

# Możliwości predykcji zużycia ciepła w lokalu mieszkalnym za pomocą modeli autoregresji

Possibilities of predicting heat consumption in a dwelling by autoregression models

GRZEGORZ BARTNICKI, BOGDAN NOWAK

DOI 10.36119/15.2021.5.1

W artykule opisano możliwości wykorzystania modeli ARIMA i XGBoost do predykcji zużycia ciepła w mieszkaniach budynku wielorodzinnego. W oparciu o dane pomiarowe zużycia ciepła w mieszkaniach dwóch zespołów budynków z okresu 2016-2020 opracowano modele ARIMA i XGBoost do predykcji zużycia ciepła w okresach miesięcznych, a do obliczeń wykorzystano środowisko R. Wyniki zaprezentowano w artykule dla wybranych mieszkań w postaci tabelarycznej i rysunków. Stwierdzono, że modele ARIMA wykazują dużą dokładność, nie są jednak skuteczne w przypadku gwałtownych zmian pojedynczych obserwacji. Do zastosowań opisanych w artykule wymagają też jeszcze dalszych badań. XGBoost jest algorytmem znacznie bardziej zaawansowanym, a w konsekwencji istnieje znacznie więcej parametrów modelu, które należy ustawić i później zoptymalizować. Ten aspekt będzie przedmiotem dalszych prac badawczych, gdyż pomimo oczekiwania takiego rezultatu, wykorzystanie tego algorytmu nie dało dużo lepszych przybliżeń wyniku do wartości rzeczywistych niż modele ARIMA.

*Słowa kluczowe: sztuczna inteligencja, uczenie maszynowe, modelowanie predykcyjne*

The article describes the possibilities of using the ARIMA and XGBoost models to predict heat consumption in dwellings in a multi-family building. Based on the measurement data from the period 2016-2020 of heat consumption in dwellings in two building complexes, ARIMA and XGBoost models were developed to predict heat consumption in monthly periods, and the R environment was used for the calculations. The results are presented in the article for selected apartments in the form of tables and figures. ARIMA models were found to be good, but not effective for rapid changes in single observations. The applications described in the article also require further research. XGBoost is a much more advanced algorithm, and consequently there are many more model parameters that need to be set and optimized later. Therefore, this aspect will be the subject of further research, because despite the expectation of good results, the use of this algorithm did not give much better prediction for rapid changes than the ARIMA models.

*Keywords: artificial intelligence, machine learning, predictive modeling*

## Wstęp

W warunkach klimatu Polski budynki wymagają wyposażenia w instalację grzewczą, zapewniającą odpowiednie warunki temperaturowe w okresie obniżonej temperatury powietrza zewnętrznego. Systemy zaopatrzenia w ciepło wykorzystują w tym celu zasoby energii pierwotnej, co z jednej strony stwarza określoną presję na środowisko, wyczerpując jego zasoby, z drugiej generuje koszty dla użytkowników tego systemu. Nadchodzące lata zapowiadają kolejne zmiany prowadzące do zmniejszania śladu węglowego, znaczącej redukcji zużycia energii, rezygnacji ze spalania węgla, wdrażania nowych technologii, wykorzystania na

szeroką skalę odnawialnych źródeł energii oraz przede wszystkim zmiany myślenia o środowisku i człowieku jako jednym z jego elementów i współużytkowników, a nie jedynym eksploatatorem bogactw naturalnych i środowiska [2, 3, 10, 12, 16]. Temu celowi służyć ma też poprawa efektywności energetycznej. W społeczeństwie, w którym wycofano z obiegu dopiero w 1996 r. (czyli tylko 25 lat temu) bilety Narodowego Banku Polskiego wg wzoru z 1948 r. o nominale 100 zł (wprowadzone do obiegu po wymianie złotego w 1950 r.), na których na rewersie kilkanaście dymiących kominów miało świadczyć o kierunku dynamicznego rozwoju polskiej gospodarki opartej na spalaniu węgla (mierzonego w tamtych czasach

m.in. ilością zużytej energii i rosnącym wydobyciem węgla przez polskie kopalnie), obecną sytuację i wyzwania stawiane na najbliższe lata należy nazwać rewolucją energetyczną, nie tylko dla przedsiębiorstw wykorzystujących lub produkujących energię czy jej nośniki, ale również dla społeczeństwa, które nie zawsze jeszcze zauważa problem konieczności szybkiej transformacji energetycznej. Nie w pełni też został wykorzystany czas na zmiany ewolucyjne, przy których łatwiej zapanować nad siłami inercji. Z punktu widzenia mieszkańców budynków jako końcowych użytkowników energii, problem ten w ostatnim roku został przysłonięty przez epidemię wywołaną przez wirus SARS-Cov-2, a częściowo przez

dr inż. Grzegorz Bartnicki – Katedra Klimatyzacji, Ogrzewnictwa, Gazownictwa i Ochrony Powietrza, Politechnika Wrocławska, <https://orcid.org/0000-0002-4482-6950>; e-mail: [grzegorz.bartnicki@pwr.edu.pl](mailto:grzegorz.bartnicki@pwr.edu.pl)

dr inż. Bogdan Nowak – Katedra Klimatyzacji, Ogrzewnictwa, Gazownictwa i Ochrony Powietrza, Politechnika Wrocławska, <https://orcid.org/0000-0002-9764-5555>; e-mail: [bogdan.nowak@pwr.edu.pl](mailto:bogdan.nowak@pwr.edu.pl)

działania osłonowe mające na celu ograniczenie wzrostu cen energii elektrycznej, jak również przez spadek popytu na energię i paliwa wywołany epidemią. Problem będzie jednak coraz bardziej zauważalny we wszystkich dziedzinach gospodarczych i społecznych, ze względu na powszechność korzystania z energii i jej nośników. Początkowy brak dużego tempa zmian ma również dobre strony, gdyż można w ten sposób wykorzystać doświadczenia wdrożone w innych krajach czy regionach i ustrzec się ich błędów oraz trudności towarzyszących fazie badań i pilotażu.

