

Potencjał algorytmów inteligencji obliczeniowej w interpretacji rozproszonych danych wodomierzowych

The potential of computational intelligence algorithms in the interpretation of distributed water meter data

PAULINA DZIMIŃSKA, JUSTYNA STAŃCZYK, KRZYSZTOF PAŁCZYŃSKI, PAWEŁ LICZNAR, TOMASZ ANDRYSIAK

DOI 10.36119/15.2024.1.5

Rozwijająca się świadomość oraz aspekty prawne związane z energochłonnością systemów dystrybucji wody, w połączeniu z starzeniem się infrastruktury wodociągowej i stresem wodnym, wymuszają poszukiwanie rozwiązań wspierających efektywniejszą kontrolę i zarządzanie infrastrukturą techniczną. Osiągnięcie standardu mądrych czy też inteligentnych systemów wodociągowych na każdym szczeblu obszaru kluczowego nadal pozostaje kwestią otwartą zarówno w warunkach krajowych jak i zagranicznych. Dotyczy to także mikroskali sieci wodociągowych, to znaczny konsumentów wody oraz stosowania inteligentnych wodomierzy z wbudowanymi algorytmami uczenia maszynowego. Artykuł przedstawia wyniki badań z wdrożenia modelu krótkoterminowej predykcji zużycia wody wraz z detekcją anomalii dla budynków wielorodzinnych. Prognoza zużycia wody, przeprowadzona w oparciu o wysokoczęstotliwościowe pomiary oraz głębokie sieci neuronowe, pozwoliła na osiągnięcie błędu predykcji poniżej 3,0%. Detekcja wykrywania anomalii, zrealizowana w oparciu o bazowy model prognostyczny, charakteryzowała się nawet 97,3% skutecznością wykrywania anomalii.

Słowa kluczowe: sieci wodociągowe; uczenie maszynowe; wodomierze; zużycie wody

The development of awareness and legislative aspects related to the energy efficiency of water distribution systems, combined with the ageing of water supply infrastructure and water stress, led to the search for solutions to support more effective control and management of technical infrastructure. Increasing the standard of smart or intelligent water supply systems at all levels of key areas is still a problem under domestic and foreign conditions. This also applies to the microscale of water supply networks, namely water consumers of water and the use of smart water meters with integrated machine learning algorithms. This article presents the results of research on the implementation of a short-term water consumption prediction model with anomaly detection for multifamily residential buildings. The prediction of water consumption, based on high-frequency measurements and deep neural networks, achieved a prediction error of less than 3.0%. Anomaly detection, based on the underlying prediction model, had up to 97.3% accuracy.

Keywords: water supply networks; machine learning; water meters; water consumption

Wprowadzenie

Wodomierze, jako urządzenia montowane w celu bilingowego pomiaru objętości zużytej wody, są stosowane w Polsce już od lat 90-tych XX w. Ich funkcjonalność kojarzona jest przede wszystkim z fakturowaniem usług dostarczania wody w wyniku tego, że stanowią podstawę do nali-

czenia opłat użytkownikom sieci wodociągowej. Główny cel nadrzędny stosowania wodomierzy, związany z przychodem przedsiębiorstw wodociągowych za świadczone usługi wodne, jest niepodważalny, uzasadniony, a nawet wymagany Rozporządzeniem Ministra Infrastruktury z dnia 12 kwietnia 2002 r. w sprawie warunków technicznych, jakim powinny

odpowiadać budynki i ich usytuowanie [1]. Zaznaczyć należy jednak, że zakres montażu wodomierzy został znacznie rozszerzony w 2021 roku po podpisaniu przez Prezydenta RP Ustawy o zmianie ustawy o efektywności energetycznej oraz niektórych innych ustaw [2], która zobowiązuje właścicieli lub zarządców budynków wielolokalowych do montażu do

mgr inż. Paulina Dżimińska <https://orcid.org/0000-0003-0427-5960> – Miejskie Wodociągi i Kanalizacja w Bydgoszczy – sp. z o.o., Bydgoszcz
dr inż. Justyna Stańczyk <https://orcid.org/0000-0002-5676-1787> – Uniwersytet Przyrodniczy we Wrocławiu, Instytut Inżynierii Środowiska, Wrocław
mgr inż. Krzysztof Pałczyński <https://orcid.org/0000-0003-3989-3400> – Politechnika Bydgoska im. Jana i Jędrzeja Śniadeckich, Wydział Telekomunikacji, Informatyki i Elektrotechniki, Bydgoszcz

prof. dr hab. inż. Paweł Licznar <https://orcid.org/0000-0002-2559-5296> – Wydział Instalacji Budowlanych, Hydrotechniki i Inżynierii Środowiska, Politechnika Warszawska, Warszawa,

dr hab. inż. Tomasz Andrysiak, prof. PBŚ <https://orcid.org/0000-0001-8138-6619> – Politechnika Bydgoska im. Jana i Jędrzeja Śniadeckich, Wydział Telekomunikacji, Informatyki i Elektrotechniki, Bydgoszcz

Autor korespondencyjny / corresponding author: justyna.stanczyk@upwr.edu.pl

stycznia 2027 r. urządzeń wodomierzowych zdalnego odczytu także dla ciepłej wody. Nierozłączne jest powiązanie wody z energią, określane w języku angielskim mianem *water-energy nexus*. Oczywista jest potrzeba precyzyjnego opomiarowania zużycia wody w dobie rosnących cen za energię elektryczną i obniżenia energochłonności obiektów technicznych, także takich jak wodociągi. Praca sieci wodociągowych, będących obiektami pracującymi pod działaniem układów pompowych, może zostać zoptymalizowana na podstawie precyzyjnych danych pomiarowych, rejestrowanych u konsumentów. Szacuje się, że energochłonność systemu dystrybucji wody będzie wzrastała. Przykładowo, dla Chin zużycie energii na potrzeby zaopatrzenia w wodę wynosiło w 2014 roku 81 mld kWh, ale prognozowano, że do 2020 roku wzrośnie do 90 mld kWh, a do 2030 nawet do 109 kWh, przy czym najbardziej energochłonnym sektorem w kontekście dostaw wody będzie sektor mieszkaniowy [3].

