



Recenzja rozprawy doktorskiej mgr inż. Pauliny Muchel zatytułowanej
„Modelowanie zmienności składu gazu ziemnego metodą sztucznych sieci
neuronowych”

zrealizowanej w Katedrze Inżynierii Chemicznej i Procesowej Wydziału Technologii Chemicznej
i Inżynierii Chemicznej Zachodniopomorskiego Uniwersytetu Technologicznego w Szczecinie
pod kierunkiem dr hab. inż. Jolanty Szoplik, prof. ZUT

Aktualność tematyki i cel pracy

Przedstawiona do recenzji rozprawa doktorska dotyczy badań nad wykorzystaniem sztucznej inteligencji do przewidywania składu gazu ziemnego transportowanego rurociągami, a pochodzącego z różnych źródeł. Konieczność dywersyfikacji źródeł energii, w tym gazu ziemnego, jest tematem poruszonym w Polsce od ponad dwóch dekad. Podjęte prace, w tym budowa terminala LNG w Świnoujściu oraz systemu gazociągów Baltic Pipe, miały zmniejszyć import gazu z dotychczasowych źródeł, a w wyniku zmian politycznych zostały wykorzystane do pełnego wyeliminowania importu gazu ziemnego z kierunku wschodniego. Skład gazu ziemnego z poszczególnych źródeł się różni, a jego mieszanie następuje w rurociągach. Oprócz zmiany źródeł gazu ziemnego w polskim systemie gazowniczym jest już obecny biometan, którego udział zgodnie z polityką zielonego ładu Unii Europejskiej ma wzrastać. W przyszłości do gazociągów może być również wtłaczany zielony wodór. Ze względu na zróżnicowanie źródeł gazu wprowadzanego do gazociągów konieczne stało się opracowanie metod pozwalających przewidzieć jego skład u odbiorcy końcowego, co przekłada się na jego kaloryczność, a ostatecznie na ilość koniecznego do dostarczenia gazu. Celem badań w tym zakresie realizowanych w ramach rozprawy doktorskiej Pani mgr inż. Pauliny Muchel jest opracowanie modelu sztucznej sieci neuronowej, która może być wykorzystana do przewidywania składu gazu w określonym miejscu sieci gazowniczej na terenie Polski. Podjętą przez Panią mgr inż. Paulinę Muchel tematykę badawczą i założony przez nią cel badań uważam za aktualne i mające duże znaczenie praktyczne.

Ocena struktury pracy

Rozprawa doktorska została przedstawiona na 188 stronach maszynopisu. Ma ona klasyczny układ, a w jej skład wchodzi kolejno: streszczenie w języku polskim i angielskim, spis oznaczeń, wprowadzenie i uzasadnienie wyboru tematyki pracy, przegląd literatury, opis problemu badawczego i hipoteza badawcza, cel i zakres pracy, opis metodyki badań oraz opis wyników badań. Praca kończy się podsumowaniem, wnioskami oraz spisem literatury. Rozprawa jest udokumentowana w postaci 46 rysunków i 16 tabel zamieszczonych w tekście głównym. Piśmiennictwo cytowane stanowi 200 właściwie dobranych pozycji literaturowych opublikowanych w latach 2001–2024, przy czym zdecydowana większość została wydana po roku 2019. Pozycje [78]–[81] mają niewystarczający opis bibliograficzny, brakuje informacji, skąd pochodzą wymienione karty oraz z którego są roku. Praca jest napisana poprawnie językowo, choć występują stosunkowo nieliczne błędy stylistyczne, interpunkcyjne i drobne błędy edytorskie. Zdecydowanie nadużywane jest słowo „natomiast” stosowane na początku zdania, nie we wszystkich przypadkach właściwie.

