarach:

- Efektywność hotelu i jego poszczególnych działów w stosunku do innych obiektów o podobnych charakterystykach i funkcjach.
- Porównanie wartości planowanych przyjętych w budżecie z ich rzeczywistymi wartościami zrealizowanymi, zarówno w określonym miesiącu, jak i sumarycznie.
- Analiza wartości zrealizowanych w poprzednim roku obrotowym oraz ich porównanie z wartościami osiągniętymi w bieżącym miesiącu i sumarycznie od początku roku.
- Porównanie wartości osiągniętych w danym miesiącu z wartościami sumarycznymi od początku roku obrotowego.

Największe bazy danych dotyczące zdarzeń w hotelarstwie są często gromadzone w systemach finansowo-księgowych (FK) oraz systemach operacyjnych (PMS i POS). Systemy te pełnią istotną rolę w codziennych operacjach hotelowych, umożliwiając zbieranie i przechowywanie różnorodnych danych dotyczących finansów, operacji, rezerwacji, sprzedaży i innych aspektów działalności hotelu. Dzięki tym systemom, które są wykorzystywane na co dzień, hotel może skutecznie zarządzać swoimi operacjami, analizować wyniki finansowe, monitorować wskaźniki wydajności i podejmować odpowiednie decyzje biznesowe oparte na rzetelnych informacjach (Enz, 2010; Kotłowska, 2014; Pomykalska & Pomykalski, 2017; Stabryła, 1998; Tyran, 2004).

3.2. Zastosowanie metod predykcji w branży hotelarskiej

W ostatnich latach firmy hotelarskie zaczęły wykorzystywać analizę predykcyjną, aby lepiej przewidywać i zaspokajać potrzeby i preferencje klientów (McKinsey Digital, 2017). Analiza predykcyjna jest nauką opartą na wykorzystaniu obfitości danych oraz zastosowaniu różnych algorytmów i technik uczenia maszynowego w celu prognozowania najbardziej prawdopodobnych przyszłych wyników. Wiele przedsiębiorstw technologicznych już doskonale przewiduje, jakie produkty konsumenci chcą nabyć, a następnie proponuje im odpowiednie rekomendacje.

Najczęstsze zastosowania uczenia maszynowego w obszarze ekonomii obejmują:

- prognozowanie ciągłych czasowych, takich jak ceny nieruchomości, ceny akcji czy stopy procentowe,
- przewidywanie sukcesu przyszłych sprzedaży detalicznych lub kampanii marketingowych, aby zasoby były wykorzystywane efektywnie,
- przewidywanie trendów klientów lub użytkowników, takich jak na platformach strumieniowych lub stronach internetowych e-commerce,
- analizowanie zbiorów danych w celu ustalenia związków między zmiennymi zależnymi i niezależnymi.

Już w 1992 r. badacze działający w obszarze turystyki, rozwijali techniki prognozowania min. na podstawie szeregów czasowych. Autorzy pracy (Athiyaman & Robertson, 1992) wskazali, że planowanie, zarówno operacyjne, jak i strategiczne, powinno opierać się na dokładnych prognozach, a planowanie w dziedzinie turystyki równie mocno zależy od dokładnych prognoz. Autorzy wskazali, że prognozowanie popytu turystycznego było w ich czasach w dużej mierze zdominowane przez stosowanie ekonometrycznych technik regresji. Wg autorów, wyniki badań dotyczących dokładności prognoz generowanych przez ekonometryczne modele regresyjne sugerują, że prognozy generowane przez te modele niekoniecznie są lepsze od prognoz generowanych przez proste techniki analizy szeregów czasowych. Badacze zaprezentowali siedem technik analizy szeregów czasowych wykorzystanych do generowania prognoz międzynarodowych rejestracji przyjazdów (tzw. meldunków) turystów z Tajlandii do Hongkongu. Ich wyniki potwierdziły, że relatywnie proste

techniki mogą być równie dokładne, a często bardziej oszczędne pod względem czasu i kosztów niż bardziej złożone metody.

Zagadnieniem związanym ze znaczeniem dokładnych prognoz dla wyników finansowych w określonych miejscach w branży hotelarskiej zajmowali się autorzy pracy (Pan & Yang, 2016), w której uwzględniono pilną potrzebę opracowywania precyzyjnych prognoz dotyczących przyszłej wydajności hotelu, aby porównać swoje obiekty i zoptymalizować działania operacyjn m.in. z powodu rosnącej konkurencji w branży. W badaniach wykorzystano modele szeregów czasowych do prognozowania tygodniowej zajętości hotelu w określonym miejscu, które uwzględniały różne źródła dużych zbiorów danych (*ang. big data*) związane z turystyką, w tym zapytania w wyszukiwarkach internetowych, ruch na stronach internetowych i cotygodniowe informacje pogodowe, które to źródła mogą dostarczyć cennych informacji na temat wzorców popytu. W szczególności były to modele ARMAX, które osiągały lepsze wyniki pod względem dokładności niż inne modele. Autorzy w badaniach uwzględnili metody wyznaczania sezonowości w zajętości hotelu, oraz zasugerowali, że tygodniowe zmienne typu dummy (zmienne reprezentujące dowolny dzień tygodnia) są skuteczniejsze niż składowe Fouriera, wykorzystane do modelowania sezonowości w hotelu. Autorzy zaznaczają, że ich rezultaty posiadają ograniczenia związane z wykorzystaniem wielu źródeł typu big data. Pomimo uwzględnienia różnych strumieni danych, redukcja błędu prognozowania była minimalna, co wskazuje na potencjalne trudności w istotnym poprawianiu dokładności prognozy poprzez integrację danych.

Znacznie późniejsze badania zaprezentowane w pracy (Lee, 2018) pokazują wyniki badania, w ramach którego rozwinięto nowatorskie podejście do krótkoterminowego prognozowania liczby rezerwacji hotelowych. W badaniu skupiono się na kluczowych cechach związanych z liczbą dokonanych rezerwacji, takich jak zmienne w czasie wskaźniki meldunków, duża zmienność w ostatecznym zapotrzebowaniu oraz wyraźnie dodatnie korelacje zależności między meldunkami dokonanymi w różnych okresach. W ramach badania przeanalizowano trzy różne modele typu Poissona, które pozwalają uwzględnić istotne cechy związane z liczbą dokonanych rezerwacji. Autorzy zaznaczają, że występowanie silnych korelacji między meldunkami dokonanymi w różnych okresach, pozwoliło na wykorzystanie ich do prognozowania przyszłych rezerwacji już na podstawie wczesnych obserwacji. W rezultacie autorzy zaproponowali nową metodę prognozowania, która bazuje na wewnętrznych związkach między rezerwacjami dokonywanymi wcześniej, a tymi dokonywanymi później. Następnie przeprowadzono walidację wyników na podstawie danych pozyskanych od przedstawiciela dużej sieci hotelowej, porównując je z wynikami uzyskanymi przez modele odniesienia.

Wyniki badania potwierdziły, że zastosowanie metody dynamicznego uaktualniania, opierającej się na międzyczasowych korelacjach, znacząco poprawiło dokładność krótkoterminowych prognoz dotyczących popytu na pokoje hotelowe.

W innym badaniu z tego okresu (Jun, Yuyan, Lingyu, & Peng, 2018), autorzy opracowali kompleksowy model prognozowania popytu, mającego kluczowe znaczenie dla planowania w sektorze turystyki. Autorzy wskazują, że zarówno administracje rządowe, jak i przedsiębiorstwa prywatne stosowały podstawowe prognozy popytu turystycznego, aby opracować odpowiednie polityki publiczne i podejmować trafne decyzje inwestycyjne w infrastrukturę. Stąd prognozowanie stało się nieodzownym elementem zarządzania turystyką. W badaniach przeprowadzonych przez autorów, zaproponowano zastosowanie połączonego modelu prognozowania popytu turystycznego, wykorzystującego sztuczną sieć neuronową oraz algorytm grupowania. Ich model uwzględniał dwa aspekty analizowanej serii danych: wzorce sekwencji i charakterystyki, które odzwierciedlały zmiany strukturalne i wzajemne korelacje szeregów czasowych. Dane treningowe zostały podzielone na jednorodne grupy, a dla każdej grupy wykorzystano dedykowany predyktor. Następnie wybrano kilka sąsiadujących ze sobą próbek, aby uchwycić bieżące zmiany w trendach danej serii. Na koniec wyniki dwóch prognoz opartych na wzorcach sekwencji i bliskich charakterystykach zostały połączone w celu określenia końcowych prognoz. Aby zweryfikować dokładność i efektywność zaproponowanego modelu, przeprowadzono porównanie z trzema innymi modelami opartymi na sieciach neuronowych oraz popularnym modelem ARIMA, używając trzech szeregów czasowych dotyczących rejestracji gości hotelowych, które cechowały się nieliniowym charakterem oraz brakiem stacjonarności. Na podstawie uzyskanych wyników badawczych wykazano, że zaproponowana przez autorów metoda łączenia konsekwentnie przewyższała referencyjne metody prognozowania. Natomiast badacze (Huang & Zheng, 2021) motywowali swoją pracę, dotyczącą doskonalenia technik prognostycznych, rosnącą konkurencją w branży hotelarskiej oraz związaną z nią koniecznością przyjęcia aktualnych praktyk zarządzania przychodami, celem maksymalizacji zysków i optymalizacji bieżącej działalności. Biorąc pod uwagę ograniczenia związane z ich dotychczasowymi badaniami, skoncentrowali się na codziennym popycie na hotele z uwzględnieniem efektu aglomeracji i zaproponowali nowatorski model oparty na uczeniu głębokim DLM-ST (*ang. Deep Learning Model with Spatial and Temporal correlations*), stanowiący rozwinięcie modelu LSTM (*ang. Long Short-Term Memory*), który wprowadził nowy element w badaniach, uwzględniając efekt aglomeracji oraz integrując mechanizm uwagi i algorytm optymalizacji bayesowskiej. Do weryfikacji wydajności modelu badacze wykorzystali dane historyczne dotyczące dziennej liczby rezerwacji (popyt dzienny) pozyskane z 210 hotelów w Xiamen, Chiny. Wyniki ich badań

wykazały, że zaproponowany model znacząco przewyższał modele referencyjne, a badanie może pomóc menedżerom hoteli udoskonalić efektowność zarządzania przychodami poprzez lepsze dopasowanie potencjalnego popytu do dostępnych zasobów. Badacze (Fiori & Foroni, 2020) prowadząc badania zauważyli, że z wyjątkiem kilku znaczących przypadków, badania dotyczące prognozowania w branżach lotniczej i hotelarskiej skupiają się głównie na tzw. punktowych prognozach rezerwacji gości (*ang. point prediction*), które jednak są ryzykowne w kontekście zarządzania przychodami, gdyż opierają się wyłącznie na prognozach punktowych. Traktując to jako wadę, badacze zaproponowali metodę stochastyczną, która pozwala na budowanie przedziałów prognozowych dla metod prognozowania rezerwacji (*ang. pickup*). Ponadto wprowadzili rozszerzenie techniki „multiplicative pickup", opartej na ogólnych modelach liniowych. Zaproponowana metoda została przetestowana na rzeczywistej liczbie rezerwacji pozyskanej z małego hotelu nad jeziorem Maggiore we Włoszech, dzięki czemu opracowano bardziej efektywne przedziały prognozowe w porównaniu do klasycznych metod analizy szeregów czasowych. Autorzy sugerują, że takie podejście może być przydatne dla menedżerów, którzy chcą podejmować bardziej świadome decyzje, planując alternatywne strategie cenowe i alokacji pokoi dla różnych scenariuszy popytu. Bardziej współczesne badanie dotyczące prognozowania popytu hotelowego przedstawiono w pracy (Viverit, Heo, Pereira, & Tiana, 2023). Opisuje ono autorski sposób prognozy, który sprawdził się w trudną i nieprzewidywalną sytuacją spowodowaną pandemią COVID-19. Autorzy zaproponowali nowe podejście, oparte na grupowaniu dat pobytów, na podstawie rezerwacji historycznych. Ich podejście różni się fundamentalnie od tradycyjnych metod prognozowania w hotelarstwie, które zakładają podobieństwo krzywych rezerwacji i wzorców w okresie historycznym. W badaniu autorzy wykorzystali algorytmy uczenia maszynowego do grupowania historycznych krzywych rezerwacji w sposób autoregresyjny, a następnie zastosowali model addytywnego przyrostu, aby prognozować dzienne obłożenie hotelu w horyzoncie czasowym do 8 tygodni. Skuteczność zaproponowanego przez nich podejścia została zweryfikowana na rzeczywistych danych rezerwacji, zebranych w trzech hotelach, a przedstawione wyniki potwierdzają, że prognozy popytu są dokładniejsze, gdy są generowane na poziomie grup, niezależnie od horyzontu prognozowania.

Anulacje rezerwacji stanowią istotny element zarządzania przychodami w hotelarstwie ze względu na ich wpływ na systemy rezerwacji pokoi. W rzeczywistości niewiele wiadomo o przyczynach, które skłaniają potencjalnych gości do anulacji rezerwacji lub o sposobach ich unikania. Celem autorów artykułu (Sánchez-Medina & C-Sánchez, 2020) było zaproponowanie metody umożliwiającej prognozowanie anulacji rezerwacji hotelowych, wykorzystując

zmniejszoną liczbę cech, w porównaniu z badaniami referencyjnymi, które autorzy uznali za najistotniejsze dla klientów dokonujących rezerwacji. W pracy zastosowano techniki uczenia maszynowego, w tym sztuczne sieci neuronowe zoptymalizowane za pomocą algorytmów genetycznych. Zaproponowana przez autorów metodologia pozwala poznać wskaźniki przyczyniające się do anulacji, a także zidentyfikować, który klient jest bardziej skłonny do rezygnacji z wizyty. Anulacjami zajmowali się również autorzy artykułu (Chen i in., 2023), którzy wskazali, że dane osobowe związane z rezerwacjami PNR (*ang. Personal Name Records*) stanowią kluczowe źródło informacji do zadań predykcyjnych w tej dziedzinie, choć często bywają niedostępne. W badaniach przeprowadzonych przez autorów stosowano bezpośrednie wprowadzenie danych PNR do zaawansowanych modeli uczenia maszynowego, aby osiągnąć efektywną predykcje liczby anulowanych rezerwacji. Jednakże, autorzy wskazali również na inną metodę polegającą na generowaniu predykcji anulowania na podstawie modelowania prawdopodobieństwa próbek. Zaproponowali metodę interpretowalnej interakcji cech w celu wzbogacenia dostępnych informacji związanych z danymi PNR. Następnie przeprowadzili badania empiryczne w celu oceny wydajności predykcyjnej obu rodzajów modeli. Badanie szczegółowo określiło, czy te dwa podejścia mogą wzajemnie się uzupełniać i poprawić predykcję anulowania rezerwacji. Autorzy zaproponowali więc model predykcyjny, który integruje sieci Bayesowskie oraz regresję Lasso. Sieci Bayesowskie zostały użyte do modelowania prawdopodobieństwa zmiennych. Zastosowano też liniowy model uczenia maszynowego, tj. regresję Lasso, ze względu na jej zdolność do redukcji nieefektywnych predyktorów oraz przejrzystość w rankingu istotności cech. Wyniki empiryczne wykazały, że zaproponowany model osiągnął lepszą wydajność predykcyjną.

Prognozowanie stanowi wyjściowy składnik cyklu zarządzania przychodami RM (*ang. Revenue Management*) w branży hotelarskiej. Dokładność prognozy ma kluczowe znaczenie dla systemów RM, aby mogły dostarczać odpowiednich rekomendacji mających na celu optymalizację przychodów. W ostatnich latach branża wskazywała na zmieniające się okna rezerwacji, spowodowane różnymi czynnikami w skali makro (np. technologicznymi i ekonomicznymi) oraz mikro (np. promocjami). Te zmiany stanowią wyzwanie dla algorytmów prognozowania w RM, zwłaszcza w obszarze technik opartych na analizie pickup. W pracy (Webb, Schwartz, Xiang, & Singal, 2020) autorzy dokonują przeglądu literatury dotyczącej prognozowania w RM w dziedzinie hotelarstwa, ze szczególnym uwzględnieniem popularnych technik używanych w praktyce. Następnie wprowadzają podejście oparte na sieciach neuronowych do środowiska zaawansowanych rezerwacji, mając na celu rozwiązanie problemów związanych ze zmianami okien rezerwacyjnych. Modele są estymowane i testowane pod kątem dokładności, a następnie ponownie testowane lata

później, gdy okna rezerwacyjne ulegają zmianie. W pracy poddano dyskusji, które modele są najbardziej odpowiednie do prognozowania w dynamicznych oknach rezerwacyjnych.

Literatura przedmiotu zawiera również badania związane ze współczynnikiem zajętości OR (*ang. occupancy rate*). Przykładowo w pracy (Pereira, 2016) autorzy rozpatrują techniki prognozowania dziennego współczynnika OR. Przeprowadzili badanie, w którym wykorzystali zarządzanie przychodami jako kluczowe narzędzie w procesie podejmowania decyzji przez menedżerów. Opracowali zaawansowane systemy zarządzania przychodami, które miały wspierać decyzje menedżerów, a niezbędnym elementem tych systemów był dokładny moduł prognozowania. Ich modele zostały opracowane w ramach struktury modelowania w przestrzeni stanu, która umożliwia radzenie sobie z problematyką sezonowości. Autorzy przeprowadzili badanie empiryczne, aby zilustrować, jak praktycy mogą zastosować i porównać wydajność różnych modeli podczas prognozowania codziennej zajętości hotelu. Wyniki wykazały, że model trygonometryczny oparty na nowej strukturze modelowania zazwyczaj przewyższał większość modeli referencyjnych. Natomiast, autorzy artykułu (Assaf & Tsionas, 2019), opisują wyzwanie zwane „klątwą wymiarowości", z jakim często spotykają się badacze pracując nad rozległymi modelami typu VAR (*ang. Vector Autoregressions*). W cytowanym artykule zaprezentowano nową metodę opartą na koncepcji regresji skompresowanej, w której zastosowano dwie nowatorskie, nieliniowe metody VAR na potrzeby prognozowania współczynnika OR dla hoteli konkurujących w wąskim obszarze geograficznym. Ich modele stały się bardziej elastyczne dzięki wykorzystaniu sieci neuronowych, natomiast skuteczność porównano z kilkoma modelami referencyjnymi, dowodząc wyższość opracowanej metody w kontekście dokładności predykcji pozostającej na stałym poziomie na niemal wszystkich horyzontach prognoz od 1 do 36 miesięcy. W artykule przeglądowym (Dowlut & Gobin-Rahimbux, 2023) autorzy wskazują, że prognoza współczynnika OR jest kluczowym elementem strategii zarządzania przychodami w branży hotelarskiej. Autorzy stwierdzają, że przewidywanie zajętości hoteli jest istotne w procesie podejmowania decyzji przez menedżerów, ponieważ dostarcza szacunków dla przyszłej wydajności biznesowej. Jednak w szybko zmieniającym się otoczeniu marketingowym w sektorze turystycznym, wspieranym przez rozwój rezerwacji online, generowanie dokładnych prognoz stało się zadaniem trudnym, wymagającym zastosowania zaawansowanych umiejętności technicznych i kosztownego oprogramowania. W zaprezentowanym przez autorów przeglądzie literatury przedstawiono informacje na temat wykorzystania technik uczenia maszynowego w przewidywaniu współczynnika OR.

Wyróżniono najnowsze trendy w tej dziedzinie z okresu pięciu lat, tj. od 2017 do 2022 roku. W podsumowaniu autorzy wskazują na algorytm LSTM jako najpopularniejszy algorytm uczenia maszynowego używany do budowy modeli prognostycznych oraz metrykę MAPE jako najpopularniejszą z siedmiu zwykle stosowanych miar oceny jakości predykcji. Stwierdzono także, że hybrydowy model CNN-LSTM (*ang. Convolutional Neural Network LSTM*) może zwiększyć dokładność, ale wymaga dalszych badań.

Relatywnie najczęściej badanym zagadnieniem, opisywanym szeroko w literaturze fachowej w branży hotelarskiej, jest analiza treści internetowych, opinii, recenzji i komentarzy, szczególnie w kontekście zarówno segmentacji gości jak i badania indeksu satysfakcji gości z usług. Ilość dostępnych tekstów w sieci stale rośnie, co stanowi wyzwanie dla organizacji, które próbują efektywnie wykorzystać te nieustrukturyzowane dane. Pomimo licznych technik opracowanych przez społeczność informatyczną, wciąż istnieje duże pole do poprawy w sposobie, w jaki organizacje wykorzystują takie dane tekstowe, zwłaszcza w kontekście wsparcia procesów decyzyjnych. W pracy (Aguwa, Olya, & Monplaisir, 2017) autorzy podkreślają, że identyfikacja, interpretacja i reagowanie na potrzeby klientów stanowią kluczowe czynniki sukcesu, a niezadowolenie gości może zaszkodzić reputacji firmy i przynieść poważne straty. W cytowanym badaniu autorzy opracowali nowe podejście do właściwego rozumienia i analizy niejednoznacznych opinii klientów, wykorzystując metody eksploracji tekstu i algorytmy uczenia bazujące na regułach asocjacyjnych. Zaproponowana metoda konwertuje dane tekstowe i jakościowe na wspólny format ilościowy, który jest następnie wykorzystywany do opracowania zintegrowanego indeksu satysfakcji klientów ICSI (*ang. Integrated Customer Satisfaction Index*). Wg autorów inne wskaźniki satysfakcji klientów nie uwzględniają kosztów rozwiązywania reklamacji i problemów klientów oraz niejasnego wpływu tych reklamacji i problemów na cały system. Dlatego, aby uwzględnić te ważne i unikalne czynniki, autorzy wprowadzili nową miarę monitorowania CTQ (*ang. Critical to Quality*), pozwalającą na ocenę w czasie rzeczywistym poprzez aktualizację odpowiednich bibliotek. Wreszcie, zaproponowano procedurę eksploracji opinii klientów i analizy sentymentu, która radzi sobie z błędami typograficznymi, obecnymi w każdej rzeczywistej bazie danych. Wyniki badania sugerują, że uwzględnienie poziomu niejasności i pozytywności komentarzy w modelu, zamiast traktowania ich jako zmiennych binarnych, prowadzi do bardziej satysfakcjonujących wyników. Natomiast, autorzy artykułu (Barnes, Mattsson, Sørensen, & Jensen, 2020) wykorzystują podejście z zakresu analizy tekstu pochodzącego z recenzji internetowych, do pomiaru interakcji między pracownikami hotelu, a gośćmi. Badacze wprowadzają koncepcyjny model mierzenia tych interakcji, wykorzystując podejście

oparte o analizy dużych zbiorów danych (*ang. big data analytics*), dzięki któremu uzyskano lepsze rezultaty niż w oparciu o tradycyjne, ograniczone na badaniach ankietowych dane. Korzystając z podejścia opartego na słownikach i dużej liczbie recenzji hoteli (n = 265 016), autorzy przetestowali swój model oraz dodatkowo określili znaczenie czynników wpływających na postrzeganie takich miar jak satysfakcja, obsługa i wartość. W pracy dowiedziono istotności różnych wymiarów wartości doświadczalnej w interakcjach między pracownikami a turystami w tworzeniu pozytywnego postrzegania przez turystów. W pracy zastosowano unikalną perspektywę analizy tekstu w odniesieniu do pięciu kluczowych typów interakcji pracowników (personalizacja, elastyczność, współtworzenie, emocje oraz wiedza i nauka). W artykule (Marcolin, Becker, Wild, Behr, & Schiavi, 2021) autorzy zaproponowali i zweryfikowali model oparty o kombinację trzech technik eksploracji tekstu: klasyfikację tekstu, analizę sentymentu i modelowanie tematów w kontekście efektywnego wykorzystania danych tekstowych w branży hotelarskiej. Opracowana metoda pozwala menedżerom analizować komentarze gości oraz porównywać konkurentów w branży hotelarskiej na podstawie SERVQUAL (*ang. Service Quality*), czyli pięciu głównych miarach oceny jakości usług. Celem autorów jest przedstawienie zautomatyzowanego procesu, który obejmuje zbieranie i analizę danych tekstowych, co przyczynia się do usprawnienia procesu podejmowania decyzji. Publikacja (Calero-Sanz, Orea-Giner, Villacé-Molinero, Muñoz-Mazón, & Fuentes-Moraleda, 2022) dotyczy narzędzi przemysłu 4.0. Umożliwiają one zautomatyzowanie procesów produkcyjnych, a podejścia oparte na sztucznej inteligencji odgrywają kluczową rolę w badaniach nad branżą turystyczną. Zastosowanie tych narzędzi do analizy treści generowanych przez użytkowników UGC (*ang. User-Generated Content*) jest istotne dla lepszego zrozumienia potrzeb, opinii i oczekiwań klientów dotyczących usług turystycznych. W ramach badania autorzy przeprowadzili analizę eksploracyjną za pomocą modeli uczenia maszynowego, aby lepiej zrozumieć rolę, jaką odgrywają rodzaje botów internetowych i typy podróżujących w ocenie hoteli na podstawie recenzji na platformie TripAdvisor dla 74 hoteli. Celem ich badania eksploracyjnego było opracowanie metodologii, skoncentrowanej na analizie recenzji internetowych, dotyczących botów internetowych obsługujących hotele, przy użyciu technik uczenia maszynowego do przetwarzania danych zebranych z TripAdvisor. Wstępne wyniki wskazują na związek między oceną hoteli na platformie TripAdvisor, a rodzajem bota internetowego wprowadzonego w hotelu. Badanie mające na celu zrozumienie, jak opinie klientów wpływają na wyniki hotelu przeprowadzili autorzy artykułu (Assaf, Josiassen, Knežević Cvelbar, & Woo, 2015). W opracowanym przez nich modelu uwzględniają zarówno poziom zadowolenia klientów, jak i liczebność skarg.

Próbowali zbadać, jak te opinie wpływają na wyniki hoteli, biorąc pod uwagę różnice w wielkości i klasie hotelu. Do pomiaru wyników hoteli zastosowali bardziej zaawansowaną metodologię niż to powszechnie stosowane w badaniach nad związkami między zadowoleniem klienta a wynikami finansowymi. Na koniec, przedstawili wyniki swoich badań i omówili, jakie implikacje mogą wynikać zarówno dla teorii, jak i praktyki zarządzania hotelami. Istotną cechą tego badania jest zastosowanie bardziej zaawansowanego podejścia do pomiaru wyników firm niż w dotychczasowych badaniach nad związkiem między zadowoleniem klienta, a wynikami działalności. Zamiast używać powszechnie stosowanych wskaźników finansowych, takich jak ROA (*ang. Return on Assets*, jest to wskaźnik rentowności, który mierzy efektywność wykorzystania aktywów przedsiębiorstwa w celu generowania zysków) czy Tobin's *q* (jest to wskaźnik oceny wartości rynkowej firmy w stosunku do wartości rynkowej jej aktywów), autorzy skupili się na luce w efektywności technicznej. To podejście ma dwie istotne zalety. Po pierwsze, pomiar efektywności technicznej uwzględnia ogólną wydajność firmy na podstawie wielu wejść i wyjść, a nie tylko pojedynczych wskaźników. Dzięki temu dostarcza kompleksowej i realistycznej oceny wyników firmy, pozwalając na pełną analizę jej osiągnięć na różnych płaszczyznach. Po drugie, pomiar efektywności technicznej pozwala określić różnicę w efektywności firmy w porównaniu do jej najlepiej działających konkurentów. To z kolei umożliwia pełne zrozumienie wyników firmy, mierząc jej wydajność w kontekście maksymalnego potencjału, jaki może osiągnąć. W rezultacie to podejście dostarcza pełnej oceny wyników każdej firmy i pozwala zrozumieć, jak sprawnie firma działa w konkurencyjnym otoczeniu rynkowym. Jest to istotna różnica w porównaniu do tradycyjnych wskaźników finansowych, a także dodaje głębi analizie wpływu opinii klientów na wyniki firmy. Również zarządzanie restauracją wymaga od przedsiębiorców reagowania na oczekiwania klientów, które stają się coraz wyższe w zależności od wielkości konkurencji na lokalnym rynku. Badanie przedstawione w pracy (Fernandes, Moro, Cortez, Batista, & Ribeiro, 2021) proponuje podejście mające na celu uproszczenie procesu podejmowania decyzji przez menedżerów restauracji poprzez połączenie opinii klientów na żywo w mediach społecznościowych i historycznych danych finansowych w modelu prognozowania sprzedaży, opartym na danych z serwisu TripAdvisor oraz modelu Bassa, znanym również jako *Model Diffusion of Innovation Bass*. Ich podejście zostało zweryfikowane za pomocą danych wewnętrznych i zewnętrznych (tj. recenzji internetowych) zebranych z sześciu restauracji. Zgromadzone dane zostały przetworzone w celu stworzenia panelu narzędziowego, który dostarcza nową wartość dla restauratorów, wykorzystując recenzje online i prognozowanie sprzedaży.

Innym obszarem prowadzonych prac badawczych, szeroko opisywanym obecnie w literaturze tematu jest eksplozja cyfrowych danych tekstowych o dużym rozmiarze TBD (*ang. textual big data*), które stworzyła możliwości dla badań naukowych również w dziedzinie konsumenckiej, w branży hotelarskiej i turystycznej. Badanie opisane w (Kwon, 2023) proponuje standardowy ramowy model badań nad TBD oraz wprowadza praktyczne metody wykorzystywane w językach R i Python, włączając w to zalety i wady różnych podejść metodologicznych. Ponadto, cytowane badanie zawiera kompleksowy przegląd literatury dotyczącej hotelarstwa i turystyki w kontekście TBD, koncentrując się na analizie sentymentu i modelowaniu tematycznym. Wyniki zaprezentowanych prac pokazują, że analiza TBD w branży hotelarskiej i turystycznej nadal jest w fazie początkowej, z niewielką uwagą poświęconą rygorystyce metodologicznej i brakiem szczegółowych opisów. Autorzy wykazali na podstawie zawartego w pracy przeglądu literaturowego, że chociaż techniki przetwarzania języka naturalnego NLP (*ang. Natural Language Processing*) rozwijają się dynamicznie z dnia na dzień, to niewiele przedsięwzięto wysiłków metodologicznych, aby rozwiązać problemy badawcze w naukach społecznych, a już na pewno w hotelarstwie i turystyce. Dlatego w pracy zaproponowali ogólny ramowy model badawczy oraz plan działania skoncentrowany na zaawansowanych technikach NLP i praktycznych metodach. Z kolei badacze w pracy (Zhang & Wu, 2023) przeprowadzili analizy dokładności prognozowania popytu na hotele, gdzie istotne źródło informacji stanowiły recenzje. Autorzy wskazują, że obecne metody prognozowania popytu na hotele często nie zapewniają wystarczającej wyjaśnialności lub nie są w stanie uchwycić lokalnych korelacji w sekwencjach danych. Dlatego w ramach prac podjęli dwie tematyki: 1) opracowanie kompleksowego algorytmu, który składa się z czterech głównych komponentów: ekspertyzy, sentymentu, popularności i nowości (*ang. framework ESPN*), mającego na celu zbadanie wpływu recenzji online na prognozowanie popytu na hotele;
2) zaproponowanie nowatorskiego modelu DA-LSTM-CNN (*ang. Dual Attention-Based Long Short-Term Memory Convolutional Neural Network*), który miał na celu zoptymalizowanie skuteczności modelu prognozowania. Badacze zebrali dane z recenzji online ze strony Ctrip.com w celu oceny proponowanego frameworku ESPN i modelu DA-LSTM-CNN. Wyniki badań empirycznych wykazały, że uwzględnienie cech pochodzących z frameworku ESPN przyczyniło się do poprawy dokładności prognozowania, a model DA-LSTM-CNN znacząco przewyższył modele uznane za najlepsze w dziedzinie. Ponadto, badacze wykorzystali studium przypadku, aby zilustrować wyjaśnialność modelu DA-LSTM-CNN, co może być przydatne w opracowywaniu przyszłych systemów prognozowania popytu na hotele.

Analizą sentymentu na podstawie recenzji internetowych, zajmowali się badacze pracy (Jain, Pamula, & Srivastava, 2021), którzy przeprowadzili badanie mające na celu określenie użyteczności, zakresu i zastosowania połączenia technik uczenia maszynowego w analizie nastrojów konsumentów CSA (*ang. Consumer Sentiment Analysis*). W ramach pracy przeprowadzono systematyczną analizę literatury w celu porównania, analizy, eksploracji i zrozumienia potrzeb oraz kierunków badawczych. Głównym celem badania było zrozumienie i analiza zastosowania technik uczenia maszynowego do analizy nastrojów konsumentów w recenzjach internetowych w dziedzinie hotelarstwa i turystyki. Natomiast, w pracy przeglądowej (Liu, Hu, Mehraliyev, & Liu, 2023) przedstawiono ocenę ówczesnej sytuacji w zaawansowaniu prac badawczych nad zastosowaniem metod uczenia głębokiego w klasyfikacji tekstu w dziedzinie turystyki i hotelarstwa oraz zaproponowano konkretne wytyczne dla przyszłych badań. Wyniki analizy literaturowej wskazują, że dotychczasowe badania głównie koncentrowały się na klasyfikacji cech tekstu, klasyfikacji ocen tekstu oraz klasyfikacji sentymentu tekstu. W większości przypadków wykorzystywane metody uczenia to znane metody, pochodzące jeszcze z XX wieku, w tym m.in. sieci neuronowe jednokierunkowe i sztuczne sieci neuronowe. Nowoczesne algorytmy uczenia głębokiego, które pojawiły się w ostatnich latach w dziedzinie nauk komputerowych, a charakteryzują się lepszą wydajnością klasyfikacyjną, nie zostały jeszcze wprowadzone na szeroką skalę w dziedzinie turystyki i hotelarstwa. Ponadto, większość danych używanych w badaniach pochodziła z ogólnodostępnych zbiorów danych ocen; tylko dwie z badanych prac ręcznie zanotowały dane pozyskane z witryn internetowych związanych z turystyką.

Według autorów artykułu (Casidy, Wymer, & O'Cass, 2018), hotelarze coraz bardziej skupiają się na budowaniu relacji z gośćmi, co prowadzi do znacznego skoncentrowania się na relacjach między konsumentem, a marką w badaniach dotyczących turystyki i hotelarstwa. Cytowane badanie analizuje wpływ PBRO (*ang. Perceived Brand Relationship Orientation*, który odnosi się do sposobu, w jaki konsumenci postrzegają swoje związki z daną marką, co może obejmować emocje, lojalność, zaufanie i inne czynniki wpływające na sposób, w jaki konsumenci identyfikują i integrują się z marką, na ich identyfikację z markami oraz oczekiwane emocje związane z markami hotelowymi, co z kolei wpływa na pożądane wyniki wydajnościowe hoteli, takie jak udział w portfelu klienta, rozmiar zbioru rozważanych opcji oraz intencja ponownego odwiedzenia. Badanie przeprowadzono, rekrutując 376 respondentów za pośrednictwem Amazon Mechanical Turk (Mturk). Wyniki wskazują, że identyfikacja konsumentów z marką oraz przewidywane emocje odgrywają rolę mediatora w związku między postrzeganą orientacją w relacji z marką, a wspomnianymi wskaźnikami wyników.

Ponadto poziom zaangażowania konsumentów w proces wyboru hotelu wpływa na te efekty: pozytywnie wpływając na związek między PBRO, a identyfikacją konsumenta z marką oraz negatywnie wpływając na efekty PBRO na przewidywane emocje.

Literatura przedmiotu obejmuje także prace dotyczące segmentacji i grupowania klientów, które odnoszą się do podziału klientów na wiele różnych podgrup według ich specyficznych cech. Dokładna segmentacja klientów jest kluczowa dla decydentów, aby w pełni zrozumieć wymagania rynkowe, a następnie tak projektować działania marketingowe, aby zaspokoić ich potrzeby. Dokładny podział klientów stanowi wyzwanie, gdy mamy do czynienia z rzeczywistymi danymi dotyczącymi wymagań klientów CRD (*ang. Customer Requirement Data*). Autorzy artykułu (Sun i in., 2021) zajmowali się opracowaniem heurystycznej metody grupowania GPHC (*ang. Gaussian Peak Heuristic Clustering*). W szczególności, w artykule tym wykorzystano metodę entropii i standaryzowany rozkład Gaussa do filtrowania i modelowania interwałowych wartości CRD. Następnie ukryty wzorzec preferencji klienta w CRD rozpoznawany był za pomocą algorytmu genetycznego i grupowania hierarchicznego. Finalny wynik grupowania CRD został uzyskany przy wykorzystaniu algorytmu k-średnich (*ang. K-means*), opartego na informacjach heurystycznych z ukrytych wzorców preferencji klientów. Eksperymenty badawcze pokazały, że wynik segmentacji klientów wygenerowany przez ich metodę był zgodny z wynikami segmentacji klientów dostarczonymi przez ekspertów. Ponadto, odporność GPHC w obliczu skomplikowanych scenariuszy segmentacji klientów została zweryfikowana za pomocą eksperymentów numerycznych. W badaniu opisanym w (Li, Chu, Tian, Feng, & Mu, 2021) autorzy zaproponowali metodę segmentacji klientów opartą na ulepszonym algorytmie k-średnich oraz algorytmie optymalizacji rojowej cząsteczek PSO (*ang. Particle Swarm Optimization*) dostosowanym do warunków (*ang. adaptive PSO*). Klasyczny algorytm PSO może zakończyć operację identyfikując lokalne ekstremum; dlatego autorzy zaproponowali algorytm adaptacyjnego uczenia ALPSO (ang. *adaptive learning PSO)* w celu poprawy dokładności optymalizacji. Na podstawie analizy optymalizacji opartej na populacji, przeprojektowano wagę bezwładności, współczynniki uczenia i metodę aktualizacji pozycji. Aby zapobiec zależności algorytmu k-średnich od początkowych centrów klastrów, wykorzystano algorytm ALPSO do optymalizacji centrów klastrów k-średnich (KM-ALPSO). W celu rozwiązania problemu grupowania rzeczywistego zestawu danych, zastosowano analizę czynnikową do wyodrębnienia zmiennych numerycznych. Następnie zaproponowano metodę pomiaru niepodobieństwa do grupowania danych mieszanych. Porównano algorytm ALPSO z kilkoma metodami aktualizacji parametrów oraz przeprowadzono eksperymenty porównawcze, w których porównano KM-ALPSO na pięciu zbiorach danych UCI. Ostatecznie ulepszony algorytm grupowania IKM-ALPSO był stosowany w segmentacji klientów.

Wszystkie wyniki pokazały, że trzy proponowane metody są lepsze od istniejących modeli. Wyniki eksperymentów dowodzą także skuteczności i praktyczności IKM-ALPSO w zakresie segmentacji klientów. Natomiast w pracy (Yadegaridehkordi i in., 2021) badacze podjęli próbę sprawdzenia zachowań podróżnych wobec ekohoteli na podstawie dostępnych recenzji internetowych na stronie TripAdvisor. W związku z tym opracowali metodę łączącą segmentację oraz technikę TOPSIS (*ang. Technique for Order of Preference by Similarity to Ideal Solution*), aby podzielić podróżnych na segmenty na podstawie dostarczonych recenzji i priorytetyzować cechy ekohotelu w oparciu o ich poziom oraz znaczenia w każdym segmencie. Dane w postaci recenzji internetowych zostały pobrane od podróżnych odwiedzających malajskie ekohotele na stronie TripAdvisor. Wyniki wykazały, że jakość snu była jednym z najważniejszych czynników wpływających na wybór ekohotelu w większości segmentów. Opracowana w ramach cytowanego badania metoda pozwoliła na analizę recenzji i ocen podróżnych na temat ekohoteli, co umożliwiło identyfikowanie przyszłego zachowania potencjalnych gości w procesie podejmowania decyzji.

W artykule (Ampountolas, 2021) autor poddał analizie efektywność różnych tradycyjnych modeli prognoz szeregów czasowych dotyczących codziennego zapotrzebowania w różnych horyzontach czasowych. Wśród tych modeli znajdują się metoda SN (*ang. Seasonal Naive*), która w prognozie bazuje wyłącznie na obserwowanej wartości z tego samego okresu w poprzednim sezonie, metoda wygładzania HW (*ang. Holt-Winters*), która opiera się na estymacji trzech komponentów (poziomu, trendu i cyklu), model ARIMA (*ang. Autoregressive Integrated Moving Average*), który łączy składniki autoregresywny i integracyjny ze składnikiem średniej ruchomej, model stanowiący rozszerzenie modelu ARIMA o składowe zewnętrzne - SARIMAX (*ang. Seasonal Autoregressive Integrated Moving Average with Exogenous Regressors*), model ANN-MLP (*ang. Artificial Neural Network Multilayer Perceptron*), czyli rodzaj sieci neuronowej wykorzystywanej w analizie szeregów czasowych do modelowania zmienności warunkowej w danych finansowych, model sGARCH (*ang. Stochastic GARCH*). stanowiący rozszerzenie modelu GARCH (*ang. Generalized Autoregressive Conditional Heteroskedasticity*), w którym zakłada się, że zmienność warunkowa jest opisana jako funkcja przeszłych błędów (reszt) modelu (składowa autoregresyjna) oraz jako funkcja przeszłych wartości zmiennej (składowa średniej ruchomej) oraz model GJR-GARCH, który jest innym rozszerzeniem modelu GARCH o komponentę GJR (*ang. Glosten-Jagannathan-Runkle*), co czyni z niego model ekonometryczny uwzględniający asymetrię w reakcji zmienności na różne czynniki. Zestaw danych wykorzystywanych w analizie autora obejmował dzienne obserwacje, dotyczące zapotrzebowania na usługi hotelowe w amerykańskim mieście metropolitalnym w okresie od 2015 do

2019 roku, a także zestaw zmiennych egzogenicznych, takich jak temperatura, dni świąteczne oraz ranking konkurencyjnych hoteli. Wyniki eksperymentów wskazują, że pod względem miary dokładności MAPE: (1) model SARIMAX ze zmiennymi egzogenicznymi przewyższał model ANN-MLP z podobnymi zmiennymi egzogenicznymi oraz pozostałe modele dla sześciu z siedmiu horyzontów prognozy, z wyjątkiem jednego; (2) modele sGARCH(4,2) oraz GJR- GARCH(4,2) wykazywały wyższą dokładność prognozowania na wszystkich horyzontach czasowych. Autor ocenił wydajność modeli poprzez przeprowadzenie porównań między rozkładami prognoz różnych modeli przy użyciu testów Diebold-Mariano oraz Harvey- Leybourne-Newbold.

3.3. Zagadnienia teoretyczne dot. wybranych metod predykcji szeregów czasowych

Zagadnienie predykcji może odnosić się do zadania matematycznej (statystycznej) analizy regresji lub do zagadnienia klasyfikacji (Amit, Mozar, & Haase, 2023; Dinov, 2023; M. Dyduch, 2010; Hadaś, 2006; Rokach, Maimon, & Shmueli, 2023).

Klasyfikacja jest jednym z głównych zagadnień w dziedzinie uczenia maszynowego, które polega na przypisywaniu obiektów lub danych do wcześniej zdefiniowanych klas lub kategorii na podstawie ich cech lub wzorców. Celem klasyfikacji jest znalezienie odpowiedniej reguły lub modelu, który pozwala przyporządkować nowe, nieznane dane do odpowiedniej kategorii na podstawie wiedzy nabytej z danych uczących. Proces klasyfikacji można porównać do podejmowania decyzji na podstawie dostępnych informacji. Na przykład, możemy mieć zbiór zdjęć zwierząt, z których każde zdjęcie ma etykietę informującą, czy to jest obraz kota, psa lub ptaka. Na podstawie cech takich jak kształt uszu, nosa, sierść czy pióra, możemy zbudować model klasyfikacji, który będzie w stanie rozpoznawać i przypisywać nowe obrazy do odpowiednich kategorii zwierząt. Algorytmy klasyfikacji w uczeniu maszynowym uczą się na podstawie danych uczących, które zawierają informacje o obiektach i ich przypisanych etykietach.

Natomiast, regresja to metoda pozwalająca na określenie związku między zmiennymi niezależnymi lub cechami, a zmienną zależną lub wynikiem. Wyniki mogą być przewidywane, gdy związek między zmiennymi niezależnymi a zmienną zależną zostanie oszacowany.

Predykcja szeregów czasowych jest kluczowym aspektem analizy danych, m.in. w dziedzinach ekonometrii, finansów, meteorologii i innych, gdzie dane są gromadzone w określonych odstępach czasu. Szeregi czasowe to sekwencje danych, zazwyczaj ułożonych w chronologicznej kolejności. Celem predykcji szeregów czasowych jest prognozowanie przyszłych wartości na podstawie wcześniejszych obserwacji.

Podstawowym i najczęściej stosowanym podejściem do predykcji szeregów czasowych

są modele liniowe, będące typowymi modelami ekonometrycznymi. Opierają się one na założeniu, że przyszłe wartości szeregu czasowego można prognozować na podstawie liniowej zależności między danymi historycznymi, a przyszłymi punktami czasowymi. Model regresji liniowej jest jednym z podstawowych przykładów modeli liniowych, gdzie przewidywana wartość jest kombinacją liniową zmiennych objaśniających. Chociaż modele liniowe są użyteczne w wielu przypadkach, mają one swoje ograniczenia, zwłaszcza gdy szeregi czasowe zawierają skomplikowane wzorce, nieliniowe zależności i sezonowość. W odpowiedzi na te wyzwania, badacze opracowali bardziej złożone modele predykcyjne (Gilli, Maringer, & Schumann, 2019), które mogą lepiej radzić sobie z różnorodnymi właściwościami szeregów czasowych. Współcześnie stosowane zaawansowane modele szeregów czasowych obejmują m.in.:

- Modele ARIMA (*ang. Autoregressive Integrated Moving Average*) (Asteriou & Hall, 2011; Mills, 1990; Shumway & Stoffer, 2017). Modele ARIMA to klasa modeli, która integruje zarówno autoregresję (AR) - opierającą się na poprzednich wartościach szeregu czasowego, jak i średnią ruchomą (MA) - opierającą się na błędach modelu. Model ARIMA może być stosowany dla danych niestacjonarnych, posiadających trendy, a także sezonowość.
- Modele GARCH (*ang. Generalized Autoregressive Conditional Heteroskedasticity*) (Bollerslev, 1986). Te modele są stosowane głównie w analizie finansowej, gdzie występują heteroskedastyczne wariancje w szeregach czasowych. Modele GARCH pozwalają na modelowanie zmieniającej się zmienności w czasie, co jest ważne w prognozowaniu ryzyka finansowego.
- Maszyny wektorów nośnych SVM (*ang. Support Vector Machines*). Jest to metoda uczenia maszynowego stosowana zarówno w zadaniach klasyfikacji, jak i regresji. W kontekście predykcji szeregów czasowych, regresja SVM może być wykorzystywana do prognozowania przyszłych wartości na podstawie wcześniejszych obserwacji. Algorytm ten poszukuje parametrów hiperpłaszczyzny, która najlepiej oddziela dane szeregu czasowego, próbując jednocześnie minimalizować błąd predykcji. Metoda ta może być

stosunkowo skuteczna w przypadku szeregów czasowych o małej liczbie wymiarów i skomplikowanych zależnościach.

- Drzewa decyzyjne (*ang. Trees*) (Rokach i in., 2023) to intuicyjna metoda uczenia maszynowego, która dzieli dane na podzbiory, tworząc hierarchiczną strukturę decyzyjną. W przypadku predykcji szeregów czasowych, drzewa decyzyjne mogą być używane do prognozowania wartości przyszłych poprzez analizę historycznych wartości i odpowiednich cech czasowych. Drzewa decyzyjne mogą być również wykorzystywane do wykrywania sezonowości i wzorców w danych szeregów czasowych.
- Zespoły klasyfikatorów (*ang. Ensamble*) (Rokach i in., 2023). Zespoły klasyfikatorów, takie jak Random Forest czy Gradient Boosting, są technikami łączenia wielu prostych modeli, aby stworzyć bardziej skuteczny model predykcyjny. W kontekście predykcji szeregów czasowych, zespoły klasyfikatorów mogą dostarczyć bardziej precyzyjne prognozy poprzez łączenie różnorodnych podejść do modelowania czasowych wzorców. Mogą one być zwłaszcza użyteczne w przypadku dużych i złożonych szeregów czasowych.
- Sieci neuronowe (Rokach i in., 2023) są obecnie jednymi z najbardziej wszechstronnych i wydajnych metod uczenia maszynowego. W przypadku predykcji szeregów czasowych, sieci neuronowe, zwłaszcza rekurencyjne sieci neuronowe RNN (*ang. Recurent Neural Networks*) i ich odmiany, takie jak sieci LSTM (*ang. Long Short-Term Memory*), zdobyły dużą popularność. Sieci neuronowe są w stanie rozpoznawać i modelować złożone wzorce czasowe, długoterminowe zależności i sezonowość w danych szeregach czasowych. Model LSTM to rodzaj rekurencyjnej sieci neuronowej, zaprojektowanej w celu lepszego dopasowania do szeregów czasowych, zawierających długoterminowe zależności. Modele te są stosowane w predykcji, zwłaszcza w przypadkach, gdy dane charakteryzują złożone wzorce czasowe.
- Metoda GPR (*ang. Gaussian Process Regression*) (Schulz, Speekenbrink, & Krause, 2018) to zaawansowana technika modelowania probabilistycznego, umożliwiająca prognozowanie wartości szeregów czasowych z jednoczesnym oszacowaniem niepewności tych prognoz. Jest szczególnie przydatna w analizie danych niestacjonarnych oraz o złożonej strukturze. Metoda ta zakłada, że obserwacje szeregu czasowego są realizacjami procesu gaussowskiego, co pozwala na elastyczne modelowanie różnorodnych wzorców w danych. GPR jest efektywna w sytuacjach, gdy dane są zaszumione lub nie w pełni zrozumiałe, oferując jednocześnie precyzyjne oszacowanie niepewności predykcji. Jest to kluczowe w kontekście podejmowania decyzji biznesowych oraz w aplikacjach wrażliwych na błędy.

W zadaniu predykcji z wykorzystaniem zaawansowanych metod uczenia maszynowego, kluczowym jest dobór optymalnych wartości tzw. hiperparametrów, które mają zasadniczy wpływ na działanie modelu (Dinov, 2023). Poniżej zestawiono nazwy najważniejszych hiperparametrów w pięciu modelach uczenia maszynowego, które zostały zastosowane w ramach prac badawczych.

- Dla algorytmu SVM:
  - Kernel - Parametr ten określa rodzaj funkcji jądra używanej do przekształcenia danych. Popularne rodzaje jądra to: liniowy, wielomianowy i radialny.
  - Cost - Parametr regularyzacji, kontrolujący kompromis między dopasowaniem do danych uczących, a minimalizacją marginesu. Wysokie wartości skupią się na dopasowaniu danych uczących, podczas gdy niskie wartości będą preferować większy margines.
- Dla drzew decyzyjnych:
  - Maksymalna głębokość drzewa, która określa maksymalną liczbę poziomów w drzewie. Głębsze drzewa mogą dopasować się lepiej do danych uczących, ale mogą prowadzić do nadmiernego dopasowania (*ang. overfitting*).
  - Minimalna liczba próbek wymagana na liściu drzewa. Zapobiega tworzeniu zbyt małych liści, które mogą prowadzić do nadmiernego dopasowania.
- Dla sieci neuronowych:
  - Liczba warstw i liczba neuronów w każdej warstwie: Określa architekturę sieci neuronowej. Większa liczba warstw i neuronów może zwiększyć zdolność modelu do modelowania skomplikowanych wzorców, ale może wymagać więcej danych uczących i wydłużyć czas uczenia.
  - Współczynnik uczenia (*ang. learning rate*): Wpływa na tempo, z jakim sieć aktualizuje wagi podczas uczenia. Odpowiednie dobranie współczynnika uczenia może przyspieszyć proces uczenia i uniknąć oscylacji lub niestabilności.
- Dla zespołów klasyfikatorów:
  - Liczba klasyfikatorów bazowych, która określa, ile takich klasyfikatorów zostanie połączonych w zespole.
  - Metoda łączenia klasyfikatorów, która określa sposób łączenia wyników klasyfikatorów bazowych, na przykład głosowanie większościowe (*ang. voting*) lub głosowanie ważone (*ang. weighted voting*).
- Dla algorytmu GPR:

- Wybór funkcji korelacji, która określa, jak zmienne są ze sobą powiązane w modelu. Przykładowe funkcje to funkcja radialna, liniowa lub wielomianowa.
- Parametry funkcji korelacji, które mają wpływ na kształt i zakres modelu GPR.

W niniejszej pracy, dobór optymalnych hiperparametrów wykonany był w procesie optymalizacji, z wykorzystaniem metody Bayesowskiej. Metoda optymalizacji Bayesowska, zwana również optymalizacją Bayesowską, jest techniką optymalizacji funkcji, która opiera się na podejściu probabilistycznym (Bull, 2011; Gelbart, Snoek, & Adams, 2014; Snoek, Larochelle, & Adams, 2012; Zulfiqar, Gamage, Kamran, & Rasheed, 2022). Metoda ta jest szczególnie przydatna w przypadku funkcji, które są kosztowne w obliczeniach i trudne do zrozumienia, jak to jest w przypadku sieci neuronowych. Optymalizacja Bayesowska ma na celu znalezienie globalnego minimum (lub maksimum) funkcji, wykorzystując minimalną liczbę prób (liczbę prób można zdefiniować na dowolnym poziomie, w zadaniu wykonano 300 iteracji). Podstawowa istota tej metody polega na aktualizowaniu wiedzy o funkcji w trakcie kolejnych iteracji. Zamiast wykonania dużej liczby prób i optymalizacji wyniku na podstawie uzyskanych rezultatów, metoda wykorzystuje modele probabilistyczne, aby określić, które próby są najbardziej wartościowe. W ten sposób metoda ta osiąga najlepsze wyniki przy minimalnym zużyciu zasobów obliczeniowych.

4. Analiza zmienności zysku operacyjnego brutto dla wybranych polskich hoteli w dziedzinie czasu i częstotliwości

Analiza zmienności cech to, syntetycznie ujmując, wyznaczenie i selekcja cech, za pomocą których będzie możliwe skuteczne przewidywanie przyszłości. Stanowi to jeden z najistotniejszych i zarazem najtrudniejszych etapów budowy narzędzia predykcyjnego (Hu, Wang, Law, & Geng, 2023).

Niniejsza dysertacja jako cechy rozpatruje historyczne wartości wskaźników finansowych dotyczących działalności siedmiu polskich hoteli. W ramach prac badawczych dla każdego z hoteli pozyskano tabele USALI z różnych horyzontów czasowych, zawierających dane finansowe zwykle od rozpoczęcia ich działalności gospodarczej. Były to:

- HDSI Katowice (KAT), dane z okresu 1.01.2014 – 31.12.2021.
- HDS Piaseczno (PIA), dane z okresu 1.01.2014 – 31.12.2021.
- HDSP Opole (OPHS), dane z okresu 1.01.2015 – 31.12.2021.
- HDSP Poznań (POZ), dane z okresu 1.01.2017 – 31.12.2021.
- HM Opole (OPHM), dane z okresu 1.02.2017 – 31.12.2021.
- HMC Częstochowa (CZE), dane z okresu 1.03.2017 – 31.12.2021.
- HM Karpacz (KAR), dane z okresu 1.05.2017 – 31.12.2021.

Ze względu na wystąpienie pandemii Covid-19 i spowodowanej nią zmianami sytuacji ekonomicznej, dane z lat 2020 – 2021 zostały wyłączone z interpretacji.

W dalszej części rozdziału zostaną przedstawione wyniki analiz, kluczowych z punktu widzenia modelowania, danych finansowych. Należy dodać, że w celu ułatwienia porównywania różnych zmiennych, identyfikacji wartości odstających, dla zapewnienia spójności i poprawności interpretacji wyników oraz ostatecznie w celu zachowania poufności działalności przedsiębiorstwa, którego dane dotyczą, wszystkie wartości zostały poddane standaryzacji metodą Z-Score. Standaryzacja wartości metodą Z-score jest techniką stosowaną w analizie danych, która przekształca zbiór danych w taki sposób, że średnia wartość wynosi zero, a odchylenie standardowe wynosi jeden. Dzięki temu można porównywać wartości między różnymi zmiennymi, eliminując jednostki miary i skalując dane w skali względnej. Aby zastosować standaryzację metodą Z-score, obliczono średnią wartość (μ) i odchylenie standardowe (σ) dla danej zmiennej. Następnie każdą wartość w zbiorze danych odjęto od średniej i podzielono przez odchylenie standardowe, zgodnie z równaniem: $Z = (x - \mu) / \sigma$. Wyższe wartości punktu Z-score wskazują, że dane są powyżej średniej, podczas gdy niższe wartości wskazują, że dane są poniżej średniej. Należy zaznaczyć, że analiza zmienności zysku

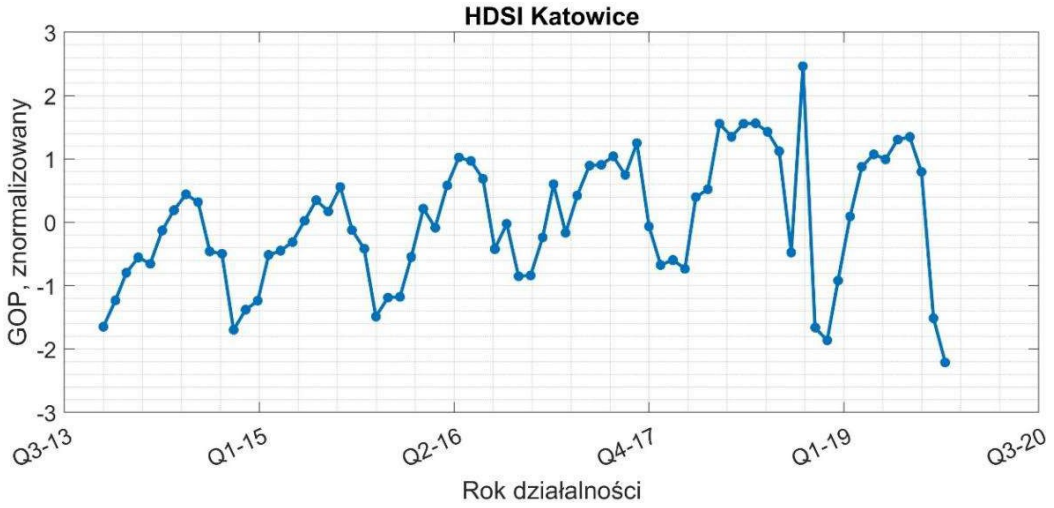
operacyjnego brutto dla badanych hoteli, jest etapem wstępnym, niejako wprowadzającym do kolejnych bardziej zaawansowanych analiz opisanych w niniejszej dysertacji. Badanie zmienności zysku operacyjnego brutto zostało przeprowadzone w czterech etapach. W etapie pierwszym (por. rozdz. 4.1) zbadano zmienność czasową zysku operacyjnego brutto, w którym zaprezentowano i syntetycznie opisano szeregi czasowe dla każdego hotelu. W etapie drugim (por. rozdz. 4.2) wyznaczono rozkłady empiryczne oraz przeprowadzono ich analizę dla każdego z hoteli, na etapie trzecim przeprowadzono analizę zmienności periodycznej (por. rozdz. 4.3). Natomiast na etapie czwartym, przeprowadzono analizę zysku operacyjnego brutto w dziedzinie częstotliwości (por. rozdz. 4.4). Cztery etapy przeprowadzonej analizy są bardzo ważne z punktu widzenia wstępnej analizy ekonomicznej, która pozwoli na zidentyfikowali wahań sezonowych oraz okresowych w badanej cesze.

4.1. Zmienność czasowa zysku operacyjnego brutto

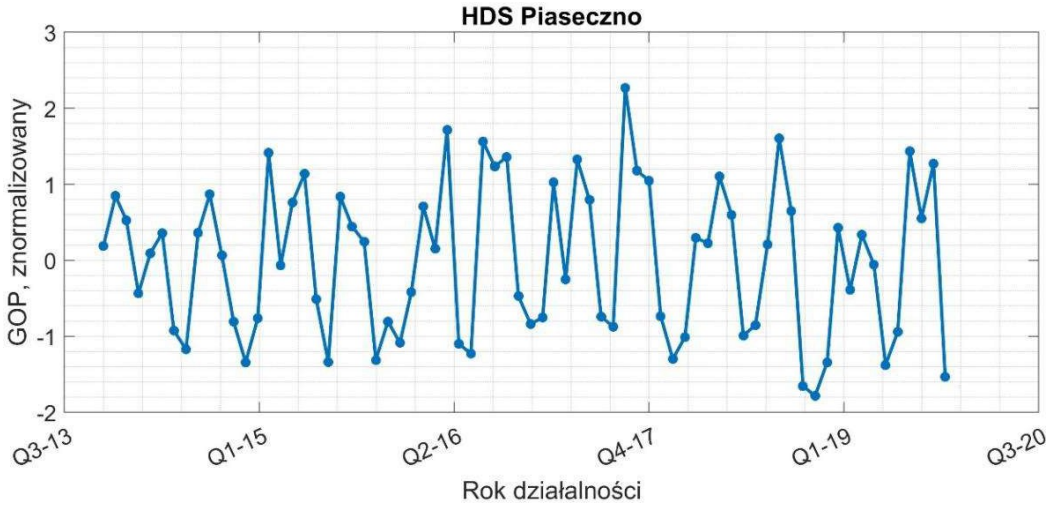
Analiza zmienności czasowej bazowała na wykreśleniu przebiegów czasowych, które pozwalają na wizualizację zmian w danych w funkcji czasu, co pomaga w identyfikacji wzorców, trendów i sezonowości, będącymi kluczowymi dla zrozumienia i prognozowania przyszłych zachowań danej zmiennej (Borzyszkowski, 2014; Butler & Schütze, 2005; Duro & Turrión-Prats, 2019; Eurostat, 2023; Petrevska, 2015; Raymond, 1976). Przebieg czasowy, nazywany również wykresem czasowym lub wykresem szeregu czasowego, pozwala na ujawnienie ogólnego kierunku zmian wartości danej zmiennej w czasie, tj. trendu. Wykres czasowy może również wykazywać regularne wzorce powtarzających się zmian, które są związane z sezonowością, np. zmiany związane z porami roku, miesiącami, dniami tygodnia itp. Oprócz sezonowości, przebieg czasowy może także ujawnić długoterminowe cykle, które powtarzają się na przestrzeni lat. Na wykresie czasowym można też zauważyć wartości odstające lub anomalie, które odbiegają od ogólnego wzorca zmian i mogą być ważnymi punktami do analizy. Ponadto przebiegi czasowe pozwalają na wizualizację stacjonarności (czyli stałych statystyki w czasie) oraz wahania w danych, które mogą być związane z różnymi czynnikami, takimi jak wzrosty sezonowe, wydarzenia lub zmiany w środowisku.

