

RECENZJA ROZPRAWY DOKTORSKIEJ

mgr inż. KLARY RAMM SZATKIEWICZ

pt.:

ANALIZA BAZ DANYCH ZAPISÓW WODOMIERZOWY DLA POTRZEB DYNAMICZNEGO MODELOWANIA ROZBIORÓW WODY W WARSZAWIE

1. PODSTAWA RECENZJI

Niniejszą recenzję opracowano na podstawie uchwały Rady Wydziału Inżynierii Środowiska, Geomatyki i Energetyki Politechniki Świętokrzyskiej podjętej w dniu 12.03.2014 r. oraz pisma Pani Dziekan Tego Wydziału dr hab. Lidii Dąbek profesor nadzwyczajnej Politechniki Świętokrzyskiej z dnia 22.05.2014 r. znak ID-510-2/13 informującego o powołaniu mnie na recenzenta rozprawy doktorskiej Pani mgr inż. Klary Ramm Szatkiewicz. Do pisma dołączony został egzemplarz rozprawy doktorskiej.

2. CHARAKTERYSTYKA PRACY ORAZ UWAGI OGÓLNE

Recenzowana praca składa się ze spisu treści, 7 rozdziałów zawierających treści merytoryczne, wykazu cytowanego piśmiennictwa, 7 załączników zawierających graficzne prezentacje wyników badań oraz streszczeń w języku polskim i angielskim. Praca nie zawiera spisu rysunków i tabel, co powszechnie jest stosowane w dysertacjach doktorskich, ale nie ma to wpływu na pozytywny odbiór strony edytorskiej przesłanego do recenzji doktoratu, który został wydany w formie zwartej przez Wydział Inżynierii Środowiska, Geomatyki i Energetyki Politechniki Świętokrzyskiej w serii Rozprawy Doktorskie. Praca liczy 199 stron. W tekście Autorka odwołuje się do 116 pozycji piśmienniczych oraz do 21 pozycji netograficznych podając adresy stron internetowych oraz daty dostępu do nich. W rozdziałach 1-7 znajduje się 35 rysunków i 7 tabel. Pozostałe 163 rysunki umieszczono w załącznikach.

Taka forma redakcji przyczyniła się do zwartości tekstu, przy czym zachowano logiczną, standardową strukturę typową dla prac naukowych, w tym rozpraw doktorskich. Strona formalna pracy nie budzi zastrzeżeń. Praca została napisana zrozumiałym językiem, a tekst zawiera bardzo małą ilość błędów redakcyjnych.

Rozdział 1., zatytułowany *Wprowadzenie*, stanowi rys historyczny rozwoju systemów wodociągowych oraz metod opomiarowania zużycia wody z zastosowaniem wodomierzy i przepływomierzy. Informacje zawarte w tym rozdziale zostały rozszerzone w rozdziale 2. *Przegląd literatury* o stosowane elektroniczne techniki rejestracji rozbiórów wody, w tym o opis urządzeń z automatycznym odczytem wodomierzy dla których stosuje się w piśmiennictwie anglojęzycznym akronim AMR (*Automatic Meter Reading*). W rozdziale 2. przedstawiono również obowiązujące w Polsce podstawy prawne kontroli wielkości zużycia wody przez odbiorców, wymagania ustanowione przez Parlament Europejski i Radę Unii Europejskiej w dniu 31 marca 2004 r. i ogłoszone w Dzienniku Urzędowym UE Nr L135 w dniu 30 kwietnia 2004 r. o przyrządach pomiarowych, zwana potocznie MID (*Measuring Instruments Directive*), stosowane strategie monitoringu zużycia wody na obszarach zurbanizowanych oraz potencjał aplikacyjny baz danych pochodzących z monitoringu zużycia wody dla celów modelowania, z zastosowaniem modeli dynamicznych, pracy sieci wodociągowej na etapie opracowywania strategii ich rozwoju, modernizacji i eksploatacji. Pomiarzy przepływu wody i zmian ciśnień prowadzone *on-line* na sieciach wodociągowych stały się w ostatnich latach standardową techniką do szybkiego wykrywania stanów awaryjnych sieci wodociągowych oraz ustalania rzeczywistych strat wody wodociągowej. Umieszczona w tym rozdziale tabela 2.3 zawiera istotny drobny błąd redakcyjny, gdyż tytuły w pierwszym wierszu zostały przesunięte w stosunku do kolumn pozostałych wierszy tej tabeli.

