

**UNIVERZITA KOMENSKÉHO V BRATISLAVE  
FAKULTA MATEMATIKY, FYZIKY  
A INFORMATIKY**

**MODELOVANIE VEKU ÁUT V PREVÁDZKE**

**Bakalárska práca**

**2011**

**Andrej Horský**

**UNIVERZITA KOMENSKÉHO V BRATISLAVE  
FAKULTA MATEMATIKY, FYZIKY  
A INFORMATIKY**

Evidenčné číslo: db1f525c-b788-4fe2-80c1-661bacb88dfd

**MODELOVANIE VEKU ÁUT V PREVÁDZKE**

**Bakalárska práca**

Študijný program : Ekonomická a finančná matematika

Študijný odbor: 9.1.9 Aplikovaná matematika

Školiace pracovisko: Katedra aplikovanej matematiky a štatistiky

Školiteľ: Mgr. Richard Kollár, PhD.

**Bratislava 2011**

**Andrej Horský**



Univerzita Komenského v Bratislave  
Fakulta matematiky, fyziky a informatiky

## ZADANIE ZÁVEREČNEJ PRÁCE

**Meno a priezvisko študenta:** Andrej Horský  
**Študijný program:** ekonomická a finančná matematika (Jednoodborové štúdium, bakalársky I. st., denná forma)  
**Študijný odbor:** 9.1.9. aplikovaná matematika  
**Typ záverečnej práce:** bakalárska  
**Jazyk záverečnej práce:** slovenský

**Názov :** Modelovanie veku áut v prevádzke

**Cieľ :** Zostavenie matematického modelu veku používaných áut pomocou systému obyčajných diferenciálnych rovníc.

**Anotácia :** Populačné matematické modely sú bežne používané v demografii. Cieľom práce je zostaviť podobný model na vek áut a skúmať možný súvis s nehodovosťou.

**Vedúci :** Mgr. Richard Kollár, PhD.

**Dátum zadania:** 27.10.2010

**Dátum schválenia:** 08.11.2010

.....  
doc. RNDr. Margaréta Halická, CSc.  
garant študijného programu

.....

študent

.....

vedúci práce

Dátum potvrdenia finálnej verzie práce, súhlas s jej odovzdaním (vrátane spôsobu sprístupnenia)

.....  
vedúci práce

## **ČESTNÉ VYHLÁSENIE**

Čestne vyhlasujem, že som túto bakalársku prácu vypracoval samostatne za pomoci konzultácií a uvedených bibliografických odkazov.

V Bratislave, jún 2011

.....  
Andrej Horský

## **POĎAKOVANIE**

Touto cestou by som sa chcel poďakovať vedúcemu bakalárskej práce, Mgr. Richardovi Kollárovi, PhD. , za odbornú pomoc a cenné rady, ktoré mi poskytol pri jej vypracovávaní. V neposlednom rade ďakujem aj mojej rodine, priateľke a spolubývajúcim za neustálu podporu a porozumenie.

## **ABSTRAKT**

HORSKÝ, Andrej: Modelovanie veku áut v prevádzke. [Bakalárska práca] – Univerzita Komenského v Bratislave. Fakulta matematiky, fyziky a informatiky; Katedra aplikovanej matematiky a štatistiky. – Vedúci: Mgr. Richard Kollár, PhD., Bratislava. 2011, s. 43

Cieľom bakalárskej práce je vytvorenie matematického a štatistického modelu, ktorý bude slúžiť na predpovedanie vekového rozloženia automobilov v dopravnej prevádzke. Najprv predstavujeme pohľad na existujúci matematický model vytvorený Alanom Greenspanom a Darrel Cohenom, ktorý sa zaoberá podobnou problematikou a rozoberáme techniky a výsledky tejto štúdie. Následne vytvárame a analyzujeme nový matematický model založený na odhade pravdepodobnosti vyradenia automobilu z prevádzky pomocou diskrétného populačného modelu s odhadom parametra, ktorý vypočítame dvoma spôsobmi. Nakoniec porovnáваме výsledky oboch prístupov a formulujeme predikcie.

Kľúčové slová : matematické modelovanie, dopravná prevádzka, vek automobilu, diskretný populačný model

## **ABSTRACT**

HORSKÝ, Andrej: Mathematical modeling of vehicle stock. [Bachelor thesis] Comenius University in Bratislava. Faculty of Mathematics, Physics and Informatics; Department of Applied Mathematics and Statistics. – Tutor: Mgr. Richard Kollár, PhD., Bratislava. 2011, p. 43

The aim of the bachelor thesis is generation of mathematical and statistical model that can be used as a forecasting tool of age distribution of motor vehicles in service. First, we present an overview of an existing mathematical model by Greenspan and Cohen that study a related problem and discuss its techniques and results. Then we design and analyze a new mathematical model based on the estimation of scrappage rate using a discrete population model with parameter fitting performed in two different ways. Finally, we compare the results of both approaches and formulate predictions.

Key words : mathematical modelling, vehicle stock, discrete population model

## Obsah

Úvod.....	10
<b>1. Analýza práce Greenspana a Cohena.....</b>	<b>12</b>
1.1. Obsah práce.....	12
1.1.1. Zber a úprava dát.....	13
1.1.2. Modelovanie stavu prevádzky.....	15
1.1.3. „Engineering scrappage (ES)“.....	16
1.1.4. „Cyclical Scrappage“.....	19
1.1.5. Explicitný príklad.....	20
<b>2. Analýza metód a výsledkov práce.....</b>	<b>21</b>
<b>3. Model.....</b>	<b>26</b>
3.1. Položenie otázok.....	26
3.2. Matematické metódy.....	26
3.2.1. Analýza dát z grafu.....	26
3.2.2. Metóda najmenších štvorcov.....	26
3.3. Analýza dát.....	27
3.3.1. Staršie automobily.....	30
3.3.1.1. Prvý prístup.....	30
3.3.1.2. Druhý prístup.....	34
3.3.2. Novšie automobily.....	35
3.3.2.1. Prvý prístup.....	35
3.3.2.2. Druhý prístup.....	38
3.4. Výsledky a ich interpretácia.....	39
3.4.1. Porovnanie oboch prístupov s reálnymi dátami.....	39
3.4.2. Predikcia vekového rozloženia do ďalších rokov.....	40
<b>Záver.....</b>	<b>42</b>
<b>Zoznam použitej literatúry.....</b>	<b>43</b>



## ZOZNAM ILUSTRÁCIÍ A TABULIEK

<b>Obrázok 1:</b> Počet automobilov zoradených podľa veku (referencia : [7, str. 3-9]) .....	13
<b>Obrázok 2:</b> Percentuálne vyjadrenie životnosti automobilov (referencia : [1, str.32]).....	17
<b>Obrázok 3 :</b> Odhady koeficientov, štandardná odchýlka, hodnota $t$ – štatistiky, Beta koeficient a elasticita (referencia : [1, str.31]) .....	18
<b>Obrázok 4 :</b> Vývoj emisných štandardov v USA pre ľahšie-úžitkové benzínové motorové vozidlá (gramy na míľu) (referencia : [8, str.3]) .....	23
<b>Obrázok 5:</b> Preškálované pravdepodobnosti vyradenia automobilu z prevádzky ( $y \rightarrow \ln(y)$ ) .....	29
<b>Obrázok 6 :</b> Grafické zobrazenie získaných pravdepodobností .....	29
<b>Obrázok 7 :</b> Prvá časť nášho rozdelenia (graf na ľavej strane) - Staré autá (age > 5) Druhá časť (graf na pravej strane) – Nové autá (age < 5) .....	30
<b>Obrázok 8 :</b> Hodnoty koeficienta $\alpha_i$ v rokoch 1998 - 2001 s jeho odhadnutou priamkou .....	31
<b>Obrázok 9 :</b> Získané pravdepodobnosti vyradenia starších automobilov z dopravnej prevádzky od roku 1999 po rok 2003 .....	33
<b>Obrázok 10 :</b> Preškálované hodnoty pravdepodobnosti vyradenia automobilov vo veku od dvoch do piatich rokov.....	35
<b>Obrázok 11 :</b> Grafy hodnôt pre roky 1999 až 2003 vygenerované našimi odhadmi pre novšie autá .....	37
<b>Obrázok 12 :</b> Novšie automobily - druhý prístup od roku 1998-2003 .....	38
<b>Obrázok 13 :</b> Graf k tabuľke 18.....	39
<b>Obrázok 14 :</b> Pravdepodobnosť vyradenia automobilu odhadnuté prvým prístupom .....	40
<b>Tabuľka 1 :</b> Počty áut vlastnenými vládou a domácnosťami v miliónoch.....	15
<b>Tabuľka 2 :</b> Odhadnuté predaje nových automobilov.....	20
<b>Tabuľka 3 :</b> Porovnanie koeficientov zo štúdie a našich .....	24
<b>Tabuľka 4 :</b> Počet áut v prevádzke v rokoch 1997 - 2001 zoradené podľa veku(v miliónoch) .....	28
<b>Tabuľka 5 :</b> Pravdepodobnosti vyradenia automobilov v rokoch 1998 - 2001 v USA .....	28
<b>Tabuľka 6 :</b> Odhadnuté hodnoty koeficientov $\alpha_i$ a $\beta_i$ pre $i = 1998, 1999, 2000, 2001$ .....	31
<b>Tabuľka 7 :</b> Odhadnuté koeficienty z lineárnej regresie (8) .....	31
<b>Tabuľka 8 :</b> Odhady pre HATbeta .....	32
<b>Tabuľka 9 :</b> Odhady pre rovnicu (7) pre roky 1999 – 2003 .....	32
<b>Tabuľka 10 :</b> Hodnoty logaritmov pravdepodobností vyradovania starších automobilov v rokoch 1999 - 2003 odhadnuté našimi rovnicami .....	32

