

Recenzja
pracy doktorskiej
mgr inż. Michała Zawodnego
pt.: „Metoda agregacji danych z dużych wolumenów, zróżnicowanego
pochodzenia i struktury w celu udoskonalenia i tworzenia modeli
transportowych”

(Promotor: dr hab. inż. Maciej Kruszyna, prof. uczelni)

Podstawą opracowania recenzji jest Zawiadomienie nr 19/05/D06/2026 z dnia 15 maja 2026 r. o wyznaczeniu Recenzenta. Recenzowana rozprawa została przygotowana na Wydziale Budownictwa Lądowego i Wodnego Politechniki Wrocławskiej. Praca liczy 157 stron i składa się z ośmiu rozdziałów, wykazu oznaczeń, bibliografii obejmującej 175 pozycji, streszczeń w języku polskim i angielskim oraz załączników. Do rozprawy dołączono pseudokod i zapis formuły zastosowanej w arkuszu kalkulacyjnym, natomiast macierze podróży dla modelu bazowego i analizowanych wariantów udostępniono w formie elektronicznej.

1. Ocena poszczególnych części rozprawy

Rozdział 1 stanowi wprowadzenie do problematyki wykorzystania danych o dużych wolumenach w modelowaniu transportu. Autor trafnie wskazuje, że rozwój automatycznych systemów pomiarowych, telefonii komórkowej, systemów biletowych, danych lokalizacyjnych i współdzielonej mobilności zwiększa zasób informacji możliwych do wykorzystania przy budowie modeli podróży. Jednocześnie różnorodność tych źródeł powoduje problemy związane z ich niekompatybilnością, odmiennym pokryciem przestrzennym i czasowym oraz zróżnicowaną reprezentatywnością. Podjęty problem jest aktualny, ważny poznawczo i ma znaczenie praktyczne.

Rozdział 2 zawiera przegląd literatury dotyczącej modelowania podróży, wykorzystania Big Data, charakterystyki wybranych źródeł informacji oraz metod symulowania danych. Za wartościowe należy uznać zestawienie w jednym miejscu

dziewięciu źródeł: GSM, GPS, Bluetooth, APC, Smart Card, ANPR, danych z mediów społecznościowych, taksówek oraz flot rowerów miejskich i UTO. Przegląd ma szeroki zakres i potwierdza duży nakład pracy włożony w zebranie materiału. Autor identyfikuje istotną potrzebę badawczą, jaką jest poszukiwanie metod integracji wielu heterogenicznych źródeł danych. Jednocześnie przegląd ma w znacznej części charakter opisowy. Brakuje bardziej krytycznego porównania istniejących metod *data fusion*, estymacji macierzy O-D, ważenia źródeł i oceny niepewności. W konsekwencji luka badawcza została sformułowana szerzej, niż wynika to bezpośrednio z przeprowadzonego przeglądu.

W **rozdziale 3** przedstawiono przedmiot, cele i tezy pracy. Głównym celem jest opracowanie metody agregacji danych pochodzących ze zróżnicowanych źródeł Big Data w celu ich wykorzystania w procesie budowy modeli transportowych. Dodatkowym celem jest stworzenie teoretycznego ośrodka miejskiego, który ma stanowić podstawowe narzędzie do weryfikacji opracowanych modeli. Sformułowano dwie tezy: pierwszą, zgodnie z którą wykorzystanie Big Data pozwala uzyskać dokładniejsze wyniki, oraz drugą, dotyczącą efektu synergicznego wynikającego z łączenia wielu źródeł. Tezy wskazują kierunek badań, jednak nie zostały dostatecznie zoperacjonalizowane. Nie określono, względem jakiej metody lub zbioru referencyjnego należy rozumieć „dokładniejsze wyniki”, ani za pomocą jakiego kryterium miałyby zostać rozpoznany efekt synergiczny. Jest to zasadniczy problem interpretacyjny powracający w dalszej części pracy.