Doktorantka w podrozdziale 2.5 *Modele komputerowe sieci wodociągowych i ich zasilanie danymi o rozbiórach wody* przedstawiła światowy stan wiedzy nad modelowaniem rozbiórów wody wodociągowej. W podrozdziale tym przedstawiono powszechnie stosowane w Polsce na etapie projektowania sieci i zbiorników wodociągowych empiryczne wzorcowe rozkłady dobowych rozbiórów wody, przedstawiające procentowy udział godzinowego rozbioru wody w rozbiorze dobowym na tle trendów badawczych nad opracowaniem stochastycznych modeli rozbiórów wody wykorzystywanych przy długoterminowych symulacjach pracy sieci wodociągowych. W pracy błędnie empiryczne rozkłady rozbiórów wody nazywa się 1-, 2- i 3-pikowe, w zależności od ilości wartości maksymalnych w rozkładzie, gdyż takie rozkłady w statystyce empirycznej nazywane są rozkładami jedno-, dwu- i trymodalnymi.

Przy tworzeniu stochastycznych modeli zużycia wody konieczna jest znajomość przestrzennego rozmieszczenia punktów poboru wody wraz z określeniem prawdopodobieństwo zainicjowania, czasu trwania i wielkości poboru wody w określonych

przedziałach czasu. Zdarzenia losowe w czasie i przestrzeni łączy się w klastry (ang. *cluster*) rozumiane jako grupy przestrzenne powiązanych ze sobą mniejszych obiektów w jeden większy obiekt. Dla tak zdefiniowanego procesu stochastycznego stosuje się modele impulsu prostokątnego Bartlett-Lewis. Najbardziej rozwinięte są stochastyczne modele prognozowania rozbiórów wody nazywane modelami *use-end*, co można przetłumaczyć jako model końcowego użytkownika. Przy opracowywaniu modeli *use-end* konieczne jest posiadanie szeregów czasowych rozbiórów wody o wysokiej rozdzielczości czasowej, pomierzonych ze stałym krokiem czasowym od 1 sekundy do 1 minuty [Alvisi S., Ansaloni N., Franchini M., 2013: *Generation of synthetic water demand time series at different temporal and spatial aggregation levels*. Urban Water Journal., vol. 10, no. 7, pp. 1-14.]. W modelach tych począwszy od budynków poprzez ulice, osiedla a kończąc na całym systemie wodociągowym tworzy się wzorce rozbiórów wody, które można wykorzystywać przy ustalaniu warunków pracy sieci wodociągowych w stanach nieustalonych. Opis tych modeli znalazł się w recenzowanej pracy, co świadczy o dobrym rozeznaniu literatury przedmiotu badań związanych z modelowaniem stochastycznym rozkładów rozbiórów wody w systemach wodociągowych.

Modelowanie warunków hydraulicznych stanów nieustalonych w sieciach wodociągowych z wykorzystaniem stochastycznych rozkładów rozbiórów wody umożliwia analizę pracy sieci wodociągowych z uwzględnieniem czynników deterministycznych jak i losowych poboru wody przez odbiorców. W podrozdziale 2.6 *Stan badań krajowych nad wzorcami rozbiórów wody* doktorantka wykazała, że w Polsce nie prowadzi się badań nad stochastycznymi modelami rozbiórów wody typu *use-end*.

W rozdziale 3 Autorka dysertacji przedstawiła cele naukowe i aplikacyjne podjętych prac badawczych. Na stronie 42 przedstawiono dwie hipotezy badawcze:

Pierwsza hipoteza zakłada, że wyspecjalizowane narzędzia przetwarzania danych z obrębu (powinno być: zakresu) analizy szeregów czasowych oraz eksploracji danych (ang. data mining), takie jak analiza widmowa i analiza skupień, mogą być bardzo przydatnymi w przetwarzaniu baz danych pochodzących z odczytów zdalnych wodomierzy i identyfikacji wzorców godzinowych rozbiórów wody dla odbiorców referencyjnych.