<b>Tabuľka 11:</b> Porovnanie nami získaných údajov (zvýraznené) a údajov z <i>Transportation energy data book</i> .....	34
<b>Tabuľka 12 :</b> Priemer odhadnutých $\alpha_i$ z tabuľky 6 .....	34
<b>Tabuľka 13 :</b> Hodnoty pravdepodobnosti vyradenia automobilu pre dlhšie obdobie .....	35
<b>Tabuľka 14 :</b> Hodnoty koeficientov $a$ , $b$ a $c$ odhadnuté pomocou interpolácie kvadratickým splajnom.....	36
<b>Tabuľka 15 :</b> Koeficienty a konštanty z lineárnej regresie, ktorou sme odhadovali koeficienty kvadratického splajnu .....	36
<b>Tabuľka 16 :</b> Hodnoty koeficientov interpolačných polynómov pre roky 1999 – 2003 .....	37
<b>Tabuľka 17 :</b> Porovnanie dát našich a reálnych .....	37
<b>Tabuľka 18 :</b> Hodnoty konštantných funkcií pre novšie automobily získané ako priemer dostupných údajov .....	38
<b>Tabuľka 19 :</b> Pravdepodobnosť vyradenie automobilu v roku 1999 prvým a druhým prístupom v porovnaní s reálnymi dátami.....	39
<b>Tabuľka 20 :</b> Predikcia druhým prístupom vekového rozloženia automobilov v dopravnej prevádzke.....	41

## Úvod

V dnešnej dobe je automobilový priemysel nevyhnutnou súčasťou makroekonomického sektoru väčšiny krajín. Automobiloví výrobcovia sa snažia vymýšľať každý deň nové, komfortnejšie typy áut, ktoré by mohli priniesť revolúciu do tohto priemyslu. Veľké toky peňazí, ktoré plávajú v tomto odvetví, sa odrážajú aj na vyspelosti krajín a automobilový priemysel výrazne koreluje s hrubým domácim produktom [1]. USA patrí medzi najväčšie krajiny, v ktorých automobilový priemysel hrá veľkú rolu v ekonomickej vyspelosti. Až 25 percent z celkovej spotreby ľudí v USA zahŕňa predaj automobilov a súčiastok do nich, kým výroba motorových vozidiel tvorí 5 percent z celkovej výroby v krajine [1]. Každý rok sa v Amerike zbierajú dáta, ktoré popisujú stav prevádzky. Tieto dáta sú pre verejnosť prístupné, avšak v posledných rokoch spoločnosti, ktoré tieto dáta uverejňovali, spoplatnili svoje štatistiky.

**Motivácia.** Cieľom našej práce je zostavenie matematického modelu, ktorý bude popisovať vývoj vekového rozloženia áut v dopravenej prevádzke ako aj trendy vo vyradovaní motorových vozidiel. Takto získaný model, popisujúci spomínané vlastnosti, nám však môže napovedať ešte omnoho viac a preto takýto model môže slúžiť aj na analýzu miery nehodovosti. Tento model bude zostavený z dvoch hlavných poznatkov a to konkrétne z počtu predaných áut (Sales) a z počtu áut, ktoré sú vyradené z prevádzky (Scrappage).

**Zozbieranie dát.** Prvým krokom, ako pri každom matematickom modelovaní, bude zozbieranie všetkých dostupných dát. Keďže európske dáta sme nezískali, budeme vychádzať z dát amerických, ktoré sú voľne publikované. Slúžia na to rôzne zdroje ako napríklad *Transportation Data Book*, ktoré vydáva *Oak Ridge National Laboratory (ORNL)*, kde sa pravidelne (každý rok) uverejňujú štatistiky pre USA hovoriace o počte áut v prevádzke, o predajnosti áut, ale aj o emisných štandardoch a rôznych vplyvoch automobilového priemyslu na životné prostredie.

**Existujúce modely.** Tejto téme sa nevenuje veľa prác, nakoľko sa v priemysle, buď používajú platené zozbierané dáta, alebo sa skúma len celkový počet automobilov bez detailného skúmania ich vekovej štruktúry. Avšak podobnými otázkami ako my sa zaoberá

práca Motor Vehicle Stocks, Scrappage, and Sales od Alan Greenspana and Darrel Cohena [1]. Autori zostavili matematický (štatistický) model, ktorý popisuje vývoj predaja automobilov v USA pomocou vyradených áut [1].

**Výsledky.** V našej práci sme pokúsili vytvoriť matematický model na základe reálnych dát o stave dopravenej prevádzky v USA. Predlohou nám bol populačný model, ktorý popisuje demografický vývoj. Zistili sme, že pomocou historických dát pre pár rokov, sme schopní zostaviť model, pomocou ktorého môžeme predikovať vekové rozloženie dopravnej prevádzky.

# 1. Analýza práce Greenspana a Cohena

## 1.1. Obsah práce

V roku 1996 predseda<sup>1</sup> inštitútu Federal Reserve Board Alan Greenspan a ekonóm Darrel Cohen zostavili matematický model, ktorý bol súčasťou práce s názvom Motor Vehicle Stocks, Scrappage, and Sales (preklad autora: Počet motorových vozidiel v prevádzke, počet vyradených áut z prevádzky a počet predaných motorových vozidiel). Tento model poukazuje na vývoj predajnosti áut v Spojených štátoch amerických. Ich práca ponúka nový rámec pre analýzu celkového predaja nových motorových vozidiel, kde táto analýza zahŕňa separované modely zmeny počtu áut v prevádzke a počtu vyradených vozidiel. Pretože takýto prístup k modelovaniu vývoja predajnosti nových vozidiel vyžaduje veľmi malé predpoklady týkajúce sa demografického vývoja, stavu ekonomiky v krajine alebo požiadaviek spotrebiteľa, je zrejmé, že takáto štúdia môže byť vhodným nástrojom na predpovedanie makroekonomického vývoja [1].

Z historických dát je známe, že vývoj predajnosti nových automobilov má cyklický charakter a preto je možné pomocou starších dát zostaviť istú závislosť predaja od počtu áut v prevádzke a počtu áut vyradených z prevádzky. Tento fakt si všimli aj autori existujúceho modelu a pomocou dostupných štatistík zostavili rovnicu na výpočet predajnosti nových automobilov v čase  $t$ :

$$\text{Sales}_t = \Delta V_t + \text{EngScrap}_t + \text{CycScrap}_t \quad (1)$$

V tejto rovnici sa nachádzajú štyri veličiny a to:

- $\text{Sales}_t$  – počet predaných áut za čas  $t$
- $\Delta V_t$  – úplný počet áut v prevádzke na konci periódy  $t$
- $\text{EngScrap}_t$  – počet „engineering scrappage“ za čas  $t$
- $\text{CycScrap}_t$  – počet „cyclical scrappage“ za čas  $t$

Pri pohľade na tieto premenné vidíme, že autori sa rozhodli scrappage, čiže počet áut vyradených z prevádzky, rozdeliť do dvoch skupín. Prvou skupinou sú tzv. „engineering scrappage(ES)“, ktorá v sebe zahŕňa autá vyradené z dôvodu opotrebovania

---

<sup>1</sup> Alan Greenspan bol predseda FRB od roku 1987 do 2006 [6].

alebo poškodenia spôsobeného používaním [1]. Druhým typom vyradených áut sú tzv. „cyclical scrappage(CS)“. Autá v tejto skupine sú vyradované predovšetkým vplyvom ekonomiky štátu, ktorá ovplyvňuje rozhodnutia ľudí vymeniť svoje automobily za novšie, šetrnejšie a lepšie, a tým pádom odráža v sebe aj faktory, ktoré ovplyvňujú samotnú ekonomiku ako cena ropy, cena nových vozidiel a opráv. Keďže ekonomika má cyklický charakter, môžeme povedať, že aj táto skupina vyradených áut je cyklická.

