Dr. MARIJANA GJUMBIR
TOMISLAV ŠIKIĆ dipl. prof. mat.
Fakultet prometnih znanosti
Zagreb, Vukelićeva 4

Znanost u prometu
Izvorni znanstveni rad
UDK: 656:519.65
Primljeno: 04.09.1991.
Prihvaćeno: 18.11.1991.

KRITERIJ ZA IZBOR ODABRANIH DVOPARAMETARSKIH MODELA I ANALIZA NJEGOVE PRIMJENLJIVOSTI

SAŽETAK

Autori predlažu kriterij koji omogućuje odabir jednog između devet dvoparametarskih modela - empirijskih formula. To su upravo oni modeli koji se odlikuju jednostavnošću te se zato vrlo često koriste u raznim primjenama pa tako i u svim oblicima prometa. Njihova uporaba ilustrirana je primjerima.

1. UVOD

Često su za primjene u prometu prikladni i jednostavni dvoparametarski modeli - empirijske formule. U ovom se radu razmatra devet dvoparametarskih modela kojih je zajedničko svojstvo:

- pogodnom transformacijom prevode se na linearni model. Tada se za procjenu njihovih parametara može primijeniti poznata metoda najmanjih kvadrata;
- postoji pouzdani kriterij na temelju kojega se može utvrditi koji od modela najbolje aproksimira promatrano prometno obilježje ili ustanoviti kako nijedan od tih devet modela ne zadovoljava.

Ako su mjerenjem, brojenjem ili na neki drugi način dobiveni podaci, npr. o vremenskoj ovisnosti nekoga prometnog obilježja, sređeni u tablici prema rastućim vrijednostima za t odnosno x, tada je za daljnje korištenje istih - npr. utvrđivanje empirijskog zakona pojave, predviđanje, odnosno određivanje vrijednosti obilježja koje nisu sadržane u tablici, pogodno raspolagati empirijskom formulom. Posebno pri radu s računalom - jedan jednostavni analitički izraz zamjenjuje i glomazne tablice podataka.

2. TEORIJSKE OSNOVE

Pretpostavlja se da je poznata tablica mjernih podataka za neko prometno obilježje u kojoj su podaci sređeni prema rastućim vrijednostima x. Prometno je obilježje označeno s y. Poznata je dakle tablica oblika

x	x_1	x_2	...	x_n
y	y_1	y_2	...	y_n

(*)

uz $x_1 < x_2 < x_3 < \dots < x_n$

Zadaća je naći jednostavni dvoparametarski model koji će adekvatno aproksimirati ove podatke.

Osim linearne funkcije

$$y = a x + b \tag{1}$$

kao najpoznatijeg i najjednostavnijega dvoparametarskog modela, razmatra se sljedećih osam empirijskih formula oblika $y = f(x; a, b)$:

Eksponencijalna funkcija

$$y = a e^{bx}, \quad a > 0 \tag{2}$$

može se koristiti i u obliku

$$y = ab^x$$

odnosno

$$y = a\, 10^{bx}$$

Odabrane su i racionalna funkcija

$$y=\frac{1}{ax+b} \tag{3}$$

logaritamska

$$y=a\ln x+b \tag{4}$$

te funkcija potencije

$$y = a\,x^b,\ a>0 \tag{5}$$

Idući je model

$$y=\frac{1}{a\ln x+b} \tag{6}$$

zatim hiperbolička funkcija

$$y=\frac{b}{x}+a \tag{7}$$

te eksponencijalna oblika

$$y= a\,e^{b/x} \tag{8}$$

i na kraju

$$y=\frac{x}{ax+b} \tag{9}$$

racionalna funkcija specijalnog oblika.

U daljnjem tekstu uvijek će modeli biti označeni brojevima 1 - 9 kao što su već obilježeni.

Tablica 1. sadrži transformacije kojima se navedene funkcije prevode u linearnu funkciju Y = K X + L, kao i veze između parametara a, b, K i L u relacijama 1 - 9.

Npr. iz modela (5) logaritmiranjem se dobiva

$$\ln y = \ln a + b \ln x$$

pa se na oblik Y = K X + L prevodi (5) zamjenom

$$Y = \ln y \quad X = \ln x$$

$$K = b \quad L = \ln a$$

Parametre K i L može se sada procijeniti nekom od poznatih metoda - metodom najmanjih kvadrata, metodom parcijalnih zbrojeva, odnosno odabranih točaka (1), (2).