Ocena części literaturowej

Studia literaturowe zostały przedstawiona w rozdziale 2, który liczy 65 stron. W podrozdziale 2.1 omówiono charakterystykę polskiego systemu przesyłowego gazu ze szczególnym uwzględnieniem jego źródeł wraz ze składem gazu ziemnego, alternatywne paliwa gazowe, parametry charakteryzujące jakość paliwa gazowego oraz zmienność obciążenia rurociągów sieci gazowej w czasie. Podrozdział 2.2 zawiera podstawowe informacje dotyczące modelowania systemów i analizy danych, natomiast podrozdział 2.3 – podstawy teorii sztucznych sieci neuronowych oraz przykłady ich wykorzystania. Za szczególnie interesujące uważam podrozdziały 2.1.1, 2.1.2 i 2.3.4, w których przedstawiono charakterystykę polskiej sieci przesyłowej gazu ziemnego, alternatywne paliwa gazowe oraz przykłady wykorzystania sztucznych sieci neuronowych. Zasadniczo ta część pracy została dobrze skonstruowana, choć mam tu dwie uwagi. W rozdziale 2 zawarte są informacje pozyskane przez Autorkę z literatury przedmiotu. W manuskrypcie można natomiast znaleźć duże fragmenty tekstu bez odpowiednich odnośników literaturowych (na przykład: str. 16, od zdania „Pozostałą część...” do końca akapitu, przy czym jest odnośnik literaturowy przy rysunku mapy, do której odnosi się opis, str. 38, dolny akapit). Zabrakło również nieco szerszego omówienia opracowanych do tej pory modeli symulacyjnych umożliwiających śledzenie składu gazu (podstawowe informacje na ten temat są podane na str. 35–38).

Ocena przeprowadzonych badań i uzyskanych wyników

Metodyka badań wstępnych i zasadniczych została opisana w podrozdziale 5.1. W badaniach wykorzystano dane historyczne określające zmianę udziałów pięciu składników gazu ziemnego, właściwości fizykochemicznych i strumienia objętościowego gazu od 1 stycznia 2017 roku do 31 grudnia 2021 roku, które pozyskano od Operatora Gazociągów Przesyłowych GAZ-SYSTEM SA. Oprócz szczegółowego opisu w podrozdziale 5.1 zamieszczono schemat metodyki przeprowadzonych badań (rys. 5.1), który w sposób bardzo przejrzysty zapoznaje czytelnika ze strukturą podejmowanych przez Doktorantkę działań badawczych.

Opierając się na badaniach wstępnych, Doktorantka określiła dane wejściowe i wyjściowe wykorzystane do budowy modelu sztucznej sieci neuronowej (podrozdział 5.2). W tym celu posłużyła się danymi o składzie gazu ziemnego w siedmiu wybranych stacjach gazowych w pobliżu punktów transgranicznych (Tietierowka, Kondratki, Wysokoje, Drozdowicze, Cieszyn i Lasów oraz terminal LNG w Świnoujściu). Na podstawie przeprowadzonej analizy składu gazu w podanych punktach Doktorantka ustaliła, że najbardziej odmienny jest skład regazyfikowanego gazu pochodzącego z terminala LNG. Z tych powodów do dalszej analizy wybrała wyniki pomiarów składu gazu w sześciu powiatach usytuowanych w północno-zachodniej Polsce (powiaty choszczeński, czarnkowsko-trzcieński, gorzowski, grodziski, gryficki oraz wejherowski), na które wpływ terminala LNG w Świnoujściu jest największy. Do statystycznej analizy danych wybrała powiat czarnkowsko-trzcieński. Opierając się na tekście rozprawy doktorskiej, czytelnik może się domyślić, dlaczego akurat ten powiat został wybrany, jednak uważam, że lepiej byłoby bardziej szczegółowo wyjaśnić ten problem w tekście.