W pierwszej kolejności rozpatrzono zmienną GOP, która jest miarą zyskowności operacyjnej, uwzględniającą wszystkie przychody i koszty związane z działalnością operacyjną hotelu, nie wliczając jednak w to kosztów finansowych, podatkowych i inwestycji. Na rysunkach 4.1 – 4.7 przedstawiono przebiegi czasowe standaryzowanych wartości zmiennej GOP dla hotelu HDSI Katowice, HDS Piaseczno, HDSP Opole, HDSP Poznań, HM Opole,

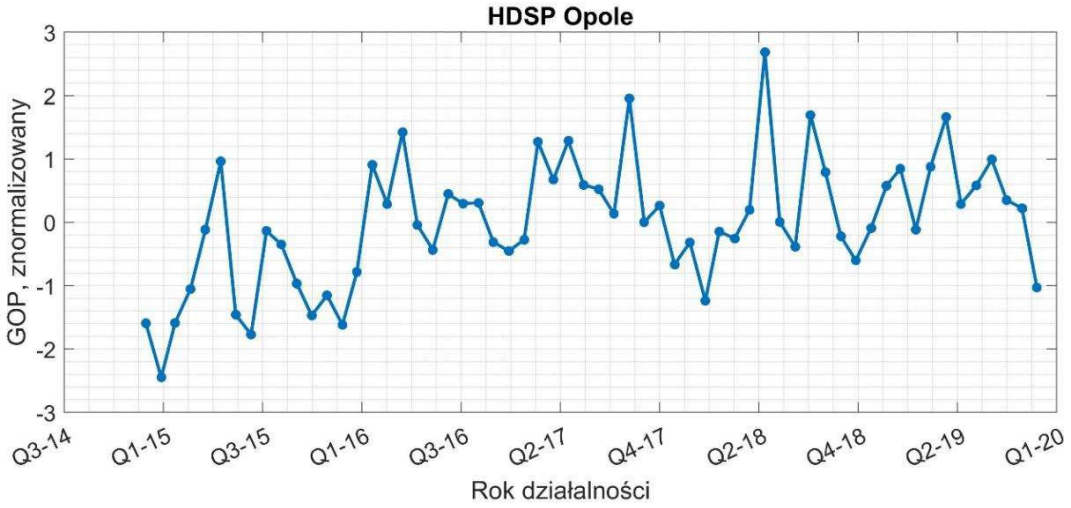
HDC Częstochowa i HM Karpacz, odpowiednio. Na rysunkach krzywą niebieską zobrazowano wartości GOP osiągnięte w kolejnych miesiącach w okresie jak podano na początku rozdziału. Dla hoteli HDSI Katowice i HDS Piaseczno było to po 72 wartości (6 lat), dla HDSP Opole były to 60 wartości (5 lat), a dla pozostałych hoteli były to po 36 wartości (3 lata).



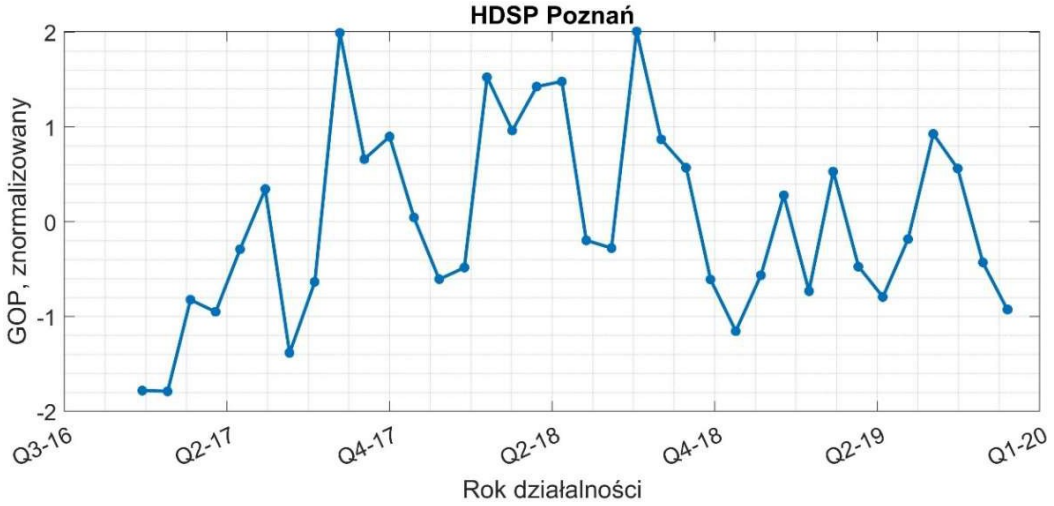
Rys. 4.1 Przebieg czasowy standaryzowanych wartości zmiennej GOP dla HDSI Katowice. Źródło: Opracowanie własne.



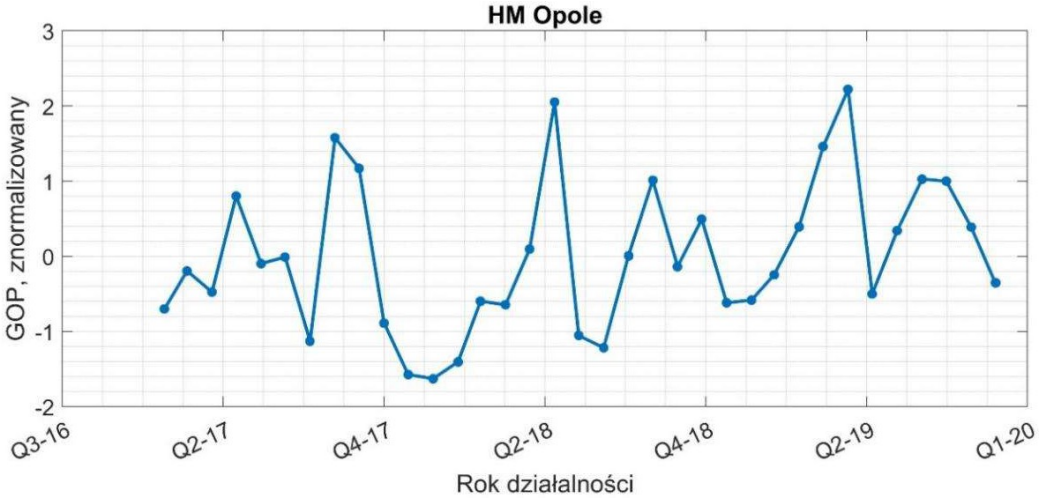
Rys. 4.2 Przebieg czasowy standaryzowanych wartości zmiennej GOP dla HDS Piaseczno. Źródło: Opracowanie własne.



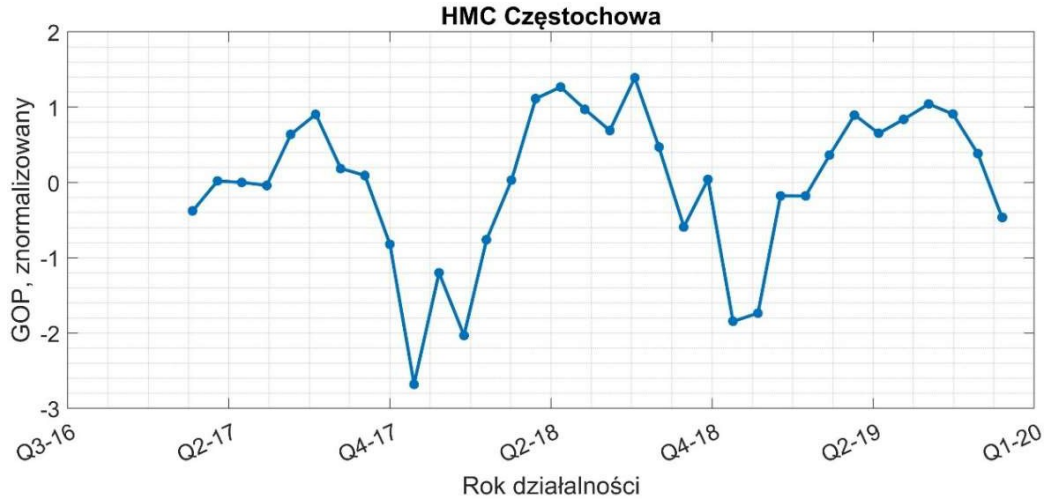
Rys. 4.3 Przebieg czasowy standaryzowanych wartości zmiennej GOP dla HDSP Opole. Źródło: Opracowanie własne.



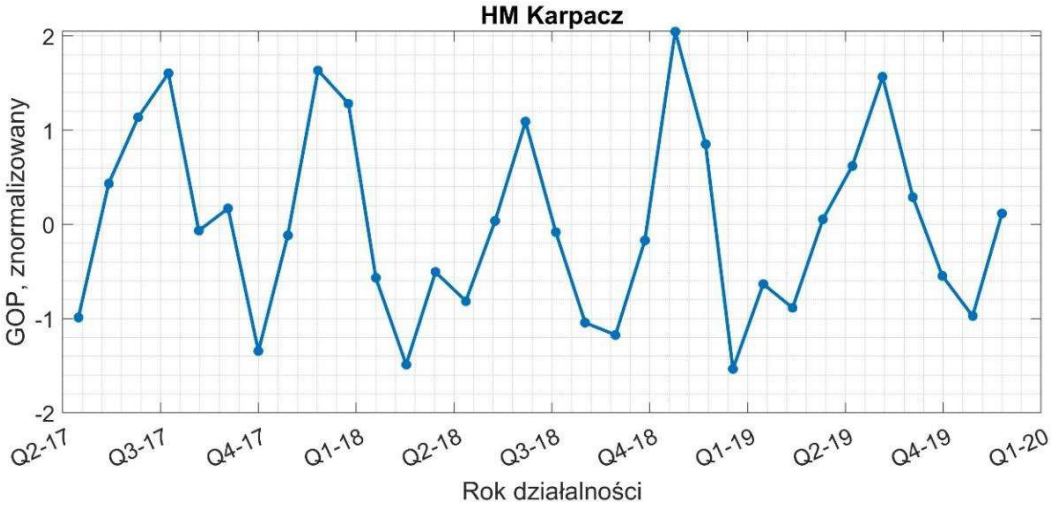
Rys. 4.4 Przebieg czasowy standaryzowanych wartości zmiennej GOP dla HDSP Poznań. Źródło: Opracowanie własne.



Rys. 4.5 Przebieg czasowy standaryzowanych wartości zmiennej GOP dla HM Opole. Źródło: Opracowanie własne.



Rys. 4.6 Przebieg czasowy standaryzowanych wartości zmiennej GOP dla HMC Częstochowa. Źródło: Opracowanie własne.



Rys. 4.7 Przebieg czasowy standaryzowanych wartości zmiennej GOP dla HM Karpacz. Źródło: Opracowanie własne.

Analiza wykresów 4.1 — 4.7 pozwala zauważyć, że każdy z hoteli charakteryzuje się odmienną dynamiką zmiennej GOP. Można również zauważyć, że dla każdego z badanych hoteli pojawiają się zmienności sezonowe. Zmienności te na obecnym etapie analiz są zidentyfikowane wyłącznie wstępnie. Aby tę zmienność określić w sposób szczegółowy, wykonano analizy periodyczności, analizę rozkładów gęstości prawdopodobieństwa oraz określono widma częstotliwościowe. W kolejnych podrozdziałach zostaną przedstawione wyniki tych analiz.

4.2. Analiza rozkładów empirycznych zysku operacyjnego brutto

W następnej kolejności wykonano analizy empirycznych rozkładów gęstości prawdopodobieństwa zmiennej GOP, tzw. histogramy, dla poszczególnych hoteli. Histogram

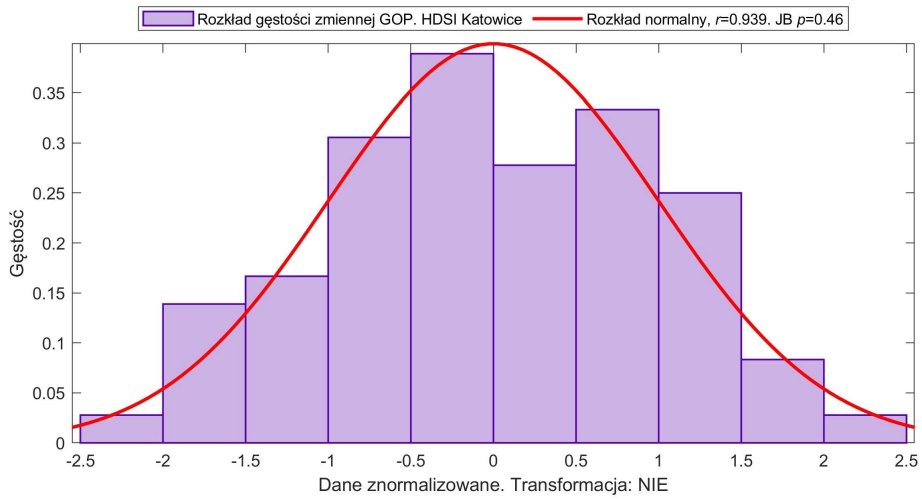
umożliwia zrozumienie charakterystyk danych oraz pozwala na identyfikację występujących w nich wzorców. Pokazuje, jak dane są rozłożone wokół różnych wartości, np. czy są skoncentrowane wokół jednej wartości (unimodalny), czy mają dwie dominujące wartości (bimodalny) lub więcej (multimodalny). Histogram pozwala zobrazować przybliżoną wartość średnią, która jest reprezentowana jako środek ciężkości histogramu, oraz medianę, która znajduje się w miejscu, gdzie wykres dzieli się na dwie równe części. Ponadto na podstawie analizy wizualnej histogramu można określić, czy dane mają skośny rozkład (czyli są przesunięte w lewo lub prawo) lub czy są bardziej równomiernie rozłożone. Histogram może ujawnić anomalie lub wartości odstające od reszty danych, są to wartości widoczne jako odosobnione szczyty na wykresie. Kolejnymi cechami, jakie histogram pozwala określić, to czy dane są skoncentrowane wokół niewielkiego zakresu wartości (mała rozpiętość) czy też mają szeroki zakres rozkładu (duża rozpiętość) oraz czy dane są symetryczne.

Na rys. 4.8–4.14 przedstawiono graficzną prezentację rozkładów gęstości wraz z dopasowaniem funkcji gęstości prawdopodobieństwa (FGP) rozkładu normalnego oraz parametrami opisującymi jego charakter w postaci statystyki Jarque'a-Berry i skośności.

Rozkład normalny, jest jednym z najważniejszych rozkładów statystycznych, który jest szeroko stosowany w analizie danych ze względu na swoje użyteczne właściwości i interpretowalność. Jest on unimodalny i symetryczny wokół średniej (μ), co oznacza, że większość danych koncentruje się wokół tego parametru. Matematyczna postać rozkładu normalnego to:

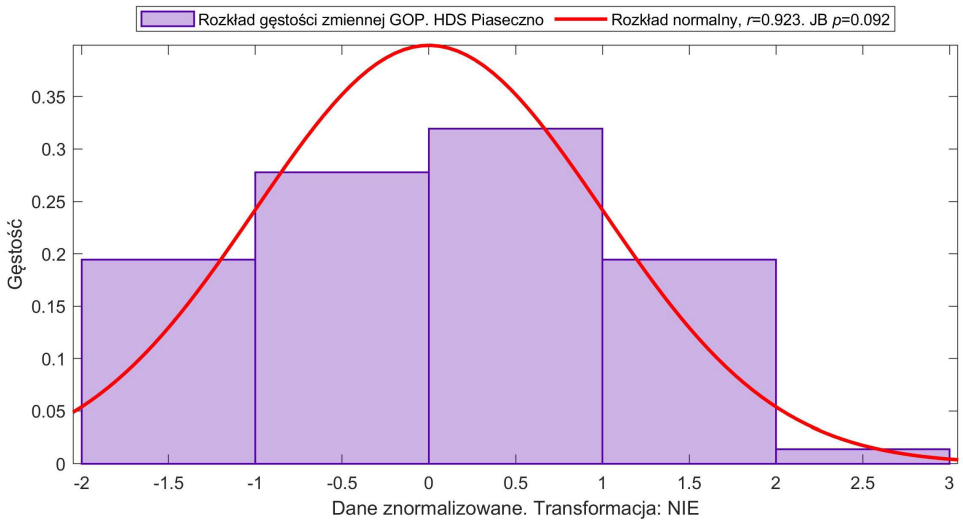
$$F(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\left(\frac{(x-\mu)^2}{2\sigma^2}\right)}, dla\ \sigma \neq$$

gdzie: $F(x)$ – funkcja gęstości prawdopodobieństwa rozkładu normalnego, μ – średnia (wartość oczekiwana), σ – odchylenie standardowe, które mierzy rozproszenie danych wokół średniej.

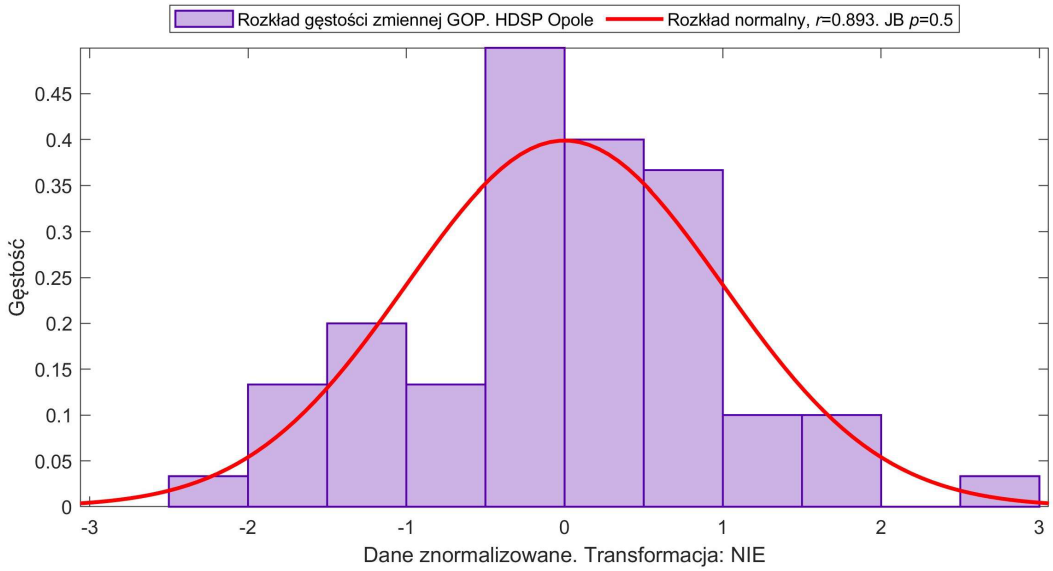


Rys.4.8 Wykres rozkładu empirycznej gęstości prawdopodobieństwa dla zmiennej GOP rejestrowanej w hotelu HDSI Katowice. Źródło: Opracowanie własne.

Rozkład zmiennej GOP zarejestrowanej w hotelu HDSI Katowice można zaliczyć do grupy rozkładów normalnych, co oceniono na podstawie wartości parametru *p*, uzyskanego w teście Jarque'a-Berry *p*=0,46 oraz na podstawie współczynnika korelacji *r* na poziomie zbliżonym do 0,939. Równie dobre dopasowanie FGP uzyskano dla rozkładu zmiennej GOP dla hotelu HDS Piaseczno, gdzie wartości *r* przekroczył wartość 0,92. Nieco niższe wartości współczynnika *r*, na poziomie 0,87 – 0,89, uzyskano dla dopasowania FGP do zmiennej GOP z hotelu HDSI Opole.

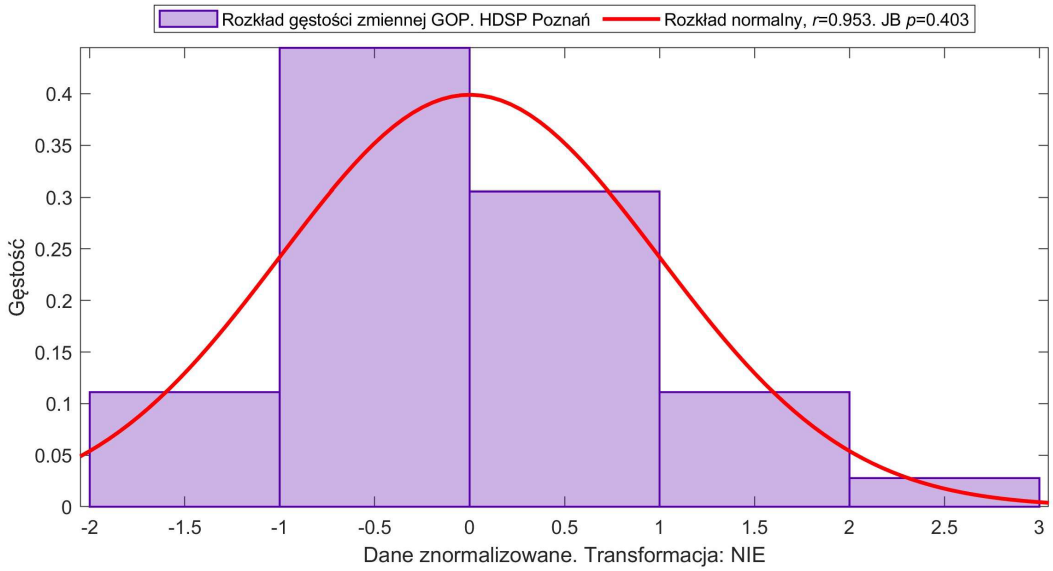


Rys.4.9 Wykres rozkładu empirycznej gęstości prawdopodobieństwa dla zmiennej GOP rejestrowanej w hotelu HDS Piaseczno. Źródło: Opracowanie własne.

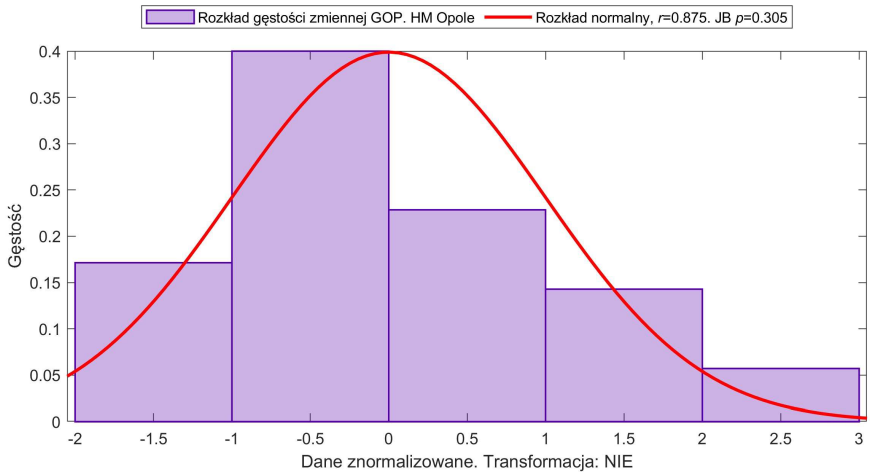


Rys.4.10 Wykres rozkładu empirycznej gęstości prawdopodobieństwa dla zmiennej GOP rejestrowanej w hotelu HDSP Opole. Źródło: Opracowanie własne.

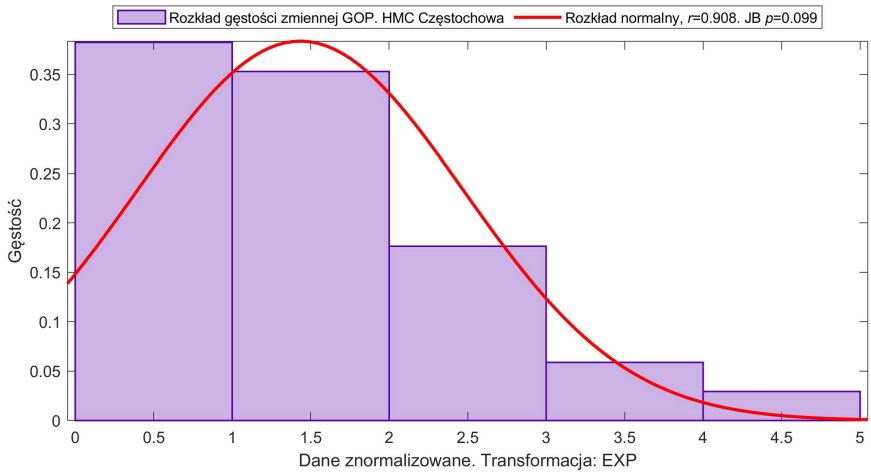
Rozkład GOP dla hotelu HDSP Opole charakteryzuje się silną reprezentacją wartości zerowych, w stosunku do pozostałych wartości. Rozkłady zmiennej GOP dla hoteli HDSP Poznań, HM Opole, HM Karpacz są prawostronnie skośne, a dopasowania krzywych FGP są zgodne z rozkładem normalnym. Rozkład dla hotelu HMC Częstochowa jest najsilniej prawostronnie skośny, ale wciąż zalicza się grupy rozkładów normalnych, co potwierdza test Jarque'a-Berry p>0,05.



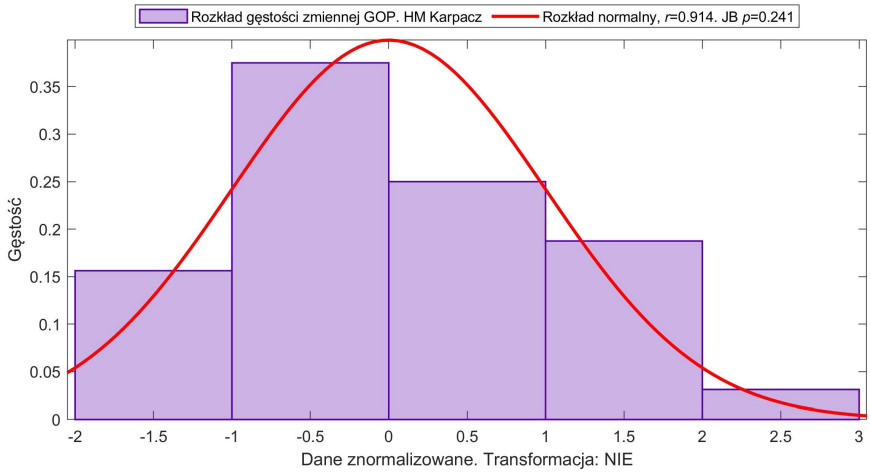
Rys.4.11 Wykres rozkładu empirycznej gęstości prawdopodobieństwa dla zmiennej GOP rejestrowanej w hotelu HDSP Poznań. Źródło: Opracowanie własne.



Rys.4.12 Wykres rozkładu empirycznej gęstości prawdopodobieństwa dla zmiennej GOP rejestrowanej w hotelu HM Opole. Źródło: Opracowanie własne.



Rys.4.13 Wykres rozkładu empirycznej gęstości prawdopodobieństwa dla zmiennej GOP rejestrowanej w hotelu HMC Częstochowa. Jedyna zmienna, którą poddano transformacji funkcją wykładniczą. Źródło: Opracowanie własne.



Rys.4.14 Wykres rozkładu empirycznej gęstości prawdopodobieństwa dla zmiennej GOP rejestrowanej w hotelu HM Karpacz. Źródło: Opracowanie własne.

Jak wspomniano powyżej, do oceny normalności rozkładów zmiennych GOP w poszczególnych hotelach zastosowano test Jarque'a-Berry (JB). Test ten jest popularnym statystycznym testem normalności, który w dysertacji posłużył również, co zostanie przedstawione w dalszych rozdziałach, do sprawdzenia czy reszty pochodzą z rozkładu normalnego. Wyniki obliczeń, w szczególności wartości parametru *p*, dla zmiennych GOP pozyskanych ze wszystkich hoteli zestawiono w tabeli 2.

Test JB opiera się na obliczeniu dwóch statystyk: skośności i kurtozy:

$$\gamma = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3}{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^{3/2}}$$

$$K = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^4}{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^{2}}$$

gdzie: n – liczba obserwacji w próbie, x_i – kolejne obserwacje danych, $\bar{x}$ – średnia arytmetyczna z danych.

Są to miary statystyczne, opisujące kształt rozkładu, które pozwalają zrozumieć, jak dane są rozłożone wokół średniej wartości. Kurtoza mierzy, jak bardzo ogony rozkładu danych są "grube" lub "ciężkie" w porównaniu z rozkładem normalnym. Formalnie, kurtoza określa stopień ogonów rozkładu, czy są bardziej rozwinięte (bardzo grube) lub mniej rozwinięte (bardzo cienkie) niż w rozkładzie normalnym. Wartość kurtozy określana jest względem liczby 3:

- Kurtoza > 3 – rozkład ma tzw. "grube ogony", co oznacza, że zawiera więcej ekstremalnych wartości niż rozkład normalny. Jest to nazywane leptokurtycznym rozkładem.
- Kurtoza = 3 – rozkład ma tę samą "grubość" ogonów co rozkład normalny. Jest to nazywane mesokurtycznym rozkładem.
- Kurtoza < 3 – rozkład ma tzw. "cienkie ogony", co oznacza, że zawiera mniej ekstremalnych wartości niż rozkład normalny. Jest to nazywane platykurtycznym rozkładem.

Skośność mierzy asymetrię rozkładu danych wokół średniej wartości.

- Skośność > 0 – rozkład jest prawostronnie skośny (skośność dodatnia), co oznacza, że ogon rozkładu jest bardziej rozwinięty w prawo, a większość danych koncentruje

się bardziej w lewo od średniej.

- Skośność = 0 – rozkład jest symetryczny wokół średniej, co oznacza, że dane są równomiernie rozłożone wokół średniej.
- Skośność < 0 – jest lewostronnie skośny (skośność ujemna), co oznacza, że ogon rozkładu jest bardziej rozwinięty w lewo, a większość danych koncentruje się bardziej w prawo od średniej.

Wartości skośności i kurtozy dla wszystkich zmiennych GOP podano w tabeli 2. Dane zawarte w tabeli 2 potwierdzają, że GOP dla hoteli HDSI Katowice i HDSP Opole mają rozkład symetryczny, o skośności ±0,04. Natomiast wartości kurtozy pokazują, że grubości ogonów są zbliżone do tych z rozkładu normalnego lub cieńsze.

Zarówno skośność jak i kurtoza zostają znormalizowane, a następnie oblicza się statystykę testową JB, która jest iloczynem kwadratów obu znormalizowanych statystyk i ma rozkład chi–kwadrat z dwoma stopniami swobody. Jeśli wartość *p* (prawdopodobieństwo) dla statystyki JB jest mniejsza niż przyjęty poziom istotności (tutaj α = 0,05), to odrzucamy hipotezę o normalności danych. W przeciwnym przypadku, nie ma podstaw do odrzucenia tej hipotezy.

$$JB = \frac{n}{6}(\gamma^2 + 0{,}25(K-3)^2)$$

gdzie: *n* – liczba obserwacji w próbie, γ – skośność, *K*– kurtoza.

W następnym kroku wykonano badania stacjonarności zmiennej GOP. W tym celu zastosowano test Dickey-Fullera, który sprawdza, czy dany szereg czasowy jest stacjonarny. Stacjonarność oznacza, że wartości szeregu czasowego nie zmieniają się systematycznie w czasie, co jest kluczowe zwłaszcza przy zastosowaniu modeli liniowych. Hipoteza zerowa (H0) dotyczy zawartości w szeregu czasowym jednostkowego pierwiastka, co świadczy o jego niestacjonarności. Hipoteza alternatywna (H1) wskazuje na szereg stacjonarny.

Test Dickey-Fullera bazuje na regresji szeregów czasowych. Standardowa forma testu zakłada, że szereg czasowy y_t jest opisany równaniem:

$$\Delta y_t = \propto + \beta t + \gamma y_{t-1} + \delta_1 \Delta y_{t-1} + \ldots + \delta_p \Delta y_{t-p} + \varepsilon_t$$

gdzie:

- Δy_t — pierwsza różnica szeregu ($y_t - y_{t-1}$),
- $\propto$ — reprezentuje wyraz wolny,
- βt — reprezentuje trend czasowy (opcjonalnie),

- γ — współczynnik, który jest kluczowy dla testu jednostkowego pierwiastka,
- δ_i — współczynniki dla poprzednich różnic,
- ε_t — składnik losowy.

W przypadku gdy wartość *p* jest mniejsza niż przyjęty poziom α= 0,05, odrzucamy hipotezę o niestacjonarności szeregu. Wartości *p* testu Dickey-Fullera dla wszystkich zmiennych GOP podano w tabeli 2. Obliczenia nie wskazały, aby w którymkolwiek przypadku zabrakło podstaw do odrzucenia hipotezy zerowej, a więc wszystkie szeregi potwierdzono jako stacjonarne.

W przypadku zmiennej GOP pozyskanej z HMC Częstochowa stwierdzono relatywnie dużą ujemną skośność, dlatego szereg ten poddano transformacji wykładniczej. W tabeli 2, w ostatniej po prawej kolumnie znajduje się informacja na temat zastosowanych transformacji.

Tabela 2. Zestawienie zbiorcze miar statystycznych obliczonych dla zmiennej GOP w kolejnych hotelach.

Hotel	Zmienna GOP				
	skośność	kurtoza	Test Dickey-Fullera. Wartość p	Test Jarque'a-Berry. Wartość p	Transformacja danych?
HDSI Katowice	-0,04	2,38	0,001	0,460	NIE
HDS Piaseczno	0,10	1,92	0,001	0,092	NIE
HDSP Opole	0,04	3,09	0,001	0,500	NIE
HDSP Poznań	0,28	2,31	0,001	0,403	NIE
HM Opole	0,43	2,48	0,001	0,305	NIE
HMC Częstochowa	-0,93	3,29	0,012	0,099	EXP
HM Karpacz	0,35	2,13	0,001	0,241	NIE

Źródło: Opracowanie własne

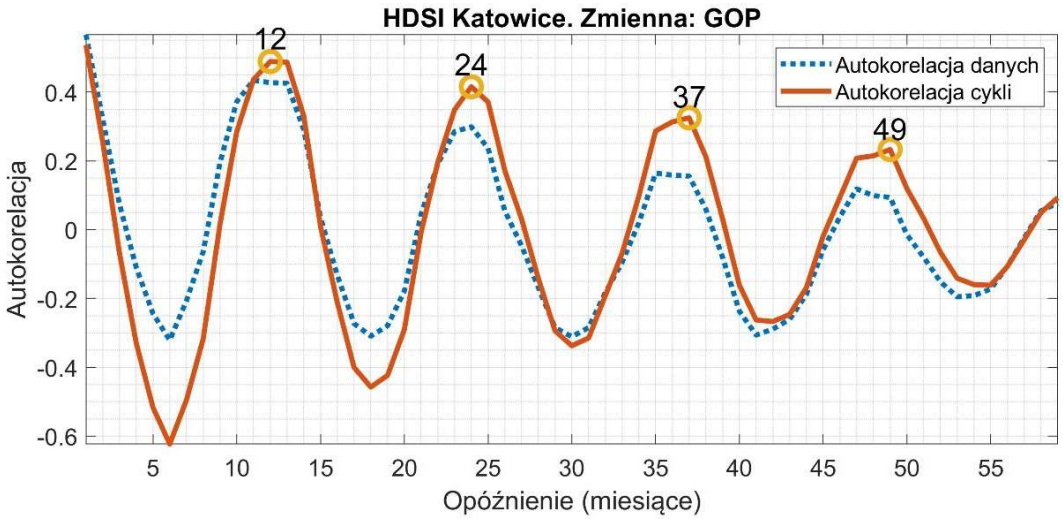
4.3. Analiza zmienności periodycznej

Analizę periodyczności wykonano na podstawie autokorelacji danych i cykli, występujących w danych. Wykres autokorelacji nazywany też korelogramem lub wykresem funkcji autokorelacji, to kluczowe narzędzie w analizie szeregów czasowych, które pomaga zrozumieć wzorce i zależności między bieżącymi, a wcześniejszymi obserwacjami w danych. Uwidacznia korelację pomiędzy danymi, a ich opóźnionymi wersjami (poprzednimi obserwacjami) w czasie. Na wykresie autokorelacji osie przedstawiają różne opóźnienia (zwane także "przesunięciami") danych względem samych siebie. Autokorelacja mierzy, jak bardzo poprzednie wartości wpływają na bieżące wartości. Kiedy wartość autokorelacji jest bliska jedności, oznacza to, że obserwacje mają silną korelację pozytywną z wcześniejszymi wartościami. Wartość bliska ujemnej jedności (-1) wskazuje silną korelację ujemną, a wartość bliska 0 oznacza brak korelacji. Na wykresie autokorelacji można też zauważyć, czy występuje jakikolwiek wzorzec powtarzający się w danych w określonym odstępie czasu, co może wskazywać na występowanie sezonowości lub okresowości. Ponadto, wykres autokorelacji pozwala na stwierdzenie czy dane

są stacjonarne (o stałych statystykach w czasie) czy też mają tendencję do wzrostu lub spadku. Stacjonarny szereg czasowy będzie charakteryzował się autokorelacją, która szybko konwerguje do zera, nie wykazując żadnych wyraźnych wzorców. Dodatkowo wykres ten dostarcza informacji o naturze szeregu czasowego. Na przykład, obserwacje z wysokimi wartościami autokorelacji, w dużych odstępach czasu, mogą wskazywać na długoterminowe zależności w danych.

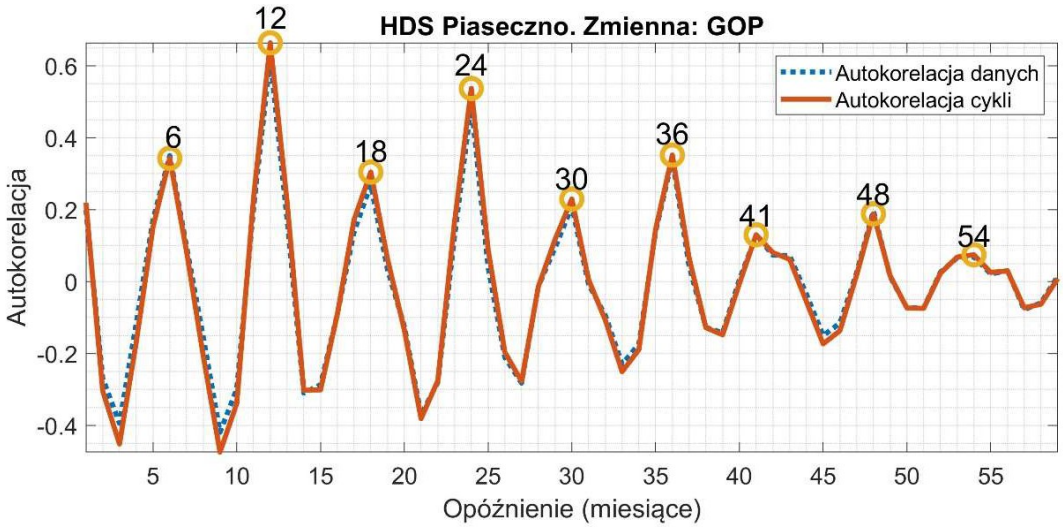
Na wykresach 4.15 – 4.21 zestawiono uzyskane wyniki analizy zmiennej GOP, kolejno dla rozpatrywanych hoteli. Krzywą kropkowaną w kolorze niebieskim zaznaczono wartości autokorelacji obliczone dla zmiennej GOP w postaci standaryzowanej, krzywą w kolorze czerwonym oznaczono wartości autokorelacji składowej sezonowej/cyklicznej, którą wyznaczono z wykorzystaniem filtru Savitzky'ego-Golaya. Natomiast kółkami w kolorze żółtym zaznaczono wartość liczbową głównych składowych cyklicznych wyrażoną w miesiącach. Liczba przy żółtym kółku determinuje wartość najistotniejszego opóźnienia.

Autokorelacja obliczona dla zmiennej GOP zarejestrowanej w hotelu HDSI Katowice wskazuje na występowanie składowej cyklicznej rocznej (12 miesięcy) i dwuletniej (24 miesięcy). Periodyczności są harmoniczne i pojawiają się w kolejnych cyklach rocznych z tendencją tłumiącą. Autokorelacja danych i cykli ma podobny przebieg, ale krzywe nie pokrywają się w pełni.



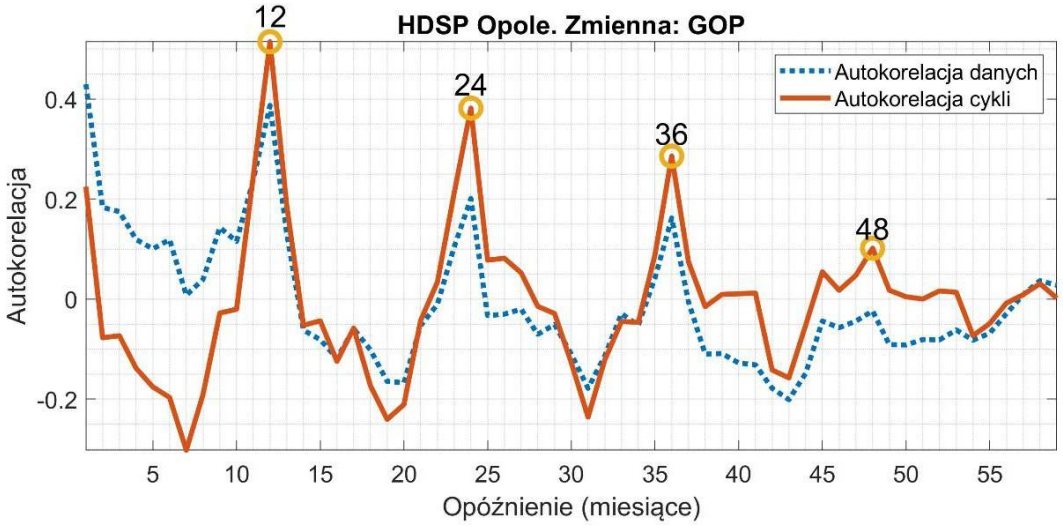
Rys. 4.15 Wykres autokorelacji standaryzowanych wartości zmiennej GOP dla HDSI Katowice. Źródło: Opracowanie własne.

W przypadku hotelu HDS Piaseczno, poza obecnością bardziej znaczących składowych rocznych (o opóźnieniu 12, 24, 36 miesięcy), występują również nieco słabsze składowe sezonowe – półroczne (o opóźnieniu 6, 18, 30 miesięcy). Krzywe autokorelacji danych i cykli pokrywają się w pełni.



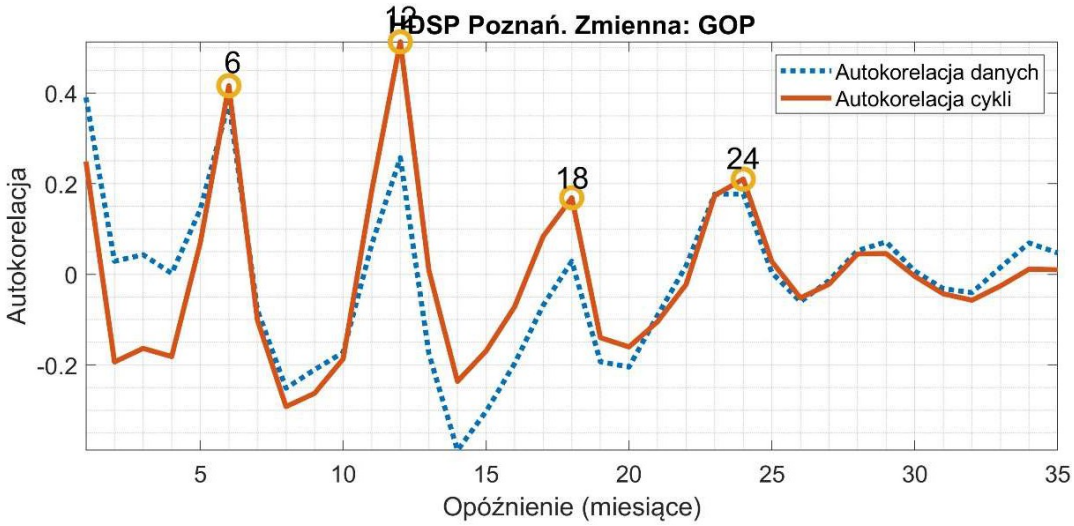
Rys. 4.16 Wykres autokorelacji standaryzowanych wartości zmiennej GOP dla HDS Piaseczno. Źródło: Opracowanie własne.

Wartości głównych składowych cykli zmiennej GOP dla hotelu HDSP Opole ma te same wartości jak w przypadku HDSI Katowice, tj. zawiera jedynie cykle roczne. Różnica obu krzywych objawia się w mniej sinusoidalnym przebiegu dla Opola. Również tutaj krzywe autokorelacji danych i cykli nie pokrywają się.



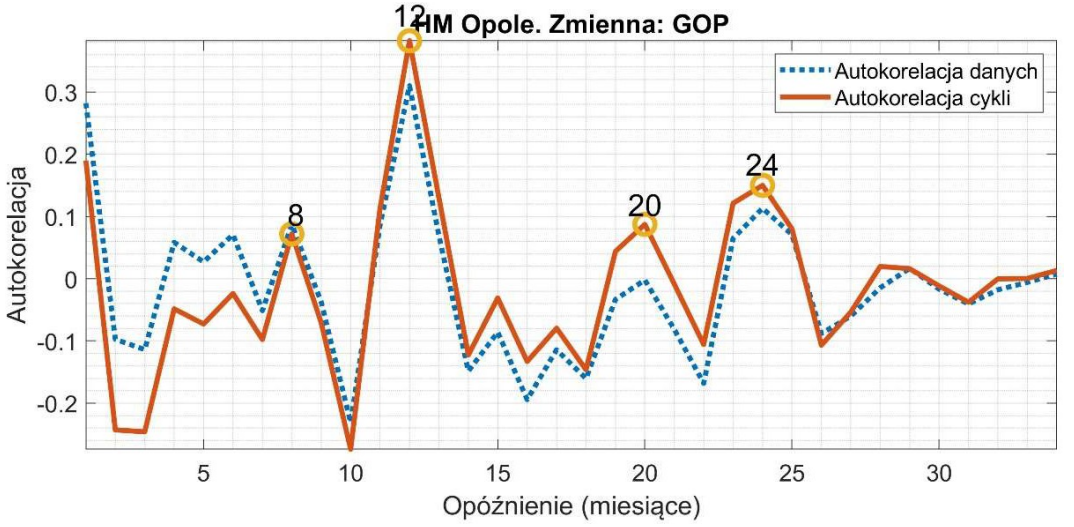
Rys. 4.17 Wykres autokorelacji standaryzowanych wartości zmiennej GOP dla HDSP Opole. Źródło: Opracowanie własne.

Z kolei autokorelacje wyznaczone dla hotelu HDSP Poznań, których przebiegi również się nie pokrywają, zawierają zarówno składowe roczne jak i półroczne.



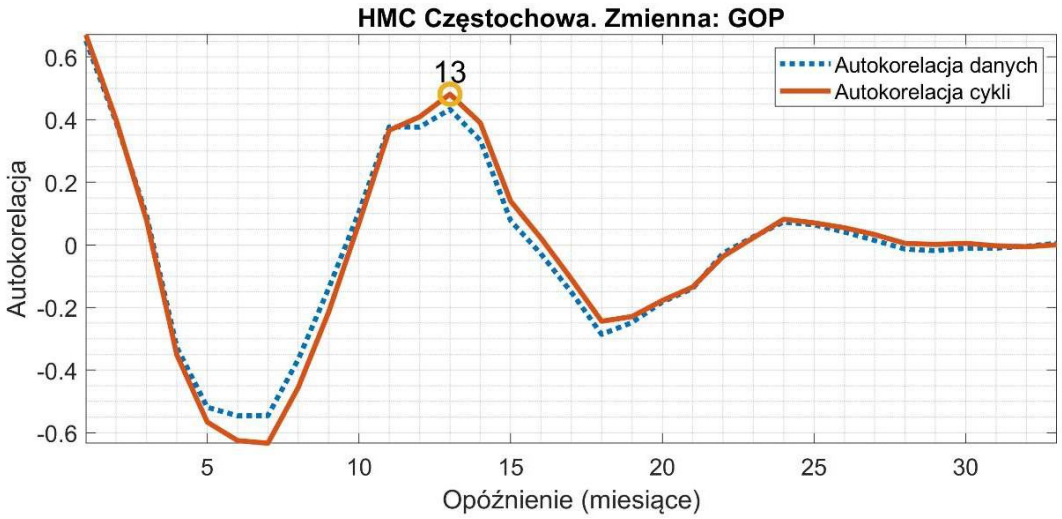
Rys. 4.18 Wykres autokorelacji standaryzowanych wartości zmiennej GOP dla HDSP Poznań. Źródło: Opracowanie własne.

W przypadku danych dla hotelu HM Opole wyznaczone krzywe wskazują na występowanie słabych składowych cykli 8– i 20–miesięcznych, silnej składowej rocznej i słabszej składowej dwuletniej.



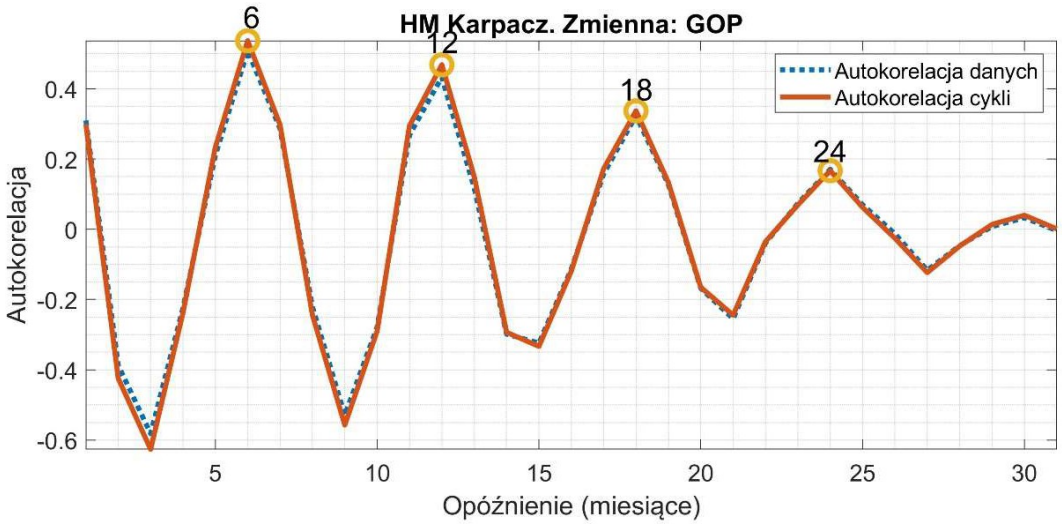
Rys. 4.19 Wykres autokorelacji standaryzowanych wartości zmiennej GOP dla HM Opole. Źródło: Opracowanie własne.

W przebiegach krzywych autokorelacji zmiennej GOP dla hotelu HMC Częstochowa algorytm wyznaczył jedynie jedną – 13–to miesięczną składową cykliczną. Krzywe danych i cykli prawie się pokrywają.



Rys. 4.20 Wykres autokorelacji standaryzowanych wartości zmiennej GOP dla HMC Częstochowa. Źródło: Opracowanie własne.

Krzywe autokorelacji obliczone dla zmiennej GOP zarejestrowanej dla hotelu HM Karpacz charakteryzują się podobnym przebiegiem jak w przypadku HDS Piaseczno, tj. posiadają cykle półroczne i roczne, z tą różnicą, że tutaj, ze względu na ograniczoną liczbę danych, możliwe było wyznaczenie cykli maksymalnie do 2 lat. Krzywe danych i cykli idealnie się pokrywają.



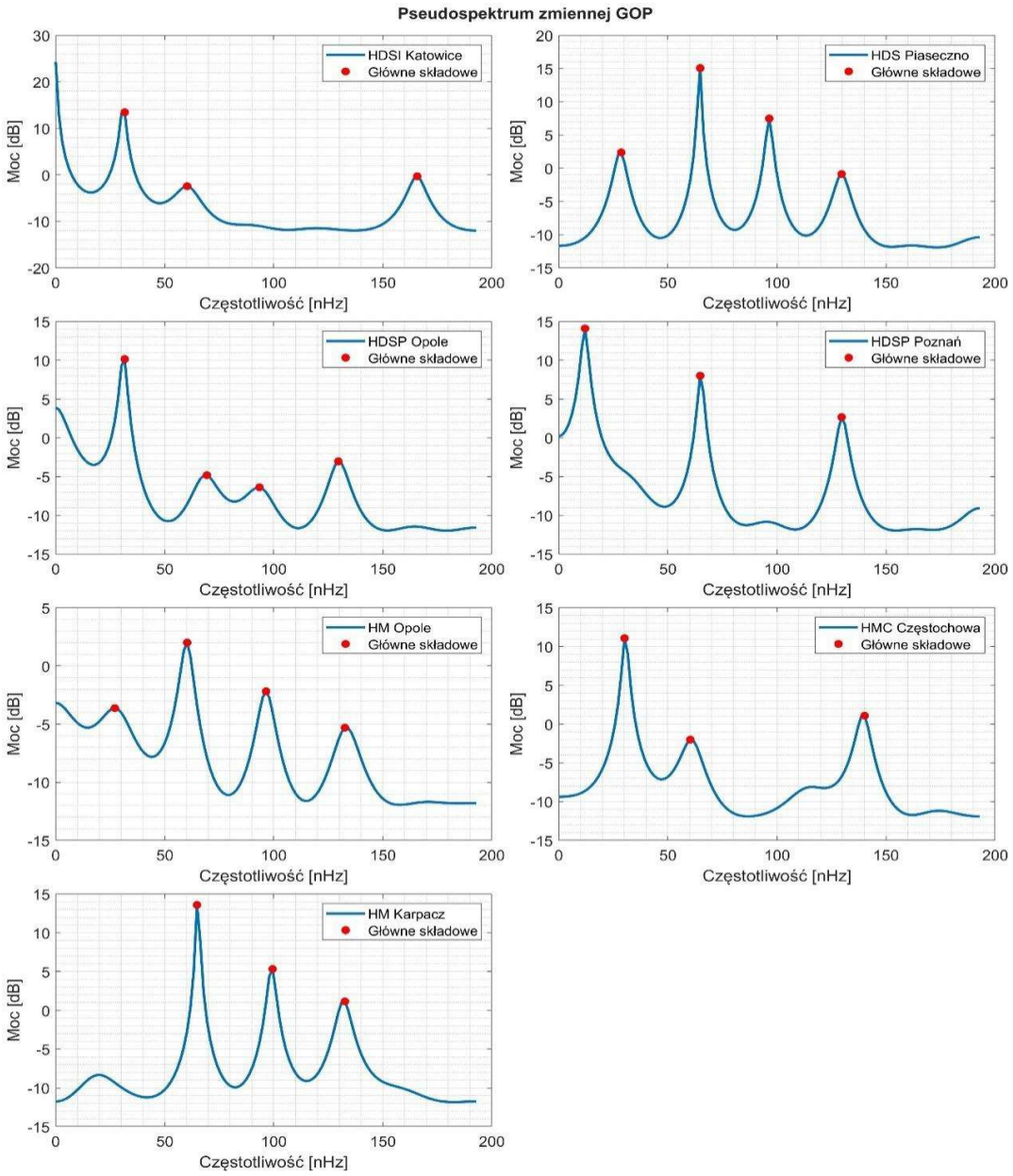
Rys. 4.21 Wykres autokorelacji standaryzowanych wartości zmiennej GOP dla HM Karpacz. Źródło: Opracowanie własne.

4.4. Analiza zmienności zysku operacyjnego brutto w dziedzinie częstotliwości

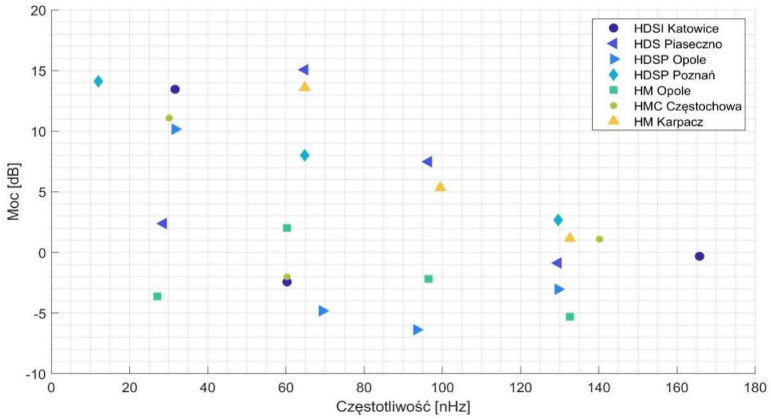
Analiza rozpatrywanych danych uwzględniła również obliczenie składowych częstotliwościowych zawartych w danych. Analiza ta, podobnie jak metoda autokorelacji, pozwala na wskazanie składników sezonowych. W badaniach zastosowano metodę PMUSIC (*ang. Pisarenko's Method of Signal Classification*), która jest popularną techniką używaną do estymacji widma sygnału (częstotliwości) oraz do identyfikacji jego składników

harmonicznych. Z wykresów pseudo spektrogramu wykonanych metodą PMUSIC można odczytać lokalizację głównych składowych częstotliwości występujących w danych. Każda wykryta częstotliwość będzie reprezentowana jako "szczyt" lub "pik" na wykresie, którego wysokość (amplituda) wskazuje, jak silna jest ta częstotliwość w szeregu czasowym. Metoda PMUSIC pozwala też na precyzyjną estymację częstotliwości, co umożliwia rozróżnienie blisko leżących częstotliwości.

Na rys. 4.22 przedstawiono wyniki analizy składowych częstotliwościowych zawartych w widmie szeregu czasowego zmiennej GOP, na kolejnych krzywych, zaznaczonych kolorem niebieskim – dla każdego z badanych hoteli. W czasie obliczeń algorytm PMUSIC wskazywał do ośmiu głównych składowych. Na wykresach kropką w kolorze czerwonym zaznaczono te składowe, których moc przekraczała wartość średniej arytmetycznej całego widma. Analiza częstotliwości występujących w widmie prążków pozwala na wskazanie podobieństw w cykliczności zmiennej GOP rejestrowanej w poszczególnych hotelach. Z wykresu wynika, że w każdym z widm występują co najmniej 3 składowe główne, a w trzech widmach – cztery. Celem porównania cykliczności występujących w zmiennej GOP poszczególnych hoteli, na rys. 4.23 przedstawiono zbiorcze zestawienie głównych składowych występujących w widmach. Na wykresie figurami geometrycznymi w różnych kolorach zaznaczono moc głównych składowych w funkcji częstotliwości, która koresponduje z cyklicznością. Na podstawie przedstawionego zestawienia zbiorczego można wywnioskować, że składowe widm GOP kolejnych hoteli różnicują się pod kątem mocy, jednak nie są zróżnicowane pod kątem częstotliwości, a jedynie w przypadku HDSI Katowice zaobserwowano występowanie wartości przekraczających 165 nHz (ok. 70 dni), co wynika pośrednio również z największego zestawu dostępnych danych. Ciekawym wydaje się brak w widmie GOP dla HM Karpacz składowych poniżej 60 nHz (powyżej 6 miesięcy), podczas gdy w pozostałych hotelach one występują. Dodatkowo ciekawym wydaje się występowanie jedynie w widmie HDSP Poznań składowej poniżej 20 nHz (ok. 6 kwartałów).



Rys. 4.22 Wyniki analizy widmowej wykorzystującej metodę PMUSIC. Źródło: Opracowanie własne.



Rys. 4.23 Zestawienie zbiorcze głównych składowych widma zmiennej GOP dla wszystkich hoteli. Źródło: Opracowanie własne.

5. Ocena współliniowości zysku operacyjnego brutto podstawą wyboru regresorów do modelu predykcyjnego

W rozdziale 5, przeprowadzono proces badań mający na celu dobór zmiennych diagnostycznych (regresorów) wykorzystanych na kolejnym etapie badań, a więc na etapie budowy modeli predykcyjnych. W tym celu przeprowadzono ocenę współliniowości zysku operacyjnego brutto ze zmiennymi USALI oraz współliniowości zmiennych USALI między sobą. Zbadano również czy regresory wymagają przekształceń, czy można je wykorzystać w postaci nieprzekształconej – badając zgodność rozkładu badanych cech z rozkładem normalnym oraz badając własność stacjonarności i heteroskedastyczności. Na koniec tego etapu badań dodatkowo przeprowadzono analizę koherencji falkowej, w celu zidentyfikowania synchronizacji w czasie.

5.1. Analiza korelacji liniowej

Po przeprowadzeniu wstępnej analizy statystycznej zysku operacyjnego brutto oraz wyciągnięciu użytecznych wniosków badawczych w zjawiskach ekonomicznych, w następnej kroku wykonano analizę korelacyjną między zmienną GOP, a pozostałymi zmiennymi, pozyskanymi z bazy danych USALI. W pracy obliczono wartości współczynnika Pearsona, który jest miarą statystyczną używaną do określenia stopnia związku (korelacji) między dwiema zmiennymi numerycznymi. Korelacja Pearsona określa, czy zmienne są ze sobą liniowo powiązane i w jakim stopniu. Wartość współczynnika mieści się w zakresie od -1 do 1, gdzie: r=1 oznacza doskonałą korelację dodatnią, czyli zmienne wzrastają razem, r=-1 oznacza doskonałą korelację ujemną, czyli jedna zmienna rośnie, gdy druga maleje, r=0 oznacza brak korelacji, czyli zmienne nie mają liniowej zależności między sobą. Wzór na obliczenie współczynnika korelacji Pearsona między dwiema zmiennymi X i Y dla n obserwacji jest następujący:

$$r = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2 \sum_{i=1}^{n}(Y_i - \bar{Y})^2}}$$

gdzie: X_i i Y_i –wartości zmiennych X i Y dla i–tej obserwacji, $\bar{X}$ i $\bar{Y}$ –średnie wartości zmiennych X i Y, n –liczba obserwacji.

Dla zmiennych USALI, które nie wykazały rozkładu normalnego zgodnie z testem Jarque'a-Berry, zastosowan współczynnik korelacji Spearmana. Współczynnik ten, w odróżnieniu od współczynnika korelacji Pearsona, opiera się na rangach zamiast na wartościach rzeczywistych,

co czyni go bardziej odpornym na wpływ obserwacji odstających. Jest szczególnie użyteczny w przypadku zmiennych porządkowych lub danych niespełniających założeń normalności.

Współczynnik Spearmana r_s oblicza się według wzoru:

$$r_s = 1 - \frac{6\sum d_i^2}{n(n^2-1)}$$

gdzie: d_i – różnica pomiędzy rangami odpowiadających sobie obserwacji dwóch zmiennych, n –liczba obserwacji.

Należy mieć na uwadze, że wysoki współczynnik korelacji nie oznacza przyczynowości między zmiennymi; jedynie sugeruje, że istnieje liniowy związek między nimi.