Povratnom zamjenom

$$a = e^L$$

$$b = K$$

dobivaju se jednostavno i vrijednosti procjena traženih parametara u modelu (5). Analognim postupcima postižu se tako i ostale veze u tablici 1.

Tablica 1. Dvoparametarski modeli i transformacije za linearizaciju

Redni broj	Model	Transformacija	Veze između parametara
1	$y=ax+b$	$Y=y$ $X=x$	$K=a$ $L=b$
2	$y=ae^{bx}$	$Y=\ln y$ $X=x$	$K=b$ $L=\ln a$
3	$y=\frac{1}{ax+b}$	$Y=1/y$ $X=x$	$K=a$ $L=b$
4	$y=a\ln x+b$	$Y=y$ $X=\ln x$	$K=a$ $L=b$
5	$y=ax^b$	$Y=\ln y$ $X=\ln x$	$K=b$ $L=\ln a$
6	$y=\frac{1}{a\ln x+b}$	$Y=1/y$ $X=\ln x$	$K=a$ $L=b$
7	$y=\frac{b}{x}+a$	$Y=y$ $X=1/x$	$K=b$ $L=a$
8	$y=ae^{b/x}$	$Y=\ln y$ $X=1/x$	$K=b$ $L=\ln a$
9	$y=\frac{x}{ax+b}$	$Y=1/y$ $X=1/x$	$K=a$ $L=b$

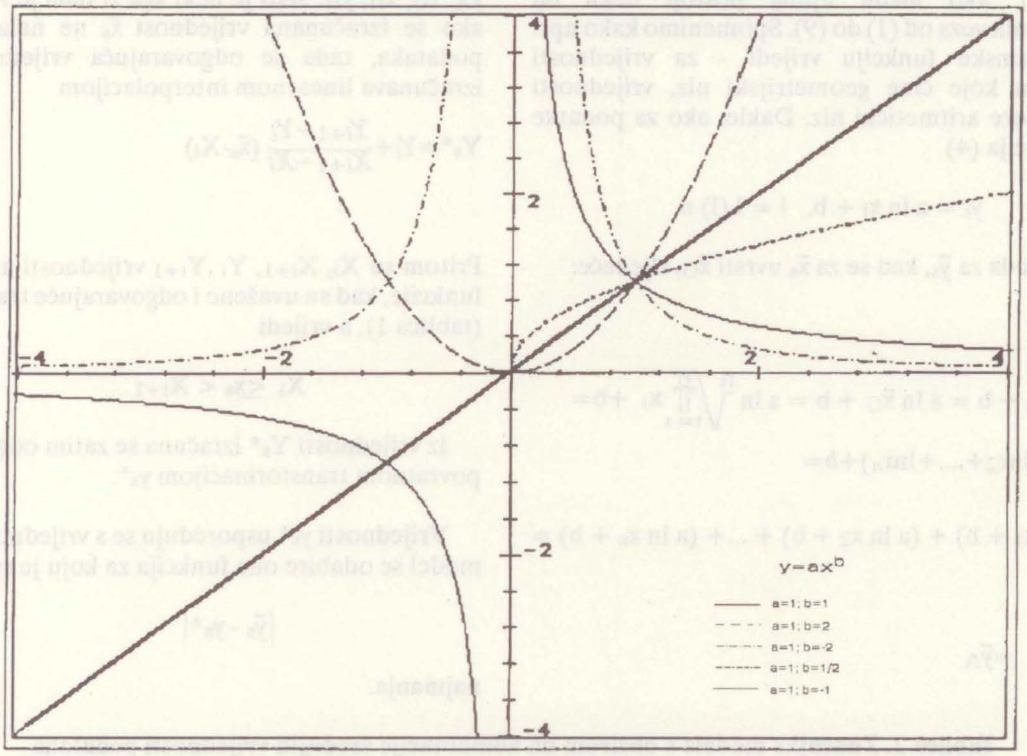
3. KRITERIJI ODABIRA MODELA

Na slici 1. predočeni su grafovi funkcije $y = a\,x^b$ za neke cjelobrojne i racionalne vrijednosti parametra b te pozitivne vrijednosti parametra a - točnije parametar a ima uvijek konstantnu vrijednost a = 1.