Rozdział 4 stanowi zasadniczą część rozprawy, ponieważ przedstawiono w nim autorską propozycję agregacji macierzy podróży pochodzących z różnych źródeł danych. Autor prawidłowo rozpoznaje podstawowy problem praktyczny: poszczególne źródła obejmują odmienne grupy użytkowników, środki transportu oraz fragmenty obszaru, a ponadto charakteryzują się różnym stopniem kompletności. W rezultacie ich bezpośrednie zestawienie nie jest możliwe bez uprzedniego przeliczenia i sprowadzenia do porównywalnej postaci. Samo podjęcie próby uporządkowania tego problemu należy ocenić pozytywnie, ponieważ wraz z rozwojem automatycznych systemów gromadzenia danych zagadnienie integracji heterogenicznych zbiorów będzie nabierało coraz większego znaczenia.

Za wartościowe uważam jednoznaczne ograniczenie metody do poziomu macierzy źródłowo - cel oraz określenie kilku możliwych sposobów jej wykorzystania. Zaproponowane rozwiązanie może służyć zarówno do utworzenia nowej macierzy na podstawie dostępnych źródeł, jak i do uzupełniania brakujących relacji bądź zastępowania wartości uznanych za nierealistyczne. Takie ujęcie nadaje metodzie elastyczny charakter i zwiększa jej potencjalną przydatność praktyczną, zwłaszcza w przypadku ośrodków dysponujących niepełnymi lub rozproszonymi danymi.

Pozytywną cechą metody jest jej modułowa konstrukcja. Liczba wykorzystywanych źródeł nie jest z góry ograniczona, a zaproponowany zapis pozwala zarówno na dodawanie nowych macierzy, jak i na wyłączanie tych, które w danym przypadku nie są

dostępne. Metoda nie jest również związana wyłącznie ze źródłami zaliczanymi do Big Data i może potencjalnie uwzględniać także wyniki klasycznych badań ruchu. Autor przyjął ponadto podejście lokalne, prowadząc obliczenia osobno dla każdej relacji między rejonami. Dzięki temu brak informacji w jednym źródle nie musi automatycznie eliminować danych dostępnych w pozostałych zbiorach. Sama intencja dostosowywania liczby uwzględnianych źródeł do zawartości danej komórki jest właściwa, nawet jeśli sposób rozróżniania braku danych i wartości zerowej wymaga dopracowania.

Na podkreślenie zasługuje przejrzysty sposób przedstawienia procedury. Autor zapisał ją kolejno w postaci wzoru obejmującego dziewięć źródeł, formuły uogólnionej, pseudokodu, schematu blokowego oraz formuły arkusza kalkulacyjnego. Takie wielostronne przedstawienie algorytmu ułatwia jego zrozumienie, kontrolę i późniejszą implementację. Czytelnik może stosunkowo łatwo odtworzyć logikę obliczeń i przenieść ją do innego środowiska programistycznego. Jest to istotna zaleta z punktu widzenia transparentności i potencjalnej reprodukowalności metody. Korzystne jest również to, że zaproponowane rozwiązanie nie wymaga specjalistycznego oprogramowania i może zostać wykorzystane w powszechnie dostępnym arkuszu kalkulacyjnym.

Interesującym elementem rozdziału jest rozdzielenie dwóch różnych zagadnień: skali reprezentacji danego źródła oraz jego wiarygodności. Pierwsze z nich Autor ujmuje za pomocą współczynnika przeliczeniowego, natomiast we wzorze 4.4 wprowadza dodatkowy współczynnik wiarygodności. Pokazuje to świadomość, że duża liczba obserwacji nie musi oznaczać wysokiej jakości danych, a poszczególne źródła mogą w różnym stopniu obciążać wynik. Wprawdzie w badaniach zasadniczych współczynników wiarygodności ostatecznie nie zastosowano, jednak sama konstrukcja metody przewiduje taką możliwość i tworzy właściwy kierunek jej dalszego rozwoju. Rozszerzenie algorytmu o wagi wyznaczone na podstawie jakości danych, zakresu przestrzennego, liczebności próby i niepewności pomiaru mogłoby istotnie zwiększyć jego wartość.