### 1.1.1. Zber a úprava dát

Dáta, pre túto štúdiu, autori získali najmä z dvoch zdrojov, ktoré každoročne publikujú podrobné edície, opisujúce stav dopravnej prevádzky. Prvým z nich je R.L. Polk & Co. Táto spoločnosť je poskytovateľom informácií o automobilovom odvetví v USA a marketingových riešení pre automobilový priemysel a pre poisťovacie spoločnosti [2]. Druhým zdrojom, ktorý slúžil na zozbieranie historických dát do práce, je organizácia AAMA, čo je akronym z amerického názvu American Automobile Manufacturers Association. Obe tieto spoločnosti uverejňujú každý rok počet zaregistrovaných áut k prvému júlu daného roku, ktoré potom vo svojich edíciách označujú názvom „Cars in Operation“, čo môžeme preložiť ako autá v prevádzke (Obr. 1).

Obrázok 1: Počet automobilov zoradených podľa veku (referencia : [7, str. 3-9])

**Table 3.5**  
Cars in Operation and Vehicle Travel by Age, 1970 and 2001

Age (years)	1970			2001			2001 Estimated vehicle travel		Average annual miles per vehicle
	Vehicles (thousands)	Percentage	Cumulative percentage	Vehicles (thousands)	Percentage	Cumulative percentage	Percentage	Cumulative percentage	
Under 1*	6,288	7.8%	7.8%	6,183	4.8%	4.8%	6.9%	6.9%	15,000
1	9,299	11.6%	19.4%	8,882	6.9%	11.7%	9.4%	16.3%	14,300
2	8,816	11.0%	30.3%	8,093	6.3%	18.0%	8.2%	24.6%	13,700
3	7,878	9.8%	40.1%	7,555	5.9%	23.9%	7.2%	31.8%	12,900
4	8,538	10.6%	50.8%	7,860	6.1%	30.0%	7.2%	39.1%	12,400
5	8,506	10.6%	61.3%	7,337	5.7%	35.7%	6.5%	45.6%	12,000
6	7,116	8.8%	70.2%	8,555	6.6%	42.3%	7.4%	53.1%	11,700
7	6,268	7.8%	78.0%	7,471	5.8%	48.1%	6.3%	59.4%	11,400
8	5,058	6.3%	84.3%	7,420	5.8%	53.9%	6.1%	65.5%	11,100
9	3,267	4.1%	88.3%	6,807	5.3%	59.2%	5.4%	71.0%	10,700
10	2,776	3.5%	91.8%	6,810	5.3%	64.5%	5.0%	76.0%	9,900
11	1,692	2.1%	93.9%	6,692	5.2%	69.7%	4.5%	80.5%	9,000
12	799	1.0%	94.9%	6,742	5.2%	74.9%	4.7%	85.2%	9,400
13	996	1.2%	96.1%	6,189	4.8%	79.7%	3.8%	88.9%	8,200
14	794	1.0%	97.1%	5,345	4.2%	83.9%	2.9%	91.8%	7,200
15 and older	2,336	2.9%	100.0%	20,773	16.1%	100.0%	8.2%	100.0%	5,300
Subtotal	80,427	100.0%		128,714	100.0%				
Age not given	22			0					
Total	80,449			128,714					

Na Obr. 1 sa nachádza ukážka z Edície 26, kde môžeme vidieť presný počet áut (Vehicles) zoradených podľa veku k 1. júlu 1970 a 2001.

Pojmy „auto zaregistrované“ k istému dátumu a „auto v prevádzke“ napovedajú, že automobily v oboch skupinách majú blízky vzťah, avšak nemožno ich pokladať za identické [1]. Hlavným dôvodom je, že autá, ktoré boli vyradené z prevádzky počas roka, nie sú vymazané z údajov, pokiaľ sa znova nezaregistrujú. Obe spoločnosti, ktoré poskytujú tieto dáta, neuverejňujú skutočný počet áut vyradených z prevádzky, ale toto číslo dostanú ako rozdiel všetkých *nových* registrácií za daný rok a úplný počet registrácií v danom roku.

Keďže tieto dva pojmy sa nezhodujú, autori tejto štúdie museli zaviesť akúsi novú rovnicu, ktorá by vyjadrovala reálny stav prevádzky v t-tom roku. Táto rovnica má za úlohu spracovať vstupy, čo sú získané dáta, takým spôsobom, aby naozaj išlo o „cars in operation“, čiže autá v prevádzke.

$$V(t-1) = R(t) - [M(t,t) + M(t-1,t)] + (.99)F(t-1) \quad (2)$$

- $V(t-1)$  – „skutočný“ stav prevádzky k prvému júlu v roku t-1;
- $R(t)$  – všetky registrácie automobilov zo zdroja R.L. Polk & Co k prvému júlu v roku t;
- $M(t,t)$  – počet nových registrácií roku t;
- $M(t-1,t)$  – počet automobilov vyrobených v roku t-1, ale zaregistrovaných v roku t;
- $F(t-1)$  – počet nových registrácií automobilov v modelovacom roku t-1 zaznamenaných počas deviatich mesiacov končiacich 30.júna v roku t-1;

Tento prepočet autori ilustrovali na príklade, ktorý by mal lepšie zobrazit' danú situáciu. Vezmime si rok 1990 a našou úlohou je zistiť počet automobilov v prevádzke v danom roku. Podľa rovnice (2) sa to rovná celkovému počtu Polk registrácií ku dňu 1.7.1991, mínus registrácie v modelovacích rokoch 1990 a 1991, plus počet všetkých nových registrácií, zaznamenaných za 9 mesiacov (vynásobené 0,99) [1].

### 1.1.2. Modelovanie stavu prevádzky

Predtým ako sa autori rozhodli vytvoriť rovnicu pre odhad stavu dopravnej prevádzky, rozdelili celkový počet áut v prevádzke do dvoch skupín. Prvou sú autá vlastniace podnikateľmi a vládou, a druhou autá, ktoré používajú domácnosti. Kvôli tomuto rozdeleniu museli autori použiť ďalší zdroj dát, a to The United States Census Bureau. Táto vládna inštitúcia je zodpovedná za zozbieravanie údajov týkajúcich sa počtu obyvateľov v USA a ekonomickej vyspelosti krajiny [3]. Okrem toho vlastní odhady na počet domácností a počty domácností, ktoré vlastní motorové vozidlá.

Year	Total	Household (Census)	Business and Government	"Smoothed"	
				Business & Government	Household
	1	2	3 = 1 - 2	4	5 = 1 - 4
1973	104.86	99.67	5.19		
1974	110.86	103.67	7.19	6.18	104.68
1975	113.84	107.67	6.17	5.98	107.86
1976	116.26	111.67	4.59	5.33	110.93
1977	120.89	115.67	5.22	5.24	115.65
1978	124.87	n.a.			
1979	129.64	n.a.			
1980	133.58	127.67	5.91	5.03	128.55
1981	136.29	132.33	3.96	4.50	131.79
1982	138.49	n.a.			
1983	141.43	n.a.			
1984	144.20	n.a.			
1985	148.62	145.00	3.62	4.62	144.00
1986	153.32	n.a.			
1987	158.10	151.83	6.27	6.09	152.01
1988	162.83	n.a.			
1989	167.22	158.83	8.39	8.89	158.33
1990	170.42	n.a.			
1991	172.85	160.83	12.02	10.20	162.65

*Tabuľka 1 : Počty áut vlastnenými vládou a domácnosťami v miliónoch*

*(referencia : [1, str.30])*

V Tabuľke 1. vidíme úplný počet áut v prevádzke v prvom stĺpci, ktorý musel byť upravený rovnicou (2). Následne v ďalšom stĺpci tabuľky sa nachádza počet domácností, ktoré vlastní auto (dáta získané z Censusu). Rozdielom prvého a druhého stĺpca autori získali počet áut vlastnených vládou. Ďalším krokom bolo tzv. „smoothing“, čo môžeme preložiť ako zhladzovanie. Autori použili na zhladzovanie diskretizáciu difúznej rovnice (Laplacovho operátora) s malým difúznym koeficientom. Tento proces sa používa v prípade, keď „surové“ dáta príliš fluktuujú v čase a nie sú dostatočne citlivé. Táto metóda tak predpokladá pekné (monotónne) správanie sledovaných ukazovateľov v počte



registrovaných áut. A tak vo štvrtom stĺpci vidíme vyhladené dáta áut vlastnených vládou. Nakoniec rozdielom úplného počtu áut v prevádzke a „vyhladených“ vládnych automobilov získali autori „vyhladené“ počty motorových vozidiel, ktoré vlastní domácnosti. Takto upravené a rozdelené autá mohli byť ďalej používané pri odhade počtu áut v prevádzke. Samotná rovnica, ktorú autori použili, obsahuje tri základné členy vid' (3).

$$V_{HH} = (\# \text{ HH's}) (\% \text{ of HH's owning vehicle}) (\text{Avg \# of vehicles per HH that own a vehicle}) \quad (3)$$

- (# HH's) – počet domácností v USA; toto číslo je odhadované Censusom po rok 1995;
- (% of HH's owning vehicle) – podiel domácností, ktoré vlastní najmenej jedno auto; taktiež získane z Censusu;
- (Avg # of vehicles per HH that own a vehicle) – tento člen odhadli autori jednoduchou rovnicou :

Avg # cars per HH = 1\*(% owning exactly 1) + 2\*(% owning exactly 2) + 3.1\*(% owning at least 3); autori predpokladajú, že priemerný počet áut vlastnených domácnosťami, ktoré majú najmenej tri autá, je 3,1; každé číslo medzi 3 a 3,2 je prijateľné;

Z takto zostavených rovníc autori mohli odhadovať počet automobilov v prevádzke a teda prvý člen v rovnici (1).