Očevidno je da nije jednostavno vizualno, tj. samo im je iskustvo pomoglo i u konačnom postavljanju kriterija za njihov izbor.

Predloženi kriterij temelji se na glavnim svojstvima promatranih funkcija. Tako je npr. dobro poznato svojstvo linearne funkcije - ako vrijednosti argumenata, tada i vrijednosti funkcije čine aritmetički niz.

Ako s $\bar{x}_A$ i $\bar{y}_A$ označimo aritmetičke sredine podataka u tablici (*) tj.



Slika 1.

iz grafičke reprezentacije podataka, zaključiti ili samo naslutiti koji će ih model adekvatno aproksimirati.

Druga je mogućnost za svaki od 9 modela procijeniti parametre a, b, te izračunati zbroj kvadrata odstupanja

$$s^2 = \sum_{i=1}^{n} (y_i - y_t)^2$$

gdje y_i označuje eksperimentalni podatak, a y_t izračunanu vrijednost s pomoću modela za odgovarajući argument x_i. Ovaj postupak zahtijeva opsežan posao i uz primjenu računala. Zato su autori ispitivali druge egzaktnije kriterije za odabir optimalnih modela. Nakon detaljne analize, uočivši zajedničke značajke predloženih devet modela, vlastito

$$\bar{x}_A = \frac{\sum_{i=1}^{n} x_i}{n} \qquad \bar{y}_A = \frac{\sum_{i=1}^{n} y_i}{n} \tag{10}$$

a s $\bar{x}_g$, $\bar{y}_G$ geometrijske sredine istih

$$\bar{x}_G = \sqrt[n]{\prod_{i=1}^{n} x_i} \qquad \bar{y}_G = \sqrt[n]{\prod_{i=1}^{n} y_i} \tag{11}$$

te s $\bar{x}_H$, $\bar{y}_H$ harmonijske sredine

$$\bar{x}_H = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}}$$

$$\bar{y}_H = \frac{n}{\sum_{i=1}^{n} \frac{1}{y_i}} \qquad (12)$$

tada poznavajući svojstva aritmetičkog, geometrijskog odnosno harmonijskog niza, može se ustanoviti povezanost između odgovarajućih sredina podataka u tablici (*), ako među njima postoji neka od funkcionalnih veza od (1) do (9). Spomenimo kako npr. za logaritamsku funkciju vrijedi - za vrijednosti argumenata koje čine geometrijski niz, vrijednosti funkcije tvore aritmetički niz. Dakle, ako za podatke vrijedi relacija (4)

$$y_i = a \ln x_i + b, \quad i = 1\,(1)\,n$$

dobiva se tada za $\bar{y}_s$, kad se za $\bar{x}_s$ uvrsti $\bar{x}_G$, sljedeće:

$$\bar{y}_s = a \ln \bar{x}_s + b = a \ln \bar{x}_G + b = a \ln \sqrt[n]{\prod_{i=1}^{n} x_i} + b =$$

$$= \frac{a}{n}(\ln x_1 + \ln x_2 + \dots + \ln x_n) + b =$$

$$= \frac{1}{n}\left[(a \ln x_1 + b) + (a \ln x_2 + b) + \dots + (a \ln x_n + b)\right] =$$

$$= \frac{1}{n}\sum_{i=1}^{n} y_i \quad = \bar{y}_A$$

Znači da geometrijskoj sredini argumenata odgovara aritmetička sredina vrijednosti funkcije, ako između tih podataka postoji funkcionalna ovisnost (4). Analogno se mogu izvesti i ostali odgovarajući odnosi u tablici 2.

Ova tablica sadrži pregled svih kombinacija srednjih vrijednosti i prijedloge za odabir najpogonijeg modela.

Kriterij se primjenjuje tako da se s pomoću podataka iz tablice (*) izračunaju vrijednosti sredina $\bar{x}_s$, $\bar{y}_s$: $\bar{x}_A$, $\bar{y}_A$, $\bar{x}_G$, $\bar{x}_H$, $\bar{y}_H$. Ako je neki $\bar{x}_s = x_i$ tada je i $y_{s^*} = y_i$, ali ako se izračunana vrijednost $\bar{x}_s$ ne nalazi u tablici podataka, tada se odgovarajuća vrijednost za y_s^* izračunava linearnom interpolacijom

$$Y_s^* = Y_i + \frac{Y_{i+1} - Y_i}{X_{i+1} - X_i}(\bar{x}_s - X_i) \qquad (13)$$

Pritom su X_i, X_{i+1}, Y_i, Y_{i+1} vrijednosti argumenata i funkcije, kad su uvažene i odgovarajuće transformacije (tablica 1), a vrijedi

$$X_i \leq x_s < X_{i+1}$$

Iz vrijednosti Y_s^* izračuna se zatim odgovarajućom povratnom transformacijom y_s^*.