Zaistniałe w tych okolicznościach potrzeby w świecie techniki i technologii prowadzą do różnego typu działań korzystających zarówno z dotychczasowych zasobów wiedzy, doświadczeń, dobrych praktyk [1, 13, 16, 18, 22], jak też przede wszystkim skłaniają do rozwoju wiedzy, wdrażania innowacji, tworzenia nowych narzędzi analitycznych czy służących poznawaniu i przewidywaniu rzeczywistości. Zauważając ten problem do systemów zaopatrzenia w ciepło budownictwa mieszkaniowego (które ma znaczący udział w krajowym zużyciu energii i emisji CO<sub>2</sub>) coraz większe znaczenie będą mieć narzędzia prognostyczne, wykorzystujące różnego typu algorytmy analizy danych historycznych i bieżących, umożliwiające prognozowanie zapotrzebowania na ciepło w różnej perspektywie czasowej [5, 7, 9, 19, 21]. W perspektywie transformacji systemów zaopatrzenia w ciepło (systemów ciepłowniczych) 2-giej i 3-ciej generacji (3GDH) w rozwiązania zero lub niskoemisyjne 5-tej a przynajmniej 4-tej generacji (4GDH, 5GDHC) [3, 13] niezbędne będą narzędzia prognostyczne wspomagające np. akumulację ciepła odpadowego, które będzie mogło być w ten sposób efektywnie wykorzystane czy transfer ciepła pomiędzy obiektami należącymi do tego samego systemu. Również tego typu narzędzia są przydatne do diagnozowania stanów awaryjnych, szukania możliwości oszczędzania energii czy zarządzania popytem. Z wyników badań [4] wynika, że w badanym przez tego Autora osiedlu zużycie ciepła przez mieszkania jest dwukrotnie większe niż wynikałoby to z teoretycznych obliczeń. Predykcja zużycia ciepła w poszczególnych lokalach czy całym budynku może zarówno służyć do doskonalenia układów sterowania pracą źródła ciepła, sygnalizować nieprawidłowości w systemie lub wskazywać tych użytkowników i te instalacje, które są odmienne, niekoniecznie z powodu uwarunkowań ustalanych przez ich

użytkownika, ale również może w efekcie niepoprawnego ich działania, prowadzącego do obniżenia efektywności energetycznej. Autor [25] podkreślił, że prognozy o charakterze naukowym muszą opierać się na dobrze sformułowanych teoriach i poddawać się falsyfikacji poprzez empiryczną weryfikację.

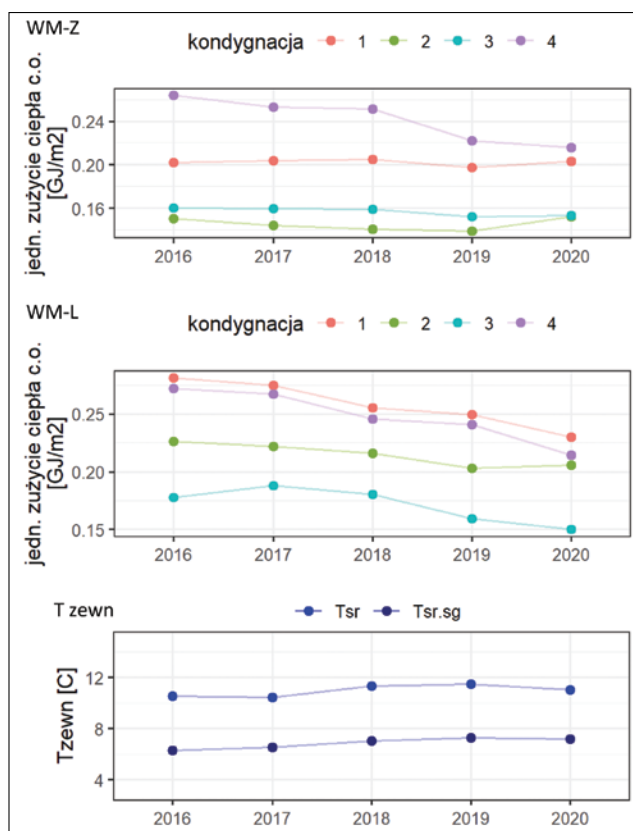
### Składnik stochastyczny w rzeczywistym zużyciu ciepła

Modele wykorzystywane do obliczania (chwilowego) zużycia ciepła wiążą tę wielkość z konstrukcją przegród budowlanych, ich powierzchnią oraz warunkami wewnętrznymi reprezentowanymi przez temperaturę powietrza wewnętrznego (zakładając brak ich zindywidualizowania) oraz warunkami pogodowymi zredukowanymi do temperatury zewnętrznej. Norma [23], która opisuje metody obliczania strat ciepła oraz sposób obliczenia zapotrzebowania na moc cieplną dla typowych przypadków w warunkach projektowych nie uwzględ-

Obecnie obserwuje się nie tylko postawy zmierzające do ograniczenia zużycia ciepła (a w konsekwencji kosztów), ale również oczekiwania dokładnie odwrotne – możliwość zapewnienia warunków podwyższonego komfortu cieplnego, co najczęściej przyjmuje formę utrzymania wyższej niż projektowa wartość temperatury powietrza wewnętrznego czy oczekiwania wydłużenie sezonu grzewczego [6] na cały rok, lub przynajmniej zapewnienie długiego okresu przejściowego, umożliwiającego korzystanie z ogrzewania w okresach, w których ono jest już wyłączane. Decydowanie przez użytkowników ogrzewania o warunkach jego działania (nie zawsze nawet świadome) prowadzi do znacznego zindywidualizowania zużycia ciepła zarówno w zakresie wielkości, jak i profilu (dobowego, miesięcznego i rocznego) zużycia ciepła na potrzeby ogrzewania pomieszczeń oraz ograniczenia powiązania tych wielkości z warunkami zewnętrznymi [1, 4, 5, 6, 17]. We współczesnych budynkach mieszkalnych,

**Rys. 1**  
Średnie jednostkowe zużycie ciepła [GJ/m<sup>2</sup>] w latach 2016-2020 (zasoby WM – L i WM – Z) dla lokali mieszkalnych zlokalizowanych na poszczególnych kondygnacjach oraz średnia temperatura zewnętrzna (1) dla całego roku; (2) tylko dla okresu sezonu grzewczego (styczeń-maj, październik-grudzień)

**Fig. 1.** Average unit heat consumption [GJ / m<sup>2</sup>] in 2016-2020 (for the building complexes WM – L and WM – Z) for dwellings located on different floors and the average outside temperature (1) for the whole year; (2) for the heating season (January-May, October-December)



nia zysków ciepła, jak było to w sposób pośredni i mocno uproszczony przyjęte np. w poprzedzającej ją normie [24], czy też innych czynników wpływających na działanie systemu zaopatrzenia w ciepło. Rzeczywiste zużycie ciepła jest procesem znacznie bardziej skomplikowanym, na który wpływ ma szereg czynników, w tym indywidualne preferencje i oczekiwania.

które spełniają wymagania aktualnych warunków technicznych lub w budynkach starszych, które zostały poddane termorenowacji, wpływ na zużycie ciepła zarówno lokalu, jak też całego budynku, poza temperaturą powietrza zewnętrznego, mają także inne parametry, takie jak zyski ciepła od nasłonecznienia, urządzeń AGD, osób korzystających z lokalu, zwierząt

