

Edward Kozłowski

Lublin University of Technology
Faculty of Management
ul. Nadbystrzycka 38, 20-618 Lublin, Poland
E-mail: e.kozlovski@pollub.pl

Anna Borucka

Military University Of Technology
ul. gen. Sylwestra Kaliskiego 2,
01-476, Warsaw, Poland
E-mail: anna.borucka@wat.edu.pl

Andrzej Świdorski

Motor Transport Institute
ul. Jagiellońska 80,
03-301 Warsaw, Poland
E-mail: andrzej.swiderski@its.waw.pl

ZASTOSOWANIE REGRESJI LOGISTYCZNEJ DO WYZNACZANIA MACIERZY PRAWDOPODOBIENSTW PRZEJŚĆ STANÓW EKSPLOATACYJNYCH W SYSTEMACH TRANSPORTOWYCH

Przedsiębiorstwa transportowe mogą być traktowane jako wyodrębniony pod względem technicznym, organizacyjnym, ekonomicznym i prawnym system transportowy. Zachowanie jakości i ciągłości realizacji zleceń przewozowych wymaga wysokiego poziomu gotowości pojazdów oraz personelu (szczególnie kierowców). Kontrolowanie i sterowanie realizowanymi zadaniami wspierane jest modelami matematycznymi, umożliwiającymi ocenę i określenie strategii dotyczącej podejmowanych działań. Wsparcie procesów zarządzania polega głównie na analizie sekwencji kolejnych, realizowanych czynności (stanów). W wielu przypadkach taki ciąg czynności jest modelowany za pomocą procesów stochastycznych, spełniających własność Markowa. Ich klasyczne zastosowanie możliwe jest tylko w przypadku, gdy warunkowe rozkłady prawdopodobieństwa przyszłych stanów są określone wyłącznie przez bieżący stan eksploatacyjny. Identyfikacja takiego procesu stochastycznego polega głównie na wyznaczeniu macierzy prawdopodobieństw przejść międzystanowych.

Niestety w wielu przypadkach analizowane ciągi czynności nie spełniają własności Markowa. Dodatkowo, na wystąpienie kolejnego stanu wpływa długość interwału czasowego pozostania systemu w określonym stanie eksploatacyjnym. W artykule przedstawiono metodę konstrukcji macierzy prawdopodobieństw przejść pomiędzy stanami eksploatacyjnymi. Wartości tej macierzy zależą od czasu przebywania obiektu w danym stanie. Celem artykułu było zaprezentowanie alternatywnej metody estymacji parametrów tej macierzy w sytuacji, gdy badany szereg nie spełnia własności Markowa. Wykorzystano w tym celu regresję logistyczną.

Słowa kluczowe: *regresja logistyczna, macierz prawdopodobieństw przejść, łańcuchy Markowa, system transportowy*

1. Wprowadzenie

Pojęcie systemu transportowego, za Grzywaczem i Burnewiczem [17] oraz Andrzejczakiem [3] postrzegane jest w niniejszym artykule jako uporządkowany układ trzech podsystemów: technicznego, organizacyjnego i ekonomiczno-prawnego, tworzący logiczną, zrównoważoną wewnątrznie całość, umożliwiającą osiągnięcie konkretnego celu. Pozwala to określić mianem systemu transportowego analizowane przedsiębiorstwo, a za cel jego funkcjonowania przyjąć realizację zadań przewozowych.

Systemy transportowe mogą być analizowane jako wielostanowe sekwencje kolejnych planowanych i nieplanowanych czynności obsługowych dokonywanych przez operatora systemu transportowego [27]. Konstrukcja modeli, które je opisują oraz pozwalają na wyznaczenie prognozy stanu użytkowanego obiektu, umożliwia planowanie strategii obsługi i sterowanie gotowością parku maszyn [9, 21.], pojazdów [13, 26] itp. Modelowanie funkcjonowania obiektów technicznych, za pomocą modeli deterministycznych nie zawsze jest możliwe, ponieważ na wyniki (realizacje) wpływają zaburzenia zewnętrzne (czynniki losowe), które uniemożliwiają dokładne przewidywanie kolejnych stanów. W takich przypadkach zachowanie systemów technicznych modelujemy za pomocą metod probabilistycznych, a w szczególności procesów stochastycznych. Ważną klasę procesów stochastycznych stanowią procesy Markowa. Pewne możliwości zastosowań tych procesów przedstawiono w pracach [7, 24]. Podstawą ich użycia jest spełnienie własności Markowa: przyszłość nie zależy od przeszłości, gdy znana jest teraźniejszość. W wielu analizach zakłada się a priori, że szereg czasowy spełnia tę własność i jej weryfikacja jest pomijana [55]. Jedyne nieliczni autorzy wskazują na konieczność jej sprawdzania [53] i eliminują przypadki które nie osiągnęły tego założenia [47]. Alternatywą dla systemów niespełniających własności Markowa są klasyczne metody niezawodnościowe, pozwalające na wyznaczenie empirycznych charakterystyk takich jak: strumień odnowy, funkcja odnowy, czas do następnego uszkodzenia czy intensywność strumienia odnowy [6, 14, 35] oraz obliczanie na ich podstawie głównych miar oceny systemu [15, 23, 42]. W literaturze, w ramach podobnych badań, prezentowane są różne modele [30, 43], w tym semi - Markowa [7, 24] a także wykorzystujące sztuczne sieci neuronowe [10, 33], algorytm faktoryzacji [31], drzewa błędów [52] czy modele niezawodnościowe [42, 36].

W wielu publikacjach nie jest brany pod uwagę czas pozostania w określonym stanie eksploatacyjnym. Niejednorodność interwału czasowego pomiędzy kolejnymi stanami również może powodować niespełnienie własności Markowa. W niniejszym artykule do oszacowania warunkowych prawdopodobieństw przebywania obiektu badań w poszczególnych stanach eksploatacyjnych zastosowano regresję logistyczną [49]. Regresja logistyczna opisuje związek między zmienną jakościową, a jedną lub większą liczbą zmiennych predykcyjnych [25, 46]. W literaturze regresja logistyczna stosowana jest w medycynie [5, 44], w tomografii komputerowej [46], do identyfikacji systemów technicznych [25], w obszarze finansów przedsiębiorstw [39, 54], bankowości [1, 34] i szeroko pojętych inwestycji [12, 29] oraz wykorzystywana jest do szacowania poziomu ryzyka [2, 8, 45], w badaniach społecznych, demograficznych [4, 41] i innych [25, 40]. W odniesieniu do systemów transportowych modele regresji logistycznej proponowane są przede wszystkim do oceny zagrożenia na drodze wynikającego z wypadków drogowych [18, 37], kształtowania wyborów trasy w sieci transportowej [32, 49] czy analizy wpływu wybranych czynników na realizację procesów przewozowych [38, 48].