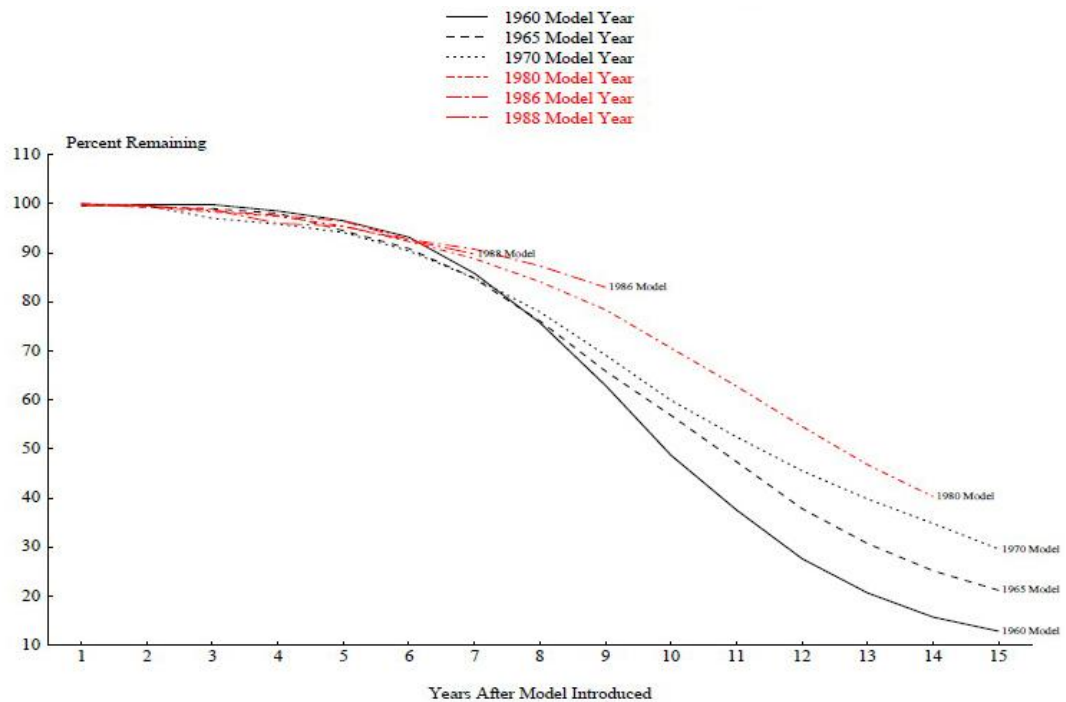
### 1.1.3. „Engineering scrappage (ES)“

Z empirických pozorovaní autori predpokladajú, že automobil za prvé tri roky svojho života nie je vyradený z prevádzky z dôvodov opotrebovania alebo kazovosti. Môžeme povedať, že v ranných obdobiach života auta je málo pravdepodobné, že by sa auto natoľko zničilo používaním, až by ho museli majitelia prestať používať. Preto v prvých rokoch pozorujeme dva hlavné dôvody, prečo sa ľudia svojich automobilov zbavujú. Za najčastejši možno považovať automobilové nehody, kedy sa musia majitelia často vzdať i svojho nového auta, pretože sú na ňom také škody, že sa nedá ďalej používať. Druhým dôvodom je tzv. „scrappage decision“, čo možno opisne preložiť, ako rozhodnutie

človeka prestať auto používať na základe nespokojnosti alebo nespoľahlivosti daného vozidla [1]. Človek si spočíta plusy a mínusy, a na základe toho sa rozhodne, či auto naďalej bude používať alebo nie.

Čo možno nevidieť na prvý pohľad je, že v tejto skupine sa nachádzajú aj rôzne faktory vplývajúce na „scrapage decision“. Takými nezávislými premennými môžu byť napríklad príjem domácnosti alebo cena ropy. Práve tieto faktory vplývajú na intenzitu používania auta a s ňou rastúcimi výdavkami na údržbu.

Obrázok 2: Percentuálne vyjadrenie životnosti automobilov (referencia : [1, str.32])



Na tomto grafe možno vidieť, že čím je auto skoršie vyrobené, tak tým má tendenciu dlhšie slúžiť. Pravdepodobne je to spôsobené novšími, a teda aj lepšimi technológiami. Taktiež môžeme pozorovať, ako autá rôznych rokov výroby slúžia ľuďom a akého veku sa „dožívajú“. Napríklad 50 percent automobilov vyrobených v roku 1960 sa aj po deviatich rokoch stále používa a vidieť, že novšie motorové vozidlá zostávajú dlhšie v dopravnej prevádzke.

Na Obr.2 môžeme vidieť graf, ktorý slúžil autorom ako predloha na zostavenie združenej regresie, ktorou sa snažili odhadnúť trend ES pre ďalšie roky.

$$\ln y_{it} = a + b_1 t^2_{60} + b_2 t^2_{61} + \dots + b_{30} t^2_{89} + ct^3 + u_{it} \quad (4)$$

- $y_{it}$  – podiel automobilov modelovacieho roku  $i$  zvyšujúcich vo veku  $t$  (získane z Polk data)
- $b_i$  – koeficienty regresie
- $t^2_i$  – stĺpcové vektory, z ktorých každý ma 309 prvkov, príklad pre rok 1960 :  $t^2_{60} = (1,4,9,\dots,144,0,\dots,0)$ ;
- $t^3$  – regresor v tvare  $(1,8,27,\dots,1728,1,8,27,\dots,1728,\dots)$

Forma tejto regresie je motivovaná funkciou hustoty normálneho rozdelenia ( $y = Ae^{-kt^2}$ ), ktorá sa v širšom slova zmysle ukazuje ako najlepšia aproximácia (pre  $t > 0$ ) grafu z Obr. 2 [1]. Dôvodom, prečo sa autori rozhodli použiť tvar tejto funkcie, je, že pre  $t$  idúce do nekonečna sa tento výraz limitne blíži k nule, čo uisťovalo autorov v tom, že si vybrali správny tvar. Avšak experimenty ukázali, že takto zvolená normálna forma vygenerovala príliš málo starých áut, t.j. chvosty neboli dostatočne hrubé. Tento problém bol vyriešený pridaním nového člena  $t^3$ . Konečný tvar združenej regresie výborne opisuje graf a všetky koeficienty vyšli signifikantné (Obr. 3).

Obrázok 3 : Odhady koeficientov, štandardná odchýlka, hodnota  $t$  – statistiky, Beta koeficient a elasticita (referencia : [1, str.31])

Series	Coefficient	Standard Error	t-Value	Beta Coefficient	Elasticity
Constant	-0.00413	0.004418	-0.935		
1) TSQUAR60	-0.01611	3.442E-4	-46.819	-0.4895	0.073955
2) TSQUAR61	-0.01568	3.442E-4	-45.570	-0.47645	0.071982
3) TSQUAR62	-0.01416	3.442E-4	-41.132	-0.43004	0.064972
4) TSQUAR63	-0.0135	3.442E-4	-39.232	-0.41018	0.06197
5) TSQUAR64	-0.0129	3.442E-4	-37.491	-0.39198	0.059221

Na Obr. 3 môžeme vidieť príklady odhadnutých hodnôt koeficientov  $b_i$  (pre  $i = 60, 61, 62, 63$  a  $64$ ).

Rovnakú regresiu autori použili pre ďalšie modelovacie roky. Avšak museli pridať predpoklad, že  $\hat{b}$  (odhad koeficientu) pre každý modelovací rok po roku 1990 sa rovná priemeru koeficientov pre roky 1980-1989. Musíme ale poznamenať, že tento

predpoklad by znamenal, že sa výdrž auta nezmenila medzi rokmi 1980-1989, a preto ak v roku 1990 nastal vývoj a životnosť stúpla, tak odhad autorov pre ES je prehnaný. Tieto fakty si autori všimli, a preto museli k odhadom koeficientov pristupovať veľmi opatrne.

Vytvorili lineárnu regresiu, ktorou sa snažili namodelovať vývoj koeficientov v čase.

$$\text{COEFF}_t = - .02 - .0007 \text{EPA}_t + .000004 \text{TEEN}_t + .0003 \text{TIME} \quad (5)$$

(-6.2)      (-1.8)                      (1.8)                      (3.4)

- COEFF – odhadnuté koeficienty
- EPA – emisné štandardy v USA; autori ich merali inverznou hodnotou k povolenému množstvu hydrokarbónu na míľu
- TEEN – počet teenagerov (16-19) v USA; autori predpokladajú, že veľa starých áut je používaných práve teenagermi
- TIME - regresor

Všimnime si, že táto rovnica dobre zapadá a koeficienty majú očakávané znamienka [1].