Dobłą stroną podrozdziału 4.2 jest próba systematycznego scharakteryzowania dziewięciu źródeł danych. Autor nie ogranicza się wyłącznie do wskazania ich nazw, lecz omawia ich zasięg, sposób rejestracji, potencjalną kompletność oraz podstawowe ograniczenia. Dostrzega między innymi punktowy charakter pomiarów Bluetooth i ANPR, ograniczenie danych APC i kart miejskich do transportu zbiorowego, zależność danych GSM od udziału operatora, selektywność danych GPS i mediów społecznościowych oraz ograniczenie danych o mikromobilności do flot współdzielonych. Zestawienie tak różnych technologii w jednym schemacie pojęciowym stanowi przydatny element pracy i potwierdza szerokie rozpoznanie problematyki przez Autora.

Rozdział 4 pokazuje również praktyczne nastawienie Doktoranta. Zaproponowane rozwiązanie jest stosunkowo proste obliczeniowo, możliwe do wdrożenia niewielkim kosztem i łatwe do modyfikacji. Nie jest to metoda zaawansowana matematycznie, lecz jej wartość może polegać przede wszystkim na uporządkowaniu procesu postępowania z niekompletnymi i niejednorodnymi macierzami. W tym sensie rozdział ten można uznać

za oryginalny demonstrator koncepcji, stanowiący podstawę do dalszego rozwinięcia w kierunku pełnoprawnej metody integracji danych.

Najważniejsze zastrzeżenia dotyczą nie tyle samej ogólnej idei zaproponowanej w rozdziale 4, ile sposobu wyznaczenia konkretnych współczynników oraz braku empirycznej oceny ich poprawności. Należy zatem rozdzielić pozytywną ocenę koncepcji, struktury i przejrzystości algorytmu od krytycznej oceny jego parametryzacji. Pomimo wskazanych uwag rozdział 4 można uznać za jeden z mocniejszych elementów rozprawy: przedstawia on autorską, implementowalną i rozwojową propozycję, której przydatność wymaga jednak dalszej kalibracji i walidacji na spójnym zbiorze danych rzeczywistych.

Rozdział 5 opisuje sposób utworzenia syntetycznego ośrodka miejskiego, poglądowego modelu w programie PTV Visum oraz przygotowania dziewięciu macierzy wejściowych. Jest to pracowita część rozprawy, a przygotowanie spójnego środowiska demonstracyjnego wymagało sprawności technicznej. Syntetyczny model może być użyteczny do pokazania działania algorytmu oraz do badań wstępnych. Nie może jednak pełnić funkcji niezależnej rzeczywistości referencyjnej. Liczby mieszkańców, współczynniki ruchliwości, atrakcyjności rejonów, zakresy losowania i układ sieci zostały przyjęte arbitralnie, bez kalibracji do rzeczywistego ośrodka. Określenie „pseudo Monte Carlo” odnosi się w istocie do pojedynczego generowania liczb pseudolosowych z zadanych przedziałów, bez wielokrotnych realizacji, analizy rozkładów wyników, stabilności i niepewności.

Pozytywnie należy ocenić uporządkowanie syntetycznego ośrodka poprzez wydzielenie rejonów centralnych, miejskich, podmiejskich i zewnętrznych oraz uwzględnienie grup mieszkańców, motywacji podróży i atrakcyjności rejonów. Autor przygotował kompletną macierz 100×100 , a następnie opracował poglądowy model sieci w programie PTV Visum, wykazując się sprawnością w przetwarzaniu macierzy i obsłudze narzędzi modelowania transportowego.