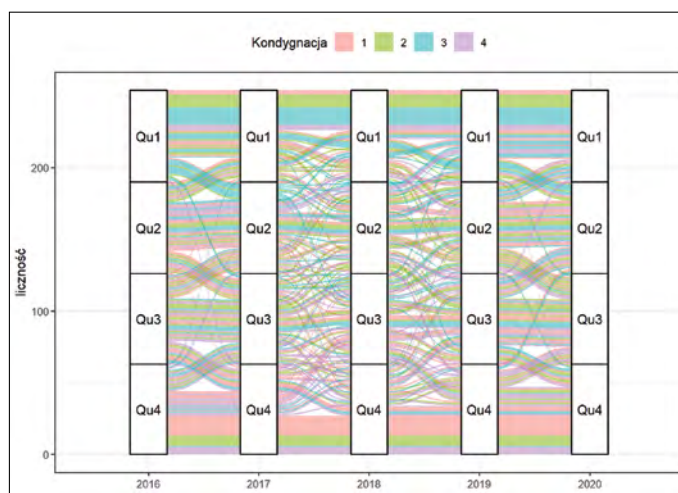
itp., sposób użytkowania lokalu, wyposażenie w dodatkowe instalacje i urządzenia (chłodzenie, oczyszczanie powietrza) i inne. Te czynniki mają również charakter losowy, tak jak warunki atmosferyczne występujące w danym okresie roku. Od marca 2020 r. na bilans energii w budynkach mieszkalnych wpływ ma również stan epidemii COVID-19. Lockdown, nauka i praca zdalna spowodowały zwiększenie liczby osób przebywających w budynku, szczególnie w okresie przedpołudniowym. Jest to widoczne w zużyciu ciepła w postaci ciepłej wody, ale również na cele ogrzewania. Należy jednak też mieć na uwadze, że część społeczeństwa (pracownicy handlu, służby zdrowia, służby mundurowe itp.) pracuje wg niezmiennych prawie zasad.

Na rys. 1 przedstawiono średnie jednostkowe zużycie ciepła wyznaczone w oparciu o pomiary przeprowadzone w dwóch grupach budynków mieszkalnych (WM-L i WM-Z) zlokalizowanych we Wrocławiu, w zależności od kondygnacji, na której znajduje się mieszkanie. Pod pierwszą kondygnacją mieszkalną (1) znajduje się nieogrzewany garaż podziemny, nad ostatnią – docieplony stropodach. Pomiary przeprowadzone zostały na przestrzeni 5 lat w mieszkaniach zlokalizowanych w zasobach wspólnot mieszkaniowych znajdujących się w bezpośrednim sąsiedztwie. Dodatkowo, na rys. 1 zamieszczono średnie wartości temperatury zewnętrznej, które zostały określone w oparciu o dane IMGW-PIB pochodzące ze stacji meteorologicznej zlokalizowanej najbliżej osiedla, którego dotyczą badania. Pomimo zauważalnego wzrostu średniej temperatury zewnętrznej w 2018 r. w porównaniu z 2017 r., w zasobach wspólnoty WM-Z spadek jednostkowego zużycia ciepła obserwuje się wyłącznie dla kondygnacji 2. W zasobach WM-L widoczny jest spadek dla mieszkań na wszystkich kondygnacjach, choć w różnej skali. Warto również zauważyć, że w 2020 r. w zasobach WM-L dla kondygnacji 2. nastąpił wzrost średniego jednostkowego zużycia ciepła, podczas gdy dla pozostałych zauważalny jest spadek. Jak zaproponowano to w [5], w budynku mieszkalnym zasadne jest i możliwe wyodrębnienie grup mieszkań jednorodnych, dla których obserwowane w czasie zmiany zużycia ciepła mają podobny charakter. W mieszkaniach ujętych w grupie mieszkań jednorodnych występowało inne zużycie ciepła wyrażone w jednostkach bezwzględnych, jednak przebieg zmienności (profil zużycia) był zbliżony. Podział na takie grupy uwzględnia kondygnację, na której znajduje się mieszkanie,

wielkość mieszkania, liczbę pokoi, preferencje mieszkańców w zakresie utrzymywanej temperatury wewnętrznej, a nawet warunki termiczne utrzymywane w sąsiadujących z nim mieszkaniach. Zapewnia równocześnie zgromadzenie większej liczby danych do analizy trendu zmian, utwierdzając w ten sposób predykcję zużycia ciepła pojedynczego mieszkania (np. w przypadku stwierdzenia awarii ciepłomierza) [5].

W oparciu o zgromadzone dane przeprowadzona została analiza trendu ("skłonności") użytkowników poszczególnych mieszkań do ograniczenia lub nadmiernego zużycia ciepła na cele ogrzewania i stabilności tej cechy w kolejnych latach. Wykorzystano w tym celu analizę kwantyli sumarycznego zużycia ciepła w sezonie. Najpierw uporządkowano zbiór mieszkań w kolejności od najmniejszego zużycia ciepła do największego i zbadano następnie przynależność każdego lokalu do poszczególnych kwantyli w kolejnych latach (Qu1: 0-25%, Qu2: 25-50%, Qu3: 50-75%, Qu4: 75-100%). W ten sposób, wykorzystując dane pomiarowe z lat 2016-2020 dla obu zespołów budynków mieszkalnych (ZM-L i ZM-Z) przeprowadzone zostało badanie stopnia zindywidualizowania zużycia ciepła na cele ogrzewania w poszczególnych mieszkaniach, odseparowując go od warunków pogodowych. Wyniki tej analizy graficznie przedstawiono na rys. 2. Zauważyć można, że lokale „bardzo oszczędne” (Qu1 2016) trafiły w kolejnym roku do grupy „oszczędne” (Qu2 2017), a nawet do grupy „podwyższone zużycie” (Qu3 2017). W kolejnych latach taka zmienność jest również obserwowana. Istnieje jednak przede wszystkim spora grupa lokali, które przez cały okres badań pozostawały w zbiorze „oszczędnych” (Qu1) i takich, które pozostawały w zbiorze „duże zużycie” (Qu4).