Vrijednosti y_s^* uspoređuju se s vrijednostima $\bar{y}_s$. Za model se odabire ona funkcija za koju je razlika

$$\left|\bar{y}_s - y_s^*\right| \qquad (14)$$

najmanja.

Tablica 2. Značajke modela s obzirom na kombinacije srednjih vrijednosti podataka

Redni broj	$\bar{x}_s$	$\bar{y}_s$	Optimalni model
1	$\bar{x}_A$	$\bar{y}_A$	$y = ax + b$
2	$\bar{x}_A$	$\bar{y}_G$	$y = ae^{bx}$
3	$\bar{x}_A$	$\bar{y}_H$	$y = \frac{1}{ax+b}$
4	$\bar{x}_G$	$\bar{y}_A$	$y = a\ln x + b$
5	$\bar{x}_G$	$\bar{y}_G$	$y = ax^b$
6	$\bar{x}_G$	$\bar{y}_H$	$y = \frac{1}{a\ln x + b}$
7	$\bar{x}_H$	$\bar{y}_A$	$y = \frac{b}{x} + a$
8	$\bar{x}_H$	$\bar{y}_G$	$y = ae^{b/x}$
9	$\bar{x}_H$	$\bar{y}_H$	$y = \frac{x}{ax+b}$

Ako je ta razlika, u odnosu na vrijednosti y_i, velik broj, tada nijedan od 9 razmatranih modela nije prikladan za aproksimaciju podataka.

Autori su programsku podršku za ovaj kriterij sastavili u programskom jeziku GfA BASIC za računalo Mega ATARI i ispitali najprije tako da su za svaki model odabrali točne funkcije. Tako je npr. za model (3) jedna od razmatranih funkcija bila

$$y=\frac{1}{2x+1}$$

te tablica podataka

x_i	1	2	3	4	5	6
y_i	1/3	1/5	1/7	1/9	1/11	1/13

Očevidno je da u ovom slučaju vrijednosti x čine aritmeitčki niz, pa vrijedi

$$x_k=\frac{x_{k-1}+x_{k+1}}{2}$$

dok za vrijednosti y vrijedi

$$y_k=\frac{2}{\frac{1}{y_{k-1}}+\frac{1}{y_{k+1}}}$$

odnosno

$$y_k=\frac{(2y_{k-1})(y_{k+1})}{y_{k-1}+y_{k+1}}$$

Za funkciju $y=\frac{1}{2x+1}$ i spomenutu tablicu, dobivene su ove numeričke vrijednosti (zaokružene na 4 decimale):

Model	$\overline{x}_s$	$\overline{y}_s$	y_s^*	$\lvert\overline{y}_s-y_s^*\rvert$
1	3.5	0.1592	0.1250	0.0342
2	3.5	0.1396	0.1250	0.0146
3	3.5	0.1250	0.1250	0.0000
4	2.9938	0.1592	0.1431	0.0161
5	2.9938	0.1396	0.1431	0.0035
6	2.9938	0.1250	0.1431	0.0181
7	2.4490	0.1592	0.1696	0.0104
8	2.4490	0.1396	0.1696	0.0300
9	2.4490	0.1250	0.1696	0.0446

Vrijednost $\overline{x}_s=\overline{x}_A=3.5$ nalazi se između vrijednosti $x=3$ i $x=4$ i nije sadržana u zadanoj tablici podataka. Zato se pripadni y_s^* računa prema (13)

$$Y_s^*=7+\frac{9-7}{4-3}(3.5-3)$$

$$Y_s^*=8$$

te je radi

$$Y=\frac{1}{y}$$

$$y_s^*=1/8=0.125=y_A^*$$

Analogno je popunjena cijela tablica. Najmanja vrijednost razlike

$$\lvert\overline{y}_s-y_s^*\rvert=0.0000$$

sugerira izbor modela (3), što je u ovom slučaju i očevidno.