Na podstawie przeprowadzonej analizy zmienności poszczególnych składników gazu Doktorantka ustaliła, że najbardziej charakterystycznymi składnikami, dzięki którym można rozpoznać rodzaj gazu transportowanego siecią, są metan, etan, propan, azot i ditlenek węgla. Do oceny stopnia zmienności zawartości danych składników w gazie ziemnym Autorka zastosowała metody opisu statystycznego obejmujące interpretację statystyk opisowych i histogramów, natomiast zmienne wejściowe i wyjściowe modelu prognostycznego dobrała metodą korelacji. Doktorantka jako dane wejściowe dobrała miesiąc, dzień miesiąca, dzień tygodnia oraz temperaturę powietrza. Na ich podstawie prognozowany był udział pięciu składników gazu ziemnego (metanu, etanu, propanu, azotu i ditlenku węgla).

W mojej ocenie badania wstępne zostały dobrze przeprowadzone i w sposób przejrzysty opisane, choć Doktorantka nie ustrzegła się drobnych błędów. W tabeli 5.2 nie została podana

jednostka udziałów poszczególnych składników. Na stronie 96 znajduje się zdanie: „W oparciu o dane przedstawione w Tabeli 5-2 oraz wyniki badań wpływu terminala LNG na jakość paliwa w polskiej sieci gazowej uzyskane przez Kumor i Porada [93] (przedstawione w części literaturowej pracy), przeprowadzono szczegółową analizę składu gazu ziemnego...”. Doktorantka mogłaby uściślić, na czym polegała szczegółowa analiza składu gazu ziemnego. Na stronie 103 pojawiło się zdanie: „Z analizy danych zamieszczonych na wykresach wynika, że w miesiącach letnich udział...”. Jak rozpoznać na zamieszczonych wykresach miesiące letnie? W zdaniu na stronie 108 dwukrotnie wymieniono dane wejściowe: „Natomiast w celu wyboru zmiennych wejściowych i wyjściowych modelu prognostycznego metodą korelacji analizowano statystyczne zależności między poszczególnymi składnikami gazu...”. Na stronie 113 dwukrotnie powtórzono wyraz „są”: „Widoczne są także słabe korelacje są między...”.

W ramach badań zasadniczych Autorka opracowała model perceptronu wielowarstwowego (MLP) w programie *STATISTICA Sieci Neuronowe* służący do prognozowania głównych składników gazu ziemnego. Model ten składa się z jednej warstwy ukrytej, w której przepływ sygnałów odbywa się od warstwy wejściowej do warstwy wyjściowej.

W celu dobrania najlepszych parametrów modelu Doktorantka przeprowadziła badania na czterech etapach, w których analizowała wpływ na wyniki prognoz składu gazu ziemnego liczności zbiorów danych użytych do trenowania modelu (etap I), liczby i typu funkcji użytej do aktywacji neuronów, metody uczenia modelu (etap II) oraz struktury modelu (etap III). Dodatkowo przeprowadziła badania nad możliwością wykorzystania opracowanego modelu MLP do prognozowania mieszaniny gazu ziemnego i wodoru (etap IV).

Do trenowania modelu Pani mgr inż. Paulina Muchel wykorzystała cztery zbiory danych o liczebności od 8760 do 35064 (godzinowe dane określające skład gazu ziemnego oraz temperaturę otoczenia z lat 2017–2020), stosując od 30 do 230 neuronów w warstwie ukrytej. Proces uczenia modelu MLP przeprowadzono metodą zmiennej metryki Broydena-Fletcher-Goldfarba-Shanno (BFGS), przyjmując funkcję liniową, logistyczną lub tangens hiperboliczny do aktywacji neuronów w warstwie ukrytej i wyjściowej. Na podstawie przeprowadzonych obliczeń Doktorantka ustaliła, że najlepsze dopasowanie danych eksperymentalnych do prognozy uzyskano, stosując do trenowania modelu zbiór o najmniejszej liczbie danych (8760), o czym świadczą największa wartość współczynników korelacji dla procesu trenowania oraz poprawne odwzorowanie zależności pomiędzy udziałami prognozowanych i eksperymentalnych składników mieszaniny gazu ziemnego. Najmniejsze wartości średniego znormalizowanego pierwiastka błędu średniokwadratowego