Zakres funkcjonalności wodomierzy, oprócz podstawowej związanej z opłatami za wodę, staje się coraz bardziej rozszerzany o możliwości wykorzystania danych o zużyciu wody do prognozowania zapotrzebowania na wodę, analiz społeczno-ekonomicznych i behawioralnych, detekcji anomalii [4], a także docelowo do optymalizacji pracy pompowni. Bardzo często określa się tego typu urządzenia mianem „smart” czy też „intelligent” i nierzadko przedrostki te są niefortunnie stosowane w odniesieniu do wodomierzy posiadających jedynie możliwość radiowego odczytu ich stanu w systemie objazdowym. Bez zaawansowanej struktury pomiarowej AMI (ang. *Advanced Metering Infrastructure*) nie ma mowy o rozwoju inteligentnych IWSs (ang. *Intelligent Water Systems*) czy też smart sieci wodociągowych SWSs (ang. *Smart Water Systems*), jednak granica pomiędzy przeciętnym systemem, a co najmniej mądrym, nie zawsze jest jednoznacznie wyznaczona. Dotyczy to nie tylko samych urządzeń pomiarowych i sieci wodociągowych, ale ogólnego nadużywania sformułowania „smart” w obszarze infrastruktury technicznej. Co bowiem stanowi o mądrości systemu dystrybucji wody czy też urządzeń monitoringowych? Odpowiedź nie jest oczywista. Li i in. [5] wskazują, że z mądrym systemem wodociągowym mamy do czynienia wówczas, gdy opóźnienie czasowe pomiędzy danymi wejściowymi a informacją zwrotną dla eksploatatora sieci, co do jej stanu, nie jest

dłuższe niż piętnaście minut. Możliwość otrzymania informacji diagnostycznie użytecznej jest jednak zależna od wysokoczęstotliwościowych danych pomiarowych [6]. Tymczasem nawet 60% rejestrowanych danych nie jest wykorzystana [7] z uwagi na brak narzędzi ich bieżącego przetwarzania, analizowania i wnioskowania. Istnieje więc potrzeba rozwijania bardziej zaawansowanych funkcjonalności wodomierzy opierających się na inteligencji obliczeniowej, która może wspomóc procesy diagnostyczne i decyzyjne związane z codzienną eksploatacją sieci wodociągowych w czasie quasi-rzeczywistym. Tymczasem, dane wodomierzowe są zazwyczaj gromadzone w postaci skumulowanego poboru wody [8], odczytywanego radiowo w systemie objazdowym, np. w miesięcznym okresie rozliczeniowym i nie mogą stanowić podstawy do analiz szeregów czasowych zużycia wody z wykorzystaniem zaawansowanych metod inteligencji obliczeniowej. Jak zaznaczają Rahim i in. [4] w ciągu ostatnich dziesięciu lat krótkoterminowe prognozy zapotrzebowania na wodę w budynkach mieszkalnych były częściej wdrażane badawczo, głównie z uwagi na coraz większą dostępność danych wysokiej częstotliwości generowanych przez wodomierze cyfrowe DWMs (ang. *Digital Water Meters*).

Każdy z elementów systemu dystrybucji wody może mieć wbudowany w algorytmie swego działania pierwiastek inteligencji obliczeniowej, w zależności od stopnia integracji technologii informacyjnych i komunikacyjnych, które zostały wdrożone w procesie zarządzania za pośrednictwem inteligentnych sieci Smart Grid, czy też systemów pomiarowych do zbierania danych w czasie rzeczywistym [5]. Na podstawie literatury, przyjąć można, że z SWSs czy też IWSs mamy do czynienia wówczas, gdy wdrożony jest innowacyjny system monitorowania, który pozwala na inteligentne zarządzanie zasobami wodnymi. Badania Li i in. [5] wykazały, że problematyka ta jest nadal poruszana w największej mierze jedynie na szczeblu naukowym, a nie przemysłowym czy też rządowym. Hamilton i in. [9] dość jednoznacznie określili, kiedy system wodociągowy można uznać za podstawowy, przeciętny, mądry i inteligentny. Wyznaczono tym samym jednoznaczne ich granice na każdym szczeblu obszarów kluczowych, tj. opomiarowania parametrów hydraulicznego przepływu w strefach DMA (ang. *District Metered Area*), na wejściu do całych sieci dystrybucji wody (ang. *Bulk Flow Measurement*), wodomie-

rzy, zarządzania ciśnieniem, wyciekami wody i infrastrukturą, bilansowania wody i kluczowych wskaźników efektywności KPIs (ang. *Key Performance Indicators*), a także zasobów ludzkich. W zależności od tego, który z obszarów kluczowych jest brany pod uwagę, jego analiza z uwzględnieniem ściśle określonych parametrów na różnych poziomach daje przesłanki do wnioskowania, czy sieć wodociągowa jest podstawowa, przeciętna, mądra czy też inteligentna. Pomijając bardzo szczegółowe wytyczne, zawarte w części tabelarycznej opracowania Hamilton i in. [9] wspólnym mianownikiem dla sieci inteligentnych są kompletne algorytmy uczenia maszynowego wdrożone w celu kontroli systemu i zarządzania wejściowymi parametrami objętości i ciśnienia wody w sieci.

W niniejszym artykule przedstawiono potencjał algorytmów inteligencji obliczeniowej w interpretacji rozproszonych danych wodomierzowych o wysokiej częstotliwości rejestracji. Podjęto próbę wdrożenia metod *machine learningu* w kontekście predykcji zapotrzebowania na wodę oraz detekcji anomalii, przeprowadzonej na podstawie prognoz w krótkim horyzoncie czasowym.

Charakterystyka obszaru badawczego

W ramach przeprowadzonych badań wykorzystano szeregi czasowe zarejestrowane dla dziewiętnastu odbiorców końcowych wody, zlokalizowanych w ramach jednej strefy DMA. Szczegółowy opis poligonu badawczego można odnaleźć w artykule Dżimińskiej i in. [6]. Analizowana strefa DMA charakteryzuje się przede wszystkim tym, że zasilana jest jednokierunkowo, a panujące w niej ciśnienie warunkuje pracę lokalnej stacji podwyższania ciśnienia (hydroforni). Strefa wodociągowa obejmuje obszarowo głównie budownictwo wielorodzinne o wysokiej zabudowie. Spośród siedemnastu budynków wielorodzinnych, trzy zasilane są dwoma osobnymi przyłączami, zatem opomiarowano je dwoma wodomierzami. Nieruchomości te w większości są zbliżone do siebie pod względem wielkości, wyrażonej liczbą mieszkań oraz liczbą mieszkańców (tabela 1). Ponad połowa z nich to budynki składające się z około 82 mieszkań, gdzie liczba mieszkańców mieści się w zakresie od 141 do 168. Natomiast mniej niż połowę stanowią obiekty mniejsze. W zbiorze znalazł się również jeden obiekt zdecydowanie większy, składający się ze 115 mieszkań,

Tab. 1 Charakterystyka badanych budynków wielorodzinnych

Parametr	Numer budynku													
Nazwa	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Liczba mieszkań	82	83	115	55	83	60	84	45	83	40	83	82	80	55
Liczba mieszkańców	168	141	239	118	bd	111	166	103	151	69	158	151	165	108

kórego szacowana liczba mieszkańców wynosiła 239 osób. Ponadto, w analizowanej strefie znajdują się dodatkowo obiekty, takie jak szkoła mała oraz szkoła duża z basenem, zasilane z odrębnych przyłączy wodociągowych [6].

Na potrzeby wykonania pomiarów, na przyłączach wodociągowych, zamontowane zostały wodomierze wyposażone w rejestrator impulsów i przetworniki ciśnienia, które dokonywały rejestracji danych dotyczących parametrów przepływu hydraulicznego z dużą częstotliwością. W przypadku impulsów z wodomierzy urządzenia rejestrowały w czasie objętość pobieranej wody przez użytkowników sieci z taktowaniem jedno – lub dziesięcio – litrowym. Natomiast pomiar ciśnienia odbywał się z częstotliwością 1-minutową. Urządzenia rejestrujące wyposażone były w modem, za pomocą którego dane w czasie rzeczywistym przesyłano do zbiorczej bazy komputerowej. Baza komputerowa wyposażona była w aplikację webową, która surowe oraz zagregowane do szeregów 15-minutowych i 1-godzinnych dane pomiarowe wizualizowała w formie wykresów możliwych do obserwacji z poziomu interfejsu użytkownika.

