

Blaski i cienie statystyki GEH do oceny poprawności modeli ruchu¹

Streszczenie: Autor omawia możliwości stosowania statystyki GEH do oceny poprawności modeli ruchu. Statystyka ta nosi nazwę od Geoffreya E. Haversa, który zastosował ją w trakcie prac nad modelem ruchu dla Londynu w latach 70. ubiegłego wieku. Jest to formuła empiryczna, która dla niskich wartości dopuszcza większe odchylenie od pomiaru, a dla wysokich mniejsze. Statystyka GEH jest zalecana przez brytyjskie wytyczne projektowania dróg i mostów i inne wytyczne do oceny poprawności modeli transportowych. W artykule podjęto dyskusję na temat konieczności stosowania różnych wartości granicznych GEH w zależności od rozkładu wielkości ruchu uzyskiwanych w pomiarach.

Słowa kluczowe: modele ruchu, ocena modeli ruchu, statystyka GEH

Wprowadzenie

Ostatnio w bardzo wielu opisach przedmiotu zamówienia na stworzenie modeli ruchu do oceny poprawności modelu zaleca się sprawdzenie wyników statystyką GEH. Przykładem może być np. SIWZ dla modelu ruchu dla województwa mazowieckiego [1]. Postawiono tam następujące wymogi:

W odniesieniu do sprawdzenia poprawności wyników rozkładu ruchu możliwy jest następujący błąd w oszacowaniu:

- w przypadku natężenia ruchu pojazdów (osobowych i towarowych):
 - $\pm 15\%$ na poszczególnych punktach kontrolnych, dla których wykonywane jest sprawdzenie, w przypadku gdy natężenie pojazdów jest na poziomie 700–2700 pojazdów/godzinę dla więcej niż 85% liczby punktów;
 - ± 100 pojazdów/godzinę na poszczególnych punktach kontrolnych, dla których wykonywane jest sprawdzenie, w przypadku, gdy natężenie pojazdów jest mniejsze od 700 pojazdów/godzinę, dla więcej niż 85% liczby punktów;
 - ± 400 pojazdów/godzinę na poszczególnych punktach kontrolnych, dla których wykonywane jest sprawdzenie, w przypadku gdy natężenie pojazdów jest większe od 2700 pojazdów/godzinę, dla więcej niż 85% liczby punktów;
 - $\pm 5\%$ dla sumy poszczególnych kordonów.

Dodatkowo uzyskane wyniki powinny spełnić warunki odnośnie wartości współczynnika GEH (obliczanego według wzoru):

$$GEH = \sqrt{\frac{(E - V)^2}{(E + V)/2}} \quad (1)$$

gdzie:

E – wartość natężenia uzyskana z modelu

V – wartość natężenia uzyskana z pomiarów ruchu;

- $GEH < 5$ – na poszczególnych punktach kontrolnych, dla których wykonywane jest sprawdzenie, dla więcej niż 85% liczby punktów,
- $GEH < 4$ – dla sumy poszczególnych kordonów.

Co to jest statystyka GEH

Jak łatwo sprawdzić w Wikipedii [3], jest to formuła używana w inżynierii ruchu i modelach ruchu. Została ona wymyślona przez Geoffreya E. Haversa i od pierwszych liter jego imion i nazwiska nosi nazwę, a powstała w trakcie prac nad modelem ruchu dla Londynu w latach 70. ubiegłego wieku. Mimo że wygląda ona podobnie do testu zgodności χ^2 , jest to formuła empiryczna, która dla niskich wartości dopuszcza większe odchylenie od pomiaru, a dla wysokich mniejsze. Ma to swój głęboki sens. Przy analizie układów ulic w mieście to, czy na arterii godzinowy potok będzie rzędu 5000 czy 6000, ma istotne znaczenie. To czy na drodze lokalnej potok będzie 100 czy 120 (taki sam błąd procentowy, jak dla arterii), nie ma żadnego znaczenia. Statystyka GEH jest zalecana przez brytyjskie wytyczne projektowania dróg i mostów [2] i inne wytyczne do oceny poprawności modeli transportowych.