**Rys. 2**  
Migracja mieszkań pomiędzy wyodrębnionymi kwantylami Q1-Q4 w latach 2016-2020 (zasoby WM - L) z zaznaczeniem kondygnacji, na której zlokalizowane jest mieszkanie  
*Fig. 2 Migration of dwellings between the separated quantiles Q1-Q4 in 2016-2020 (building complex WM - L) with the floor on which the dwelling is located*



Zarówno wyniki analizy przedstawione na rys. 1, jak i na rys. 2 potwierdzają, że w badanych zespołach budynków mieszkalnych zużycie ciepła na potrzeby ogrzewania pomieszczeń w znacznym stopniu zależy od indywidualnych oczekiwań użytkowników, a powiązanie tej wielkości z warunkami pogodowymi (temperaturą zewnętrzną) – choć wciąż bardzo istotne – podlegać może pewnemu ograniczeniu (tym bardziej, im ta temperatura jest wyższa [6]). Można też w tych budynkach wyodrębnić grupy mieszkań, w których zużycie ciepła ma podobny charakter. Dopuszczenie takiej możliwości pozwala na wykorzystanie większej liczby danych dla predykcji zużycia ciepła pojedynczego mieszkania niż gdyby tylko korzystać z jego danych [5].

Odrębnym zagadnieniem jest model, który może być wykorzystany do predykcji zużycia ciepła na ogrzewanie w mieszkaniu budynku wielorodzinnego. Rozwój technik komputerowych i narzędzi matematycznych udostępnił różne pakiety obliczeniowe wykorzystujące bardziej lub mniej złożone algorytmy. Wybór modelu i jego parametrów zależy też od charakteru opisywanego zjawiska, jak i rodzaju danych wykorzystywanych do obliczeń. Aby porównać te możliwości, wykorzystano do obliczeń środowisko R oraz modele ARIMA i XGBoost.

## ARIMA – dane historyczne jako szereg czasowy

Modele ARIMA są autoregresyjnymi zintegrowanymi modelami średniej ruchomej do prognozowania szeregów czasowych. Składają się z procesu autoregresyjnego (AR), gdzie każda wartość jest liniową kombinacją poprzednich wartości), procesu średniej ruchomej (MA) i stopnia integracji (I). Przy predykcji dane są traktowane jako szereg czasowy. Głównym źródłem informacji są dane historyczne uporządkowane

w czasie. Modele ARIMA umożliwiają wprowadzenie dodatkowych regresorów, np. temperatury zewnętrznej czy wielkości zużycia z mieszkań podobnych.

Modele tego typu znajdują z powodzeniem zastosowanie w prognozowaniu również innych mediów, jak choćby zużycie c.w. [11] czy pobór paliwa gazowego [7].

W ramach badań opracowane zostały trzy modele ARIMA. W pierwszym (oznaczonym jako Predict.1) wprowadzono dodatkowe regresory w postaci zużycia ciepła w pięciu mieszkaniach podobnych. O stopniu podobieństwa zdecydowała wartość współczynnika korelacji wyznaczona dla zdekomponowanych szeregów czasowych. W modelu tym również została zastosowana transformacja Boxa-Coxa, przy czym wybór wartości  $\lambda$  opierał się na wizualnej ocenie dopasowania modelu do zbioru uczącego. W przyszłych badaniach planowane jest rozwinięcie zagadnienia wyboru najkorzystniejszej wartości  $\lambda$ .

W drugim modelu (oznaczonym jako Predict.2) pominięto regresory zewnętrzne, zaś transformacja Boxa-Coxa przeprowadzona była z tymi samymi parametrami jak w modelu pierwszym (Predict.1).

Ostatnie rozwiązanie (oznaczone jako Predict.3) to najprostszy model ARIMA, przy czym zrezygnowano zarówno z transformacji danych wejściowych, a także z regresorów zewnętrznych.

W wszystkich trzech modelach szereg czasowy został zdefiniowany z sezonowością roczną, a wybór parametrów modelu ARIMA ( $p$ ,  $q$ ,  $d$ ) został przeprowadzony w oparciu o algorytm automatycznego wyznaczania tych parametrów. Wszystkie analizy zostały wykonane w środowisku R, modele obrazowano stosując pakiet *features*.

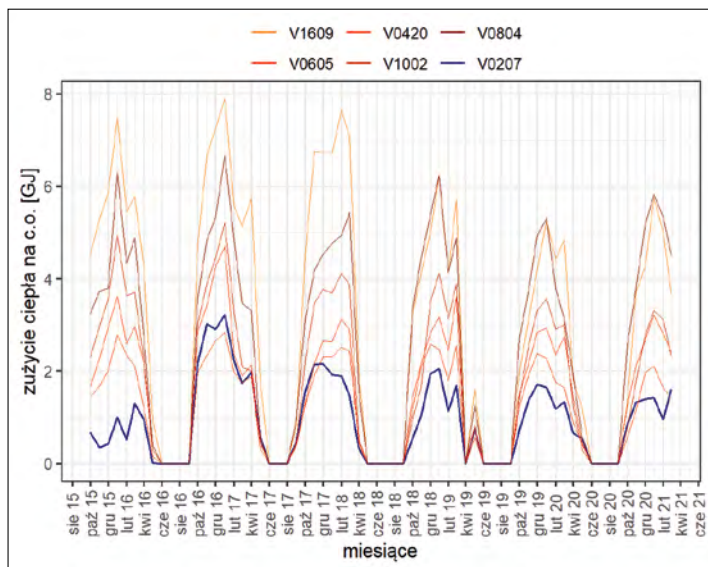
**Tab. 1** Rzeczywiste i prognozowane zużycie ciepła w trzech okresach rozliczeniowych (miesiącach) uzyskane w oparciu o różne modele ARIMA

**Tab. 1.** Real heat consumption and forecast in three periods (months) obtained on the basis of various ARIMA models

Ozn	Okres	Rzeczywiste zużycie ciepła [GJ]	Wartości przewidywane w oparciu o Predict.1 [GJ]	Wartości przewidywane w oparciu o Predict.2 [GJ]	Wartości przewidywane w oparciu o Predict.3 [GJ]	Dokładność najlepszej predykcji
V0624	01.2021	1,3	1,35	1,17	<b>1,32</b>	2%
	02.2021	0,91	1,26	0,76	<b>0,85</b>	-7%
	03.2021	0,82	<b>0,65</b>	0,54	0,6	-21%
V0803	01.2021	3,29	3,39	2,94	<b>3,35</b>	2%
	02.2021	3,15	<b>2,8</b>	2,5	2,63	-11%
	03.2021	2,91	2,31	2,62	<b>3,03</b>	4%
V1621	01.2021	3,82	3,25	3,6	3,62	-5%
	02.2021	1,67	2,77	2,35	<b>2,35</b>	41%
	03.2021	0,4	2,02	<b>1,96</b>	1,96	390%
V0207	01.2021	1,42	1,84	<b>1,6</b>	1,88	12%
	02.2021	0,94	1,54	<b>1,06</b>	1,17	13%
	03.2021	1,6	1,2	1,22	<b>1,54</b>	-4%