W pracy pokazano istnienie zależności pomiędzy czasem trwania stanu eksploatacyjnego, a wartością prawdopodobieństwa przejścia do kolejnego stanu. Celem dokładnej i wyczerpującej analizy problemu w pierwszej kolejności dokonano wprowadzenia, prezentującego metody matematyczne przywoływane w artykule. W rozdziale drugim

przedstawiono definicje dotyczące łańcuchów Markowa i sposób weryfikacji własności Markowa. W rozdziale trzecim zaprezentowano metodę estymacji macierzy prawdopodobieństwa przejść za pomocą regresji logistycznej. Następnie zaprezentowano przykład wdrożenia proponowanej metody z wykorzystaniem danych empirycznych dla wybranego środka transportu, realizującego zadania przewozowe w ramach systemu transportowego (przedsiębiorstwa). W końcowym etapie dokonano omówienia otrzymanych wyników, podsumowania przeprowadzonych analiz i wskazano kierunki dalszych badań.

2. Łańcuchy Markowa

Stanem obiektu określa się jego charakterystyczną cechę, właściwość techniczną, która przyporządkowuje go do danego systemu eksploatacji [50], jest to wektor, którego składowymi są wartości fizyczne, opisujące obiekt z punktu widzenia danego badania [28]. W literaturze stan obiektu technicznego definiowany jest jako rezultat jednego i tylko jednego zdarzenia w ciągu doświadczeń spośród skończonego lub przeliczalnego zbioru parami wyłączających się zdarzeń [11, 54]. Do analizy systemów technicznych wykorzystujemy narzędzia rachunku prawdopodobieństwa i statystyki matematycznej [22, 53]. Niech (Ω, \mathcal{F}, P) będzie przestrzenią probabilistyczną, N – zbiorem liczb naturalnych, S – przestrzenią stanów analizowanego zjawiska.

Definicja 1 Rodzinę $\{X_t\}_{t \in N}$ zmiennych losowych $X_t: \Omega \rightarrow S$ dla dowolnego $t \in N$ nazywamy procesem stochastycznym z czasem dyskretnym [51, 54].

W pracy analizujemy stany eksploatacyjne, w których przebywają pojazdy. Zbiór stanów eksploatacyjnych S jest to zbiór wartości procesu stochastycznego $\{X_t\}_{t \in N}$. W każdej chwili $t \in N$ obiekt znajduje się w jednym z możliwych stanów oraz $X_t(\omega) = x_t$, tzn. w przypadku zajścia zdarzenia losowego ω w momencie t , system znajduje się w stanie $x_t \in S$. W prowadzonych badaniach przyjmujemy, że zbiór stanów S jest zbiorem skończonym oraz $S = \{s_1, s_2, \dots, s_k\}$, $k \in N$, $2 \leq k < \infty$. Wielkość $P(X_t = s_i) = p_i(t)$ oznacza prawdopodobieństwo, że system w momencie $t \in N$ znajduje się w stanie $s_i \in S$, $1 \leq i \leq k$ oraz $\sum_{i=1}^k p_i(t) = 1$.

Definicja 2 Proces stochastyczny $\{X_t\}_{t \in N}$ z czasem dyskretnym nazywa się łańcuchem Markowa jeżeli dla każdego $n \in N$, dowolnych momentów $t_1, t_2, \dots, t_n \in N$ spełniających warunek $t_1 < t_2 < \dots < t_n$ oraz dowolnych $x_1, x_2, \dots, x_n \in S$, zachodzi równość [26, 47]:

$$P(X_{t_n} = x_n | X_{t_{n-1}} = x_{n-1}, X_{t_{n-2}} = x_{n-2}, \dots, X_{t_1} = x_1) = P(X_{t_n} = x_n | X_{t_{n-1}} = x_{n-1}) \quad (1)$$

Z definicji łańcucha Markowa wynika, że rozkład warunkowy zmiennej losowej X_n dla danych wartości $X_{t_0}, X_{t_1}, \dots, X_{t_{n-1}}$ zależy tylko od ostatniej znanej wartości $X_{t_{n-1}}$. Zazwyczaj zakłada się, że odstępy pomiędzy t_i i t_{i+1} są jednakowe [16]. Poniżej przyjmujemy, że $t_n = n \in N$. Jeżeli $\{X_t\}_{t \in N}$ jest niejednorodnym łańcuchem Markowa, to dla dowolnego $t \in N$ oraz $1 \leq i, j \leq k$, wielkość:

$$P(X_t = s_j | X_{t-1} = s_i) = p_{ij}(t) \quad (2)$$

nazywamy prawdopodobieństwa przejścia ze stanu s_i w momencie $t-1$ do stanu s_j w momencie t . Zatem dla łańcuchów spełniających własność Markowa (1) warunkowe rozkłady prawdopodobieństwa przyszłych stanów procesu są zdeterminowane wyłącznie przez jego bieżący stan oraz moment t , bez względu na przeszłość (są warunkowo niezależne od stanów przeszłych). Macierz $P(t) = [p_{ij}(t)]_{1 \leq i, j \leq k}$ spełniającą warunek $\sum_{j=1}^k p_{ij}(t) = 1$ dla $t \in N$ oraz $1 \leq i \leq k$ nazywamy macierzą prawdopodobieństwa przejść w jednym kroku w chwili t [7, 47, 51].

Definicja 3 Łańcuch Markowa $\{X_t\}_{t \in N}$ jest jednorodnym łańcuchem Markowa, jeżeli prawdopodobieństwa przejścia $p_{ij}(t)$ nie zależą od momentu $t \in N$.

Zatem dla jednorodnego łańcucha Markowa $p_{ij}(t) = p_{ij}$ dla $1 \leq i, j \leq k$ oraz dowolnego momentu $t \in N$, natomiast macierz $P = [p_{ij}]_{1 \leq i, j \leq k}$ spełniającą warunek $\sum_{j=1}^k p_{ij} = 1$, $1 \leq i \leq k$ nazywamy macierzą prawdopodobieństwa przejścia w jednym kroku. Dla jednorodnego łańcucha Markowa prawdopodobieństwa przejścia ze stanu s_i w momencie t do stanu s_j w momencie $t + n$ wyznaczamy ze wzoru [13, 24]:

$$P(X_{t+n} = s_j | X_t = s_i) = p_{ij}^{(n)} \quad (3)$$

gdzie $[p_{ij}^{(n)}]_{1 \leq i, j \leq k} = P^n$, $n \in N$ jest macierzą prawdopodobieństwa przejścia w n krokach.

Definicja 4 Jeżeli $\{X_t\}_{t \in N}$ jest jednorodnym łańcuchem Markowa oraz istnieje rozkład $\pi = (\pi_1, \pi_2, \dots, \pi_k)$, gdzie $\pi_i \geq 0$, $1 \leq i \leq k$ oraz $\sum_{i=1}^k \pi_i = 1$, spełniający równanie:

$$\pi P = \pi \quad (4)$$

to rozkład π nazywamy rozkładem stacjonarnym jednorodnego łańcucha Markowa.

Własność stacjonarności oznacza, że jeżeli w pewnym momencie $n \in N$ łańcuch osiągnie rozkład stacjonarny, to dla każdej kolejnej chwili większej od n rozkład pozostanie taki sam. Rozkład stacjonarny wyznaczamy rozwiązując układ równań [16, 22]:

$$\sum_{j=1}^k \pi_j \cdot p_{ij} = \pi_i \quad (5)$$

$$\sum_{i=1}^k \pi_i = 1 \quad (6)$$

oraz $\pi_i \geq 0$ dla $1 \leq i \leq k$.