(nRMSEśr) Doktorantka uzyskała dla liczby neuronów w warstwie ukrytej od 35 do 45. Wykazała również, że najlepszą metodę trenowania BFGS uzyskuje się przy przyjętej logistycznej funkcji aktywacji neuronów w warstwie wyjściowej i warstwie ukrytej. Ustaliła też, że model MLP z pięcioma neuronami w warstwie wyjściowej pozwala uzyskać wynik prognozy składu gazu z mniejszym znormalizowanym pierwiastkiem błędu średniokwadratowego niż pięć modeli MLP z jednym neuronem wyjściowymi.

Trzy pierwsze etapy badań przeprowadzone przez Doktorantkę pozwoliły jej ustalić, że do prognozowania podstawowych składników gazu ziemnego (metan, etan, propan, azot i ditlenek węgla) przy założonym typie dnia tygodnia (roboczy, świąteczny, weekendowy), miesiącu, dniu miesiąca i temperaturze należy zastosować 37 neuronów w warstwie ukrytej, przeprowadzając proces uczenia metodą zmiennej metryki Broydena-Fletcher-Goldfarb-Shanno (BFGS), stosując logistyczną funkcję aktywacji neuronów.

Opracowany model MLP do prognozowania składu gazu ziemnego uważam za największe osiągnięcie naukowe rozprawy doktorskiej Pani mgr inż. Pauliny Muchel. Jak wykazała Doktorantka, opracowany model MLP może być również z powodzeniem stosowany w prognozowaniu mieszaniny gazu ziemnego i wodoru po rozszerzeniu warstwy wyjściowej o jeden neuron (dodatkowa dana wyjściowa) oraz po przeprowadzeniu ponownego wytrenowania modelu przy użyciu składu mieszaniny gazu, która ma być prognozowana. Opracowany przez Panią mgr inż. Paulinę Muchel model MLP może być wykorzystywany do przewidywania ciepła spalania gazu ziemnego, co może być pomocne przy bilansowaniu systemu gazowego.

Moja ocena przeprowadzonych przez Panią mgr inż. Paulinę Muchel badań jest pozytywna, jednak podczas czytania manuskryptu nasunęło mi się kilka pytań i uwag wymienionych poniżej.

1. Średni błąd względny ciepła spalania obliczony na podstawie składu gazu prognozowanego z użyciem modelu MLP i rzeczywistego składu gazu wynosi 1,57%. Doktorantka podczas realizacji rozprawy doktorskiej współpracowała z Operatorem Gazociągów Przesyłowych GAZ-SYSTEM S.A. Czy z praktycznego punktu widzenia wykorzystania opracowanego modelu do prognozowania składu gazu ziemnego taki poziom błędu jest dopuszczalny? Czy dopuszczalny błąd prognozy może być większy?
2. Na stronie 139 zamieszczone jest zdanie: „Z danych zamieszczonych na Rys. 5-25 i Rys. 5-26 wynika, że najwyższe współczynniki RT i najniższe błędy SSET charakteryzują modele MLP wytrenowane za pomocą metody BFGS i przyjętej logistycznej funkcji aktywacji neuronów w warstwie ukrytej i wyjściowej, niezależnie od liczby neuronów w warstwie ukrytej.” Analizując

rysunki 5–25 i 5–26, nie zauważyłem wyraźnych różnic między seriami danych oznaczonymi „BFGS log-log” i „BFGS tan-log” oraz między „grad sprz log-log” i „grad sprz tan-log”. W jaki sposób Autorka sprawdziła istotność różnic?