Rejestracja danych o poborze wody realizowana była przez okres ponad ośmiu miesięcy poczynwszy od 1 stycznia 2021 r. [6]. Warto podkreślić, że badania prowadzone były jeszcze w okresie obowiązywania ograniczeń związanych z pandemią COVID-19. Zgodnie z początkowymi przypuszczeniami, zarejestrowane charakterystyki znacząco odbiegały od standardowych warunków zużycia wody [10]. Wpływ panującego w tym czasie lockdown'u uwidocznił się w szczególności w przypadku szkoły wyposażonej w basen pływacki, która była w tym czasie zamknięta oraz pawilonów handlowo-usługowych, które mogły funkcjonować z pewnymi ograniczeniami. Z tego też powodu wykluczono niefunkcjonujące wówczas obiekty cukierni czy też domu kultury podczas wdrożenia prognoz i detekcji anomalii. Na zmianę charakterystyki poboru wody przez mieszkańców zabudowy wielorodzinnej miała również wpływ pandemia COVID-19, gdyż w owym czasie, duża część społeczeństwa pracowała zdalnie. Wyraźny wpływ pandemii na zmiany w poborze wody był

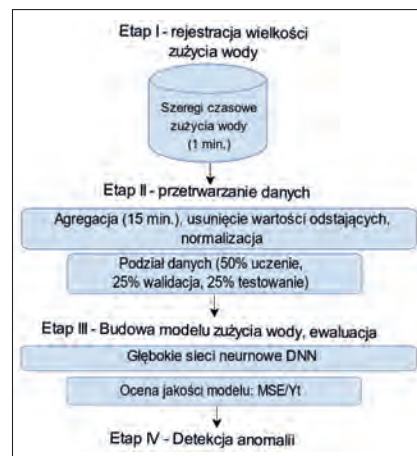
obserwowany zarówno w kraju jak i za granicą [10, 11].

Metodyka i zakres badań

Do wdrożenia prognoz zapotrzebowania na wodę oraz detekcji anomalii wykorzystuje się między innymi metody oparte na sieciach neuronowych ANNs (ang. Artificial Neural Networks), lasach losowych RF (ang. Random Forest), metody stochastyczne oraz tworzące hybrydy, będące połączeniem kilku technik analitycznych. ANNs są często wdrażane do przewidywania zużycia wody w różnych skalach wielkości jej poboru i z różnym horyzontem prognoz. Prognozy krótko-, średnio- i długoterminowe dostarczane przedsiębiorstwom wodociągowym pomagają zapewniać niezbędną wydajność dostaw, przynosząc bieżące korzyści operacyjne. Persantez i in. [12] w kontekście detekcji anomalii wykazali walory wdrożeniowe RF i ANNs, bazując na danych pochodzących z precyzyjnych wodomierzy cyfrowych. Wdrożono klasyfikację w oparciu o algorytm różnicujący zużycie na poszczególne klastry stanu na podstawie średniego godzinowego zużycia wody z uwzględnieniem pory dnia. Bardziej rozbudowane narzędzia analityczne dedykowane detekcji wycieków mają wdrożone podejście oparte o modele uczenia maszynowego ML (ang. Machine Learning). Jak wskazują wyniki badań, sztuczne sieci neuronowe oraz nadzorowany model ML mogą dokonać klasyfikacji anomalii, wykorzystując dwa typy szeregów czasowych, tj. względnie normalnych warunków pracy oraz zawierających awarie [13]. Nierzadko jednak część badań opiera się w odniesieniu jedynie do całej strefy DMA w świetle braku wysokoczęstotliwościowych danych wodomierzowych [14].

W ramach przeprowadzonych badań do interpretacji, predykcji i detekcji anomalii wykorzystano dane wodomierzowe z pojedynczych budynków wielorodzinnych, a w tym celu wdrożono głębokie uczenie maszynowe (ang. Deep Learning), ściślej model liniowej głębokiej sieci neuronowej DNN (ang. Deep Neural Networks). Pierwszym etapem badań było wstępne przetworzenie danych (tzw. *pre-processing*) realizujące m.in. agregację danych, a następnie utworzono model

predykcji krótkoterminowej zużycia wody, zweryfikowano jego jakość i przy jego użyciu dokonano detekcji anomalii. Schemat zastosowanej metodyki przedstawiono na rysunku 1.



Rys. 1
Schemat metodyczny badań

Pre-processing danych

Analizę możliwości wykorzystania metod *machine learningu* do prognozowania zapotrzebowania na wodę rozpoczęto od wykonania *pre-processingu* danych, ponieważ ich oryginalna postać zawierała rejestrację impulsów generowanych przez wodomierze po przepłynięciu odpowiedniej objętości wody (jednego lub dziesięciu litrów) wraz z adekwatnym stemplem czasowym. Dane źródłowe, czyli obserwacje o długości przedziałów czasowych pomiędzy impulsami z wodomierzy, zostały zagregowane do rozdzielczości: piętnastu minut, trzydziestu minut oraz jednej godziny. Dopiero w tej postaci mogły reprezentować zużycie wody w założonej jednostce czasu i zostać poddane dalszej analizie. Pośrednim celem tego zabiegu było określenie zdolności predykcyjnych i diagnostycznych algorytmów bazujących na danych zagregowanych na różnych poziomach czasowych. Jednakże, biorąc pod uwagę potencjał aplikacyjny podejmowanych analiz w praktyce, uwagę skoncentrowano na predykcji i detekcji anomalii bazujących na danych zagregowanych do 15-minutowego interwału pomiarowego. Powyższy interwał pomiarowy jest swoistym kompromisem pomiędzy generalizacją trudnego do zamodelowania i niedeterministycznego zróżnicowania poborów wody w skali pojedynczych minut czy też nawet sekund, a zarazem postulatem, aby w przypadku mądrych sieci wodociągowych opóźnienie czasowe pomiędzy powstaniem informacji o jej stanie, a przesłaniem jej do eksploatatora nie było dłuższe niż 15 minut [5].

W kolejnym etapie badań wykonano szereg zabiegów mających na celu eksplorację i odpowiednie przygotowanie danych. Pierwszą z nich było usunięcie wartości odstających czyli wyraźnie mniejszych lub większych od standardowych ich zakresów. Do tego celu posłużono się testami określającymi zgodność dystrybucji badanych rozkładów do dystrybucji rozkładu normalnego (rozkładu Gaussa), m.in. testem Shapiro-Wilka. Jak się okazało, większość badanych rozkładów była zgodna z rozkładem normalnym. Zatem można było na tej podstawie dokonać usunięcia z danych wartości odstających. Wstępnie oczyszczone dane poddano standaryzacji za pomocą techniki ang. z-score, która polega na wyznaczeniu średniej wartości oraz odchylenia standardowego i zakodowaniu sygnałów w sposób określony wzorem [15] (1):

$$z_i = (x_i - \mu) / \sigma \quad (1)$$

gdzie:

- z_i – zmienna znormalizowana,
- x_i – zmienna obserwowana, niestandardyzowana,
- μ – średnia wartość oczekiwana,
- σ – odchylenie standardowe.