Ważną rolę w badaniu procesów, przy wykorzystaniu łańcuchów Markowa, pełnią jego własności graniczne, a szczególnie granice prawdopodobieństw $p_j(n)$ oraz $p_{ij}^{(n)}$ przy $n \rightarrow \infty$, które opisują probabilistyczne zachowanie procesu po długim czasie [16, 22].

Twierdzenie 1 (ergodyczne) Niech $\{X_t\}_{t \in N}$ będzie jednorodnym łańcuchem Markowa o skończonej liczbie stanów $k < \infty$ ($k = \#S = \#\{i: s_i \in S\}$), to:

- istnieje wektor $\pi = (\pi_1, \pi_2, \dots, \pi_k)$, taki że $\pi_i > 0$ dla $1 \leq i \leq k$ oraz $\sum_{i=1}^k \pi_i = 1$;
- dla dowolnych $1 \leq i, j \leq k$

$$\pi_j = \lim_{n \rightarrow \infty} p_{ij}^{(n)} ;$$

- wektor π jest rozwiązaniem równania (6).

Poniżej przedstawiony zostanie sposób oszacowania macierzy prawdopodobieństwa przejścia dla jednorodnego łańcucha Markowa oraz sposób weryfikacji własności Markowa. Niech $\{x_t\}_{0 \leq t \leq n}$ będzie realizacją łańcucha Markowa. Wielkość $n_i = \#\{t: x_t = s_i, 0 \leq t \leq n\}$ oznacza liczbę momentów, dla których system przebywał w stanie s_i dla $1 \leq i \leq k$, gdzie $\sum_{i=1}^k n_i = n$, natomiast wielkość $n_{ij} = \#\{t: x_t = s_i, x_{t+1} = s_j, 0 \leq t \leq n-1\}$ oznacza liczbę przejść ze stanu s_i do stanu s_j dla $1 \leq i, j \leq k$ oraz $\sum_{j=1}^k n_{ij} = n_i$. Przy założeniu, że spełniona jest własność Markowa, szacujemy macierz prawdopodobieństwa przejść. Estymator przejścia ze stanu s_i do stanu s_j wyznaczamy ze wzoru $\hat{p}_{ij} = \frac{n_{ij}}{n_i}$ dla $1 \leq i, j \leq k$.

W celu weryfikacji własności Markowa stosujemy test zgodności χ^2 . Na poziomie istotności α , $\alpha \in (0,1)$ tworzymy hipotezę roboczą:

$H_0: P(X_t = x | X_{t-1} = y, X_{t-2} = z) = P(X_t = x | X_{t-1} = y)$ (łańcuch $\{X_t\}_{t \in N}$ spełnia własność Markowa)

oraz hipotezę alternatywną:

$H_1: P(X_t = x | X_{t-1} = y, X_{t-2} = z) \neq P(X_t = x | X_{t-1} = y)$ (łańcuch $\{X_t\}_{t \in N}$ nie spełnia własności Markowa),

gdzie $x, y, z \in S$. Za miarę rozbieżności pomiędzy rozkładami $P(X_t = x | X_{t-1} = y, X_{t-2} = z)$ oraz $P(X_t = x | X_{t-1} = y)$ wybieramy statystykę testową:

$$\chi_e^2 = \sum_{i=1}^k \sum_{j=1}^k \sum_{v=1}^k \frac{(n_{ijv} - n_{ij}\hat{p}_{jv})^2}{n_{ij}\hat{p}_{jv}} \quad (7)$$

która ma rozkład χ^2 o k^3 stopniach swobody.

Wielkość $n_{ijv} = \#\{t: x_t = s_i, x_{t+1} = s_j, x_{t+2} = s_v, 0 \leq t \leq n-2\}$ oznacza liczbę przejść ze stanu s_i do stanu s_j , a następnie do stanu s_v dla $1 \leq i, j, v \leq k$. Z tablic dla rozkładu χ^2 o k^3 stopniach swobody wyznaczamy kwantyl rzędu $1 - \alpha$, który oznaczamy jako $\chi^2(1 - \alpha, k^3)$. Jeżeli $\chi_e^2 < \chi^2(1 - \alpha, k^3)$, to na poziomie istotności α nie ma podstaw do odrzucenia hipotezy roboczej H_0 , zatem przyjmujemy, że łańcuch $\{X_t\}_{t \in N}$ spełnia własność Markowa. Jeżeli natomiast $\chi_e^2 \geq \chi^2(1 - \alpha, k^3)$, to na poziomie istotności α hipotezę roboczą H_0 odrzucamy na korzyść hipotezy alternatywnej, zatem łańcuch $\{X_t\}_{t \in N}$ nie spełnia własności Markowa.

3. Regresja logistyczna

W wielu przypadkach proces stochastyczny $\{X_t\}_{t \in N}$ nie spełnia własności Markowa. Realizacja procesu $\{X_t\}_{t \in N}$ zależy od dodatkowych czynników. W systemach transportowych, logistycznych czas pozostania w określonym stanie bezpośrednio wpływa na prawdopodobieństwo przejścia do stanów pozostałych. Poniżej autorzy zastosowali regresję logistyczną do zdefiniowania macierzy prawdopodobieństwa przejścia, która zależy od czasu przebywania obiektu w danym stanie. W rozważanym przypadku, zmienna losowa $X_t, t \in N$ opisująca stan systemu, może przyjąć k możliwych realizacji. Ponieważ rozważamy momenty, dla których stan systemu zmienia się, to jeżeli w momencie $t \in N$ system znajdował się w stanie $s_i \in S$, to w momencie $t + \tau$ system może przebywać w stanach $S \setminus \{s_i\}$ (zmienna losowa $X_{t+\tau}$ może przyjąć $k - 1$ możliwych realizacji). Wyznaczenie wartości prawdopodobieństw przejść umożliwi wielomianowa regresja logistyczna [19, 25, 44, 46]. Jeden z poziomów należy przyjąć jako odniesienie. Dla każdego stanu $s_i \in S$, $1 \leq i \leq k$ wyznaczamy prawdopodobieństwa przejścia do pozostałych stanów:

$$P(X_{t+\tau} = s_j | X_t = s_i) = p_{ij}(\tau) \quad (8)$$

gdzie $t, \tau \in N$ oraz $s_j \in S \setminus \{s_i\}$. Ze zbioru $S \setminus \{s_i\}$ wybieramy stan odniesienia $s_q \in S \setminus \{s_i\}$ oraz wyznaczamy logarytm szans dla pozostałych stanów:

$$\ln \frac{P(X_{t+\tau} = s_j | X_t = s_i)}{P(X_{t+\tau} = s_q | X_t = s_i)} = \beta_{ij}^0 + \beta_{ij}^1 \tau \quad (9)$$

dla $s_j \in S \setminus \{s_i, s_q\}$. Wartości parametrów strukturalnych w modelu (9) wyznaczamy korzystając z metody największej wiarygodności [20, 25, 49]. Do oceny istotności parametrów modelu stosuje się test Walda.