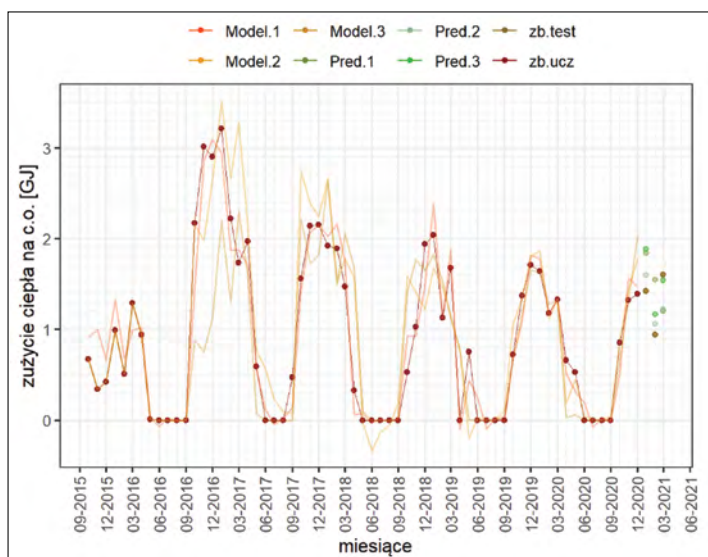
**Rys. 3** Porównanie wybranego mieszkania (V0207) z mieszkaniem należącym do tej samej kategorii profilu zużycia ciepła na potrzeby c.o.

**Fig. 3** Comparison of the selected apartment (V0207) with apartments belonging to the same category of heat consumption profile for central heating



**Rys. 4** Trzy warianty modeli ARIMA dla wybranego mieszkania (V0207) wraz z wynikami predykcji

**Fig. 4** Three variants of ARIMA models for a selected apartment (V0207) with prediction results



Dodatkowo zestawiono jeszcze summaryczne zużycie ciepła w okresie trzech miesięcy (dla których opracowane były

prognozy) z wynikami uzyskanymi z modeli predykcyjnych. I tak:

- V0624 – najlepsze dopasowanie modelem Predict.1 – dokładność 7%,
- V0803 – najlepsze dopasowanie modelem Predict.3 – dokładność – 4%,
- V1621 – najlepsze dopasowanie modelem Predict.2 – dokładność 34%,
- V0207 – najlepsze dopasowanie modelem Predict.2 – dokładność – 2%.

Dokładność modeli dla okresu 2-3 miesięcy można uznać za dobrą, a nawet bardzo dobrą. Od modeli tego typu trudno zresztą oczekiwać, aby uzyskano prognozę równą wartości rzeczywistej. Dla pojedynczego mieszkania jest to trudne, gdyż czynniki losowe mogą powodować znaczące zmiany warunków użytkowania mieszkania czy zachowań lokatorów. O ile też postawiony cel badawczy ma swoje uzasadnienie, trudno uzasadnić potrzebę tak trafnych prognoz dla pojedynczego mieszkania, nawet dla zastosowań opisanych w [5]. Systemem zaopatrzenia w ciepło wielorodzinnego

budynku mieszkalnego rozpatrywany zawsze jest jako całość, a w takiej skali osiągnięta dokładność wydaje się wystarczająca.

Modele ARIMA wykazują dużą dokładność, nie są jednak skuteczne w przypadku gwałtownych zmian pojedynczych obserwacji, co uwidoczniło się na przykładzie mieszkania V1621. W dalszych badaniach należy rozważyć celowość ustalania prognozowanego zużycia ciepła w oparciu o średnią lub średnią ważoną z wyników uzyskanych z poszczególnych modeli. Wartościowe będzie również zaproponowanie procedury wyboru stopnia skomplikowania modelu ARIMA. Wstępne analizy wskazują, że istotnym czynnikiem wpływającym na wybór może być wartość współczynnika korelacji dla mieszkań podobnych. Na tym etapie badań i w oparciu o uzyskane wyniki, można wstępnie już stwierdzić, że zwiększanie liczby mieszkań podobnych w macierzy regresorów zewnętrznych powyżej 5 nie przynosi istotnej poprawy prognoz.

### XGBoost – zaawansowane narzędzie

XGBoost (Extreme Gradient Boosting), który jest pochodną klasycznych drzew decyzyjnych i lasów losowych, jest dosyć nowym algorytmem (2014/2016 r.) wykorzystywanym do predykcji [8, 25]. Wprowadzony został bardzo szybko do programów i środowisk obliczeniowych, w tym do środowiska R, ze względu na jakość i szybkość uzyskiwanych w ten sposób prognoz. Algorytm został stworzony jako klasyfikator i narzędzie regresji, nie zaś jako system prognozowania szeregów czasowych. Badania i ich zastosowania wykazały jednak, że dzięki odpowiedniej obróbce wstępnej danych i ich przygotowaniu, możliwe jest użycie tego algorytmu również do predykcji [20, 25]. W badanym przypadku dotyczącym zużycia ciepła na potrzeby c.o. dane historyczne zostały potraktowane jako zestaw danych nie indeksowanych w czasie, a okres rozliczeniowy (miesiąc) jest traktowany jako jedna z wielu zmiennych objaśniających. Takie ujęcie powoduje, że procedura jest bardziej odporna na brakujące pojedyncze dane oraz umożliwia łatwiejsze prognozowanie wartości dla pojedynczych okresów, które nie sąsiadują ze sobą (brakujące pomiary dla określonego miesiąca (np. w efekcie awarii ciepłomierza w mieszkaniu [5])).

XGBoost jest algorytmem znacznie bardziej zaawansowanym, a w konsekwencji istnieje znacznie więcej parametrów modelu, które należy ustawić, a dalej

– zoptymalizować. Ten aspekt będzie przedmiotem dalszych prac badawczych. Prezentowane na rys. 5 wyniki zostały uzyskane po zastosowaniu strojenia modelu. Zdefiniowane zostały zakresy dla 7 wybranych parametrów modelu, a następnie na tak zakreślonej przestrzeni losowo wybierane było  $n$  zestawów parametrów. Tych parametrów nie może być zbyt mało, bo wówczas wylosowane mogą zostać zestawy z jednego "narożnika" zdefiniowanej dla modelu przestrzeni. Większa liczba iteracji daje natomiast szansę na rozłożenie wylosowanych zestawów parametrów po całej przestrzeni. Zwiększanie liczby iteracji znacząco wpływa jednak na czas prowadzonych obliczeń. Do badań przyjęto 3 modele, dla różnych iteracji:

- xgb.1 – 60 iteracji,
- xgb.2 – 560 iteracji,
- xgb.3 – 860 iteracji.