#### 1.1.4. „Cyclical Scrappage“

Ďalšou zložkou celkového počtu vyradených áut z prevádzky je tzv. „non-engineering component“ a teda cyclical scrappage (CS). Ako sme už uviedli, táto časť má procyklický charakter a v nasledujúcej lineárnej regresii je tento charakter zastúpený prítomnosťou miery nezamestnanosti a jej neskorších hodnôt.

Ďalšie nezávislé premenné, ktoré sa autori rozhodli pridať do regresie, sú napríklad cena ropy (PG), cena nových vozidiel (PN) alebo cena opráv (PR). Tieto faktory nemajú cyklický charakter, ale ich vplyv na rozhodnutie človeka vyradiť automobil je markantný.

$$\text{CYCSCRAP}_t = - 4.3 - .38 \text{RU}_t + .31 \text{RU}_{t-1} + 9.3 (\text{PR/PN})_t - 0.05 \text{PG}_t \quad (6)$$

(-1.6)      (-2.5)                      (2.2)                      (2.8)                      (-4.0)

### 1.1.5. Explicitný príklad

Na lepšie porozumenie významu a funkčnosti vypracovaného modelu, sa na koniec autori rozhodli uviesť konkrétny príklad. Regresie a modely, ktoré uviedli vo svojej štúdií, im poskytujú vhodné podklady na generovanie údajov pre roky 1992-1996. Z tohto dôvodu si zvolili rok končiaci 30. júna 1995 za svoj modelovací.

Prvým krokom bolo určenie všetkých parametrov z rovnice (1):

$$\text{Sales}_{1995} = \Delta V_{1995} + \text{EngScrap}_{1995} + \text{CycScrap}_{1995}$$

- $\Delta V_{1995}$  – rozdiel všetkých automobilov v prevádzke v roku 1995 a 1994. Tento výpočet im implikoval nárast o 3,4 milióna automobilov za jeden rok.
- $\text{EngScrap}_{1995}$  – odhad tohto parametra bol vypočítaný združenou regresiou z kapitoly 1.3. Výsledok prebehnutých výpočtov bol 11,1 milióna automobilov
- $\text{CycScrap}_{1995}$  – počet automobilov zahrnutých v tejto skupine bol, pre rok 1995, 0,8 milióna

Súčtom všetkých získaných hodnôt dostaneme celkový počet nových automobilov, ktoré boli predané v roku 1995. Samozrejme ide o odhad, ktorý musí byť porovnaný s reálnymi dátami. Po sčítaní  $3,4 + 11,1 + 0,8$  prišli autori k výsledku 15,3 milióna jednotiek, ktorý v porovnaní so skutočnou hodnotou 15,3 milióna, je bezchybný.

Year ending June 30	Change in Stock of Vehicles	Engineering Scrappage	Cyclical Scrappage	Predicted New Sales	Actual Sales
	(1)	(2)	(3)	4 = 1+2+3	
1992	2.3	10.5	0.3	13.1	12.8
1993	1.6	11.0	0.5	13.1	13.6
1994	2.9	11.0	1.0	14.9	14.9
1995	3.4	11.1	0.8	15.3	15.3
1996	3.0	11.5	0.6	15.2	15.5

Tabuľka 2 : Odhadnuté predaje nových automobilov

(referencia : [1, str.17])

V Tabuľke 2 môžeme vidieť ďalšie hodnoty z odhadnutých rokoch v porovnaní s existujúcimi dátami.

## 2. Analýza metód a výsledkov práce

V tejto časti by sme sa chceli venovať chybám a nedostatkom, ktoré sme si všimli v štúdií od Alana Greenspana a Darrela Cohena. Nanešťastie je v ich práci badateľné, že obaja majú vo svojich pracovných oboroch menší kontakt s matematikou ako s ekonómiou. Motivácia pre ich snahu je veľmi jasná a dobrá, ale prevedenie už horšie.

**Základná rovnica.** Keď sa pozrieme na rovnicu (1), z ktorej autori vychádzali, tak vidíme jednoduchý vzorec. Počet nových predaných áut možno spočítať ako súčet zmeny počtu áut v prevádzke a áut vyradených z prevádzky. Čo je však zaujímavé, je prítomnosť člena  $CycScrap_t$ , ktorý má slúžiť na odhad počtu áut vyradených z prevádzky vplyvom zmeny ekonomiky a jej vplyvu na ceny áut, súčiastok, benzínu a opráv. Avšak je zrejmé, že úlohou tejto premennej je „pohltiť“ všetky drobné výkyvy a chyby, ktoré vznikli pri výpočte  $EngScrap_t$ . Pre odhad  $CycScrap$  bola vytvorená lineárna regresia, ktorá sa dá veľmi jednoducho namodelovať tak, aby výsledky vyšli ako potrebovali autori a teda vidíme, že celkový prínos tejto premennej v explicitnom príklade je malý. Na druhej strane nemožno poprieť, že takto pridaný člen sa veľmi často využíva, a ak jeho namodelovanie lineárnou regresiou vyšlo singifikantné ako v tomto prípade, nie je dôvod na jeho odstránenie.

**Engineering Scrapage.** Ďalší problém, ktorý sa nachádza v tejto štúdií a pokladáme ho za markantný, je odhad a vysvetlenie  $EngScrap_t$  z rovnice (1). Autori po získaní a vykreslení grafu, ktorý môžeme vidieť na Obr.3, sa snažili vytvoriť takú regresiu, ktorá by najlepšie opisovala daný graf. Rozhodli sa vychádzať z funkcie hustoty normálneho rozdelenia ( $y = Ae^{-kt^2}$ ) ako vidíme v rovnici (4) :  $\ln y_{it} = a + b_1 t_{60}^2 + b_2 t_{61}^2 + \dots + b_{30} t_{89}^2 + ct^3 + u_{it}$ . Avšak z dôvodu nedostatku starých áut museli pridať člen  $t^3$ , i keď oveľa logickejšie vyzerá byť doplnenie člena, ktorý by pokračoval v analógií funkcie hustoty normálneho rozdelenia, a teda  $t^4$ . Autori tento svoj čin neopisujú a preto sa môžeme iba domnievať, že s regresorom  $t^3$  im vychádzali krajšie výsledky. Posledný člen združenej regresie  $u_{it}$  nie je v celej práci nijako vysvetlený, čo je z pohľadu čitateľa veľmi záhadné, a preto nevieme, čo hľadať za touto premennou. Ako uvádzajú autori, každý vektor  $t_i^2$  má 309 členov. Skladá sa z druhých mocnín čísiel 1 až 12 a núl, ktoré dopĺňajú zvyšné miesto (napr.  $t_{60}^2 = (1,4,9,\dots,144,0,\dots,0)$ ). Ako sme už viac krát spomínali, autori tejto štúdie nevysvetľovali viacero faktov a tieto vektory nie sú výnimkou. Prečo ich zvolili práve takto nevieme, a dokonca sme neprišli na žiadne racionálne vysvetlenie. Preto

podľa nášho názoru by sa mali vektory  $t^2$  skladať z inak usporiadaných čísiel 1 až 12. Napríklad pre rok 1960 by vektor  $t^2_{60}$  mal mať tvar (0, 0, 0, 1, 4, ..., 12, ..., 0), pretože autori pozerali na 15 ročný horizont a prvé tri roky ich z hľadiska vyradovania nezaujímali (viď kapitola 1.2.3). Potom nasledujúci vektor  $t^2_{61}$  by bol skoro identický až na posunutie o dvanásť miest nižšie (ak si vektory  $t^2$  predstavíme ako stĺpcové). Takto by sme posúvali druhé mocniny čísiel 1 až 12, až kým by sme neprišli k roku 1980, kedy roky, ku ktorým mali autori dostupné dáta, sa začali skracovať. Z tohto dôvodu vektor  $t^2_{80}$  by z nášho hľadiska mal tvar (0, ..., 0, 1, 4, 9, 16, 25, 36, 49, 64, 81, 100, 121, 0, ..., 0). Potom vektor pre rok 1981 len druhé mocniny po číslo 10 a takto by sme pokračovali po rok 1986, ktorý by sa podľa našej analógie mal skladať z piatich mocnín (1, 4, 9, 16, 25). Avšak, po preskúmaní grafu z Obr. 3, sme dospeli k názoru, že v tomto roku museli autori získať o údaj viac, a preto aj tento vektor obsahoval mocnín šesť. Nakoniec by sme vytvorili zvyšné vektory až po rok 1989, ktorý je konečným rokom pre autorov (práca bola napísaná v roku 1996) a obsahoval by práve tri mocniny<sup>2</sup>. Dôležitou kontrolou, či sme postupovali pri vytváraní vektorov správne, je súčet všetkých nenulových elementov. Od roku 1960 po rok 1979 obsahoval každý vektor 12 druhých mocnín (20 rokov x 12 elementov = 240). Potom sa dáta, ktoré mali autori k dispozícii, začali skracovať a vektory mali postupne 11, 10, 9 až 6 (rok 1987) nenulových prvkov. Následne prišiel zlom, kedy znova v roku 1986 mali údaje pre 6 ročné obdobie. Od tohto roku to však išlo o jeden údaj nižšie, až po rok 1989, kedy mali údaje 3 (11 + 10 + ... + 6 + 6 + ... + 3 = 69). Takto sme získali 309 prvkové vektory, ktoré podľa nášho názoru dávajú väčší zmysel a popisujú dostupné údaje, ktoré mali autori štúdie.