Prawdopodobieństwa przejścia dla stanów $s_j \in S \setminus \{s_i, s_q\}$ wyznaczamy ze wzoru:

$$p_{ij}(\tau) = \frac{e^{\beta_{ij}^0 + \beta_{ij}^1 \tau}}{1 + \sum_{\substack{1 \leq v \leq k \\ v \neq i \\ v \neq q}} e^{\beta_{iv}^0 + \beta_{iv}^1 \tau}} \quad (10)$$

natomiast dla stanu odniesienia s_q prawdopodobieństwo wynosi:

$$p_{iq}(\tau) = \frac{1}{1 + \sum_{\substack{1 \leq v \leq k \\ v \neq i \\ v \neq q}} e^{\beta_{iv}^0 + \beta_{iv}^1 \tau}} \quad (11)$$

Ze wzorów (10)-(11) otrzymujemy, że logarytm ilorazu szans dla dwóch dowolnych poziomów $s_j, s_v \in S \setminus \{s_i, s_q\}$ jest równy:

$$\ln \frac{P(X_{t+\tau}=s_j|X_t=s_i)}{P(X_{t+\tau}=s_v|X_t=s_i)} = \ln \frac{p_{ij}(\tau)}{p_{iv}(\tau)} = (\beta_{ij}^0 - \beta_{iv}^0) + (\beta_{ij}^1 - \beta_{iv}^1)\tau. \quad (12)$$

4. Estymacja macierzy prawdopodobieństwa przejść dla wybranego środka transportu

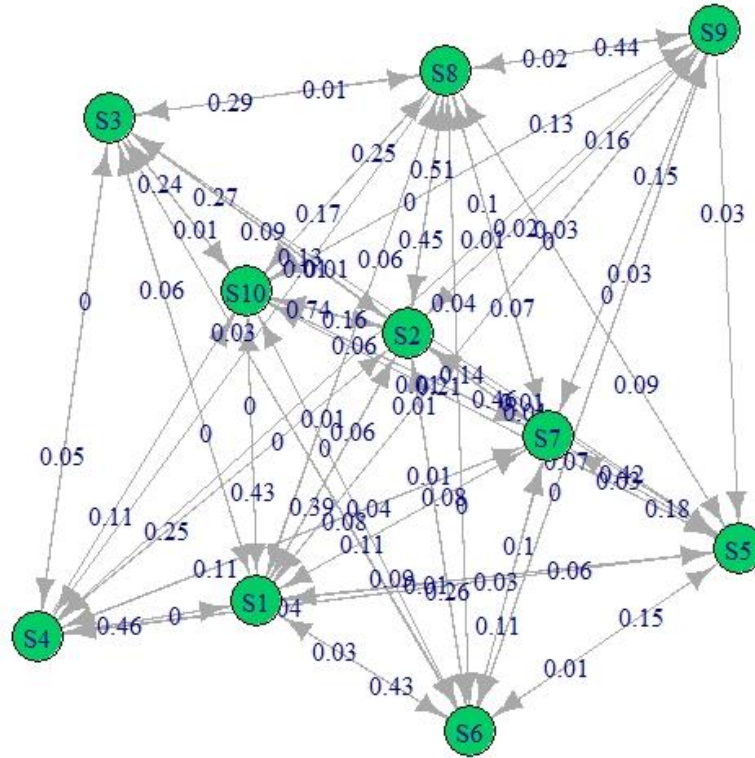
Podmiotem badania był belgijski oddział dystrybucji działający na rzecz sieci hipermarketów. Usługi transportowe realizowane są codziennie, 24 godziny na dobę, dlatego istotne jest właściwe harmonogramowanie przewozów, uwzględniające dostępność zatrudnionego personelu (szczególnie kierowców), a także gotowość pojazdów.

W badaniu wykorzystano dane pochodzące z funkcjonującego w przedsiębiorstwie systemu informatycznego do zarządzania flotą, który integruje, przetwarza i archiwizuje odczyty z zamontowanego na pojeździe nadajnika GPS, tachografu, magistrali CAN (*Controller Area Network*) oraz komputera pokładowego. Pozwala to na pozyskanie danych dotyczących kierowcy i pojazdu w czasie rzeczywistym, umożliwia śledzenie pozycji i ruchu samochodów, wizualizację lokalizacji pojazdów i naczep na mapie, monitoring czasu jazdy i odpoczynku itp. Informacje dotyczyły 69 pojazdów ciężarowych marki Iveco Stralis EEV 460. Zgromadzone dane uporządkowano oraz analizowano 10 stanów eksploatacyjnych realizowanych przez pojazdy ciężarowe. Czynności te wyszczególniono w tab. 1.

Tabela 1. Wyróżnione w badaniu stany eksploatacyjne

| Lp. | Nazwa stanu eksploatacyjnego |
|-----|------------------------------|
| S1 | Dyspozycyjność |
| S2 | Jazda |
| S3 | Czynności manipulacyjne |
| S4 | Naprawa |
| S5 | Obsługa |
| S6 | Parkowanie |
| S7 | Postój |
| S8 | Wyładunek |
| S9 | Tankowanie |
| S10 | Załadunek |

Zaprezentowane w artykule badanie przeprowadzono dla jednego, losowo wybranego pojazdu. Dokonano sprawdzenia własności Markowa. Wykorzystano w tym celu test χ^2 . Statystyka testu wyniosła 2672,74, a $p - value = 2.2 * 10^{-16}$. Oznacza to, że na poziomie istotności $\alpha = 0.001$ hipotezę roboczą należy odrzucić na korzyść hipotezy alternatywnej, zatem analizowany proces stochastyczny nie spełnia własności Markowa. Niemniej jednak oszacowano (w celach porównawczych) macierz prawdopodobieństw przejść realizacji procesu $\{X_t\}_{1 \leq t \leq n}$, $n = 6822$, którą w sposób graficzny zaprezentowano na rys.1, natomiast wartości tej macierzy przedstawiono w tab. 2



Rysunek 1. Graf przejść międzystanowych wg łańcucha Markowa

Tabela 2. Macierz prawdopodobieństw przejść dla łańcucha Markowa

| | S1 | S2 | S3 | S4 | S5 | S6 | S7 | S8 | S9 | S10 |
|-----|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| S1 | 0 | 0.392 | 0.002 | 0.002 | 0.012 | 0.032 | 0.113 | 0.009 | 0.007 | 0.431 |
| S2 | 0.060 | 0 | 0.128 | 0.003 | 0.015 | 0.015 | 0.142 | 0.453 | 0.022 | 0.162 |
| S3 | 0.065 | 0.274 | 0 | 0.003 | 0.009 | 0.029 | 0.085 | 0.294 | 0 | 0.241 |
| S4 | 0.456 | 0.246 | 0.053 | 0 | 0.035 | 0 | 0.105 | 0 | 0 | 0.105 |
| S5 | 0.056 | 0.416 | 0.011 | 0.034 | 0 | 0.146 | 0.180 | 0.090 | 0 | 0.067 |
| S6 | 0.433 | 0.264 | 0.082 | 0 | 0.014 | 0 | 0.111 | 0.005 | 0.005 | 0.087 |
| S7 | 0.084 | 0.456 | 0.003 | 0.044 | 0.027 | 0.103 | 0 | 0.074 | 0.001 | 0.208 |
| S8 | 0.062 | 0.510 | 0.008 | 0.009 | 0.005 | 0.044 | 0.097 | 0 | 0.019 | 0.247 |
| S9 | 0.035 | 0.163 | 0 | 0.012 | 0.035 | 0.035 | 0.151 | 0.442 | 0 | 0.128 |
| S10 | 0.004 | 0.735 | 0.008 | 0 | 0.012 | 0.001 | 0.064 | 0.173 | 0.003 | 0 |

Rozwiązując równanie (4) oszacowano prawdopodobieństwa graniczne. Wartości tych prawdopodobieństw przedstawia tab. 3.