Autori su ispitali ukupno četiri kriterija:

"Kriterij I." - se temelji na trima vrijednostima iz tablice zadanih podataka. Odabiru se tri para podataka: početni (x_1, y_1), neki srednji (x_s, y_s) i krajnji (x_n, y_n). Umjesto relacije (13) primjenjuje se linearna interpolacija. Takav je kriterij razmatran u knjizi (3) i primijenjen na 7 od ispitivanih 9 modela. No, već su i njegovi autori istaknuli kako se radi samo o vrlo gruboj procjeni, budući da se ne uzimaju u obzir svi podaci. Zato smo "proširili" ovaj kriterij na sve podatke i dobili "kriterij III". Ustanovili smo da također ne zadovoljava radi primjene linearne interpolacije. Kod "kriterija II". uvrstili smo stvarne funkcijske veze i transformacije prema tablici 1, ali smo računali srednje vrijednosti samo s tri para vrijednosti. Ni taj se kriterij nije pokazao uvijek pouzdanim.

Na kraju smo došli i do vlastitoga "kriterija IV"., ovdje predloženog, kod kojeg se uzimaju u obzir svi podaci i relacija (13) za interpolaciju. Također je uočeno i razmotreno svih devet mogućih kombinacija srednjih vrijednosti.

Na primjeru funkcije $y=5e^{3x}$ ilustrirana je jedna primjena spomenutih 4 kriterija, za x = 1 (1) 6. Odmah se može uočiti nepouzdanost I. i III. kriterija iz tablice razlika.

Model	Kriterij I.	Kriterij II.
1	163722828.19	163722828.19
2	245567.18	2.24×10^{-5}
3	426943.83	76987.03
4	164130651.15	164128706.59
5	162255.79	172537.30
6	19120.87	3642.10
7	164148503.36	164148275.18
8	180108.00	180354.05
9	1268.66	281.61

Model	Kriterij III.	Kriterij IV.
1	57156411.83	57156411.83
2	245567.18	$\boxed{4.91 \times 10^{-5}}$
3	426572.12	76615.31
4	57543279.98	57543237.69
5	141300.97	141678.01
6	39703.98	36345.01
7	57564254.44	57560365.33
8	162275.43	171074.29
9	18729.52	3652.86

4. PRIMJENA

Na tri konkretna primjera ilustrirana je primjena opisanoga kriterija. Radi se o podacima koji su dobiveni brojenjem, mjerenjem, pa prema tomu među njima ne postoje prave funkcijske veze, već treba pronaći adekvatne empirijske formule - modele.

Primjer 1. Kretanje opremljenosti telefonima može se pratiti prema broju telefona na tisuću stanovnika. Prema podacima (4), (5), (6) sastavljena je tablica za Jugoslaviju, za razdoblje 1975-1988. godina.

Godina	Broj telefona na tisuću stanovnika
1975.	61
1976.	66
1977.	71
1978.	79
1979.	86
1980.	96
1981.	103
1982.	112
1983.	123
1984.	132
1985.	144
1986.	155
1987.	167
1988.	180

Da bi se utvrdilo je li trend rasta jedan od 9 dvoparametarskih modela, testirani su podaci prema opisanom kriteriju.

Dobivena je sljedeća tablica razlika:

Redni broj modela	(y_s-y_s*)
1	5.
2	$\boxed{1.1358281824}$
3	7.0359064402
4	16.154271349
5	9.9356587783
6	3.9524036927
7	31.015623266
8	24.853310899
9	18.925673005

Može se zaključiti da će eksponencijalni trend

$$y = a\,e^{bx}$$

adekvatno aproksimirati ove podatke.

Metodom najmanjih kvadrata dobiva se za parametre procjene:

$$a = 56.35866113$$
$$b = 0.08456621$$

a veličina zbroja kvadrata odstupanja jednaka je

$$\sum_{i=1}^{n}(y_i-y_t)^2 = 41.130954714$$

te potvrđuje adekvatnost odabranog trenda.

Ovaj se trend može prihvatiti npr. za prognoziranje tendencije rasta, utvrđivanje "zakonitosti" razvitka telefonije i slično.