Zgodnie z [2] przyjmuje się, że dla 85% analizowanych wartości potoków godzinowych wartości GEH powinny być niższe niż 5. Jeśli występujące wartości GEH są większe niż 10, to istnieje wysokie prawdopodobieństwo błędu czy to w modelu, czy we wprowadzonych danych ruchowych.

Co oznacza współczynnik GEH o zadanej wartości

Tyle teoria – spróbujmy przyjrzeć się teraz nieco praktyce. Co to oznacza, że współczynnik GEH ma być mniejszy od 5, jakie odchylenia od pomierzonych natężeń ruchu są dopuszczalne w modelu. Spójrzmy na nieco inaczej zapisaną formę porównania potoków pomierzonych i policzonych:

$$GEH = \sqrt{\frac{2 \times (M - C)^2}{M + C}} \quad (2)$$

gdzie:

M – wartość natężenia uzyskana z modelu

C – wartość natężenia uzyskana z pomiarów ruchu.

Jeżeli przyjmiemy, że

$$M = C + \mu \quad (3)$$

gdzie:

M, C – jak we wzorze (2),

μ – maksymalny dopuszczalny błąd

¹ ©Transport Miejski i Regionalny, 2016.

to możliwe są następujące przekształcenia:

$$GEH = \sqrt{\frac{2 \times (C + \mu - C)^2}{C + \mu + C}} \quad (4)$$

$$GEH^2 = \frac{2 \times \mu^2}{2C + \mu} \quad (5)$$

$$GEH^2 \times 2C + GEH^2 \times \mu = 2 \times \mu^2 \quad (6)$$

$$2 \times \mu^2 - GEH^2 \times \mu - 2C GEH^2 = 0 \quad (7)$$

Mamy klasyczne równanie kwadratowe, które rozwiązujemy:

$$\mu_1 = \frac{GEH^2 + \sqrt{GEH^2 \times (GEH^2 + 16 \times C)}}{2 \times 2} \quad (8)$$

$$\mu_2 = \frac{GEH^2 - \sqrt{GEH^2 \times (GEH^2 + 16 \times C)}}{2 \times 2} \quad (9)$$

Możemy zatem dla wskazanej wartości GEH – np. 5, pokazać, w jakim przedziale powinny zawierać się prawidłowe wartości wyliczonych potoków ruchu dla różnych wielkości ruchu pomierzonego. Można też sprawdzić, jak mają się do siebie warunki poprawności modelu przytoczone na początku artykułu.