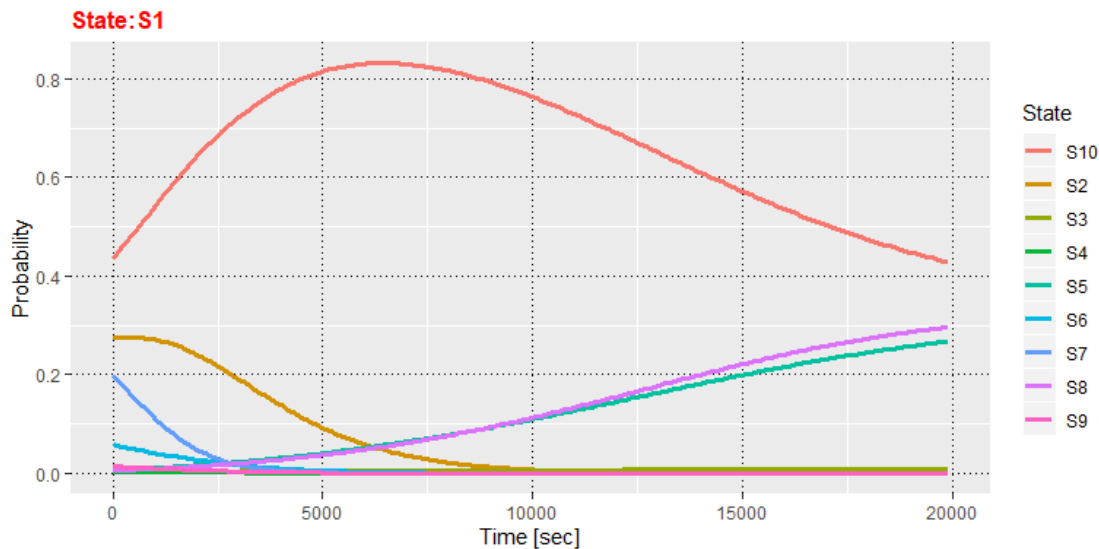
Tabela 3. Wartości granicznych prawdopodobieństw przejść dla łańcucha Markowa

| | S1 | S2 | S3 | S4 | S5 | S6 | S7 | S8 | S9 | S10 |
|---------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| π_i | 0.064 | 0.337 | 0.050 | 0.008 | 0.013 | 0.031 | 0.099 | 0.212 | 0.013 | 0.173 |

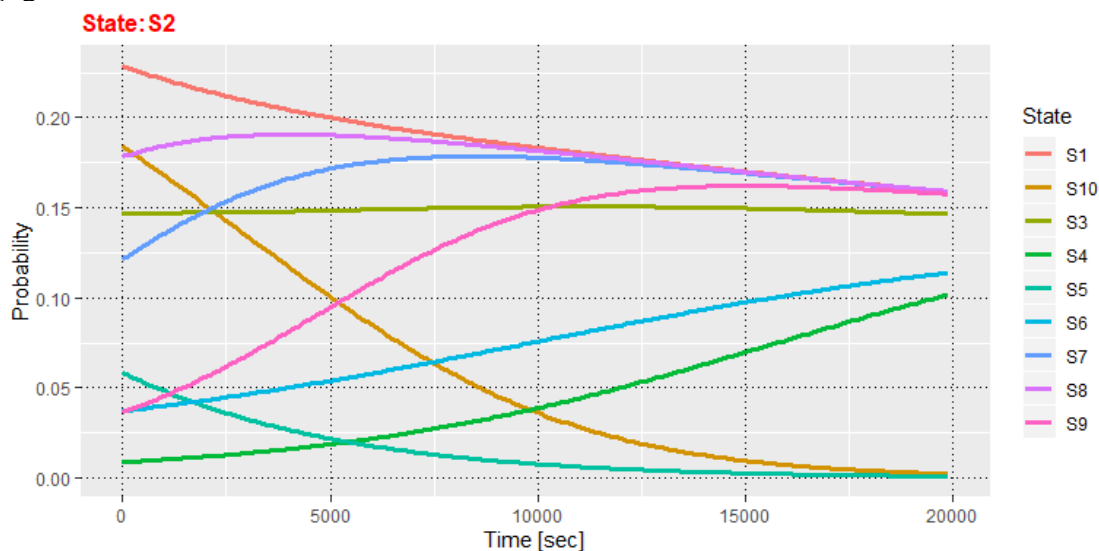
Ponieważ dla analizowanych danych własność Markowa nie została spełniona, dokonano oszacowania parametrów macierzy prawdopodobieństw przejść, wykorzystując w tym celu wielomianowy model regresji logistycznej. Zbadano wpływ czasu pobytu w określonym stanie eksploatacyjnym na prawdopodobieństwo przejścia do pozostałych

stanów. Założono, że prawdopodobieństwo w momencie $t + \tau$ jest wartością warunkowo zależną w od stanu w jakim obiekt przebywał w chwili t oraz od długości czasu τ jego trwania, a także, że po czasie τ system do niego nie powraca. Dla każdego stanu wyznaczono 8 równań regresji logistycznej postaci (9), które opisują zależności dla dziewięciu możliwych przejść. Istotność parametrów strukturalnych zbadano za pomocą testu Walda. W wyniku zastosowania wielomianowej regresji logistycznej dla każdego ze stanów systemu otrzymano macierz prawdopodobieństwa przejścia, która zależy od czasu przebywania τ .