**Reprodukcia modelovania.** Z dôvodu mnohých nejasností, ktoré popisujeme vyššie, sme sa pokúsili proces vytvárania a odhadovania  $EngScrap_t$  zreprodukovať a následne odhadnúť aj pre ďalšie roky. Prvý krok, ktorý sme museli podniknúť, je odhad koeficientov  $b_i$  z rovnice (4). Tieto koeficienty slúžia ako isté meradlo, čo sa stalo v danom roku  $i$ . Autori ich získali zo združenej regresie (4) a následne vytvorili lineárnu regresiu (5) ( $COEFF_t = - .02 - .0007 EPA_t + .000004 TEEN_t + .0003 TIME$ ). Koeficienty lineárnej regresie (5) odhadli pomocou metódy najmenších štvorcov. Našou úlohou teda bolo získať potrebné dáta (EPA – inverzná hodnota povoleného množstva gramov hydrokárbonu podľa emisných štandardov v USA, TEEN – počet teenagerov v USA, TIME – bližšie

---

<sup>2</sup> Prečo práve tri mocniny? Ako sme už uviedli roky 1990, 1991, 1992 autorov nezaujímali, pretože hľadeli na obdobie po prvých troch rokoch. Keďže údaje, ktoré vtedy boli schopní získať, sú po rok 1995, tak tieto tri mocniny korešpondujú k rokom 1993, 1994 a 1995.

nešpecifikovaný regresor), ktoré by sa zhodovali s dátami, ktoré mali k dispozícii autori štúdie. Ako kontrolu sme dáta dosadili do lineárnej regresie (5) a následne naše výsledky porovnali s výsledkami, ktoré uvádzajú autori (Obr. 3).

**Dáta.** EPA, emisné štandardy v USA, sme získali z publikácie autorov Asif Faiz, Christopher S. Weaver, Michael P. Walsh, ktorí vo svojom diele *Air Pollution From Motor Vehicles : Standards and Technologies For Controlling Emissions* na strane číslo 3 v tabuľke 1.1 (Obr. 4) uvádzajú vývoj povoleného množstva hydrokarbónu v USA. Ako vidíme v tejto tabuľke sa nenachádzajú údaje pre všetky roky, avšak zistili sme, že tieto nepopísané roky sa dátovo zhodujú s posledným popísaným rokom. Napríklad v rokoch 1982 - 1993 boli v USA také isté emisné štandardy ako v roku 1981 a teda 0,41 gramu hydrokarbónu na míľu. Takto získané dáta sme ešte nemohli vložiť do regresie (5), ale ako Greenspan s Cohenom uvádzajú, musíme spraviť inverzné hodnoty.

Obrázok 4 : Vývoj emisných štandardov v USA pre ľahšie-úžitkové benzínové motorové vozidlá (gramy na míľu) (referencia : [8, str.3])

**Table 1.1 Progression of U.S. Exhaust Emission Standards for Light-Duty Gasoline-Fueled Vehicles (grams per mile)**

Model year	Carbon monoxide	Hydrocarbons	Nitrogen oxides
Pre-1968 (uncontrolled)	90.0	15.0	6.2
1970	34.0	4.1	—
1972	28.0	3.0	—
1973-74	28.0	3.0	3.1
1975-76	15.0	1.5	3.1
1977	15.0	1.5	2.0
1980	7.0	0.41	2.0
1981	3.4	0.41	1.0
1994-96 (Tier 1)	3.4 (4.2)	0.25 <sup>a</sup> (0.31)	0.4 (0.6)
2004 (Tier 2) <sup>b</sup>	1.7 (1.7)	0.125 <sup>a</sup> (0.125)	0.2 (0.2)

TEEN, čiže počet teenagerov v USA, autori brali do úvahy od 16 do 19 rokov života. Potrebne dáta sme získali z internetového zdroja <sup>3</sup>, ktorý nám poskytoval výber vekového rozmedzia.

Po vyhľadání vstupných údajov bolo našou úlohou dáta dosadiť do regresie (5) a vypočítať koeficienty  $b_i$ . Na overenie správnosti výpočtov a dát sme vypočítali najskôr koeficienty pre roky, ktoré sa už nachádzali v štúdií.

<sup>3</sup> Internetový zdroj : <http://www.bls.gov/cps/demographics.htm>



Tabuľka 3 : Porovnanie koeficientov zo štúdie a našich

	Greenspan	Horský
1970	-0,01023	-0,01128
1971	-0,01041	-0,0108
1972	-0,00983	-0,01039
1973	-0,01022	-0,00995
1974	-0,01	-0,00952
1975	-0,00982	-0,00936
1976	-0,00888	-0,00901
1977	-0,00784	-0,00869
1978	-0,00763	-0,00839
1979	-0,00733	-0,00811
1980	-0,00852	-0,00913
1981	-0,00825	-0,00899
1982	-0,00778	-0,00887
1983	-0,00682	-0,00878
1984	-0,00666	-0,00868
1985	-0,00626	-0,00842
1986	-0,006	-0,00809
1987	-0,0067	-0,00775
1988	-0,00757	-0,00753
1989	-0,00926	-0,00737

V Tabuľke 3 môžeme vidieť porovnanie koeficientov  $b_i$  pre rok 1970 – 1989. V stĺpci Greenspan sa nachádzajú údaje zo štúdie a v stĺpci s názvom Horský môžeme vidieť hodnoty koeficientov, ktoré sme vypočítali dosadením našich dát do lineárnej regresie (5) odhadnutej autormi štúdie. Koeficienty  $b_i$  získané zo štúdie začali ku koncu (1986 - 1989) nečakane stúpať. Tento fakt považujeme za veľmi zvláštny, keďže ani počet TEENagerov a ani Emisné štandardy (EPA) sa v tých rokoch žiadnym spôsobom nezmenili. Preto môžeme predpokladať, že autori spravili vo svojich odhadoch alebo dátach chybu.

**Výsledky reprodukcie.** Po dlhom skúmaní a porovnávaní dát sme prišli k záveru, že žiadnym spôsobom nie sme schopní dostať také isté hodnoty koeficientov  $b_i$  ako sa nachádzajú v štúdiu. Prečo je tomu tak nevieme povedať presne, ale za najpravdepodobnejšie pokladáme rozdielnosť dát, ktoré mali k dispozícii autori štúdie, a s ktorými sme pracovali my.

V tejto Analýze existujúceho modelu sme sa pokúsili ukázať náš pohľad na prácu, ktorá je jediným vypracovaným odborným referátom na tému Modelovania dopravnej prevádzky. Ako sme uviedli, v tejto práci sa nachádza viacero nedostatkov a z tohto dôvodu možno povedať, že naša motivácia je o to zaujímavejšia.

## 3. Model

### 3.1. Položenie otázok

Téma modelovania počtu áut v dopravnej prevádzke nie je vôbec obsiahnutá. Existujúci model, ktorý na túto tému pojednáva, sme popísali, rovnako ako aj jeho chyby a nedostatky. Práve odhad a predpovedanie počtu vyradených áut z prevádzky sa stáva hlavnou otázkou, na ktorú sa budeme snažiť v tejto práci odpovedať. Ak by sme teda mali sformulovať presné znenie otázky, ktorej odpoveď je našim cieľom, znela by : ***Aký bude vývoj počtu áut vyradených z prevádzky v čase?*** Motiváciou pre nás je v prvom rade nedostatok informácií o týchto trendoch a v druhom rade neexistencia riešenia takto postaveného problému.

### 3.2. Matematické metódy

#### 3.2.1. Analýza dát z grafu

Po získaní relevantných dát bude našou prvou úlohou *analýza dát z grafu*. Odhadnúť správne trendy bude nevyhnutné pre ďalšie výpočty. Greenspanova a Cohenova práca nám napovedá, že trendy, ktoré budeme pozorovať na grafoch z našich dát, budú mať exponenciálne tvary. Z tohto dôvodu predpokladáme, že budeme musieť pretransformovať naše grafy do logaritmickej kriviek, aby sme boli schopní lepšie sledovať zmeny a trendy.

#### 3.2.2. Metóda najmenších štvorcov

Metódu najmenších štvorcov (MNS) použijeme na odhadovanie prechodových koeficientov na dvoch časových podintervaloch životnosti, na ktoré rozdelíme naše dáta. To, ako sa rozhodneme tieto podintervaly zvoliť, bude záležať od tvaru vygenerovaných kriviek, ale z pozorovaní a už získaných poznatkov z Greenspanovej a Cohenovej práce predpokladáme, že budú tieto podintervaly dva. Získané hodnoty z časových podintervalov lineárne extrapolujeme do ďalších rokov. Všetky výpočty a odhady nám budú neskôr slúžiť na vybudovanie diferenčného populačného modelu pre autá v dopravnej prevádzke, kde jednotlivé prechodové koeficienty („prežitie“ automobilu daného veku z jedného roku na druhý) sú odhadnuté vyššie spomínanými analýzami.