Druga hipoteza zakłada, że godzinowe wzorce rozbiórów wody, realizowane przez określone rodzaje odbiorców, wykazują duże podobieństwo wzajemne w obrębie homogenicznych grup odbiorców, co uzasadnia w praktyce modelowania komputerowego systemów wodociągowych, stosowanie rozkładów rozbiórów dla tzw. odbiorców referencyjnych.

Przy tak postawionych hipotezach można stwierdzić, że podjęte badania w większym stopniu były ukierunkowane na opracowanie empirycznych rozkładów rozbiórów wody dla poszczególnych klastrów (obiektów referencyjnych) niż opracowanie dla nich modeli stochastycznych rozbioru wody, co potwierdza akapit znajdujący się na stronie 42: *Celem*

praktycznym pracy jest opracowanie wzorców rozbiórów godzinowych dla kilku typów odbiorców referencyjnych.

Podjęcie się badań nad opracowaniem empirycznych modeli rozkładu rozbiórów wody w zależności od zróżnicowania struktury punktów poboru wody wodociągowej w klastrze ma duże znaczenie aplikacyjne i poznawcze. Znajomość empirycznych modeli rozkładów wody dla różnych obiektów referencyjnych pozwala na opracowanie racjonalnych strategii rozwoju sieci wodociągowej na obszarach zurbanizowanych. Jest to szczególnie ważne na etapie opracowywania miejscowych planów zagospodarowania przestrzennego. Znajomość empirycznych lub stochastycznych modeli rozbioru wody dla nowych planowanych obiektów pozwala prognozować skutki rozwoju tkanki miejskiej na hydrauliczne warunki pracy sieci wodociągowej, a pośrednio kanalizacji, po oddaniu nowych obiektów do użytkowania. W Polsce, w odróżnieniu do innych krajów, takich analiz się nie prowadzi głównie ze względu na brak danych wejściowych do programów modelujących pracę sieci wodociągowej. Wykazanie indywidualnych cech rozkładu poboru wody w obiektach o zróżnicowanej strukturze przyborów sanitarnych oraz czasu ich użytkowania stanowi punkt wyjścia do szczegółowych analiz nad stworzeniem stochastycznych modeli rozkładu rozbioru wody klastrów obejmujących wydzielone strefy. Z tego względu poznanie różnic w empirycznych rozkładów rozbioru wody dla różnych obiektów referencyjnych stanowi istotny cel poznawczy recenzowanej pracy, choć Autorka nie wyartykułowała w dostateczny sposób tego celu w tekście dysertacji.

Rozdział 4 zawiera informacje o poligonie badań, badawczych szeregach czasowych rejestrowanych stanów wodomierzowych wykorzystanych w pracy, sposobie ich zbierania oraz zastosowanych metodach wstępnej obróbki danych. Na stronie 43 znajduje się akapit:

Poligonem badawczym pracy jest sieć wodociągowa miasta Warszawy, będąca zarazem największym systemem wodociągowym w Polsce. Na poligonie tym prowadzono rejestracje rozbiórów wody dla uprzednio wytypowanych odbiorców, dzięki instalacji na ich przyłączach wodomierzy ze zdalnym odczytem. W pracy wykorzystuje się materiał pomiarowy zgromadzony zarówno w latach 2008 i 2009, kiedy to przeprowadzono kampanię pomiarową dla wybranych obiektów, na potrzeby kalibracji tworzonego wówczas modelu matematycznego sieci wodociągowej, jak również dodatkowe rejestracje z roku 2012, wykonane specjalnie dla celów niniejszych studiów.