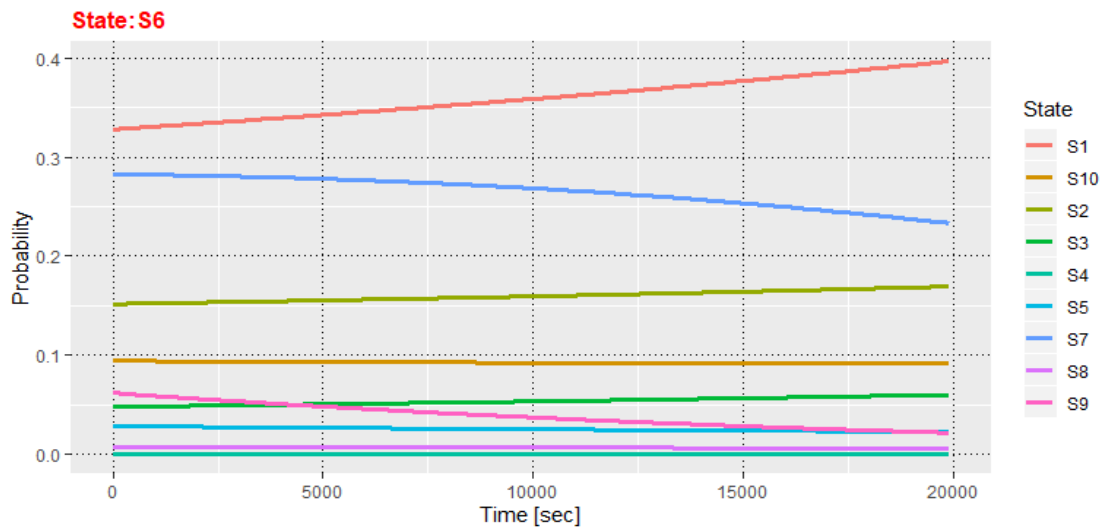
Następnie wyznaczono wykresy obrazujące zmianę prawdopodobieństwa przejścia w zależności od czasu τ pozostania w danym stanie eksploatacyjnym. Dla wybranych stanów: dyspozycyjność, jazda, parkowanie, postój, tankowanie, przedstawiono zależności prawdopodobieństwa przejść od czasu przebywania (rys. 2 – rys. 6).



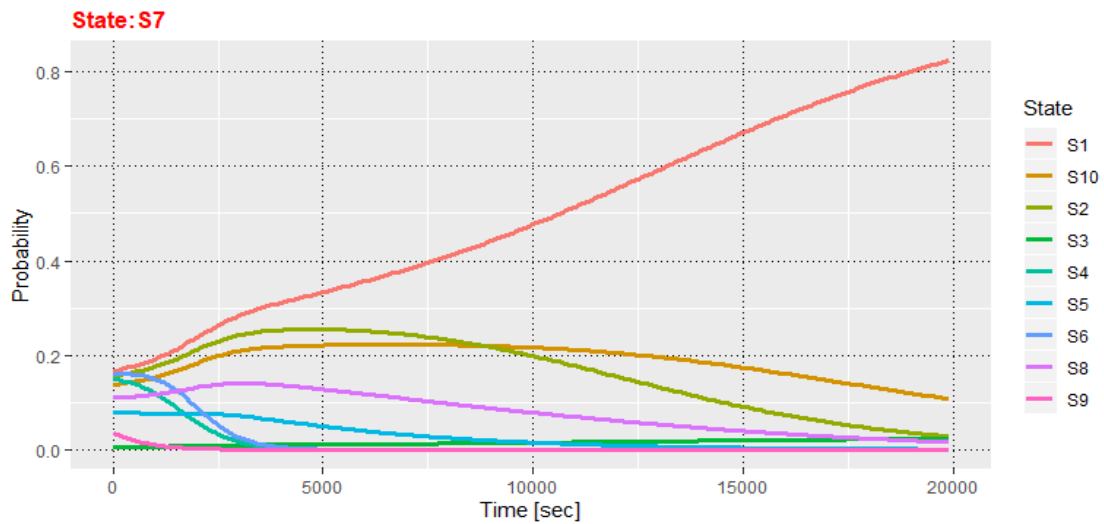
Rysunek 2. Zależność prawdopodobieństwa przejścia ze stanu Dyspozycyjność od czasu jego trwania



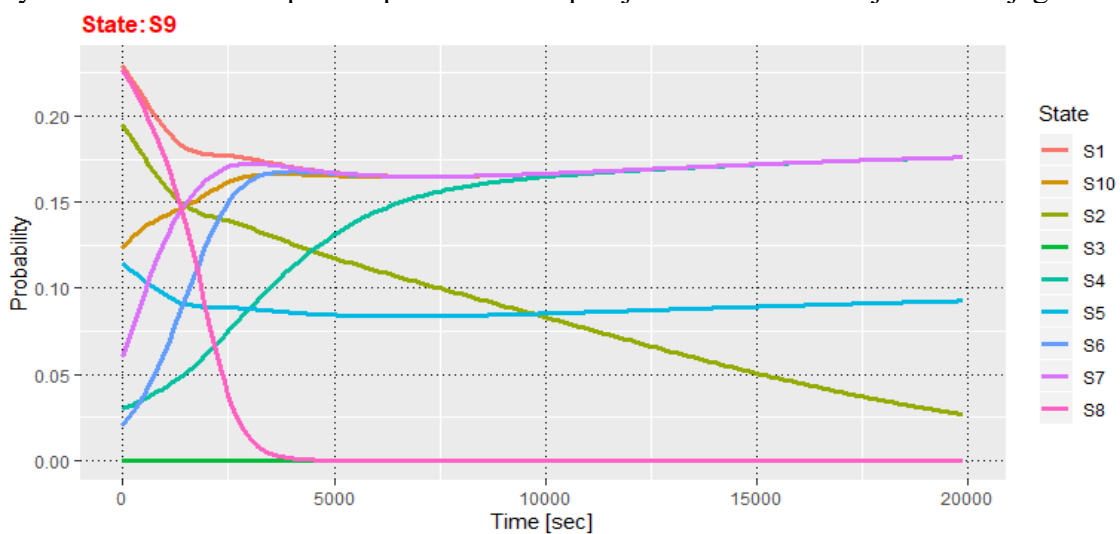
Rysunek 3. Zależność prawdopodobieństwa przejścia ze stanu Jazda od czasu jego trwania



Rysunek 4. Zależność prawdopodobieństwa przejścia ze stanu Parkowanie od czasu jego trwania



Rysunek 5. Zależność prawdopodobieństwa przejścia ze stanu Postój od czasu jego trwania



Rysunek 6. Zależność prawdopodobieństwa przejścia ze stanu Tankowanie od czasu jego trwania

Powyższe rysunki prezentują zależność wartości prawdopodobieństwa przejścia z danego stanu eksploatacyjnego do kolejnych, w zależności od czasu jaki pojazd w nim spędza. Z rysunków widzimy, że wartości tych prawdopodobieństw nie są stałe, co pokazuje niezasadność stosowania klasycznego podejścia, przy oszacowaniu macierzy prawdopodobieństw przejść jak dla łańcucha Markowa. Podejście zaproponowane przez autorów pokazuje sposób wyznaczenia macierzy przejść dla przypadku, gdy czas pozostania systemu w określonym stanie istotnie wpływa na wartości elementów tej macierzy. Zmienność wartości prawdopodobieństw przejść jest uzasadniona i oddaje specyfikę realizacji procesów transportowych, które zdeterminowane są częściowo uregulowaniami prawnymi dotyczącymi na przykład czasu pracy kierowcy, a także terminami wynikającymi z wdrożonej w przedsiębiorstwie strategii eksploatacyjnej, regulującej czasookresy napraw i przeglądów.