### 3.3. Analýza dát

Na začiatku celého modelovania bolo prioritné získať relevantné dáta, z ktorých by sme mohli vychádzať. Vedeli sme pomocou Greenspanovej a Cohenovj štúdie, že spoločnosť R.L. Polk & Co. publikuje potrebné dáta. Bohužiaľ, tieto údaje sú spoplatnené a verejne neprístupné od roku 2001. Avšak po dlhodobejšom bádání sme natrafili na Edície, ktoré vydáva *Oak Ridge National Laboratory(ORNL)*, čo sú mnoho účelné vedecko-technické národné laboratóriá, ktoré slúžia Ministerstvu Energetiky v USA [4]. V týchto edíciách, ktoré sa nazývajú *Transportation Energy Data Book*, sa nachádzajú podrobné informácie týkajúce sa StavuDopravy a Energie v USA. Pre našu prácu sme vybrali iba dáta opisujúce počet áut v prevádzke (Tabuľka 4). Na lepšie priblíženie obsahu Tabuľky 4 sa pokúsime opísať napríklad rok 1998. V danom roku bolo 8,049 milióna ročných áut, 7,564 mil. dvojročných atď. Takto vypísané dáta sme mohli ďalej spracovávať a vyhodnocovať.

Keďže našim cieľom bolo odhadovanie vyradovania áut z prevádzky, museli sme zo získaných dát vypočítať tzv. *Scrapage Rate*<sup>4</sup>. Ako vidíme v Tab.4, za najrelevantnejšie dáta môžeme považovať autá, ktoré majú 1 až 14 rokov. Automobily mladšie ako 1 rok a staršie ako 14 rokov sme nebrali do úvahy. Samotný výpočet môžeme demonštrovať na príklade áut vyrobených v roku 1996, ktoré v roku 1997 mali 1 rok a v roku 1998 analogicky 2. V prvom prípade je to počet 7,696 miliónov, v druhom 7,564 miliónov.

Rozdielom týchto dvoch hodnôt dostaneme počet áut, ktoré počas roku 1998 boli vyradené z prevádzky ( $7,696 - 7,564 = 0,132$ ). Môžeme povedať, že na konci roku 1998 bolo automobilov o 132 tisíc menej ako na konci roku predošlého. Ak podelíme tento počet všetkými jednoročnými autami v roku 1997 dostaneme pravdepodobnosť toho, že auto bude vyradené z prevádzky ( $0,132 / 7,696 = 0,017151767$ ). Postupnými, rovnakými výpočtami sme dostali všetky pravdepodobnosti, a tak sme zostavili súhrnnú tabuľku (Tabuľka 5) a vykreslili grafy.

---

<sup>4</sup> Pravdepodobnosť toho, že auto vyradíme z dopravnej prevádzky

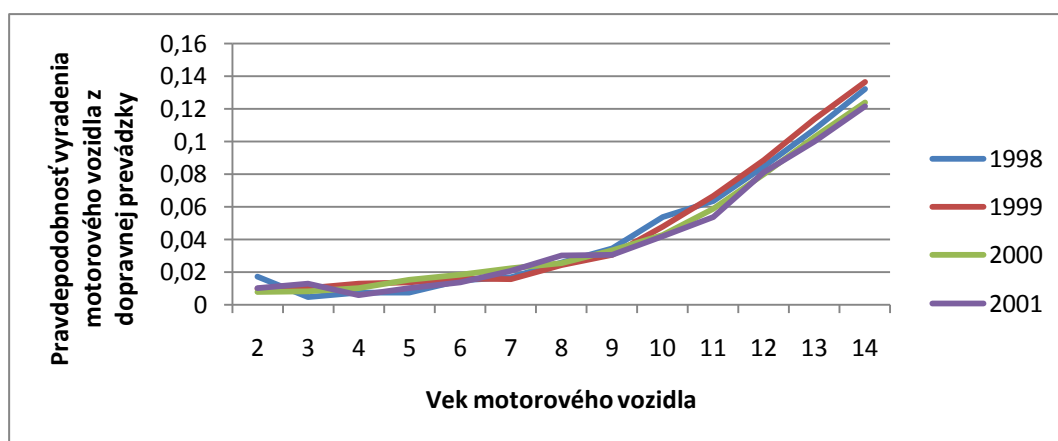
Tabuľka 4 : Počet áut v prevádzke v rokoch 1997 - 2001 zoradené podľa veku(v miliónoch)

Age (years)	1997	1998	1999	2000	2001
<b>Under 1</b>	5,622	5,619	6,219	6,665	6,183
<b>1</b>	7,696	8,049	7,714	8,177	8,882
<b>2</b>	8,968	7,564	7,971	7,655	8,093
<b>3</b>	7,938	8,926	7,488	7,906	7,555
<b>4</b>	8,013	7,878	8,811	7,413	7,860
<b>5</b>	7,430	7,953	7,771	8,675	7,337
<b>6</b>	7,665	7,320	7,826	7,628	8,555
<b>7</b>	7,821	7,536	7,204	7,650	7,471
<b>8</b>	8,479	7,620	7,354	7,021	7,420
<b>9</b>	8,463	8,187	7,387	7,109	6,807
<b>10</b>	7,944	8,008	7,797	7,071	6,810
<b>11</b>	7,504	7,439	7,475	7,338	6,692
<b>12</b>	6,469	6,870	6,780	6,876	6,742
<b>13</b>	5,342	5,774	6,089	6,084	6,189
<b>14</b>	3,365	4,636	4,987	5,334	5,345
<b>15 and older</b>	15,954	16,589	17,996	19,119	20,773
<b>TOTAL</b>	<b>124,673</b>	<b>125,968</b>	<b>126,869</b>	<b>127,721</b>	<b>128,714</b>

Tabuľka 5 : Pravdepodobnosti vyradenia automobilov v rokoch 1998 - 2001 v USA

Age (years)	1998	1999	2000	2001
<b>2</b>	0,017151767	0,009690645	0,007648	0,010272716
<b>3</b>	0,004683318	0,010047594	0,008155	0,013063357
<b>4</b>	0,007558579	0,012883711	0,010016	0,005818366
<b>5</b>	0,007487832	0,013582127	0,015435	0,01025226
<b>6</b>	0,014804845	0,015968817	0,018402	0,013832853
<b>7</b>	0,016829746	0,015846995	0,022489	0,020582066
<b>8</b>	0,025700038	0,024150743	0,025403	0,030065359
<b>9</b>	0,034438023	0,030577428	0,033315	0,030479989
<b>10</b>	0,053763441	0,047636497	0,042778	0,042059361
<b>11</b>	0,06356999	0,066558442	0,058869	0,053599208
<b>12</b>	0,084488273	0,088587176	0,080134	0,081221041
<b>13</b>	0,107435461	0,113682678	0,102655	0,09991274
<b>14</b>	0,13216024	0,136300658	0,123994	0,121466141

Obrázok 5 : Grafické zobrazenie získaných pravdepodobností



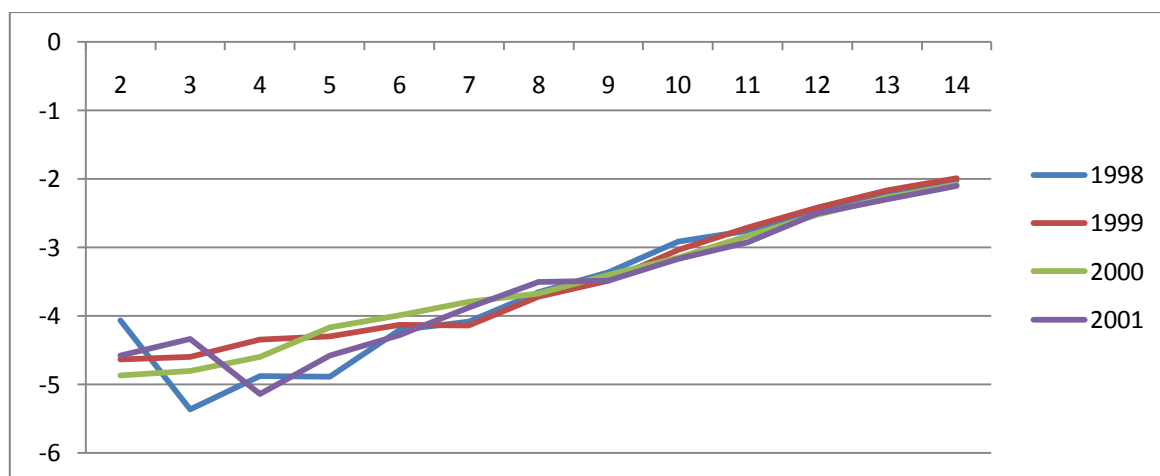
Na Obr.5 vidíme exponenciálne rastúce štyri krivky, ktoré nám zobrazujú pravdepodobnosti vyradenia automobilu z prevádzky v štyroch po sebe idúcich rokoch. Po dôkladnejšom preskúmaní týchto priamok sme sa rozhodli k preškálovaniu osi-y (pravdepodobnosť vyradenia motorového vozidla z dopravnej prevádzky), aby sme boli schopní lepšie rozoznať možné disturbancie. Z dôvodu podobnosti nášho grafu s grafom exponenciálnej funkcie ( $y = Ce^{ax}$ ) bolo jednoznačne najlepšie rozhodnutie nahradiť y funkciou prirodzeného logaritmu  $\ln(y)$ , kedy nám zo spomínanej exponenciálnej funkcie vznikla lineárna funkcia.