W tabeli 3 przedstawiono wyniki obliczeń analizy korelacyjnej. Występowanie statystycznie istotnej (p<0,05) korelacji r powyżej 0,4 zaznaczono literą ‘P’ oraz kolorem jasnozielonym. Brak korelacji lub jej istotności oznaczono literą ‘F’ oraz kolorem białym. W dalszej części prac badawczych wykonano analizę korelacji między zmiennymi USALI skorelowanymi z GOP, celem ich późniejszej eliminacji w modelu ekonometrycznym.

Zmienne oznaczamy liczbą porządkową. Przykładowo, oznaczenie X5 odnosi się do cechy nr 5, tj. OCC.

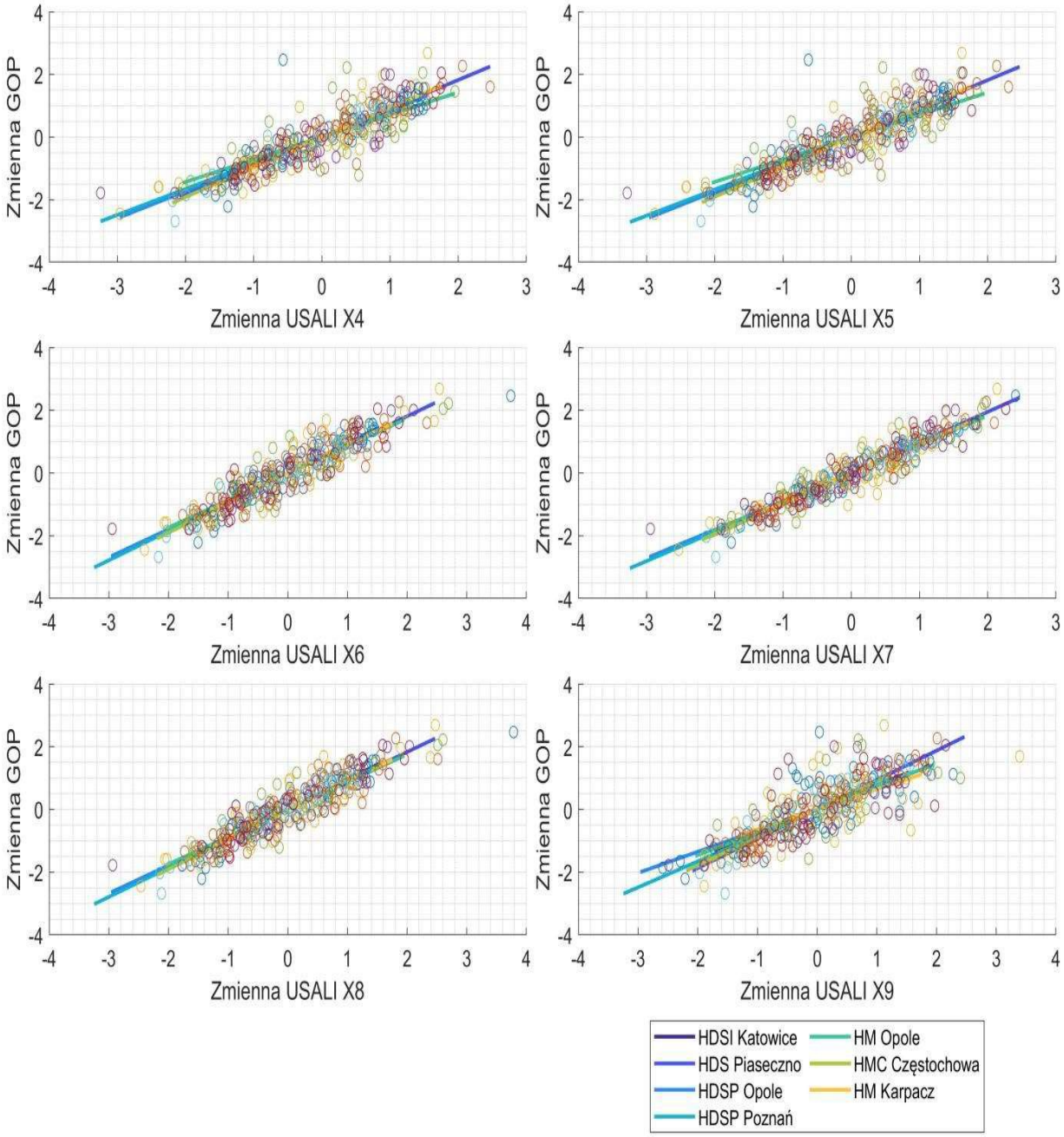
Tabela 3. Wyniki obliczeń zależności korelacyjnej zmiennej GOP do pozostałych zmiennych tablicy USALI dla wszystkich hoteli. Źródło: Opracowanie własne.

l.p.	ZMIENNA USALI	GOP						
		KAT	PIA	OP DS	POZ	OP HM	CZE	KAR
X1	OKRES	F	F	P	F	F	F	F
X2	ADR	F	F	P	F	P	F	P
X3	ZDOL	F	F	P	F	F	F	F
X4	SPRZ	P	P	P	P	P	P	P
X5	OCC	P	P	P	P	P	P	P
X6	REV	P	P	P	P	P	P	P
X7	PRZYCHODY OPERACYJNE:	P	P	P	P	P	P	P
X8	**Pokoje**	P	P	P	P	P	P	P
X9	**Gastronomia**, w tym:	P	P	P	P	P	P	P
X10	śniadania pakietowe	P	P	P	P	P	P	P
X11	kolacje pakietowe	F	P	F	F	F	F	P
X12	żywność	F	P	F	P	P	P	F
X13	napoje	F	P	P	P	P	P	F
X14	żywność bankiety	P	P	F	P	F	F	F
X15	napoje bankiety	F	P	F	P	F	F	F
X16	wynajem sal	F	P	F	P	P	F	F
X17	**Pozostałe Centra**, w tym:	P	F	P	P	F	F	P
X18	parking/garaż	P	F	P	P	P	P	P
X19	sklep DeSilva	P	P	F	P	F	F	F
X20	pranie dla gości	P	F	F	F	F	P	F
X21	transport gości	F	F	F	F	F	F	F
X22	inne	F	F	F	F	F	F	F

l.p.	ZMIENNA USALI	GOP						
		KAT	PIA	OP DS	POZ	OP HM	CZE	KAR
X23	KOSZT OPERACJNE:	P	P	P	P	F	P	P
X24	**Koszty Operacyjne – Pokoje**, w tym:	P	P	P	P	F	P	F
X25	koszty osobowe	F	F	P	F	F	F	F
X26	zużycie materiałów, w tym:	F	F	F	P	F	P	F
X27	materiały eksploatacyjne i artykuły spożywcze	F	P	F	P	F	P	P
X28	kosmetyki i środki czystości	F	F	F	P	F	P	F
X29	bielizna pościelowa, ręczniki, szlafroki, inne	F	F	F	F	F	P	F
X30	usługi obce, w tym:	F	P	F	P	F	P	F
X31	usługi sprzątania	F	P	F	F	F	F	F
X32	usługi prania	F	P	F	P	P	P	P
X33	usługi prania odzieży służbowej	F	F	F	F	F	F	F
X34	wynajem sprzętu	F	F	F	F	F	F	F
X35	inne usługi	F	F	F	F	F	F	F
X36	pozostałe koszty	F	F	F	F	F	F	F
X37	**Koszty Operacyjne Gastronomia**, w tym:	P	P	P	P	F	P	P
X38	koszty surowca, w tym:	P	P	P	P	P	P	P
X39	koszty surowca żywność	P	P	P	P	P	P	P
X40	koszty surowca napoje	F	P	F	P	P	F	F
X41	zastawa stołowa, szkło, naczynia gastronomiczne	F	F	F	F	F	F	F
X42	drobne wyposażenie nisko–cenne	F	F	P	F	F	F	P
X43	bielizna gastronomiczna	F	F	F	F	F	F	F
X44	karty dań, rachunki kelnerskie	F	F	F	F	F	F	F
X45	środki czystości	F	F	F	F	F	F	F
X46	inne materiały (serwetki, podkłady, Tea–light, opakowania), artykuły biurowe	F	F	F	F	F	F	F
X47	usługi sprzątania	F	F	F	F	F	F	F
X48	pranie bielizny gastronomicznej	F	F	F	F	F	F	F
X49	wynajem lokali użytkowych i sprzętu	F	F	F	F	F	F	F
X50	inne usługi (rozrywkowe, obsługa gastronomiczna)	F	F	F	F	F	F	F
X51	**Koszty Pozostałe Centra**	P	F	F	F	F	F	P
X52	WYNIK OPERACYJNY	P	P	P	P	P	P	P
X53	**Gastronomia**	P	P	P	P	P	P	P
X54	**Pozostałe centra**	P	F	F	F	F	F	F
X55	KOSZTY NIEPODZIELONE	F	F	F	F	F	F	F
X56	**Administracja i Dyrekcja**, w tym:	F	F	F	F	F	F	F
X57	zużycie materiałów	F	F	F	F	F	P	F
X58	kontrakty na usługi ochrony	P	F	F	F	F	F	P
X59	opłaty RTV, prawa autorskie, abonamenty	F	F	F	P	F	F	F
X60	wynajem sprzętu biurowego	F	F	F	F	F	F	F
X61	opłaty telekomunikacyjne	F	F	F	F	F	F	F
X62	opłaty bankowe	F	F	F	F	F	F	F

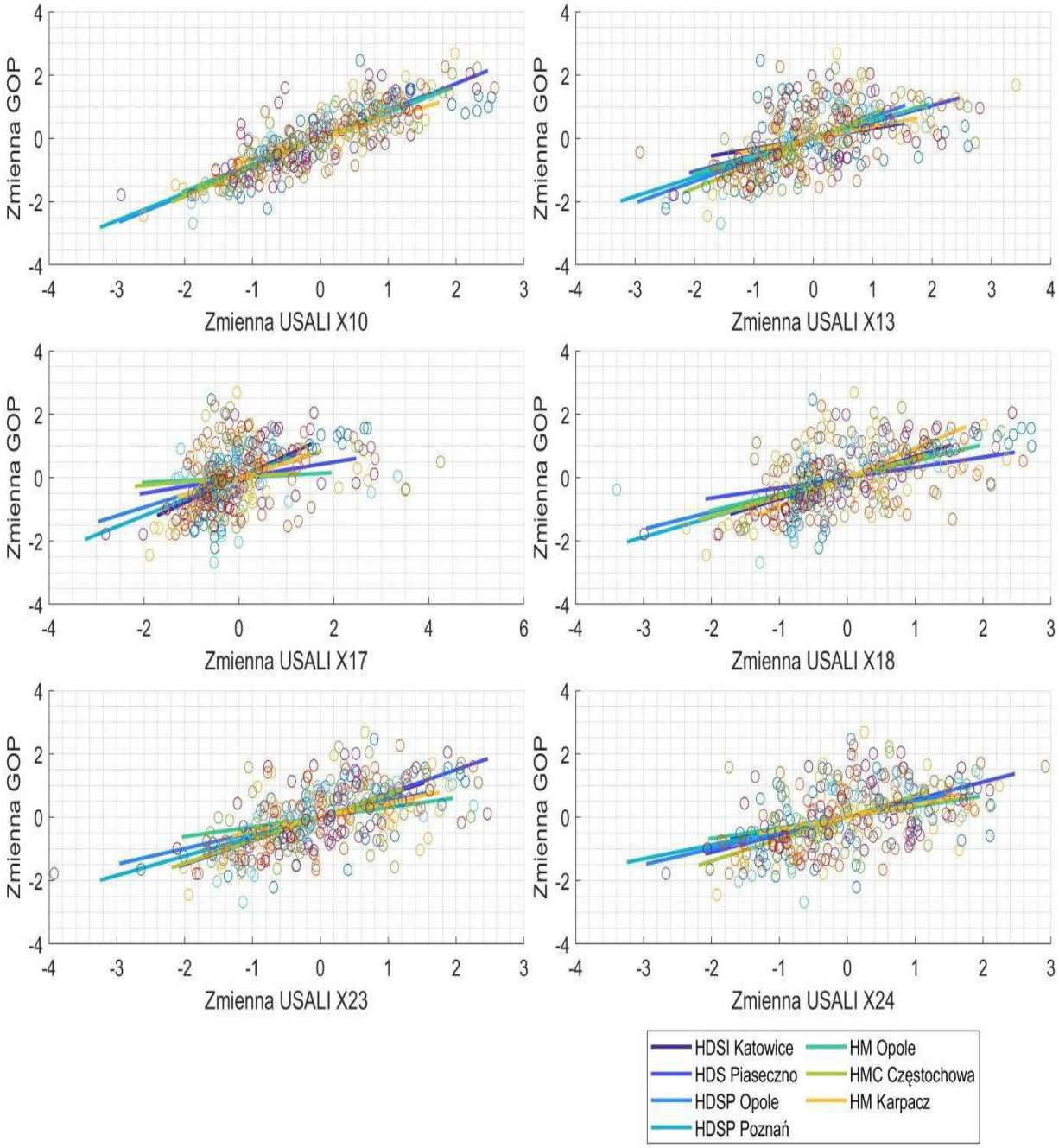
l.p.	ZMIENNA USALI	GOP						
		KAT	PIA	OP DS	POZ	OP HM	CZE	KAR
X63	prowizje kart kredytowych	F	F	F	F	F	F	F
X64	pozostałe usługi (w tym biurowe, poczta i kurier, prawne, tłumacz, serwis informatyczny, rezerwa)	F	F	F	F	F	F	F
X65	pozostałe koszty, w tym:	F	F	F	F	F	F	F
X66	**Dział Techniczny**, w tym:	F	F	F	F	F	F	F
X67	materiały do konserwacji, napraw i eksploatacji	F	F	F	F	F	F	F
X68	środki czystości, sprzęt komputerowy, materiały biurowe, inne nisko–cenne	F	F	F	F	F	F	F
X69	wynajem sprzętu	F	F	F	F	F	F	F
X70	usługi IT	F	F	F	F	F	F	F
X71	wywóz nieczystości	F	F	F	F	F	F	F
X72	dezynsekcja i deratyzacja	F	F	F	F	F	F	F
X73	utrzymanie zieleni	F	F	F	F	F	F	F
X74	przeglądy, konserwacje	F	F	F	F	F	F	F
X75	**Media**, w tym:	F	F	F	F	F	F	F
X76	zużycie energii elektrycznej	F	F	F	F	F	F	F
X77	zużycie wody	P	F	F	F	F	P	P
X78	**Dział Sprzedaży**, w tym:	F	F	P	P	F	P	P
X79	zużycie materiałów reklamowych, w tym:	F	F	F	F	F	F	F
X80	zużycie materiałów biurowych i innych	F	F	F	F	F	F	F
X81	usługi reklamowe Internet	F	F	F	F	F	F	F
X82	usługi reklamowe inne	F	F	F	F	F	F	F
X83	dystrybucja: prowizje biur podróży	F	F	P	P	F	P	P
X84	delegacje, udział w targach	F	F	F	F	F	F	F
X85	inne usługi (w tym kurier, spedycja, tłumaczenia, YP, STR)	F	F	F	F	F	F	F
X86	koszty reprezentacji	F	F	F	F	F	F	F

W celu porównania poszczególnych hoteli względem siebie pod kątem zróżnicowania współzależności poszczególnych zmiennych USALI i zmiennej GOP, w ramach prac badawczych wykonano analizy rozrzutu. W szczególności zestawiono wykresy rozrzutu wybranych zmiennych USALI względem zmiennej GOP dla wszystkich hoteli. Na rys 5.1–5.7 zaprezentowano uzyskane wyniki.



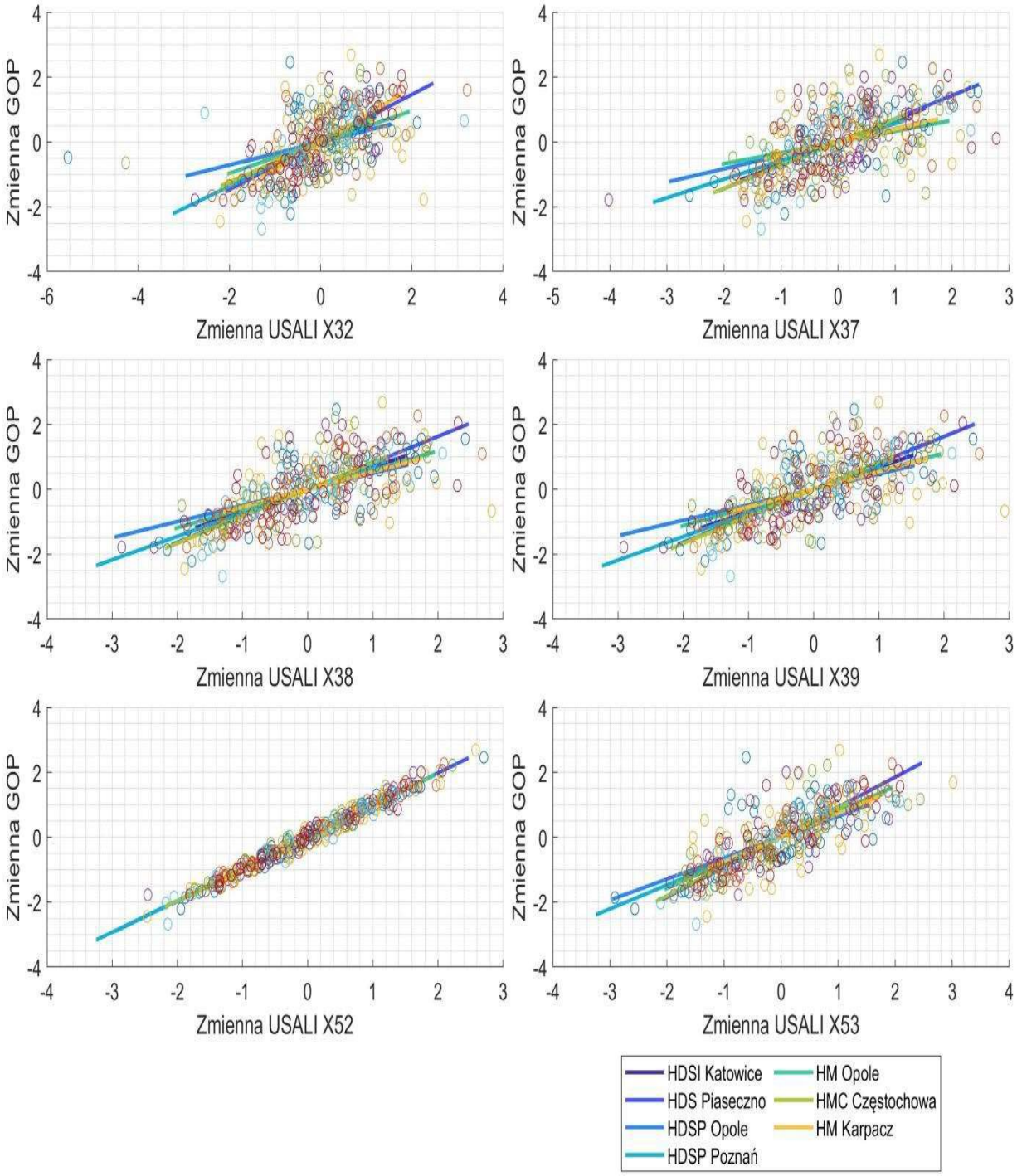
Rys. 5.1 Wizualizacja rozrzutu zmiennych X4 do X9 względem zmiennej GOP, pozyskanych dla wszystkich hoteli. Źródło: Opracowanie własne.

Wykresy rozrzutu pozwalają na wizualną weryfikację, że w danych występują zróżnicowania w danych pozyskanych dla poszczególnych hoteli. Dodatkowo widoczny jest kierunek korelacji między wizualizowanymi zmiennymi. Zależności zaprezentowane na rys. 5.1 nie wskazują na występowanie znaczących różnic między hotelami.



Rys. 5.2 Wizualizacja rozrzutu zmiennych X10, X13, X17, X18, X23 i X24 względem zmiennej GOP, pozyskanych dla wszystkich hoteli. Źródło: Opracowanie własne.

Zależności zaprezentowane na rys. 5.2 wskazują na występowanie pewnego zróżnicowania w stopniu skorelowania między porównywanymi hotelami, które objawia się w postaci stopnia nachylenia prostej, dla zmiennych X13, X17, X18, X23 i X24.



Rys. 5.3 Wizualizacja rozrzutu zmiennych X32, X37, X38, X39, X52 i X53 względem zmiennej GOP, pozyskanych dla wszystkich hoteli. Źródło: Opracowanie własne.

Zależności zaprezentowane na rys. 5.3 wskazują na występowanie pewnego zróżnicowania w stopniu skorelowania między porównywanymi hotelami, które objawia się w postaci stopnia nachylenia prostej, dla zmiennych X32, X37, X38, X39 i X53.

Zmienne USALI X6, X7, X8, X10 i X52 wskazują na identyczne zależności dla poszczególnych hoteli, co uwidacznia się w postaci prawie idealnie nałożonych na siebie prostych.

Histogramy wraz z dopasowaniem funkcji gęstości prawdopodobieństwa dla zmiennych

USALI pozostałych hoteli, skorelowanych ze zmienną GOP zgodnie z danymi przedstawionymi w Tabeli zostały zamieszczone w Załączniku Z1.

5.2. Wyznaczenie zmiennych USALI z uwzględnieniem własności współliniowości

W ramach prac badawczych, dla każdego hotelu, dla zmiennych USALI silnie skorelowanych z GOP, czyli takich dla których uzyskiwano wartość współczynnika *r* powyżej 0,8, obliczono ich korelacje wzajemne. Następnie te zmienne, które wzajemnie korelowały w sposób statystycznie istotny, przy dla $r > 0,4$, porównano oraz wskazano te, które słabiej korelowały ze zmienną GOP. Zmienne o słabszej korelacji względem GOP pominięto w dalszych rozważaniach i analizach. W tabeli 4 przedstawiono zestawienie zbiorcze wyselekcjonowanych, na podstawie powyżej opisanego kryterium, zmiennych USALI dla poszczególnych hoteli.

Tabela 4. Zestawienie zbiorcze wyselekcjonowanych zmiennych USALI dla poszczególnych hoteli. Źródło: Opracowanie własne.

HDSI Katowice	HDS Piaseczno	HDSP Opole	HDSP Poznań	HM Opole	HMC Częstochowa	HM Karpacz
X9	X11	X2	X2	X2	X11	X2
X14	X12	X3	X3	X4	X15	X11
X17	X13	X9	X11	X13	X17	X15
X20	X15	X13	X12	X14	X19	X16
X24	X16	X17	X15	X16	X20	X21
X51	X19	X18	X16	X29	X22	X25
X52	X23	X24	X21	X35	X25	X30
X58	X27	X42	X22	X36	X26	X34
X75	X31	X52	X25	X37	X28	X35
X77	X32	X53	X26	X47	X33	X36
	X40	X78	X28	X49	X34	X47
	X52		X30	X50	X35	X48
			X35		X48	X49
			X36		X49	X73
			X37		X60	X74
			X50		X73	X79
			X51		X74	
			X57		X79	
			X76			
			X81			

W tabelach 5 - 11 zestawiono dla poszczególnych hoteli wyniki obliczeń statystycznych opisujących zmienne najsilniej skorelowane z GOP, ale nieskorelowane wzajemnie.

W statystykach tych uwzględniono wartość skośności i kurtozy oraz wartości *p* testu Jarque'a-Berry, Dickey-Fullera i Breuscha-Pagana.

Test Breuscha-Pagana jest testem statystycznym służącym do wykrywania heteroskedastyczności w danych, czyli sytuacji, w której wariancja składnika losowego nie jest stała, a zależy od wartości zmiennej objaśniającej lub kombinacji zmiennych objaśniających. Jest to istotne w kontekście modeli regresji liniowej, gdzie jednym z kluczowych założeń jest homoskedastyczność, czyli równość wariancji reszt niezależnie od wartości zmiennej objaśniającej. Naruszenie tego założenia może prowadzić do nieskutecznych lub błędnie ocenionych wyników estymacji. W procedurze obliczeniowej zakłada się następujący model regresji liniowej:

$$y_i = \beta_0 + \beta_1 x_{1,i} + \beta_2 x_{2,i} + \ldots + \beta_k x_{k,i} + \varepsilon_i$$

gdzie ε_i — to składnik losowy o średniej zero, którego wariancja jest badana pod kątem stałości, β_k — współczynniki regresji, które mierzą wpływ zmiennych objaśniających (czyli niezależnych) na zmienną objaśnianą (zależną).

Test Breuscha-Pagana sprawdza, czy wariancja składnika losowego zależy liniowo od zmiennych objaśniających modelu regresji. A zatem najpierw estymuje się model regresji i oblicza reszty. Następnie wykonuje się regresję pomocniczą, badając, czy kwadraty reszt zależą od zmiennych objaśniających. Ostatecznie obliczona statystyka testowa (oparta na współczynniku determinacji z tej pomocniczej regresji) jest porównywana z wartością krytyczną rozkładu chi-kwadrat.

Statystyka Breuscha-Pagana opiera się na liczbie obserwacji oraz na współczynniku determinacji (R^2) regresji pomocniczej. Statystykę wylicza się ze wzoru:

$$BP = nR^2$$

gdzie: n — liczba obserwacji.

Jeśli statystyka jest większa niż wartość krytyczna lub gdy wartość $p < 0.05$, odrzucamy hipotezę zerową, co sugeruje heteroskedastyczność. W przeciwnym razie nie ma podstaw do odrzucenia hipotezy zerowej, co wskazuje na brak heteroskedastyczności.

Transformacji danych dokonywano, jeśli zmienna miała rozkład różny od normalnego lub była niestacjonarna. Wtedy dla szeregów o dodatniej skośności dokonywano transformacji polegającej na różnicowaniu, w tabelach oznaczone jako DIFF. Nie stosowano logarytmowania, ponieważ dane wejściowe były standaryzowane, więc zawierały wartości ujemne, co wyklucza

możliwość poddania ich logarytmowaniu. Dla szeregów o dodatniej skośności dane transformowano funkcją wykładniczą (EXP). Jeśli dane miały rozkład normalny i były stacjonarne nie poddawano ich transformacji, co w tabeli oznaczono znakiem -.

Tabela 5. Zestawienie zbiorcze miar statystycznych obliczonych dla wyselekcjonowanych zmiennych USALI w HDSI Katowice. Źródło: Opracowanie własne.

Zmienna USALI	HDSI Katowice					
	Skośność	Kurtoza	Test Jarque'a-Berry. Wartość *p*	Test Breuscha-Pagana. Wartość *p*	Test Dickey-Fullera. Wartość *p*	Trans-formacja
X9	-0,24	3,09	0,500	0,020	0,001	-
X14	0,29	2,26	0,162	0,719	0,001	-
X17	1,43	3,85	0,002	0,000	0,011	DIFF
X20	2,14	6,96	0,001	0,070	0,004	DIFF
X24	0,29	2,23	0,144	0,321	0,001	-
X51	1,55	5,82	0,001	0,407	0,001	DIFF
X52	0,15	2,51	0,500	0,000	0,001	-
X58	0,77	3,24	0,029	0,887	0,027	DIFF
X75	-0,43	6,54	0,001	0,172	0,001	EXP
X77	0,39	2,58	0,197	0,035	0,001	-

Na podstawie przeprowadzonej analizy, której wyniki zawarto w tabelach 5-11, można określić właściwości zmiennych USALI pod względem zgodności rozkładu z rozkładem normalnym oraz stacjonarności i homoskedastyczności dla poszczególnych hoteli. W przypadku hotelu HDSI Katowice, test Jarque'a-Berry wskazuje na normalność rozkładu dla zmiennych X9, X14, X24, X52 i X77 ($p > 0,05$), co sugeruje brak istotnych odchyleń od rozkładu normalnego. Z kolei zmienne X17, X20, X51, X58 i X75 wykazują istotne odchylenia od normalność ($p \leq 0,05$), co wymaga zastosowania odpowiednich transformacji.

Analiza homoskedastyczności przeprowadzona za pomocą testu Breuscha-Pagana sugeruje, że zmienne X14, X20, X24, X51, X58 oraz X75 są homoskedastyczne ($p > 0,05$), co oznacza stałość wariancji w modelu. Natomiast zmienne X9, X17, X52 oraz X77 wykazują heteroskedastyczność ($p \leq 0,05$), co wskazuje na zróżnicowanie wariancji i może wymagać uwzględnienia modyfikacji w analizie. Wszystkie zmienne w analizie spełniają warunek stacjonarności zgodnie z wynikami testu Dickey-Fullera ($p \leq 0,05$).

Dla zmiennych X17, X20, X51, X58 i X75 zastosowano transformacje, takie jak różnicowanie (oznaczone jako DIFF) lub przekształcenie wykładnicze (EXP), w celu spełnienia kryteriów stacjonarności oraz normalności rozkładu.

Tabela 6. Zestawienie zbiorcze miar statystycznych obliczonych dla wyselekcjonowanych zmiennych USALI w HDS Piaseczno. Źródło: Opracowanie własne.

Zmienna USALI	HDS Piaseczno					
	Skośność	Kurtoza	Test Jarque'a-Berry. Wartość p	Test Breuscha-Pagana. Wartość p	Test Dickey-Fullera. Wartość p	Trans-formacja
X11	1,29	4,42	0,002	0,200	0,001	DIFF
X12	0,47	2,71	0,140	0,668	0,001	-
X13	0,08	3,24	0,500	0,343	0,001	-
X15	0,65	3,00	0,052	0,294	0,001	-
X16	0,66	3,57	0,039	0,110	0,001	DIFF
X19	0,39	3,41	0,201	0,027	0,001	-
X23	0,21	2,38	0,324	0,205	0,001	-
X27	0,43	2,56	0,141	0,661	0,001	-
X31	0,17	2,18	0,195	0,016	0,001	-
X32	0,23	3,54	0,371	0,114	0,001	-
X40	0,30	2,74	0,428	0,959	0,001	-
X52	0,08	1,86	0,079	0,026	0,001	-

W przypadku hotelu HDS Piaseczno, test Jarque'a-Berry wskazuje na normalność rozkładu dla wszystkich rozpatrywanych zmiennych USALI za wyjątkiem X11 i X16, które wymagały zastosowania odpowiednich transformacji, tutaj różnicowania (DIFF).

Analiza homoskedastyczności przeprowadzona za pomocą testu Breuscha-Pagana sugeruje, że zmienne X19, X31 i X52 są heteroskedastyczne, a pozostałe są homoskedastyczne. Również w tym hotelu, wszystkie zmienne USALI spełniają warunek stacjonarności zgodnie z wynikami testu Dickey-Fullera.

Tabela 7. Zestawienie zbiorcze miar statystycznych obliczonych dla wyselekcjonowanych zmiennych USALI w HDSP Opole. Źródło: Opracowanie własne.

Zmienna USALI	HDSP Opole					
	Skośność	Kurtoza	Test Jarque'a-Berry. Wartość p	Test Breuscha-Pagana. Wartość p	Test Dickey-Fullera. Wartość p	Trans-formacja
X2	0,85	4,76	0,007	0,786	0,001	DIFF
X3	-1,87	4,60	0,001	0,000	0,018	EXP
X9	0,37	3,92	0,092	0,943	0,001	-
X13	0,89	4,38	0,010	0,363	0,001	DIFF
X17	0,57	3,70	0,061	0,218	0,001	-
X18	0,00	3,00	0,500	0,808	0,001	-
X24	0,31	2,71	0,445	0,682	0,005	-
X42	-0,26	5,06	0,012	0,921	0,001	EXP
X52	-0,07	3,15	0,500	0,765	0,001	-
X53	0,35	3,00	0,442	0,964	0,001	-
X78	-0,18	2,73	0,500	0,062	0,002	-

W przypadku hotelu HDSP Opole, dla czterech zmiennych USALI należało odrzucić

hipotezę zerową testu Jarque'a-Berry, mianowicie X2, X3, X13 i X42, oraz poddać je odpowiednim transformacjom. Ponadto, wszystkie zmienne USALI są stacjonarne, a jedynie zmienna X3 charakteryzuje się brakiem homoskedastyczności.

Tabela 8. Zestawienie zbiorcze miar statystycznych obliczonych dla wyselekcjonowanych zmiennych USALI w HDSP Poznań. Źródło: Opracowanie własne.

Zmienna USALI	Skośność	Kurtoza	HDSP Poznań			
			Test Jarque'a-Berry. Wartość p	Test Breuscha-Pagana. Wartość p	Test Dickey-Fullera. Wartość p	Trans-formacja
X2	-1,18	4,51	0,010	0,033	0,001	EXP
X3	-1,19	4,62	0,009	0,035	0,001	EXP
X11	-0,16	4,17	0,160	0,369	0,001	-
X12	0,40	2,27	0,230	0,242	0,001	-
X15	-0,07	4,79	0,046	0,196	0,001	EXP
X16	-0,30	4,48	0,067	0,257	0,001	-
X21	-1,89	7,87	0,001	0,178	0,001	EXP
X22	-0,31	3,02	0,500	0,774	0,004	-
X25	0,31	2,64	0,500	0,466	0,001	-
X26	0,17	1,83	0,155	0,417	0,001	-
X28	0,79	3,86	0,046	0,004	0,002	DIFF
X30	-0,75	3,28	0,078	0,083	0,001	-
X35	-1,88	8,39	0,001	0,125	0,001	EXP
X36	-0,81	3,74	0,048	0,102	0,001	EXP
X37	-0,83	3,82	0,042	0,074	0,001	EXP
X50	-0,18	2,57	0,500	0,085	0,001	-
X51	0,07	1,94	0,245	0,388	0,001	-
X57	-1,44	4,06	0,007	0,002	0,004	EXP
X76	-0,67	6,01	0,005	0,133	0,001	EXP
X81	-0,31	4,76	0,041	0,295	0,001	EXP

W przypadku hotelu HDSP Poznań, dla którego uwzględniano największą liczbę zamiennych USALI, ponad połowa z nich miała rozkład niespełniający kryteriów normalności i była poddana odpowiednim transformacjom. Zmienne X2, X3, X28 i X57 miały charakter heteroskedastyczny. Natomiast z testu Dickey-Fullera wynika, że również tutaj wszystkie szeregi są stacjonarne.

Tabela 9. Zestawienie zbiorcze miar statystycznych obliczonych dla wyselekcjonowanych zmiennych USALI w HM Opole. Źródło: Opracowanie własne.

Zmienna USALI	HM Opole					
	Skośność	Kurtoza	Test Jarque'a-Berry. Wartość p	Test Breuscha-Pagana. Wartość p	Test Dickey-Fullera. Wartość p	Trans-formacja
X2	3,08	13,09	0,001	0,386	0,001	DIFF
X4	-0,16	2,29	0,500	0,483	0,005	-
X13	1,06	4,03	0,020	0,326	0,001	DIFF
X14	-0,36	2,33	0,328	0,726	0,001	-
X16	0,48	2,52	0,252	0,056	0,003	-
X29	-2,23	10,67	0,001	0,933	0,001	EXP
X35	0,19	2,24	0,453	0,438	0,001	-
X36	0,25	2,32	0,468	0,572	0,001	-
X37	0,06	2,05	0,346	0,332	0,001	-
X47	1,60	6,32	0,001	0,690	0,001	DIFF
X49	0,43	2,47	0,296	0,853	0,001	-
X50	0,43	3,16	0,416	0,990	0,001	-

W przypadku hotelu HM Opole, test Jarque'a-Berry wykazał brak normalności rozkładu zmiennych X2, X13, X29 i X47, które poddano odpowiednim transformacjom. Ponadto wszystkie zmienne były stacjonarne i homoskedastyczne.

Tabela 10. Zestawienie zbiorcze miar statystycznych obliczonych dla wyselekcjonowanych zmiennych USALI w HMC Częstochowa. Źródło: Opracowanie własne.

Zmienna USALI	HMC Częstochowa					
	Skośność	Kurtoza	Test Jarque'a-Berry. Wartość p	Test Breuscha-Pagana. Wartość p	Test Dickey-Fullera. Wartość p	Trans-formacja
X11	0,13	2,43	0,500	0,424	0,002	-
X15	-0,79	5,17	0,013	0,358	0,001	EXP
X17	0,84	2,41	0,051	0,771	0,011	-
X19	0,50	2,58	0,255	0,498	0,002	-
X20	0,40	2,40	0,320	0,551	0,001	-
X22	0,31	2,77	0,500	0,854	0,002	-
X25	1,41	5,17	0,004	0,959	0,001	DIFF
X26	1,00	3,74	0,029	0,024	0,010	DIFF
X28	0,25	4,87	0,039	0,109	0,001	DIFF
X33	0,50	2,58	0,245	0,468	0,002	-
X34	-0,05	2,10	0,402	0,986	0,006	-
X35	-0,03	2,18	0,500	0,924	0,005	-
X48	-0,82	2,79	0,064	0,847	0,010	-
X49	-0,13	2,62	0,500	0,937	0,006	-
X60	3,45	16,86	0,001	0,379	0,001	DIFF
X73	-0,42	3,78	0,211	0,297	0,001	-
X74	-0,56	3,21	0,209	0,512	0,001	-
X79	-0,08	3,32	0,500	0,262	0,001	-

W przypadku hotelu HMC Częstochowa, test Jarque'a-Berry wykazał brak normalności rozkładu zmiennych X15, X25, X26, X28 i X60, które poddano odpowiednim transformacjom. Ponownie wszystkie zmienne były stacjonarne wg testu Dickey-Fullera, a wg testu Breuscha-Pagana jedynie zmienna X26 charakteryzowała się heteroskedastycznością.

Tabela 11. Zestawienie zbiorcze miar statystycznych obliczonych dla wyselekcjonowanych zmiennych USALI w HM Karpacz. Źródło: Opracowanie własne.

Zmienna USALI	HM Karpacz					
	Skośność	Kurtoza	Test Jarque'a-Berry. Wartość p	Test Breuscha-Pagana. Wartość p	Test Dickey-Fullera. Wartość p	Trans-formacja
X2	0,17	2,50	0,500	0,368	0,009	-
X11	1,28	3,38	0,016	0,911	0,004	DIFF
X15	1,14	4,24	0,016	0,125	0,002	DIFF
X16	0,71	2,46	0,089	0,479	0,001	-
X21	0,74	3,48	0,081	0,030	0,001	-
X25	0,68	2,98	0,120	0,565	0,001	-
X30	0,38	1,95	0,141	0,493	0,001	-
X34	1,10	4,10	0,020	0,047	0,002	DIFF
X35	0,58	2,90	0,210	0,257	0,001	-
X36	0,54	2,75	0,246	0,197	0,001	-
X47	0,15	3,78	0,500	0,961	0,001	-
X48	0,47	2,12	0,148	0,929	0,001	-
X49	0,45	2,38	0,259	0,650	0,001	-
X73	0,66	2,18	0,082	0,333	0,001	-
X74	1,29	4,44	0,010	0,166	0,001	DIFF
X79	0,66	2,89	0,136	0,093	0,001	-

W przypadku hotelu HM Karpacz, test Jarque'a-Berry wykazał brak normalności rozkładu zmiennych X11, X15, X34 i X74, które poddano odpowiednim transformacjom. Ponownie wszystkie zmienne były stacjonarne wg testu Dickey-Fullera, a wg testu Breuscha-Pagana jedynie zmienne X21 i X34 charakteryzowały się heteroskedastycznością.

5.3. Analiza koherencji falkowej

Celem identyfikacji współzależności między zmiennymi, które stanowią szeregi czasowe w różnych skalach czasowych, zastosowano metodę koherencji falkowej (*ang. wavelet coherence*). W metodzie tej wykorzystuje się analizę widmową i transformację falkową (Ramsey, 1999; Ramsey & Lampart, 1998; Schleicher, 2002) do obliczenia współczynnika koherencji, który mierzy stopień współzależności między dwoma sygnałami w różnych częstotliwościach (Foufoula-Georgiou & Kumar, 1994; Grinsted, Moore, & Jevrejeva, 2004; Maraun, Kurths, & Holschneider, 2007; Torrence & Compo, 1998; Torrence & Webster, 1999). Koherencja jest miarą, która wskazuje, jak bardzo dwie zmienne są powiązane w dziedzinie

czasowo-częstotliwościowej. W trakcie procedury obliczania współczynnika koherencji, w pierwszej kolejności, szeregi czasowe poddaje się ciągłej transformacji falkowej. Zastosowanie falki Morleta'a o ω_0= 6 stanowi optymalny wybór, który zapewnia właściwą równowagę pomiędzy lokalizacją czasową i częstotliwościową (Kurths & Maraun, 2004):

$$\Psi(t) = \pi^{-1/4} e^{i\omega_0 t} e^{-1/2 t^2}$$

gdzie:

ω_0 – częstość kołowa,

t – czas.

Dla szeregu czasowego $x(t)$ otrzymujemy jego transformatę falkową $X(t,\tau)$:

$$X(t,\tau) = \int [x(t)^* \psi_x^*(t-\tau)]\, dt$$

Dla szeregu czasowego $y(t)$ otrzymujemy jego transformatę falkową $Y(t,\tau)$:

$$Y(t,\tau) = \int [y(t)^* \psi_y^*(t-\tau)]\, dt$$

Indeks górny * oznacza sprzężenie zespolone, t – czas, τ – skala.

Następnie obliczane jest widmo krzyżowe $C_{\mathrm{xy}}(a,b)$ między transformatami falkowymi $X(t,\tau)$ i $Y(t,\tau)$:

$$C_{xy}(t,\tau) = \frac{|X^*(t,\tau)\, Y(t,\tau)\,|}{X(t,\tau)^* X(t,\tau)^{**} Y(t,\tau)^* Y(t,\tau)^*}$$

Ostatecznie współczynnik koherencji falkowej $WC(t,\tau)$ wyznacza się w oparciu o korelację pomiędzy widmami szeregów $\mathrm{C_{xy}}(t,\tau)$, zgodnie z równaniem:

$$WC(t,\tau) = \frac{\left|C_{xy}(t,\tau)\right|^2}{X(t,\tau)^* X(t,\tau)^{**} Y(t,\tau)^* Y(t,\tau)^*}$$

Metoda koherencji falkowej oferuje kilka zalet. Po pierwsze umożliwia wykrywanie współzależności w różnych skalach czasowych, pozwalając na identyfikację zarówno krótkoterminowych, jak i długoterminowych wzorców współzależności między szeregami czasowymi, co jest szczególnie przydatne w przypadku analizy danych niestacjonarnych, gdzie współzależności mogą się zmieniać w zależności od skali czasowej. Kolejną zaletą metody jest możliwość przeprowadzenia lokalnej analizy czasowej, która pozwala na dokładne zbadanie

struktury danych w różnych momentach czasowych, zamiast ograniczać się tylko do ogólnej analizy na całym przedziale czasowym. Daje to bardziej szczegółowe informacje na temat zmian i dynamiki współzależności między szeregami czasowymi.

W ekonomii, koherencja falkowa pomaga zrozumieć, czy dwa wskaźniki ekonomiczne mają tendencję do oscylowania w podobnym rytmie, czy też działają niezależnie od siebie. Wyższa koherencja między dwoma wskaźnikami oznacza większe prawdopodobieństwo, że są one ze sobą powiązane, podczas gdy niższa koherencja wskazuje na większą niezależność tych wskaźników. Analiza koherencji falkowej pozwala również na badanie dynamiki tych powiązań w różnych okresach. Można zidentyfikować, czy relacje między wskaźnikami ekonomicznymi są stabilne i trwałe w długim okresie, czy też ulegają zmianom w zależności od konkretnych warunków ekonomicznych.

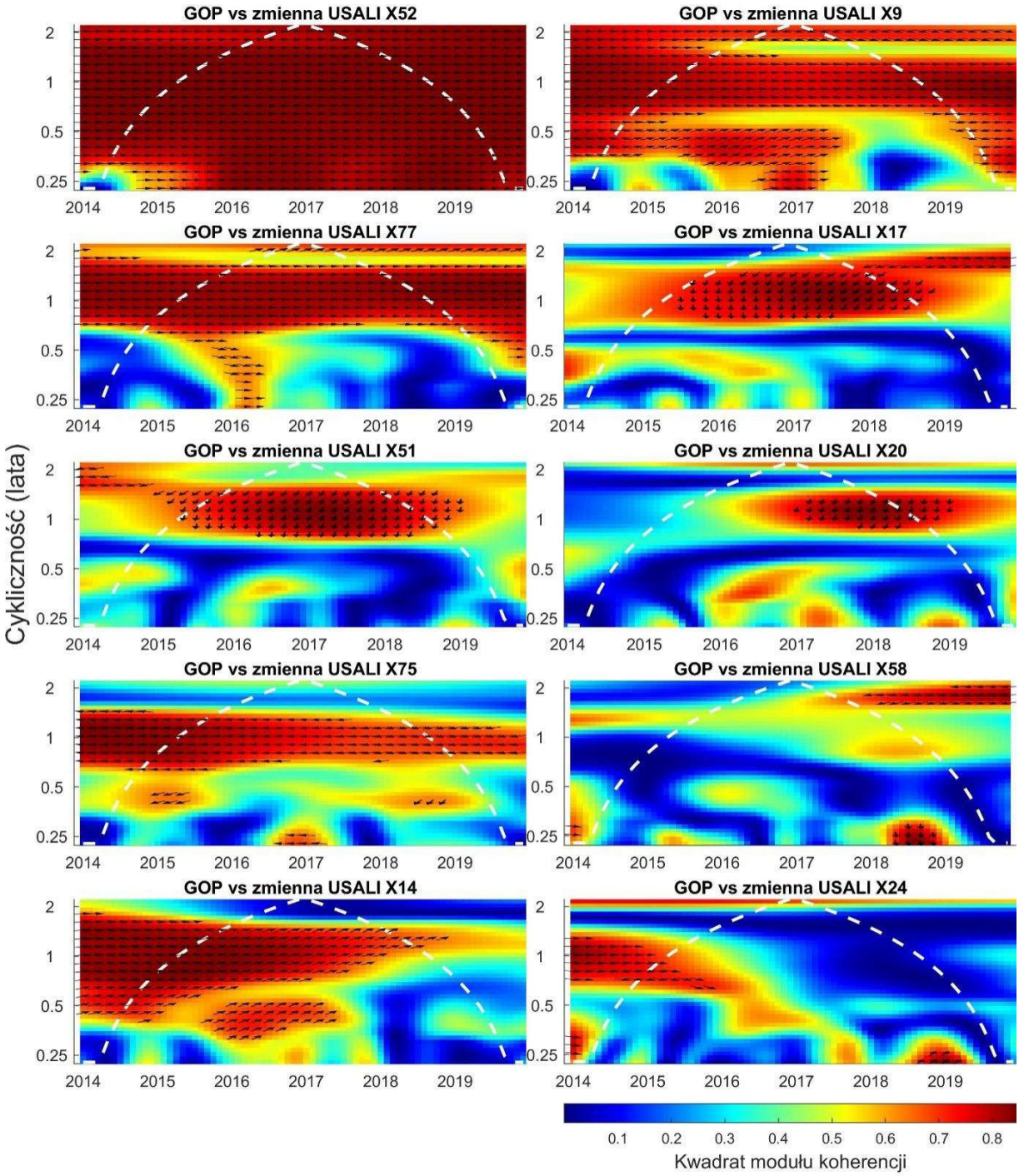
Wykorzystanie analizy koherencji falkowej w prezentowanych pracach badawczych miało na celu zidentyfikowanie synchronizacji w czasie, czyli określeniu fazy lub antyfazy pomiędzy badanymi zmiennymi oraz zidentyfikowanie wyprzedzeń lub opóźnień pomiędzy nimi. Przy wykorzystaniu tej analizy zastały obliczone kąty osi koherencji zmiennej GOP z cechami wykazującymi statystycznie istotną koherencję.

Na rys 5.4 przedstawiono dwuwymiarową wizualizację wartości kwadratu modułu koherencji falkowej $WC(t,\tau)$ między wartościami GOP i wyselekcjonowanymi, skorelowanymi z GOP zmiennymi USALI, zebranymi dla hotelu HDSI Katowice. Kod kolorystyczny dla mocy obejmuje zakres od niebieskiego do ciemno czerwonego. Oś Y mierzy sezonowość (opóźnienie, wyprzedzenie), a oś X przedstawia badany okres czasu. Obszary w kolorze czerwonym, identyfikują silną współzależność między dwoma szeregami czasowymi, a w kolorze niebieskim jej brak. Wartości kwadratu modułu koherencji mogą się zawierać w zakresie od 0 do 1, gdzie 0 oznacza brak współzależności, a 1 wskazuje na pełną współzależność między dwoma sygnałami w danej częstotliwości. Należy przy tym pamiętać, że analiza ta nie określa związku w rozumieniu przyczynowo skutkowym, pokazuje jedynie sam fakt występowania zależności między zmiennymi.

Ponadto na wykresach zobrazowano za pomocą strzałek fazową relację między dwoma szeregami czasowymi w różnych skalach czasowych. Strzałki wskazują kierunek i siłę fazy pomiędzy analizowanymi szeregami danych. Obecność strzałek odpowiada sile koherencji powyżej 0,7, a kierunek wskazuje na fazową relację. Skierowanie strzałki w dowolnym kierunku może wskazywać na synchronizację lub przesunięcie fazowe między szeregami danych w danym punkcie czasowym. Przesunięcie to można interpretować jako wyprzedzenie (jeśli dodatnie) lub opóźnienie (jeśli ujemne) reakcji jednej zmiennej względem drugiej. Kierunkowe strzałki wskazują zatem na relacje fazowe między zmiennymi. Strzałki skierowane

w prawo symbolizują, że zmienne są zsynchronizowane. Jeśli strzałka jest skierowana w górę, oznacza to, że zmienna USALI jest opóźniona w stosunku do zmiennej GOP. Natomiast strzałka skierowana w dół wskazuje, że zmienna USALI jest w fazie przewodzącej. Przeciwnie, strzałki skierowane w lewo oznaczają, że badane zmienne są w przeciwfazie. Wyrażenie "w fazie" oznacza, że zmienne oddziałują na siebie cyklicznie, natomiast "w przeciwfazie" oznacza, że zmienne oddziałują na siebie antycyklicznie.

Szczegółowa analiza wykresów koherencji falkowej, zawartych na rys. 5.4, które obliczono dla hotelu HDSI Katowice pozwala na stwierdzenie, że zmienna GOP jest w pełni skorelowana ze zmienną USALI X52 (*wynik operacyjny*) oraz w bardzo dużym zakresie ze zmienną X9 (*przychody operacyjne z departamentu gastronomia*); koherencja ta jest w fazie (strzałki skierowane w prawo). Natomiast zmienne X77 (*koszty niepodzielone − zużycie wody*), X51 (*koszty operacyjne pozostałych centrów*), X17 (*przychody operacyjne pozostałych centrów*) i X75 (*koszty niepodzielone − media*) charakteryzują się silną koherencją o cyklu rocznym, przy czym zmienna X77 jest w fazie z GOP, zmienne X17 i X51 wskazują na opóźnienia względem GOP, natomiast zmienna X75 jest w przeciwfazie do GOP. Analiza wykresu dla zmiennej X20 (*przychody operacyjne pozostałych centrów − pranie dla gości*) wykazuje silną, opóźnioną koherencję w cyklu rocznym jedynie w okresie po 2017 r. Natomiast zmienna X58 (*koszty niepodzielone − kontrakty na usługi ochrony*) wykazuje silną koherencję w cyklu kwartalnym jedynie w latach 2018 r. — 2019 r., jednak koherencja była opóźniona o ok. miesiąc (strzałki skierowane w dół). Zmienna X24 (*koszty operacyjne w departamencie pokoje*) podobnie jak zmienna X58 były również w koherencji z GOP w latach 2018 r. — 2019 r. oraz dodatkowo również wykazywała koherencję w fazie w cyklu rocznym do 2016 r. Zmienna USALI X14 (*przychody operacyjne z departamentu gastronomia − żywność bankiety*) charakteryzuje się koherencją w fazie lub delikatnie wyprzedzającą w cyklach rocznych oraz 6-cio miesięcznych w latach 2016 r. — 2017 r.



Rys. 5.4 Wizualizacja kwadratu modułu koherencji falkowej dla zmiennej GOP względem dziesięciu zmiennych USALI, pozyskanych w rozpatrywanym okresie działalności hotelu HDSI Katowice. Źródło: Opracowanie własne.

Zbiorcze zestawienia dla pozostałych hoteli zawarto w Załączniku Z2.

Podsumowując, w rozdziale 5 zaprezentowano proces selekcji zmiennych do modelu predykcyjnego GOP, w którym kluczową rolę odegrała analiza współliniowości oraz szczegółowe testy statystyczne. Na wstępie zidentyfikowano zmienne USALI silnie skorelowane z GOP, po czym usunięto te, które wykazywały wzajemną istotną korelację,

minimalizując w ten sposób ryzyko współliniowości. W dalszej części omówiono analizy statystyczne, zawierające wskaźniki skośności, kurtozy oraz wyniki testów Jarque'a-Berry, Dickey-Fullera i Breuscha-Pagana, które pozwoliły lepiej scharakteryzować wybrane zmienne. Dodatkowo, przy użyciu metody koherencji falkowej, określono współzależności czasowe pomiędzy cechami w różnych skalach, co wzbogaciło model o informacje o relacjach dynamicznych. Takie podejście zapewniło solidną podstawę dla modelowania GOP, opartego na optymalnie dobranych zmiennych oraz odpowiednim uwzględnieniu struktur czasowych danych.

6. Analiza zmienności predyktorów objaśniających zysk operacyjny brutto

W rozdziale 6 przeprowadzono badania mające na celu analizę zmienności zmiennych diagnostycznych (predyktorów), zidentyfikowanych w procesie doboru zmiennych i wybranych do dalszych analiz w rozdziale 5. Analiza zmienności tych zmiennych obejmowała w szczególności analizę wprowadzającą, polegającą na prezentacji oraz syntetycznym opisie zmienności szeregów w dziedzinie czasu, a następnie badanie i analizę periodyczności oraz analizę zmiennych objaśniających w dziedzinie częstotliwości. Analizie poddano zmienne dla wszystkich hoteli, jednak szczegółowy opis, zamieszczony w poniższych rozdziałach, dotyczy hotelu HDSI Katowice. Wyniki uzyskane dla pozostałych hoteli zostały zawarte w Załącznikach Z3–Z5.

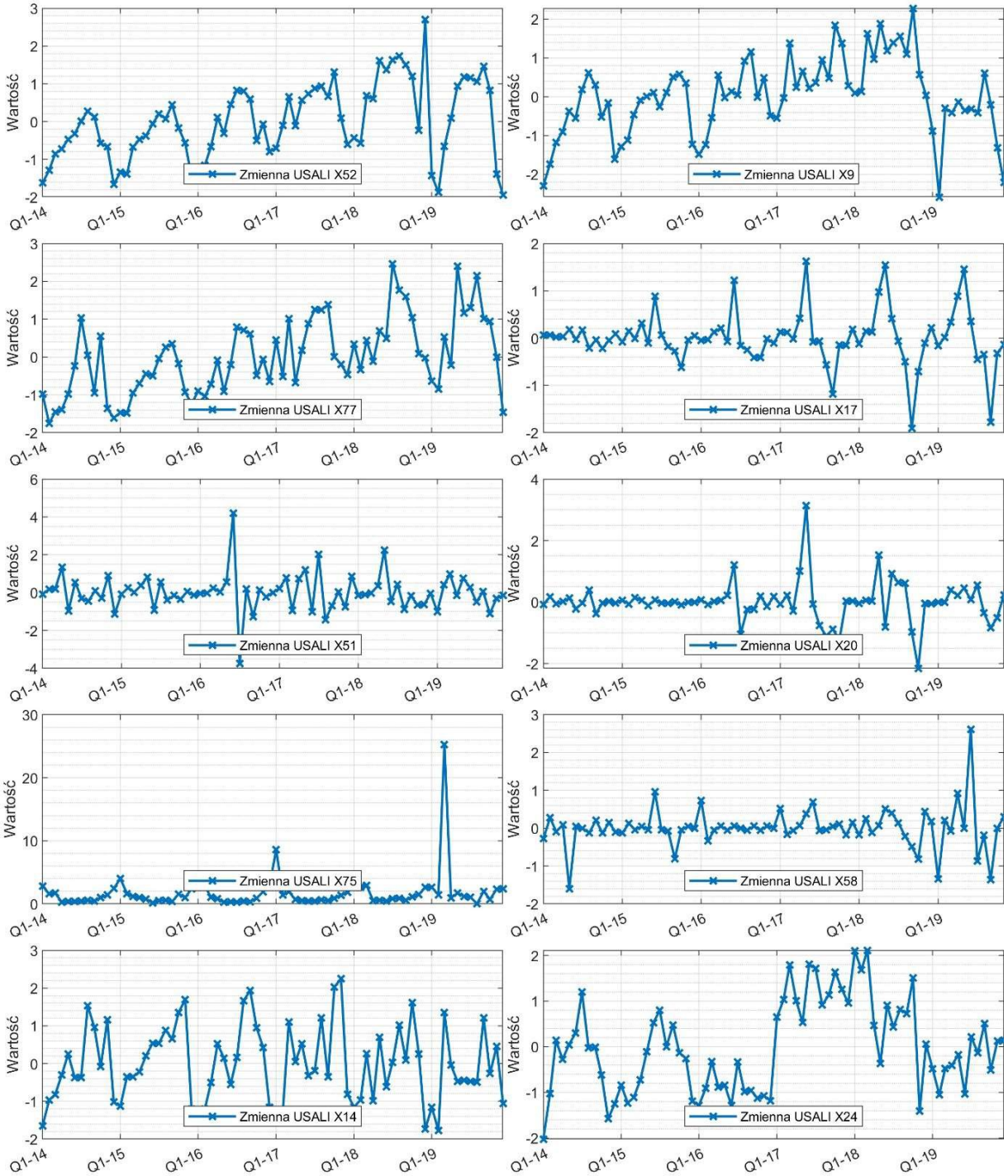
6.1. Analiza zmienności szeregów czasowych predyktorów

Na rysunku 6.1 przedstawiono przebiegi czasowe standaryzowanych wartości zmiennych skorelowanych ze zmienną GOP, tj. zmiennych wymienionych w tabeli nr 4, dla hotelu HDSI Katowice. Analiza krzywych zobrazowanych na wykresach pokazuje, jak bardzo dynamika zmian jest podobna do tej zaobserwowanej w zmiennej GOP.

Anormalny przyrost wartości w grudniu 2018 r., widoczny na wykresie zmiennej GOP, widoczny jest również na wykresie zmiennej X52 (*wynik operacyjny*). Natomiast dla tego samego miesiąca, w wykresie zmiennej X9 (*przychody operacyjne z departamentu gastronomia*) widoczny jest anormalny spadek wartości. Duży przyrost wartości pojawił się w przebiegu zmiennej X75 (*koszty niepodzielone − media*) w okolicy lutego 2019 r. oraz w połowie 2019 r. w przebiegu zmiennej X58 (*koszty niepodzielone − kontrakty na usługi ochrony*).

Ponadto na wykresach zmiennych X52 (*wynik operacyjny*), X9 (*przychody operacyjne z departamentu gastronomia*) i X77 (*koszty niepodzielone − zużycie wody*) można obserwować trendy wzrostowe oraz występowanie cykli. Cykle obserwowano również w przebiegach zmiennych X17 (*przychody operacyjne pozostałych centrów*), X14 (*przychody operacyjne z departamentu gastronomia −żywność bankiety*) i X24 (*koszty operacyjne w departamencie pokoje*), które jednak nie uwidaczniały stałego trendu wzrostowego.

Zbiorcze zestawienia przebiegów czasowych cech dla pozostałych hoteli zawarto w Załączniku Z3.



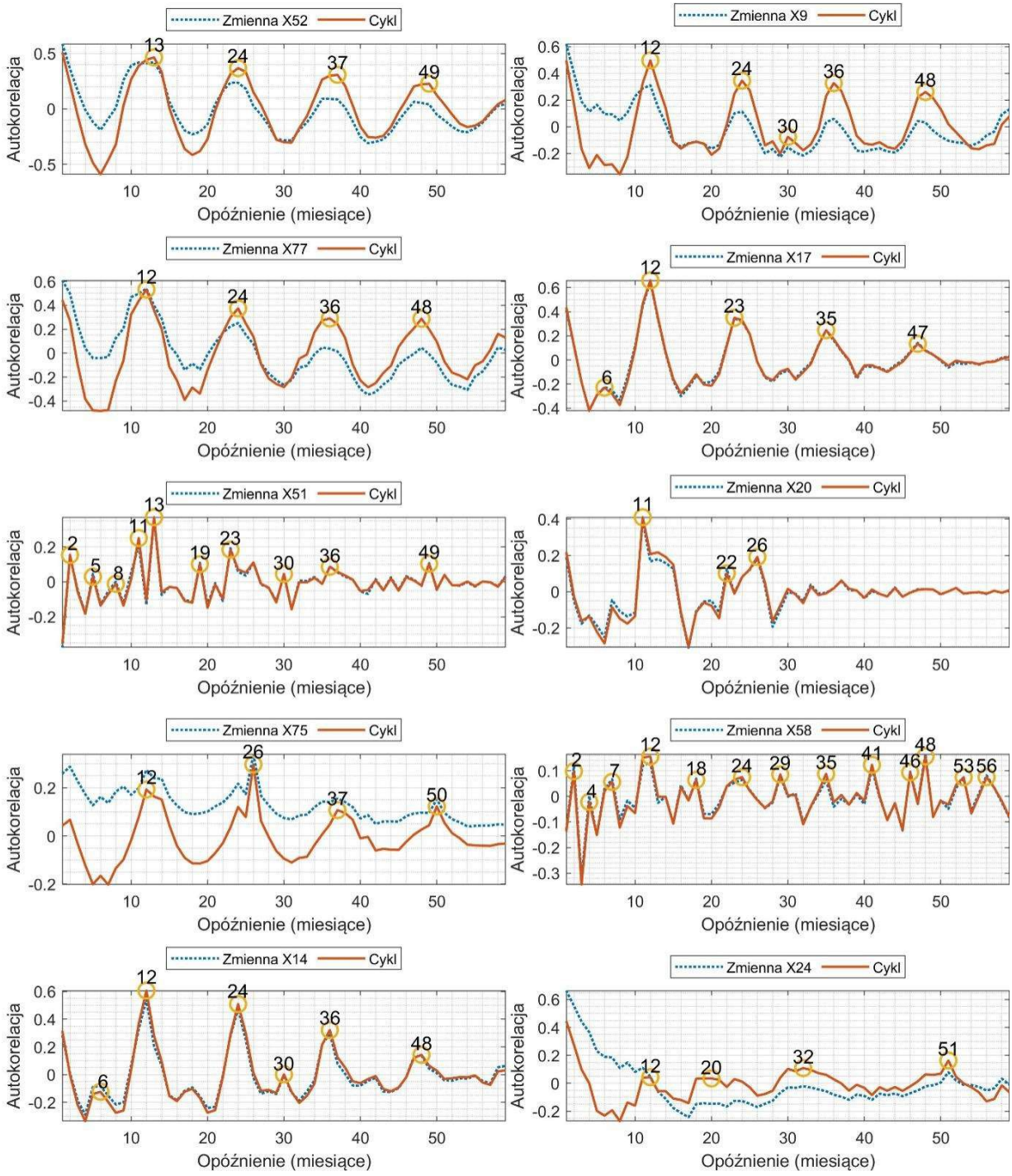
Rys. 6.1 Przebiegi czasowe standaryzowanych wartości zmiennych USALI skorelowanych z GOP dla HDSI Katowice. Źródło: Opracowanie własne.