Wartościowym elementem jest również zgromadzenie i sprowadzenie do wspólnej postaci dziewięciu macierzy reprezentujących bardzo różne technologie pomiarowe. Pomimo dyskusyjnego sposobu ich przekształcania, wykonane prace wymagały znacznego nakładu obliczeniowego i organizacyjnego. Mapy ciepła oraz wizualizacja potoków na sieci poprawiają czytelność analiz i pozwalają obserwować wpływ poszczególnych źródeł na strukturę macierzy wynikowej. Rozdział ten należy zatem traktować jako dobrze przygotowane środowisko demonstracyjne do badania działania algorytmu, choć nie jako model skalibrowany do rzeczywistego ośrodka miejskiego.

Największe zastrzeżenia budzi jednak sposób przygotowania macierzy wejściowych. Dane pochodzą z różnych miast, okresów i systemów pomiarowych, a następnie zostały przeskalowane i przypisane do rejonów sztucznego ośrodka. W przypadku danych wrocławskich rejonu uporządkowano według wielkości potoków i przypisano do kategorii rejonów modelu syntetycznego. Pozostałe macierze

umieszczano w wybranych fragmentach macierzy 100×100 , kierując się ich ogólną charakterystyką. Trzy źródła odtworzono na podstawie map ciepła i opisów literaturowych. Po takich przekształceniach zestawy nie opisują tego samego systemu transportowego, czasu ani populacji. Mogą służyć do demonstracji pracy algorytmu, ale nie do oceny jego zgodności z rzeczywistym przebiegiem podróży.

Istotne wątpliwości dotyczą również wartości części współczynników. Na przykład na s. 44 Autor podaje 2,37 mld mil pokonywanych dziennie (co samo w sobie jest wartością podejrzaną), 5567 mil przypadających na jednego podróżnego, a następnie wyprowadza z tych danych 425 723 podróży i współczynnik 1,237%. Iloraz odległości całkowitej i odległości przypadającej na osobę daje jednak liczbę osób, a nie podróży. Następnie liczba podróży rowerami i UTO zostaje podzielona przez liczbę osób, co jest działaniem niespójnym wymiarowo. Prawdopodobnie pomieszane zostały dane roczne z dziennymi. Ponieważ wartości źródłowe są w algorytmie mnożone przez odwrotności współczynników, podobne błędy mogą zostać znacząco wzmocnione w macierzy wynikowej. Analogiczne wyjaśnienia są potrzebne dla współczynników dotyczących GPS, APC, kart miejskich i ANPR.

Rozdział 6 został zatytułowany „Walidacja metody”. Autor przedstawia dziewięć wariantów zestawienia źródeł oraz porównuje ich wyniki z syntetyczną macierzą bazową. Pozytywnie oceniam próbę pokazania konsekwencji różnych konfiguracji danych i wykorzystanie programu PTV Visum do wizualizacji zmian potoków. Analizy te dostarczają informacji o zachowaniu algorytmu w warunkach braku części danych i mogą być potraktowane jako eksperyment symulacyjny lub analiza demonstracyjna. Nie stanowią jednak walidacji w ścisłym znaczeniu, ponieważ brakuje niezależnych obserwacji referencyjnych dotyczących tego samego ośrodka.

Podstawową miarą porównania jest różnica pomiędzy sumami liczby podróży. Wariant obejmujący wszystkie źródła różni się od modelu bazowego o 2,22%. Taka zgodność sum nie oznacza jednak zgodności macierzy: przeszacowania i niedoszacowania poszczególnych relacji mogą się wzajemnie kompensować. Należało zastosować miary błędów obliczane dla komórek macierzy, na przykład MAE, RMSE, medianę błędów względnego, korelację lub porównanie rozkładów długości i kierunków podróży. Niejasna jest także interpretacja odchylenia standardowego w tabeli 6.2. Najniższe odchylenia uzyskują warianty najbardziej odległe od modelu bazowego pod względem liczby podróży, co raczej potwierdza, że nie jest to miara trafności metody.