W ramach prac analitycznych opracowano serie pomiarowe godzinnych rozbiórów wody zarejestrowanych w 7 budynkach wielorodzinnych należących do warszawskich spółdzielni mieszkaniowych, 6 kamienicach i blokach mieszkalnych, w których mieściły się sklepy i (lub) punkty usługowe oraz mieszkania użytkowane jako biura, 4 obiektach handlowo-usługowych, w tym małym osiedlowym pawilonie, jak również największym w Warszawie centrum handlowym, 2 biurowcach, 1 szkole podstawowej, 1 ośrodku sportu i rekreacji posiadającym basen kąpielowy oraz 1 domku jednorodzinnym.

W rozdziale 5 przedstawiono opis zastosowanych metod analitycznych. Szeregi czasowe opracowywano z zastosowaniem analizy widmowej a uzyskane empiryczne rozkłady rozbiórów wody były porównywane między sobą z zastosowaniem analizy skupień. W rozdziale tym zabrakło opisu zastosowanych metod określania empirycznych rozkładów rozbiórów wody. Tukey wykazał, że wieloelementowe zbiory danych uzyskane w ramach monitoringu nie można analizować z zastosowaniem tradycyjnych metod statystycznych [Turkey J. W., 1977: *Exploratory data analysis*. Reading, PA: Addison-Wesley]. W takich przypadkach powinno się analizować cechy statystycznych uzyskiwanych zbiorów danych. Powszechnie do tego celu wykorzystuje się wykresy pudełkowe (ang. *box plot*) wykreślane w oparciu o wyznaczone wartości: minimalną, maksymalną, medianę, kwantyl dolny 25% i kwantyl górny 75%. Ten typ wykresów jest przez Autorkę stosowany na rysunkach umieszczonych w załącznikach, lecz w pracy zabrakło informacji, że jest to jedna z powszechnie stosowanych technik *data mining* eksploracji dużych zbiorów danych.

W analizie danych strumieniowych, do których zalicza się wyniki systematycznie wykonywanych pomiarów poboru wody, bardzo często obserwuje się zjawisko zmiany cech statystycznych szeregu czasowego zwane w literaturze przedmiotu dryftem lub zamienne dryfem [Quevedo, J., Pascual, J., Puig, V., Saludes, J., Sarrate, R., Escobet, A., Espin, S., Roquet, J., 2013: *Flowmeter data validation and reconstruction methodology to provide the annual efficiency of a water transport network: the ATLL case study in Catalonia*. Water Science & Technology: Water Supply, (doi:10.2166/ws.2013.203)]. Szkoda, że Autorka nie przedstawiła w pracy przykładowych szeregów czasowych w funkcji czasu próbkowania. Na podstawie analizy długich szeregów czasowych można ustalić pojawianie się mutacji ustalonych dryftu (całkowitej zmiany cech statystycznych zbioru wyników) w odróżnieniu od okresowych fluktuacji (wahań okresowych nie wykazujących trwałych tendencji zmian) zwanych również dryftem cyklicznym. Do ustalania fluktuacji wykorzystuje się analizę widmową i periodogram, co było przedmiotem badań analitycznych przeprowadzonych przez Doktorantkę.

Rozdział 6 prezentuje wyniki badań oraz logiczne wywody przeprowadzone w oparciu o uzyskany uporządkowany materiał analityczny. Autorka dysertacji wydzieliła część graficzną opracowanych zbiorów do załączników, a w rozdziale 6 znalazły się jedynie wybrane rysunki stanowiące element prowadzonych rozważań nad dyskutowanymi wynikami. Przeprowadzono dyskusję wyników analizy widmowej badanych szeregów czasowych oraz uzyskanych empirycznych godzinowych rozkładów rozbiórów wody w różnych obiektach referencyjnych. Wyciągnięte na tej podstawie wnioski, przedstawione w rozdziale 7, mają swoje uzasadnienie w przeprowadzonej przez Doktorantkę dyskusji i nie budzą zastrzeżeń.