Normalizacja ma na celu zapewnienie numerycznie poprawnych warunków podczas określania i propagacji gradientu błędu w trakcie tworzenia modeli predykcyjnych. W kolejnym kroku znormalizowany zbiór danych został podzielony na trzy podzbiory: uczący (50%), testowy (25%) i walidacyjny (25%). Na koniec etapu przygotowania danych do modelowania, szeregi czasowe podzielone zostały na 24-godzinne okna nakładające się na siebie. Z założenia, dokonywano predykcji wartości prognozowanej na podstawie wektora wejściowego stanowiącego dane okno czasowe.

Predykcja zużycia wody

Do realizacji krótkoterminowej predykcji wybrano model liniowej głębokiej sieci neuronowej DNN. Modele z zakresu głębokiego uczenia DL, stosowane w ramach uczenia maszynowego, w swej idei mają za zadanie odwzorować pewne złożone zjawiska za pomocą struktur warstwowych (hierarchicznych), w ramach których najniższe warstwy sieci reprezentują ich najprostsze cechy, a najwyższe odzwierciedlają ogólną postać danych zjawisk w relacji z warstwami niższego rzędu. Głębokie modele uczenia charakteryzują się więcej niż dwiema nieliniowościami (warstwami przetwarzania) pomiędzy sygnałem wejściowym a wyjściowym. Spo-

śród szeregu dostępnych i zrealizowanych w ramach badań modeli, wybrano DNN ze względu na wysoką wydajność obliczeniową oraz zdolności predykcyjne dające w efekcie najmniejszy błąd prognozy. Architektura sieci składała się z dwóch warstw liniowych, przy czym warstwa wejściowa sieci zawierała 24 neurony, a na jej wyjściu znajdował się jeden neuron dokonujący predykcji rozbiórki wody. Sieć posiadała jedną warstwę ukrytą złożoną z 10 neuronów wykorzystujących ReLU (ang. Rectified Linear Unit) jako funkcję aktywacji. Funkcja ta ma za zadanie wspomóc sieci neuronowe podczas modelowania z wykorzystaniem głębokich sieci neuronowych, w kontekście zanikających gradientów i niewielkich zmian wag w warstwach neuronów, które usytuowane są najdalej w odniesieniu do centrum sieci neuronowej. Zanikający gradient stanowi podstawowy problem towarzyszący implementacji głębokich sieci neuronowych, któremu zapobiec można wykorzystując właśnie typ neuronów ReLU. Ponadto, jak zaznaczają Guo i Liu [16] konwencjonalne sieci neuronowe mają ograniczoną efektywność ze względu na nieliniowy charakter zmienności zapotrzebowania na wodę, podczas gdy głębokie sieci neuronowe odznaczają się lepszą jakością predykcji, są bardziej elastyczne i skuteczne.

Ewaluacja prognozy

Przystępując do ewaluacji prognoz poborów wody miano pełną świadomość braku możliwości obiektywnego porównania uzyskanych wyników predykcji do wyników badań krajowych i zagranicznych innych autorów z uwagi na różnorodność stosowanych w tym celu metryk. Zauważono przy tym, że w piśmiennictwie najczęściej używanym błędem do oceny jakości modeli prognostycznych rozbiórki wody jest MAPE (ang. Mean Absolute Percentage Error) [17, 18, 19]. Błąd ten choć wydaje się być niezależny, według autorów sprzyja danym o niskiej dynamice, a ponadto jest metryką relatywną. Z uwagi na występowanie zerowych wartości poborów wody, zwłaszcza w godzinach nocnych, w badaniach nie było możliwości stosowania do oceny jakości prognoz miar o charakterze relatywnym, w przypadku których istniałaby konieczność dzielenia różnic pomiędzy wartościami obserwowanymi i prognozowanymi przez obserwowane zerowe wartości przepływów.

Z uwagi na powyższe czynniki, zgodnie z rekomendacjami Siwonina i in. [20], w odniesieniu do oceny jakości prognoz

zużycia wody posłużono się znormalizowanym wskaźnikiem błędu średniokwadratowego MSE (ang. Mean Square Error) odniesionym do niezerowej wartości średniej przepływu, według wzorów (2) i (3):

$$MSE = \frac{1}{n} \sum_{t=1}^n (Y_t - \hat{Y}_t)^2 \quad (2)$$

$$\frac{MSE}{\bar{Y}_t} = \frac{\frac{1}{n} \sum_{t=1}^n (Y_t - \hat{Y}_t)^2}{\frac{1}{n} \sum_{t=1}^n Y_t} \quad (3)$$

gdzie:

$t = 1, 2, 3, \dots, n$ – wskaźnik czasu dla prognozowania zbioru n – elementów tworzącego szereg czasowy,

Y_t – wektor obserwacji zużycia wody, tj. wartości opomiarowanego zużycia wody (zmienna podlegająca prognozie),

$\hat{Y}_t$ – wektor predykcji zużycia wody wygenerowany przez ewaluowany model, tj. prognozowanych wartości zużycia wody.

Detekcja anomalii

Przed przystąpieniem do kolejnego zagadnienia, a mianowicie detekcji anomalii, konieczne było zdefiniowanie pojęcia anomalii w odniesieniu do poboru wody. Na zjawisko anomalii może mieć wpływ szereg różnych czynników, począwszy od awarii lub wycieku w instalacji wodociągowej, a skończywszy na wpływie specyficznych czynników losowych jakim była pandemia COVID-19 [10]. W niniejszych rozważaniach anomalia rozumiana jest jako zarejestrowana wartość wyraźnie odstająca od wartości generowanej przez predykcję. Zatem podejmuje się próbę wykrywania anomalii, powstałych w wyniku wycieku czy awarii instalacji u odbiorcy końcowego wody. Podczas prowadzonych badań, w żadnym z obiektów nie zarejestrowano rzeczywistych zjawisk anomalnych, dlatego istniała konieczność ich symulacji poprzez wygenerowanie awaryjnych poborów zużycia wody. Przyjęto, że wartość anomalna to taka, która stanowi od 5% do 20% większy przepływ wody, natomiast duża awaria na instalacji będzie stanowiła przepływ $\geq 20\%$ w stosunku do okresu bezawaryjnego.

Bazę wejściową do detekcji anomalii stanowiły szeregi czasowe zagregowane do wartości 15-minutowych. Do detekcji anomalii wykorzystano algorytm, którego działanie opiera się na bazowym modelu predykcyjnym, opisanym powyżej. O istnieniu anomalii świadczy wysoka wartość błędu między predykcją a bieżącą

obserwacją. Z założenia niski błąd prognozy świadczy o standardowych warunkach pracy wodociągu, natomiast wysoki o rozpoczęciu anomalii. Stworzony autorski algorytm miał za zadanie dokonać klasyfikacji, czy bieżący pomiar należy do przedziału, w którym nie występują anomalie oraz określić, czy błąd predykcji zawiera się w przedziale ustalonego rozstępu kwartylowego od mediany błędu w okresie poprzedzających 24 godzin. Jeżeli błąd absolutny był poniżej zdefiniowanego błędu minimalnego i zachowane były powyższe warunki, pomiar nie był klasyfikowany jako anormalny.

