

Dęblin, 21 września 2019 r.

dr hab. inż. Aneta Krzyżak, prof. LAW
Katedra Płatowca i Silnika
Wydział Lotnictwa
Lotnicza Akademia Wojskowa

Recenzja rozprawy doktorskiej

mgra inż. Przemysława Palasza pt.: „Metoda analizy niezawodności mierników energii cieplnej w zależności od warunków pracy”

1. Informacje ogólne i struktura pracy

Praca została napisana pod kierunkiem promotora prof. dra hab. inż. Józefa Żurka oraz promotora pomocniczego dra inż. Radosława Przysowy.

Praca została napisana na 120 stronach i zawiera 6 rozdziałów z opisem zagadnienia. Ponadto w pracy znajdują się: streszczenie w języku polskim, spis treści, spis ważniejszych oznaczeń i symboli użytych w pracy oraz spis literatury. W pracy znajduje się spis rysunków (39 rysunków), a także spis tabel (13 tabel). W treści pracy zawarte są liczne wzory matematyczne.

Po wprowadzeniu, w którym krótko zarysowano genezę badań omówionych w pracy oraz zawarto bardzo zwięzły opis o stanie literatury światowej, w rozdziale pierwszym opisano sposób pozyskiwania i przygotowania danych źródłowych. W rozdziale drugim zaprezentowano wyniki statystyki opisowej wybranych czynników z bazy danych. W kolejnym rozdziale omówione zostały wyniki analizy prawdopodobieństwa wystąpienia awarii oraz analizy przeżycia. Rozdział piąty dotyczy zastosowania metody nauczania maszynowego w celu opracowania algorytmu prognozowania czasu i liczności wystąpienia awarii mierników ciepła. Ostatni rozdział to podsumowanie.

2. Tematyka rozprawy pracy i zakres badań

Tematyka rozprawy wpisuje się w światowe trendy działań podejmowanych na rzecz zrównoważonego zarządzania energią w ujęciu całościowym z uwzględnieniem energii cieplnej. Jest to bardzo obszerne zagadnienie obejmujące zarówno procesy wytwarzania energii m.in. pod względem efektywności, sposoby jej wykorzystywania, poprzez aspekty techniczne

urządzeń przekazujących i odbiorczych, jak i obsługę całego procesu zarządzania np. z zastosowaniem „systemów inteligentnych” opartych na automatyce i informatyce. Jednym z ogniw tego rozbudowanego łańcucha są mierniki energii ciepła.

W rozdziale pierwszym pracy doktorskiej rozwinięte zostały niektóre aspekty powyższego zagadnienia. Doktorant przedstawił koncepcję Inteligentnych miast, opisał rolę bezprzewodowej sieci czujników, jak również samych inteligentnych mierników (budowa, zasada działania, ocena stanu technicznego, kalibracja, wymiana, najczęstsze awarie) będących obiektem badań w pracy doktorskiej. Wymienił również akty prawne, które mają wpływ na zarządzanie siecią mierników energii elektrycznej. Ponadto w tym samym rozdziale przedstawiona została hipoteza pracy o charakterze bardzo ogólnym, że w zasadzie użycie słowa „teza” nie stanowi błędu. Brzmi ona: „Istnieje możliwość oceny niezawodności i prognozowania zużycia liczników ciepła przy pomocy wybranych modeli i narzędzi, z uwzględnieniem warunków pracy”. Jako najważniejsze cele pracy, cele naukowe, wskazano wykonanie analizy niezawodności oraz „opracowanie metody predykcji awarii mierników ciepła z wykorzystaniem metod stochastycznych i nauczania maszynowego jak również optymalizacja tej metody”. Dalej wymienione zostały cele szczegółowe w zasadzie będące rozszerzonymi celami naukowymi.

Kolejny drugi rozdział zawiera opis sposobu pozyskania bazy danych oraz jej charakterystykę. Wymienione i omówione zostały parametry (cechy mierników), które zdaniem doktoranta są ważne z punktu widzenia eksploatacji mierników. Baza danych oparta jest na przeszło pięćdziesięciu tysiącach mierników i obejmuje informacje za dziewięć lat eksploatacji, co dało liczbę 367 000 rekordów. Przedstawiono sposób przygotowania bazy danych do dalszej analizy, poprzez eliminację danych o charakterze błędów przypadkowych, a także standaryzację danych. Wspomniano, że okres pomiędzy kolejnymi wymianami mierników w przypadku ich bezawaryjnego działania wynosi 10 lat.