3. UWAGI DYSKUSYJNE

Poniżej przedstawiono uwagi dyskusyjne, na które oczekuję odpowiedzi w czasie publicznej obrony rozprawy doktorskiej:

- a) W tabelach 4.1-4.4 przedstawiono charakterystykę obiektów badawczych. W opisach tych zabrakło danych o liczbie poszczególnych typów przyborów sanitarnych: umywalek, zlewów, misek ustępowych itd. W jakim stopniu dane te mogłyby wpłynąć na wnioski końcowe przeprowadzonych badań?
- b) Na rysunku 6.5 przedstawiono dendrogram dla zbioru rozkładów godzinowych wody dla obiektu MW1. Na osi pionowej podano numer porządkowy doby prowadzenia obserwacji. Czy wprowadzenie w miejsce numeru porządkowego doby kodu zawierającego dane: dzień (poniedziałek, wtorek,...), numer tygodnia obserwacji, dzień powszedni (P) lub dzień świąteczny (S) pozwoliłoby lepiej rozpoznać występowanie fluktuacji w analizowanych zbiorach danych? Czy możliwe jest wprowadzenie innych metod kodowania danych przy badaniu fluktuacji i dryftu szeregów czasowych?
- c) Jakie inne metody *data mining* można byłoby zastosować do badania szeregów czasowych danych pochodzących z monitoringu poboru wody?
- d) Czy czas próbkowania danych i klasa wodomierza wpływa na jakość uzyskanych pomiarów i wyniki przeprowadzonych analiz?

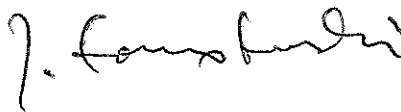
5. WARTOŚCI POZNAWCZE PRACY

Do wartości naukowych i aplikacyjnych pracy zaliczam:

- a) wykazanie celowości zastosowanie metod *data mining* przy opracowywaniu baz danych pochodzących z eksploatacyjnego monitoringu rozbioru wody w celu rozpoznawanie deterministycznej i stochastycznej natury rozbiorów wody;
- b) potwierdzenie, że analiza widmowa szeregów czasowych rozbiorów wody pozwala na wykrywanie fluktuacji w szeregu czasowego pomiarów poboru wody dla jednorodnych i zróżnicowanych obiektów referencyjnych (klastrow punktów poboru wody wodociągowej);
- c) wykazanie, że analiza skupień pozwala na rozpoznawanie podobieństw empirycznych rozkładów czasowych poboru wody przez różnych odbiorców wody;
- d) wykazanie, że przy opracowywaniu modeli empirycznych i probabilistycznych rozkładów poboru wody zachodzi konieczność oddzielnego opracowywania danych uzyskanych z pomiarów w dni powszednie i dni wolne od pracy, gdyż szeregi czasowe dla tych przedziałów czasowych różnią się statystyczną i probabilistyczną strukturą danych.

6. WNIOSEK KOŃCOWY

Zawarte w recenzji uwagi nie wpływają na pozytywną ocenę recenzowanej dysertacji. Rozprawa mgr inż. Klary Ramm Szatkiewicz pt.: *Analiza baz danych zapisów wodomierzy dla potrzeb dynamicznego modelowania rozbiorów w Warszawie* zawiera oryginalny materiał eksperymentalny o dużych walorach poznawczych i aplikacyjnych. Autorka wykazała się znajomością metod *data mining* stosowanych w eksploracji dużych zbiorów danych, umiejętnością prowadzenia kameralnych prac badawczych oraz formułowania wniosków i uogólnień na podstawie analizy danych pochodzących z monitoringu rozbiorów wody w dużym systemie wodociągowym. W mojej opinii praca przedłożona do recenzji w pełni odpowiada wymaganiom stawianym rozprawom doktorskim określonym w art.13 ust.1 *Ustawy o stopniach naukowych i tytule naukowym oraz stopniach i tytule w zakresie sztuki* z dnia 14 marca 2003 r. (Dz. U. Nr 65, poz. 595, z późn. zm.) i na tej podstawie wnoszę o dopuszczenie mgr inż. Klary Ramm Szatkiewicz do publicznej obrony przed Radą Wydziału Inżynierii Środowiska, Geomatyki i Energetyki Politechniki Świętokrzyskiej w Kielcach.



Wrocław, czerwiec 2014 r.