Dobór dziewięciu wariantów ma charakter ekspercki, ale nie wyczerpuje możliwych kombinacji źródeł. Wariant dziewiąty utworzono na podstawie wyników wariantów wcześniejszych. Takie postępowanie może być przydatne eksploracyjnie, lecz nie stanowi niezależnego testu. W konsekwencji nie można uznać, że sama przewaga wariantu obejmującego wszystkie źródła dowodzi efektu synergicznego. Równie prawdopodobny jest efekt uśredniania albo przypadkowego kompensowania błędów wynikających z arbitralnych współczynników.

Rozdziały 7 i 8 obejmują dyskusję, ocenę metody, realizację celu i kierunki dalszych prac. Pozytywnie należy ocenić, że Autor dostrzega potrzebę zastosowania wag wiarygodności, przeprowadzenia badań na danych z jednego rzeczywistego ośrodka oraz rozszerzenia metody. Szczególnie ważne jest stwierdzenie na s. 118, że metoda zostanie ostatecznie przetestowana na danych rzeczywistych. Potwierdza ono świadomość ograniczeń przyjętego podejścia. Pozostaje jednak w sprzeczności z wcześniejszym określeniem analiz jako walidacji oraz z jednoznacznym uznaniem tez za potwierdzone. W mojej ocenie przedstawione wyniki potwierdzają możliwość technicznej implementacji koncepcji, nie zaś dokładność ani uniwersalną skuteczność metody.

Bibliografia obejmuje 175 pozycji, w większości anglojęzycznych i uwzględnia wiele aktualnych publikacji. Jej szeroki zakres jest wartością rozprawy. Autor wykonał znaczną pracę w celu rozpoznania bardzo zróżnicowanych technologii pozyskiwania danych.

2. Uwagi szczegółowe

- S. 33–34 – pojęcie „dokładniejszych wyników” nie zostało zdefiniowane. Należałoby wskazać wartość referencyjną, metodę porównawczą i miarę dokładności.
- S. 38–41 oraz Załączniki 1–2: algorytm traktuje każdą wartość zerową jako brak danych. Konieczne jest rozróżnienie prawidłowego zera strukturalnego, zera wynikającego z braku detekcji oraz braku obserwacji.
- S. 42 – zalecenie stosowania wartości współczynników podanych w rozprawie w innych ośrodkach, jeżeli nie ma współczynników lokalnych, jest zbyt daleko idące i nie zostało empirycznie uzasadnione.
- S. 43 – współczynnik dla Bluetooth łączy wykrywalność urządzenia z udziałem modalnym, nie uwzględniając między innymi napętnienia pojazdów, lokalizacji detektorów i odmiennej skłonności użytkowników do korzystania z urządzeń.
- S. 43–44 – przeliczenie danych APC z jednej godziny przez przemnożenie ich przez 17 godzin kursowania zakłada stałe natężenie przez cały dzień i wymagałoby uzasadnienia albo korekty profilem dobowym.
- S. 44 – wyliczenie współczynnika 1,237% dla rowerów i UTO jest niespójne wymiarowo; wartości 2,37 mld mil, 5567 mil oraz 425 723 powinny zostać zweryfikowane także pod względem okresu odniesienia.
- S. 44 – 81 tys. użytkowników w populacji 1,8 mln stanowi 4,5%, a nie 5,06%. Następnie przyjęcie wartości 37% wymaga jednoznacznego uzasadnienia.
- S. 46–48 – zakresy losowania liczby mieszkańców, ruchliwości i atrakcyjności rejonów są arbitralne i nie mają wiele wspólnego z zastosowanymi danymi; nie przeprowadzono analizy wrażliwości wyniku na te założenia.
- S. 58–66 – sortowanie, redukcja i przypisanie macierzy z różnych miast do rejonów syntetycznych prowadzi do utraty ich pierwotnego znaczenia przestrzennego.
- S. 68 – określenie „walidacja” jest nieadekwatne. Odpowiedniejsze byłoby określenie „eksperyment symulacyjny” albo „analiza działania metody”.