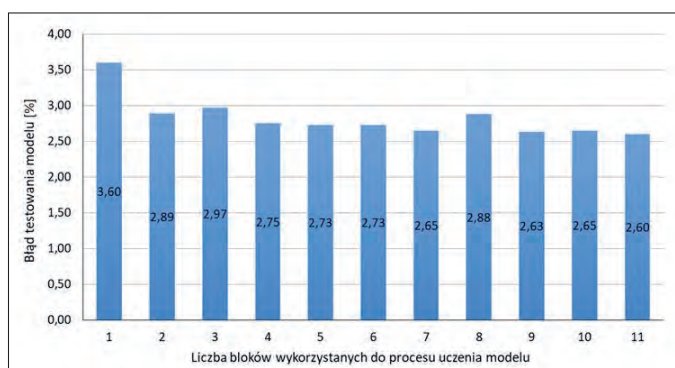
Do identyfikacji anomalii wykorzystywano jednowymiarowe kryterium kwartylowe, które wykorzystywane jest do budowy wykresów pudełkowych wprowadzonych przez Tukeya [21]. W tym przypadku konieczne było obliczenie pierwszego (Q1) i trzeciego kwartyla (Q3), jak również rozstępu międzykwartylowego (IQR) równego $IQR = Q3 - Q1$. Dzięki tej operacji możliwe stało się określenie zjawisk anormalnych, ponieważ przyjęto, że wszystkie wartości wykraczające poza przedział $(Q1 - 1,5 \cdot IQR, Q3 + 1,5 \cdot IQR)$ są wartościami anormalnymi. Oprócz kryterium kwartylowego, założono dodatkowy parametr decydujący o wystąpieniu zjawisk anormalnych, a mianowicie liczba próbek. Z uwagi na stosowanie algorytmu na zagregowanym 15-minutowym szeregu czasowym i możliwe chwilowo występującej pojedynczej wartości odstającej, z uwagi choćby na znane doskonale eksploatorom sieci zjawiska wzmożonego poboru na przykład w przerwach transmisji ważnych meczy piłkarskich, przyjęto dodatkowe kryterium występowania trzech kolejno występujących po sobie w przedziałach 15-minutowych podejrzeń anomalii do ostatecznego zdiagnozowania zaistnienia anomalii odpowiadającej awarii, wyciekowi na instalacji odbiorczej.

Wyniki predykcji zużycia wody i detekcji anomalii

Prognozy zużycia wody zostały przeprowadzone z wykorzystaniem dwuwarstwowej sieci neuronowej z funkcją aktywacji typu liniowego (ReLU). Błąd predykcji został określony dla trzech przedziałów czasowych agregacji, tj. jednogodzinnej, trzydziesto – oraz piętnastominutowej przy jednoczesnym założeniu, że bieżąca prognoza zostanie zrealizowana na podstawie ostatnich dwudziestu czterech wartości zużycia, co wynikało z doświadczeń z wcześniejszych badań Pałczyńskiego i in. [22, 23, 24].

Z uwagi na to, że dla budynków wielorodzinnych często operuje się zagregowanymi poborami wody do dłuższych przedziałów czasu, nawet godzinowych, w ramach prowadzonych badań przeanalizowano możliwość predykcji dla szeregu przyjętych poziomów agregacji poboru wody wynoszących od bazowych 15-minut poprzez 30 minut do 1 godziny. Ponadto, zweryfikowano czy stosowanie do uczenia sztucznych sieci neuronowych szeregów czasowych poborów wody z poszczególnych budynków wielorodzinnych pozwala na utworzenie modeli o podobnych parametrach. Za pomocą testu ANOVA sprawdzono, czy utworzony model sieci neuronowej dla konkretnego budynku jest na tyle uniwersalny, że może być z powodzeniem zastosowany dla innego budynku wielorodzinnego. Inaczej tę kwestię ujmując, zweryfikowano jaka jest przewidywana jakość predykcji modelu dla budynku mieszkalnego wielorodzinnego, którego szeregów czasowych nie wykorzystano w procesie uczenia sieci neuronowej, a uczenie sieci zrealizowano na podstawie innych budynków mieszkalnych wielorodzinnych. Praktyczną przesłanką do tego było zagadnienie weryfikacji, na ile można próbować opracować uniwersalny model sieci neuronowej do prognozowania poborów wody dla całego zbioru budynków wielorodzinnych. W tym celu eksperyment składał się z wylosowania pewnej liczby sygnałów z budynków wielorodzinnych do wytrenowania sieci neuronowej i dokonania ewaluacji dla pozostałych obiektów, z których pomiary nie zostały wykorzystane w procesie uczenia. Eksperymenty powtórzono trzydzieści razy, wyniki uśredniono, a ich przykładowe wyniki dla 15-minutowej rozdzielczości szeregów czasowych rozbiórów wody przedstawiono na rysunku 2.

Rys. 2
Wyniki jakości predykcji dla obiektów (budynków wielorodzinnych) w obrębie poligonu badawczego strefy DMA



W wyniku przeprowadzonych analiz stwierdzono, że jakość predykcji przeprowadzonej na podstawie szeregu czasowego o rozdzielczości 15-minutowej tylko jednego budynku wynosi 3,60%, ale wy-

korzystanie danych z dwóch budynków do treningu obniżyło błąd predykcji do 2,89%. Zastosowanie wszystkich jedenastu budynków pozwoliło na uzyskanie dalszego, ale już bardzo ograniczonego spadku błędu predykcji do wartości 2,60%.

Testy Tukey'a oraz analiza ANOVA wykazały, że istotny jest rozkład błędów prognozowania w przypadku zastosowania w procesie uczenia danych zarejestrowanych dla jednego i dwóch budynków. Nie wykazano natomiast różnicy między jakością predykcji dla dowolnych dwóch sygnałów zużycia wody dla budynków wykorzystanych do uczenia sieci, co oznacza, że trenowanie modelu (uczenie sieci neuronowej) na danych zgromadzonych z dwóch różnych budynków jest statystycznie istotne, lecz dodatkowe zwiększenie ilości wykorzystywanych w tym celu zbiorów danych z kolejnych budynków wielorodzinnych nie powodowało znaczącej różnicy. Wyniki badań dowodzą, że model sieci neuronowej, który został wytrenowany na danych o zużyciu wody z dwóch budynków wielorodzinnych może być z powodzeniem wdrożony do prognoz poboru wody dla innych budynków wielorodzinnych.