6.2. Analiza periodyczności predyktorów

Wyniki analizy autokorelacji zmiennych USALI dla hoteli HDSI Katowice, przedstawiono na wykresach zgrupowanych na rys. 6.2. Interpretacja jest podobna do rysunków w rozdziału 4.3. Zdecydowana większość zmiennych wykazuje opóźnienia roczne, dwuletnie i trzyletnie.

Zmienne X51, X20, X24 i X58 wykazują chaotyczne cykle, a zmienne X51 i X58 dodatkowo słabe cykle półroczne i ich harmoniczne.

Zbiorcze zestawienia przebiegów czasowych zmiennych dla pozostałych hoteli zawarto w Załączniku Z4.



Rys. 6.2 Wykresy autokorelacji standaryzowanych wartości zmiennych USALI skorelowanych z GOP dla HDSI Katowice. Źródło: Opracowanie własne.

6.3. Analiza predyktorów w dziedzinie częstotliwości

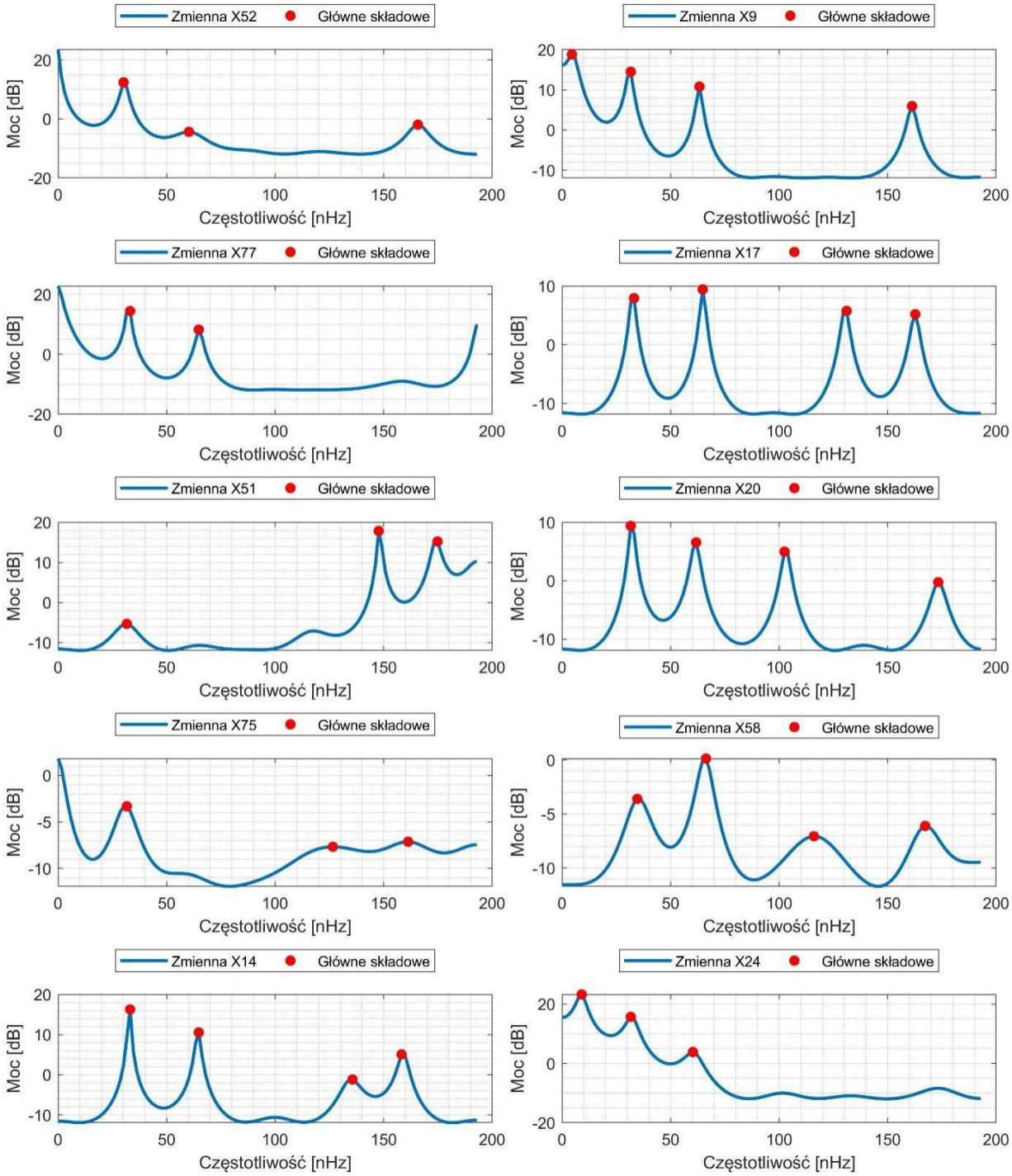
Wyniki analizy składowych częstotliwościowych zawartych w szeregach czasowych zmiennych, skorelowanych ze zmienną GOP zostały przedstawione na rys. 6.3. Interpretacja wyników analizy jest tożsama z tą, zawartą w rozdziale 4.4.

Ten typ analizy pozwala na pokazanie, że zmienna X52 (*wynik operacyjny*) i X77 (*koszty niepodzielone − zużycie wody*) charakteryzuje się bardzo podobnym składowymi częstotliwościowymi jak w widmie zmiennej GOP.

Zdecydowanie więcej silnych składowych częstotliwościowych zawierają widma zmiennych X9 (*przychody operacyjne z departamentu gastronomia*), X17 (*przychody operacyjne pozostałych centrów*) i X20 (*przychody operacyjne pozostałych centrów − pranie dla gości*). Natomiast zmienne X14 (*przychody operacyjne z departamentu gastronomia − żywność bankiety*) i X24 (*koszty operacyjne w departamencie pokoje*) charakteryzują się niskimi wartościami estymowanych składowych widma.

Zbiorcze zestawienia widm dla pozostałych hoteli zawarto w Załączniku Z5.

Rozdział ten koncentrował się na analizie dynamicznych cech predyktorów, kluczowych dla modelowania GOP. Na początku przeprowadzono analizę zmienności szeregów czasowych predyktorów, co pozwoliło na identyfikację ich charakterystyki w zakresie zmienności i stabilności. Następnie, wykorzystując metodę autokorelacji, oceniono periodyczność predyktorów, uwzględniając zarówno ogólne zależności czasowe, jak i autokorelację składników sezonowych, co pozwoliło na wykrycie sezonowych wzorców w danych Dodatkowo, poprzez analizę w dziedzinie częstotliwości za pomocą metody PMUSIC, zidentyfikowano dominujące komponenty częstotliwościowe w predyktorach, co umożliwiło pełniejsze scharakteryzowanie ich struktury czasowej. Wyniki tych analiz dostarczyły cennych informacji o regularnych wzorcach i dynamice predyktorów, które stanowią kluczowy element budowy skutecznego modelu predykcyjnego.



Rys. 6.3 Wyniki analizy widmowej standaryzowanych wartości zmiennych USALI skorelowanych z GOP obliczonych dla HDSI Katowice. Źródło: Opracowanie własne.

7. Opracowanie modelu predykcyjnego zysku operacyjnego brutto

7.1. Modelowanie zysku operacyjnego brutto z wykorzystaniem modelu ekonometrycznego

Wykonano badania zmierzające do opracowania modelu ekonometrycznego opisującego zmienną GOP na podstawie wszystkich cech USALI z nią skorelowanych, ale nieskorelowanych ze sobą, które zestawiono w sposób porównawczy w tabeli 4.

Ogólna formuła modelu ekonometrycznego – równania regresji linowej ma następującą postać:

$$\hat{y} = a + \sum_{i}^{C} b_i x_i$$

gdzie: a – przesunięcie, które jest interpretowane jako wartość oczekiwana zmiennej zależnej $\hat{y}$, gdy wszystkie zmienne niezależne x_i wynoszą zero, b_i – i–ta estymata wartości współczynników nachylenia prostej regresyjnej, powiązanych z kolejnymi cechami x_i, gdzie i $\in [1 - C]$, gdzie C – liczba cech uwzględnianych w regresji.

W procesie budowy modelu ekonometrycznego powszechną praktyką jest testowanie hipotez dotyczących parametrów modelu w celu oceny ich istotności statystycznej, co można przeprowadzić na przykład za pomocą testów t-Studenta. W ramach testu t każdy parametr modelu podlega osobnemu testowi, który porównuje wartość estymowaną parametru do zera. Przed przystąpieniem do testów ustala się poziom istotności, czyli wartość *p*, przy której uznajemy wynik za istotny. Najczęściej stosowane poziomy istotności to 0,05 i 0,01, co oznacza, że hipotezę zerową odrzuca się, gdy wartość *p* jest mniejsza niż 0,05 lub 0,01. Po przeprowadzeniu testów statystycznych należy zinterpretować wyniki: jeśli wartość *p* jest mniejsza niż przyjęty poziom istotności, możemy odrzucić hipotezę zerową i uznać parametr za statystycznie istotny; w przeciwnym razie brak jest podstaw do odrzucenia hipotezy zerowej. Test t to test dwustronny, który opiera się na wartościach t oraz liczbie stopni swobody *df.* Pozwala to na wykonanie oceny, czy estymaty są statystycznie istotne na określonym poziomie istotności (przyjęto α=0,05). Procedura oblicza macierz kowariancji CovMatrix na podstawie estymat *b* uzyskanych z modelu regresji, danych rzeczywistych *y* oraz zmiennych X. Następnie obliczane są wartości statystyki t i wartości *p* – do testowania hipotezy statystycznej.

$$Reszty = y - \hat{y}$$

$$df = n - p - 1$$

$$CovMatrix = \frac{(Reszty * Reszty)}{df}$$

$$t = \frac{b}{\sqrt{diag(CovMatrix)}}$$

$$p = 2 * (1 - tcdf(abs(t), df))$$

gdzie: *df* – liczba stopni swobody, *tcdf* – funkcja rozkładu t–Studenta, diag(CovMatrix) – wartości na przekątnej macierzy kowariancji.

Hipoteza zerowa H_i=0 estymaty b_i występuje, gdy wartość $p > 0.05$ i oznacza brak istotności statystycznej, w przeciwnym przypadku estymata jest istotna.

Wartość SE to błąd standardowy (*ang. Standard Error*), który mierzy, jak dokładne jest oszacowanie współczynnika w stosunku do danych. Mniejszy SE oznacza większą precyzję oszacowania, co wskazuje, że wynik jest bardziej stabilny i wiarygodny. SE jest obliczany jako odchylenie standardowe podzielone przez pierwiastek kwadratowy z liczby obserwacji.

Wartość t, to wartość używana do sprawdzenia istotności współczynnika. Jest to stosunek oszacowanego współczynnika do jego błędu standardowego. Im wyższa bezwzględna wartość statystyki t, tym bardziej prawdopodobne, że oszacowany współczynnik istotnie różni się od zera (czyli ma znaczący wpływ na zmienną zależną).

Wartość *p* wskazuje prawdopodobieństwo, że statystyka t uzyskałaby taką wartość przypadkowo, gdyby faktyczny wpływ zmiennej był równy zero. Niska wartość *p* (np. poniżej 0,05) sugeruje, że współczynnik jest istotny statystycznie, co oznacza, że zmienna może mieć rzeczywisty wpływ na zmienną zależną.

Badania przeprowadzono dla wszystkich hoteli. Końcowe wyniki procesu doboru modelu ekonometrycznego metodą propagacji wstecznej dla poszczególnych hoteli znajdują się poniżej. Natomiast szczegółowe obliczenia dotyczące modelu początkowego wraz z estymatami, wartościami SE, *t* oraz wartościami *p* – zamieszono w Załączniku Z6.

W otrzymanych wynikach badań zestawiono rezultaty procesu doboru modeli dla poszczególnych hoteli (tabela 12 –18). W Załączniku Z6 podano formuły modeli początkowych, uwzględniających wszystkie zmienne USALI, które dla danego hotelu przedstawia tabela 4, wraz z ich estymatami oraz podano podsumowanie wyników doboru

modelu początkowego. Poniżej zapisano jedynie formułę modelu końcowego (ostatecznego), wartości estymat i podsumowanie wyników doboru modelu końcowego. Kluczowe elementy w podsumowaniu obejmują:

- Liczbę obserwacji – łączna liczba punktów danych, które były użyte do dopasowania modelu.
- Liczbę stopni swobody – liczba wolnych parametrów, które model może dopasować.
- Błąd średniokwadratowy – miara przeciętnego kwadratu odchyleń przewidywanych wartości od rzeczywistych.
- R-kwadrat – współczynnik determinacji, który wskazuje, jaki procent zmienności w danych można wyjaśnić modelem.
- R-kwadrat skorygowany – R-kwadrat uwzględniający liczbę predyktorów i dopasowanie modelu do liczby obserwacji.
- Statystykę F – test F wskazujący, czy model jest istotny statystycznie w porównaniu do modelu stałego.

Uzyskane wartości R-kwadrat (tabela 12–18), zawsze wynoszące co najmniej 0,98, wskazują, że niemal cała zmienność danych została wyjaśniona przez model we wszystkich modelach końcowych. Ponadto, wartości *p* były zawsze bliskie zeru, co potwierdza statystyczną istotność modeli. We wszystkich analizowanych modelach ostatecznych uwzględniono zmienną X52 (wynik operacyjny brutto), której współczynnik wahał się od 0,985 do 1,110. Oznacza to, że wzrost wyniku operacyjnego o jeden punkt skutkował równoczesnym wzrostem wartości GOP. Natomiast pozostałe zmienne znacząco różnicowały poszczególne modele, co zostanie opisane pod definicją poszczególnych modeli.

Model ekonometryczny dla HDSI Katowice

$$GOP = Wyraz\ wolny + X52 + X9 + X75 + X14$$

Wartości estymat, ich parametry statystyczne oraz nazwa zmiennej diagnostycznej:

Wyraz wolny: **0,034** (SE = 0,016; t = 2,11; p = 0,038)

X52: **1,007** (SE = 0,024; t = 42,66; p < 0,0001)

X9: **-0,072** (SE = 0,029; t = -2,48; p = 0,016)

X75: **-0,020** (SE = 0,005; t = -4,27; p < 0,0001)

X14: **0,066** (SE = 0,020; t = 3,38; p = 0,0012)

W oszacowanym modelu wyraz wolny jest dodatni i statystycznie istotny, co sugeruje, że przy zerowych wartościach pozostałych zmiennych wartość GOP jest dodatnia. Zmienne, poza X52, które są w modelu istotne to:

X9 – przychody operacyjne gastronomii, które oddziałują ujemnie na GOP – jednostkowy wzrost tej zmiennej powoduje niewielki spadek GOP, tj. o 0,072.

X75 – koszty niepodzielone mediów wykazują ujemny wpływ, czyli wzrost o jednostkę skutkuje obniżeniem GOP o 0,020.

X14 – przychody operacyjne żywność mają umiarkowany, dodatni wpływ na GOP – jednostkowy wzrost tej zmiennej prowadzi do wzrostu GOP o 0.066.

Podsumowując, model wskazuje, że oprócz dodatniego wyrazu wolnego, najważniejszym czynnikiem podnoszącym GOP jest wynik operacyjny (X52), podczas gdy przychody operacyjne gastronomii (X9) oraz koszty mediów (X75) działają w kierunku obniżenia GOP. Przychody operacyjne z żywności (X14) przyczyniają się natomiast do jego wzrostu, choć w umiarkowanym zakresie.

Tabela 12. Zestawienie zbiorcze wyników końcowego modelu regresji liniowej dla hotelu HDSI Katowice. Źródło: Opracowanie własne.

Podsumowanie wyników końcowego modelu regresji liniowej HDSI Katowice			
Liczba obserwacji	72	Błąd średniokwadratowy	0,118
Stopnie swobody błędu	67	R-kwadrat (R-kwadrat skorygowane)	0,987 (0,986)
Statystyka F względem modelu stałego	1250	*p*	< 0,0001

Model ekonometryczny dla HDS Piaseczno

$$GOP = Wyraz\ wolny + X52 + X23 + X32 + X15 + X19$$

Wartości estymat, ich parametry statyczne oraz nazwa zmiennej diagnostycznej:

Wyraz wolny: **-1,82e-16** (SE = 0,013; t = -1,38e-14; p = 1)

X52: **0,985** (SE = 0,026; t = 37,86; p < 0,0001)

X23: **-0,148** (SE = 0,026; t = -5,74; p < 0,0001)

X32: **0,084** (SE = 0,023; t = 3,61; p = 0,0006)

X15: **0,075** (SE = 0,017; t = 4,34; p < 0,0001)

X19: **0,039** (SE = 0,017; t = 2,32; p = 0,023)

Wyraz wolny modelu wyjaśniającego wartość GOP w hotelu HDS Piaseczno nie jest istotny statystycznie, co wskazuje, że przy zerowych wartościach pozostałych zmiennych przewidywana wartość GOP nie różni się istotnie od zera.

Pozostałe istotne zmienne to:

X23 – koszty operacyjne, które oddziałują negatywnie; jednostkowy wzrost kosztów operacyjnych wiąże się ze spadkiem GOP o 0,148.

X32 – koszty operacyjne pokoje, usługi prania, mają dodatni wpływ; wzrost tej zmiennej o jednostkę skutkuje wzrostem GOP o 0,084.

X15 – przychody operacyjne gastronomia – napoje, bankiety, wpływają dodatnio, powodują wzrost GOP o 0,075.

X19 – przychody operacyjne pozostałe centra również oddziałują dodatnio, lecz efekt tej zmiennej jest najmniejszy wśród pozytywnych wpływów i wynosi 0,039.

A zatem, model wskazuje, że kluczowym czynnikiem zwiększającym wynik operacyjny jest X52 (wynik operacyjny), którego wpływ jest niemal jednopunktowy. Z kolei koszty operacyjne (X23) wykazują wyraźny efekt obniżający GOP. Dodatkowe dodatnie efekty pochodzą z przychodów operacyjnych gastronomii (X15) oraz pozostałych centrów (X19) oraz z kosztów operacyjnych związanych z pokojami i usługami prania (X32), które jednak działają w umiarkowanym zakresie.

Tabela 13. Zestawienie zbiorcze wyników końcowego modelu regresji liniowej dla hotelu HDS Piaseczno. Źródło: Opracowanie własne.

Podsumowanie wyników końcowego modelu regresji liniowej HDS Piaseczno			
Liczba obserwacji	72	Błąd średniokwadratowy	0,112
Stopnie swobody błędu	66	R-kwadrat (R-kwadrat skorygowane)	0,988 (0,988)
Statystyka F względem modelu stałego	1130	*p*	< 0,0001

Model ekonometryczny dla HDSP Opole

$$GOP = Wyraz\ wolny + X52 + X24 + X78$$

Wartości estymat, ich parametry statyczne oraz nazwa zmiennej diagnostycznej:

Wyraz wolny: **4,20e-16** (SE = 0,015; t = 2,79e-14; p = 1)

X52: **1,110** (SE = 0,020; t = 54,82; p < 0,0001)

X24: **-0,046** (SE = 0,022; t = -2,13; p = 0,038)

X78: **-0,164** (SE = 0,022; t = -7,40; p < 0,0001)

W estymowanym modelu końcowym pojawiły się jedynie trzy istotne zmienne X52, X24 i 78. Wyraz wolny nie jest statystycznie istotny, co sugeruje, że przy zerowych wartościach zmiennych objaśniających przewidywana wartość GOP nie odbiega istotnie od zera.

X24 – koszty operacyjne pokoje oddziałują ujemnie na GOP – jednostkowy wzrost kosztów operacyjnych związanych z pokojami powoduje obniżenie GOP o 0,046.

X78 – koszty niepodzielone dział sprzedaży wykazują istotny ujemny wpływ, przy czym wzrost tej zmiennej o jednostkę skutkuje spadkiem GOP w umiarkowanie większym stopniu niż obserwowany przy kosztach operacyjnych pokoi, tj. o 0,164.

Podsumowując, model wskazuje, że głównym czynnikiem zwiększającym wynik operacyjny jest zmienna X52 (wynik operacyjny), natomiast koszty operacyjne pokojów (X24) oraz koszty niepodzielone działu sprzedaży (X78) działają negatywnie, przy czym efekt obniżający GOP jest bardziej wyraźny dla kosztów niepodzielonych działu sprzedaży.

Tabela 14. Zestawienie zbiorcze wyników końcowego modelu regresji liniowej dla hotelu HDSP Opole. Źródło: Opracowanie własne.

Podsumowanie wyników końcowego modelu regresji liniowej HDSP Opole			
Liczba obserwacji	60	Błąd średniokwadratowy	0,117
Stopnie swobody błędu	56	R-kwadrat (R-kwadrat skorygowane)	0,987 (0,986)
Statystyka F względem modelu stałego	1430	*p*	< 0,0001

Model ekonometryczny dla HDSP Poznań

$$GOP = Wyraz\ wolny + X52 + X5 + X53 + X38 + X78 + X24$$

Wartości estymat, ich parametry statyczne oraz nazwa zmiennej diagnostycznej:

Wyraz wolny: **0,017** (SE = 0,068; t = 0,25; p = 0,804)

X52: **1,002** (SE = 0,052; t = 19,22; p < 0,0001)

X5: **0,116** (SE = 0,047; t = 2,46; p = 0,020)

X53: **0,115** (SE = 0,037; t = 3,08; p = 0,0045)

X38: **-0,085** (SE = 0,029; t = -2,90; p = 0,0071)

X78: **-0,035** (SE = 0,010; t = -3,57; p = 0,0013)

X24: **-0,202** (SE = 0,026; t = -7,92; p < 0,0001)

Wyniki analizy wskazują na następujący wpływ poszczególnych zmiennych:

Wyraz wolny nie jest statystycznie istotny, co sugeruje, że przy zerowych wartościach zmiennych objaśniających przewidywana wartość GOP nie odbiega istotnie od zera.

X5 – OCC ma dodatni, choć stosunkowo niewielki wpływ na GOP – wzrost o jednostkę powoduje umiarkowany przyrost wyniku operacyjnego o 0,116.

X53 – wynik operacyjny gastronomia wpływa dodatnio na GOP – przy jednostkowym wzroście tej zmiennej następuje niewielki wzrost wyniku operacyjnego o 0,115.

X38 – koszty operacyjne gastronomii – koszty surowca oddziałują ujemnie na GOP – wzrost tej zmiennej o jednostkę powoduje obniżenie wyniku operacyjnego o 0,085.

X78 – koszty niepodzielone dział sprzedaży mają negatywny wpływ na GOP – wzrost tej zmiennej skutkuje spadkiem wyniku operacyjnego, choć efekt jest mniejszy niż w przypadku niektórych innych kosztów i wynosi -0,035.

X24 – koszty operacyjne pokoje wykazują istotny, ujemny wpływ na GOP – wzrost tej zmiennej o jednostkę wiąże się z umiarkowanym obniżeniem wyniku operacyjnego o 0,202.

Podsumowując można zapisać, że w modelu kluczowym czynnikiem dodatnio wpływającym na wynik operacyjny jest zmienna X52, której efekt jest niemal jednostkowy. Dodatkowo, niewielki, ale istotny pozytywny wpływ na GOP mają zmienne X5 oraz X53. Z kolei koszty operacyjne, związane z gastronomią (X38) oraz pokojami (X24), a także koszty niepodzielone działu sprzedaży (X78), wywierają negatywny wpływ, przy czym największe obniżenie GOP obserwuje się przy wzroście kosztów operacyjnych pokoi.

Tabela 15. Zestawienie zbiorcze wyników końcowego modelu regresji liniowej dla hotelu HDSP Poznań. Źródło: Opracowanie własne.

Podsumowanie wyników końcowego modelu regresji liniowej HDSP Poznań			
Liczba obserwacji	36	Błąd średniokwadratowy	0,118
Stopnie swobody błędu	29	R-kwadrat (R-kwadrat skorygowane)	0,988 (0,986)
Statystyka F względem modelu stałego	413	*p*	< 0,0001

Model ekonometryczny dla HM Opole

$$GOP = Wyraz\ wolny + X52 + X38 + X18 + X32$$

Wartości estymat, ich parametry statyczne oraz nazwa zmiennej diagnostycznej:

Wyraz wolny: **-0,120** (SE = 0,059; t = -2,05; p = 0,049)

X52: **1,102** (SE = 0,045; t = 24,59; p < 0,0001)

X38: **-0,115** (SE = 0,036; t = -3,14; p = 0,0037)

X18: **-0,159** (SE = 0,039; t = -4,02; p = 0,0004)

X32: **0,088** (SE = 0,039; t = 2,29; p = 0,029)

Wyniki estymacji modelu HM Opole wskazują na następujący wpływ poszczególnych składników:

Wyraz wolny: Przyjmuje istotnie ujemną wartość, co oznacza, że przy zerowych wartościach zmiennych objaśniających przewidywany wynik operacyjny byłby ujemny.

X38 (koszty operacyjne gastronomii – koszty surowca): Oddziałuje ujemnie na GOP – wzrost tej zmiennej powoduje umiarkowany spadek wyniku operacyjnego o 0,115.

X18 (przychody operacyjne parkingu/garażu): Wpływa negatywnie na GOP – jej wzrost wiąże się z nieco większym obniżeniem wyniku operacyjnego niż obserwowany w przypadku X38.

X32 (koszty operacyjne pokoi, usługi prania): Ma dodatni, umiarkowany wpływ – wzrost tej zmiennej przyczynia się do wzrostu GOP o 0,088.

Podsumowując, kluczowym czynnikiem podnoszącym wynik operacyjny jest zmienna X52, której dodatni wpływ jest bardzo silny. Z kolei wzrost zmiennych X38 oraz X18 powoduje obniżenie GOP, przy czym efekt X18 jest nieco silniejszy. Dodatni, lecz umiarkowany wpływ zmiennej X32 również przyczynia się do wzrostu wyniku operacyjnego. Ujemny wyraz wolny sugeruje, że przy zerowym poziomie zmiennych objaśniających wynik operacyjny byłby ujemny.

Tabela 16. Zestawienie zbiorcze wyników końcowego modelu regresji liniowej dla hotelu HM Opole. Źródło: Opracowanie własne.

Podsumowanie wyników końcowego modelu regresji liniowej HM Opole			
Liczba obserwacji	35	Błąd średniokwadratowy	0,154
Stopnie swobody błędu	30	R-kwadrat (R-kwadrat skorygowane)	0,979 (0,976)
Statystyka F względem modelu stałego	353	*p*	< 0,0001

Model ekonometryczny dla HMC Częstochowa

GOP = Wyraz wolny + X52 + X39 + X23 + X26 + X78 + X32 + X18 + X77 + X64 + X29

Wartości estymat, ich parametry statyczne oraz nazwa zmiennej diagnostycznej:

Wyraz wolny: **0,085** (SE = 0,045; t = 1,89; p = 0,071)

X52: **1,061** (SE = 0,062; t = 17,01; p < 0,0001)

X39: **-0,215** (SE = 0,068; t = -3,15; p = 0,0044)

X23: **0,293** (SE = 0,059; t = 4,92; p < 0,0001)

X26: **0,081** (SE = 0,036; t = 2,27; p = 0,033)

X78: **-0,101** (SE = 0,031; t = -3,24; p = 0,0036)

X32: **0,041** (SE = 0,019; t = 2,21; p = 0,037)

X18: **-0,058** (SE = 0,027; t = -2,17; p = 0,040)

X77: **-0,078** (SE = 0,031; t = -2,55; p = 0,018)

X64: **0,067** (SE = 0,015; t = 4,53; p < 0,0001)

X29: **0,049** (SE = 0,019; t = 2,55; p = 0,018)

W estymowanym modelu ekonometrycznym dla hotelu HMC Częstochowa, wyraz wolny nie osiąga istotności statystycznej, co sugeruje, że przy jednoczesnym uwzględnieniu wpływu pozostałych zmiennych jego wartość nie różni się istotnie od zera. Pozostałe zmienne to:

X39 – koszty operacyjne gastronomii, żywność: Wzrost kosztów w tej kategorii o jednostkę skutkuje spadkiem GOP o około 0,215 jednostki.

X23 – koszty operacyjne: Wzrost tej zmiennej o jednostkę wiąże się ze wzrostem GOP o około 0,293 jednostki.

X26 – koszty operacyjne pokoje (zużycie materiałów): Jednostkowy wzrost tej zmiennej powoduje wzrost GOP o około 0,081 jednostki.

X78 – koszty niepodzielone działu sprzedaży: Wzrost o jednostkę prowadzi do obniżenia GOP o około 0,101 jednostki.

X32 – koszty operacyjne pokoje, usługi prania: Wzrost o jednostkę wiąże się ze wzrostem GOP o około 0,041 jednostki.

X18 – przychody operacyjne parking garaż: Wzrost tej zmiennej o jednostkę powoduje spadek GOP o około 0,058 jednostki.

X77 – koszty niepodzielone (zużycie wody): Jednostkowy wzrost tej zmiennej skutkuje obniżeniem GOP o około 0,078 jednostki.

X64 – koszty niepodzielone (administracja, pozostałe usługi): Wzrost tej zmiennej o jednostkę wiąże się ze wzrostem GOP o około 0,067 jednostki.

X29 – koszty operacyjne pokoje (pościel, ręczniki itd.): Wzrost o jednostkę powoduje wzrost GOP o około 0,049 jednostki.

Podsumowując, model wskazuje, że wzrost wartości zmiennych X52, X23, X26, X32, X64 oraz X29 przyczynia się do zwiększenia GOP, podczas gdy wzrost zmiennych X39, X78, X18 oraz X77 prowadzi do jego obniżenia. Wyniki te podkreślają kluczowy wpływ zarówno kosztów operacyjnych, jak i niepodzielonych kosztów administracyjnych oraz sprzedażowych na kształtowanie wyniku operacyjnego hotelu.

Tabela 17. Zestawienie zbiorcze wyników końcowego modelu regresji liniowej dla hotelu HMC Częstochowa. Źródło: Opracowanie własne.

Podsumowanie wyników końcowego modelu regresji liniowej HMC Częstochowa			
Liczba obserwacji	34	Błąd średniokwadratowy	0,12
Stopnie swobody błędu	23	R-kwadrat (R-kwadrat skorygowane)	0,99 (0,986)
Statystyka F względem modelu stałego	228	*p*	< 0,0001

Model ekonometryczny dla HM Karpacz

$$GOP = Wyraz\ wolny + X52 + X32 + X77 + X78 + X51$$

Wartości estymat, ich parametry statyczne oraz nazwa zmiennej diagnostycznej:

Wyraz wolny: **0,003** (SE = 0,018; t = 0,14; p = 0,887)

X52**:** **1,101** (SE = 0,040; t = 27,62; p < 0,0001)

X32: **0,101** (SE = 0,044; t = 2,28; p = 0,031)

X77: **-0,180** (SE = 0,046; t = -3,93; p = 0,0006)

X78: **0,084** (SE = 0,016; t = 5,10; p < 0,0001)

X51: **-0,064** (SE = 0,023; t = -2,80; p = 0,0095)

W analizowanym modelu końcowym, wyraz wolny nie wykazuje istotności statystycznej. Istotny wpływ na wartość GOP mają następujące zmienne:

X32 – koszty operacyjne pokoi, w tym usługi prania: Wzrost tej zmiennej o jednostkę wiąże się ze wzrostem GOP o około 0,101 jednostki (p = 0,031).

X77 – koszty niepodzielone związane z zużyciem wody: Wzrost tej zmiennej o jednostkę powoduje spadek GOP o około 0,180 jednostki (p = 0,0006).

X78 – koszty niepodzielone działu sprzedaży: Wzrost tej zmiennej o jednostkę skutkuje

wzrostem GOP o około 0,084 jednostki ($p < 0,0001$).

X51 – koszty operacyjne w pozostałych centrach: Wzrost tej zmiennej o jednostkę wiąże się z obniżeniem GOP o około 0,064 jednostki ($p = 0,0095$).

Podsumowując, wyniki modelu wskazują, że najważniejszym czynnikiem podnoszącym GOP jest wynik operacyjny (X52) oraz że zarządzanie kosztami operacyjnymi pokoi (X32) oraz kosztami niepodzielonymi działu sprzedaży (X78) przyczynia się do wzrostu wyniku operacyjnego hotelu, natomiast koszty związane z zużyciem wody (X77) oraz operacyjne koszty w pozostałych centrach (X51) mają działanie negatywne na GOP

Tabela 18. Zestawienie zbiorcze wyników końcowego modelu regresji liniowej dla hotelu HM Karpacz. Źródło: Opracowanie własne.

Podsumowanie wyników końcowego modelu regresji liniowej HM Karpacz			
Liczba obserwacji	32	Błąd średniokwadratowy	0,102
Stopnie swobody błędu	26	R-kwadrat (R-kwadrat skorygowane)	0,991 (0,99)
Statystyka F względem modelu stałego	587	*p*	< 0,0001

Aby ocenić jakość dopasowania końcowego modelu predykcyjnego wykonano analizę rozkładów reszt, w tym zbadano ich zgodność z rozkładem normalnym i symetryczność. Ponadto reszty testowano pod kątem występowania autokorelacji i stałości wariancji składnika losowego (błędów) w modelu regresji w stosunku do wartości zmiennych niezależnych.

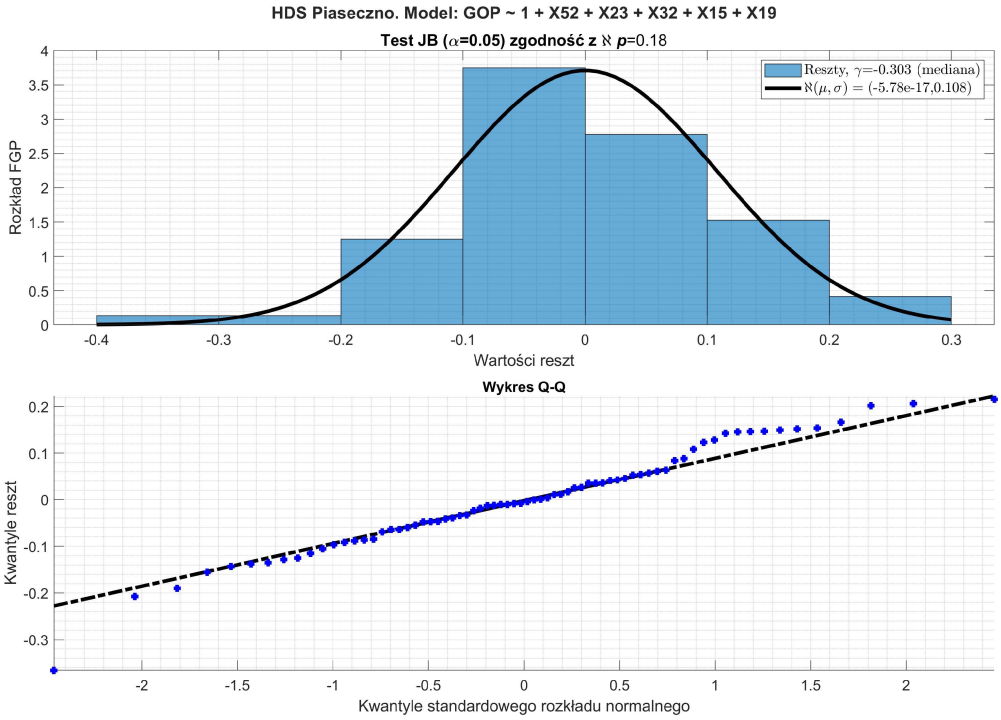
W tabeli 26 zawarto zestawienie zbiorcze wyników procesu doboru modelu ekonometrycznego, dla wszystkich hoteli. W szczególności tabela zawiera wyniki analizy reszty, a w tym: wartość miary symetrii (skośność) i wartości *p* z testów na zgodność z rozkładem normalnym (test Jarque'a-Berry), homoskedastyczność (test ARCH Engle'a) i autokorelację rzędu 1-go (test Durbina-Watsona).

Analiza wartości skośności rozkładów obliczonych reszt wykazała wartości większe niż moduł z 1 jedynie dla hotelu HMC Częstochowa i HM Karpacz. W pozostałych przypadkach skośność jest bliższa wartości zerowej. Analiza reszt za pomocą testu Jarque'a-Berry wykazała brak zgodności z rozkładem normalnym dla hoteli HMC Częstochowa i HM Karpacz. Z kolei w analizie za pomocą testu ARCH Engle'a obliczono, że reszty uzyskane dla hoteli HDS Piaseczno i HDSP Opole wykazują heteroskedastyczność. Ponadto wykazano, że żaden z modeli nie wykazuje autokorelacji reszt.

Tabela 19. Zestawienie zbiorcze wyników procesu doboru końcowego modelu ekonometrycznego, dla wszystkich hoteli. Źródło: Opracowanie własne.

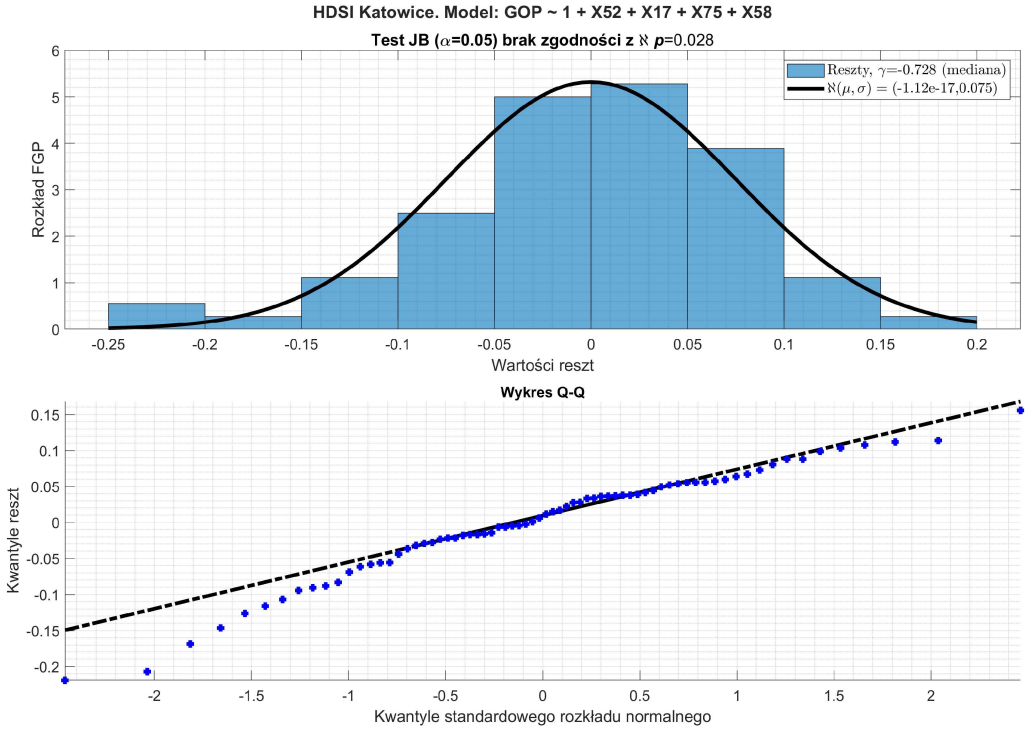
Wyniki analizy reszt z modelowania liniowego - model końcowy				
Hotel	Skośność	Zgodność z rozkładem א wartość *p*	Homo-skedastyczność wartość *p*	Autokorelacja wartość *p*
HDSI Katowice	-0,728	0,028	0,136	0,440
HDS Piaseczno	-0,303	0,175	0,007	0,568
HDSP Opole	0,513	0,099	0,003	0,463
HDSP Poznań	-0,468	0,343	0,758	0,866
HM Opole	0,161	0,500	0,222	0,136
HMC Częstochowa	-1,019	0,018	0,481	0,910
HM Karpacz	1,400	0,001	0,783	0,925

Dodatkowym sposobem oceny jakości predyktora szeregu czasowego jest wykres zależności kwantyli reszt do kwantyli rozkładu normalnego. Na rysunkach 7.1 – 7.7 przedstawiono wyniki analizy reszt uzyskanych w wyniku doboru modelu ekonometrycznego dla zmiennej GOP, zarejestrowaną w kolejnych hotelach. Każdy z rysunków zawiera dwa wykresy. Wykres u góry przedstawia rozkład reszt wraz z dopasowaniem FGP o rozkładzie normalnym $\aleph(\mu, \sigma)$. W tytule podano formułę modelu i wynik zgodności z rozkładem normalnym, a w legendzie podano wartość skośności γ i parametry dopasowanej funkcji $\aleph(\mu, \sigma)$. Kolejny wykres stanowi wizualizację zależności kwantyli reszt do kwantyli rozkładu normalnego.

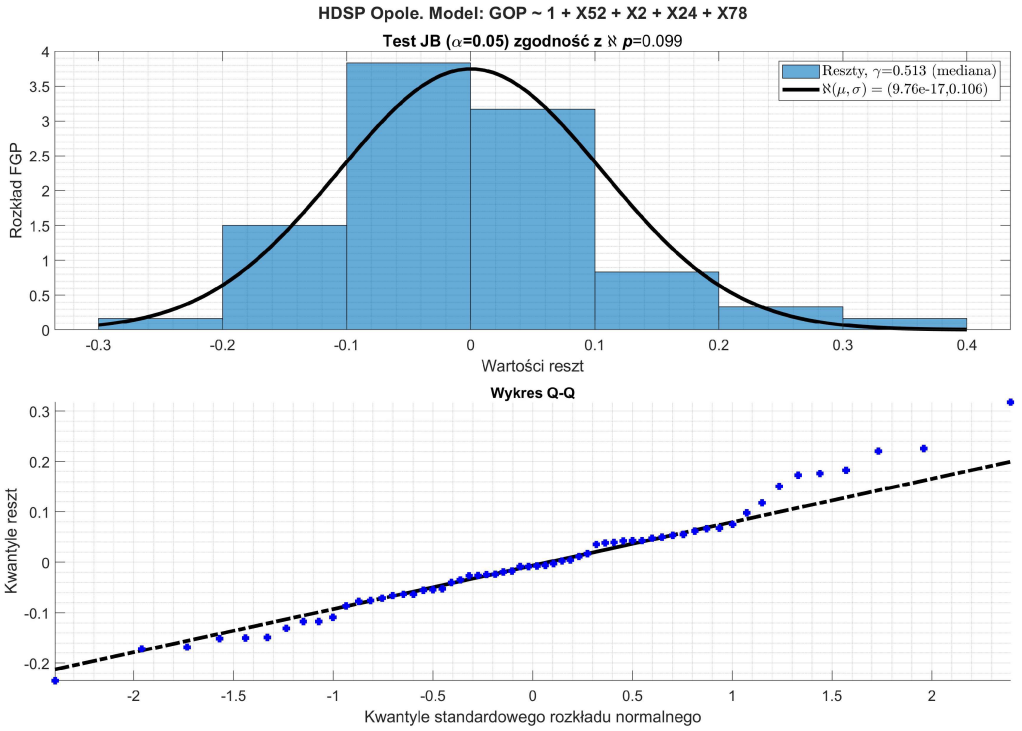


óln

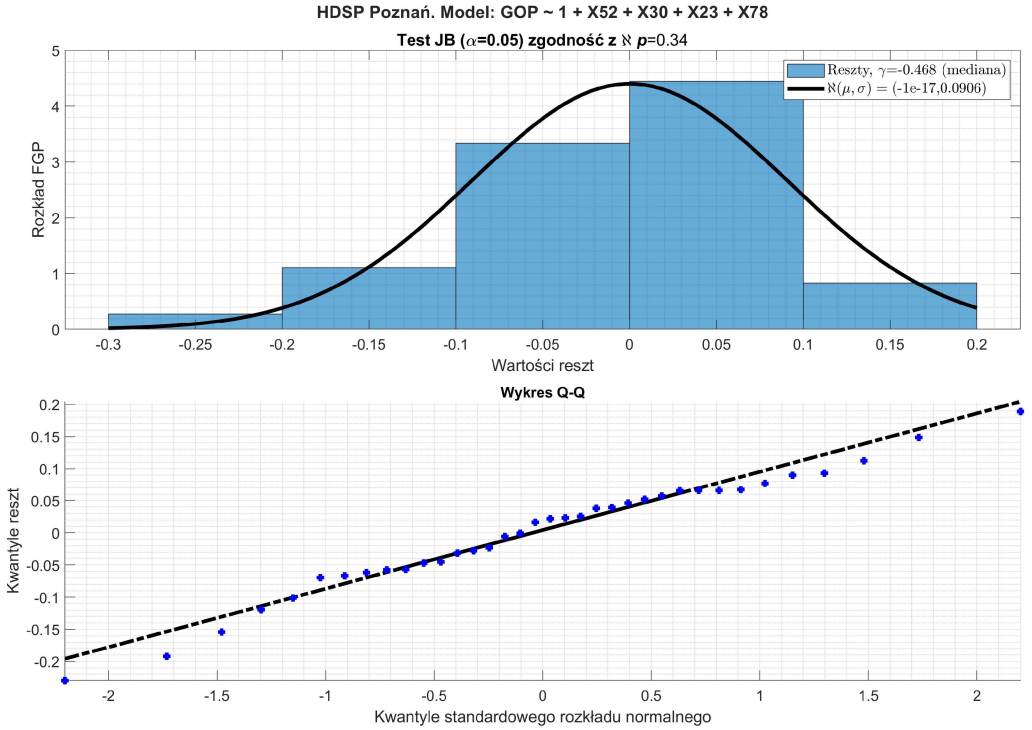
Rys. 7.1 Zestawienie porównawcze rozkładu gęstości prawdopodobieństwa reszt wraz z wykresem Q–Q, uzyskane w wyniku dopasowania modelu ekonometrycznego dla hotelu HDS Piaseczno. Źródło: Opracowanie własne.



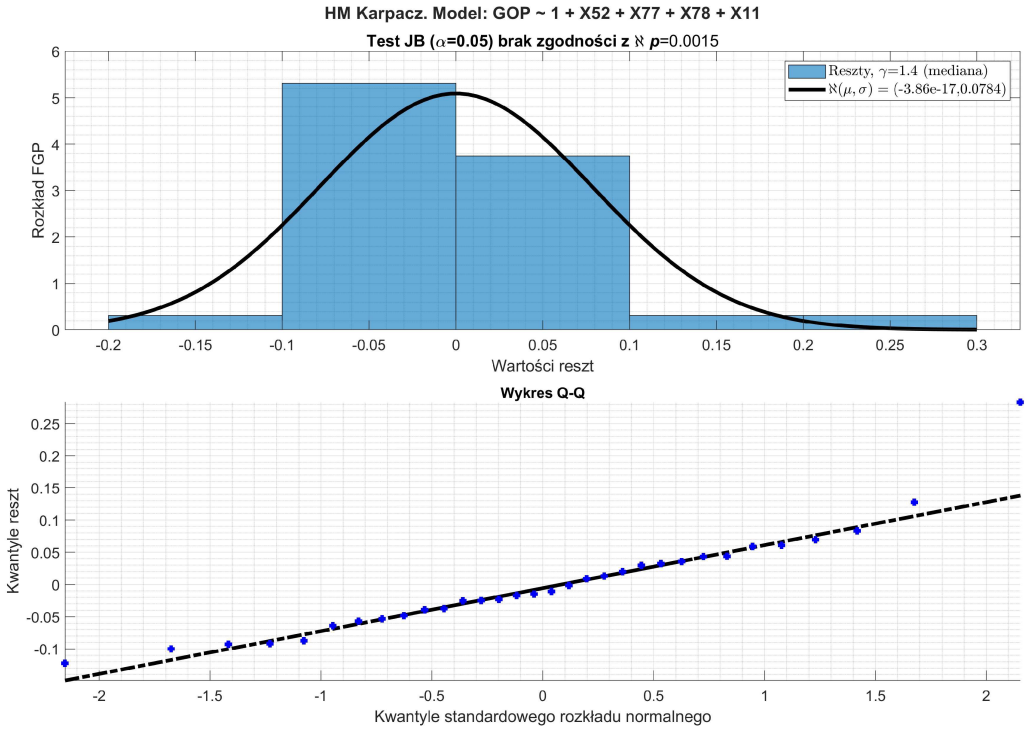
Rys. 7.2 Zestawienie porównawcze rozkładu gęstości prawdopodobieństwa reszt wraz z wykresem Q–Q, uzyskane w wyniku dopasowania modelu ekonometrycznego dla hotelu HDSI Katowice. Źródło: Opracowanie własne.



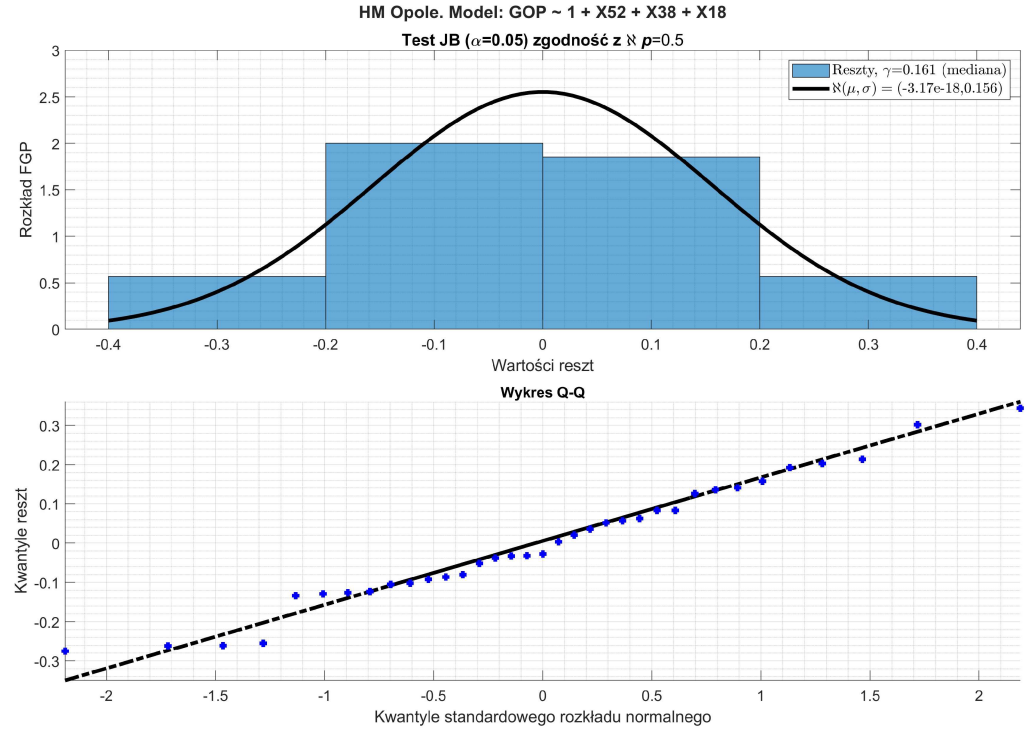
Rys. 7.3 Zestawienie porównawcze rozkładu gęstości prawdopodobieństwa reszt wraz z wykresem Q–Q, uzyskane w wyniku dopasowania modelu ekonometrycznego dla hotelu HDSP Opole. Źródło: Opracowanie własne.



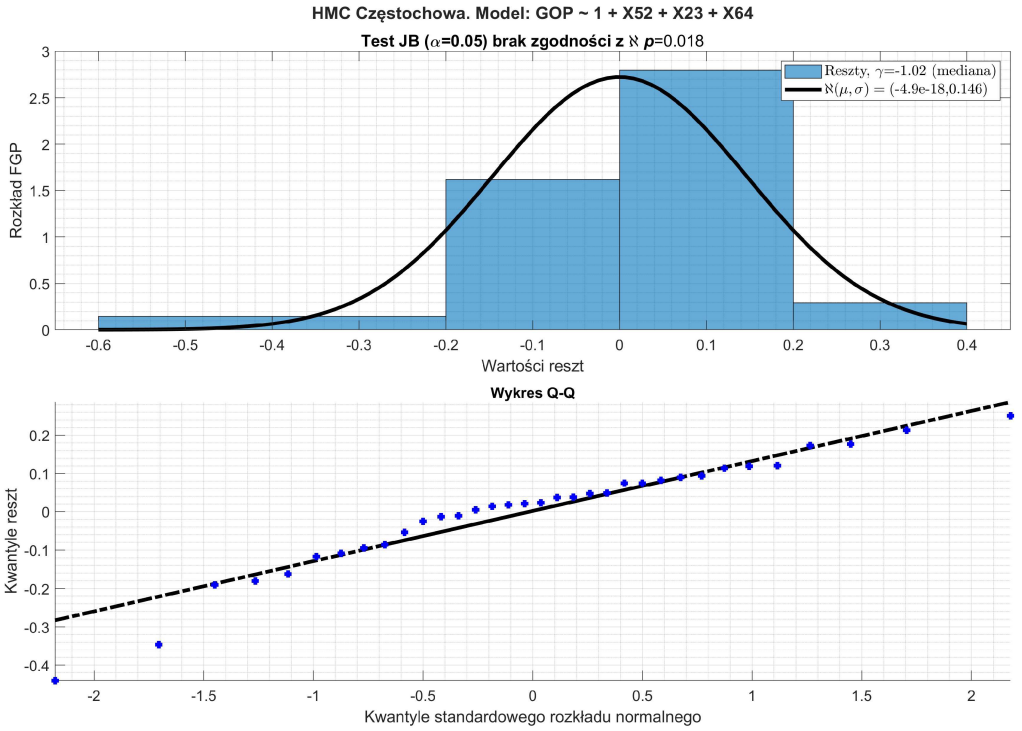
Rys. 7.4 Zestawienie porównawcze rozkładu gęstości prawdopodobieństwa reszt wraz z wykresem Q–Q, uzyskane w wyniku dopasowania modelu ekonometrycznego dla hotelu HDSP Poznań. Źródło: Opracowanie własne.



Rys. 7.5 Zestawienie porównawcze rozkładu gęstości prawdopodobieństwa reszt wraz z wykresem Q–Q, uzyskane w wyniku dopasowania modelu ekonometrycznego dla hotelu HM Karpacz. Źródło: Opracowanie własne.



Rys. 7.6 Zestawienie porównawcze rozkładu gęstości prawdopodobieństwa reszt wraz z wykresem Q–Q, uzyskane w wyniku dopasowania modelu ekonometrycznego dla hotelu HM Opole. Źródło: Opracowanie własne.



Rys. 7.7 Zestawienie porównawcze rozkładu gęstości prawdopodobieństwa reszt wraz z wykresem Q–Q, uzyskane w wyniku dopasowania modelu ekonometrycznego dla hotelu HMC Częstochowa. Źródło: Opracowanie własne.

7.2. Proces budowy modelu predykcyjnego z wykorzystaniem nieliniowych algorytmów regresyjnych i uczenia maszynowego

W zadaniu przyjęto założenie, że analizie zostaną poddane jedynie wyselekcjonowane, wzajemnie nieskorelowane zmienne USALI, które wykazują współzależność z GOP, czyli te które podano w tabeli 4, lecz bez transformacji. Modele nieliniowe nie wymagają spełnienia standardowych założeń takich jak homoskedastyczność czy normalność rozkładu zmiennych, a także nie narzucają konieczności stacjonarności danych. Dzięki tej odporności, modele nieliniowe mogą efektywnie uchwycić złożone zależności, nawet w przypadku danych niespełniających tradycyjnych założeń ekonometrycznych.

Ogólny proces budowy efektywnego modelu prognozującego jest procesem iteracyjnym, obejmującym szereg kroków. Kroki te zwizualizowano w postaci schematu blokowego na rys. 7.8. Pierwszym z nich jest zgromadzenie i przygotowanie danych, włączając w to selekcję odpowiednich zmiennych (zwanych dalej cechami), usuwanie błędów w danych (tzw. czyszczenie danych) i normalizację. Następnie, dane są dzielone na zestawy treningowy (walidacyjny) i testowy w celu oceny wydajności modelu. W kolejnym kroku następuje wybór odpowiedniego algorytmu uczenia maszynowego, takiego jak sieci neuronowe, drzewa decyzyjne czy maszyny wektorów nośnych i inicjalizacja jego parametrów. Następnie model jest trenowany na danych treningowych, a proces ten obejmuje optymalizację funkcji kosztu

poprzez dostosowywanie jego wag i parametrów. Podczas treningu modelu zwykle stosowane są metody pozwalające na ocenę wydajności i regularyzację modelu, np. walidowanie krzyżowe (*ang. cross-validation*), bootstrappingowe *(ang. bootstrap validation*), walidowanie z wykorzystaniem walidacji krzyżowej stratyfikowanej (*ang. stratified cross-validation*).

Po zakończeniu treningu modelu, przystępuje się do jego oceny za pomocą danych walidacyjnych lub testowych. W zależności od zadania: klasyfikacji, grupowania czy regresji, stosuje się różne miary oceny jakości. W zagadnieniach klasyfikacji i grupowania ocena może obejmować metryki, takie jak dokładność (*ang. accuracy*), precyzja (*ang. precision*), czułość (*ang. recall*), specyficzność (*ang. specificity*), F1-score, krzywa ROC (*ang. receiver operating characteristic*), miara Kappa (*ang. Cohen's kappa*) itp. Natomiast, miary R^2 oraz R^2 skorygowane (adjR^2) służą do oceny jakości dopasowania modelu regresji do rzeczywistych szeregów czasowych. Obydwie miary pozwalają zrozumieć, jak dobrze model wyjaśnia zmienność danych oraz jak dobrze przewiduje rzeczywiste obserwacje. Miara R^2 jest obliczana w celu określenia, jak duża część zmienności w zmiennej zależnej (danych rzeczywistych) jest wyjaśniana przez model regresji. Jej wartość zawiera się w zakresie od 0 do 1. Wartość równa 1 oznacza, że model doskonale przewiduje zmienność danych, natomiast wartość równa 0 oznacza, że model nie wyjaśnia w ogóle zmienności danych. Adj R^2 to zmodyfikowana wersja miary R^2, która uwzględnia liczbę zmiennych objaśniających w modelu. Poprawia ona niedoskonałość R^2, która może rosnąć w miarę dodawania kolejnych zmiennych do modelu, nawet jeśli te zmienne nie mają istotnego wpływu na przewidywania.

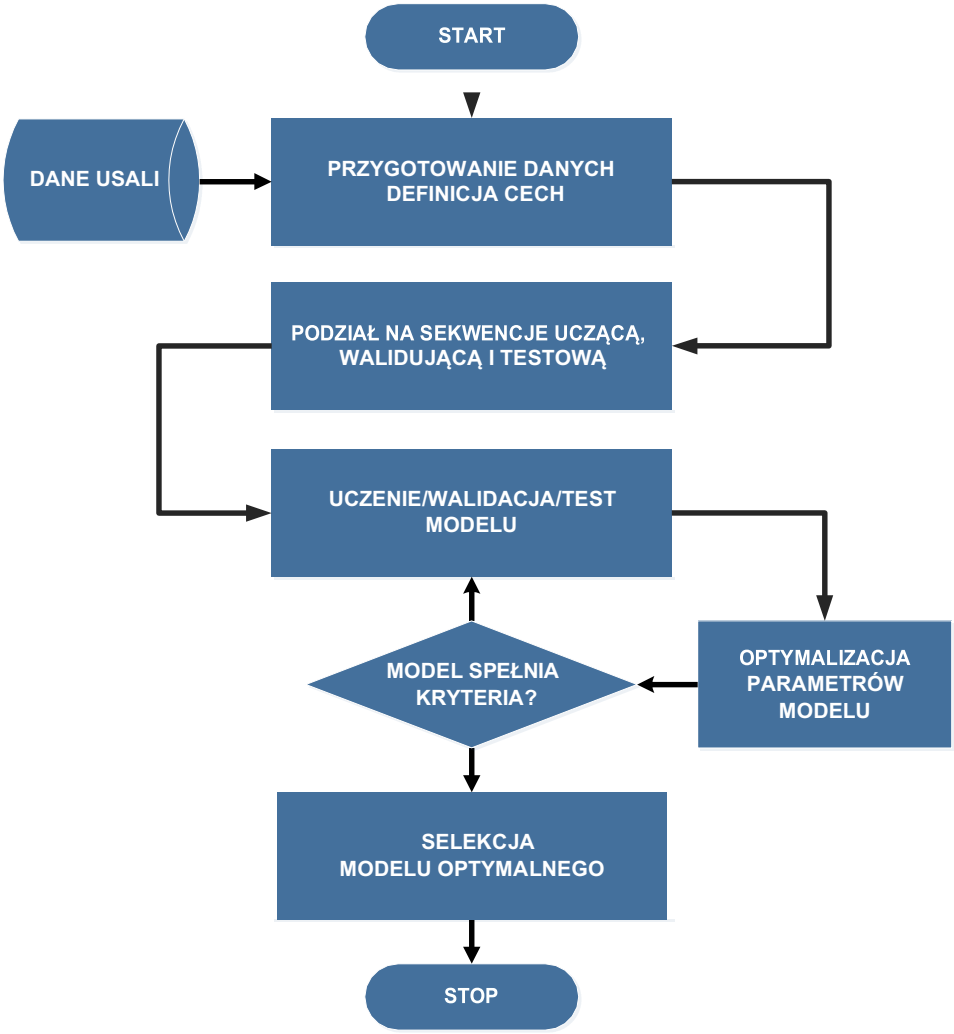
$$R^2 = 1 - \frac{SSE}{SST}$$

$$SSE = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

$$SST = \sum_{i=1}^{n} (y_i - \bar{y})^2$$

$$adjR^2 = 1 - \frac{(1 - R^2)(1 - n)}{n - p - 1}$$

gdzie: n – liczba obserwacji pojedynczej cechy, p – liczba cech w modelu.



Rys. 7.8 Schemat blokowy procesu przygotowania i modelowania danych. Źródło: Opracowanie własne.

W przypadku niezadowalających wyników, możliwe jest dostosowanie parametrów modelu lub zmiana architektury w celu dalszej optymalizacji. Gdy osiągnięta jest satysfakcjonująca wydajność modelu, proces zostaje zakończony, a opracowany model można wykorzystać do prognozowania nowych danych.