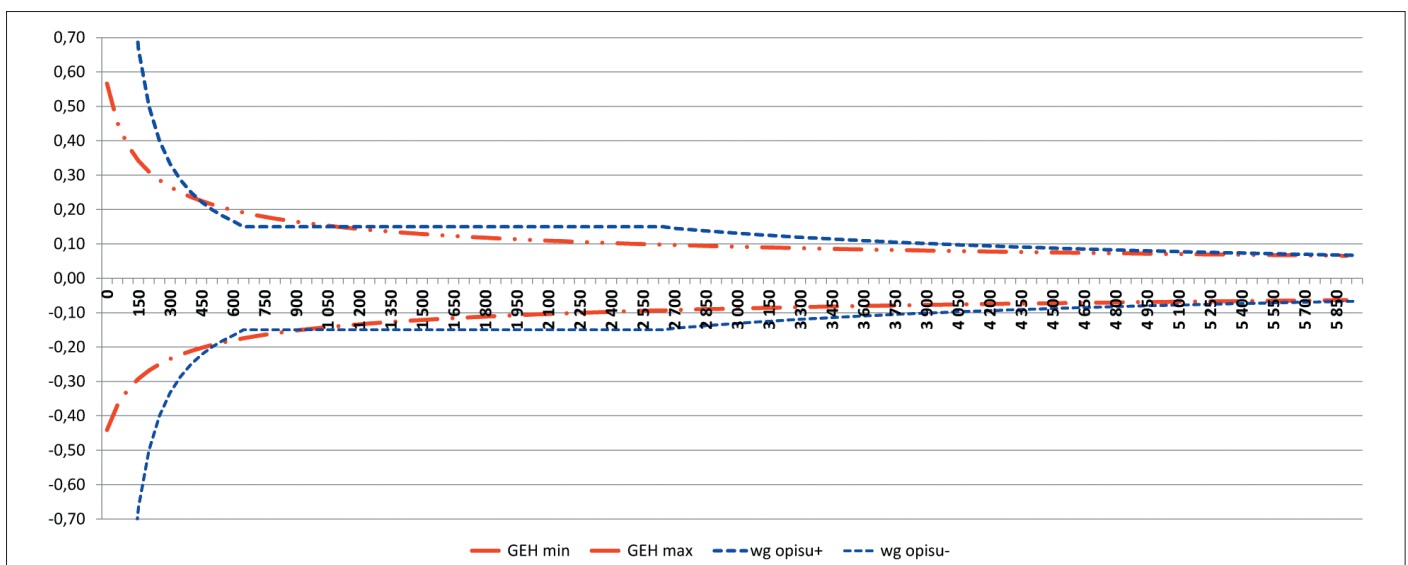
Na rysunku 1 pokazano zakres dopuszczalnych błędów względnych wyliczonych wg statystyki GEH (kolorem czerwonym) i wg warunków z OPZ (kolorem niebieskim). Widać wyraźnie, że oczekiwania opisane w warunkach przetargu są niespójne. Dla potoków rzędu 2700 pojazdów na godzinę jedna metoda oceny (GEH) za prawidłowe uznaje potoki od 2450 do 2950, natomiast druga 2300 do 3100. W przypadkach wątpliwości nie wiadomo, która metoda jest „ważniejsza”.

Należy też zwrócić uwagę, iż warunek $GEH < 5$ dotyczy potoków godzinowych. Zakres sensownego działania statystyki GEH przy podanej wartości granicznej 5 jest ograniczony.

Po poprzednim GPR (Generalnym Pomiarze Ruchu) autor niniejszego artykułu pracował nad aktualizacją krajowego modelu ruchu do pomiarów GPR2010. W ocenie poprawności modelu porównywano potoki dla 3088 odcinków jednokierunkowych. Porównywano potok wyliczony w kierunku do połowy SDRR uzyskanego z pomiaru. Na pewnym etapie budowy modelu porównanie statystyką GEH dało prawidłowe wyniki ($GEH \leq 5$) dla 35% odcinków. Przyjęcie założenia, że porównujemy ruch w godzinie szczytu, a wskaźnik godziny szczytu wynosi 10% – czyli porównujemy wartości $0,1 \times M$ i $0,1 \times C$, spowodowało natychmiastowy wzrost liczby prawidłowych odcinków do 78%. Wskaźnik godziny szczytu 7% to już 85% prawidłowych odcinków – stosując kryteria przytoczone we wprowadzeniu, można model uznać za poprawny. Jest to jednak zbyt daleko idące uproszczenie. W praktyce wskaźniki godzin szczytu są różne i konieczne byłoby wprowadzanie ich indywidualnie dla odcinków pomiarowych czy klas odcinków pomiarowych.