Odnosząc powyższe wyniki do badań innych autorów, warto odnotować, że Guo i Liu [16] wykazali, że błąd MAPE dla konwencjonalnych i głębokich sieci neuronowych do predykcji realizowanej w 15-minutowym horyzoncie czasowym może wynosić odpowiednio 2,59 i 2,02%. Zaznaczyć należy, że w przywołanej pracy badania prowadzono na bazie poborów wody rejestrowanych dla znacznie większego obszaru wydzielonej strefy DMA, natomiast modele predykcyjne, prezentowane w niniejszym artykule, utworzono w mikroskali w odniesieniu do pojedynczych budynków wielorodzinnych. W przy-

padku analizy danych o poborach wody na poziomie budynku dane powinny charakteryzować się odpowiednio wysoką rozdzielczością czasową, o interwale nie większym od godzinowego. W badaniach

własnych i przy użyciu pełnej bazy pomiarowej poborów wody z wszystkich 11 budynków, dla godzinowej agregacji sygnału zużycia wody uzyskano błąd testowania, wyraźnie wyższy na poziomie 9,7%, dla 30-minutowej agregacji danych błąd ten zmalał do wartości 4,7% a dla 15-minutowej agregacji uzyskano już wspomniany niski błąd 2,6%, porównywalny z wynikami prac Guo i Liu [16]. Na tej podstawie można wnioskować, że im rejestracja zużycia wody odbywa się z wyższą częstotliwością, tym modele prognostyczne mogą osiągać lepszą jakość, dlatego też zaleca się rejestrację poborów wody z interwałem co najmniej 15-minutowym, o ile oprócz celów bilingowych mają zostać one wykorzystane do tworzenia modeli predykcyjnych ukierunkowanych na wykrywanie anomalii, takich jak wycieki wody.

Zastosowana w badaniach sieć neuronowa generowała akceptowalne wartości predykcji, zatem wszelkie wysokie wartości błędu powinny sugerować występowanie sytuacji anomalnych. Do przeprowadzenia badań nad detekcją anomalii sieć neuronową wytrenowano z użyciem zbioru złożonego z poborów zarejestrowanych dla 11 budynków wielorodzinnych. Ze względu na brak zarejestrowanych rzeczywistych awarii i wycieków wody z instalacji, konieczne było ich wygenerowanie zgodnie z przyjętymi w metodyce badań założeniami. Przygotowane syntetyczne szeregi czasowe z występującymi nietypowymi wartościami poboru wody zbadano za pomocą algorytmu określającego błąd pomiędzy wartością zaobserwowaną (symulowanym szeregiem czasowym) a wartością predykowaną poboru wody, generowaną przez sieć neuronową.

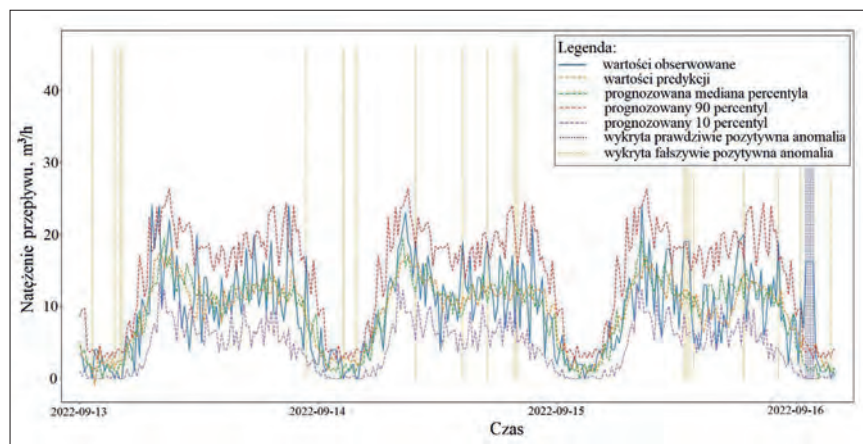
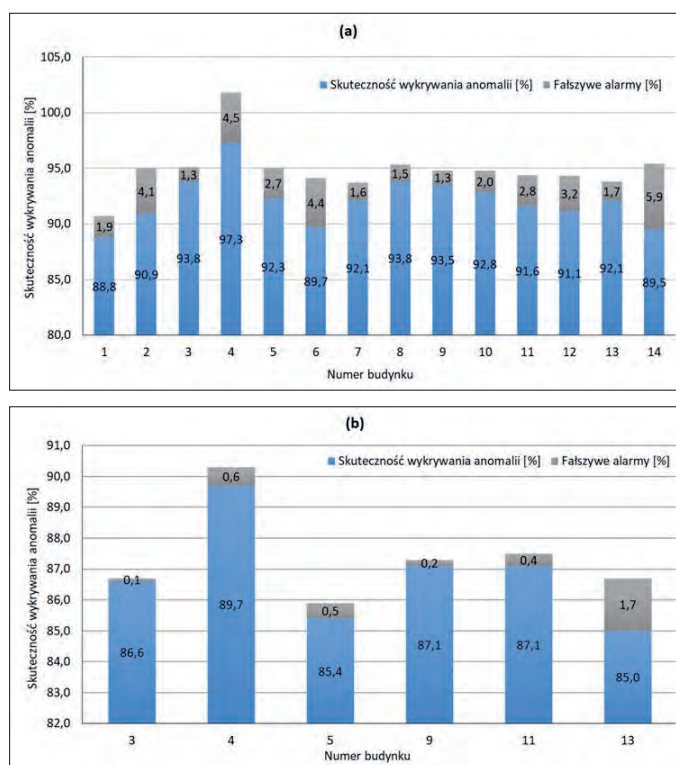
Prezentowany na rysunku 3 wykres przedstawia przykładowy wynik prowadzonych badań detekcji anomalii dla sze-

regu czasowego, składającego się z trzech dób dla wybranego budynku wielorodzinnego. Jest to wizualizacja działania algorytmu, na której widoczny jest szereg parametrów biorących udział w analizie i detekcji anomalii. Wartości obserwowane oznaczono niebieską ciągłą linią, natomiast wynik predykcji sieci neuronowej zdefiniowano jako pomarańczowa linia przerywana. Najważniejszym elementem wykresu są pionowe linie kropkowane, które odpowiadają wykrytym anomalom. Zauważono, że w analizowanym szeregu występuje duża liczba pojedynczych próbek błędnie uznanych za podejrzane, które mogą odpowiadać za generowanie fałszywych alarmów, jednak przyjęcie odpowiedniego poziomu czułości tego algorytmu nie musi powodować wzbudzenia

alarmu. Pionowa niebieska linia kropkowana przedstawia anomalne próbki, które zostały uznane przez algorytm za podejrzane i pozwoliły na skuteczne wykrycie faktycznie symulowanej anomalii, która trwa przez co najmniej 45 minut.

Dokładność algorytmu w zakresie wykrywania anomalii określona została z wykorzystaniem tablicy pomyłek (ang. confusion matrix). Tablica pomyłek, inaczej zwana macierzą błędów, jest najczęściej wykorzystywanym narzędziem do oceny jakości klasyfikacji binarnej. Porównuje ona rzeczywiste wartości docelowe z przewidywanymi przez model i odpowiednio klasyfikuje etykietami na predykowaną klasę pozytywną albo predykowaną klasę negatywną. O skuteczności modelu klasyfikacji decydują dwie wartości

Rys. 4
Skuteczność wykrywania anomalii wraz z fałszywymi alarmami, opracowane na podstawie tablicy pomyłek dla ilości podejrzeń n=2 (a) oraz n=3 (b)



Rys. 3
Przykładowy wykres 15-minutowego natężenia przepływu wody dla budynku wielorodzinnego w trakcie detekcji anomalii

TP (prawdziwie pozytywna – ang. True Positives) i TN (prawdziwie negatywna – ang. True Negative). W macierzy błędów znajdują się także statystyki określane jako fałszywie pozytywna (FP) i negatywna (FN), które określane są odpowiednio za błąd pierwszego i drugiego rodzaju.