7.2.1. Analiza zmiennych USALI jako predyktorów w modelach nieliniowych

Istnieje szereg standardowych algorytmów pozwalających na określenie ważności i przydatności danej cechy w procesie budowy modelu predykcyjnego. W badaniach wykorzystano dwie metody:

1. Metoda MRMR (*ang. Minimum Redundancy and Maximum Relevance*) (Ding & Peng, 2005; Radovic, Ghalwash, Filipovic, & Obradovic, 2017; Zhao, Anand, & Wang, 2019) jest algorytmem służącym do uszeregowania cech (zmiennych objaśniających) w zadaniach

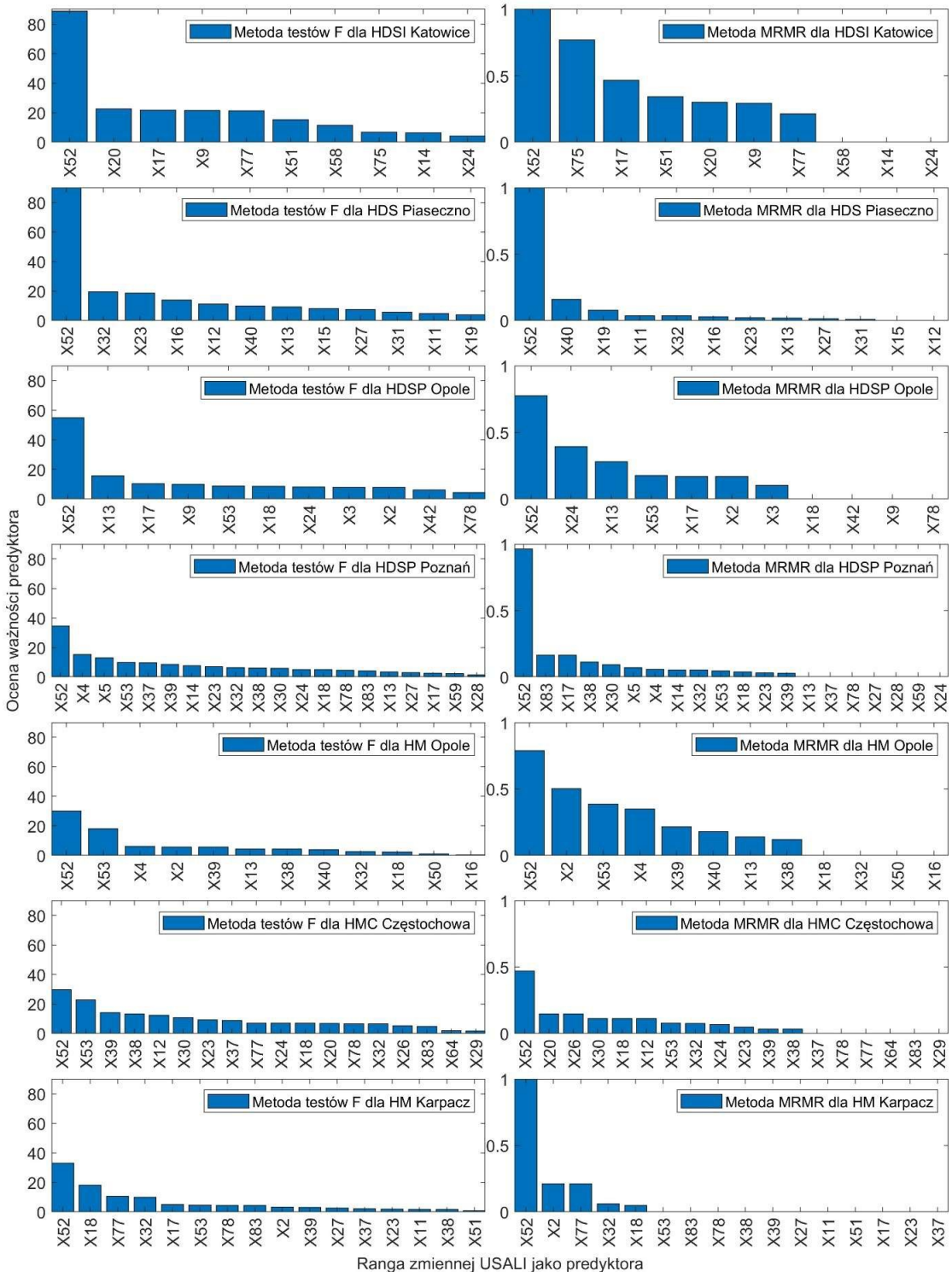
regresji i klasyfikacji. Jej główną zasadą jest minimalizacja redundantnych informacji pomiędzy cechami oraz maksymalizacja trafności cech w odniesieniu do zmiennej zależnej. Celem metody MRMR jest wybranie najbardziej istotnych cech, które jednocześnie są jak najmniej redundantne w kontekście zmiennej zależnej. Dzięki temu redukuje się wymiarowość danych i wybiera najlepsze cechy, co może poprawić wydajność modelu i uniknąć przetrenowania (*ang. overfitting*). Wadą metody MRMR jest wysoki koszt obliczeniowy, szczególnie dla dużych zbiorów danych, brak uwzględnienia interakcji między cechami oraz duża zależność od zastosowanej metryki.

2. Metoda jednoczynnikowego rankingu cech dla regresji z wykorzystaniem testów F jest techniką stosowaną do wyboru najbardziej istotnych cech (zmiennych objaśniających) w zadaniach regresji (Petković, Kocev, & Džeroski, 2020; Song, Jiang, & Liu, 2017). W metodzie tej dla każdej cechy oblicza się statystykę testu F (*ang. Fisher-Snedecorratio*), która mierzy istotność wpływu danej cechy na zmienną zależną. Celem tej metody jest identyfikacja tych cech, które wykazują największy wpływ na zmienność zmiennej zależnej, czyli wybranie cech, które mają istotne statystycznie znaczenie i wpływają na wynik modelu regresji. Zaletą ten metody jest jej stosunkowo prosta implementacja programistyczna. Do wad zalicza się, podobnie jak w przypadku metody MRMR, brak uwzględnienia interakcji między cechami oraz ryzyko wystąpienia przetrenowania.

Należy mieć na uwadze, że rozpatrywano jedynie zmienne skorelowane ze zmienną wyjściową, tj. GOP, podane w tabeli 4. Na rys. 7.9 przedstawiono zestawienie zbiorcze analizy zmiennych z wykorzystaniem obu metod oceny ważności. W lewej kolumnie wizualizowano wyniki dla metody MRMR, w prawej kolumnie wyniki dla metody testu F. W kolejnych wierszach zawarto wyniki dla poszczególnych hoteli. Niebieskie słupki na wykresach odpowiadają poziomowi istotności danej zmiennej, której oznaczenie znajduje się na osiach odciętych (OX). Słupki są uporządkowane w sposób monotoniczne malejący.

Z wykresu można odczytać, że dla HDSI Katowice i HDS Piaseczno uzyskano najwyższe poziomy ważności przekraczające 80, obliczone metodą testów F, każdorazowo dla zmiennej X52 (*wynik operacyjny*). Pozostałe zmienne uzyskały znacznie niższe ważności o porównywalnych poziomach. Natomiast metoda MRMR wykazała, że najwyższe rangi – bliskie jedności, uzyskano dla HDSI Katowice, HDS Piaseczno, HDSP Poznań i HM Karpacz. W pozostałych hotelach cechy uzyskują niższe oceny wartości. We wszystkich przypadkach najważniejsza rangą jest zmienna USALI X52.

Warto nadmienić, że wyniki obliczeń nie w pełni pokrywają się z tymi, które wykazał analiza, wykonana za pomocą modelowania ekonometrycznego, czyli rozpatrywane w pracy metody wykazują jako ważne w predykcji GOP inne typy zmiennych, za wyjątkiem zmiennej X52.



Rys. 7.9 Zestawienie zbiorcze wyników analizy ważności zestawu zmiennych USALI skorelowanych ze zmienną GOP wyselekcjonowanych w poprzednich badaniach, wyznaczone z wykorzystaniem metody MRM i testu F. Źródło: Opracowanie własne.

7.2.2. Modelowanie zysku operacyjnego brutto z wykorzystaniem nieliniowych algorytmów regresyjnych

Na potrzeby opracowania modelu nieliniowego utworzono macierz predyktorów dla wszystkich zmiennych wyznaczonych we wcześniejszych badaniach. Podobnie jak w przypadku modelu ekonometrycznego, kolejne wiersze stanowią dane miesięczne, które poddano standaryzacji.

W zadaniu przebadano pięć klasycznych algorytmów uczenia maszynowego, które wymieniono łącznie z zastosowanymi w pracy nazwami skrótowymi w tabeli 20.

Tabela 20. Nazwy i oznaczenia algorytmów uczenia maszynowego zastosowanych w analizach. Źródło: Opracowanie własne.

L.P.	Nazwa algorytmu	Oznaczenie
1	Zbiory algorytmów (*ang. Ensemble*)	EN
2	Maszyna wektorów wspierających (*ang. Support Vctor Machine*)	SVM
3	Drzewa decyzyjne (*ang. Tree*)	TREE
4	Sieć neuronowa (*ang. Neural Network*)	NN
5	Regresja z użyciem procesu Gaussa (*ang. Gaussian Proces Regression*)	GPR

Wartości parametrów modeli uczenia maszynowego zastosowanych pracy są następujące:

1. EN – Model regresji utworzony za pomocą zespołowego algorytmu uczenia maszynowego (*ang. ensemble learning*) wykorzystuje algorytm LSBoost (*ang. Least Squares Boosting*). Algorytm uczący zbudowany jest ze zbioru drzew decyzyjnych, które są trenowane przez 24 cykle na podstawie danych treningowych. Każde drzewo ma minimalny rozmiar liścia równy 1 (min. jedna obserwacja na liść). Algorytm wybiera w każdym cyklu losowo 16 zmiennych z dostępnych zmiennych objaśniających (tj. dane wejściowe/cechy/predyktory). Wartość współczynnika uczenia dla metody LSBoost, który kontroluje szybkość nauki modelu w kolejnych iteracjach wynosi 0,2984.
2. SVM – Model regresji zbudowany w oparciu u maszynę wektorów nośnych (*ang. Support Vector Machine*) wykorzystuje liniową funkcję jądra, o skali równej 1, co oznacza, że algorytm poszukuje hiperpłaszczyzny separującej dane treningowe. W badaniach zastosowano parametr regularyzacji równy 0,1575 (parametr wyznaczony empirycznie), który kontroluje wpływ błędów trenowania na model. Im mniejsza wartość, tym większe stanowi on ograniczenie dla modelu. Ponadto, w zadaniu przyjęto szerokość marginesu akceptowalnego dla błędów predykcji na poziomie ε = 0,001298 (parametr wyznaczony empirycznie).

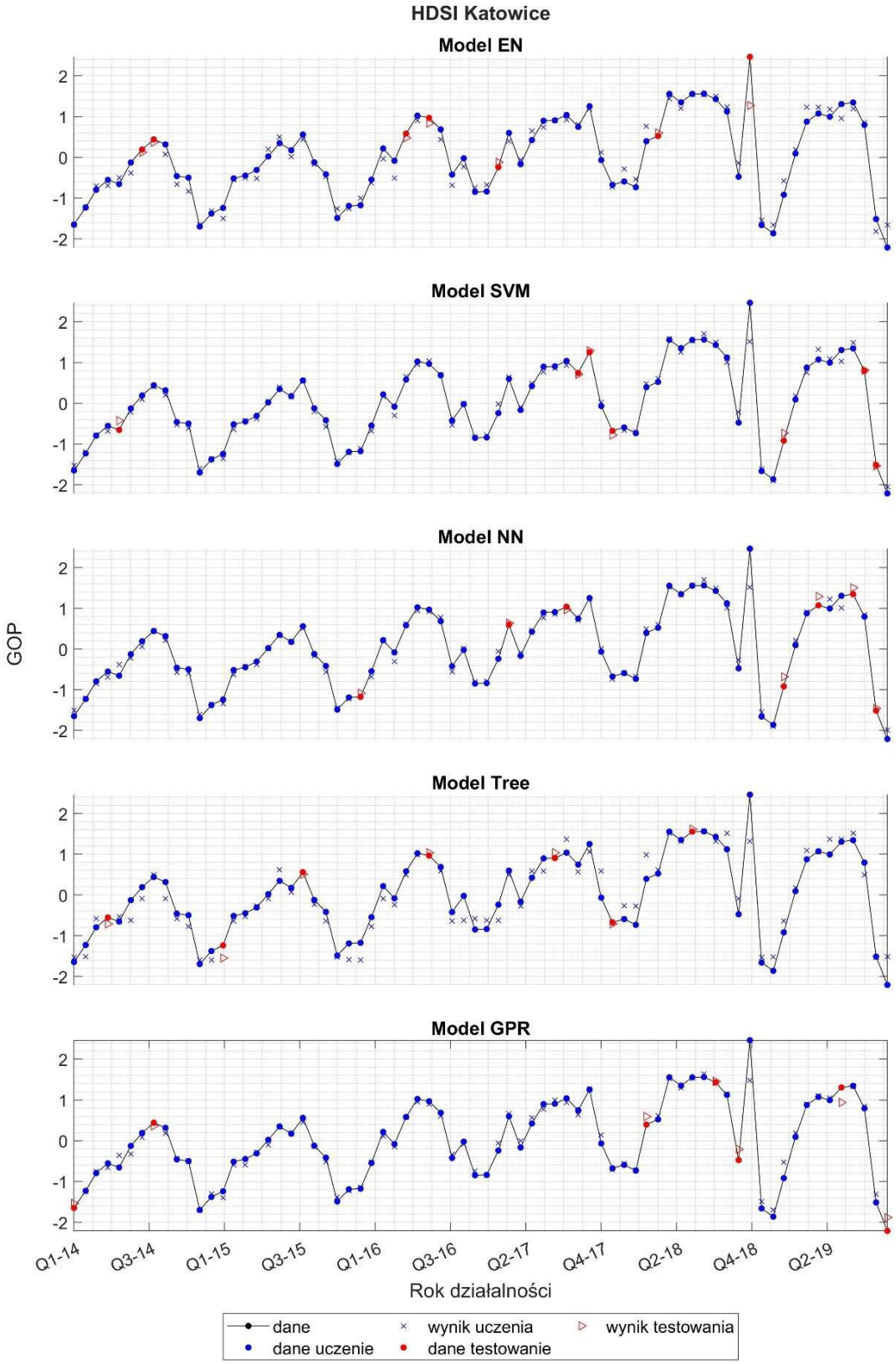
3. TREE – Model regresji oparty o pojedyncze drzewo decyzyjne o minimalnym rozmiarze liścia równym 6, co oznacza, że każdy liść drzewa musi mieć co najmniej 6 obserwacji.
4. NN – Model regresji utworzony z wykorzystaniem sztucznej sieci neuronowej składającej się z dwóch warstw bez funkcji aktywacji, co oznacza liniową naturę modelu. Zastosowano regularyzację L2 wraz z parametrem λ = 0,004513 (parametr wyznaczony empirycznie). Proces uczenia sieci trwał maksymalnie 1000 epok.
5. GPR – Model regresji utworzony za pomocą procesu Gaussa opartym o model jądra eksponencjalnego z diagonalnie skalowalnymi długościami skali (*ang. Automatic Relevance Determination*) o wielkości σ = 9,830 (parametr wyznaczony empirycznie).

Wartości parametrów wszystkich rozpatrywanych modeli dobrano w osobnym procesie optymalizacji, wykonanym w 300–tu iteracjach, dążąc do minimalizacji wartości błędu średnio–kwadratowego (*ang. Mean Square Error* – MSE).

Aby uwiarygodnić uzyskiwane wyniki, proces uczenia i testowania wykonano w 100 iteracjach. Dla każdej z iteracji, zestaw danych dzielono losowo na sekwencje uczącą i testującą w stosunku procentowym 90/10, przy czym 90% stanowiły dane uczące, a 10% – dane wykorzystane na potrzeby testowania. Celem uniknięcia przeuczenia każdorazowo zastosowano metodę walidacji krzyżowej o parametrze k=5.

Na rys. 7.10 przedstawiono zestawienie porównawcze przykładowych wyników dopasowania nieliniowych modeli predykujących zmienną GOP dla kolejnych hoteli. Poszczególne wykresy obrazują wyniki zastosowanych modeli: EN, SVM, NN, TREE, GPR. Na każdym wykresie kolorem czarnym zaznaczono dane rzeczywiste, kolorem niebieskim – dane uczące, a kolorem czerwonym – dane testujące. Kolorową kropką zaznaczono dane wejściowe, krzyżykiem bądź trójkątem – dane predykowane.

Analiza wizualna uzyskanych wyników pozwala na stwierdzanie, że zarówno procedura uczenia, jak i testu daje dobre rezultaty, czyli dane pokrywają się lub są zbliżone.

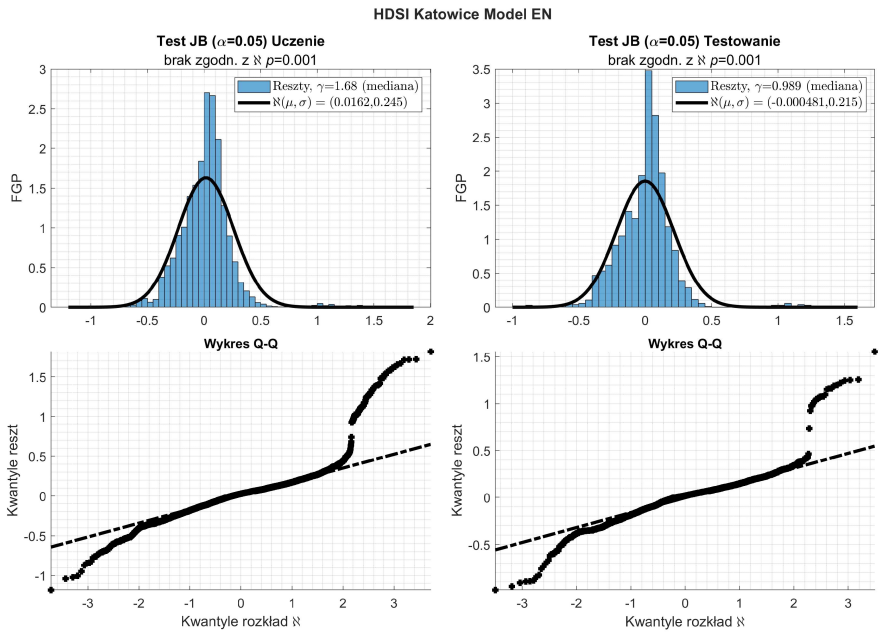


Rys. 7.10 Zestawienie porównawcze przykładowych wyników dopasowania nieliniowych modeli predykujących zmienną GOP dla hotelu HDSI Katowice. Źródło: Opracowanie własne.

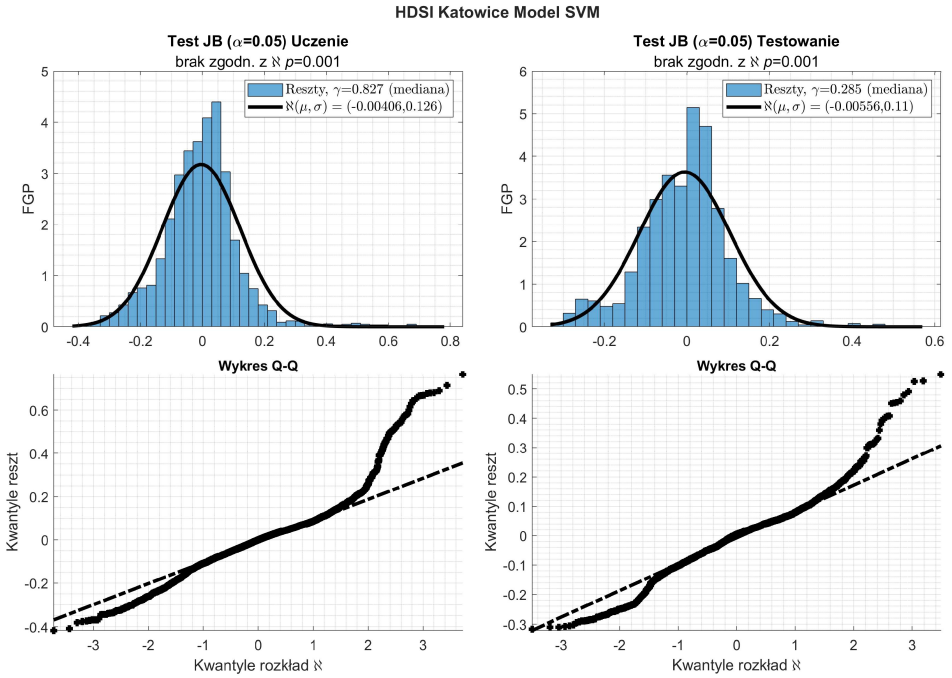
Aby dokonać rzetelnej oceny jakości dopasowania modeli wykreślono rozkłady reszt, obliczono ich zgodności z rozkładem normalnym i symetryczność oraz wykreślono wykresy zależności kwantyli reszt do kwantyli rozkładu normalnego. W dalszej części rozdziału przedstawiono wyniki analiz jakości dopasowania tych modeli, które uzyskano w procesie 100 iteracji uczenia/testowania.

Na rysunkach 7.11 – 7.15 przedstawiono wyniki analizy reszt uzyskanych w wyniku przeprowadzenia 100 iteracji procesu uczenia i testowania modelu predykującego zmienną GOP, zarejestrowaną w hotelu HDSI Katowice, z wykorzystaniem modeli EN, SVM, TREE, NN i GPR odpowiednio. Rysunki zawierają cztery wykresy. Dwa z nich przedstawiają rozkłady reszt wraz z dopasowaniem FGP o rozkładzie normalnym, procentowym wynikiem zgodności kolejnych iteracji wykonanym testem Jarque'a-Berry oraz wartość mediany obliczoną dla skośności γ – dla procesu uczenia i testowania odpowiednio. Kolejne dwa wykresy stanowią wizualizację zależności kwantyli reszt do kwantyli rozkładu normalnego dla procesówuczenia i testowania odpowiednio. Analiza wykresów Q–Q wykazała, że w większości modeli reszty wskazują na duże niedoszacowanie dla wartości bardzo małych i przeszacowanie dla wartości bardzo dużych. Lepsze dopasowanie dla wartości bardzo małych uzyskano przy zastosowaniu modeli TREE, SVM i NN. W następnej kolejności reszty poddano kolejnym badaniom, w tym obliczono skośność, homoskedastyczność i autokorelację, wyniki zawarto w tabelach 28 i 29.

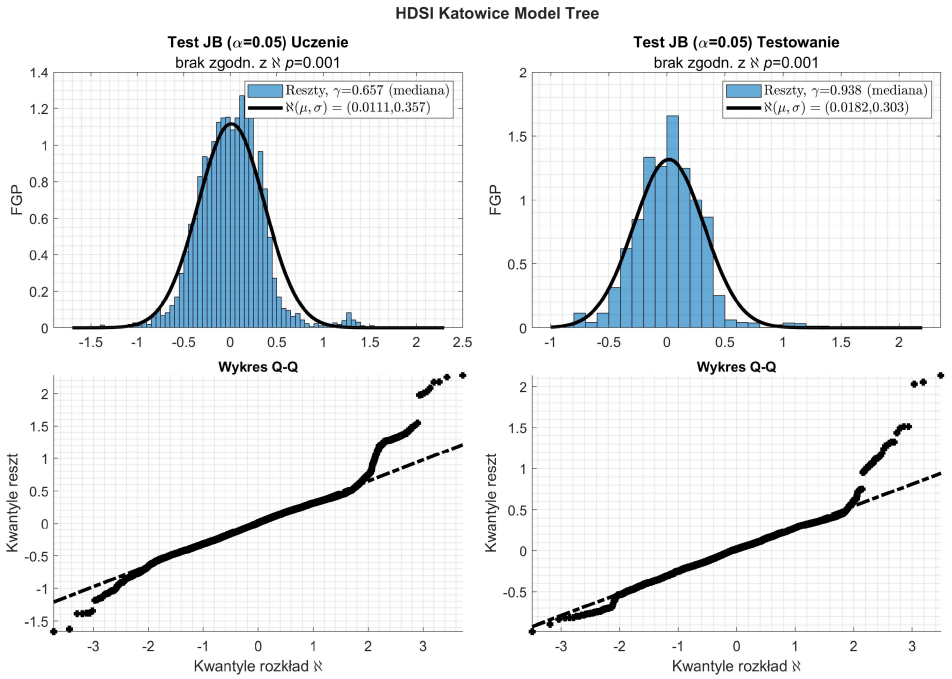
Wyniki analizy reszt dla pozostałych hoteli zestawiono w Załączniku Z7.



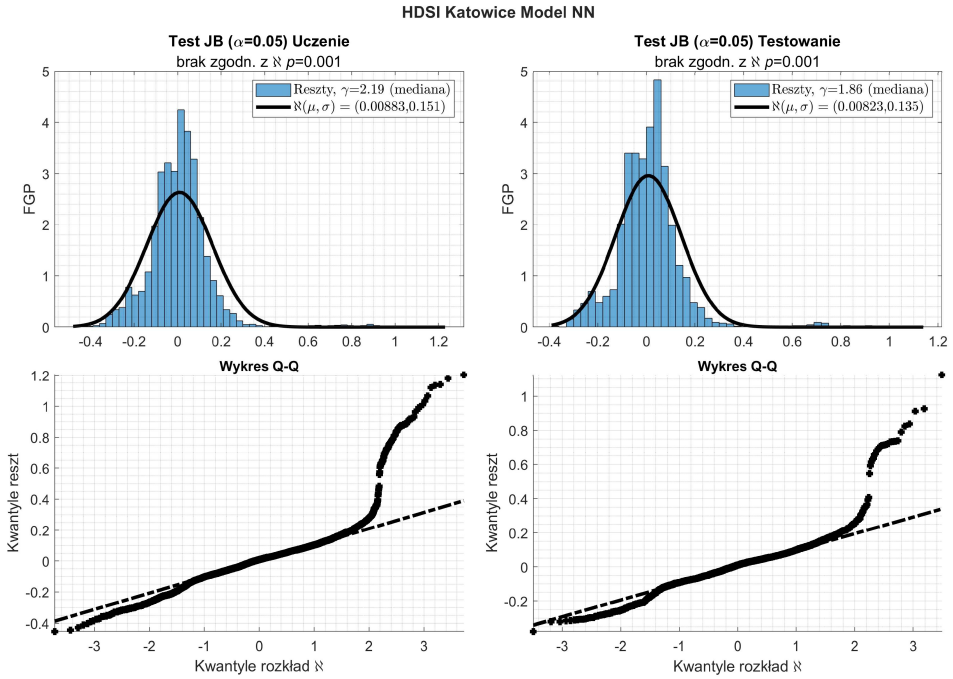
Rys. 7.11 Zestawienie porównawcze rozkładów gęstości prawdopodobieństwa reszt wraz z wykresami Q–Q, przeprowadzone w 100 iteracjach uczenia i testowania modelu predykującego GOP dla hoteli HDSI Katowice, z wykorzystaniem algorytmu EN. Źródło: Opracowanie własne.



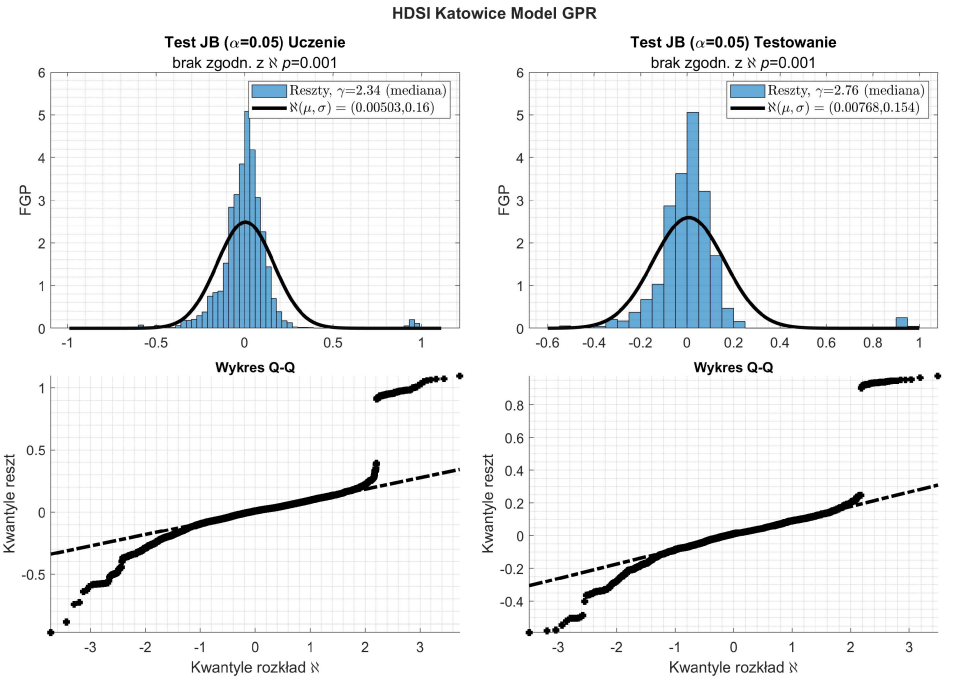
Rys. 7.12 Zestawienie porównawcze rozkładów gęstości prawdopodobieństwa reszt wraz z wykresami Q–Q, przeprowadzone w 100 iteracjach uczenia i testowania modelu predykującego GOP dla hoteli HDSI Katowice, z wykorzystaniem algorytmu SVM. Źródło: Opracowanie własne.



Rys. 7.13 Zestawienie porównawcze rozkładów gęstości prawdopodobieństwa reszt wraz z wykresami Q–Q, przeprowadzone w 100 iteracjach uczenia i testowania modelu predykującego GOP dla hoteli HDSI Katowice, z wykorzystaniem algorytmu TREE. Źródło: Opracowanie własne.



Rys. 7.14 Zestawienie porównawcze rozkładów gęstości prawdopodobieństwa reszt wraz z wykresami Q–Q, przeprowadzone w 100 iteracjach uczenia i testowania modelu predykującego GOP dla hoteli HDSI Katowice, z wykorzystaniem algorytmu NN. Źródło: Opracowanie własne.



Rys. 7.15 Zestawienie porównawcze rozkładów gęstości prawdopodobieństwa reszt wraz z wykresami Q–Q, przeprowadzone w 100 iteracjach uczenia i testowania modelu predykującego GOP dla hoteli HDSI Katowice, z wykorzystaniem algorytmu GPR. Źródło: Opracowanie własne.

W tabeli 21 i 22 zestawiono w sposób zbiorczy wyniki analizy modelowania z wykorzystaniem modeli uczenia maszynowego typu EN, SVM, NN, TREE i GPR, dla wyników procesu trenowania i testowania odpowiednio. W szczególności są to wartości mediany ze 100 iteracji miar R^2, adjR^2 oraz miary symetrii (skośności) dla procesu uczenia i dla procesu testowania danego modelu odpowiednio w tabelach 21 i 22. Ponadto w tabelach

zestawiono procentową wielkość (jako procent ze 100 iteracji) zgodności z rozkładem normalnym (test Jarque'a-Berry), homoskedastyczności (test ARCH Engle'a) i braku autokorelacji rzędu 1-go (Test Darwina-Whatsona), które obliczono dla reszt uzyskanych w obu procesach uczenia i testowania kolejnych modeli odpowiednio w tabelach 21 i 22. Kolorem zielonym zaznaczono wiersze z najlepszymi dla każdego hotelu wynikami modelowania.

Tabela 21. Zestawienie zbiorcze wyników procesu trenowania modeli uczenia maszynowego typu EN, SVM, NN, TREE, GPR, dla wszystkich hoteli. Źródło: Opracowanie własne.

Hotel	Algorytm	Proces trenowania w 100 iteracjach					
		adjR^2 (mediana)	R^2 (mediana)	Skośność (mediana)	% zgodności reszt z rozkładem א	% zgodności reszty homoske-dastyczne	% zgodności reszty nieskorelowane
HDSI Katowice	EN	0,941	0,927	1,422	96	16	90
	SVM	0,984	0,980	0,763	80	28	86
	NN	0,978	0,972	2,063	89	25	89
	TREE	0,876	0,845	0,608	95	30	79
	GPR	0,972	0,965	2,446	98	8	88
HDS Piaseczno	EN	0,947	0,931	0,359	95	62	19
	SVM	0,972	0,963	0,440	100	59	92
	NN	0,973	0,964	0,368	96	91	95
	TREE	0,916	0,890	0,630	99	41	70
	GPR	0,966	0,955	0,542	91	31	12
HDSP Opole	EN	0,911	0,878	0,424	97	46	82
	SVM	0,961	0,947	0,284	94	79	52
	NN	0,968	0,956	0,228	97	90	49
	TREE	0,795	0,720	0,365	100	60	78
	GPR	0,949	0,931	0,788	100	7	74

Hotel	Algorytm	Proces trenowania w 100 iteracjach					
		adjR^2 (mediana)	R^2 (mediana)	Skośność (mediana)	% zgodności reszt z rozkładem א	% zgodności reszty homoske-dastyczne	% zgodności reszty nieskorelowane
HDSP Poznań	EN	0,831	0,155	0,372	99	86	69
	SVM	0,892	0,462	0,735	100	60	71
	NN	0,920	0,599	0,736	98	64	79
	TREE	0,708	NaN	0,208	95	99	45
	GPR	0,943	0,714	0,588	92	57	48
HM Opole	EN	0,887	0,774	0,315	98	83	79
	SVM	0,902	0,804	0,292	97	92	93
	NN	0,911	0,821	0,464	98	73	96
	TREE	0,687	0,374	0,236	99	90	89
	GPR	0,911	0,822	0,487	100	79	70
HMC Częstochowa	EN	0,814	NaN	0,687	96	49	85
	SVM	0,867	0,390	0,276	99	89	74
	NN	0,877	0,435	0,377	98	76	74
	TREE	0,541	NaN	0,579	97	66	70
	GPR	0,920	0,630	0,928	98	41	96
HM Karpacz	EN	0,910	0,669	0,314	98	90	89
	SVM	0,944	0,793	0,305	97	91	87
	NN	0,950	0,818	0,332	97	92	85
	TREE	0,629	NaN	0,272	97	96	91
	GPR	0,955	0,835	0,520	94	60	83

Analiz wskaźników jakości, rozkładów reszt i parametrów symetryczności, zestawionych tabelarycznie dla wyników uczenia pozwoliła na wskazanie co następuje:

- W przypadku hotelu HDSI Katowice najlepiej sprawdził się model SVM, osiągając najwyższe wartości adjR^2 (0,984) i R^2 (0,980), co wskazuje na bardzo dobre dopasowanie do danych. Najsłabiej wypadł model TREE, którego wartości adjR^2 i R^2 wynosiły odpowiednio 0,876 i 0,845, znacząco odstając od pozostałych modeli. Dla hotelu HDS Piaseczno najlepsze wyniki osiągnął model NN z adjR^2 równym 0,973 oraz R^2 równym 0,964. Jest to wyraźnie najwyższy poziom dopasowania dla tego zestawu danych. Model TREE ponownie okazał się najsłabszy, z adjR^2 na poziomie 0,916 i R^2 równym 0,890. W hotelu HDSP Opole ponownie najlepsze wyniki uzyskał model NN, którego wartości adjR^2 i R^2 wyniosły odpowiednio 0,968 i 0,956. TREE okazał się tu również najsłabszy, uzyskując adjR^2 na poziomie 0,795 i R^2 równym 0,720, co było znacząco niższe niż w przypadku pozostałych modeli. W przypadku HDSP Poznań najlepszym modelem okazał się GPR, mimo że jego R^2 (0,714) było stosunkowo niskie. TREE ponownie uzyskał najgorszy wynik, z adjR^2 równym 0,708 i wartością R^2 nieokreśloną (NaN). Dla hotelu HM Opole najlepiej wypadł model NN, osiągając adjR^2 równe 0,911 i R^2 na poziomie 0,821. Model TREE zajął ostatnie miejsce z adjR^2 wynoszącym 0,687 i R^2 na poziomie 0,374. W hotelu HMC Częstochowa model GPR uzyskał najlepsze wyniki (adjR^2: 0,920, R^2: 0,630), natomiast TREE ponownie najsłabiej z dopasowaniem na poziomie adjR^2 równym 0,541 i wartością R^2 równą NaN. Na koniec, dla hotelu HM Karpacz, najlepiej sprawdził się model GPR z adjR^2 na poziomie 0,955 i R^2 równym 0,835, podczas gdy TREE uzyskał najgorszy wynik, osiągając adjR^2 równy 0,629 i brak wyniku dla R^2. Podsumowując, modele NN i SVM okazały się najskuteczniejsze, osiągając wysokie wartości dopasowania w większości analizowanych hoteli. Model TREE natomiast wielokrotnie wypadał najsłabiej, uzyskując najniższe wartości adjR^2 i R^2 lub ich brak, co sugeruje jego niższą trafność w dopasowaniu do danych w porównaniu do innych modeli.
- Skośność przekracza wartość 1 dla kilku modeli, co wskazuje na wyraźniejszą asymetrię w rozkładzie reszt: HDSI Katowice – EN (1,422), NN (2,063), GPR (2,446) oraz HMC Częstochowa – EN (0,687), GPR (0,928). Modele TREE i SVM wielokrotnie uzyskiwały najniższe wartości, wskazując na bardziej symetryczny rozkład reszt, podczas gdy model GPR zazwyczaj charakteryzował się najwyższą skośnością, co sugeruje większą asymetrię w rozkładzie reszt.
- Analizując procenty zgodności reszt z rozkładem normalnym, dla wyników uczenia można zauważyć, że modele z wysoką zgodnością (powyżej 90%) występują w kilku przypadkach,

HDSI Katowice: Najwyższą zgodność osiągnął model GPR (98%), a także EN (96%) i TREE (95%). HDS Piaseczno: Model SVM uzyskał pełną zgodność (100%), natomiast pozostałe modele także osiągnęły wysokie wyniki, bliskie 95% (EN – 95%, NN – 96%, TREE – 99%). HDSP Opole: Modele TREE i GPR uzyskały najwyższą zgodność (100%), z dobrą zgodnością dla EN i NN (97%). HDSP Poznań: Najlepiej wypadł model SVM (100%), a modele TREE i NN również uzyskały wysoką zgodność, odpowiednio 95% i 98%. HM Opole: Model GPR osiągnął pełną zgodność (100%), z wysokimi wynikami dla NN (98%), EN (98%), i TREE (99%). HMC Częstochowa: Model SVM osiągnął 99% zgodności, a pozostałe modele, takie jak NN i TREE, również miały dobre wyniki, odpowiednio 98% i 97%. HM Karpacz: Najwyższe wyniki zgodności osiągnęły modele GPR (94%) i EN (98%). W sumie, najwyższą zgodność z rozkładem normalnym uzyskują modele SVM i GPR, z wartościami sięgającymi nawet 100% w kilku hotelach. Modele te pokazują dobrą stabilność pod kątem zgodności z rozkładem normalnym, co sugeruje ich większą przydatność w przypadku danych o rozkładzie bliskim normalnemu.

- Najlepszą zgodność rozkładu reszt z rozkładem normalnym (100%) oraz idealną symetrię ($\gamma = 0$) uzyskano dla algorytmu SVM w predykcji GOP dla HDS Piaseczno. Dobry rezultat ($\gamma = -0{,}01$, 95%) uzyskano również dla algorytmu SVM w hotelu HM Opole. Najczęściej wysoką zgodność rozkładu reszt z rozkładem normalnym (wartość medianowana poziomie 87%) uzyskiwano dla algorytmu TREE. Rozpatrując te wartości dla poszczególnych hoteli, to najczęściej normalne rozkłady reszt uzyskiwano dla hotelu HM Opole. Najgorszą asymetrię rozkładu reszt (γ = 2,446) uzyskano dla algorytmu GPR w hotelu HDSI Katowice.
- Modele, które częściej wykazują heteroskedastyczność (czyli niskie procenty zgodności z homoskedastycznością), to: HDSI Katowice: Model GPR wykazuje heteroskedastyczność z wynikiem 8%, a także EN (16%) i SVM (28%). HDS Piaseczno: Model GPR (31%) oraz TREE (41%) częściej wskazują na heteroskedastyczność. HDSP Opole: GPR wykazuje heteroskedastyczność (7%). HDSP Poznań: TREE pokazuje wyraźną heteroskedastyczność (99%), a także GPR z wynikiem 57%. HM Opole: GPR (79%) oraz TREE (90%) wskazują na niższą zgodność z homoskedastycznością. HMC Częstochowa: GPR wykazuje niższą zgodność z homoskedastycznością (41%). HM Karpacz: GPR (60%) i TREE (96%) są częściej heteroskedastyczne. Ogólnie rzecz biorąc, model GPR wykazuje heteroskedastyczność w większości hoteli, a model TREE również często odznacza się niższą zgodnością z homoskedastycznością, co może sugerować, że reszty w tych modelach są bardziej podatne na zmienność w zależności od wartości predykcyjnych.

- Analiza pod kątem występowania autokorelacji reszt, czyli najniższej zgodności reszt z brakiem autokorelacji, trzy najgorsze wyniki osiągnęły: HDS Piaseczno, EN – 19% zgodności, wskazując na wysoką autokorelację reszt. HDSP Opole, NN – 49% zgodności, co również oznacza znaczną autokorelację i HDSP Poznań, TREE – 45% zgodności, także wskazując na wyraźną autokorelację. Te modele mają najmniejsze wskaźniki zgodności z brakiem autokorelacji, co sugeruje, że reszty w ich przypadku mogą być silnie skorelowane.

Tabela 22. Zestawienie zbiorcze wyników procesu testowania modeli uczenia maszynowego typu EN, SVM, NN, TREE, GPR, dla wszystkich hoteli. Źródło: Opracowanie własne.

Hotel	Algorytm	Proces testowania w 100 iteracjach					
		adjR^2 (mediana)	R^2 (mediana)	Skośność (mediana)	% zgodności reszt z rozkładem א	% zgodności reszty homo-skedastyczne	% zgodności reszty nieskore-lowane
HDSI Katowice	EN	0,960	0,920	0,723	96	57	74
	SVM	0,988	0,976	0,603	96	76	66
	NN	0,987	0,975	0,653	99	67	69
	TREE	0,916	0,832	0,541	95	63	65
	GPR	0,985	0,970	1,028	98	46	71
HDS Piaseczno	EN	0,947	0,868	0,660	99	60	28
	SVM	0,981	0,951	0,419	99	89	76
	NN	0,979	0,946	0,399	99	92	63
	TREE	0,941	0,853	0,459	97	80	50
	GPR	0,970	0,924	0,517	95	71	29
HDSP Opole	EN	0,924	0,785	0,546	97	69	51
	SVM	0,976	0,933	0,455	99	92	39
	NN	0,975	0,929	0,392	95	94	42
	TREE	0,815	0,475	0,657	99	66	65

Hotel	Algorytm	Proces testowania w 100 iteracjach					
		adjR^2 (mediana)	R^2 (mediana)	Skośność (mediana)	% zgodności reszt z rozkładem א	% zgodności reszty homo-skedastyczne	% zgodności reszty nieskore-lowane
	GPR	0,957	0,879	1,175	95	35	41
HDSP Poznań	EN	0,858	NaN	0,390	97	95	NaN
	SVM	0,904	NaN	0,513	100	86	0
	NN	0,934	NaN	0,486	98	93	0
	TREE	0,753	NaN	0,247	96	99	0
	GPR	0,967	NaN	0,405	95	94	0
HM Opole	EN	0,902	NaN	0,323	97	97	0
	SVM	0,910	NaN	0,343	98	91	0
	NN	0,922	NaN	0,391	98	83	0
	TREE	0,734	NaN	0,387	100	96	0
	GPR	0,931	NaN	0,272	98	100	0
HMC Częstochowa	EN	0,813	NaN	0,547	97	84	0
	SVM	0,873	NaN	0,430	97	91	0
	NN	0,902	NaN	0,426	99	89	0
	TREE	0,701	NaN	0,575	100	85	0
	GPR	0,923	NaN	0,652	100	76	0
HM Karpacz	EN	0,925	NaN	0,535	98	95	0
	SVM	0,943	NaN	0,350	98	88	0
	NN	0,956	NaN	0,501	96	89	0
	TREE	0,774	NaN	0,457	100	91	0
	GPR	0,960	NaN	0,626	98	87	0

Analiza wskaźników jakości, rozkładów reszt i parametrów symetryczności, zestawionych tabelarycznie dla wyników testowania pozwoliła na wskazanie co następuje:

- Najwyższe wartości adjR^2 i R^2 uzyskano dla HDSI Katowice, modele SVM, NN i GPR, każdorazowo przekraczające 0,97. Najgorsze R^2 obliczono dla HDSP Opole, model TREE (R^2=0,475) oraz HMS Częstochowa - model TREE (adjR^2 = 0,701). Dla kilku hoteli, ze względu na zbyt mały rozmiar zestawu danych nie udało się obliczyć wartości R^2, a w tabeli zawarto wtedy NaN.
- Wartość skośności nie przekracza jedności w większości przypadków za wyjątkiem: HDSI Katowice z algorytmem GPR - skośność: 1,028 i HDSP Opole z algorytmem GPR - skośność: 1,175. Te przypadki wskazują na większą asymetrię rozkładu reszt niż w pozostałych modelach, co może wpływać na interpretację wyników oraz zgodność z założeniami normalności rozkładu reszt.
- Analiza reszt pod kątem zgodności z rozkładem normalnym wykazała wszędzie dużą zgodność, powyżej 95%.
- Modele, w których reszty są heteroskedastyczne (zgodność z homoskedastycznością poniżej 60%) to: HDSI Katowice z algorytmem EN - zgodność homoskedastyczności: 57% i z algorytmem GPR - zgodność homoskedastyczności: 46%, HDS Piaseczno z algorytmem GPR - zgodność homoskedastyczności: 50% oraz HDSP Opole z algorytmem GPR - zgodność homoskedastyczności: 35%. Te modele wskazują na potencjalne problemy z założeniem homoskedastyczności, co może wpływać na stabilność estymacji.
- Natomiast analiza autokorelacji reszt wykazała, że modele, dla których nie udało się obliczyć wartości R^2, osiągały zerową zgodność reszt nieskorelowanych. Najmniej skorelowane reszty uzyskano dla hotelu HDSI Katowice, dla modelu EN, GPR i SVM, gdzie procent zgodności przekroczył 71%.

Rozdział przedstawił proces modelowania zysku operacyjnego brutto za pomocą dwóch podejść: modelu ekonometrycznego oraz nieliniowych algorytmów regresyjnych i uczenia maszynowego. Na początku opisano modelowanie ekonometryczne GOP, obejmujące wyniki selekcji zmiennych metodą eliminacji wstecznej dla poszczególnych hoteli. Wskazano model początkowy i końcowy, prezentując wartości estymowanych parametrów i odpowiednie miary oceny procesu doboru modelu ekonometrycznego. W dalszej części rozdziału opisano proces budowy modelu predykcyjnego z wykorzystaniem nieliniowych algorytmów regresyjnych i technik uczenia maszynowego. Przedstawiono wyniki analizy ważności wybranych zmiennych USALI, korelujących z GOP, określone metodą MRM i testem F. Następnie, w oparciu o wyselekcjonowane zmienne, zbudowano nieliniowe modele GOP, których rezultaty przedstawiono w podziale na procesy uczenia i testowania, w formie tabelarycznej i graficznej. Całość zapewnia kompleksowe porównanie wyników obu podejść modelowania GOP, oferując cenne wnioski na temat ich skuteczności oraz charakterystyki zastosowanych predyktorów.

8. Podsumowanie i wnioski

W ramach przeprowadzonych prac badawczych dokonano obszernego przeglądu literatury związanej z rachunkowością, wskaźnikami ekonomicznymi oraz możliwościami ich przewidywania w kontekście branży hotelarskiej. Oceniono, że mimo licznych prac naukowych nad tymi zagadnieniami, istnieje potrzeba dalszego doskonalenia narzędzi i metod zarządzania hotelami, ze szczególnym naciskiem na maksymalizację zysków.

Kolejnym etapem prac było zebranie i analiza danych finansowych z siedmiu różnych hoteli w Polsce. Badania obejmowały wskaźniki finansowe zgodne z bazą USALI. Wykorzystano różnorodne techniki analizy danych, takie jak analiza czasowa, periodyczność, rozkłady funkcji gęstości prawdopodobieństwa, analiza sezonowości i cykliczności w dziedzinie częstotliwościowej, analiza współzależności między wskaźnikami, a zyskiem operacyjnym brutto (GOP), analiza korelacji liniowej oraz analiza koherencji falkowej w dziedzinie czasowo-częstotliwościowej. Ponadto, oceniono istotność poszczególnych wskaźników dla możliwości predykcyjnych GOP za pomocą zaawansowanych technik statystycznych, takich jak *Minimum Redundancy Maximum Relevance* i test Fishera-Snedecora F.

Następnie, na podstawie wyników analiz, wyselekcjonowano zestaw wskaźników istotnie skorelowanych z zmienną GOP, ale nieskorelowanych wzajemnie. Te istotne zmienne posłużyły do stworzenia modelu ekonometrycznego oraz modeli sztucznej inteligencji, które miały na celu przewidywanie wartości GOP, dla każdego hotelu oddzielnie.

Na podstawie przeprowadzonych analiz wyników predykcji z wykorzystaniem modelu ekonometrycznego i modelu uczenia maszynowego, wykazano że obie metody są skuteczne w predykowaniu zmiennej GOP. Porównanie modeli ekonometrycznych i uczenia maszynowego jest możliwe, choć każde z tych podejść ma odmienny charakter i zastosowanie. Modele ekonometryczne, bazując na założeniach przyczynowości, są bardziej interpretowalne i transparentne. Z kolei metody sztucznej inteligencji to podejście oparte na danych, które pozwala wykrywać skomplikowane wzorce w dużych zbiorach danych dzięki modelom nieliniowym, takim jak sieci neuronowe. Choć dają precyzyjne prognozy, modele SI są zazwyczaj mniej interpretowalne, co sprawia, że bywają postrzegane jako „czarne skrzynki". Różnice w interpretowalności i zastosowaniu obu podejść widać również w sposobie walidacji. Ekonometria polega na analizie statystycznej reszt i testach założeń, takich jak heteroskedastyczność, autokorelacja czy normalność reszt, co pomaga w ocenie poprawności modelu. W SI walidacja polega na analizie metryk wydajności (jak dokładność) na zbiorach treningowych, walidacyjnych i testowych, gdzie liczy się głównie dokładność prognozy. Mimo

tych różnic oba podejścia mogą się uzupełniać: SI może wskazać nieoczywiste zależności w dużych zbiorach danych, które następnie można pogłębić i interpretować przy użyciu modeli ekonometrycznych.

Założone cele pracy zrealizowano w pełni, na podstawie przeprowadzonych analiz empirycznych oraz uzyskanych wyników możemy potwierdzić każdą z postawionych na wstępie hipotez badawczych, co prowadzi do następujących wniosków:

- Potwierdzenie Hipotezy 1: Przeprowadzone analizy statystyczne wykazały istotną zależność między wybranymi wskaźnikami USALI, a zyskiem operacyjnym brutto w rozpatrywanych hotelach. Analiza korelacji oraz modelowanie regresyjne jednoznacznie wskazały, że wybrane wskaźniki USALI, wykazują korelację z wartością GOP przekraczającą 0,8, co potwierdza hipotezę o ich kluczowym wpływie na wyniki finansowe przedsiębiorstw hotelarskich. Ponadto wykazano, że wynik operacyjny (zmienna X52) koreluje z GOP we wszystkich hotelach, natomiast pozostałe zmienne USALI korelują w sposób zróżnicowany.
- Potwierdzenie Hipotezy 2: Badanie potwierdziło, że zastosowanie nowoczesnych technologii sztucznej inteligencji pozwala na uzyskanie porównywalnie precyzyjnych prognoz zysku operacyjnego brutto (GOP), w porównaniu do metod statystycznych opartych na klasycznych technikach statystycznych (ekonometrycznych). Wykorzystanie modeli uczenia maszynowego, takich jak sieci neuronowe i maszyny wektorów nośnych, eliminuje konieczność spełnienia założeń wymaganych przez klasyczne metody statystyczne, przy jednoczesnym zapewnieniu porównywalnej lub wyższej dokładności wyników.
- Potwierdzenie Hipotezy 3: Wyniki analiz potwierdziły skuteczność wyselekcjonowanego zestawu wskaźników finansowych jako predyktorów zysku operacyjnego brutto w rozpatrywanych hotelach. Wskaźniki te, dzięki silnej korelacji z GOP, mogą być stosowane do skutecznej predykcji wyników operacyjnych w branży hotelarskiej. Przeprowadzone badania predykcyjne wskazały, że uwzględnienie tych wskaźników znacząco poprawia dokładność przewidywań, co potwierdza hipotezę.
- Potwierdzenie Hipotezy 4: Analiza szeregów czasowych oraz analiza sezonowości potwierdziły istnienie cyklicznych wzorców w wynikach operacyjnych hoteli w Polsce. Wyniki wskazują na istotne wzorce sezonowe związane z określonymi okresami roku, które wpływają na poziom GOP.

Poniżej określono możliwe dalsze kierunki prowadzenia prac badawczych:

- Doskonalenie modeli predykcyjnych. Kontynuacja prac nad rozwojem zaawansowanych modeli uczenia maszynowego w kontekście GOP może przyczynić się do bardziej

efektywnego zarządzania hotelami, zwiększenia przychodów i poprawy doświadczeń klientów. Jest to obszar o dużym potencjale, który wymaga interdyscyplinarnych badań i stałego dostosowywania do zmieniających się warunków rynkowych. Możliwie kierunki prac badawczych to:

- Badanie możliwości zastosowania głębokich sieci neuronowych, takich jak konwolucyjne sieci neuronowe (CNN), sieci głębokie (*ang. Deep Learning*) lub rekurencyjne sieci neuronowe (RNN), do analizy danych związanymi z GOP. Eksperymenty z różnymi architekturami sieci, funkcjami aktywacji i technikami regularyzacji w celu poprawy precyzji prognoz.
- Poszukiwanie nowych źródeł danych, które mogą być używane do trenowania modeli, takich jak dane pogodowe, dane z mediów społecznościowych lub dane z okolicznych atrakcji turystycznych. Integracja danych z wielu źródeł w celu uzyskania bardziej kompleksowego obrazu i poprawy jakości prognoz GOP.
- Zastosowanie algorytmów RL (*ang. Reinforcement Learning*) do optymalizacji zarządzania dostępnymi pokojami w hotelu na podstawie prognoz GOP. Opracowanie agentów RL, którzy podejmują decyzje w czasie rzeczywistym, aby zoptymalizować przychody hotelu.
- Badanie, w jaki sposób interakcje klientów z systemem rezerwacji i innymi usługami hotelu mogą być wykorzystane do doskonalenia modeli GOP. Opracowanie personalizowanych prognoz opartych na zachowaniach klientów.
- Rozważenie aspektu zrównoważonego zarządzania hotelem poprzez uwzględnienie prognoz GOP w kontekście optymalizacji zużycia energii i innych zasobów.
- Opracowanie technik zarządzania niepewnością w prognozach GOP, co może być istotne ze względu na zmienne warunki rynkowe. Eksploracja podejść takich jak bayesowskie sieci neuronowe lub metody bootstrapingu.

- Tworzenie narzędzi do wizualizacji prognoz GOP, które pomogą menedżerom hoteli lepiej zrozumieć i wykorzystać te prognozy w codziennym zarządzaniu.
- Praca nad modelami, które są w stanie uczyć się na bieżąco, korzystając z danych strumieniowych, takich jak rezerwacje hotelowe dokonywane przez klientów w czasie rzeczywistym.
- Porównanie różnych technik i modeli prognozowania GOP pod kątem ich skuteczności i efektywności. Opracowanie metryk oceny wydajności, które odzwierciedlają rzeczywisty wpływ prognoz na zarządzanie hotelem.

- Implementacja w praktyce. Przeniesienie wyników badań do praktyki oraz wprowadzenie nowych narzędzi i metod zarządzania w hotelach, a także monitorowanie ich skuteczności, może otworzyć wiele perspektywicznych kierunków dalszych prac. Oto kilka możliwych obszarów, które można by eksplorować:
  - Dalsze badania nad rozwojem innowacyjnych narzędzi i systemów zarządzania hotelami, uwzględniając najnowsze technologie. Tworzenie interaktywnych platform i aplikacji mobilnych, które ułatwiają zarządzanie hotelem oraz komunikację z klientami. Badania nad sztuczną inteligencją i automatyzacją procesów, aby zoptymalizować operacje hotelu.
  - Badanie efektywności nowych narzędzi i metod zarządzania w kontekście zmniejszenia kosztów operacyjnych hotelu. Analiza, jak te narzędzia mogą przyczynić się do zwiększenia wydajności personelu oraz poprawy obsługi klienta.
  - Badania nad tym, w jaki sposób nowe technologie i metody zarządzania mogą wpłynąć na poprawę doświadczeń gości w hotelach. Monitorowanie opinii i satysfakcji klientów po wprowadzeniu zmian, aby dostosowywać strategię zarządzania.
  - Badania nad wpływem nowych narzędzi i praktyk zarządzania na zrównoważony rozwój hoteli, włączając aspekty ekologiczne i społeczne. Opracowywanie strategii zmierzających do zminimalizowania wpływu działalności hotelowej na środowisko.
  - Badanie potencjalnych zagrożeń związanych z wprowadzaniem nowych narzędzi i technologii w hotelach oraz opracowywanie strategii zarządzania ryzykiem. Monitorowanie i ocena systemów bezpieczeństwa danych w kontekście przechowywania informacji o klientach.
  - Badanie, jak wprowadzenie nowych narzędzi wpływa na potrzeby szkoleniowe personelu hotelowego. Opracowywanie programów szkoleniowych, które pomogą pracownikom zrozumieć i efektywnie wykorzystywać nowe narzędzia.

- Porównywanie wyników wdrożonych nowych narzędzi i praktyk zarządzania z wynikami innych hoteli w branży. Badania porównawcze jako narzędzie do oceny skuteczności wprowadzonych zmian.
- Badanie, jak nowoczesne technologie, takie jak Internet rzeczy (*ang. IoT*) czy analiza big data, mogą być wykorzystane do dalszego usprawnienia zarządzania hotelami.

9. Literatura

Aguwa, C., Olya, M. H., & Monplaisir, L. (2017). Modeling of fuzzy-based voice of customer for business decision analytics. *Knowledge-Based Systems*, *125*, 136–145. https://doi.org/10.1016/j.knosys.2017.03.019

Amit, K., Mozar, S., & Haase, J. (2023). Advances in Cognitive Science and Communications. W A. Kumar, S. Mozar, & J. Haase (Red.), *Advances in Cognitive Science and Communications Selected Articles from the 5th International Conference on Communications and Cyber-Physical Engineering (ICCCE 2022)*. Singapore: Springer Nature Singapore. https://doi.org/10.1007/978-981-19-8086-2

Ampountolas, A. (2021). Modeling and forecasting daily hotel demand: a comparison based on SARIMAX, Neural Networks, and GARCH models. *Forecasting 2021, Vol. 3, Pages 580-595*, *3*(3), 580–595. https://doi.org/10.3390/FORECAST3030037

Assaf, A. G., Josiassen, A., & Cvelbar, L. K. (2012). Does Triple Bottom Line reporting improve hotel performance? *International Journal of Hospitality Management*, *31*(2), 596–600. https://doi.org/10.1016/J.IJHM.2011.08.005

Assaf, A. G., Josiassen, A., Knežević Cvelbar, L., & Woo, L. (2015). The effects of customer voice on hotel performance. *International Journal of Hospitality Management*, *44*, 77–83. https://doi.org/10.1016/j.ijhm.2014.09.009

Assaf, A. G., & Tsionas, M. G. (2019). Forecasting occupancy rate with Bayesian compression methods. *Annals of Tourism Research*, *75*, 439–449. https://doi.org/10.1016/j.annals.2018.12.009

Asteriou, D., & Hall, S. G. (2011). ARIMA Models and the Box–Jenkins Methodology. W *Applied Econometrics* (ss. 265–286). Palgrave MacMillan.

Athiyaman, A., & Robertson, R. W. (1992). Time series forecasting techniques: short-term planning in tourism. *International Journal of Contemporary Hospitality Management*, *4*(4). https://doi.org/10.1108/09596119210018864

Atkinson, H., & Brander Brown, J. (2001). Rethinking performance measures: Assessing progress in UK hotels. *International Journal of Contemporary Hospitality Management*, *13*(3), 128–136. https://doi.org/10.1108/09596110110388918/FULL/XML

Barburski, J. (2010). Ekonometryczny pomiar efektywności ekonomicznej instytucji finansowych. Stochastyczny model graniczny kosztów. *Bank i Kredyt*, *41*(1), 31–56.

Barnes, S. J., Mattsson, J., Sørensen, F., & Jensen, J. F. (2020). Measuring employee-tourist encounter experience value: a big data analytics approach. *Expert Systems with*

Applications, *154*, 113450. https://doi.org/10.1016/j.eswa.2020.113450

Bednarska, M. A. (2014). Analiza działalności operacyjnej w hotelarstwie dla potrzeb wyceny nieruchomości. *Biuletyn Stowarzyszenia Rzeczoznawców Majątkowych Województwa Wielkopolskiego*.

BnB_Holiday. (2021). ARPAR Calculator. Pobrano 21 wrzesień 2023, z https://bnbholiday.com/calculators/arpar

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*(3), 307–327. https://doi.org/10.1016/0304-4076(86)90063-1

Borzyszkowski, J. (2014). Zjawisko sezonowości w turystyce – istota problemu i działań krajowych podmiotów polityki turystycznej na rzecz jego ograniczania. *Rozprawy Naukowe Akademii Wychowania Fizycznego we Wrocławiu*, *45*, 167–176.

Bragg, S. M. . (2011). *Business ratios and formulas : a comprehensive guide*. Wiley.

Bull, A. D. (2011). Convergence rates of efficient global optimization algorithms. *Journal of Machine Learning Research*, *12*, 2879–2904.

Butler, D., & Schütze, M. (2005). Integrating simulation models with a view to optimal control of urban wastewater systems. *Environmental Modelling and Software*, *20*(4 SPEC. ISS.), 415–426. https://doi.org/10.1016/j.envsoft.2004.02.003

Calero-Sanz, J., Orea-Giner, A., Villacé-Molinero, T., Muñoz-Mazón, A., & Fuentes-Moraleda, L. (2022). Predicting a new hotel rating system by analysing UGC content from Tripadvisor: machine learning application to analyse service robots Influence. W *Procedia Computer Science* (T. 200, ss. 1078–1083). https://doi.org/10.1016/J.PROCS.2022.01.307

Casidy, R., Wymer, W., & O'Cass, A. (2018). Enhancing hotel brand performance through fostering brand relationship orientation in the minds of consumers. *Tourism Management*, *66*, 72–84. https://doi.org/10.1016/j.tourman.2017.11.008

Chen, S., Ngai, E. W. T., Ku, Y., Xu, Z., Gou, X., & Zhang, C. (2023). Prediction of hotel booking cancellations: Integration of machine learning and probability model based on interpretable feature interaction. *Decision Support Systems*, *170*, 113959. https://doi.org/10.1016/J.DSS.2023.113959

Cote, R. (2001). *Accounting for Hospitality Managers*. Educational Inst of the Amer Hotel.