### Dyskusja wyników

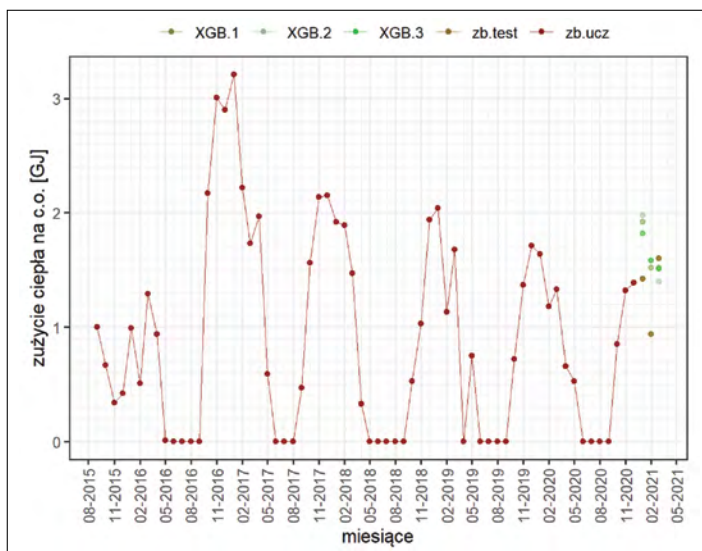
Przedstawione w tab. 1 i 2 wyniki wskazują, że oszacowanie zużycia ciepła w mieszkaniu budynku wielorodzinnego za pomocą predykcji modelami ARIMA i XGBoost może zarówno przekraczać rzeczywiście zmierzone wartości, jak i to, że takie oszacowanie może być mniejsze. W zależności od przyjętego algorytmu modelu, jak i jego parametrów uzyskano różną dokładność modelu. Model ARIMA okazuje się mało skuteczny w przypadku gwałtownych zmian pojedynczych obserwacji. Wykorzystanie algorytmu XGBoost nie dało również oczekiwanej poprawy prognozy, choć wartości uzyskane z predykcji są nieznacznie bardziej zbliżone do wartości rzeczywistych wykorzystanych do oceny jej trafności. Należy pamiętać, że w tego typu modelowaniu nie można

**Tab. 2** Rzeczywiste i prognozowane zużycie ciepła w trzech okresach rozliczeniowych (miesiącach) uzyskane w oparciu o różne modele XGBoost

**Tab. 2.** Real heat consumption and forecast in three periods (months) obtained on the basis of various XGBoost models

Ozn.	Okres	Rzeczywiste zużycie ciepła [GJ]	Wartości przewidywane w oparciu o XGB.1 [GJ]	Wartości przewidywane w oparciu o XGB.2 [GJ]	Wartości przewidywane w oparciu o XGB.3 [GJ]	Dokładność najlepszej predykcji
V0624	01.2021	1,3	1,24	1,23	<b>1,27</b>	-2%
	02.2021	0,91	<b>1,04</b>	1,05	1,08	14%
	03.2021	0,82	<b>0,85</b>	0,86	0,88	4%
V0803	01.2021	3,29	<b>3,05</b>	2,95	2,97	-7%
	02.2021	3,15	2,63	<b>2,92</b>	2,55	-7%
	03.2021	2,91	<b>2,2</b>	1,78	2,14	-24%
V1621	01.2021	3,82	3,09	<b>3,11</b>	3,08	-19%
	02.2021	1,67	<b>2,57</b>	2,6	2,59	54%
	03.2021	0,4	1,86	1,86	<b>1,85</b>	363%
V0207	01.2021	1,42	1,92	1,98	<b>1,82</b>	28%
	02.2021	0,94	<b>1,52</b>	1,58	1,58	62%
	03.2021	1,6	<b>1,52</b>	1,4	1,51	-5%

**Rys. 5** Model XGBoost dla wybranego mieszkania (V0207) wraz z wynikami predykcji  
**Fig. 5** XGBoost model for a selected apartment (V0207) with prediction results



mówić o najlepszym modelu opisującym analizowany szereg czasowy w ogóle, ale co najwyżej o modelu najlepszym z przyjętych do analizy i oceny. Model będzie zawsze tylko pewnym przybliżeniem opisywanego zjawiska i przede wszystkim należy go oceniać z punktu widzenia oczekiwanej dokładności uzyskanej predykcji. Z punktu widzenia celu, jakim jest prognozowanie zużycia ciepła w poszczególnych mieszkaniach budynku wielorodzinnego, taka dokładność nie musi wcale osiągać wartości poniżej tzw. dopuszczalnego błędu obliczeń inżynierskich (w czasach braku współcześnie stosowanych komputerów i narzędzi obliczeniowych przyjmowanego jako 10%), chociaż wartość 390% czy 363% musi zastanawiać. Przypadek ten dotyczy jednak bardzo małego rzeczywistego zużycia ciepła w mieszkaniu, w którym wcześniej tak małych wskazań raczej nie obserwowano, i dla miesiąca, który charakteryzuje się dość zmiennymi warunkami temperaturowymi. Uzyskany z prognozy większy wynik mieści się w wartościach obserwowanych czy prognozowanych dla innych lokali dla tego okresu sezonu grzewczego. Należy mieć na uwadze, że tego typu prognozy są i będą przede wszystkim przydatne z punktu widzenia analizy całego systemu tj. całego budynku. Ważne jest więc, aby dla wszystkich mieszkań nie utrzymywał się ten sam trend błędu np. jako duże zawyżenie lub jako duże niedoszacowanie dla wszystkich mieszkań równocześnie.

Opracowane modele i uzyskana ich jakość pokazują, że zasadne jest kontynuowanie tego typu badań oraz szukanie określonego czynnika, zmiennej lub parametru modelu, który może być predyktorem takiego gwałtownego wahnięcia prognozy (i znaczącej rozbieżności pomiędzy predykcją i wartością rzeczywistą, rzędu kilkuset procent). Może to być błąd modelu, ale również błąd danych pomiarowych. Najważniejszym wnioskiem jest jednak to, że już najprostsze modele wykazują dużą przydatność (np. przy rozliczaniu zużycia ciepła [5] i uzupełnianiu brakujących danych w przypadku awarii układów pomiarowych, analizie zjawisk zachodzących w systemie zaopatrzenia w ciepło, diagnostyce poprawności działania systemu itp.), nie wymagają zbyt wielu danych wejściowych, a dane historyczne niemal zawsze są dostępne. Modele tego typu nie angażują czasowo, nie wymagają też dużej mocy obliczeniowej. Dodatkowo, są dość proste w implementacji w różnych środowiskach obliczeniowych /aplikacjach (Matlab, R, Python itp. ).