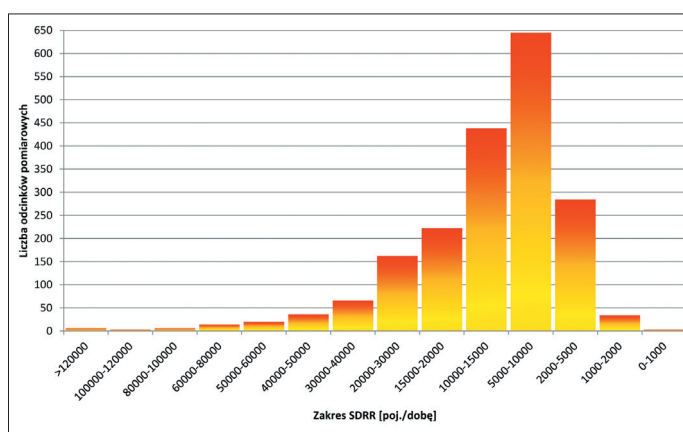
Jeśli spojrzymy na wyniki GPR 2015[4], to wartości SDRR (Średni Dobowy Ruch Roczny) zarejestrowanych na drogach krajowych wahają się od 739 do 142 269. Rozkład wielkości SDRR na odcinkach pomiarowych przedstawia się jak na rysunku 2. W badaniu uwzględniono 1953 odcinki pomiarowe. Jak łatwo zauważyć, niemal na połowie odcinków natężenia zawierają się pomiędzy 5000 a 10 000 pojazdów na dobę. Wartość średnia w punktach pomiarowych (bez ważenia długością odcinków) to 13 724 pojazdy/dobę, a mediana 10 023 pojazdy na dobę. Jeśli przyjmiemy założenie, że porównujemy ruch w jednym kierunku (połowa SDRR), to dla średniej jest to wartość około 6900 a dla mediany 5000 pojazdów/dobę.

Jeśli przyjmiemy, że dla średniego ruchu dopuszczamy odchylenie $\pm 15\%$ to, graniczna wartość GEH wynosi 12. Przy potokach zbliżonych do maksymalnych (rzędu 71 000) dopuszczalny błąd wynosi wtedy $\pm 4,5\%$. Jeśli GEH ustawimy wg mediany, to odchylenie -15% będzie dla wartości 11. Wtedy dla największych potoków dopuszczamy od-



Rys. 1. Dopuszczalne błędy względne potoków modelowych w zależności od potoków pomierzonych

Źródło: opracowanie własne



Rys. 2. Rozkład liczby odcinków pomiarowych z różnymi wielkościami SDRR
Źródło: opracowanie własne na podstawie [4]

chylenie od $-4,1\%$ do $+4,2\%$. Zatem można zastosować statystykę GEH do oceny poprawności modelu krajowego liczonego dla ruchu dobowego przy przyjęciu założenia, że ponad 85% analizowanych odcinków musi mieć wartość GEH poniżej 11 lub 12.

Podobnie możemy ustalić graniczny GEH do oceny ruchu ciężkiego. Średni ruch ciężki to 2300 pojazdów/dobę – 1150 w jednym kierunku. Dla $GEH=5$ dopuszczalny błąd to od $-14,2\%$ do $+15,3\%$.

Zatem, chcąc stosować statystykę GEH do ocen poprawności modelu ruchu dla potoków dobowych w Polsce, należy uznać za prawidłowe odcinki, dla których GEH dla dobowych potoków ruchu ogółem w jednym kierunku jest mniejszy od 11 lub 12, natomiast przy kontroli ruchu ciężkiego dla uznania wyniku za prawidłowy GEH dla potoku dobowego ruchu ciężkiego w jednym kierunku powinien być mniejszy niż 5.

Zasada ta obowiązuje także dla potoków godzinowych. Jeśli konieczne jest sprawdzenie poprawności modelu osobno dla ruchu samochodów osobowych i ciężarowych, to kryteria GEH dla ruchu ciężarowego powinny być ostrzejsze – GEH rzędu 3 – tak, żeby średnie pomierzone wartości godzinowego ruchu ciężkiego mogły być odchyłone o ok. $+15\%$.

Podobnie wygląda sprawa z zastosowaniem statystyki GEH do oceny zgodności potoków modelowych i pomierzonych w transporcie zbiorowym. Należy sprawdzić, jak rozkładają się wielkości pomierzone i stosownie do tego dobrać wartości GEH.