Wyniki statystyki opisowej danych zawarto w rozdziale trzecim. Wskazują one na dalsze odrzucanie hipotez statystycznych opartych na założeniach dotyczących rozkładu normalnego. Wykonana została również analiza wzajemnego wpływu czynników poprzez zastosowanie macierzy korelacji. Doktorant stwierdził, że wiele z czynników nie wykazuje korelacji a zatem, że można je uznać za liniowo niezależne. Dodatkowo wykonano eksploracyjną analizę danych, której wyniki zaprezentowano za pomocą wykresu punktowego wskazującego na zależność awarii od skumulowanego zużycia ciepła i rodzaju komunikacji z urządzeniem przez firmę obsługującą. Zaprezentowano również gęstość rozkładu kilku wybranych parametrów za pomocą tzw. wykresów skrzypcowych. Statystyczna analiza danych wskazała słusność

kroków podejmowanych w celu rezygnowania z mierników ciepła, z których odczyt danych wykonywany jest ręcznie. Ponadto wyniki analizy posłużyły doktorantowi do zastanowienia się nad opcją wymiany mierników zgodnie ze stanem technicznym, co wymagałoby dalszych analiz.

W dalszej kolejności (rozdział czwarty) doktorant podjął się przeprowadzenia analizy niezawodności mierników uwzględniając parametr o nazwie „skumulowane zużycie”. Analiza niezawodności została wykonana w oparciu o rozkład Weibull’a. Zaprezentowano wykresy częstotliwości, gęstości oraz prawdopodobieństwa wystąpienia określonej wartości skumulowanego zużycia ciepła dla całej populacji mierników oraz dla mierników, które uległy awarii. Dalej opisano proces Markowa określając prawdopodobieństwo przejścia pomiędzy kolejnymi odczytami wartości zużycia ciepła. Wykonano analizę z podziałem na sposób odczytywania danych z mierników na ręczny lub zdalny, a następnie porównano wyniki wykazując, że liczniki ze zdalną komunikacją wykazywały mniejszą awaryjność niż te z koniecznością odczytu manualnego. Metodą Kaplana-Meiera oszacowano prawdopodobieństwo „przeżycia” mierników w okresie ich obserwacji.

Rozdział piaty obejmuje opis kolejnych działań podjętych w celu opracowania metody przewidywania wystąpienia awarii mierników w kolejnych okresach. Doktorant wykorzystał trzy algorytmy: las losowy, maszyna wektorów nośnych o sztuczna sieć neuronowa. W celu określenia przydatności tych metod zastosowano dwie metryki. Jedna mierząca zdolność algorytmu do odróżniania klas tj. pozytywnej (brak awarii) i negatywnej (awaria) od siebie (AUC – Area Under a Curve) a druga oceniająca algorytm w obu klasach jednocześnie (MCC – Matthews Correlation Coefficient). Otrzymano przewidywalność awarii na poziomie 90%. Dalsze postępowanie związane z wykonaniem optymalizacji algorytmów z użyciem tzw. hyperparametrów, wykazało podwyższenie przewidywalności do poziomu przekraczającego 96%. W końcowej części rozdziału znajduje się opis wykorzystania algorytmu przewidywania wystąpienia awarii do ekonomicznej analizy przeprowadzania prewencyjnych wymian mierników.

Ostatni szósty rozdział to podsumowanie zawierające stwierdzenie, że cele zostały osiągnięte, ocenę rozwiązania omawianego w pracy doktorskiej, wnioski oraz kierunki dalszych badań. Doktorant, w związku z obserwowalnym wzrostem liczby urządzeń chłodzących, wskazuje na zasadność podjęcia podobnych działań w stosunku do mierników chłodu. Podobnie doktorant widzi zastosowanie opracowanych algorytmów przewidywania wystąpienia awarii do liczników ciepłej i zimnej wody, czy też liczników energii elektrycznej.

3. Uwagi i pytania

Poniżej zamieszczam wybrane uwagi i pytania, które pojawiły się w trakcie analizy recenzowanej rozprawy. Są one zapisane w kolejności pojawiania się ich podczas analizy rozprawy. Pominięto drobne uwagi związane z niespójnością niektórych zdań, czy edycji tekstu.