Primjer 2. Broj registriranih putničkih automobila u Hrvatskoj prema izvoru (6) kretao se u razdoblju 1970-1988. godina kao u tablici:

Godina	Broj registriranih putničkih automobila u Hrvatskoj
1970.	184 068
1971.	222 737
1972.	256 546
1973.	286 589
1974.	324 008
1975.	368 788
1976.	405 471
1977.	463 606
1978.	521 722
1979.	549 258
1980.	584 173
1981.	633 753
1982.	679 162
1983.	652 705
1984.	679 317
1985.	652 625
1986.	712 795
1987.	716 268
1988.	756 119

Tablica razlika

Model	Razlika
1.	41378.53
2.	81161.60
3.	125985.55
4.	48157.51
5.	8621.93
6.	35934.76
7.	166035.34
8.	126943.63
9.	82797.67

pokazuje da je bolje potražiti neki drugi dvoparametarski model, odnosno troparametarski.

Primjer 3. Za neko prometno obilježje dobiveni su mjerni podaci kao u tablici

x_i	13	13.4	16	23	26	34
y_i	1.0	1.3	2.4	8.2	11.0	26.0

Tablica razlika u ovom slučaju

Model	Razlike
1	1.85667
2	1.26584
3	2.29180
4	2.67440
5	0.36370
6	1.50774
7	3.39612
8	0.31254
9	1.00330

sadrži dvije razlike koje su mali brojevi, ali se međusobno malo razlikuju. Takva razlika može nastati i uslijed zaokruživanja tijekom analize, pa se preporuča, npr. metodom najmanjih kvadrata, provesti daljnje ispitivanje.

Tako se za model 5 dobivaju vrijednosti

$$a = 0.00022752413387$$
$$b = 3.317556346$$
$$s^2 = 2.5747959365$$

a za model 8

$$a = 149.88370562$$
$$b = -65.031042533$$
$$s^2 = 17.087939742$$

Zato se ipak može dati prednost modelu 5, tj. empirijskoj formuli

$$y = a\,x^b$$

5. ZAKLJUČAK

U ovom radu opisani kriterij predstavlja posebno pogodno sredstvo u praksi.

Omogućuje:

- izabrati najprikladniji između devet ponuđenih modela,

- utvrditi kako nijedan od devet modela nije pogodan za aproksimaciju podataka,

- zaključiti da treba provesti daljnje ispitivanje (npr. metodom najmanjih kvadrata), ako su razlike u (14) mali brojevi koji se međusobno vrlo malo razlikuju. No ni u ovom slučaju ne procjenjuju se parametri za svih devet modela te zbroj kvadrata odstupanja, već samo za dva do tri od njih, što znatno skraćuje analizu,

- ustanoviti da među promatranim obilježjima postoji i empirijska zakonitost, pa model postaje "zakon".

Istodobno je ovaj kriterij dostatno jednostavan, provediv i primjenom džepnih programirajućih računala. Uz stanovito iskustvo bit će pouzdano oruđe u rukama inženjera koji se često sreće s problemima modeliranja, ispitivanja tendencija razvoja, odnosno optimiranja.

SUMMARY

CRITERION FOR THE SELECTION OF TWO-PARAMETER MODELS AND REVIEW OF THEIR USABILITY

The authors set forth the criterion (i.e. the standard) which makes it possible to select one out of nine two-parameter models, i.e. empiric equations. This refers us to the very simple models which are rather frequently used in different applications including all aspects of traffic. Their utilization has been illustrated by examples.

LITERATURA

[1] **J.H. MATHEWS:** Numerical Methods for Computer Science, Engineering and Mathematics. Englewood Cliffs, N.J., Prentice Hall, 1987.

[2] **W. BOEHM, G. GOSE, J. KAHMANN:** Methoden der Numerischen Mathematik. Braunschweig, Vieweg, 1985.

[3] **B.P. DEMIDOVIČ, I.A. MARON, E.Z. ŠUVALOVA:** Čislennye metody analiza. Moskva, Nauka, 1967.

[4] Poslovni izvještaj i statistika Zajednice JPTT. Beograd, publikacije po godinama izdanja, od 1979. do 1989. godine.

[5] Zbornik PTT prometa '87. Zagreb, Fakultet prometnih znanosti, 1988.

[6] Statistički godišnjak Jugoslavije, Beograd, Savezni zavod za statistiku, 1989.