Przedstawione rozwiązania są pomocne dodatkowo do opracowania metody pozwalającej na ocenę gotowości systemu do realizacji zadań transportowych. Stany eksploatacyjne można sklasyfikować jako stany zdatności oraz niezdatności i wyznaczyć współczynnik gotowości technicznej jako sumę odpowiednich prawdopodobieństw stanów niezawodnościowych.

5. Zakończenie

W artykule dokonano estymacji macierzy prawdopodobieństw przejść do wyróżnionych stanów eksploatacyjnych, w których przebywał badany pojazd. Popularne w takich oszacowaniach jest zastosowanie łańcuchów Markowa, które wymaga spełnienia warunku o braku pamięci analizowanego procesu. W przedstawionym przypadku własność ta nie została spełniona. Dodatkowo pokazano, że prawdopodobieństwo przejścia do danego stanu eksploatacyjnego jest warunkowo zależne od stanu, w jakim przebywał obiekt oraz od długości czasu jaki w nim pozostawał. Dlatego do ich oszacowania zaproponowano alternatywną metodę. Wykorzystano w tym celu wielomianowy model regresji logistycznej. Otrzymano prawdopodobieństwa przejść, których wartości dla danego stanu różniły się w zależności od długości czasu pobytu pojazdu w stanie wcześniejszym.

Uzyskane wyniki porównano z wartościami otrzymanymi wg łańcucha Markowa - dla którego są one stałe - pokazując, że zastosowanie tak obliczonej macierzy prawdopodobieństw przejść, przy niespełnieniu warunku Markowa, może prowadzić do błędnych wniosków.

Zaproponowany model regresji logistycznej pozwala na krótkookresowe prognozy w zakresie realizacji procesu transportowego. Ocena prawdopodobieństwa przejścia w zależności od wcześniej realizowanych czynności stanowi wsparcie procesu harmonogramowania zadań przewozowych, a także planowania w zakresie obsługi technicznej pojazdów.

W ramach dalszych badań, zaproponowaną metodę warto rozszerzyć o wyznaczenie granicznych wartości estymatorów i ocenę prawdopodobieństw przejść dla poszczególnych stanów w długim przedziale czasu. Pozwoli to na kompleksową ocenę funkcjonowania systemu, a także na wyznaczenie poziomu gotowości do realizacji zadań transportowych. Podział stanów eksploatacyjnych na stany zdatności i niezdatności umożliwi obliczenie współczynnika gotowości technicznej jako sumy odpowiednich prawdopodobieństw granicznych stanów niezawodnościowych. Przedstawione rozwiązanie może także znaleźć zastosowanie przy modelowaniu cykli jezdnych pojazdów ciężarowych, które bezpośrednio odwzorowują rzeczywiste warunki pracy silnika, bądź też podzespołów na hamowni podwoziowej.

Bibliografia

1. Abid L, Masmoudi A, Zouari-Ghorbel S. The Consumer Loan's Payment Default Predictive Model: an Application of the Logistic Regression and the Discriminant Analysis in a Tunisian Commercial Bank. *Journal of the Knowledge Economy* 2018; 9(3): 948-962.
2. Adha M R, Nurrohmah S, Abdullah S. Multinomial Logistic Regression and Spline Regression for Credit Risk Modelling, *Journal of Physics: Conference Series* 2018; 1108.
3. Andrzejczak K. Metody prognozowania w modelowaniu eksploatacji środków transportu. Poznań: Wydawnictwo Politechniki Poznańskiej, 2013.
4. Ari E. Using Multinomial Logistic Regression to Examine the Relationship Between Children's Work Status and Demographic Characteristics, *Ekonomi ve Yönetim Araştırmaları Dergisi* 2016; 4(1): 77-93.
5. Bertens I C, Karel K G, Rutten F H, van Mourik Y, Hoes A W, Reitsma J B. A nomogram was developed to enhance the use of multinomial logistic regression modeling in diagnostic research. *Journal of Clinical Epidemiology* 2016; 71: 51-57.
6. Bobrowski D. Modele i metody matematyczne teorii niezawodności w przykładach i zadaniach. Warszawa: Wydawnictwa Naukowo – Techniczne, 1985.
7. Borucka A, Niewczas A, Hasilova K. Forecasting the readiness of special vehicles using the semi-Markov model. *Eksploatacja i Niezawodność – Maintenance and Reliability* 2019; 21 (4): 662–669.
8. Calster B V, Hoorde V K, Vergouwe Y, Bobdiwala S, Condous G., Kirk E., Bourne T, Steyerberg E W. Validation and updating of risk models based on multinomial logistic regression. *Diagnostic and Prognostic Research* 2017; 1:2.
9. Cartella F, Lemeire J, Dimiccoli L, Sahli H. Hidden Semi-Markov Models for Predictive Maintenance, *Mathematical Problems in Engineering* 2015; 278120.
10. Ferdowsi A, Challita U, Saad W. Deep Learning for Reliable Mobile Edge Analytics in Intelligent Transportation Systems: An Overview, *IEEE Vehicular Technology Magazine* 2019; 14(1): 62-70.
11. Fisz M. Rachunek prawdopodobieństwa i statystyka matematyczna. Warszawa: PWN, 1967.
12. Gautam V, Vaidya R R. Evidence on the determinants of investment-cash flow sensitivity *Indian Economic Review* December 2018; 53(1–2): 229-244.
13. Girtler J, Ślęzak M. Application of the theory of semi-Markov processes to the development of a reliability model of an automotive vehicle, *Archiwum Motoryzacji* 2012; 2: 15-27.
14. Gniedenko B.W., Bielajew J.K., Sołowiew A.D. Metody matematyczne w teorii niezawodności. Warszawa: Wydawnictwa Naukowo-Techniczne, 1965.
15. Gola A. Reliability analysis of reconfigurable manufacturing system structures using computer simulation methods, *Eksploatacja i Niezawodność – Maintenance and Reliability* 2019, 21 (1): 90-102.
16. Grabski F. Teoria semi-Markowskich procesów eksploatacji obiektów technicznych. Gdynia: Zeszyty Naukowe Wyższej Szkoły Marynarki Wojennej 75A, 1982.
17. Grzywacz W., Burnewicz J. *Ekonomika transportu*. Gdańsk: Wydawnictwa Komunikacji i Łączności, 1989.
18. Halim H, Ramli M I, Adisasmita S A, Aly S H, Prasetyonternational J A. Relationship model between accident factors and the traffic accident severity using logistic regression model, *Journal of Engineering and Science Application* 2018; 4(2), 169-181.
19. Hastie T., Tibshirani R., Friedman J., *The elements of statistical learning*, Springer-Verlag New York Inc., 2009.