1. Hipoteza, główne cele naukowe oraz szczegółowe cele pracy są bardzo podobne w brzmieniu merytorycznym i wykazują niewielkie uzasadnienie podjęcia stosownych działań dowodowych wobec hipotezy.
2. Czy średnia wartość bieżącego odczytu (75105,08) przyjętego na etapie odrzucania wyników odstających (rozdział 2.4 str. 26) zawierała wartości odstające? Jeśli tak, to może należało przyjąć inny próg odcięcia, nieco niższy. Ponadto nie jest wyjaśnione w jaki sposób ustalono próg odcięcia danych.
3. W rozdziale 3.2.1 rysunki wyprzedzają opis w tekście, co utrudnia analizę treści.
4. Przy omawianiu wyników badań doktorant stosował formułę gramatyczną opartą na czasownikach w pierwszej osobie liczby mnogiej „widzimy”, „policzymy” itp. Z uwagi na przyjętą zasadę, że praca doktorska jest samodzielną pracą naukową, takie sformułowania raczej nie budują tej świadomości u czytającego.
5. Doktorant z bardzo dużej bazy danych wyodrębnił 16 parametrów, o których w rozdziale 2.3 doktorant pisał, jako o parametrach poddanych analizie. W tym miejscu zostały one dość dobrze scharakteryzowane. Niestety w dalszej części pracy, oprócz dwóch miejsc w tabelach z metryką Fishera i jednego w macierzy korelacji, nie omówiono wpływu niektórych czynników na wyniki analizy niezawodności. Dotyczy to takich parametrów jak: piętro, typ powierzchni użytkowej, typ pomieszczenia, czy współczynnik wyrównawczy. Czy wobec tego była konieczna prezentacja tych parametrów, skoro nie wykazano ich w analizie i nie służyły ani w algorytmie prognozowania awarii ani w podbudowie wniosków?
6. Mimo iż doktorant uprzedzająco napisał „mimo, że taka miejscowość może w ogóle nie istnieć”, to jednak wykonał analizę statystyczną i analizę niezawodności na średniej kodów pocztowych. O ile można niestandardowo podejść do geografii w statystyce opisowej, o tyle opieranie wnioskowania na nieistniejących miejscach geograficznych budzi moje wątpliwości. O wiele więcej korzyści, aczkolwiek wymagałoby to większych nakładów pracy, można byłoby wynieść opierając analizę niezawodności o mapę i lokowanie wyników w konkretnych miejscach, bez sformułowań typu „gdzieś pomiędzy”, co z uwagi

na możliwość bezpośredniego sąsiedzowania rejonów o kodach stosunkowo od siebie odległych wartościami, może okazać się chybione.

7. Przy omawianiu macierzy korelacji (rozdział 3.1), jak również w kilku miejscach dalej, doktorant używa pojęcia istotności bez przywołania wyników takiej analizy, co osłabia siłę i wydźwięk zarówno samego pojęcia, jak i umiejętności posługiwania się wynikami analizy istotności czynników zmiennych.
8. W tym samym miejscu (rozdział 3.1) zostały użyte określenia „relatywnie mała” (chodzi o wartość współczynnika korelacji $r = 0,26$, którą można, jak autor pisze, pomijać), „niewielka korelacja” (wartość $r = 0,39$) oraz „bardzo silna korelacja” przy wartości 0,99, co z kolei również uprawnia doktoranta do odrzucenia czynnika w dalszej analizie. Należy zatem doprecyzować, jakie wartości współczynnika korelacji i na jakiej podstawie uznano za istotne. Jest to ważne z punktu widzenia planowania eksperymentu, gdzie jednym z etapów jest tzw. eliminacja szumów, tj czynników nieistotnych, których pomijanie w dalszej analizie nie będzie generowało większej niepewności. Pomimo zauważenia braku liniowej korelacji, nie wpisano jednoznacznego stwierdzenia o wyłączeniu niektórych czynników z dalszej analizy. Jak dalej zauważyłam, te czynniki, pomimo odrzucenia, jednak się dalej w treści pojawiały.
9. W odniesieniu do omówienia danych zawartych na rysunku 3.2 oraz w tabeli 3.2 nasunęły się dwa pytania. Pierwsze: skoro w rozdziale 2.4 str. 26 była mowa o odseparowywaniu punktów odstających, to dlaczego zostały one uwzględnione w dalszej analizie? Drugie: jeśli posiłkując się skalą z rysunku, odejmiemy liczbę 120 liczników od ogólnej liczby liczników bez komunikacji, to okazuje się, że stanowią one tylko około 0,4% wszystkich liczników, zatem skąd pojawił się wniosek, że mają one „duży wpływ na uszkodzenia na poziomie 30%”?
10. W rozdziale 4.1 wyodrębniono podrozdział 4.1.1., co zapewne wynika z pomyłki, ponieważ nie ma podrozdziału 4.1.2.
11. Przeprowadzenie analizy niezawodności z wykorzystaniem rozkładu Weibull’a prowadzi doktoranta do wniosku, że „na przestrzeni 10 lat, intensywność uszkodzeń liczników ciepła jest słabo zależna od czasu użytkowania” (rozdział 4.1. str. 48). Czy zatem nie byłoby właściwe odniesienie się na wyznaczonej funkcji prawdopodobieństwa rozkładu dla skumulowanego zużycia ciepła do (stosowanego w analizie niezawodności z użyciem rozkładu Weibull’a) skumulowanego prawdopodobieństwa wystąpienia uszkodzenia na poziomie $p = 1 - 1/e$, co w przybliżeniu daje wartość 0,632? Oszacowanie wartości skumulowanego zużycia, które odpowiada wartości $p = 0,632$, w dalszym odniesieniu do