Ding, C., & Peng, H. (2005). Minimum redundancy feature selection from microarray gene expression data. *Journal of Bioinformatics and Computational Biology*, *3*(2), 185–205. https://doi.org/10.1142/S0219720005001004

Dinov, I. D. (2023). *Data Science and Predictive Analytics*. Cham: Springer International

Publishing. https://doi.org/10.1007/978-3-031-17483-4

Dowlut, N., & Gobin-Rahimbux, B. (2023). Forecasting resort hotel tourism demand using deep learning techniques – A systematic literature review. *Heliyon*, *9*(7), e18385. https://doi.org/10.1016/j.heliyon.2023.e18385

Duro, J. A., & Turrión-Prats, J. (2019). Tourism seasonality worldwide. *Tourism Management Perspectives*, *31*, 38–53. https://doi.org/10.1016/J.TMP.2019.03.010

Dyduch, M. (2010). Współczynniki transformaty falkowej jako narzędzie generujące prognozę przedziałową szeregów czasowych. *Prace Naukowe / Akademia Ekonomiczna w Katowicach*, 35–45.

Dyduch, R. (2017). Financial efficiency of structured bonds. *Wiadomości Statystyczne. The Polish Statistician*, *9*(676), 63–78. https://doi.org/10.5604/01.3001.0014.1049

Enz, C. A. (2010). *The Cornell School of Hotel Administration handbook of applied hospitality strategy*. SAGE.

Eurostat. (2023). Seasonality in the tourist accommodation sector. Pobrano 23 wrzesień 2023, z https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Seasonality_in_the_tourist_accommodation_sector

Farrell, M. T. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society. Series A (General)*, *120*(3), 253–290.

Fernandes, E., Moro, S., Cortez, P., Batista, F., & Ribeiro, R. (2021). A data-driven approach to measure restaurant performance by combining online reviews with historical sales data. *International Journal of Hospitality Management*, *94*, 102830. https://doi.org/10.1016/J.IJHM.2020.102830

Filip, P. (2011). Rentowność i efetywność działania przedsiębiorstwa. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, (182).

Fiori, A. M., & Foroni, I. (2020). Prediction accuracy for reservation-based forecasting methods applied in Revenue Management. *International Journal of Hospitality Management*, *84*, 102332. https://doi.org/10.1016/j.ijhm.2019.102332

Foufoula-Georgiou, E., & Kumar, P. (1994). *Wavelets in geophysics*. Academic Press Inc.

Gallego, I., González-Rodríguez, M. R., & Font, X. (2023). A multi-criteria, composite index methodology to measure the suitability of target markets for the hotel industry. *Tourism Management Perspectives*, *47*(August 2022). https://doi.org/10.1016/j.tmp.2023.101104

Gawlik, A., & Puciato, D. (2016). *Nowoczesne finanse przedsiębiorstwa hotelowego*. CeDeWu.

Gelbart, M. A., Snoek, J., & Adams, R. P. (2014). Bayesian optimization with unknown

constraints. W *Uncertainty in Artificial Intelligence - Proceedings of the 30th Conference, UAI 2014* (ss. 250–259). AUAI Press.

Gilli, M., Maringer, D., & Schumann, E. (2019). *Econometric models*. *Numerical Methods and Optimization in Finance*. Academic Press. https://doi.org/10.1016/B978-0-12-815065-8.00028-5

Grinsted, A., Moore, J. C., & Jevrejeva, S. (2004). Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear Processes in Geophysics*, *11*(516), 561–566. https://doi.org/https://doi.10.5194/npg-11-5d-2004

Guilding, C., & Ji, K. M. (2022). Accounting Essentials for Hospitality Managers, Fourth Edition. *Accounting Essentials for Hospitality Managers, Fourth Edition*, 1–380. https://doi.org/10.4324/9781003183334/ACCOUNTING-ESSENTIALS-HOSPITALITY-MANAGERS-CHRIS-GUILDING-KATE-MINGJIE-JI/ACCESSIBILITY-INFORMATION

Gursoy, D., Buttle, F., & Bowie, D. (2022a). Hospitality marketing: Principles and practices. W *Hospitality Marketing: Principles and Practices*. Taylor and Francis. https://doi.org/10.4324/9781003186342

Gursoy, D., Buttle, F., & Bowie, D. (2022b). *Hospitality marketing: Principles and practices*. *Hospitality Marketing: Principles and Practices*. Taylor and Francis. https://doi.org/10.4324/9781003186342

Hadaś, M. (2006). Zastosowanie sieci falkowo-neuronowej do predykcji ekonomicznych szeregów czasowych. *Prace Naukowe Akademii Ekonomicznej we Wrocławiu*, (nr 1112), 69–80.

Harris, P., & Hazzard, P. A. (1992). *Managerial Accounting in the Hospitality Industry*.

Harris, Peter, & Mongiello, M. (2012). *Accounting and financial management: Developments in the international hospitality industry*. *Accounting and Financial Management: Developments in the International Hospitality Industry*. Taylor and Francis. https://doi.org/10.4324/9780080492483

HFTP. (2018). Uniform system of accounts for the lodging Industry. Pobrano 14 wrzesień 2023, z https://www.hftp.org/hospitality_resources/usali_guide/

HospitalityNet. (2022). What is GOPPAR and how do you calculate it? Pobrano 21 wrzesień 2023, z https://www.hospitalitynet.org/explainer/4112220.html

Hu, T., Wang, H., Law, R., & Geng, J. (2023). Diverse feature extraction techniques in internet search query to forecast tourism demand: An in-depth comparison. *Tourism Management Perspectives*, *47*, 101116. https://doi.org/10.1016/J.TMP.2023.101116

Hua, N., Morosan, C., & DeFranco, A. (2015). The other side of technology adoption: examining the relationships between e-commerce expenses and hotel performance. *International Journal of Hospitality Management*, *45*, 109–120. https://doi.org/10.1016/j.ijhm.2014.12.001

Huang, L., & Zheng, W. (2021). Novel deep learning approach for forecasting daily hotel demand with agglomeration effect. *International Journal of Hospitality Management*, *98*, 103038. https://doi.org/10.1016/j.ijhm.2021.103038

Jagels, M., & Ralston, C. E. . (2007). *Hospitality management accounting*. John Wiley & Sons.

Jain, P. K., Pamula, R., & Srivastava, G. (2021). A systematic literature review on machine learning applications for consumer sentiment analysis using online reviews. *Computer Science Review*, *41*, 100413. https://doi.org/10.1016/j.cosrev.2021.100413

Jajuga, K. (1956-)., & Jajuga, T. (1954-). (2006). *Inwestycje: instrumenty finansowe, aktywa niefinansowe, ryzyko finansowe, inżynieria finansowa*. Wydawnictwo Naukowe PWN.

Johnes, J. (2004). *Efficiency measurement*. *International Handbook on the Economics of Education*.

Jones, P., & Lockwood, A. (1999). The management of hotel operations, 182.

Jun, W., Yuyan, L., Lingyu, T., & Peng, G. (2018). Modeling a combined forecast algorithm based on sequence patterns and near characteristics: An application for tourism demand forecasting. *Chaos, Solitons and Fractals*, *108*, 136–147. https://doi.org/10.1016/j.chaos.2018.01.028

Kim, W. G., & Ham, S. (2006). The impact of information technology implementation on service quality in the hotel industry. *Information Technology in Hospitality*, *4*(4), 143–151. https://doi.org/10.3727/154595306779868430

Klimczuk-Kochańska, M. (2016). Wkład Petera F. Druckera w rozwój teorii zarządzania. *Zarządzanie, organizacje i organizowanie – przegląd perspektyw teoretycznych*.

Kotłowska, M. (2014). Finansowe aspekty modelowania procesów w przedsiębiorstwie. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, *335*(126–139).

Kucharczyk, E., & Mitura, M. (2010). *Ekonomia i prawo w hotelarstwie*. Difin.

Kurths, D., & Maraun, J. (2004). Cross wavelet analysis: significance testing and pitfalls. *Nonlinear Processes in Geophysics*, *11*, 505–514. https://doi.org/1607-7946/npg/2004-11-505

Kwon, W. (2023). Reading customers' minds through textual big data: challenges, practical guidelines, and proposals. *International Journal of Hospitality Management*, *111*,

103473. https://doi.org/10.1016/J.IJHM.2023.103473

Le, T. H., Chuc, A. T., & Taghizadeh-Hesary, F. (2019). Financial inclusion and its impact on financial efficiency and sustainability: Empirical evidence from Asia. *Borsa Istanbul Review*, *19*(4), 310–322. https://doi.org/10.1016/J.BIR.2019.07.002

Lee, M. (2018). Modeling and forecasting hotel room demand based on advance booking information. *Tourism Management*, *66*, 62–71. https://doi.org/10.1016/j.tourman.2017.11.004

Li, Y., Chu, X., Tian, D., Feng, J., & Mu, W. (2021). Customer segmentation using K-means clustering and the adaptive particle swarm optimization algorithm. *Applied Soft Computing*, *113*(B), 107924. https://doi.org/10.1016/J.ASOC.2021.107924

Liu, J., Hu, S., Mehraliyev, F., & Liu, H. (2023). Text classification in tourism and hospitality – a deep learning perspective. *International Journal of Contemporary Hospitality Management*. https://doi.org/10.1108/IJCHM-07-2022-0913

Maraun, D., Kurths, J., & Holschneider, M. (2007). Nonstationary Gaussian processes in wavelet domain: synthesis, estimation, and significance testing. *Physical review. E, Statistical, nonlinear, and soft matter physics*, *75*(1 Pt 2), 1–14. https://doi.org/10.1103/PHYSREVE.75.016707

Marcolin, C. B., Becker, J. L., Wild, F., Behr, A., & Schiavi, G. (2021). Listening to the voice of the guest: A framework to improve decision-making processes with text data. *International Journal of Hospitality Management*, *94*, 102853. https://doi.org/10.1016/j.ijhm.2020.102853

Mauri, A. G. (2012). *Hotel revenue management: principles and practices*. *Journal of Revenue and Pricing Management*. Pearson.

McKinsey Digital. (2017). Advanced analytics in hospitality. Pobrano 18 wrzesień 2023, z https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/advanced-analytics-in-hospitality#]

Milewska, M., & Włodarczyk, B. (2018). *Hotelarstwo. Organizacja i funkcjonowanie obiektów noclegowych*. Polskie Wydawnictwo Ekonomiczne.

Mills, T. C. (1990). *Time series techniques for economists*. Cambridge University Press. Mitura, E., & Koniuszewska, E. (2006). *Ekonomika i organizacja pracy w hotelarstwie*. Difin.

Nam, J., Bon Sesay, J., Wynne, K., & Zhang, G. (2020). Financial efficiency and accounting quality: The impact of institutional micro-factors on FDI. *Journal of Policy Modeling*, *42*(2), 451–465. https://doi.org/10.1016/J.JPOLMOD.2019.12.004

Olsen, M. D., & Szabo, K. (2000). Experience-based travel: How technology is changing the

hospitality industry. *The Cornell Hotel and Restaurant Administration Quarterly*, *41*(1), 30–40. https://doi.org/10.1016/S0010-8804(00)88883-7

Pan, B., & Yang, Y. (2016). Forecasting destination weekly hotel occupancy with big data. *Journal of Travel Research*, *56*(7), 957–970. https://doi.org/10.1177/0047287516669050

Pavlatos, O., & Paggios, I. (2009). A survey of factors influencing the cost system design in hotels. *International Journal of Hospitality Management*, *28*(2), 263–271. https://doi.org/10.1016/j.ijhm.2008.09.002

Pereira, L. N. (2016). An introduction to helpful forecasting methods for hotel revenue management. *International Journal of Hospitality Management*, *58*, 13–23. https://doi.org/10.1016/j.ijhm.2016.07.003

Petković, M., Kocev, D., & Džeroski, S. (2020). Feature ranking for multi-target regression. *Machine Learning*, *109*(6), 1179–1204. https://doi.org/10.1007/S10994-019-05829-8/TABLES/7

Petrevska, B. (2015). Effects of tourism seasonality at local level. *Analele Stiintifice ale Universitatii Al I Cuza din Iasi - Sectiunea Stiinte Economice*, *62*(2), 241–250. https://doi.org/10.1515/aicue-2015-0016

Pomykalska, B., & Pomykalski, P. (2017). *Analiza finansowa przedsiębiorstwa*.
Wydawnictwo Naukowe PWN.

Qualpro. (2021). Revenue Management – różnice między dużym, a małym hotelem. Pobrano 21 wrzesień 2023, z https://qualpro.co/roznice_miedzy_malym_a_duzym_hotelem/ Rachuba, J. (2020). Revenue Management w hotelu - teoria a praktyka. Pobrano 21 wrzesień
2023, z https://lepszyhotel.pl/revenue-management-w-hotelu/

Radovic, M., Ghalwash, M., Filipovic, N., & Obradovic, Z. (2017). Minimum redundancy maximum relevance feature selection approach for temporal gene expression data. *BMC Bioinformatics*, *18*(1), 1–14. https://doi.org/10.1186/s12859-016-1423-9

Ramsey, J. B. (1999). The contribution of wavelets to the analysis of economic and financial data. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, *357*(1760), 2593–2606. https://doi.org/10.1098/RSTA.1999.0450

Ramsey, J. B., & Lampart, C. (1998). The Decomposition of Economic Relationships by Time Scale Using Wavelets: Expenditure and Income. *Studies in Nonlinear Dynamics & Econometrics*, *3*(1). https://doi.org/10.2202/1558-3708.1039

Raymond, R. (1976). Seasonality in Tourism: A Guide to the Analysis of Seasonality and Trends for Policy Making. *Journal of Travel Research*, *14*(4).

https://doi.org/10.1177/004728757601400470

Rokach, L., Maimon, O., & Shmueli, E. (Red.). (2023). *Machine Learning for Data Science Handbook*. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-24628-9

Sadkowski, W., & Sala, K. (2022). Zarządzanie jakością w usługach hotelarskich. *Zeszyty Naukowe Wydziału Zarządzania GWSH*, *18*, 48–52. https://doi.org/10.53259/2022.18.05

Sánchez-Medina, A. J., & C-Sánchez, E. (2020). Using machine learning and big data for efficient forecasting of hotel booking cancellations. *International Journal of Hospitality Management*, *89*, 102546. https://doi.org/10.1016/J.IJHM.2020.102546

Sarnecki, C. (2023a). Jak ocenić sprawność operacyjną hotelu? – czyli rachunkowość zarządcza w hotelarstwie. Pobrano 22 wrzesień 2023, z https://secco.com.pl/Jak-ocenic-sprawnosc-operacyjna-hotelu-czyli-rachunkowosc-zarzadcza-w-hotelarstwie-

Sarnecki, C. (2023b). Rachunkowość zarządcza w hotelarstwie. Cechy charakterystyczne branży hotelarskiej, historia USALI i układ Sprawozdania Operacyjnego.

Schleicher, C. (2002). *An Introduction to Wavelets for Economists*. *Bank of Canada Working Paper 2002-3*.

Schmidgall, R. S. ., & DeFranco, A. L. . (2023). *Hospitality Industry Managerial Accounting*. AHLA : National Restaurant Association Solutions. LLC & American Hotel & Lodging Educational Institute.

Schulz, E., Speekenbrink, M., & Krause, A. (2018). A tutorial on Gaussian process regression: Modelling, exploring, and exploiting functions. *Journal of Mathematical Psychology*, *85*, 1–16. https://doi.org/10.1016/J.JMP.2018.03.001

Shumway, R. H., & Stoffer, D. S. (2017). *Time Series Analysis and Its Applications: With R Examples.*

Slattery, P. (2002). Finding the Hospitality Industry. *The Journal of Hospitality Leisure Sport and Tourism*, *1*(1), 19–28. https://doi.org/10.3794/JOHLSTE.11.7

Snoek, J., Larochelle, H., & Adams, R. P. (2012). Practical bayesian optimization of machine learning algorithms. *Advances in Neural Information Processing Systems*, *4*, 2951–2959.

Song, Q. J., Jiang, H. Y., & Liu, J. (2017). Feature selection based on FDA and F-score for multi-class classification. *Expert Systems with Applications*, *81*, 22–27. https://doi.org/10.1016/J.ESWA.2017.02.049

Stabryła, A. (1998). Rachunek produktywności w diagnostyce ekonomicznej przedsiębiorstw. *Przegląd Organizacji*, *6*(29–32).

Straus, M. (2020). Wskaźniki hotelowe. Jak ocenić finanse obiektu noclegowego? Pobrano 21

wrzesień 2023, z https://maciejstraus.pl/wskazniki-hotelowe/

Sun, Z. H., Zuo, T. Y., Liang, D., Ming, X., Chen, Z., & Qiu, S. (2021). GPHC: A heuristic clustering method to customer segmentation. *Applied Soft Computing*, *111*, 107677. https://doi.org/10.1016/j.asoc.2021.107677

Szmidt, C. (2017). *Zarządzanie i ekonomia. Teraźniejszość i przyszłość*. Księgarnia PWN.

Torrence, C., & Compo, G. P. (1998). A Practical Guide to Wavelet Analysis. *Bulletin of the American Meteorological Society*, *79*(1), 61–78. https://doi.org/10.1023/A:1018662703668

Torrence, C., & Webster, P. J. (1999). Interdecadal changes in the ENSO-monsoon system. *Journal of Climate*, *12*(8 PART 2), 2679–2690. https://doi.org/10.1175/1520-0442(1999)012<2679:icitem>2.0.co;2

Tsionas, M., & Assaf, A. G. (2021). Revenue functions for hotels within clusters. *International Journal of Hospitality Management*, *98*(July), 103016. https://doi.org/10.1016/j.ijhm.2021.103016

Tylińska, R., & Wajgner, M. (2012). *Ekonomia i prawo w hotelarstwie. Podręcznik*. Księgarnia PWN.

Tyran, M. R. (2004). *Wskaźniki finansowe*. Oficyna Ekonomiczna.

Viverit, L., Heo, C. Y., Pereira, L. N., & Tiana, G. (2023). Application of machine learning to cluster hotel booking curves for hotel demand forecasting. *International Journal of Hospitality Management*, *111*, 103455. https://doi.org/10.1016/J.IJHM.2023.103455

Wang, Z., & Sarkis, J. (2017). Corporate social responsibility governance, outcomes, and financial performance. *Journal of Cleaner Production*, *162*, 1607–1616. https://doi.org/10.1016/J.JCLEPRO.2017.06.142

Webb, T., Schwartz, Z., Xiang, Z., & Singal, M. (2020). Revenue management forecasting: the resiliency of advanced booking methods given dynamic booking windows. *International Journal of Hospitality Management*, *89*(March), 102590. https://doi.org/10.1016/j.ijhm.2020.102590

Yadegaridehkordi, E., Nilashi, M., Nasir, M. H. N. B. M., & Ibrahim, O. (2018). Predicting determinants of hotel success and development using Structural Equation Modelling (SEM)-ANFIS method. *Tourism Management*, *66*(2018), 364–386. https://doi.org/10.1016/j.tourman.2017.11.012

Yadegaridehkordi, E., Nilashi, M., Nizam Bin Md Nasir, M. H., Momtazi, S., Samad, S., Supriyanto, E., & Ghabban, F. (2021). Customers segmentation in eco-friendly hotels using multi-criteria and machine learning techniques. *Technology in Society*,

65(February), 101528. https://doi.org/10.1016/j.techsoc.2021.101528

Yuan, S., Wu, Z., & Liu, L. (2022). The effects of financial openness and financial efficiency on Chinese macroeconomic volatilities. *The North American Journal of Economics and Finance*, *63*, 101819. https://doi.org/10.1016/J.NAJEF.2022.101819

Zhang, D., & Wu, C. (2023). What online review features really matter? An explainable deep learning approach for hotel demand forecasting. *Journal of the Association for Information Science and Technology*, *74*(9), 1100–1117. https://doi.org/10.1002/ASI.24807

Zhao, Z., Anand, R., & Wang, M. (2019). Maximum relevance and minimum redundancy feature selection methods for a marketing machine learning platform. *Proceedings - 2019 IEEE International Conference on Data Science and Advanced Analytics, DSAA 2019*, 442–452. https://doi.org/10.1109/DSAA.2019.00059

Zulfiqar, M., Gamage, K. A. A., Kamran, M., & Rasheed, M. B. (2022). Hyperparameter optimization of bayesian neural network using bayesian optimization and intelligent feature engineering for load forecasting. *Sensors (Basel, Switzerland)*, *22*(12). https://doi.org/10.3390/S22124446

10. Spis tabel

Tabela 1. Ogólny format Sprawozdania Operacyjnego w języku polskim i angielskim
Tabela 2. Zestawienie zbiorcze miar statystycznych obliczonych dla zmiennej GOP w kolejnych hotelach
Tabela 3. Wyniki obliczeń zależności korelacyjnej zmiennej GOP do pozostałych zmiennych tablicy USALI dla wszystkich hoteli
Tabela 4. Zestawienie zbiorcze wyselekcjonowanych zmiennych USALI dla poszczególnych hoteli
Tabela 5. Zestawienie zbiorcze miar statystycznych obliczonych dla wyselekcjonowanych zmiennych USALI w HDSI Katowice
Tabela 6. Zestawienie zbiorcze miar statystycznych obliczonych dla wyselekcjonowanych zmiennych USALI w HDS Piaseczno
Tabela 7. Zestawienie zbiorcze miar statystycznych obliczonych dla wyselekcjonowanych zmiennych USALI w HDSP Opole
Tabela 8. Zestawienie zbiorcze miar statystycznych obliczonych dla wyselekcjonowanych zmiennych USALI w HDSP Poznań
Tabela 9. Zestawienie zbiorcze miar statystycznych obliczonych dla wyselekcjonowanych zmiennych USALI w HM Opole
Tabela 10. Zestawienie zbiorcze miar statystycznych obliczonych dla wyselekcjonowanych zmiennych USALI w HMC Częstochowa
Tabela 11. Zestawienie zbiorcze miar statystycznych obliczonych dla wyselekcjonowanych zmiennych USALI w HM Karpacz
Tabela 12. Zestawienie zbiorcze wyników końcowego modelu regresji liniowej dla hotelu HDSI Katowice
Tabela 13. Zestawienie zbiorcze wyników końcowego modelu regresji liniowej dla hotelu HDS Piaseczno
Tabela 14. Zestawienie zbiorcze wyników końcowego modelu regresji liniowej dla hotelu HDSP Opole
Tabela 15. Zestawienie zbiorcze wyników końcowego modelu regresji liniowej dla hotelu HDSP Poznań
Tabela 16. Zestawienie zbiorcze wyników końcowego modelu regresji liniowej dla hotelu HM Opole
Tabela 17. Zestawienie zbiorcze wyników końcowego modelu regresji liniowej dla hotelu HMC Częstochowa
Tabela 18. Zestawienie zbiorcze wyników końcowego modelu regresji liniowej dla hotelu HM Karpacz
Tabela 19. Zestawienie zbiorcze wyników procesu doboru końcowego modelu ekonometrycznego, dla wszystkich hoteli
Tabela 20. Nazwy i oznaczenia algorytmów uczenia maszynowego zastosowanych w analizach
Tabela 21. Zestawienie zbiorcze wyników procesu trenowania modeli uczenia maszynowego typu EN, SVM, NN, TREE, GPR, dla wszystkich hoteli
Tabela 22. Zestawienie zbiorcze wyników procesu testowania modeli uczenia maszynowego typu EN, SVM, NN, TREE, GPR, dla wszystkich hoteli

11. Spis rysunków

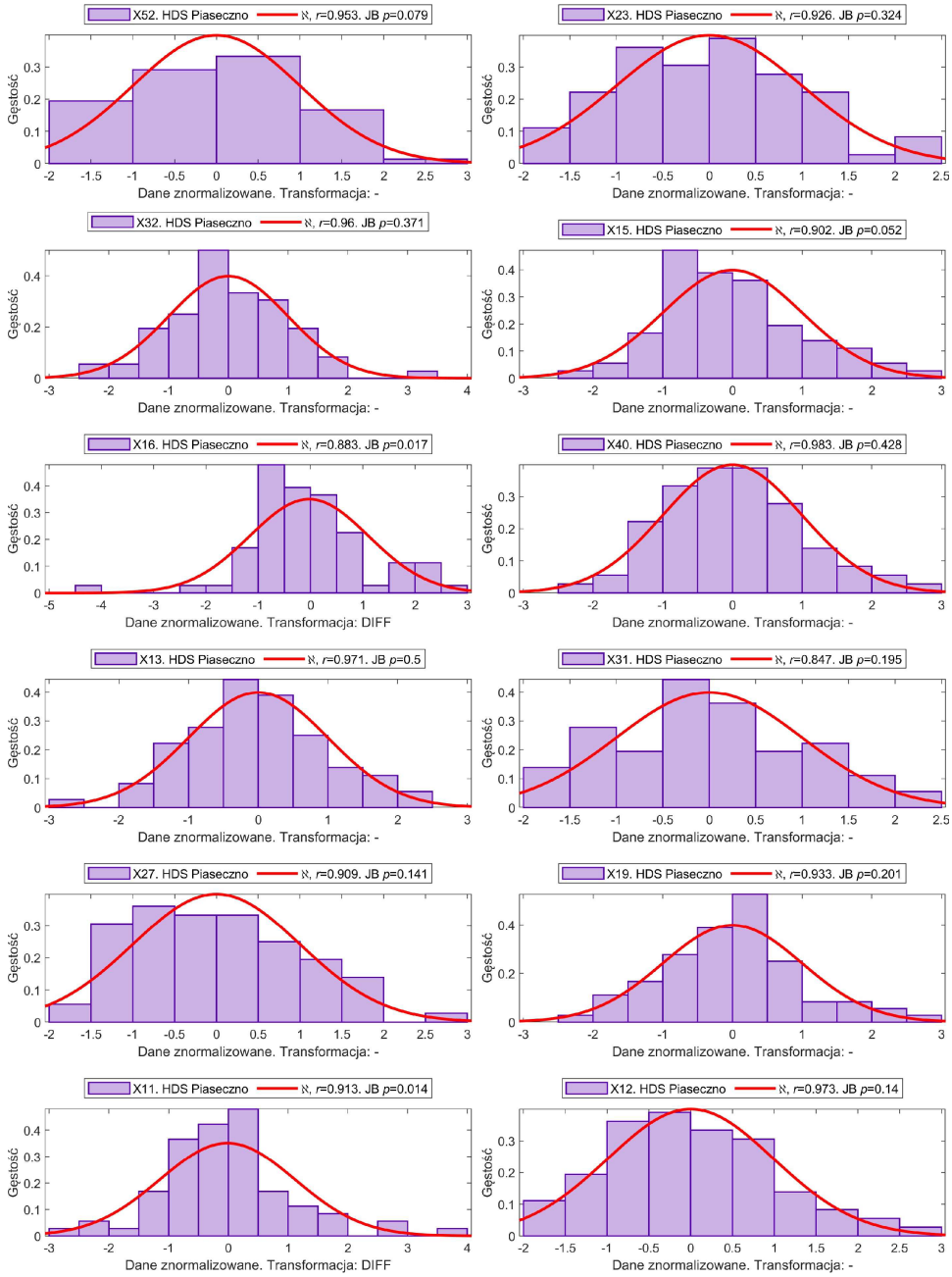
Rys. 4.1 Przebieg czasowy standaryzowanych wartości zmiennej GOP dla HDSI Katowice
Rys. 4.2 Przebieg czasowy standaryzowanych wartości zmiennej GOP dla HDS Piaseczno
Rys. 4.3 Przebieg czasowy standaryzowanych wartości zmiennej GOP dla HDSP Opole
Rys. 4.4 Przebieg czasowy standaryzowanych wartości zmiennej GOP dla HDSP Poznań
Rys. 4.5 Przebieg czasowy standaryzowanych wartości zmiennej GOP dla HM Opole
Rys. 4.6 Przebieg czasowy standaryzowanych wartości zmiennej GOP dla HMC Częstochowa
Rys. 4.7 Przebieg czasowy standaryzowanych wartości zmiennej GOP dla HM Karpacz
Rys.4.8 Wykres rozkładu empirycznej gęstości prawdopodobieństwa dla zmiennej GOP rejestrowanej w hotelu HDSI Katowice
Rys.4.9 Wykres rozkładu empirycznej gęstości prawdopodobieństwa dla zmiennej GOP rejestrowanej w hotelu HDS Piaseczno
Rys.4.10 Wykres rozkładu empirycznej gęstości prawdopodobieństwa dla zmiennej GOP rejestrowanej w hotelu HDSP Opole
Rys.4.11 Wykres rozkładu empirycznej gęstości prawdopodobieństwa dla zmiennej GOP rejestrowanej w hotelu HDSP Poznań
Rys.4.12 Wykres rozkładu empirycznej gęstości prawdopodobieństwa dla zmiennej GOP rejestrowanej w hotelu HM Opole
Rys.4.13 Wykres rozkładu empirycznej gęstości prawdopodobieństwa dla zmiennej GOP rejestrowanej w hotelu HMC Częstochowa
Rys.4.14 Wykres rozkładu empirycznej gęstości prawdopodobieństwa dla zmiennej GOP rejestrowanej w hotelu HM Karpacz
Rys. 4.15 Wykres autokorelacji standaryzowanych wartości zmiennej GOP dla HDSI Katowice
Rys. 4.16 Wykres autokorelacji standaryzowanych wartości zmiennej GOP dla HDS Piaseczno
Rys. 4.17 Wykres autokorelacji standaryzowanych wartości zmiennej GOP dla HDSP Opole
Rys. 4.18 Wykres autokorelacji standaryzowanych wartości zmiennej GOP dla HDSP Poznań
Rys. 4.19 Wykres autokorelacji standaryzowanych wartości zmiennej GOP dla HM Opole
Rys. 4.20 Wykres autokorelacji standaryzowanych wartości zmiennej GOP dla HMC Częstochowa
Rys. 4.21 Wykres autokorelacji standaryzowanych wartości zmiennej GOP dla HM Karpacz
Rys. 4.22 Wyniki analizy widmowej wykorzystującej metodę PMUSIC.
Rys. 4.23 Zestawienie zbiorcze głównych składowych widma zmiennej GOP dla wszystkich hoteli.
Rys. 5.1 Wizualizacja rozrzutu zmiennych X4 do X9 względem zmiennej GOP, pozyskanych dla wszystkich hotel
Rys. 5.2 Wizualizacja rozrzutu zmiennych X10, X13, X17, X18, X23 i X24 względem zmiennej GOP, pozyskanych dla wszystkich hoteli
Rys. 5.3 Wizualizacja rozrzutu zmiennych X32, X37, X38, X39, X52 i X53 względem zmiennej GOP, pozyskanych dla wszystkich hoteli
Rys. 5.4 Wizualizacja kwadratu modułu koherencji falkowej dla zmiennej GOP względem dziesięciu zmiennych USALI, pozyskanych w rozpatrywanym okresie działalności hotelu HDSI Katowice
Rys. 6.1 Przebiegi czasowe standaryzowanych wartości zmiennych USALI skorelowanych z GOP dla HDSI Katowice
Rys. 6.2 Wykresy autokorelacji standaryzowanych wartości zmiennych USALI skorelowanych z GOP dla HDSI Katowice
Rys. 6.3 Wyniki analizy widmowej standaryzowanych wartości zmiennych USALI skorelowanych z GOP obliczonych dla HDSI Katowice
Rys. 7.1 Zestawienie porównawcze rozkładu gęstości prawdopodobieństwa reszt wraz z wykresem Q–Q, uzyskane w wyniku dopasowania modelu ekonometrycznego dla hotelu HDS Piaseczno

Rys. 7.2 Zestawienie porównawcze rozkładu gęstości prawdopodobieństwa reszt wraz z wykresem Q–Q, uzyskane w wyniku dopasowania modelu ekonometrycznego dla hotelu HDSI Katowice
Rys. 7.4 Zestawienie porównawcze rozkładu gęstości prawdopodobieństwa reszt wraz z wykresem Q–Q, uzyskane w wyniku dopasowania modelu ekonometrycznego dla hotelu HDSP Poznań
Rys. 7.5 Zestawienie porównawcze rozkładu gęstości prawdopodobieństwa reszt wraz z wykresem Q–Q, uzyskane w wyniku dopasowania modelu ekonometrycznego dla hotelu HM Karpacz
Rys. 7.6 Zestawienie porównawcze rozkładu gęstości prawdopodobieństwa reszt wraz z wykresem Q–Q, uzyskane w wyniku dopasowania modelu ekonometrycznego dla hotelu HM Opole
Rys. 7.7 Zestawienie porównawcze rozkładu gęstości prawdopodobieństwa reszt wraz z wykresem Q–Q, uzyskane w wyniku dopasowania modelu ekonometrycznego dla hotelu HMC Częstochowa
Rys. 7.8 Schemat blokowy procesu przygotowania i modelowania danych.
Rys. 7.9 Zestawienie zbiorcze wyników analizy ważności zestawu zmiennych USALI skorelowanych ze zmienną GOP wyselekcjonowanych w poprzednich badaniach, wyznaczone z wykorzystaniem metody MRM i testu F.
Rys. 7.10 Zestawienie porównawcze przykładowych wyników dopasowania nieliniowych modeli predykujących zmienną GOP dla hotelu HDSI Katowice
Rys. 7.11 Zestawienie porównawcze rozkładów gęstości prawdopodobieństwa reszt wraz z wykresami Q–Q, przeprowadzone w 100 iteracjach uczenia i testowania modelu predykującego GOP dla hoteli HDSI Katowice, z wykorzystaniem algorytmu EN
Rys. 7.12 Zestawienie porównawcze rozkładów gęstości prawdopodobieństwa reszt wraz z wykresami Q–Q, przeprowadzone w 100 iteracjach uczenia i testowania modelu predykującego GOP dla hoteli HDSI Katowice, z wykorzystaniem algorytmu SVM
Rys. 7.13 Zestawienie porównawcze rozkładów gęstości prawdopodobieństwa reszt wraz z wykresami Q–Q, przeprowadzone w 100 iteracjach uczenia i testowania modelu predykującego GOP dla hoteli HDSI Katowice, z wykorzystaniem algorytmu TREE
Rys. 7.14 Zestawienie porównawcze rozkładów gęstości prawdopodobieństwa reszt wraz z wykresami Q–Q, przeprowadzone w 100 iteracjach uczenia i testowania modelu predykującego GOP dla hoteli HDSI Katowice, z wykorzystaniem algorytmu NN
Rys. 7.15 Zestawienie porównawcze rozkładów gęstości prawdopodobieństwa reszt wraz z wykresami Q–Q, przeprowadzone w 100 iteracjach uczenia i testowania modelu predykującego GOP dla hoteli HDSI Katowice, z wykorzystaniem algorytmu GPR

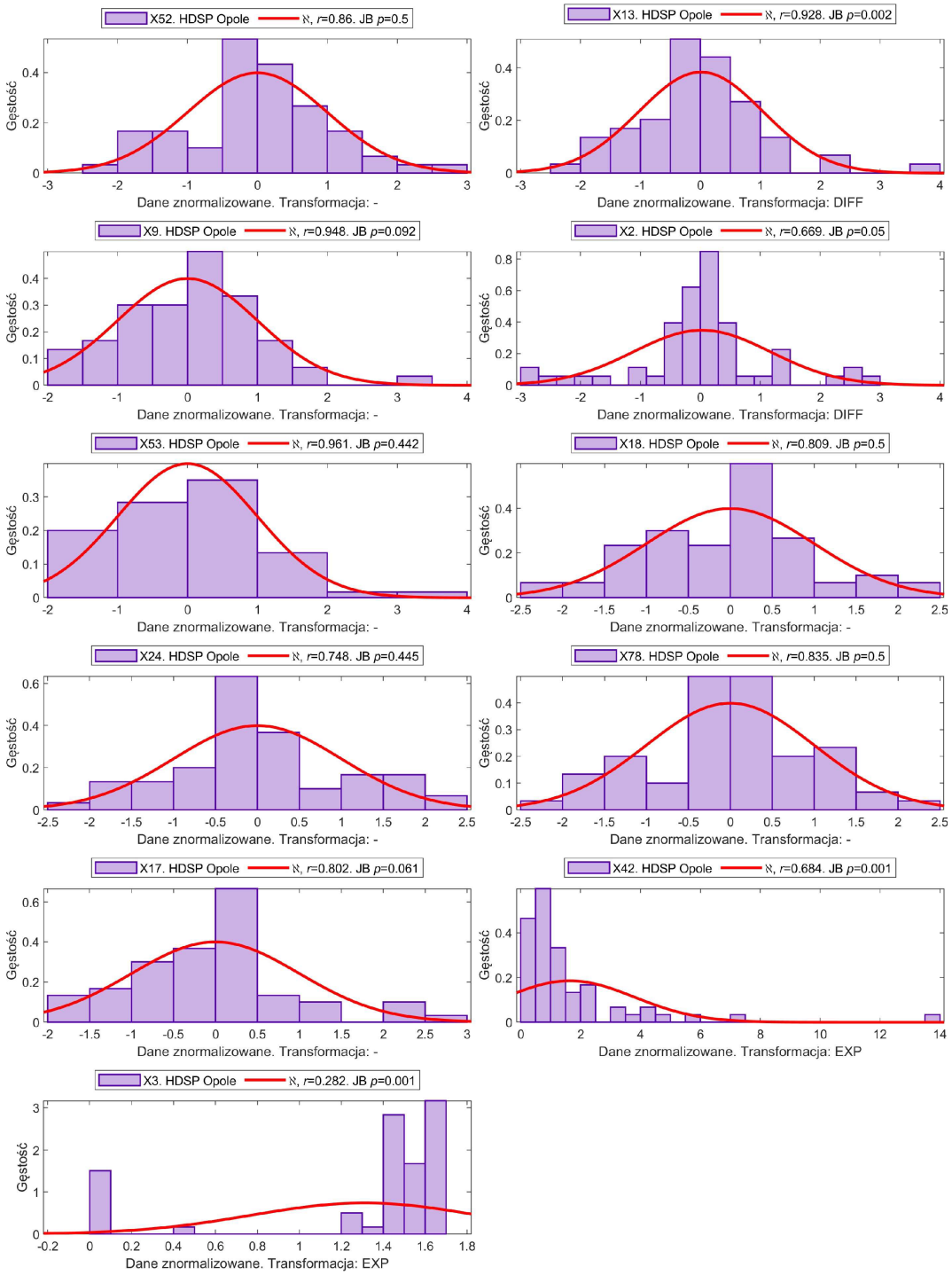
Załączniki

Załącznik Z1

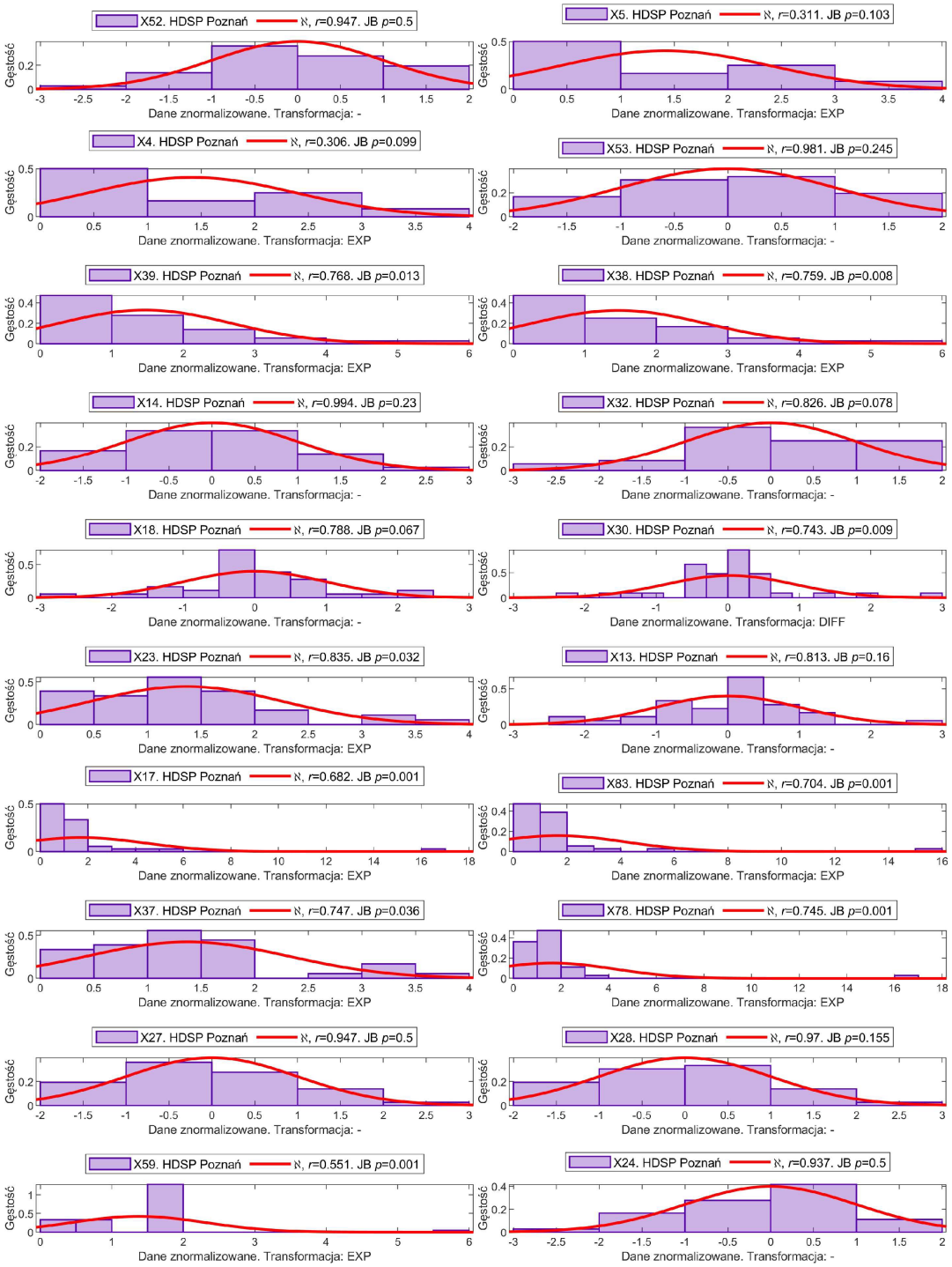
Na rysunkach Z1.1 – Z1.6 przedstawiono graficzną prezentację rozkładów gęstości wraz z dopasowaniem FGP rozkładu normalnego, obliczone dla rozpatrywanych w pracy zmiennych USALI, dla poszczególnych hoteli (HDS Piaseczno, HDSI Opole, HDSI Poznań, HM Opole, HDC Częstochowa, HM Karpacz).



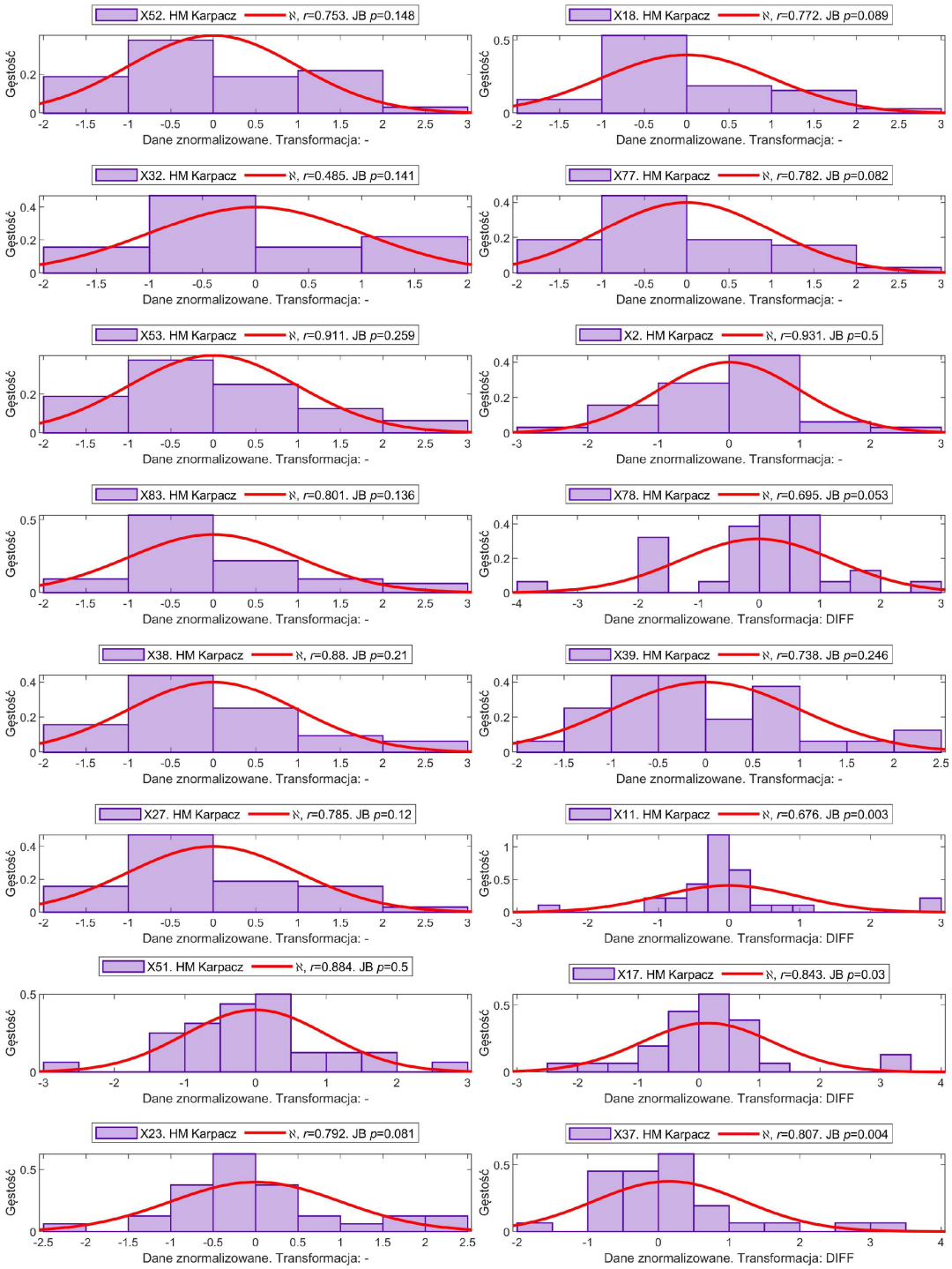
Rys. Z1.1 Zestawienie zbiorcze rozkładów empirycznej gęstości prawdopodobieństwa rozpatrywanych w pracy zmiennych USALI. Wykresy dotyczą danych pozyskanych z hotelu HDS Piaseczno. Źródło: Opracowanie własne.



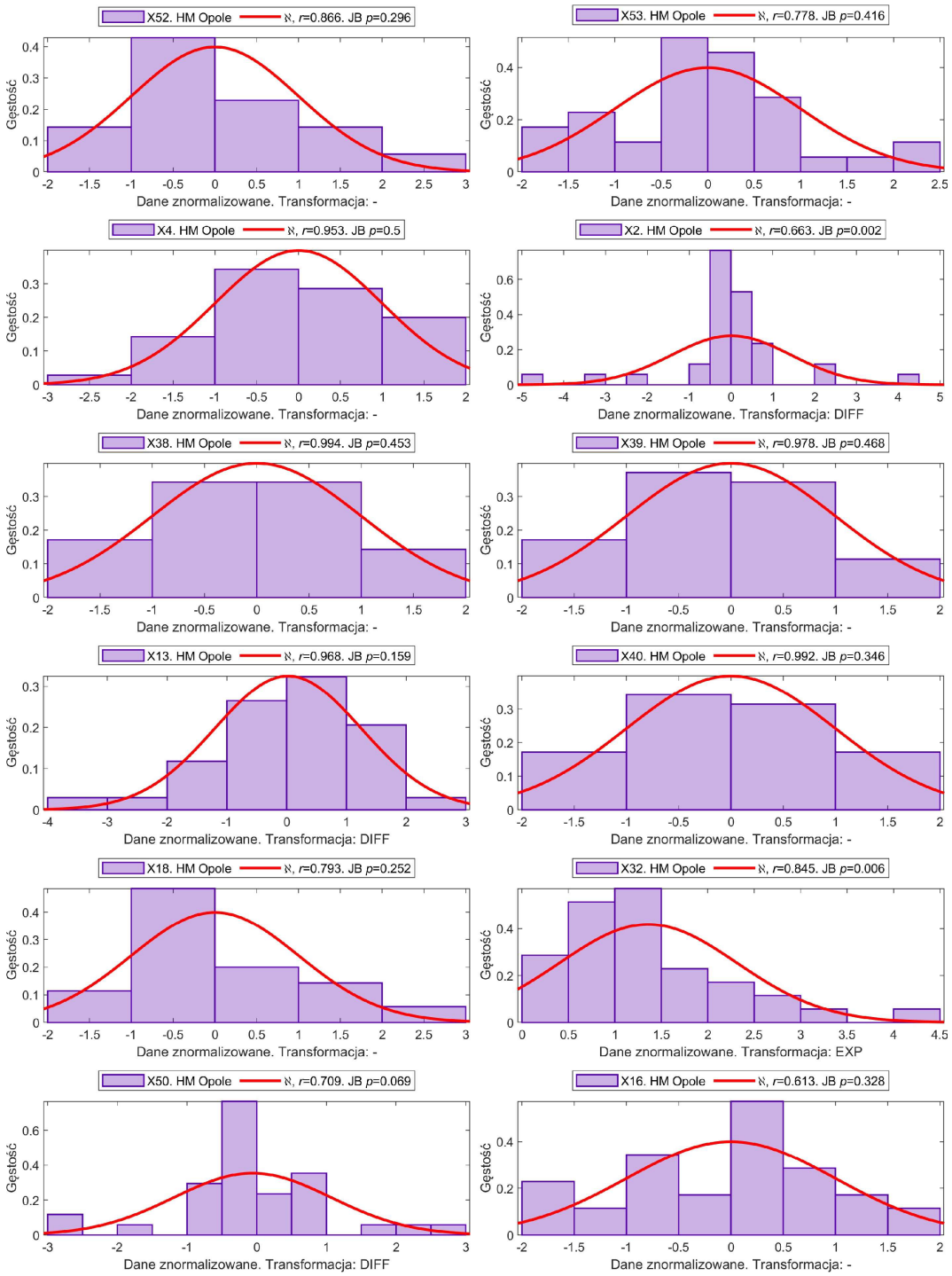
Rys. Z1.2 Zestawienie zbiorcze rozkładów empirycznej gęstości prawdopodobieństwa rozpatrywanych w pracy zmiennych USALI. Wykresy dotyczą danych pozyskanych z hotelu HDSP Opole. Źródło: Opracowanie własne.



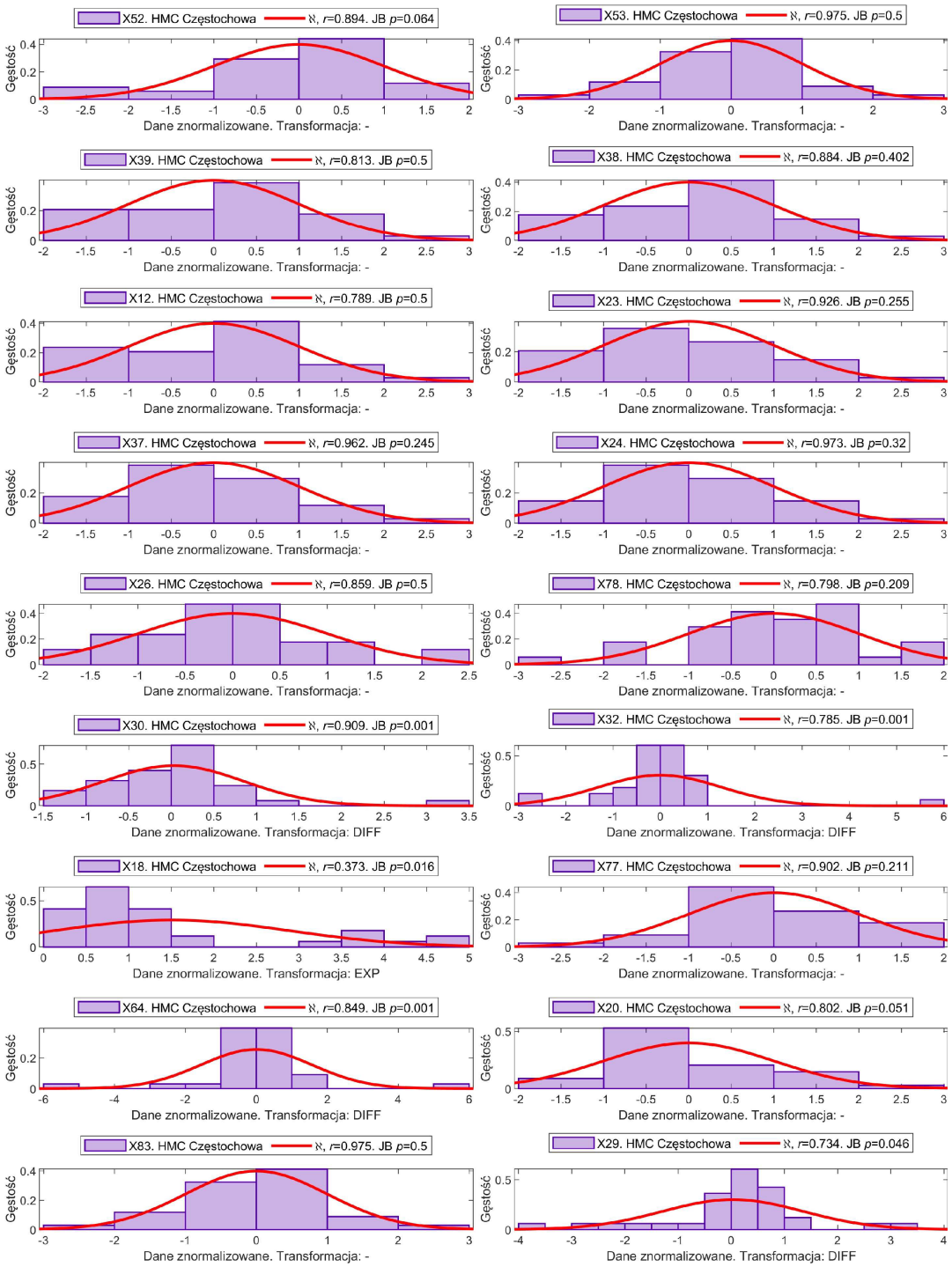
Rys. Z1.3 Zestawienie zbiorcze rozkładów empirycznej gęstości prawdopodobieństwa rozpatrywanych w pracy zmiennych USALI. Wykresy dotyczą danych pozyskanych z hotelu HDSP Poznań. Źródło: Opracowanie własne.



Rys. Z1.4 Zestawienie zbiorcze rozkładów empirycznej gęstości prawdopodobieństwa rozpatrywanych w pracy zmiennych USALI. Wykresy dotyczą danych pozyskanych z hotelu HM Karpacz. Źródło: Opracowanie własne.



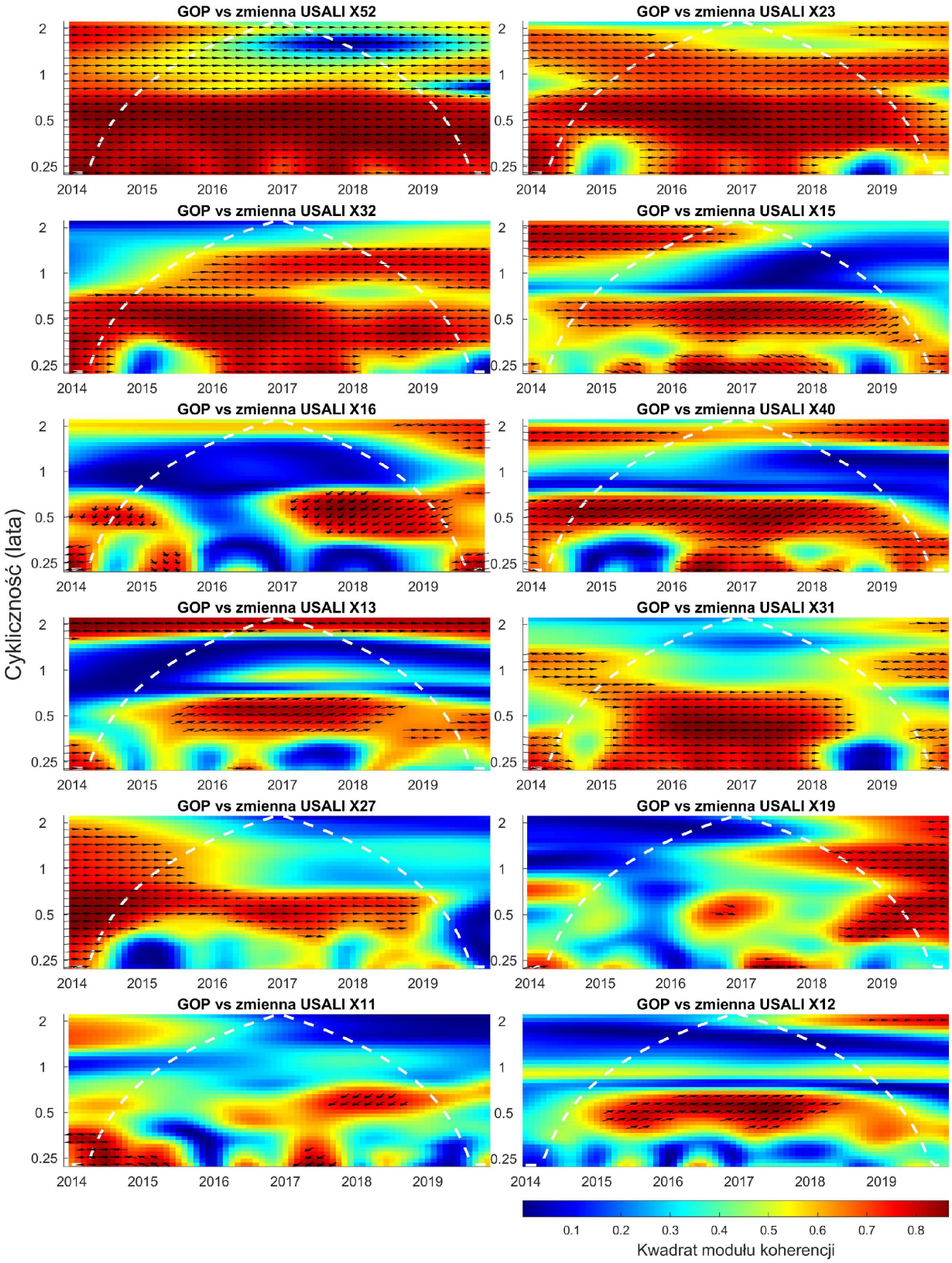
Rys. Z1.5 Zestawienie zbiorcze rozkładów empirycznej gęstości prawdopodobieństwa rozpatrywanych w pracy zmiennych USALI. Wykresy dotyczą danych pozyskanych z hotelu HM Opole. Źródło: Opracowanie własne.



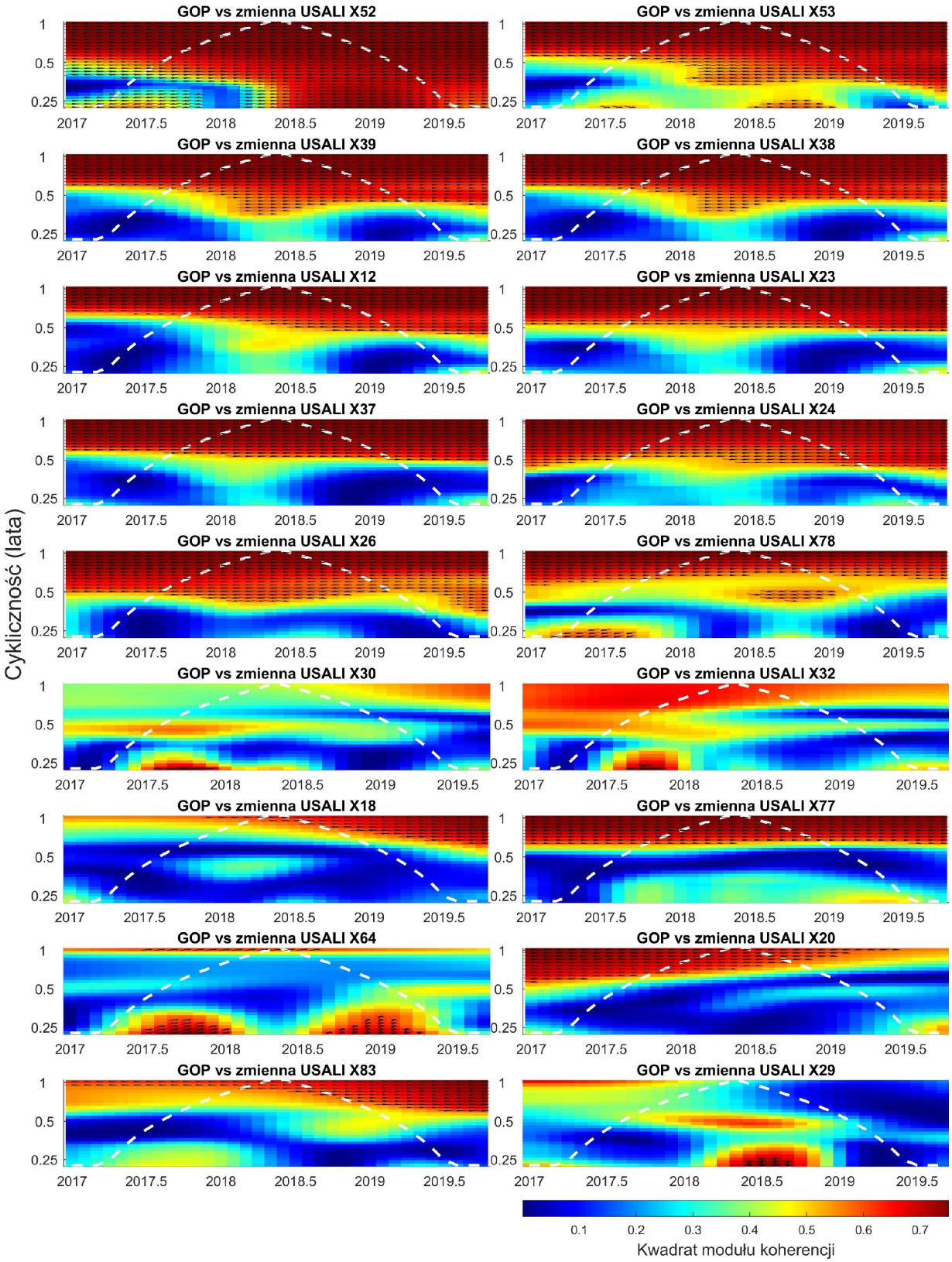
Rys. Z1.6 Zestawienie zbiorcze rozkładów empirycznej gęstości prawdopodobieństwa rozpatrywanych w pracy zmiennych USALI. Wykresy dotyczą danych pozyskanych z hotelu HMC Częstochowa. Źródło: Opracowanie własne.