3. Zastanowiło mnie również podejście zastosowane przez Doktorantkę do badań nad możliwością zastosowania modelu MLP do prognozowania mieszaniny gazu ziemnego i wodoru. Autorka wykorzystwała do trenowania i testowania modelu zbiory danych zawierające metan, etan, propan, azot i ditlenek węgla oraz zakładając stałe stężenie wodoru (5%), a zawartość pozostałych składników przeliczając. Tym samym wartość średniego znormalizowanego pierwiastka błędu średniokwadratowego dla prognozy wodoru (tab. 5–14) właściwie wynosi zero. Powstaje pytanie, czy w świetle przeprowadzonych przez Doktorantkę badań zastosowanie opracowanego modelu MLP nie ogranicza się do przypadku, w którym szósty składnik na wyjściu będzie miał stałe stężenie.

Pewien niedosyt pozostawia szata graficzna rozprawy, której lepsze opracowanie mogłoby poprawić przejrzystość pracy. Na wykresach stosowane są znaczniki o tym samym kształcie dla różnych serii danych, które zróżnicowane są jedynie barwą, co bardzo utrudnia analizę wykresów (np. rys. 5–15 i 5–16, rozróżnienie serii to nie lada wyzwanie). Zauważyłem również kilka innych drobnych błędów, np. na stronie 145: „...udziały metanu lub etanu lub propanu lub azotu lub ditlenku węgla...” – zapis zdania z ograniczoną ilością „lub” byłby lepszy; strona 162 – powtórzenie „dzień tygodnia”.

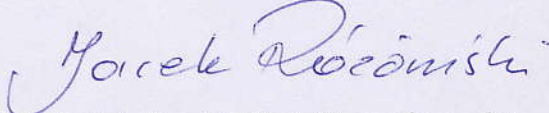
Wniosek końcowy

Rozprawa doktorska Pani mgr inż. Pauliny Muchel pokazuje, że Autorka posiada dużą wiedzę z zakresu użycia sztucznej inteligencji do prognozowania procesów zachodzących w instalacjach przemysłowych. W mojej ocenie opracowany przez Doktorantkę model MLP stanowi oryginalne rozwiązanie problemu naukowego. Na podkreślenie zasługuje duża liczba przeprowadzonych czasochłonnnych obliczeń, które umożliwiły Autorce opracować model sztucznych sieci neuronowych do prognozowania składu gazu ziemnego. Określony przez Doktorantkę cel badań został zrealizowany, a przedstawione w rozprawie doktorskiej wyniki w istotny sposób mogą przyczynić się do lepszego prognozowania składu gazu ziemnego, a w przyszłości gazów pochodzących ze źródeł odnawialnych. Część uzyskanych wyników została już przez Doktorantkę opublikowana w renomowanych czasopismach. Trzy prace ukazały się w czasopismach znajdujących się na liście Journal Citation Reports (JCR). Dwie prace zostały wydane w prestiżowym czasopiśmie „Energy”, co

świadczy o zainteresowaniu specjalistów zajmujących się transportem gazu badaniami Pani mgr inż. Pauliny Muchel.

Podsumowując, recenzowana rozprawa reprezentuje dobry poziom naukowy, zawiera elementy nowości naukowej, a wymienione w recenzji uwagi i pytania nie umniejszają mojej wysokiej oceny recenzowanej pracy. W mojej ocenie rozprawa doktorska Pani mgr inż. Pauliny Muchel spełnia wymogi określone w art. 187 Ustawy z dnia 20 lipca 2018 r. – Prawo o szkolnictwie wyższym i nauce (Dz.U. 2018, poz. 1668 z późniejszymi zmianami). W związku z powyższym wnioskuję do Rady Dyscypliny Inżynierii Chemicznej Uniwersytetu Technologicznego w Szczecinie o przyjęcie pracy oraz dopuszczenie Pani mgr inż. Pauliny Muchel do dalszych etapów przewodu doktorskiego.

Poznań, 22 sierpnia 2024 roku


dr hab. inż. Jacek Różański, prof. uczelni