czasu wystąpienia awarii, może wskazać, że wartość skumulowanego zużycia ciepła na poziomie blisko 40 000 kWh (patrz rys. 4.7.) odpowiada okresowi sześciu lub siedmiu lat (patrz wartości średniej i mediany tab. 3.1.). Zatem przytoczony wniosek mógłby być niezasadny. Niestety nie ma w pracy bardziej szczegółowego przedstawienia samych danych, chociażby w postaci graficznej np. zależności skumulowanego zużycia ciepła w momencie wystąpienia awarii od czasu jaki upłynął od chwili wprowadzenia miernika do użytkowania do czasu wystąpienia awarii, bym mogła ocenić to sama. W tym kontekście, jak również mając na względzie, że zużywanie ciepła przez różnych odbiorców nie zawsze jest analogiczne w czasie, stwierdzenie (str. 44), że „Skumulowane zużycie ciepła jest cechą miernika, która może być traktowana jako zegar eksploatacyjny czyli ekwiwalent czasu pracy” mogło być przedwczesne. Pewne zaskoczenie nad prawdziwością tego stwierdzenia wystąpiło już wcześniej (str.36), gdzie doktorant napisał „liczba awarii jest słabo zależna od czasu eksploatacji licznika”, co de facto odnosiło się nie do czasu a do skumulowanego zużycia ciepła (rys. 3.2). Podobnie opisano analizę przeżycia metodą Kaplana-Meiera (str. 61-62). Czy zatem założenie o tożsamości czasu eksploatacji ze skumulowanym zużyciem ciepła doktorant wcześniej weryfikował analitycznie? Proszę o wskazanie dowodu na potwierdzenie tego założenia, tym bardziej, że prawdopodobieństwa przejść między kolejnymi stanami uzyskane w procesie Markowa (odniesienie do kolejnego odczytu wykonywanego co rok) pod koniec dziesięcioletniego okresu obserwacji wskazują raczej na duży spadek tego prawdopodobieństwa, blisko 25%.

12. Rozdział 5.3.1 jest zatytułowany „Wybór parametrów i redukcja wymiarów”. O ile w rozdziale tym jest mowa o konkretnym wyliczeniu i wyborze parametrów o „istotnym” znaczeniu, chociaż w efekcie końcowym nie odrzucono żadnego parametru, o tyle redukcja wymiarów jest opisana tylko teoretycznie w oparciu o bibliografię. Metody tej nie zastosowano do omawianej bazy danych.

4. Konkluzja

Ocena ogólna pracy jest pozytywna. Praca wpisuje się w dyscyplinę inżynieria mechaniczna. Praca zawiera oryginalne badania o charakterze inżynierskim oparte na numerycznej analizie bazy danych eksploatacyjnych oraz występowania awarii różnego typu liczników ciepła montowanych w domach i mieszkaniach w Szwajcarii. W pracy opisano autorską metodę prognozowania wystąpienia uszkodzenia liczników ciepła. W oparciu o tę metodę można określić wystąpienie awarii w odniesieniu do skumulowanego zużycia ciepła, co w mojej ocenie stanowi wartościową część pracy. Algorytm prognozujący jest oparty

o metody samouczące się i jest podatny na uzupełnianie i rozbudowywanie wejściowej bazy danych, co może poprawić, i tak już bardzo wysoką, skuteczność prognozowania. Wyniki podjętych analiz i opracowany model przewidywania awarii z powodzeniem mogą być, i zapewne będą, zastosowane w sferze gospodarczej w usprawnieniu metod nadzorowania procesu eksploatacji liczników ciepła, co może przełożyć się na wskaźniki ekonomiczne firm zajmujących się obsługą tego typu liczników. Cele stawiane w pracy zostały osiągnięte, czym pozytywnie potwierdzono stawianą hipotezę.

Uważam, że recenzowana rozprawa doktorska mgra inż. Przemysława Pałasza pt.: „Metoda analizy niezawodności energii cieplnej w zależności od warunków pracy” spełnia warunki stawiane rozprawom doktorskim i po spełnieniu innych warunków formalnych wnoszę o jej publiczną obronę.

Podstawa opracowania recenzji:

- zlecenie opracowania recenzji rozprawy doktorskiej przez Zastępcę Przewodniczącego Rady Naukowej Instytutu Technicznego Wojsk Lotniczych prof. dr hab. inż. Józefa Żurka (pismo z dn. 12.07.2019 r., nr wych. 1888/19),
- ustawa Prawo o szkolnictwie wyższym i nauce (Dz. U. 2018 poz. 1668 z późn. zm.),
- ustawa Przepisy wprowadzające ustawę - Prawo o szkolnictwie wyższym i nauce (Dz. U. 2018 poz. 1669).