Załącznik Z2

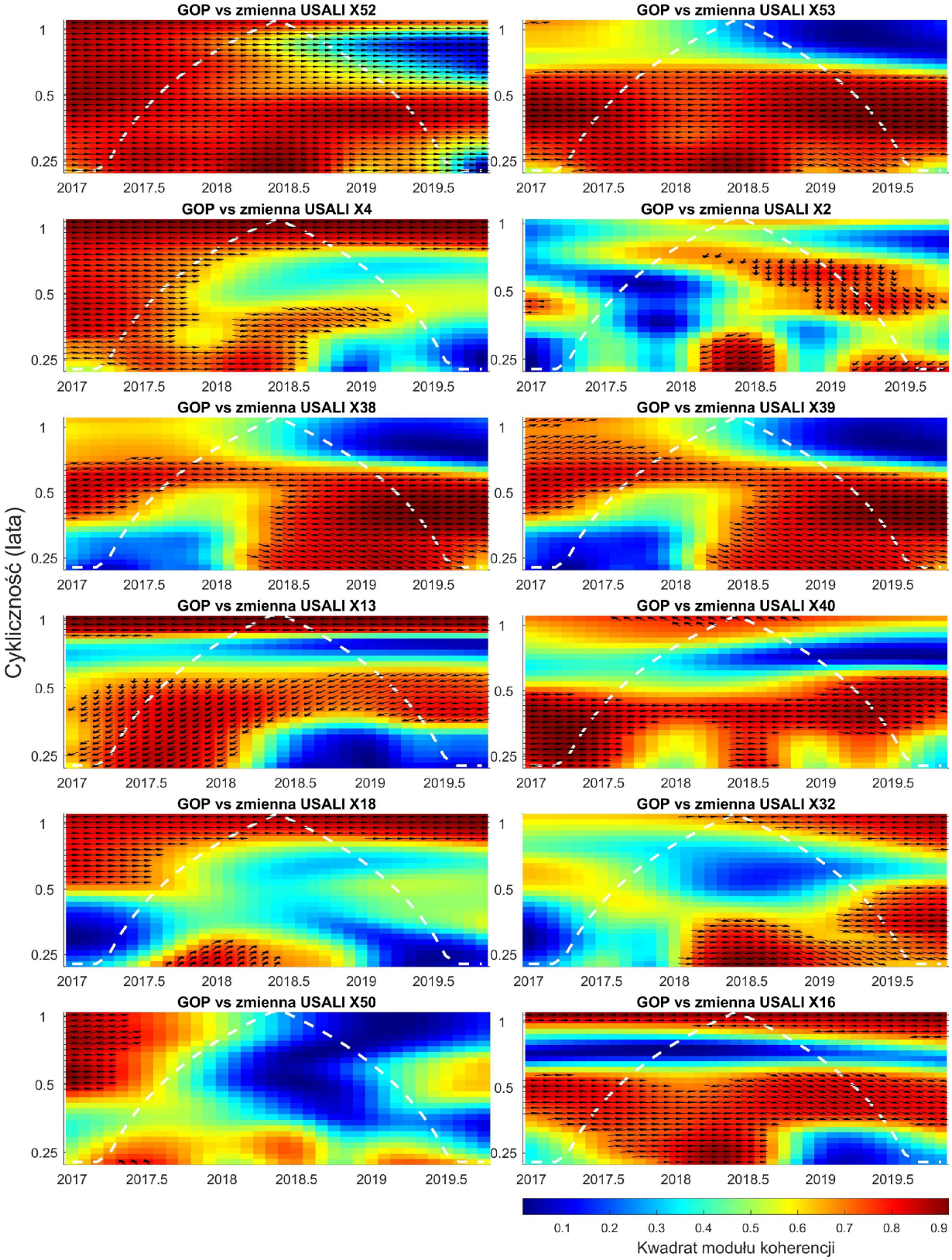
Na rysunkach Z2.1 – Z2.6 przedstawiono zbiorcze zestawienie wyników analizy koherencji falkowej między zmienną GOP, a zmiennymi istotnie z nim skorelowanymi, dla poszczególnych hoteli (HDS Piaseczno, HDSI Opole, HDSI Poznań, HM Opole, HDC Częstochowa, HM Karpacz).



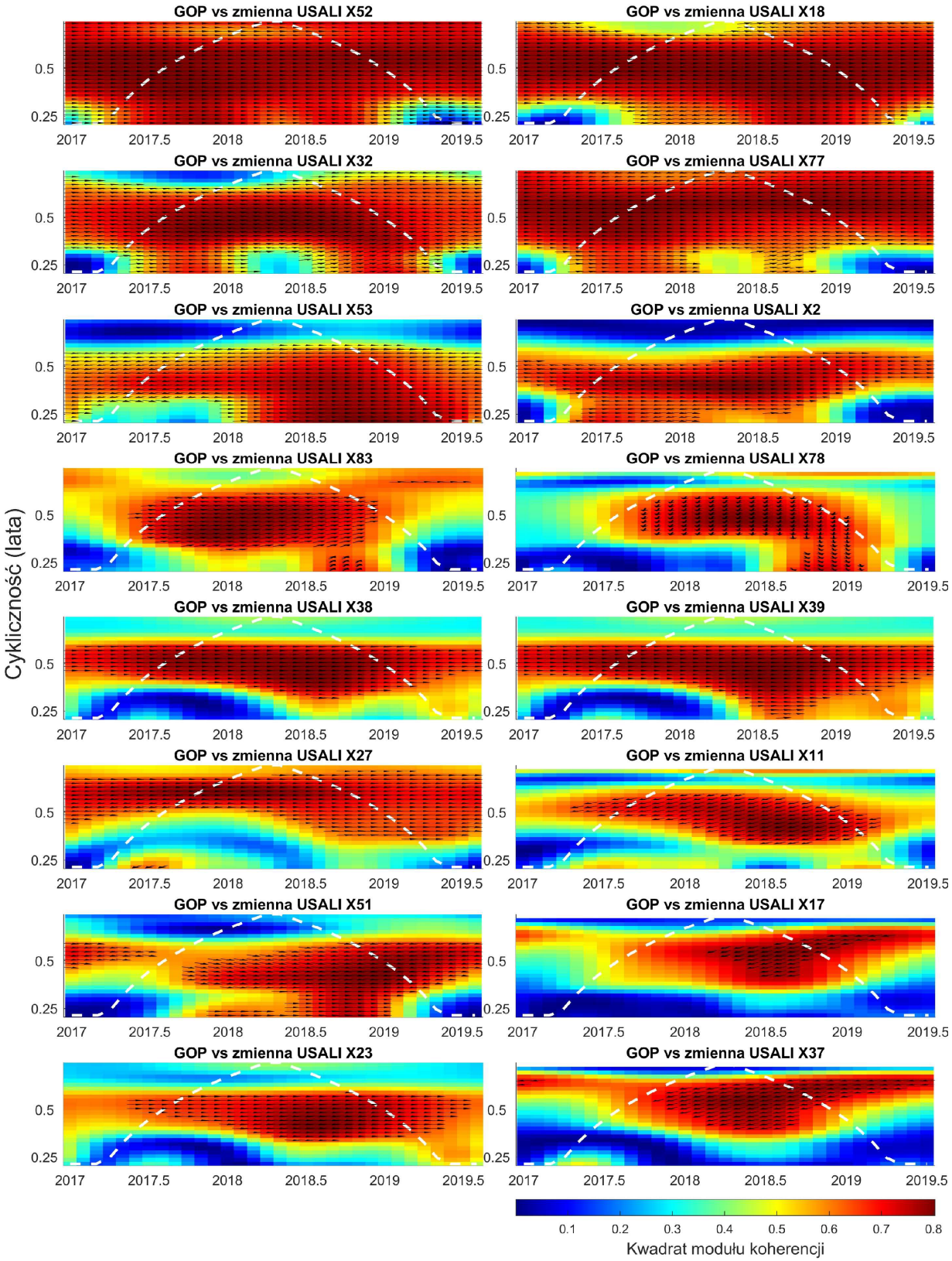
Rys. Z2.1 Zestawienie zbiorcze kwadratu modułu koherencji falkowej dla zmiennej GOP względem zmiennych z nią skorelowanych. Wykresy dotyczą danych pozyskanych z hotelu HDS Piaseczno. Źródło: Opracowanie własne.



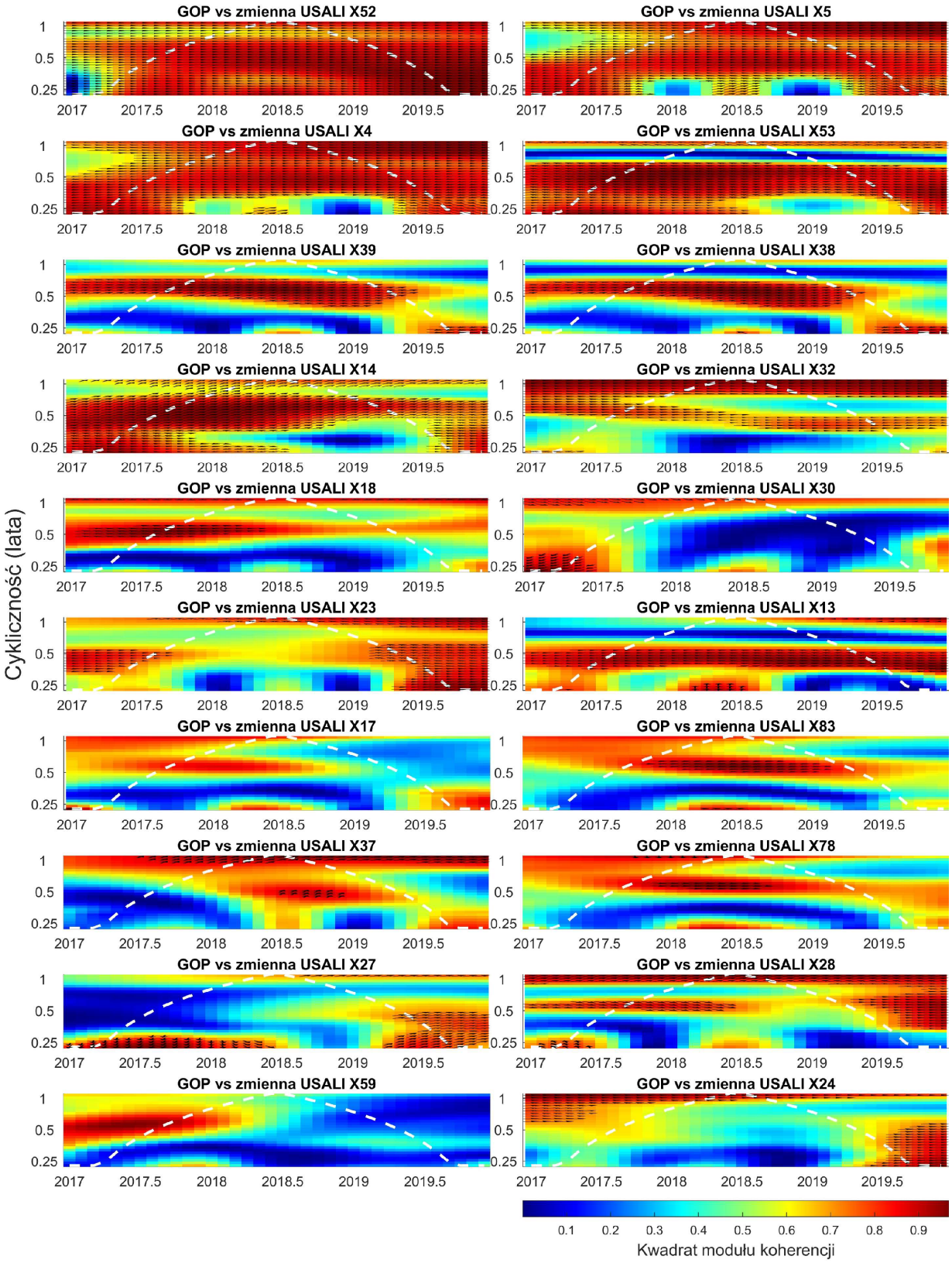
Rys. Z2.2 Zestawienie zbiorcze kwadratu modułu koherencji falkowej dla zmiennej GOP względem zmiennych z nią skorelowanych. Wykresy dotyczą danych pozyskanych z hotelu HMC Częstochowa. Źródło: Opracowanie własne.



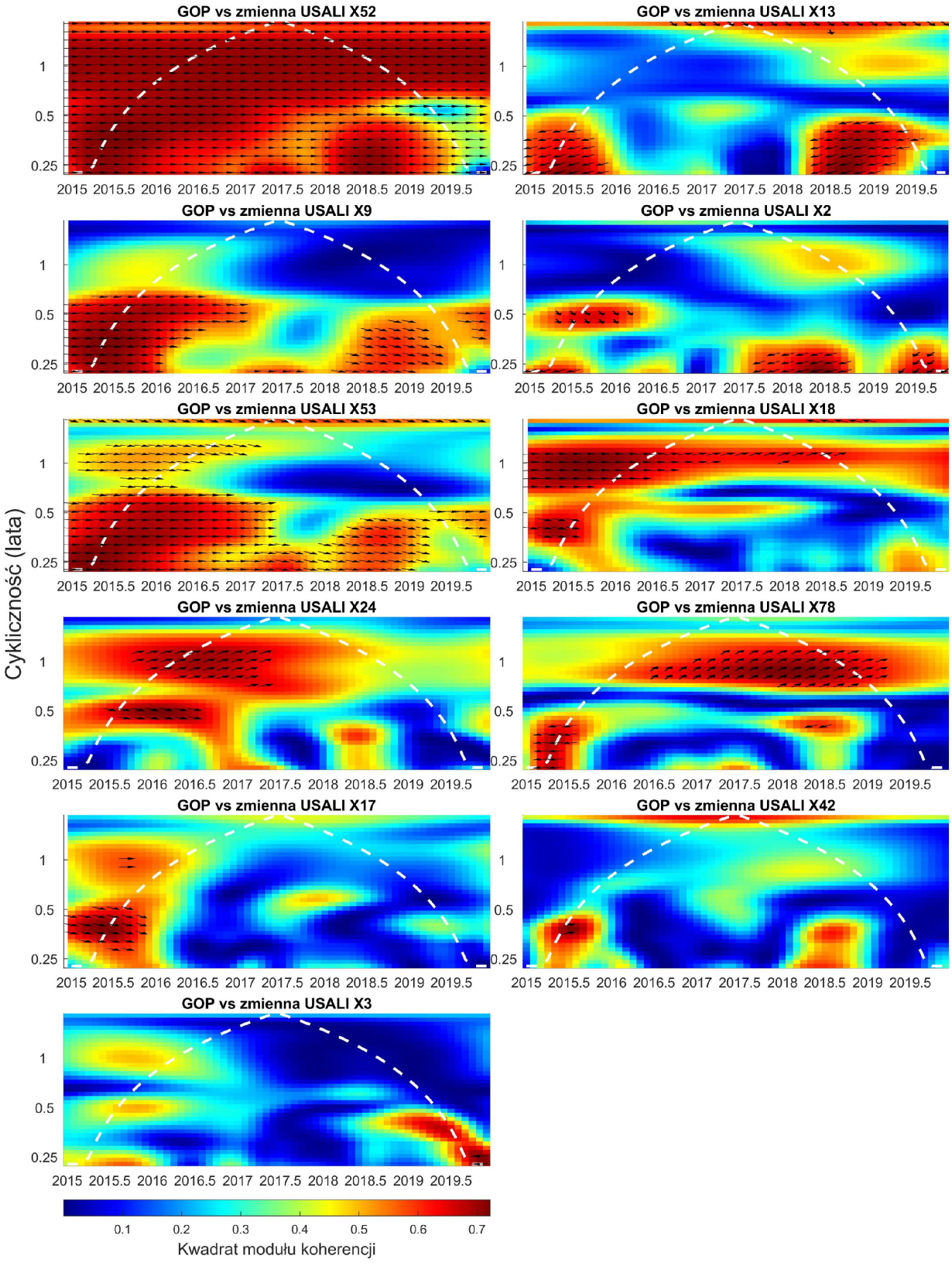
Rys. Z2.3 Zestawienie zbiorcze kwadratu modułu koherencji falkowej dla zmiennej GOP względem zmiennych z nią skorelowanych. Wykresy dotyczą danych pozyskanych z hotelu HM Opole. Źródło: Opracowanie własne.



Rys. Z2.4 Zestawienie zbiorcze kwadratu modułu koherencji falkowej dla zmiennej GOP względem zmiennych z nią skorelowanych. Wykresy dotyczą danych pozyskanych z hotelu HM Karpacz. Źródło: Opracowanie własne.



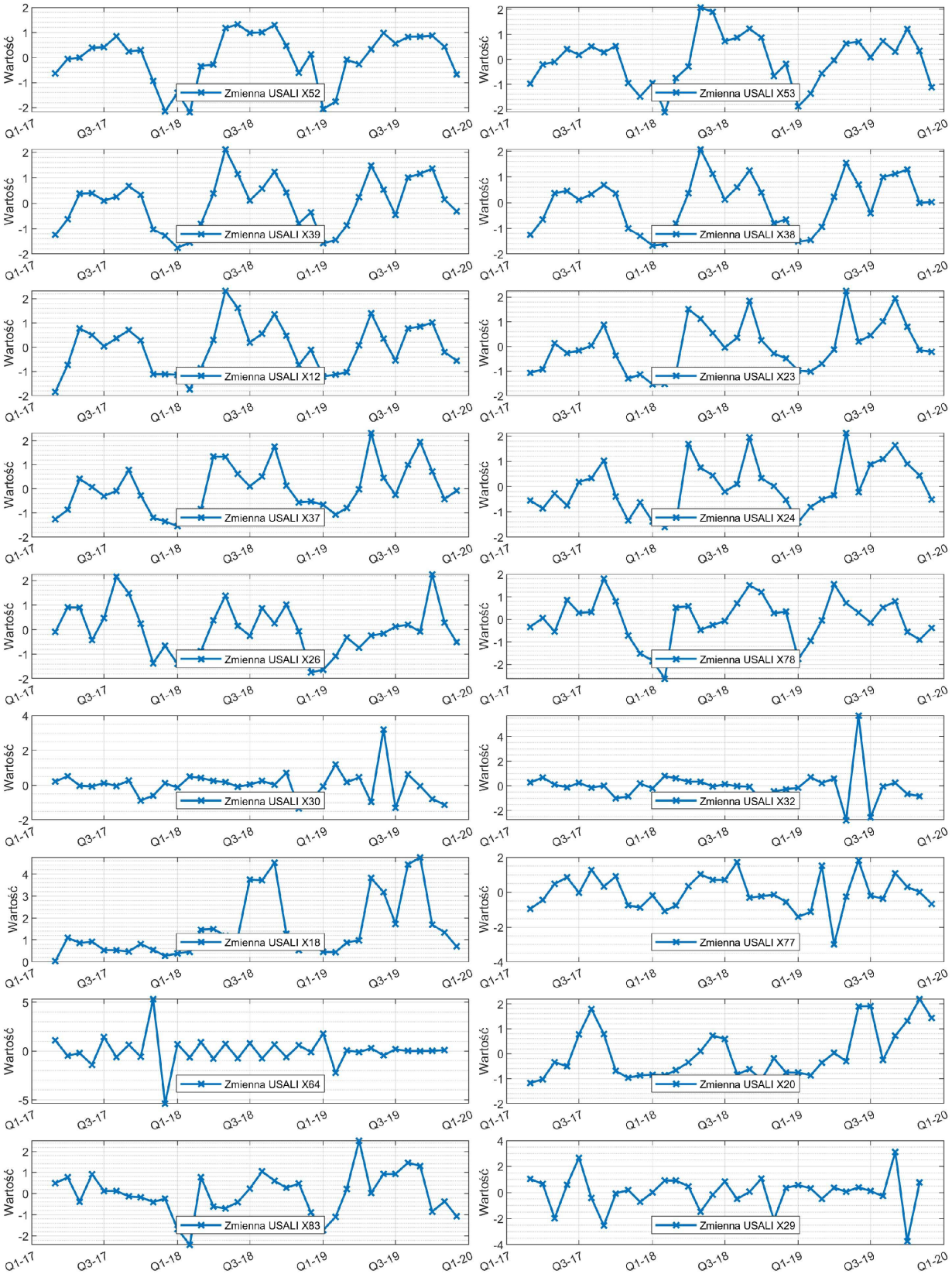
Rys. Z2.5 Zestawienie zbiorcze kwadratu modułu koherencji falkowej dla zmiennej GOP względem zmiennych z nią skorelowanych. Wykresy dotyczą danych pozyskanych z hotelu HDSP Poznań. Źródło: Opracowanie własne.



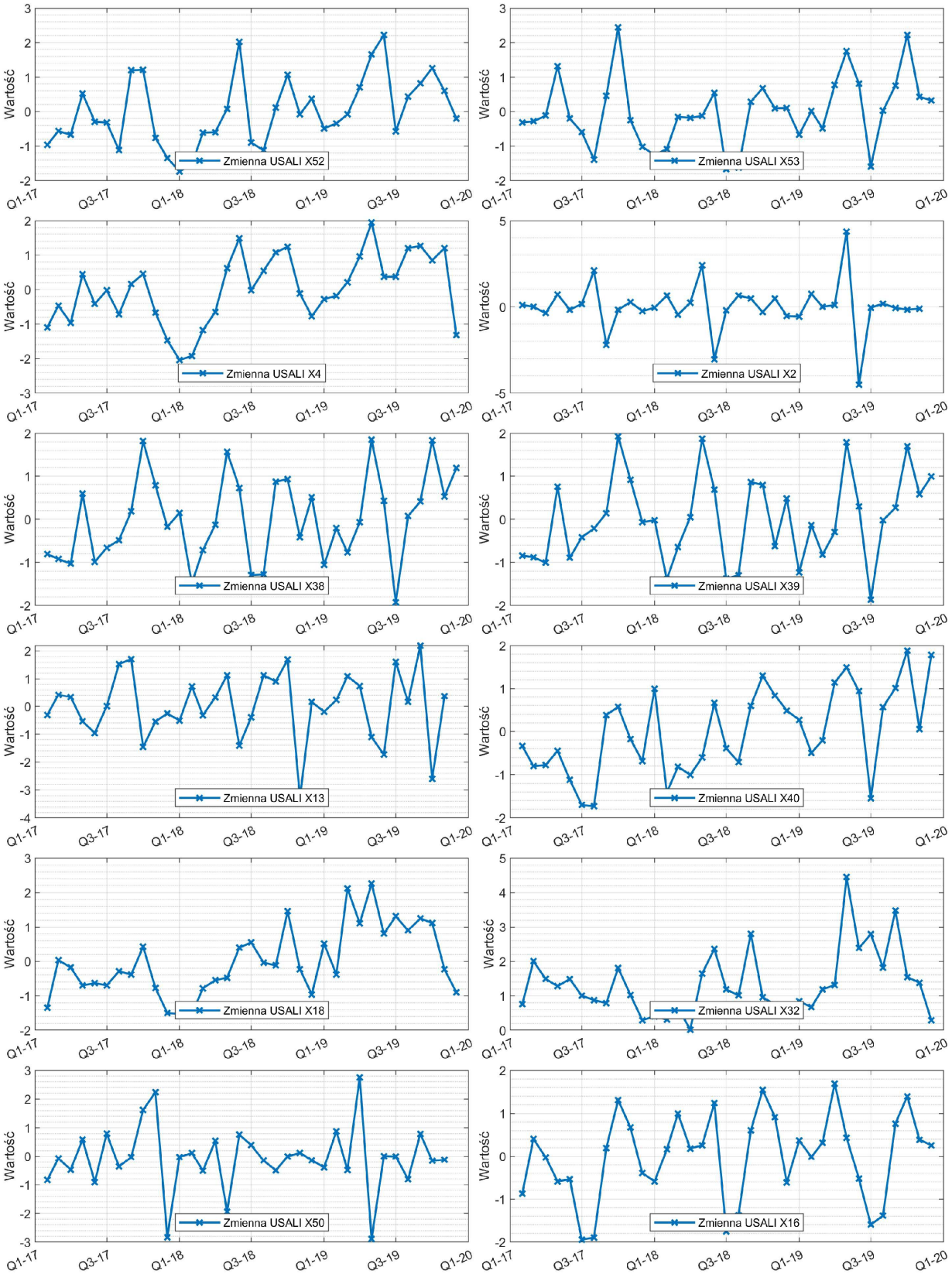
Rys. Z2.6 Zestawienie zbiorcze kwadratu modułu koherencji falkowej dla zmiennej GOP względem zmiennych z nią skorelowanych Wykresy dotyczą danych pozyskanych z hotelu HDSP Opole. Źródło: Opracowanie własne.

Załącznik Z3

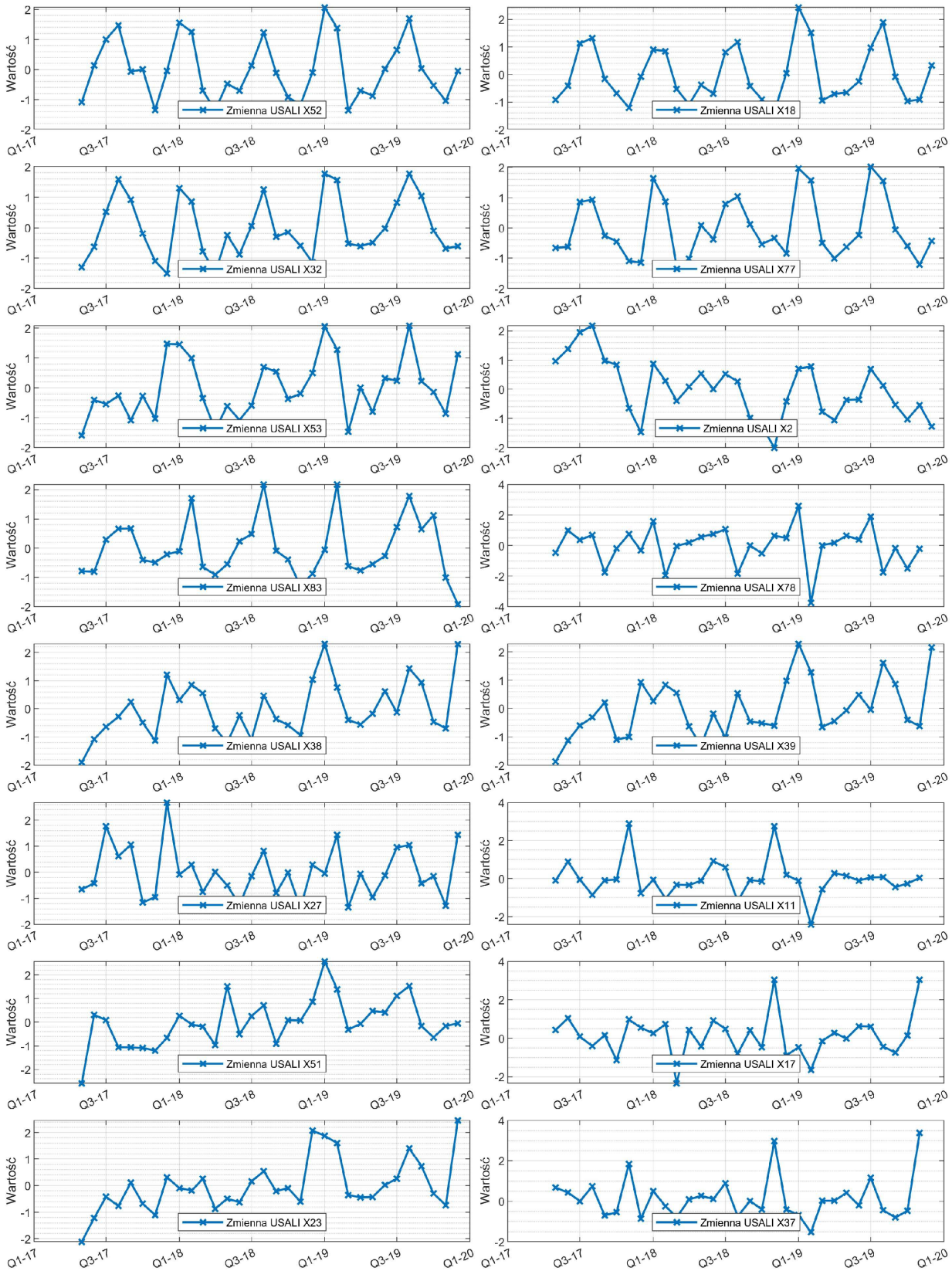
Na rysunkach Z3.1– Z3.6 przedstawiono przebiegi czasowe ustandaryzowanych wartości rozpatrywanych w pracy zmiennych USALI dla hoteli: HM Karpacz, HDC Częstochowa, HM Opole, HDSI Poznań, HDSI Opole, HDS Piaseczno.



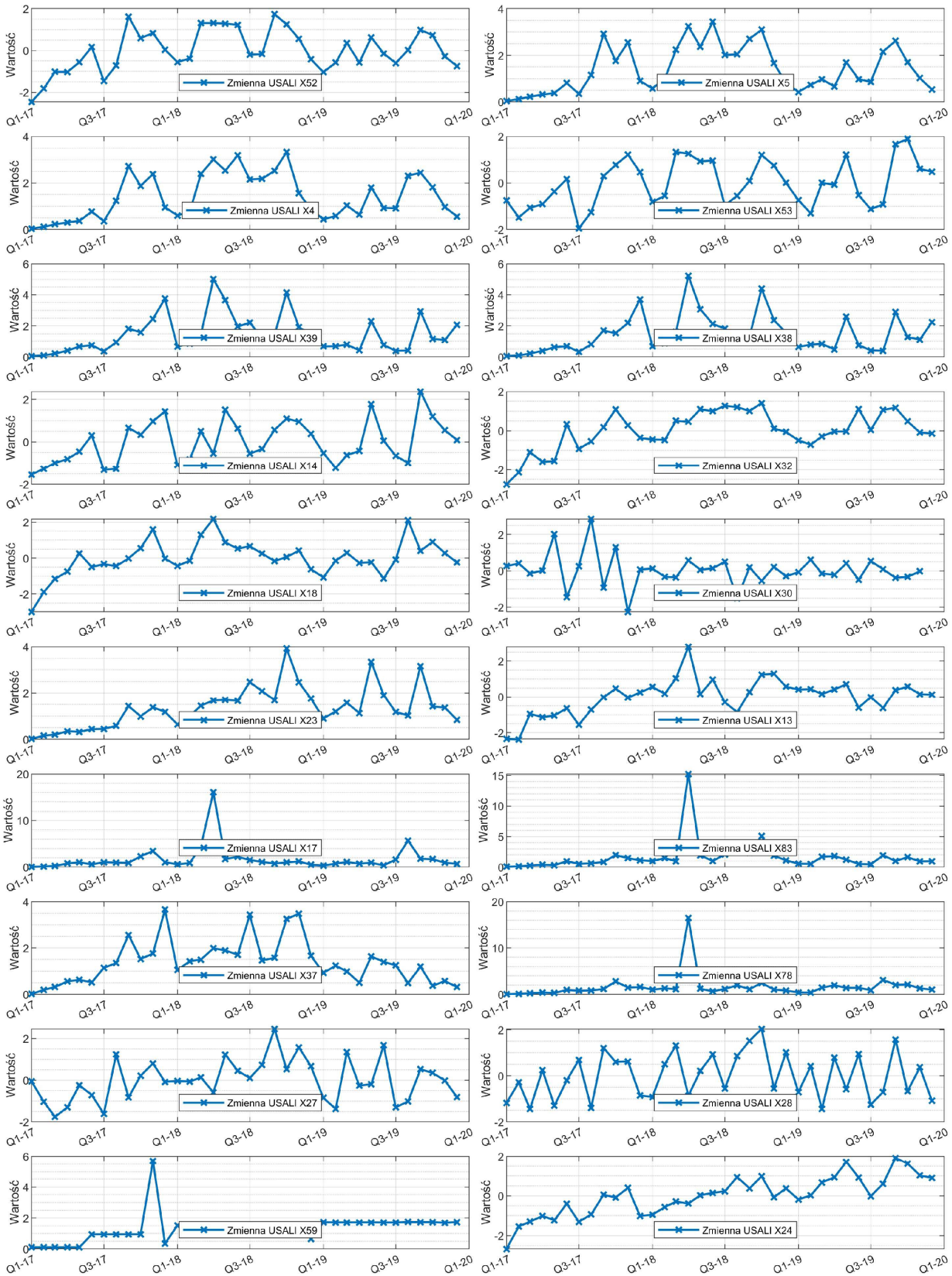
Rys. Z3.1 Zestawienie zbiorcze przebiegów czasowych rozpatrywanych w pracy zmiennych USALI. Wykresy dotyczą danych pozyskanych z hotelu HMC Częstochowa. Źródło: Opracowanie własne.



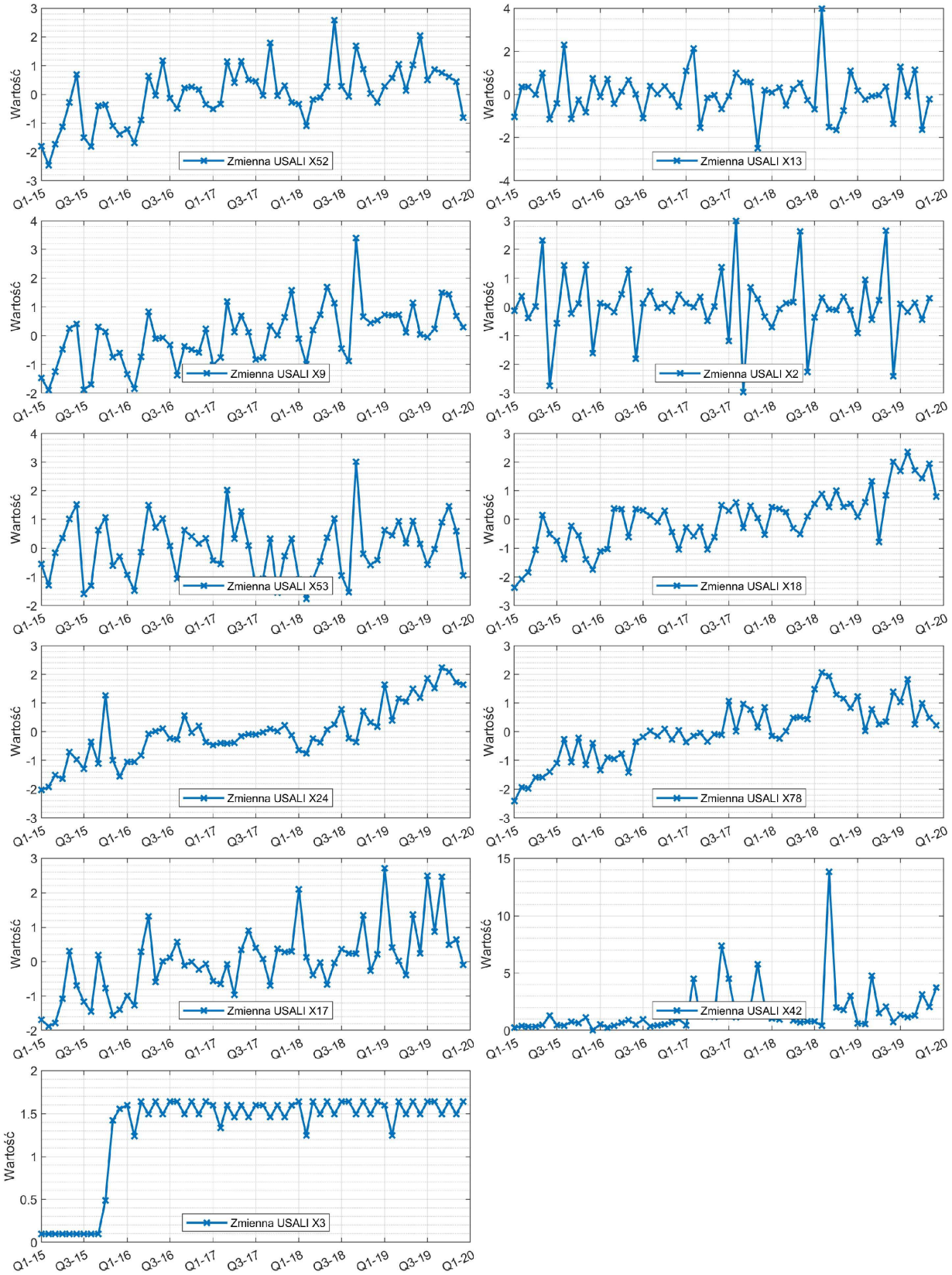
Rys. Z3.2 Zestawienie zbiorcze przebiegów czasowych rozpatrywanych w pracy zmiennych USALI. Wykresy dotyczą danych pozyskanych z hotelu HM Opole. Źródło: Opracowanie własne.



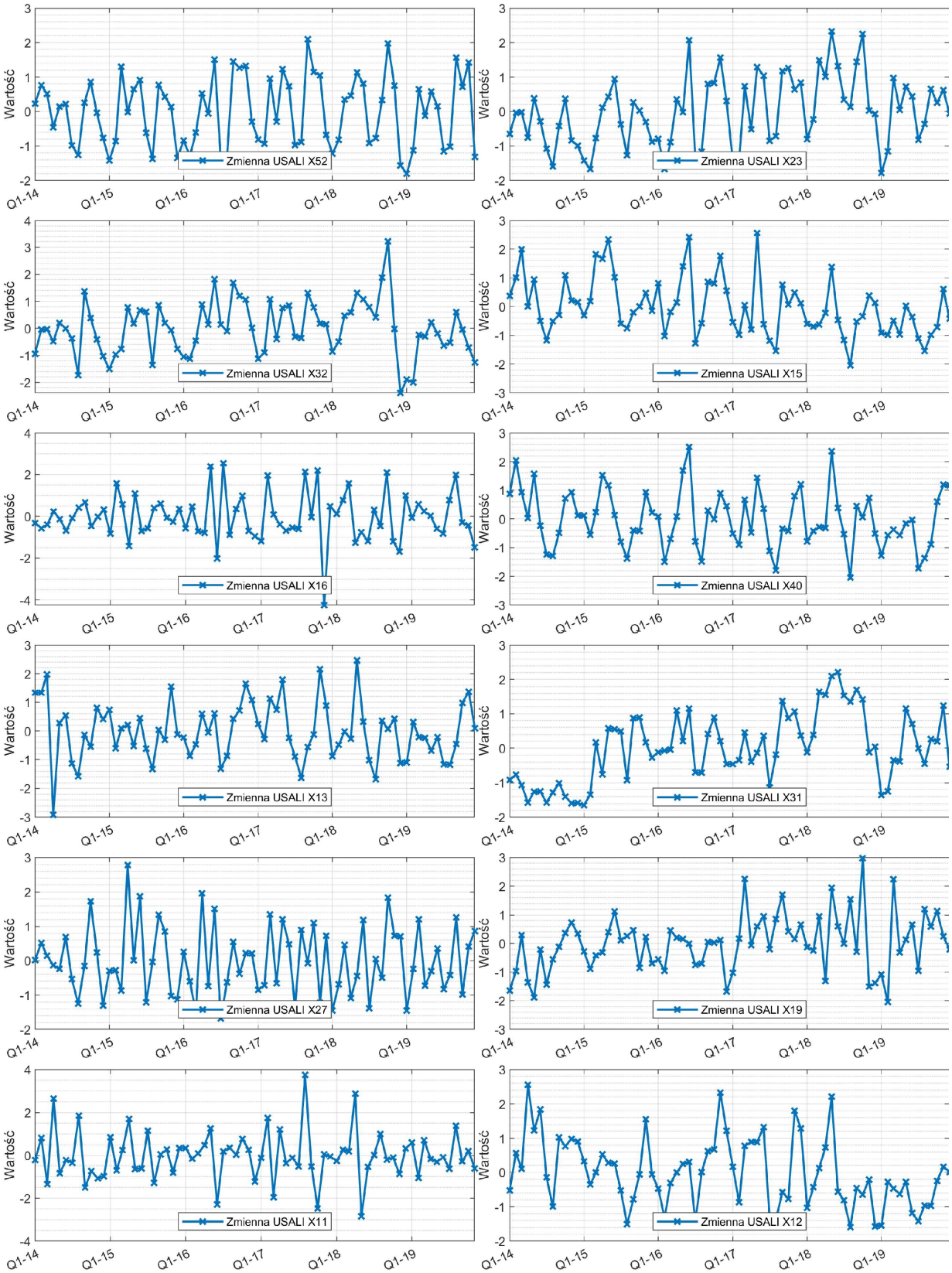
Rys. Z3.3 Zestawienie zbiorcze przebiegów czasowych rozpatrywanych w pracy zmiennych USALI. Wykresy dotyczą danych pozyskanych z hotelu HM Karpacz. Źródło: Opracowanie własne.



Rys. Z3.4 Zestawienie zbiorcze przebiegów czasowych rozpatrywanych w pracy zmiennych USALI. Wykresy dotyczą danych pozyskanych z hotelu HDSP Poznań. Źródło: Opracowanie własne.



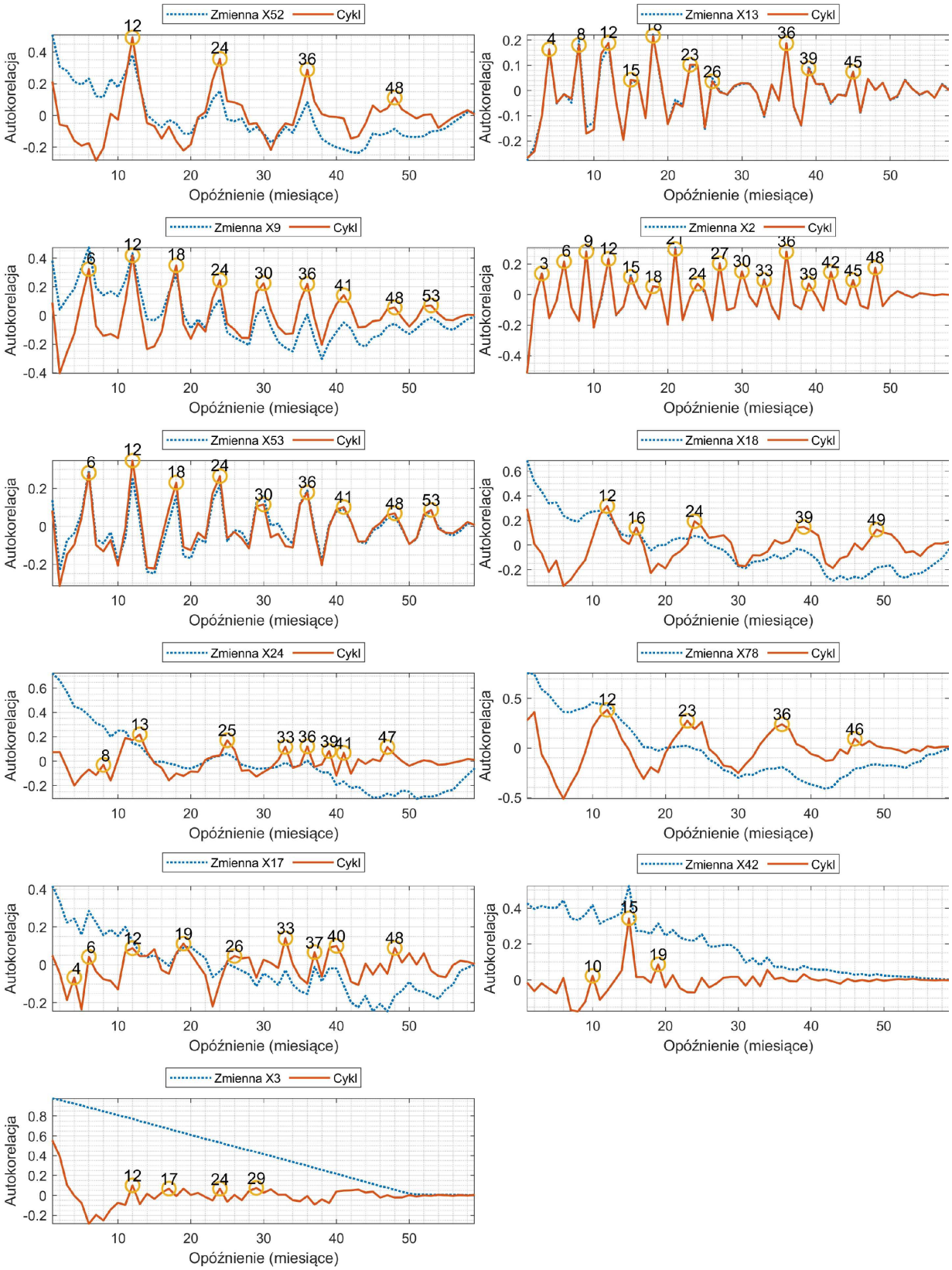
Rys. Z3.5 Zestawienie zbiorcze przebiegów czasowych rozpatrywanych w pracy zmiennych USALI. Wykresy dotyczą danych pozyskanych z hotelu HDSP Opole. Źródło: Opracowanie własne.



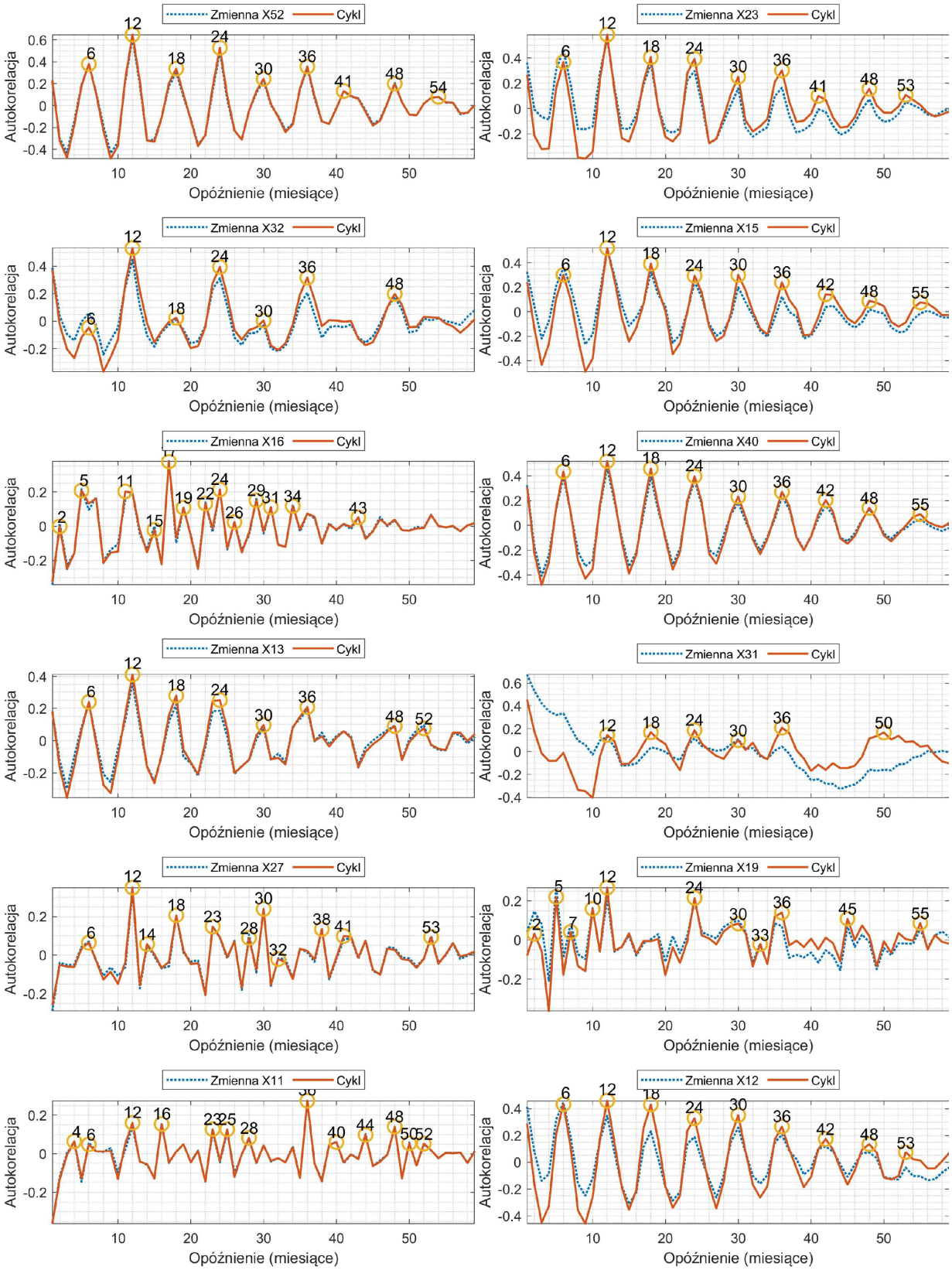
Rys. Z3.6 Zestawienie zbiorcze przebiegów czasowych rozpatrywanych w pracy zmiennych USALI. Wykresy dotyczą danych pozyskanych z hotelu HDS Piaseczno. Źródło: Opracowanie własne.

Załącznik Z4

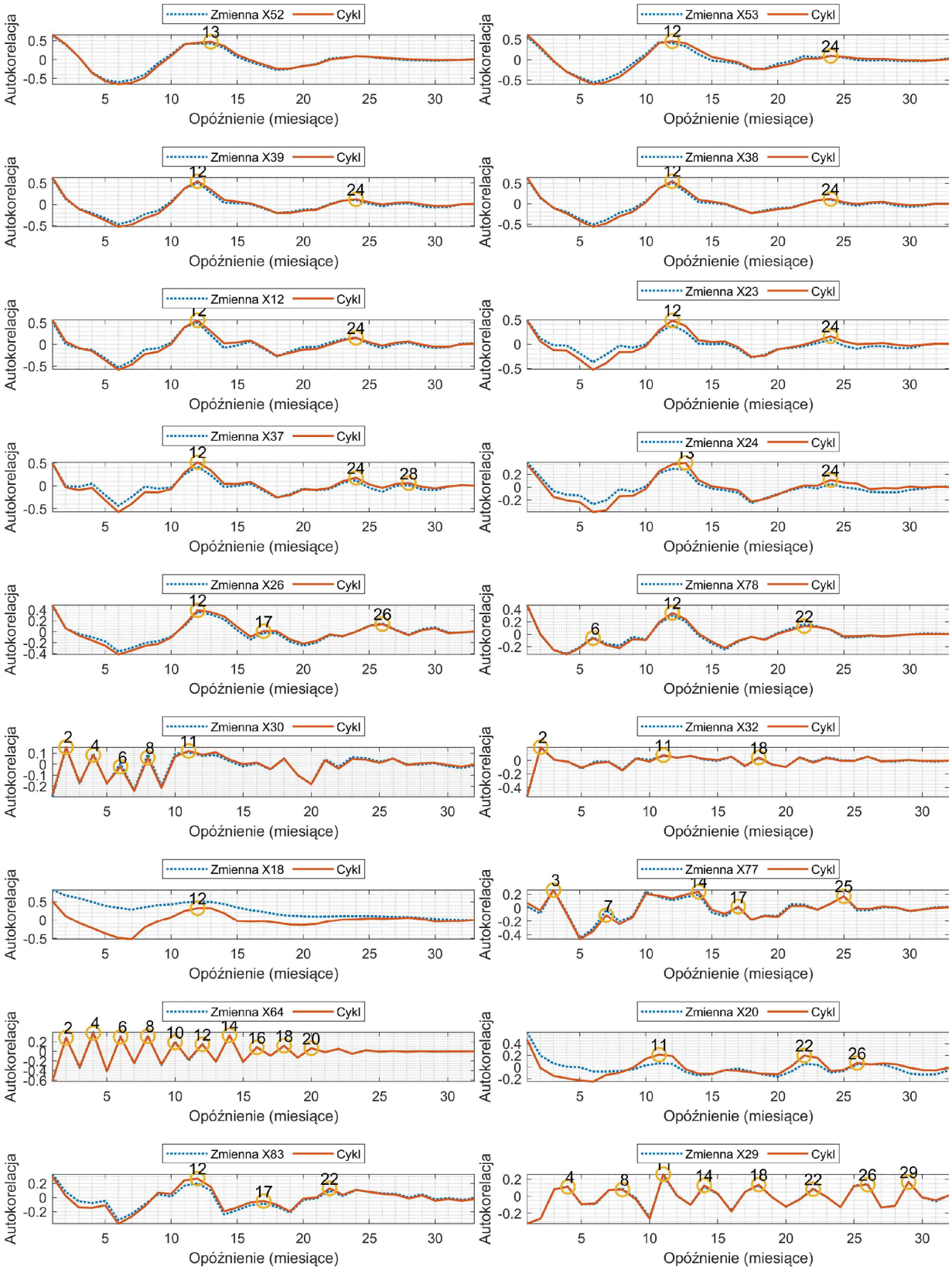
Na rysunkach Z4.1– Z4.6 przedstawiono wykresy autokorelacji rozpatrywanych w pracy zmiennych USALI dla hoteli: HM Karpacz, HDC Częstochowa, HM Opole, HDSI Poznań, HDSI Opole, HDS Piaseczno.



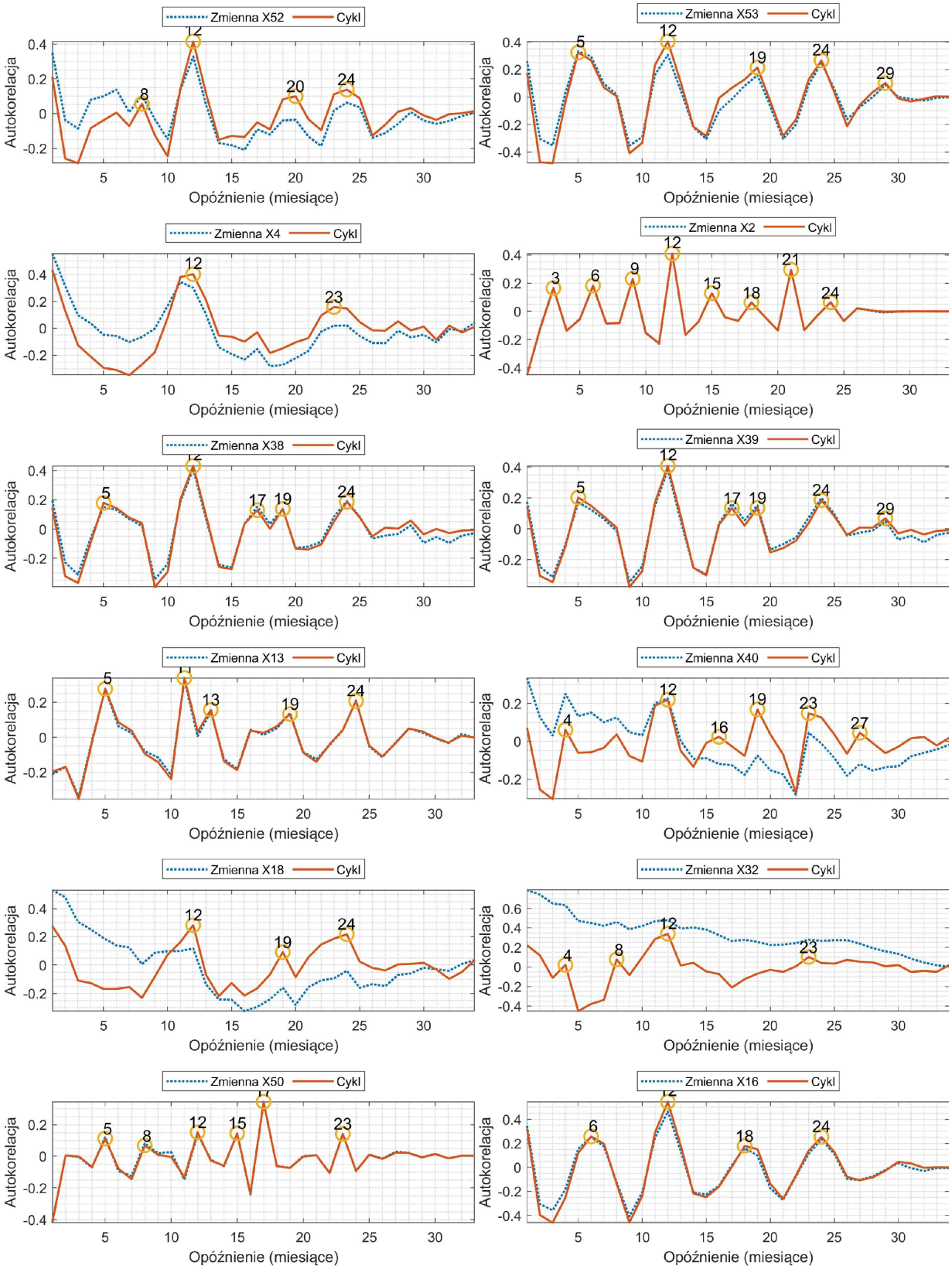
Rys. Z4.1 Zestawienie zbiorcze wykresów autokorelacji rozpatrywanych w pracy zmiennych USALI. Wykresy dotyczą danych pozyskanych z hotelu HDSP Opole. Źródło: Opracowanie własne.



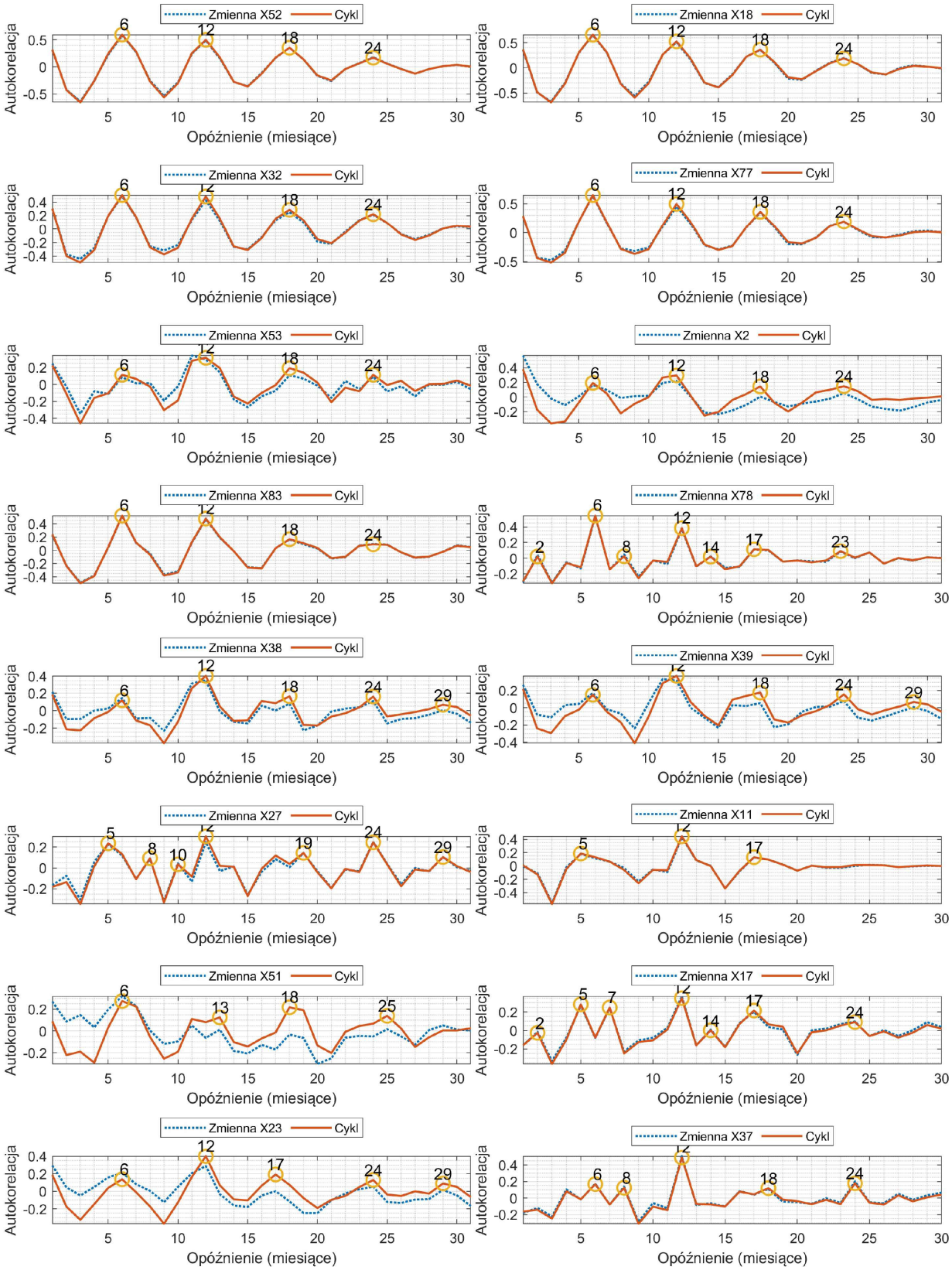
Rys. Z4.2 Zestawienie zbiorcze wykresów autokorelacji rozpatrywanych w pracy zmiennych USALI. Wykresy dotyczą danych pozyskanych z hotelu HDS Piaseczno. Źródło: Opracowanie własne.