Porównując modele ARIMA i XGBoost, te drugie mają jeszcze tą dodatkową zaletę, że zawierają pakiety i procedury, które pozwalają ocenić i zwizualizować wpływ poszczególnych zmiennych na wynik końcowy (na wynik prognozy). Jest to ważne zagadnienie w uczeniu maszynowym, które co do zasady uniemożliwia lub utrudnia interpretację wyników (inaczej niż przykładowo w regresji liniowej). Poza prawie oczywistą zależnością zużycia ciepła od temperatury zewnętrznej, w ten sposób można zidentyfikować i wykazać inne parametry wpływające na zużycie ciepła w mieszkaniu budynku wielolokalowego (kondygnacja budynku, "skłonności" użytkowników poszczególnych mieszkań do ograniczenia lub nadmiernego zużycia ciepła i przynależność lokalu do określonego kwantyla Q1-Q4 itp.). Taka funkcjonalność modelu może być np. przydatna na etapie kodyfikowania (np. zapisania określonego algorytmu w regulaminach rozliczania kosztów, tak aby np. skłaniały też one użytkowników do racjonalnego użytkowania energii i jej oszczędzania).

Istotne przy budowie modelu jest przyjęcie jego parametrów, które w przypadku modelu XGBoost są określane zmiennymi objaśniającymi, a w modelu ARIMA – regresorami. W obu przypadkach opisanych w artykule modeli zostały uwzględnione jako taki parametr tylko mieszkania podobne (5 lokali mieszkalnych). W przypadku XGBoost (ponieważ algorytm ten nie zachowuje indeksowania w czasie) uwzględniono dodatkowo również miesiąc, którego dotyczy zużycie ciepła. Nie uwzględniono natomiast czynnika jakim jest temperatura zewnętrzna. Wynikało to z założenia, że jeśli temperatura zewnętrzna ma istotny wpływ na zużycie ciepła na ogrzewanie pomieszczeń, to został on już zdyskontowany w zużyciu ciepła w mieszkaniach podobnych. Wprowadzenie dodatkowo temperatury zewnętrznej jako parametru modelu (obok mieszkań podobnych) byłoby działaniem nieprawidłowym, gdyż z punktu widzenia metodyki tworzenia modeli, nie wprowadza się dodatkowej zmiennej silnie skorelowanej z inną zmienną, już obecną w modelu. Wprowadzenie do modelu dwóch zmiennych silnie ze sobą skorelowanych, zbytecznie komplikuje model, a dodatkowo często może przyczynić się do problemów z takim modelem (przypadkowe błędy, przeuczenie itp.). Uzasadnieniem takiego podejścia jest też i to, aby model był oparty na parametrach już dostępnych (np. dla administratora nieruchomości), a w szczególności również łatwo dostępnych w przyszłości. O ile łatwo dostępne są

wartości temperatury zewnętrznej w ujęciu historycznym, to nie można już tego samego powiedzieć o ich prognozie dla przyszłego okresu. Również inne parametry mogą być trudne do precyzyjnego ustalenia. Dla zarządcy nieruchomości problemem może być nawet ustalenie liczby osób w lokalu, nie wspominając o charakterze użytkownika mieszkania (np. na potrzeby własne, wynajem, częściowa działalność usługowa, akademik itp.).

## Wnioski

Różnego typu układy pomiarowe montowane w systemach zaopatrzenia w ciepło umożliwiają zbieranie i gromadzenie danych historycznych opisujących zużycie ciepła przez cały system lub poszczególne jego elementy (w przypadku budynków mieszkalnych – mieszkania i lokale).

Dostęp do tych danych pomiarowych jest coraz łatwiejszy, problemem jest natomiast brak ich wykorzystywania przez zarządców nieruchomości, zainteresowanych wyłącznie bieżącym konserwowaniem systemu i rozliczeniem kosztów jego utrzymania pomiędzy użytkownikami. Zużycie ciepła w lokalach mieszkalnych zapisywane jest w pamięci ciepłomierzy. Szczytowane jest w określonej częstotliwości przez zarządcę i wykorzystywane do rozliczeń kosztów, a później zazwyczaj archiwizowane. Z danych indeksowanych w czasie można natomiast wyznaczać trend oraz profil sezonowości i te dane w sposób zindywidualizowany uwzględnić w prognozie dla danego mieszkania. Trend i sezonowość dla konkretnego mieszkania mogą być zupełnie inne niż wynika to z uśrednionego profilu dla danej klatki schodowej, piętra czy całego budynku. Wykorzystując dostępne narzędzia matematyczne do analizy szeregów czasowych możliwe jest prognozowanie w oparciu o taki zbiór danych historycznych brakującego lub przyszłego zużycia ciepła, co pozwala na zarządzanie popytem, jak i daje np. możliwość tworzenia algorytmów do regulacji systemem, aby jak najbardziej odpowiadał potrzebom użytkowników, a jednocześnie gwarantował odpowiednią efektywność energetyczną. W przyszłości mogą być więc tego typu modele wykorzystywane też do całonocnego zarządzania energią systemu i do transferu ciepła/energii pomiędzy lokalami z nadwyżką ciepła (zyski ciepła) i lokalami z jego niedoborem. Opisano w artykule tylko przykładowe modele, dla konkretnej grupy budynków i mieszkań. Jak podkreślono to w dyskusji wyników, wymagają jeszcze dalszych prac i badań,

tak aby można było też określić metody ustalania parametrów modelu, jego złożoność czy rodzaj algorytmu, który powinien być wykorzystany do jak najlepszego opisu szeregu czasowego o określonych cechach, przedstawiające zużycie ciepła w mieszkaniu na cele ogrzewania. Już teraz można jednak stwierdzić celowość kontynuacji tego typu badań, jak i kierunków wykorzystania ich wyników. Pomimo tego, że prowadzić to może do jeszcze większego skomplikowania układów regulacji elementów instalacji, to pozwoli jednocześnie dostosować ich działanie do indywidualnych potrzeb budynku i poszczególnych jego użytkowników. Ten kierunek badań widoczny jest zarówno w publikacjach dotyczących szeroko rozumianej inżynierii środowiska np. [11, 20], jak i w pracach dotyczących systemów zaopatrzenia w ciepło np. [3, 4, 5, 6, 7, 9]. Predykcja zużycia ciepła w lokalach mieszkalnych może również być wykorzystywana do diagnostyki stanów awaryjnych i szybkiego wykrywania nieprawidłowości w działaniu instalacji czy jej układów pomiarowych. Nietypowość może być zarówno cechą indywidualną danego lokalu czy grupy lokali, ale również np. sygnalizować niską efektywność energetyczną rozwiązań zastosowanych na etapie projektu lub budowy instalacji lub po prostu jej awarie.