Różny GEH dla potoków odcinkowych i ekranowych

Pojawia się także pytanie, na ile ma sens wprowadzanie różnych wartości granicznych współczynnika GEH dla potoków odcinkowych i ekranowych – tak jak to napisano w cytowanej SIWZ. Jak łatwo zaobserwować na rysunku 1, w miarę wzrostu wartości dopuszczalny błąd względny jest coraz mniejszy. Stąd dla ruchów ekranowych, które są sumą ruchu na kilku/kilkunastu odcinkach, z samej natury GEH dopuszczalne odchylenia będą mniejsze. Stąd nie wydaje się celowe wprowadzenie takiego dodatkowego obostrzenia.

Inne metody oceny

Zdaniem autora nie należy mieszać metod oceny. Jeśli decydujemy się na statystykę GEH, to nie stosujemy metody zmiennych warunków dla różnych przedziałów natężeń ruchu – tak jak to zrobiono w SIWZ przytoczonej we wprowadzeniu. Jeśli natomiast chcemy zastosować metodę różnych dokładności dla różnych potoków ruchu, trzeba pamiętać o dwóch bardzo istotnych elementach. Jeden to dostosowanie przedziałów do wielkości uzyskiwanych w pomiarach – np. wartościami granicznymi może być percentyl 25 i percentyl 75. Drugim istotnym elementem jest ciągłość funkcji. O tym warunku np. nie pamiętali autorzy cytowanej SIWZ. Granica lewostronna dla wartości 700 to 100, natomiast prawostronna 0,15 x 700 czyli 105. Podobnie w punkcie 2700 – granica lewostronna to 405, natomiast prawostronna 400. W tym przypadku różnica jest niewielka, ale autor opracowania natknął się na SIWZ, w której ktoś przesunął granicę wymogów z 2700 na 2500 bez zmiany warunków, i wtedy różnica była większa lewostronna 375 a prawostronna 400.

Podsumowanie

Ustalając jakiegokolwiek warunki dla poprawności wyliczonych wartości, należy określić je w sposób adekwatny do oczekiwanych wyników i ich rozkładu. Konieczne jest uwzględnienie reguł matematyki – w szczególności faktu, że jeśli funkcja określająca warunki ma różną formę dla różnych przedziałów, to konieczne jest sprawdzenie jej ciągłości. Statystyka GEH spełni warunek ciągłości. Warunek powinien być jeden – albo statystyka GEH, albo określona dokładność dla różnych przedziałów.

Zawsze należy bardzo dokładnie zdefiniować, jaki ruch ma być sprawdzany – ruch dla godziny czy dla doby, w jednym kierunku czy w obu, dla sumy ruchu czy dla kategorii pojazdów. W zależności od wyboru mogą być stawiane różne wymogi.

Warunkiem dodatkowym powinno być sprawdzenie korelacji liniowej wartości pomierzonych i obliczonych ze sprawdzeniem wskaźnika r^2 .

Literatura

1. Specyfikacja istotnych warunków zamówienia (SIWZ) w-z-rp(trm)-0712.17.1; Mazowieckiego Biura Planowania Regionalnego w Warszawie w postępowaniu o udzielenie zamówienia publicznego prowadzonym w trybie przetargu nieograniczonego na wykonanie modeli podróży w województwie mazowieckim w ramach projektu „Trendy rozwojowe Mazowsza”.
2. *Design Manual for Roads and Bridges*, volume 12; <https://www.gov.uk/guidance/standards-for-highways-online-resources#the-design-manual-for-roads-and-bridges>
3. https://en.wikipedia.org/wiki/GEH_statistic
4. http://www.gddkia.gov.pl/userfiles/articles/g/generalny-pomiar-ruchu-w-2015_15598//SYNTEZA/WYNIKI_GPR2015_DK.pdf