Wyniki przeprowadzonej analizy skuteczności algorytmu w zakresie wykrywania anomalii zostały przedstawione na rysunku 4. Prezentowane wyniki uzyskano porównując szeregi czasowe obserwowanych i prognozowanych poborów wody o rozdzielczości czasowej 15 minut i przy przyjętej liczbie dwóch (n=2) i trzech (n=3) kolejno występujących po sobie okresów z podejrzeniem anomalii.

Dla przyjętej liczby podejrzeń anomalii $n=2$ i $n=3$, udało się uzyskać efektywność detekcji wynoszącej kolejno 88-97% oraz 85-90%. Fałszywe alarmy stanowiły 1,3-5,9% dla $n=2$ oraz 0,1-1,7% dla $n=3$. Choć skuteczność wykrywania anomalii dla kolejnych dwóch podejrzeń jest większa, generuje większy odsetek fałszywych alarmów w porównaniu z kolejnymi trzema podejzreniami. Wyniki te odpowiadają jakościowo wynikom prezentowanym przez Fuentes i Mauricio [25]. Badacze stosując metodę k – najbliższego sąsiedztwa kNN (ang. k Nearest Neighbours) oraz twierdzenie Czebyszewa skonstruowali przedział ufności, według którego co najmniej 75% wartości rejestrowanego zużycia wody mieści się w odległości dwóch odchyłeń standardowych od średniej. Wartości zużycia, które nie mieściły się w tym przedziale uznawali za anomalie. W efekcie końcowym uzyskali oni dokładność modeli na poziomie 84-98%. Porównując doniesienia literaturowe, otrzymane w ramach niniejszych badań wyniki detekcji uznano za satysfakcjonujące i możliwe do praktycznego wdrożenia.

Podsumowanie

Od wodomierzy tworzących strukturę monitoringową mądrych czy też inteligentnych sieci wodociągowych i nadzorujących ich systemów informatycznych oczekuje się dokonywania pomiarów z precyzyjnym interwałem pomiarowym, bieżącej interpretacji zarejestrowanych parametrów, umiejętności prognozowania zużycia wody oraz detekcji anomalii. Docelowo, służyć one mają także nadzornemu sterowaniu pracą pompowni wody, tak by zachować ich jak największą wydajność i sprawność przy jednoczesnym obniżeniu energochłonności układu. Gospodarka wodomierzowa nie powinna już być związana tylko z kwestią fakturowania i bilingowania poboru wody, ponieważ istnieje prawny wymóg wyposażania budynków w inteligentne wodomierze od 2027 roku w związku ze zmianą ustawy o efektywności energetycznej [2]. Inteligentne sieci wodociągowe w swych zamierzeniach to te, których zarządzanie jest wspomagane algorytmami uczenia maszynowego na tyle elastycznymi, by adaptować się do dynamicznych warunków przepływu wody w systemach jej dystrybucji z zachowaniem wrażliwości na detekcję stanów anormalnych w czasie pracy rzeczywistym.

Wyniki przeprowadzonych badań wskazują na możliwość implementacji algorytmów głębokich sieci neuronowych

(DNN) w prognozowaniu zapotrzebowania na wodę. Dla budynków mieszkalnictwa wielorodzinnego stosując sieci DNN uzyskano błąd prognozy wynoszący poniżej 3,0%, przy jednoczesnej demonstracji możliwości predykcji zużycia wody na podstawie modelu sieci wytrenowanej na szeregach czasowych rozbiórów wody z innego budynku o podobnym wzorcu dobowego zapotrzebowania na wodę.

W świetle przeprowadzonych badań, optymalna agregacja zużycia wody powinna wynosić 15 minut, ponieważ tak szczegółowe dane pozwalają na osiągnięcie wyraźnie lepszego jakościowo modelu do prognozy poboru wody o błędzie 2,6%, aniżeli 30-minutowa agregacja (błąd 4,7%) czy też 60-minutowa (błąd 9,7%). W związku z tym, jeżeli wodomierze mają docelowo być wykorzystane nie tylko do celów bilingowych, ale także wspomóc na przykład diagnozę stanu instalacji wodnych, powinny umożliwić przeprowadzenie precyzyjnego czasowo pomiaru. Wyniki badań wykazują także, że dobry jakościowo model prognostyczny może być z powodzeniem wdrożony jako podstawa detekcji anomalii o skuteczności jej wykrywania na poziomie 88,8 – 97,3% w zależności od wielkości budynku wielorodzinnego, ale także wielkości samej awarii. Autorzy badań chcieliby jednak zwrócić uwagę na dość szeroko rozumiane i czasami zamiennie stosowane pojęcia awarii, wycieku czy też anomalii. Brak jest w literaturze ściśle określonych definicji tych zjawisk, a tylko poziom wrażliwości modelu klasyfikacyjnego jest pewnym wyznacznikiem, czy aktualnie panujące w sieci warunki są standardowe czy też nie.

Wodomierze zdalne, w większości przedsiębiorstw wodociągowych, powszechnie stosowane są w celach przede wszystkim rozliczania poborów wody oraz bilansowych, do określania wielkości strat wody w wydzielonych strefach DMA. Przeprowadzone przez autorów badania dowiodły, że dzięki zwiększeniu rozdzielczości ich odczytów i w połączeniu z inteligencją obliczeniową mogą znaleźć znacznie szersze zastosowanie, poprawiając bezpieczeństwo użytkowników końcowych wody w kontekście potencjalnych awarii i wycieków z instalacji w budynkach. W celu udoskonalania eksploatacji sieci wodociągowych i przechodzenia od sieci typu mądrych (ang. smart) do sieci inteligentnych (ang. intelligent) wyzwaniem przyszłości nadal pozostaje w pełni zintegrowany system monitoringowy, obejmujący nie tylko strefowe uzdatnianie pomiarowe, ale także dane wodo-

mierzowe, który będzie wspierał procesy decyzyjne, wspomagał automatyzację pracy urządzeń nadrzędnych, takich jak pompownie, dokonywał detekcji i sygnalizował pojawienie się stanów anormalnych w krótkim czasie.

Podziękowanie

Autorzy pracy składają podziękowania Miejskim Wodociągom i Kanalizacji w Bydgoszczy Sp. z o.o. za udostępnienie poligonu badawczego i współpracę. Wyniki badań pozyskano w ramach realizacji projektu POIR.01.01.01-00-0633/21 System ekspertowy rozbiór wody, finansowanego przez NCBiR w ramach Programu Operacyjnego Inteligentny Rozwój 2014-2020; Oś priorytetowa: Wsparcie prowadzenia prac B+R przez przedsiębiorstwa; Działanie: Projekty B+R przedsiębiorstw; Poddziałanie: Badania przemysłowe i prace rozwojowe realizowane przez przedsiębiorstwa.