Opisane w artykule badania potwierdzają, że przy prognozowaniu zużycia ciepła w lokalach mieszkalnych należy brać pod uwagę nie tylko temperaturę zewnętrzną, ale również inne parametry, które wpływają na bilans energii. Dla uzyskania dobrej jakości modelu duże znaczenie może mieć wyodrębnienie grup mieszkań, w których zużycie ciepła ma podobny charakter oraz poprawne przypisanie mieszkań, dla którego budowany jest model do wyodrębnionej grupy. Jak zilustrowano to na rys. 2, zawsze należy się liczyć, że charakterystyka zużycia ciepła w danym lokalu może ulegać zmianom, a lokal może "przenieść się" do innej grupy.

W badaniach opisanych w artykule skupiono się na modelach najprostszych, trzeba mieć na uwadze, że są też inne, jeszcze bardziej złożone algorytmy,

zaczynając od modeli ARFIMA-GARCH a na głębszych sieciach neuronowych kończąc. W zasadzie i one nie są trudne do zaimplementowania (dostępne są pakiety darmowe oraz takie, które są przystosowane do szybkiego użycia). Zwykle jednak interpretacja wyników i wyjaśnienie wyników w oparciu o dane wejściowe jest utrudnione, co stanowi pewną barierę w ich wdrażaniu.

#### LITERATURA:

- [1] Adamski M., Myszkowska A., Rozliczanie kosztów ogrzewania na przykładzie wybranego mieszkania, *Budownictwo i Inżynieria Środowiska* 9/2018, s. 7-14.
- [2] Ashina S., Nakata T. (2008) Energy-efficiency strategy for CO<sub>2</sub> emissions in a residential sector in Japan. *Applied Energy*, 85 (2-3), pp. 101-114.
- [3] Baldvinsson I., Development of Community Heat Supply System towards Increased Sustainability, Doctoral Thesis, Department of Management Science & Technology Graduate School of Engineering, Thoku University, August 9, 2016.
- [4] Bandurski K., Wpływ użytkowników na bilans energetyczny budynków mieszkalnych – badania i modelowanie, Rozprawa doktorska, Wydział Inżynierii Środowiska i Energetyki Politechniki Poznańskiej, Poznań 2021.
- [5] Bartnicki G., Nowak B., Predykcja zużycia ciepła w lokalu mieszkalnym na podstawie danych historycznych z ciepłomierzy mieszkaniowych, *INSTAL* 12/2020, s. 15-20. DOI: 10.36119/15.2020.12.2
- [6] Bartnicki G., Nowak B., Koniec sezonu grzewczego a efektywność energetyczna instalacji odbioru ciepła, *INSTAL* 4/2020, s. 2-11. DOI: 10.36119/15.2020.4.1
- [7] Bartnicki G., Nowak B., Model ARIMA w prognozowaniu zużycia gazu w cyklach miesięcznych. *Zeszyty Naukowe Instytutu Gospodarki Surowcami Mineralnymi i Energią PAN*. 2018, nr 103, s. 145-158. DOI: 10.24425/123712
- [8] Chen T., Guestrin C. (2016), XGBoost: A Scalable Tree Boosting System [w:] Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining – KDD '16, ACM Press, New York 2016, s. 785-794.
- [9] Dell'Isola M., Fico G., Canale L., Palella B. I., Puglisi G., An IoT Integrated Tool to Enhance User Awareness on Energy Consumption in Residential Buildings, *Atmosphere* 10/2019 (743). DOI:10.3390/atmos10120743
- [10] Dudziński K., Nowa dyrektywa wzmacnia możliwości oszczędnego gospodarowania ciepłem w budynkach mieszkalnych, *INSTAL* 2/2019, s. 26-28.
- [11] Gajda J., Bartnicki G., Burnecki K., Modeling of water usage by means of ARFIMA-GARCH

processes, *Physica A: Statistical Mechanics and its Applications*, Vol. 512, 15 December 2018, s. 644-657

- [12] Harvey L.D.D., *A Handbook on Low-Energy Buildings and District-Energy Systems*, Earthscan, London – Sterling VA 2006.
- [13] IEA, *Promoting Energy Efficiency Investments. Case studies in the residential sector*, OECD/IEA and AFD, 2008
- [14] Krzyśko M., Wołyński W., Górecki T., Skorzybut M., *Systemy uczące się. Rozpoznawanie wzorców, analiza skupień i redukcja wymiarowości*, WNT, Warszawa 2008.
- [15] Levermore G.J., *Building energy management systems: applications to low-energy HVAC and natural ventilation control*, E & FN Spon, 2000.
- [16] Michalak P., *Badania efektywności energetycznej budynku użyteczności publicznej wykorzystującego odnawialne źródła energii*, Rozprawa doktorska, Wydział Inżynierii Mechanicznej i Robotyki AGH, Kraków 2009.
- [17] Michnikowski P., Szczechowiak E., *Rozliczanie kosztów ogrzewania lokali uzależnione od rzeczywistego zużycia energii dla ogrzewania w budynkach wielorodzinnych*, *Rynek Energii* 6/2013.
- [18] Nikolaou T., Kolokotsa D., Stavrakakis G., *Review on methodologies for energy benchmarking, rating and classification of buildings*, *Advances in Building Energy Research* 5/2011, s. 53-70. DOI:10.1080/17512549.2011.582340
- [19] Nowak B., Bartnicki G., *Prognozowanie przedziału czasowego z maksymalnym w ciągu doby zużyciem gazu przez kotłownię*. *Zeszyty Naukowe Instytutu Gospodarki Surowcami Mineralnymi i Energią PAN*. 2019, nr 109, s. 93-109. DOI: 10.24425/znigsme.2019.130166
- [20] Osman A. A., Ahmed A. N., Chow M. F., Huang Y. F., El-Shafie A., *Extreme gradient boosting (Xgboost) model to predict the groundwater levels in Selangor Malaysia*, *Ain Shams Engineering Journal*, dostęp online, 22 January 2021
- [21] Ozmen, A., Yilmaz, Y. i Weber, G.W., *Natural gas consumption forecast with MARS and CMARS models for residential users*, *Energy Economics* 70/2018, s. 357-381. DOI: 10.1016 / j.eneco.2018.01.022
- [22] Pelsmakers S., *The environmental design pocketbook*, RIBA Publishing, London 2015.
- [23] PN-EN 12831:2006 *Instalacje ogrzewcze w budynkach – – Metoda obliczania projektowego obciążenia cieplnego*.
- [24] PN-B-03406:1994 *Ogrzewnictwo – – Obliczanie zapotrzebowania na ciepło pomieszczeń o kubaturze do 600 m<sup>3</sup>*
- [25] Wójcik F., *Prognozowanie dziennych obrotów przedsiębiorstwa za pomocą algorytmu XGBoost – studium przypadku*, *Studia Ekonomiczne w Katowicach*, Nr 375, Katowice 2018, s. 121-140.