20. James G., Witten D., Hastie T., Tibshirani R., *An introduction to statistical learning*, Springer-Verlag GmbH, 2013.
21. Jasiulewicz-Kaczmarek M., Gola A., *Maintenance 4.0 Technologies for Sustainable Manufacturing – an Overview*, IFAC PapersOnLine 2019; 52-10: 91-96.
22. Jaźwiński J, Grabski F. *Niektóre problemy modelowania systemów transportowych*. Radom: Instytut Technologii Eksploatacji, 2003.
23. Jia X, Cui L, Xing L. New insights into reliability problems for supply chains management based on conventional reliability model. *Eksploatacja i niezawodność – Maintenance and Reliability* 2018; 20 (3): 465–470.
24. Knopik L, Migawa K. Semi-Markov system model for minimal repair maintenance. *Eksploatacja i Niezawodność – Maintenance and Reliability* 2019; 21 (2): 256–260.
25. Kozłowski E, Mazurkiewicz D, Żabiński T, Prucnal S, Sęp J. Assessment model of cutting tool condition for real-time supervision system. *Eksploatacja i Niezawodność – Maintenance and Reliability* 2019; 21 (4): 679–685.
26. Landowski B. Example of applying markov decision process o model vehicle maintenance process. *Journal of KONES Powertrain and Transport* 2013;20(4): 209-218
27. Lewitowicz J, Zieja M, Szelmanowski A, Pazur A. New Measures of Operational Readiness for Multi-states Avionics Integrated Systems with Reduced Efficiency. In: Hanus R., Mazur D., Kreischer C. (eds) *Methods and Techniques of Signal Processing in Physical Measurements. Lecture Notes in Electrical Engineering*, 2019; 548: 165-180.
28. Lewitowicz J., Kustroń K. *Podstawy eksploatacji statków powietrznych*. Warszawa: Wydawnictwo ITWL, 2003.
29. Linhares F S, Costa F M D, Beiruth A X. Earnings management and investment efficiency. *Revista Brasileira de Gestão de Negócios* 2018; 20(2): 295-310.
30. Macián V, Tormos B, Herrero J. Maintenance management balanced scorecard approach for urban transport fleets. *Eksploatacja i Niezawodność – Maintenance and Reliability* 2019; 21 (2): 226–236.
31. Młynarski S, Pilch R, Smolnik M, Szybka J. Methodology of network systems reliability assessment on the example of urban transport. *Eksploatacja i Niezawodność – Maintenance and Reliability* 2018; 20 (2): 278–283.
32. Murat Y S, Uludag N. Route choice modelling in urban transportation networks using fuzzy logic and logistic regression methods, *Journal of Scientific & Industrial Research*; 67(01): 19-27.
33. Nabian M A, Meidani H. Deep Learning for Accelerated Seismic Reliability Analysis of Transportation Networks, *Computer-Aided Civil and Infrastructure Engineering* 2018; 33: 443–458.
34. Nguyen T. *Understanding Municipal Bond Ratings Using Ordered Multinomial Logistic Regression*. Senior Independent Study Theses 2018; 8171.
35. Nicholson A, Schmöcker J, Bell M, Iida Y. *Assessing Transport Reliability: Malevolence and User Knowledge*. In *The Network Reliability of Transport*, Emerald Group Publishing Limited 2003; 1-22.
36. Niewczas A, Rymarz J, Debicka E. Stages of operating vehicles with respect to operational efficiency using city buses as an example. *Eksploatacja i Niezawodność – Maintenance and Reliability* 2019; 21 (1): 21-27.
37. Nowakowska M. Spatial and temporal aspects of prior and likelihood data choices for Bayesian models in road traffic safety analyses. *Eksploatacja i Niezawodność – Maintenance and Reliability* 2017; 19 (1): 68–75.
38. Odeck J, Bråthen S. A meta-analysis of DEA and SFA studies of the technical efficiency of seaports: A comparison of fixed and random-effects regression models. *Transportation Research Part A: Policy and Practice* 2012; 46(10): 1574-1585.

39. Oliva F L. A maturity model for enterprise risk management, *International Journal of Production Economics* 2016; 173: 66-79.
40. Osuji M N, Mejeha R O, Simonyan J B, Gbolagun A S. Determinants of Enterprise Preference of Cassava-based farmers by Gender in Southeast, Nigeria. *Research Journal of Food and Nutrition* 2018; 2(4): PP 45-52.
41. Park M, Wu A D. Investigating differential options functioning using multinomial logistic regression. *International Journal of Quantitative Research in Education* 2017; 4(1-2): 94 - 119.
42. Paś J, Klimczak T. Selected issues of the reliability and operational assessment of a fire alarm system. *Eksplatacja i Niezawodność – Maintenance and Reliability* 2019; 21 (4): 553–561.
43. Paś J, Rosiński A. Selected issues regarding the reliability-operational assessment of electronic transport systems with regard to electromagnetic interference. *Eksplatacja i Niezawodność – Maintenance and Reliability* 2017; 19 (3): 375–381.
44. Peter D, Hart A. Multinomial Logistic Regression Examination of TV Time and Two Different Measures of Obesity in U.S. Adults. *American Journal of Public Health Research* 2018; 6(5): 222-226.
45. Raco A, Pesce A, Frascchetti F, D'Andrea G, Polli F M, Acqui M, Frati A. Risk of Postoperative Performance Status Worsening after Resection of Lesions Involving the Motor Pathway: A Multinomial Logistic Regression Model. *Journal of Neurological Surgery, Part A: Central European Neurosurgery* 2018; 79(06): 453-463.
46. Rymarczyk T.; Kozłowski E.; Kłosowski G.; Niderla K. Logistic Regression for Machine Learning in Process Tomography. *Sensors* 2019, 19, 3400.
47. Sanusi W, Jemain A A, Zin W Z, Zahari M. The Drought Characteristics Using the First-Order Homogeneous Markov Chain of Monthly Rainfall Data in Peninsular Malaysia, *Water Resources Management* 2015; 29(5): 1523–1539.
48. Schwartz J, Guasch J L, Wilmsmeier G, Stokenberga A. Logistics, Transport and Food Prices in LAC: Policy Guidance for Improving Efficiency and Reducing Costs. *Latin America and Caribbean Region Environment and Water Resources occasional paper series* 2009; 2.
49. Shao C, Xia J C, Lin T G, Goulias K G, Chen C. Logistic regression models for the nearest train station choice: A comparison of captive and non-captive stations, *Case Studies on Transport Policy* 2015; 3(4): 382-391.
50. Sowa A. Stan obiektu jako wieloznaczne pojęcie we współczesnej eksploatacji technicznej. *Czasopismo Techniczne. Mechanika* 2012; 109(7-M): 269-278.
51. Stewart W.J. Probability, Markov Chains, Queues, and Simulation, 2009.
52. Szkoda M, Kaczor G. Reliability and availability assessment of diesel locomotive using fault tree analysis. *Archives of Transport* 2016; 40(4): 65-75.
53. Wang Z, Huang X, Liang Y J. Oil-gas reservoir lithofacies stochastic modeling based on one- to three-dimensional Markov chains. *Journal of Central South University* 2018; 25(6): 1399–1408.
54. Zhu Y, Xie C, Sun B, Wang G-J, Yan X-G. Predicting China's SME Credit Risk in Supply Chain Financing by Logistic Regression, Artificial Neural Network and Hybrid Models. *Sustainability* 2016; 8(5): 1-17.
55. Ziółkowski J., Borucka A. Markov model in logistic management of enterprise. *Journal of Konbin* 2016; 38(1): 271-290.