Temat zaprezentowany podczas II Konferencji Naukowo-Technicznej „Nauka-Technologia-Środowisko” w dniach 27-29 września 2023 r. w Wiśle. Konferencja finansowana przez Ministra Edukacji i Nauki w ramach programu „Dokonała nauka” – moduł „Wsparcie konferencji naukowych” (projekt nr DNK/SP/546599/2022).



LITERATURA

- [1] Rozporządzenie Ministra Infrastruktury z dnia 12 kwietnia 2002 r. w sprawie warunków technicznych, jakim powinny odpowiadać budynki i ich usytuowanie (Dz.U. 2002 nr 75 poz. 690).
- [2] Ustawa z dnia 20 kwietnia 2021 r. o zmianie ustawy o efektywności energetycznej oraz niektórych innych ustaw (Dz.U. 2021 poz. 868).
- [3] G. He, Y. Zhao, J. Wang, H. Li, Y. Zhu i S. Jiang, "The water-energy nexus: energy use for water supply in China". *International Journal of Water Resources Development*, pp. 587-604, 35(4), 2019. <https://doi.org/10.1080/07900627.2018.1469401>
- [4] M. S. Rahim, K. A. Nguyen, R. A. Stewart, D. Giurco i M. Blumenstein, "Machine learning and data analytic techniques in digital water metering: A review", *Water*, 12(1), 294, 2020. <https://doi.org/10.3390/w12010294>
- [5] J. Li, X. Yang i R. Sitzenfrei, "Rethinking the framework of smart water system: A review", *Water*, 12(2), 412, 2020. <https://doi.org/10.3390/w12020412>
- [6] P. Dżimińska, J. Stańczyk, S. Drzewiecki, P. Licznar, "Wykorzystanie systemu rejestracji danych z dużą częstotliwością do analizy nierównomierności zużycia wody", *INSTAL* 1/2022, DOI:10.36119/15.2022.1.3

- [7] A. Cahn, "An overview of smart water networks", *Journal-American Water Works Association*, 106(7), 68-74, 2014. <https://doi.org/10.5942/jawwa.2014.106.0096>.
- [8] J. Wawrzosek, S. Ignaciuk, J. Stańczyk i J. Kajewska-Szkudlarek, "Water Consumption Variability Based on Cumulative Data From Non-simultaneous and Long-term Measurements," *Water Resources Management*, pp. 2799–2812, 35 2021. DOI: 10.1007/s11269-021-02868-6
- [9] S. Hamilton, B. Charalambous i G. Wyeth, "Improving Water Supply Networks: Fit for Purpose Strategies and Technologies", 2021.
- [10] P. Dzięmińska, S. Drzewiecki, M. Ruman, K. Kosek, K. Mikołajewski i P. Licznar, "The Use of Cluster Analysis to Evaluate the Impact of COVID-19 Pandemic on Daily Water Demand Patterns," *Sustainability*, p. 5772, 13(11) 2021. DOI: 10.3390/su13115772.
- [11] J. Cahill, C. Hoolehan, R. Lawson i A. L. Browne, (2022) "COVID-19 and water demand: A review of literature and research evidence" *Wiley Interdisciplinary Reviews: Water*, 9(1), e1570, 2022. <https://doi.org/10.1002/wat2.1570>
- [12] J. E. Pesantez, E. Z. Berglund i N. Kaza, "Smart meters data for modeling and forecasting water demand at the user-level", *Environmental Modelling & Software*, 125, 104633, 2020. <https://doi.org/10.1016/j.envsoft.2020.104633>.
- [13] X. Fan, X. Zhang, i X. B. Yu, "Machine learning model and strategy for fast and accurate detection of leaks in water supply network", *Journal of Infrastructure Preservation and Resilience*, 2(1), 1-21, (2021). <https://doi.org/10.1186/s43065-021-00021-6>.
- [14] S. Spedaletti, M. Rossi, G. Comodi, L. Cioccolanti, D. Salvi i M. Lorenzetti, "Improvement of the energy efficiency in water systems through water losses reduction using the district metered area (DMA) approach", *Sustainable Cities and Society*, 77, 103525, (2022). <https://doi.org/10.1016/j.scs.2021.103525>
- [15] D. T. Larose, "Metody i modele eksploracji danych", Wydawnictwo Naukowe PWN, 2012.
- [16] G. Guo, S. Liu, Y. Wu, J. Li, R. Zhou i X. Zhu, "Short-term water demand forecast based on deep learning method", *Journal of Water Resources Planning and Management*, 144(12), 04018076, 2018. [https://doi.org/10.1061/\(ASCE\)WR.1943-5452.0000992](https://doi.org/10.1061/(ASCE)WR.1943-5452.0000992)
- [17] M. Bakker, H. Van Duist, K. Van Schagen, J. Vreeburg i L. Rietveld, "Improving the performance of water demand forecasting models by using weather input", *Procedia Engineering*, 70, 93-102, 2014. <https://doi.org/10.1016/j.proeng.2014.02.012>
- [18] A. Candelieri i F. Archetti, "Identifying typical urban water demand patterns for a reliable short-term forecasting—the icewater project approach", *Procedia Engineering*, 89, 1004-1012, 2014.
- [19] Y. Xu, J. Zhang, Z. Long, H. Tang i X. Zhang, "Hourly urban water demand forecasting using the continuous deep belief echo state network", *Water*, 11(2), 351, 2019. <https://doi.org/10.3390/w11020351>
- [20] Z. Siwoń, W. Ciezak i J. Ciezak, "Modele neuronowe szeregów czasowych godzinowego poboru wody w osiedlach mieszkaniowych", *Ochrona Środowiska*, 33(2), 23-26, 2011.
- [21] J. W. Tukey, "Exploratory data analysis", Wydawnictwo Addison-Wesley Pub. Co., 1997.
- [22] K. Pałczyński, T. Andrysiak, M. Czyżewska, M. Kierul i T. Kierul, "Neural Network Models for the Prediction of Time Series Representing Water Consumption: A Comparative Study", In *New Advances in Dependability of Networks and Systems: Proceedings of the Seventeenth International Conference on Dependability of Computer Systems DepCoS-RELCOMEX*, June 27–July 1, 2022, Wrocław, Poland, Cham: Springer International Publishing, pp. 186-196, 2022.
- [23] K. Pałczyński, T. Andrysiak, M. Czyżewska, M. Kierul i T. Kierul, "Universality of Forecasting Models on Water Consumption Prediction Tasks", *Proceedings of the 3rd Polish Conference on Artificial Intelligence PP-RAI'2022*, April 25-27, 2022, Gdynia, Poland.
- [24] K. Pałczyński, T. Andrysiak, T., Głowacki, M. Kierul i T. Kierul, "Analysis of Long-Range Forecast Strategies for IoT on Urban Water Consumption Prediction Task", In *International Joint Conference 15th International Conference on Computational Intelligence in Security for Information Systems (CISIS 2022)* 13th International Conference on EUropean Transnational Education (ICEUTE 2022) *Proceedings*. Cham: Springer Nature Switzerland, pp. 3-11, 2022.
- [25] H. Fuentes i D. Mauricio, "Smart water consumption measurement system for houses using IoT and cloud computing", *Environmental Monitoring and Assessment*, 192, 1-16, 2020. <https://doi.org/10.1007/s10661-020-08535-4>