- S. 115–116 – różnica sumy podróży nie jest wystarczającą miarą podobieństwa macierzy, ponieważ nie uwzględnia rozmieszczenia błędów w relacjach O-D.
- S. 117 – nie opisano dokładnie, dla jakiego zbioru wartości obliczano odchylenie standardowe ani dlaczego miałyby ono mierzyć zgodność z modelem bazowym.
- S. 118 i 122 – wskazanie przyszłego testu na danych rzeczywistych potwierdza, że w rozprawie nie przeprowadzono pełnej walidacji empirycznej.
- Załącznik 2 – formuła arkusza kalkulacyjnego nie zabezpiecza przypadku, w którym wszystkie źródła mają wartość zero; w takim przypadku mianownik przyjmuje wartość zerową. Pseudokod zawiera natomiast odpowiednią obsługę tego przypadku.

3. Zagadnienia do dyskusji podczas publicznej obrony

Przedstawione powyżej uwagi i wątpliwości dotyczą zarówno podstaw metodycznych rozprawy, jak i sposobu interpretacji uzyskanych wyników. Część z nich może zostać wyjaśniona podczas publicznej obrony, zwłaszcza w zakresie świadomości ograniczeń zaproponowanego podejścia oraz rzeczywistego zakresu osiągnięcia naukowego. Oczekuję zatem, że Doktorant odniesie się do poniższych zagadnień, wykazując umiejętność krytycznej oceny własnych badań i wskazując sposób dalszego rozwinięcia oraz empirycznej weryfikacji metody.

1. Na czym, zdaniem Doktoranta, polega oryginalny wkład zaproponowanej metody w porównaniu ze zwykłym przeskalowaniem i średnią arytmetyczną dostępnych estymat macierzy O-D?
2. Czy Doktorant podtrzymuje określenie analiz z rozdziału 6 jako walidacji? Proszę wskazać, jaka była niezależna wartość referencyjna i które elementy badania spełniają warunek empirycznej walidacji.
3. W jaki sposób w danych źródłowych należy rozróżniać prawidłową wartość zero od braku obserwacji? Jaki kierunek i skalę błędu może powodować wyłączenie wszystkich zer z obliczeń?
4. Jak zmieniłyby się wyniki po zmianie przyjętych współczynników przeliczeniowych? Proszę zaproponować schemat analizy wrażliwości i sposób szacowania niepewności macierzy wynikowej.
5. Jakie miary należy zastosować do porównania dwóch macierzy O-D, aby uniknąć kompensowania się przeszacowań i niedoszacowań?
6. Jak powinno wyglądać badanie na danych dotyczących jednego rzeczywistego ośrodka, które pozwoliłoby wiarygodnie ocenić dokładność i efekt synergiczny?

4. Ocena redakcji pracy

Układ rozprawy jest zasadniczo logiczny: od wprowadzenia i przeglądu literatury, przez cele, opis metody i modelu syntetycznego, do prezentacji wariantów, dyskusji i wniosków. Praca zawiera liczne schematy, mapy ciepła i wyniki symulacji przeprowadzonych w programie PTV Visum, które ułatwiają śledzenie toku postępowania. Pozytywnie należy ocenić zamieszczenie wykazu oznaczeń oraz przedstawienie algorytmu w kilku równoważnych formach.

Jednocześnie rozprawa wymagałaby staranniejszej redakcji językowej i terminologicznej. Występują powtórzenia, niezręczne sformułowania oraz zbyt kategoryczne oceny, takie jak „wynik satysfakcjonujący”, „realistyczna macierz” lub „potwierdzenie tezy”, niepoparte odpowiednimi kryteriami. Termin „walidacja” został użyty niezgodnie z zakresem wykonanych analiz. W opisie obliczeń pojawiają się błędy jednostek i okresów odniesienia. Część rozdziału 6 jest nadmiernie rozbudowana przez powtarzalny opis kolejnych zrzutów ekranowych, podczas gdy zasadnicza analiza statystyczna jest bardzo krótka.