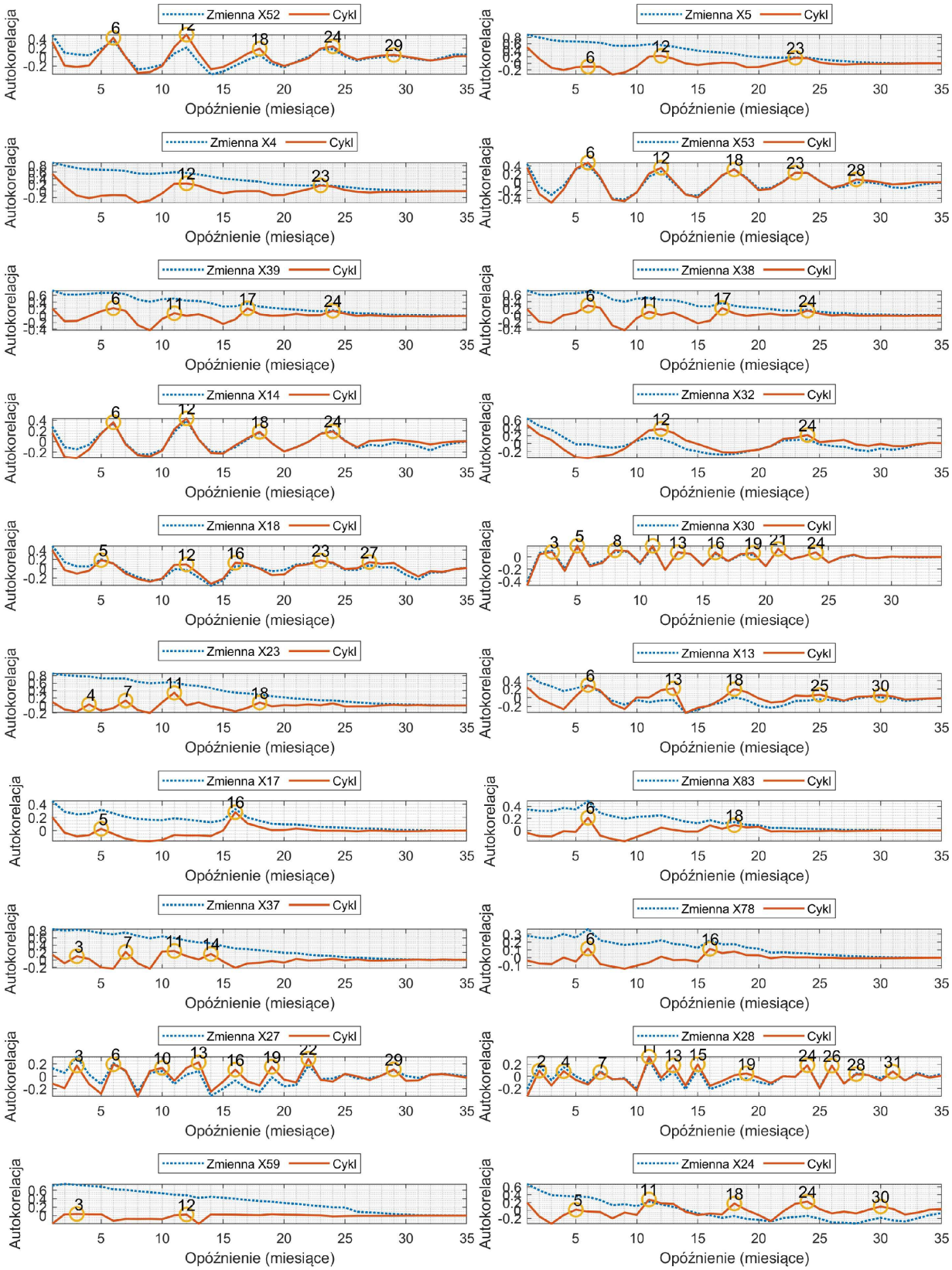
Rys. Z4.3 Zestawienie zbiorcze wykresów autokorelacji rozpatrywanych w pracy zmiennych USALI. Wykresy dotyczą danych pozyskanych z hotelu HMC Częstochowa. Źródło: Opracowanie własne.



Rys. Z4.4 Zestawienie zbiorcze wykresów autokorelacji rozpatrywanych w pracy zmiennych USALI. Wykresy dotyczą danych pozyskanych z hotelu HM Opole. Źródło: Opracowanie własne.



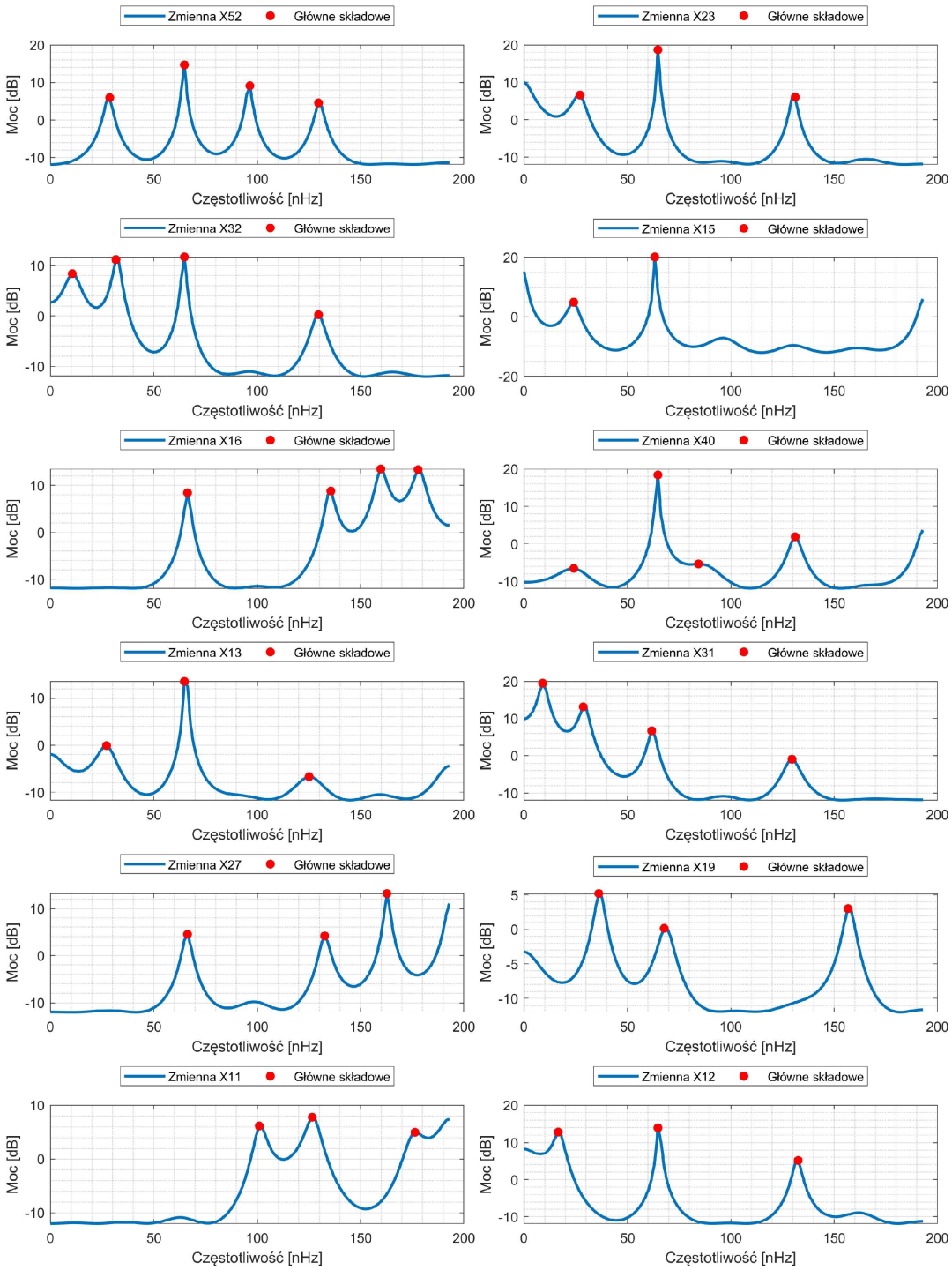
Rys. Z4.5 Zestawienie zbiorcze wykresów autokorelacji rozpatrywanych w pracy zmiennych USALI. Wykresy dotyczą danych pozyskanych z hotelu HM Karpacz. Źródło: Opracowanie własne.



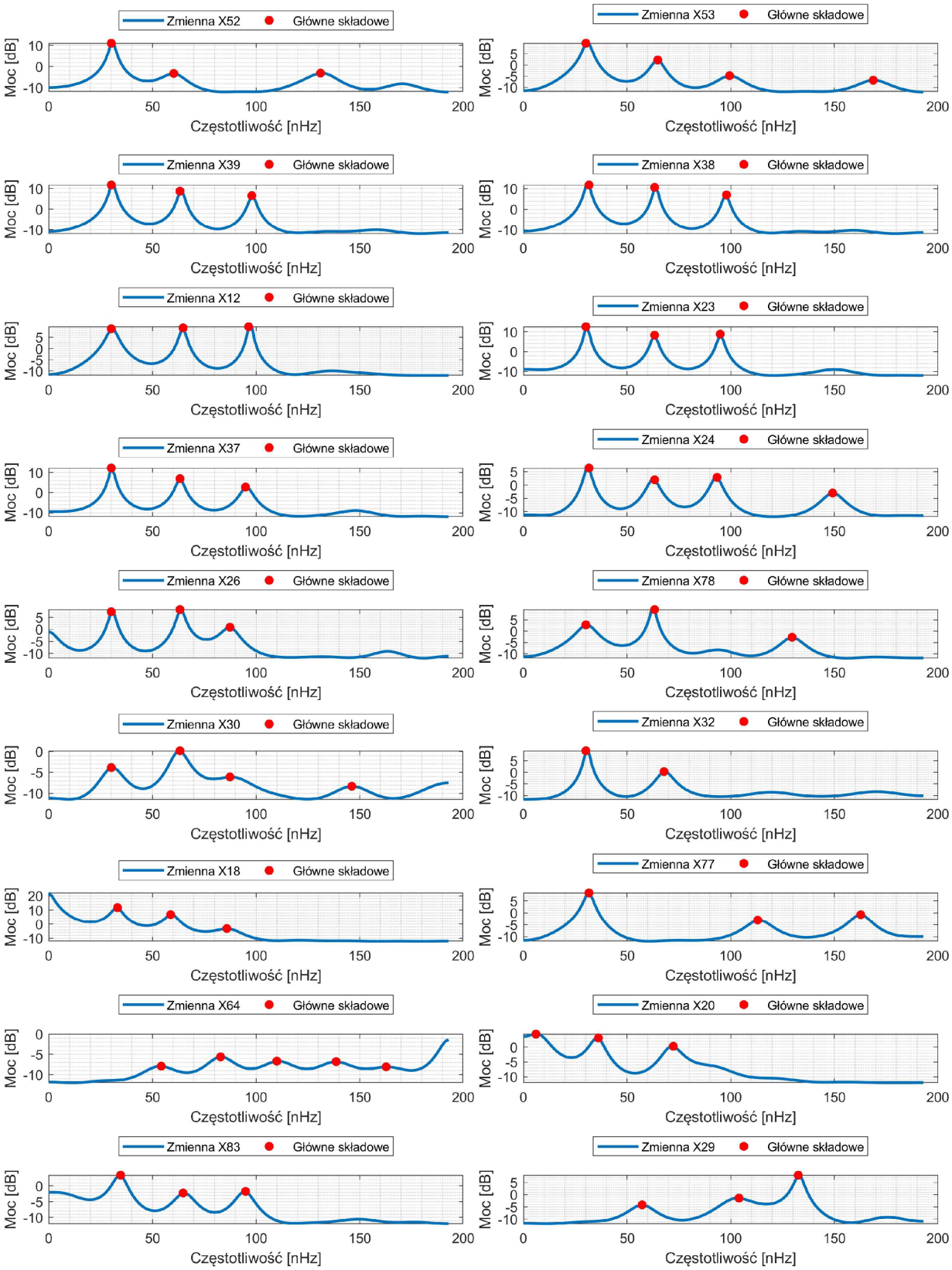
Rys. Z4.6 Zestawienie zbiorcze wykresów autokorelacji rozpatrywanych w pracy zmiennych USALI. Wykresy dotyczą danych pozyskanych z hotelu HDP Poznań. Źródło: Opracowanie własne.

Załącznik Z5

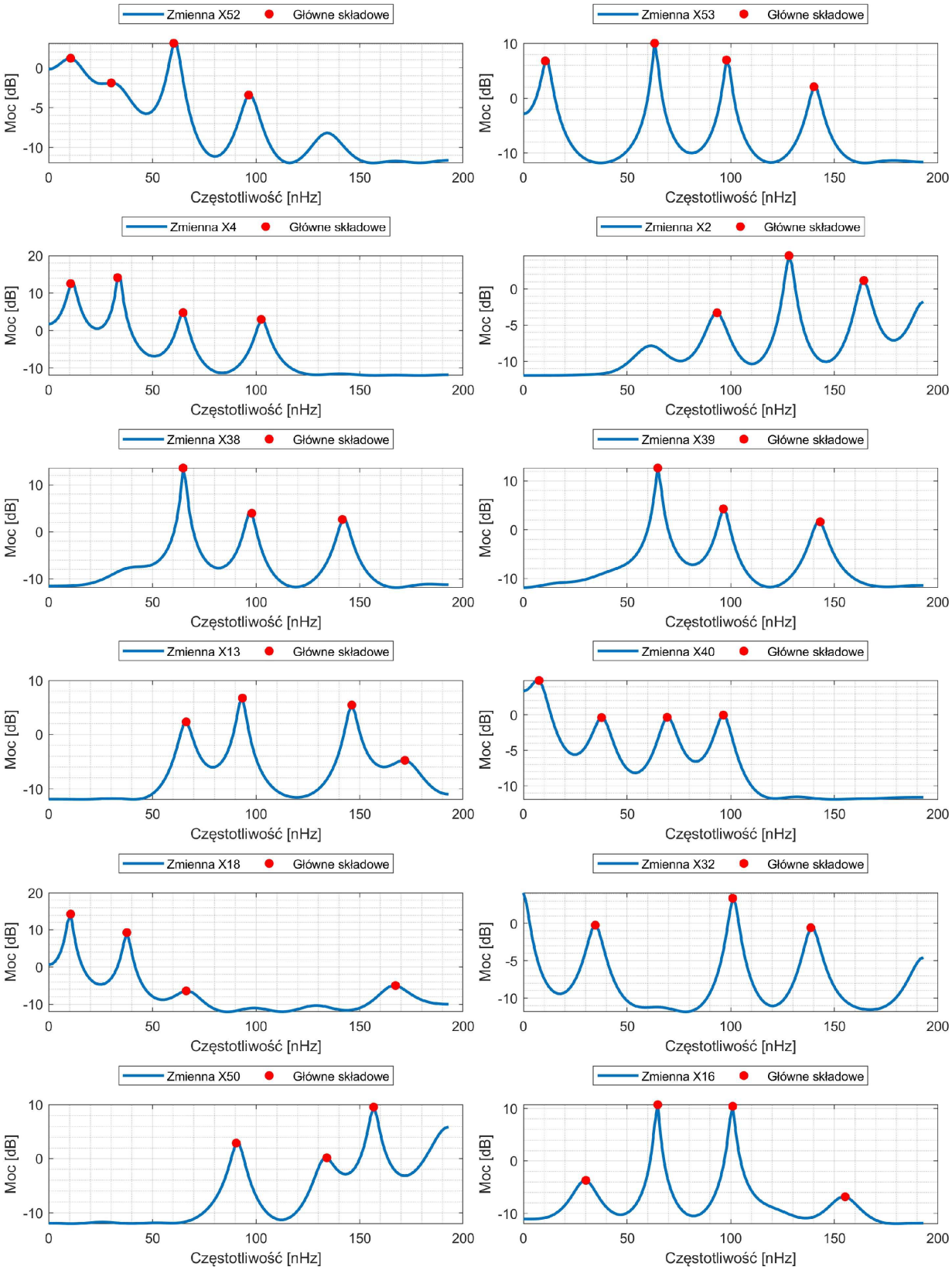
Na rysunkach Z5.1– Z5.6 przedstawiono wykresy analizy PMUSIC, obrazujące składowe częstotliwościowe zawarte w widmach rozpatrywanych w pracy zmiennych USALI dla hoteli: HM Karpacz, HDC Częstochowa, HM Opole, HDSI Poznań, HDSI Opole, HDS Piaseczno.



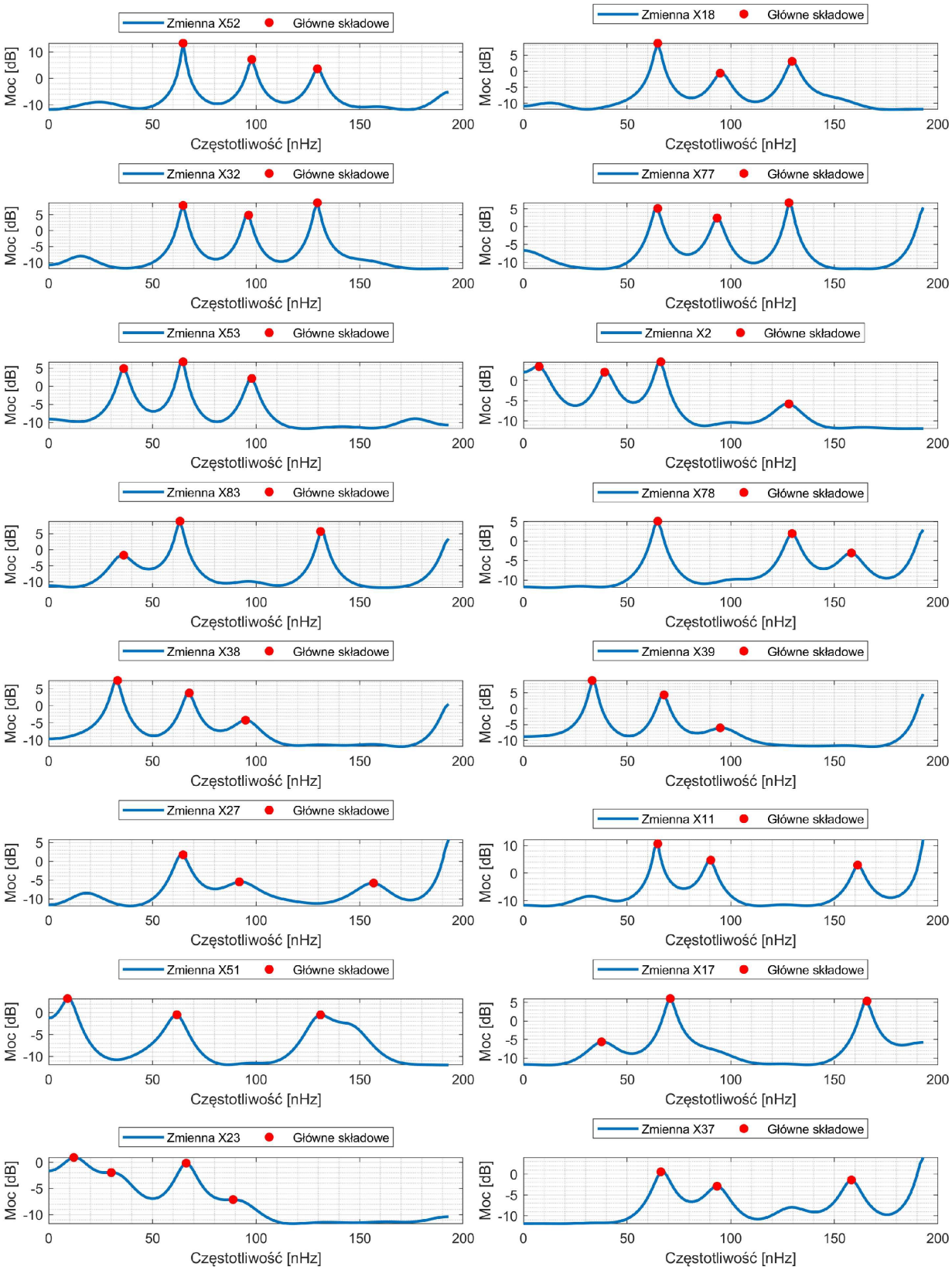
Rys. Z5.1 Zestawienie zbiorcze widmo PMUSIC obliczonych dla rozpatrywanych w pracy zmiennych USALI. Wykresy dotyczą danych pozyskanych z hotelu HDS Piaseczno. Źródło: Opracowanie własne.



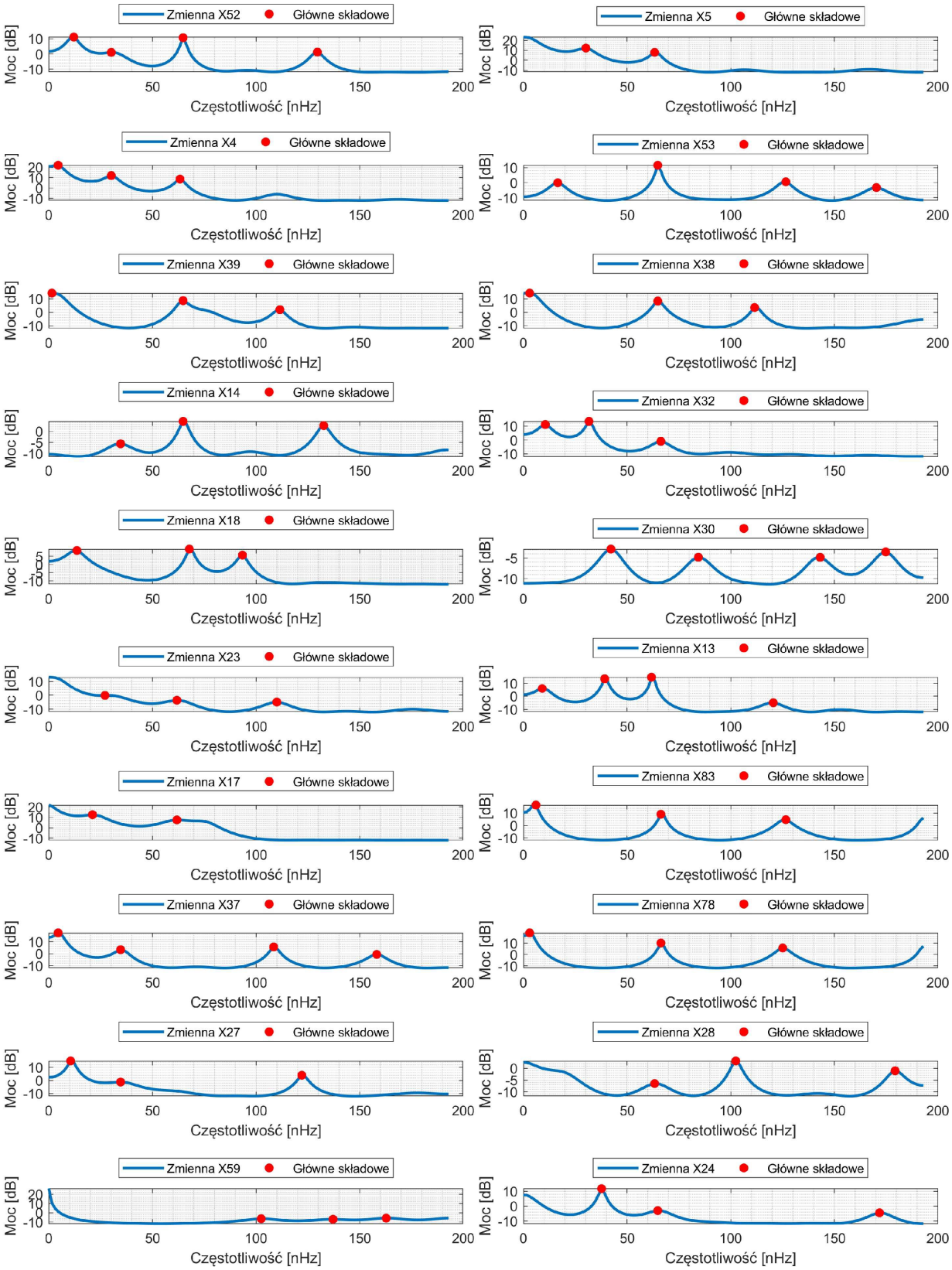
Rys. Z5.2 Zestawienie zbiorcze widmo PMUSIC obliczonych dla rozpatrywanych w pracy zmiennych USALI. Wykresy dotyczą danych pozyskanych z hotelu HMC Częstochowa. Źródło: Opracowanie własne.



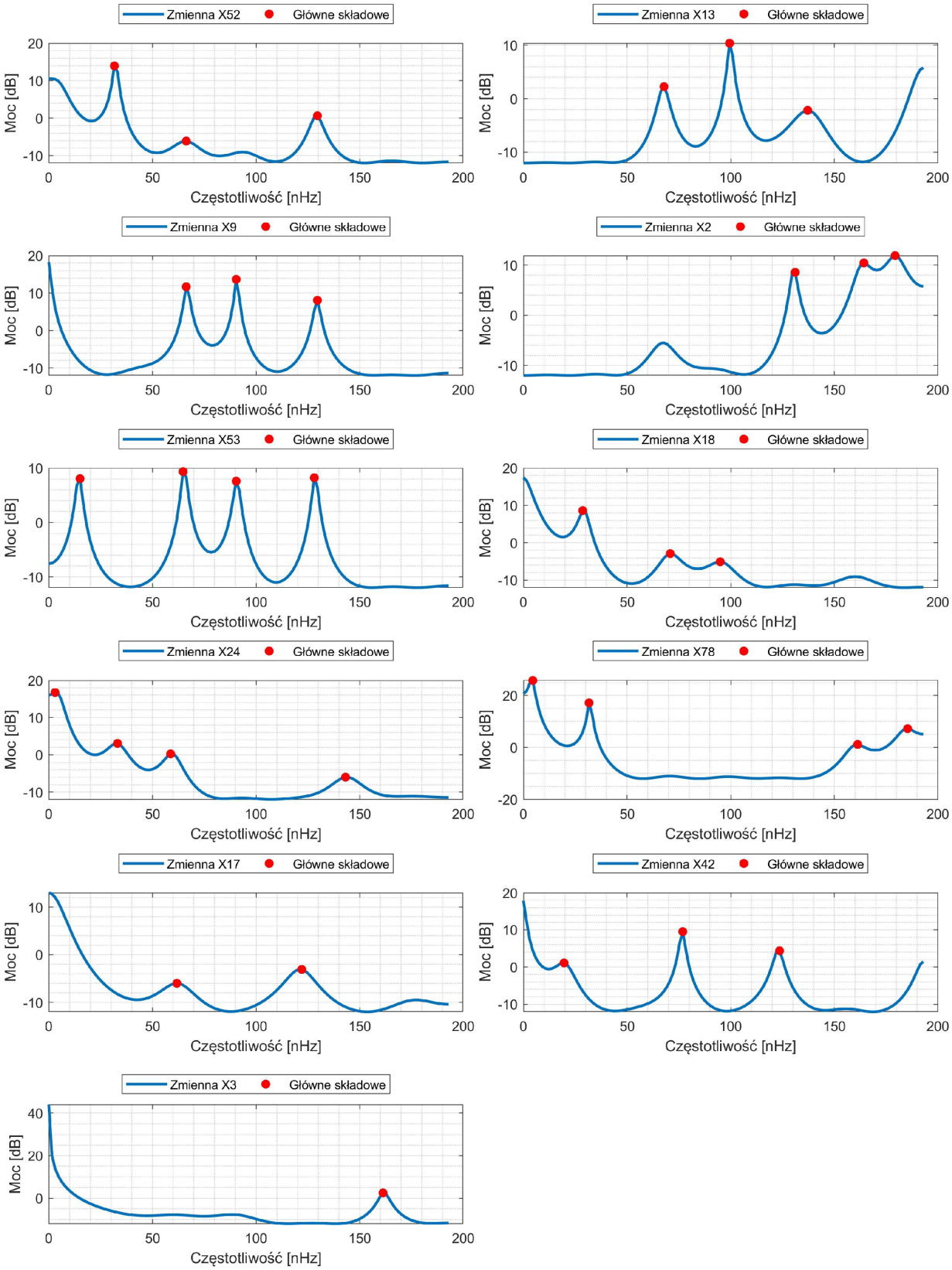
Rys. Z5.3 Zestawienie zbiorcze widmo PMUSIC obliczonych dla rozpatrywanych w pracy zmiennych USALI. Wykresy dotyczą danych pozyskanych z hotelu HM Opole. Źródło: Opracowanie własne.



Rys. Z5.4 Zestawienie zbiorcze widmo PMUSIC obliczonych dla rozpatrywanych w pracy zmiennych USALI. Wykresy dotyczą danych pozyskanych z hotelu HM Karpacz. Źródło: Opracowanie własne.



Rys. Z5.5 Zestawienie zbiorcze widmo PMUSIC obliczonych dla rozpatrywanych w pracy zmiennych USALI. Wykresy dotyczą danych pozyskanych z hotelu HDSP Poznań. Źródło: Opracowanie własne.



Rys. Z5.6 Zestawienie zbiorcze widmo PMUSIC obliczonych dla rozpatrywanych w pracy zmiennych USALI. Wykresy dotyczą danych pozyskanych z hotelu HDSP Opole. Źródło: Opracowanie własne.

Załącznik Z6

Poniżej przedstawiono równania początkowych modeli ekonometrycznych wraz z estymatami poszczególnych współczynników dla poszczególnych hoteli. W tabelach Z6.1 – Z6.7 zawarto zestawienie zbiorcze dotyczące szacowania modeli regresji.

Model początkowy HDSI Katowice:

$$GOP = \text{Wyraz wolny} + X52 + X9 + X77 + X17 + X51 + X20 + X75 + X58 + X14 + X24$$

Wartości estymat:

· Wyraz wolny: 0,035 (SE = 0,016; t = 2,19; p = 0,033)
· X52: 0,990 (SE = 0,030; t = 32,48; p < 0,0001)
· X9: -0,051 (SE = 0,033; t = -1,56; p = 0,124)
· X77: 0,013 (SE = 0,023; t = 0,57; p = 0,569)
· X17: 0,003 (SE = 0,035; t = 0,09; p = 0,928)
· X51: 0,024 (SE = 0,017; t = 1,42; p = 0,161)
· X20: -0,044 (SE = 0,026; t = -1,66; p = 0,103)
· X75: -0,021 (SE = 0,005; t = -4,31; p < 0,0001)
· X58: 0,029 (SE = 0,029; t = 0,99; p = 0,327)
· X14: 0,059 (SE = 0,021; t = 2,80; p = 0,007)
· X24: -0,025 (SE = 0,018; t = -1,40; p = 0,165)

Tabela Z6.1. Zestawienie zbiorcze wyników początkowego modelu regresji liniowej dla hotelu HDSI Katowice. Źródło: Opracowanie własne.

Podsumowanie wyników początkowego modelu regresji liniowej HDSI Katowice			
Liczba obserwacji	72	Błąd średniokwadratowy	0,116
Stopnie swobody błędu	61	R-kwadrat (R-kwadrat skorygowane)	0,988 (0,986)
Statystyka F względem modelu stałego	517	p	< 0,0001

Model początkowy HDS Piaseczno:

$$GOP = Wyraz\ wolny + X52 + X23 + X32 + X15 + X16 + X40 + X13 + X31 + X27 + X19 + X11 + X12$$

Wartości estymat:
· Wyraz wolny: 0,001 (SE = 0,013; t = 0,063; p = 0,950)
· X52: 0,991 (SE = 0,028; t = 35,11; p < 0,0001)
· X23: -0,162 (SE = 0,037; t = -4,33; p < 0,0001)
· X32: 0,072 (SE = 0,025; t = 2,89; p = 0,0053)
· X15: 0,082 (SE = 0,025; t = 3,28; p = 0,0017)

· X16: 0,022 (SE = 0,014; t = 1,60; p = 0,116)
· X40: -0,016 (SE = 0,030; t = -0,52; p = 0,607)
· X13: -0,006 (SE = 0,023; t = -0,28; p = 0,779)
· X31: 0,028 (SE = 0,025; t = 1,11; p = 0,271)
· X27: 0,006 (SE = 0,018; t = 0,33; p = 0,741)
· X19: 0,037 (SE = 0,017; t = 2,11; p = 0,039)
· X11: -0,007 (SE = 0,014; t = -0,47; p = 0,637)
· X12: 0,028 (SE = 0,021; t = 1,35; p = 0,181)

Tabela Z6.2. Zestawienie zbiorcze wyników początkowego modelu regresji liniowej dla hotelu HDS Piaseczno. Źródło: Opracowanie własne.

Podsumowanie wyników początkowego modelu regresji liniowej HDS Piaseczno			
Liczba obserwacji	72	Błąd średniokwadratowy	0,114
Stopnie swobody błędu	59	R-kwadrat (R-kwadrat skorygowane)	0,989 (0,987)
Statystyka F względem modelu stałego	454	p	< 0,0001

Model początkowy HDSP Opole:

GOP = Wyraz wolny + X52 + X13 + X9 + X2 + X53 + X18 + X24 + X78 + X17 + X42 + X3

Wartości estymat:
· Wyraz wolny: -0,051 (SE = 0,056; t = -0,92; p = 0,365)
· X52: 1,148 (SE = 0,033; t = 34,94; p < 0,0001)
· X13: 0,021 (SE = 0,017; t = 1,18; p = 0,246)
· X9: -0,051 (SE = 0,035; t = -1,44; p = 0,157)
· X2: 0,018 (SE = 0,016; t = 1,14; p = 0,261)
· X53: 0,006 (SE = 0,033; t = 0,17; p = 0,867)
· X18: -0,032 (SE = 0,031; t = -1,04; p = 0,304)
· X24: -0,038 (SE = 0,027; t = -1,40; p = 0,167)
· X78: -0,175 (SE = 0,033; t = -5,22; p < 0,0001)
· X17: 0,024 (SE = 0,025; t = 1,00; p = 0,324)
· X42: 0,014 (SE = 0,009; t = 1,61; p = 0,114)
· X3: 0,021 (SE = 0,039; t = 0,52; p = 0,604)

Tabela Z6.3. Zestawienie zbiorcze wyników początkowego modelu regresji liniowej dla hotelu HDSP Opole. Źródło: Opracowanie własne.

Podsumowanie wyników początkowego modelu regresji liniowej HDSP Opole			
Liczba obserwacji	60	Błąd średniokwadratowy	0,115
Stopnie swobody błędu	48	R-kwadrat (R-kwadrat skorygowane)	0,988 (0,987)
Statystyka F względem modelu stałego	398	p	< 0,0001

Model początkowy HDSP Poznań:

$$GOP = Wyraz\ wolny + X52 + X5 + X4 + X53 + X39 + X38 + X14 + X32 + X18 + X30 + X23 + X13 + X17 + X83 + X37 + X78 + X27 + X28 + X59 + X24$$

Wartości estymat:
- Wyraz wolny: 0,027 (SE = 0,101; t = 0,27; p = 0,791)
- X52: 1,065 (SE = 0,082; t = 13,03; p < 0,0001)
- X5: 0,515 (SE = 0,244; t = 2,11; p = 0,052)
- X4: -0,447 (SE = 0,279; t = -1,60; p = 0,130)
- X53: 0,182 (SE = 0,092; t = 1,98; p = 0,067)
- X39: 0,331 (SE = 0,186; t = 1,78; p = 0,096)
- X38: -0,425 (SE = 0,206; t = -2,07; p = 0,056)
- X14: -0,035 (SE = 0,098; t = -0,36; p = 0,726)
- X32: -0,029 (SE = 0,067; t = -0,43; p = 0,676)
- X18: -0,075 (SE = 0,069; t = -1,09; p = 0,292)
- X30: 0,032 (SE = 0,043; t = 0,74; p = 0,472)
- X23: 0,017 (SE = 0,086; t = 0,20; p = 0,842)
- X13: -0,053 (SE = 0,064; t = -0,83; p = 0,420)
- X17: 0,044 (SE = 0,052; t = 0,85; p = 0,410)
- X83: 0,034 (SE = 0,061; t = 0,56; p = 0,582)
- X37: 0,040 (SE = 0,066; t = 0,61; p = 0,552)
- X78: -0,093 (SE = 0,078; t = -1,19; p = 0,251)
- X27: -0,031 (SE = 0,037; t = -0,85; p = 0,409)
- X28: -0,029 (SE = 0,031; t = -0,95; p = 0,359)
- X59: -0,032 (SE = 0,043; t = -0,75; p = 0,463)
- X24: -0,101 (SE = 0,084; t = -1,20; p = 0,251)

Tabela Z6.4. Zestawienie zbiorcze wyników początkowego modelu regresji liniowej dla hotelu HDSP Poznań. Źródło: Opracowanie własne.

Podsumowanie wyników początkowego modelu regresji liniowej HDSP Poznań			
Liczba obserwacji	36	Błąd średniokwadratowy	0,112
Stopnie swobody błędu	15	R-kwadrat (R-kwadrat skorygowane)	0,994 (0,985)
Statystyka F względem modelu stałego	116	p	< 0,0001

Model początkowy: HM Opole

$$GOP = Wyraz\ wolny + X52 + X53 + X4 + X2 + X38 + X39 + X13 + X40 + X18 + X32 + X50 + X16$$

Wartości estymat:
- Wyraz wolny: -0,131 (SE = 0,063; t = -2,08; p = 0,049)
- X52: 1,114 (SE = 0,100; t = 11,18; p < 0,0001)
- X53: 0,098 (SE = 0,073; t = 1,34; p = 0,194)
- X4: -0,086 (SE = 0,063; t = -1,37; p = 0,185)

· X2: 0,001 (SE = 0,034; t = 0,02; p = 0,983)
· X38: 0 (SE = 0; t = NaN; p = NaN)
· X39: -0,106 (SE = 0,052; t = -2,02; p = 0,055)
· X13: 0,057 (SE = 0,025; t = 2,27; p = 0,034)
· X40: -0,038 (SE = 0,040; t = -0,94; p = 0,359)
· X18: -0,127 (SE = 0,050; t = -2,52; p = 0,019)
· X32: 0,095 (SE = 0,042; t = 2,23; p = 0,036)
· X50: 0,006 (SE = 0,027; t = 0,22; p = 0,828)
· X16: -0,044 (SE = 0,037; t = -1,17; p = 0,255)

Tabela Z6.5. Zestawienie zbiorcze wyników początkowego modelu regresji liniowej dla hotelu HM Opole. Źródło: Opracowanie własne.

Podsumowanie wyników początkowego modelu regresji liniowej HM Opole			
Liczba obserwacji	35	Błąd średniokwadratowy	0,149
Stopnie swobody błędu	23	R-kwadrat (R-kwadrat skorygowane)	0,985 (0,978)
Statystyka F względem modelu stałego	138	p	< 0,0001

Model początkowy HMC Częstochowa:

$$GOP = \text{Wyraz wolny} + X52 + X53 + X39 + X38 + X12 + X23 + X37 + X24 + X26 + X78 + X30 + X32 + X18 + X77 + X64 + X20 + X83 + X29$$

Wartości estymat:
· Wyraz wolny: 0,079 (SE = 0,070; t = 1,12; p = 0,282)
· X52: 1,126 (SE = 0,165; t = 6,83; p < 0,0001)
· X53: 0,015 (SE = 0,186; t = 0,08; p = 0,935)
· X39: -0,503 (SE = 0,346; t = -1,45; p = 0,167)
· X38: 0,296 (SE = 0,359; t = 0,83; p = 0,422)
· X12: -0,053 (SE = 0,170; t = -0,31; p = 0,758)
· X23: 0,323 (SE = 0,581; t = 0,56; p = 0,587)
· X37: -0,016 (SE = 0,463; t = -0,04; p = 0,972)
· X24: -0,006 (SE = 0,238; t = -0,02; p = 0,982)
· X26: 0,087 (SE = 0,056; t = 1,56; p = 0,139)
· X78: -0,085 (SE = 0,059; t = -1,44; p = 0,172)
· X30: -0,037 (SE = 0,077; t = -0,48; p = 0,639)
· X32: 0,069 (SE = 0,051; t = 1,37; p = 0,192)
· X18: -0,053 (SE = 0,044; t = -1,20; p = 0,250)
· X77: -0,107 (SE = 0,042; t = -2,57; p = 0,021)
· X64: 0,056 (SE = 0,021; t = 2,68; p = 0,017)
· X20: -0,036 (SE = 0,048; t = -0,75; p = 0,467)
· X83: -0,056 (SE = 0,050; t = -1,13; p = 0,276)
· X29: 0,051 (SE = 0,026; t = 1,93; p = 0,073)

Tabela Z6.6. Zestawienie zbiorcze wyników początkowego modelu regresji liniowej dla hotelu HMC Częstochowa. Źródło: Opracowanie własne.

Podsumowanie wyników początkowego modelu regresji liniowej HMC Częstochowa			
Liczba obserwacji	34	Błąd średniokwadratowy	0,137
Stopnie swobody błędu	15	R-kwadrat (R-kwadrat skorygowane)	0,991 (0,981)
Statystyka F względem modelu stałego	96,2	p	< 0,0001

Model początkowy HM Karpacz:

$$GOP = Wyraz\ wolny + X52 + X18 + X32 + X77 + X53 + X2 + X83 + X78 + X38 + X39 + X27 + X11 + X51 + X17 + X23 + X37$$

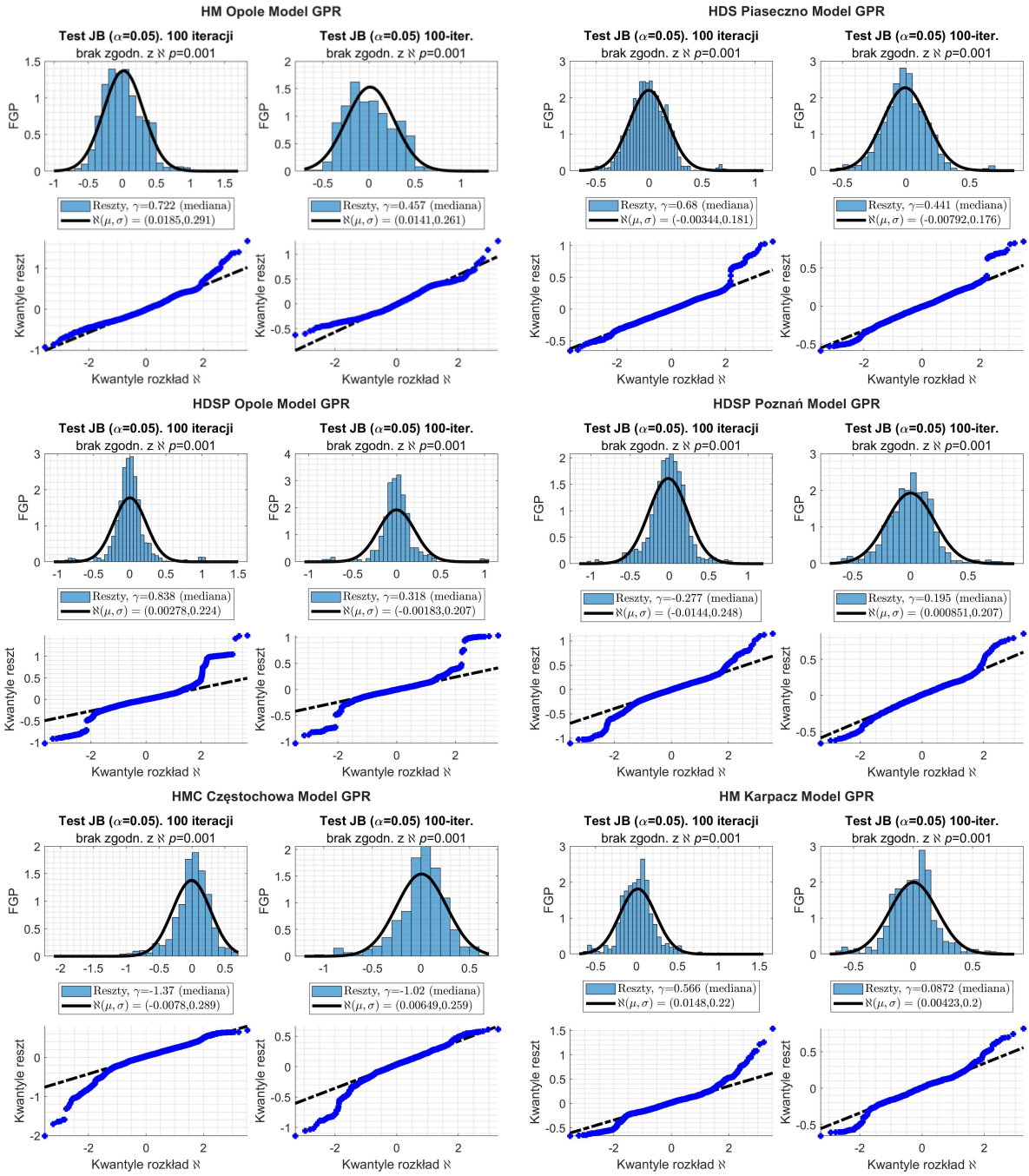
Wartości estymat:
- Wyraz wolny: -0,007 (SE = 0,021; t = -0,33; p = 0,746)
- X52: 1,138 (SE = 0,203; t = 5,61; p < 0,0001)
- X18: -0,083 (SE = 0,119; t = -0,69; p = 0,499)
- X32: 0,103 (SE = 0,070; t = 1,48; p = 0,161)
- X77: -0,155 (SE = 0,075; t = -2,07; p = 0,056)
- X53: 0,017 (SE = 0,119; t = 0,14; p = 0,887)
- X2: 0,026 (SE = 0,087; t = 0,30; p = 0,772)
- X83: -0,012 (SE = 0,063; t = -0,19; p = 0,850)
- X78: 0,051 (SE = 0,038; t = 1,33; p = 0,205)
- X38: 0,160 (SE = 0,164; t = 0,98; p = 0,344)
- X39: -0,158 (SE = 0,161; t = -0,98; p = 0,341)
- X27: 0,008 (SE = 0,035; t = 0,24; p = 0,817)
- X11: 0,016 (SE = 0,037; t = 0,42; p = 0,684)
- X51: -0,029 (SE = 0,039; t = -0,74; p = 0,469)
- X17: -0,007 (SE = 0,047; t = -0,15; p = 0,882)
- X23: -0,015 (SE = 0,066; t = -0,23; p = 0,823)
- X37: 0,042 (SE = 0,048; t = 0,86; p = 0,402)

Tabela Z6.7. Zestawienie zbiorcze wyników początkowego modelu regresji liniowej dla hotelu HM Karpacz. Źródło: Opracowanie własne.

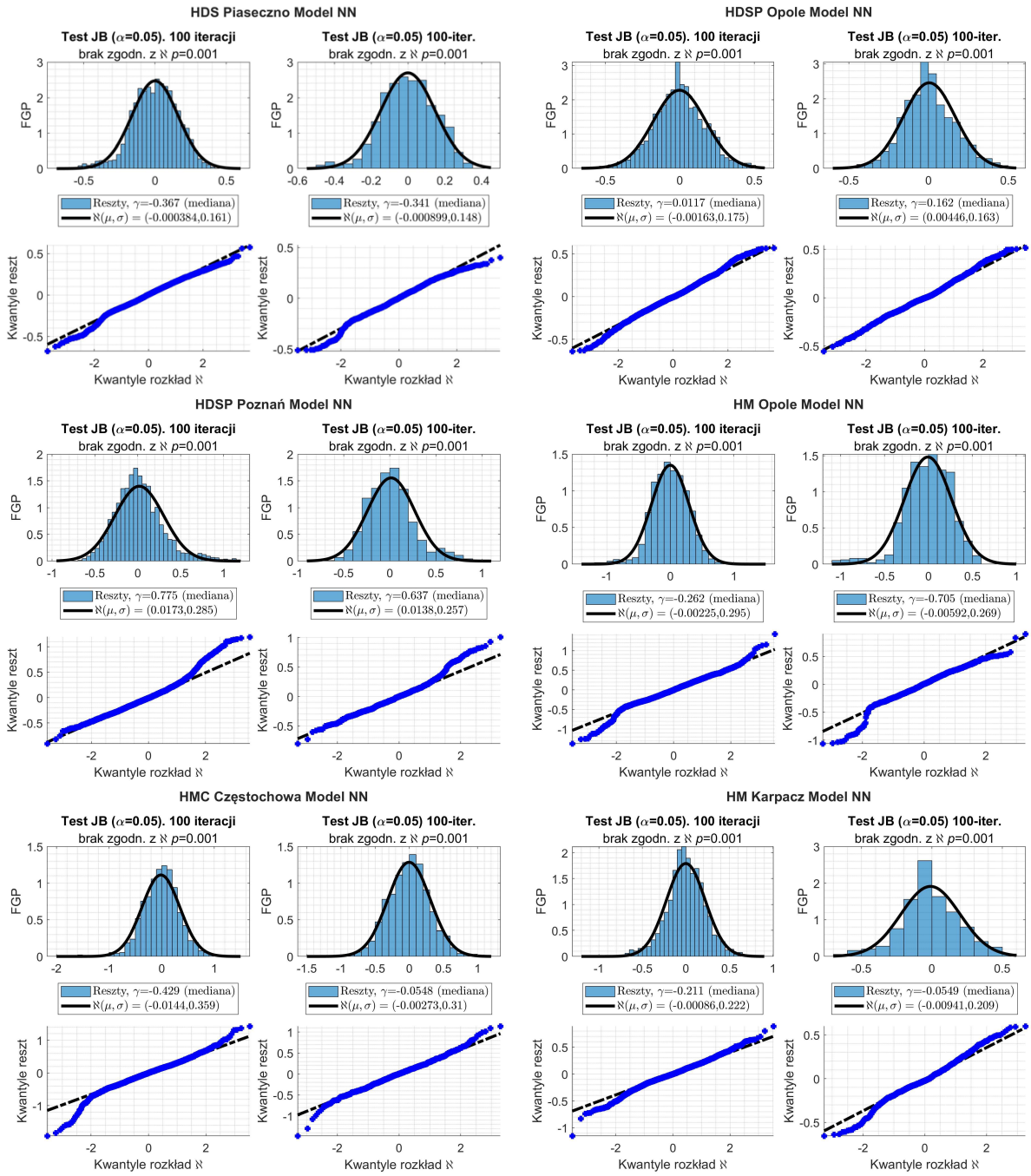
Podsumowanie wyników początkowego modelu regresji liniowej HM Karpacz			
Liczba obserwacji	32	Błąd średniokwadratowy	0,113
Stopnie swobody błędu	15	R-kwadrat (R-kwadrat skorygowane)	0,994 (0,987)
Statystyka F względem modelu stałego	150	p	< 0,0001

Załącznik Z7

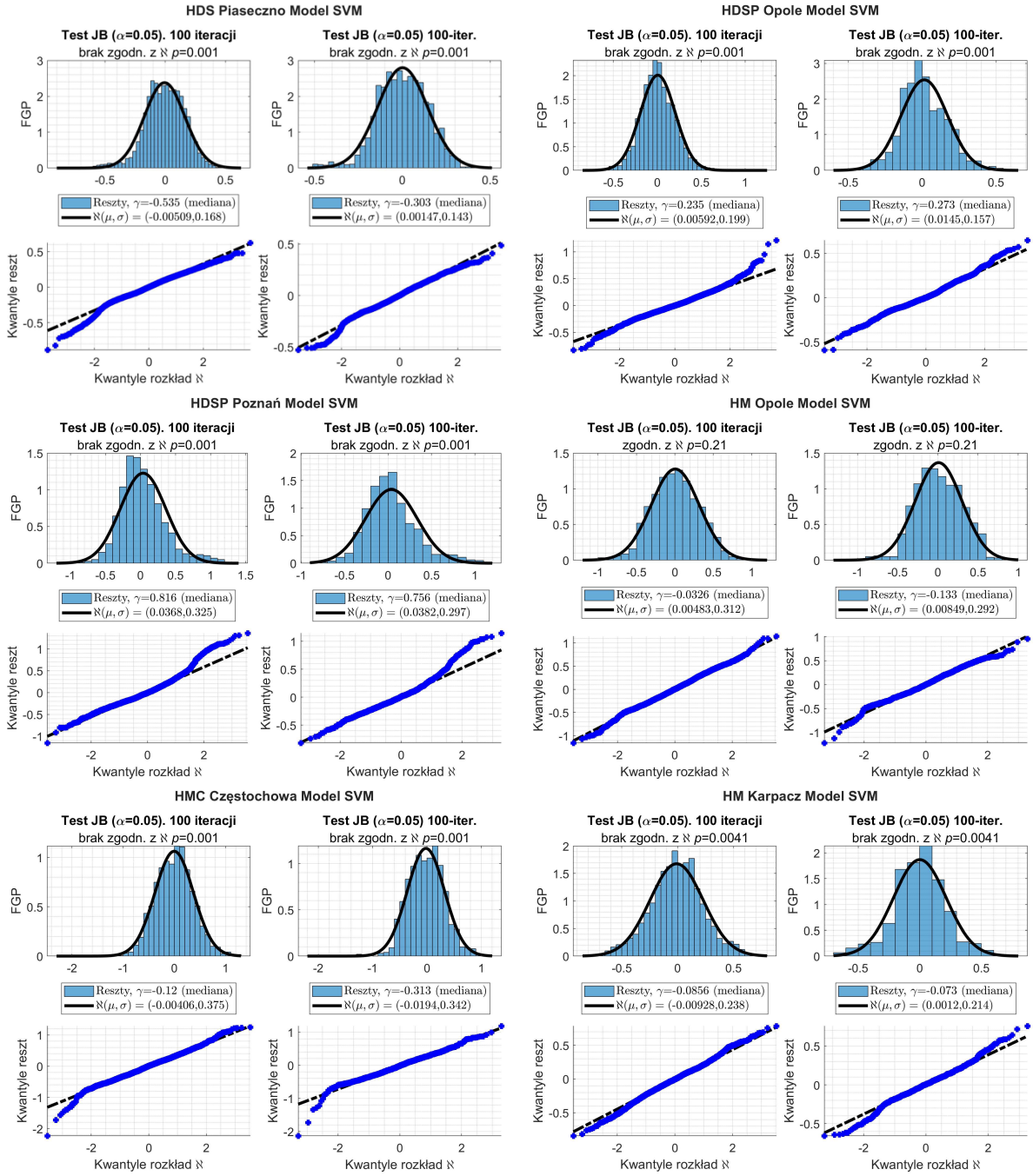
Na rysunkach Z7.1 – Z7.5 przedstawiono zbiorcze zestawienie wyników 100 iteracji odpowiednio uczenia i testowania modeli predykujących zmienną GOP dla poszczególnych hoteli (HM Karpacz, HDC Częstochowa, HM Opole, HDSI Poznań, HDSI Opole, HDS Piaseczno) kolejno dla rozpatrywanych algorytmów: GPR, NN, TREE, SVM i EN.



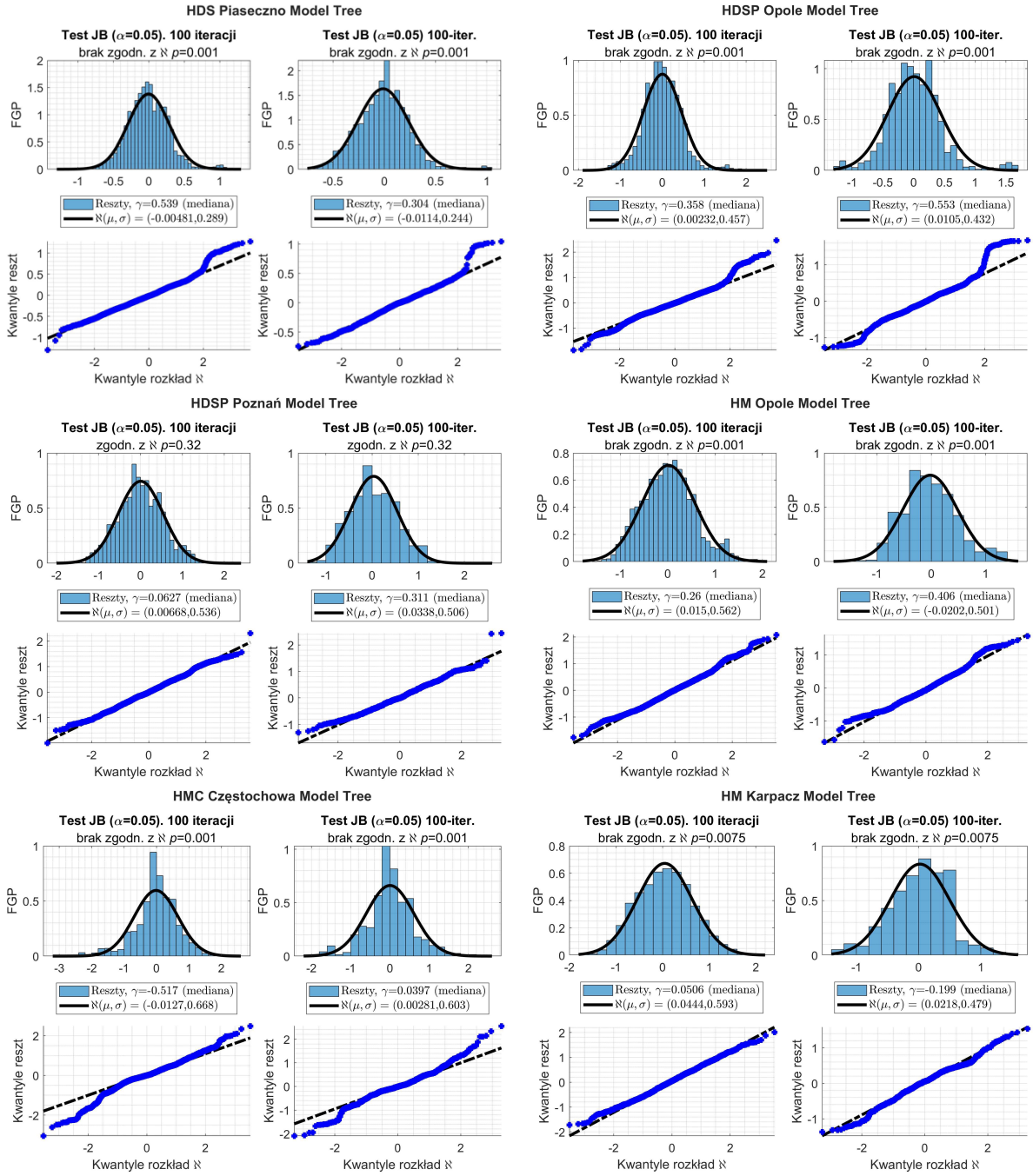
Rys. Z7.1 Zestawienie porównawcze rozkładów gęstości prawdopodobieństwa reszt wraz z wykresami Q–Q, przeprowadzone w 100 iteracjach odpowiednio uczenia i testowania modelu predykującego GOP. Analizy dotyczyły hoteli HM Karpacz, HDC Częstochowa, HM Opole, HDSI Poznań, HDSI Opole i HDS Piaseczno z wykorzystaniem algorytmu GPR. Źródło: Opracowanie własne.



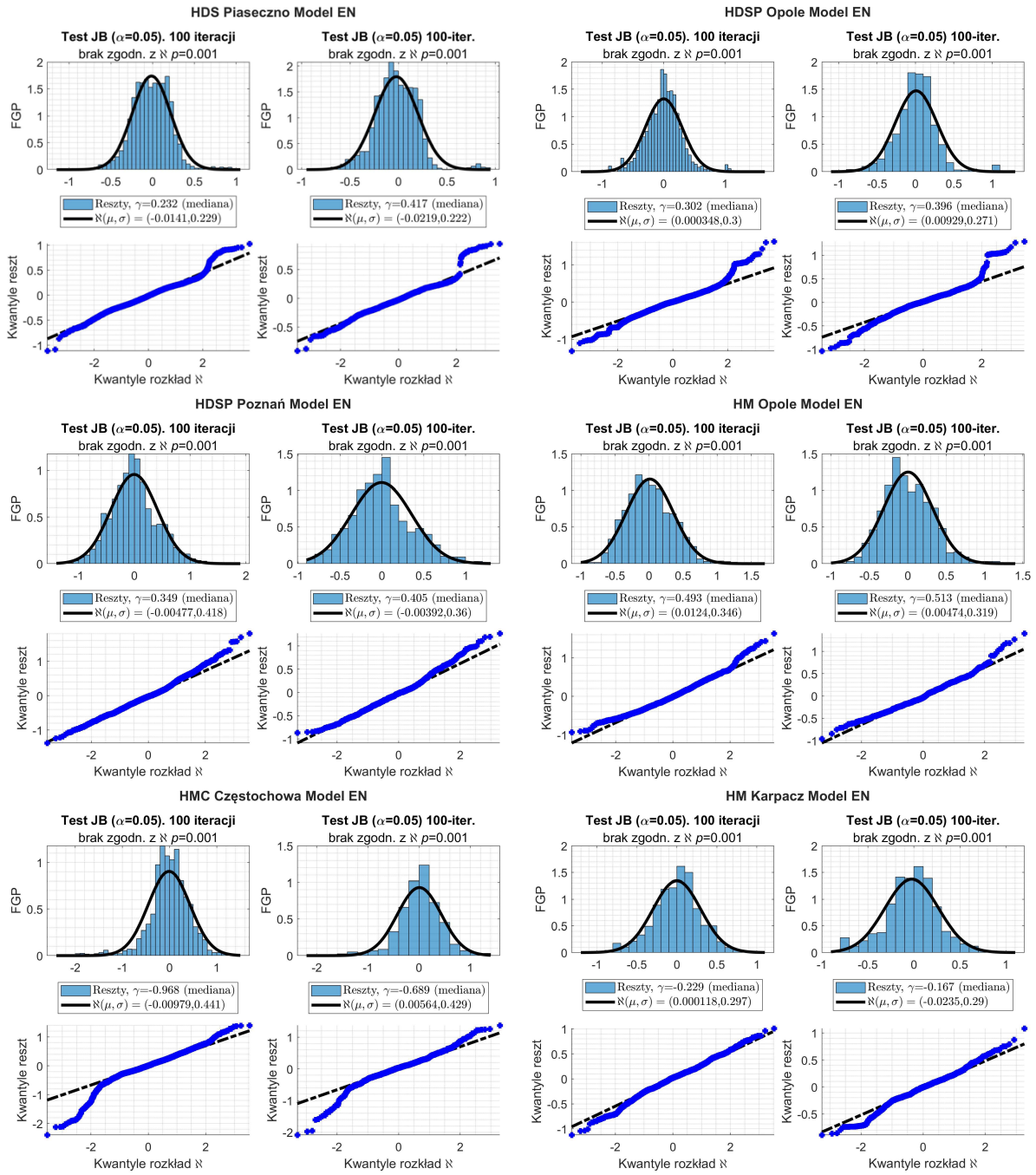
Rys. Z7.2 Zestawienie porównawcze rozkładów gęstości prawdopodobieństwa reszt wraz z wykresami Q–Q, przeprowadzone w 100 iteracjach odpowiednio uczenia i testowania modelu predykującego GOP. Analizy dotyczyły hoteli HM Karpacz, HDC Częstochowa, HM Opole, HDSI Poznań, HDSI Opole i HDS Piaseczno z wykorzystaniem algorytmu NN. Źródło: Opracowanie własne.



Rys. Z7.3 Zestawienie porównawcze rozkładów gęstości prawdopodobieństwa reszt wraz z wykresami Q–Q, przeprowadzone w 100 iteracjach odpowiednio uczenia i testowania modelu predykującego GOP. Analizy dotyczyły hoteli HM Karpacz, HDC Częstochowa, HM Opole, HDSI Poznań, HDSI Opole i HDS Piaseczno z wykorzystaniem algorytmu SVM. Źródło: Opracowanie własne.



Rys. Z7.4 Zestawienie porównawcze rozkładów gęstości prawdopodobieństwa reszt wraz z wykresami Q–Q, przeprowadzone w 100 iteracjach odpowiednio uczenia i testowania modelu predykującego GOP. Analizy dotyczyły hoteli HM Karpacz, HDC Częstochowa, HM Opole, HDSI Poznań, HDSI Opole i HDS Piaseczno z wykorzystaniem algorytmu TREE. Źródło: Opracowanie własne.



Rys. Z7.5 Zestawienie porównawcze rozkładów gęstości prawdopodobieństwa reszt wraz z wykresami Q–Q, przeprowadzone w 100 iteracjach odpowiednio uczenia i testowania modelu predykującego GOP. Analizy dotyczyły hoteli HM Karpacz, HDC Częstochowa, HM Opole, HDSI Poznań, HDSI Opole i HDS Piaseczno z wykorzystaniem algorytmu EN. Źródło: Opracowanie własne.