5. Ogólna ocena pracy

Podjęty w rozprawie problem jest aktualny i ważny dla rozwoju współczesnego modelowania transportowego. Rosnąca liczba automatycznych źródeł danych rzeczywiście wymaga metod integracji, kontroli jakości i ograniczania obciążeń charakterystycznych dla pojedynczych technologii. Autor wykazał się szeroką orientacją w dostępnych źródłach informacji oraz umiejętnością przygotowania macierzy, implementacji algorytmu i wykorzystania programu PTV Visum. Opracował przejrzystą, modułową procedurę obliczeniową, którą można uznać za autorską koncepcję i punkt wyjścia do dalszych badań.

Największą wartością rozprawy nie jest jednak – wbrew niektórym stwierdzeniom Autora – udowodniona dokładność metody, lecz stworzenie demonstratora integracji heterogenicznych i niekompletnych macierzy podróży. W tym znaczeniu praca ma charakter *proof of concept*. Pokazuje możliwość wykonania operacji i pozwala zidentyfikować problemy wymagające rozwiązania: sposób wyznaczania współczynników, ocenę wiarygodności źródeł, rozróżnienie wartości zerowych od braków danych, analizę niepewności oraz potrzebę walidacji empirycznej.

Zastrzeżenia dotyczące danych wejściowych i interpretacji wyników budzą wątpliwości. W rozprawie nie przeprowadzono walidacji na niezależnych danych dotyczących jednego rzeczywistego ośrodka. Współczynniki przeliczeniowe są częściowo arbitralne, a niektóre zostały wyznaczone błędnie. Zastosowane miary nie pozwalają wykazać zgodności przestrzennej macierzy ani efektu synergicznego. Dlatego twierdzenie o potwierdzeniu obu tez należy uznać za zbyt daleko idące.

6. Wniosek końcowy i jego uzasadnienie

Za elementy uzasadniające dopuszczenie rozprawy do publicznej obrony uznaję:

1. podjęcie aktualnego i istotnego problemu integracji heterogenicznych źródeł danych wykorzystywanych w modelowaniu transportowym;
2. szerokie rozpoznanie dziewięciu współczesnych źródeł informacji o mobilności oraz zestawienie ich możliwości i ograniczeń;
3. opracowanie autorskiej, modułowej i przejrzystej procedury agregacji niekompletnych macierzy podróży;
4. przedstawienie procedury w postaci matematycznej, algorytmicznej i implementacyjnej;
5. przygotowanie środowiska symulacyjnego i przeprowadzenie wielowariantowego eksperymentu pokazującego zachowanie proponowanego algorytmu;
6. zidentyfikowanie kierunków dalszego rozwoju, w szczególności potrzeby ważenia źródeł i przeprowadzenia empirycznej walidacji.

Przedstawione w recenzji uwagi krytyczne wymagają szczegółowego odniesienia się do nich podczas publicznej obrony. Dotyczą przede wszystkim braku walidacji empirycznej, jakości części danych i współczynników, niewłaściwego sposobu oceny dokładności oraz zbyt kategorycznej interpretacji wyników. Nie przekreślają jednak faktu opracowania autorskiej koncepcji i wykazania jej technicznej wykonalności.

Biorąc pod uwagę aktualność podjętego problemu, wykonany przegląd źródeł danych, opracowanie własnego schematu agregacji oraz jego wdrożenie, wyrażam opinię, że recenzowana rozprawa – przy wszystkich wskazanych zastrzeżeniach – spełnia wymagania stawiane rozprawom doktorskim w stopniu uzasadniającym dopuszczenie jej do publicznej obrony. Wnoszę zatem o dopuszczenie mgr inż. Michała Zawodnego do publicznej obrony rozprawy doktorskiej w dyscyplinie inżynieria lądowa, geodezja i transport.

prof. dr hab. inż. Andrzej Szarata