$$y = Ce^{ax}$$

$$\ln(y) = \ln C + \ln e^{ax}$$

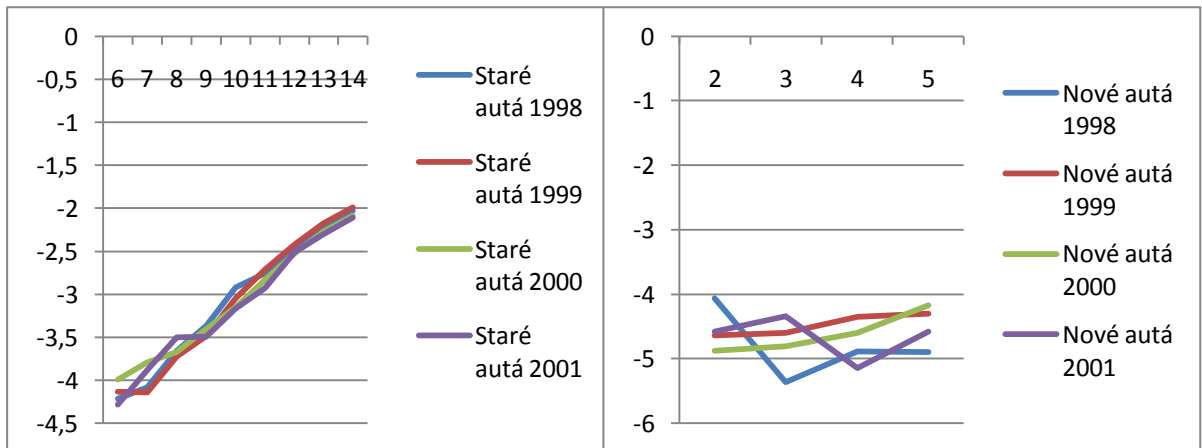
$$\ln(y) = ax + \ln C \quad (7)$$

Obrázok 6: Preškálované pravdepodobnosti vyradenia automobilu z prevádzky ( $y \rightarrow \ln(y)$ )



Graf z Obr.6 splnil naše očakávania a poskytuje nám lepšie zobrazenie zmien v údajoch. Prvá vec, ktorú môžeme pozorovať na grafe z Obr.6, je rozdiel tvaru funkcie päť a menej ročných automobilov a automobilov, ktoré sú v prevádzke dlhšie ako päť rokov. Z tohto dôvodu sme sa rozhodli, že pre ďalšie pozorovania rozdelíme dáta na dve časti.

Obrázok 7 : Prvá časť nášho rozdelenia (graf na ľavej strane) - Staré autá (age > 5)  
Druhá časť (graf na pravej strane) – Nové autá (age < 5)



### 3.3.1. Staršie automobily

Ako prvé sme sa rozhodli odhadovať staršie autá, keďže graf, ktorý popisuje túto skupinu je veľmi podobný priamke, ktorej všeobecný tvar (po preškálovaní) je  $\ln(y) = \alpha x + \ln C$ . Vidieť, že takáto priamka bude mať veľmi podobný tvar pre všetky roky (1998-2001) a predpokladáme, že sa bude líšiť iba v absolútnom člene  $\ln C$ , ktorý sa bude meniť v závislosti od času.

#### 3.3.1.1. Prvý prístup

Na odhadnutie prechodových koeficientov sme sa rozhodli zvoliť dva prístupy. Prvý prístup sa skladá z odhadnutia pomocou metódy najmenších štvorcov oboch koeficientov  $\alpha_i$  aj  $\beta_i$ , kde  $\beta_i$  reprezentuje absolútny člen  $\ln(C)$  z (7) pre i-ty rok. Keďže máme k dispozícii dáta  $\ln(y)$  (os y) a aj x (vek auta), tak nie je problém z takejto lineárnej regresii vyzistiť odhadnuté hodnoty koeficientov pomocou Metódy najmenších štvorcov (Tab. 6).

	1998	1999	2000	2001
<b>HATalfa</b>	0,2885115	0,2977559	0,2508861	0,2663676
<b>HATbeta</b>	-5,9664248	-6,070633	-5,5907298	-5,7943604

Tabuľka 6 : Odhadnuté hodnoty koeficientov  $\alpha_i$  a  $\beta_i$  pre  $i = 1998, 1999, 2000, 2001$

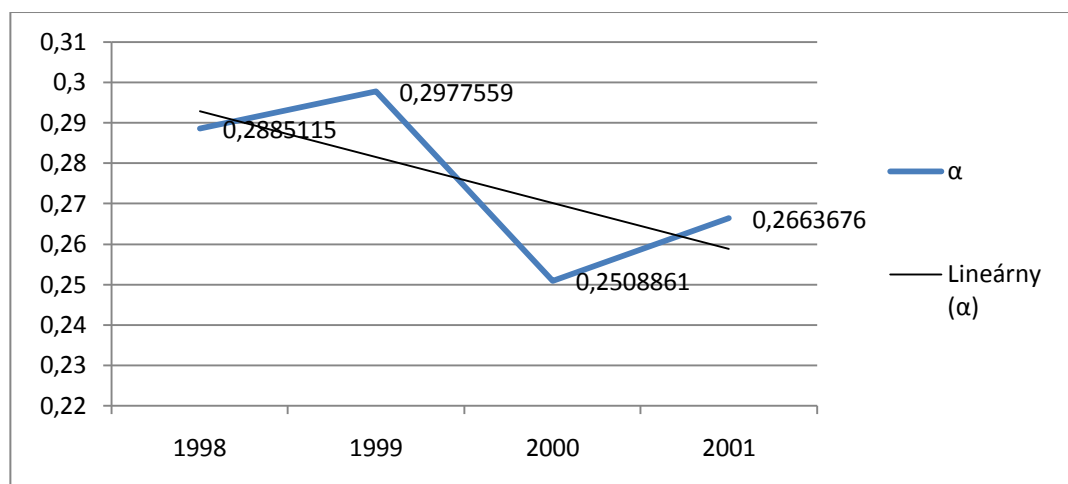
S takto odhadnutými koeficientmi máme k dispozícii 4 lineárne rovnice pre 4 roky, ku ktorým sme boli schopní nájsť dáta. Avšak aby sme vedeli odhadnúť vývoj tejto prvej časti, musíme pomocou metódy najmenších štvorcov odhadnúť vývoj koeficientov  $\alpha_i$  a  $\beta_i$  pre ďalšie roky. Znova sme vytvorili novú lineárnu regresiu (8), v ktorej musíme odhadnúť koeficienty  $\gamma$  a  $\delta$ , pomocou ktorých potom budeme schopní zistiť koeficientu  $\alpha_i$  pre nasledujúce roky.

$$HATalfa(Y) = \gamma Y + \delta \quad (8)$$

V regresii (8) nám  $Y$  predstavuje rok a  $HATalfa(Y)$  odhadnuté hodnoty  $\alpha_i$  z tabuľky 6 v roku  $Y$ . S takýmito vstupnými dátami sme pomocou softvéru R zistili odhadnuté hodnoty koeficientov  $\gamma$  a  $\delta$  pre regresiu (8).

	$\delta$	$\gamma$
<b>HATalfa</b>	22,9305152	-0,01133015

Tabuľka 7 : Odhadnuté koeficienty z lineárnej regresie (8)



Obrázok 8 : Hodnoty koeficienta  $\alpha_i$  v rokoch 1998 - 2001 s jeho odhadnutou priamkou



Rovnakú analógiu sme použili pre HATbeta a odhadli sme aj pre tento koeficient lineárnu regresiu, ktorej výsledky môžeme vidieť v tabuľke 8.

	koeficient	konštanta
<b>HATbeta</b>	0,09960964	-205,025012

*Tabuľka 8 : Odhady pre HATbeta*

Keď sme odhadli koeficienty v lineárnych regresiach, ktorými popisujem vývoj koeficientov z rovnice (7), boli sme schopní určiť  $\alpha_i$  a  $\beta_i$  pre ďalšie roky.

	$\beta_i$	$\alpha_i$
<b>1999</b>	-5,90534182	0,28154535
<b>2000</b>	-5,80573218	0,2702152
<b>2001</b>	-5,70612254	0,25888505
<b>2002</b>	-5,6065129	0,2475549
<b>2003</b>	-5,50690326	0,23622475

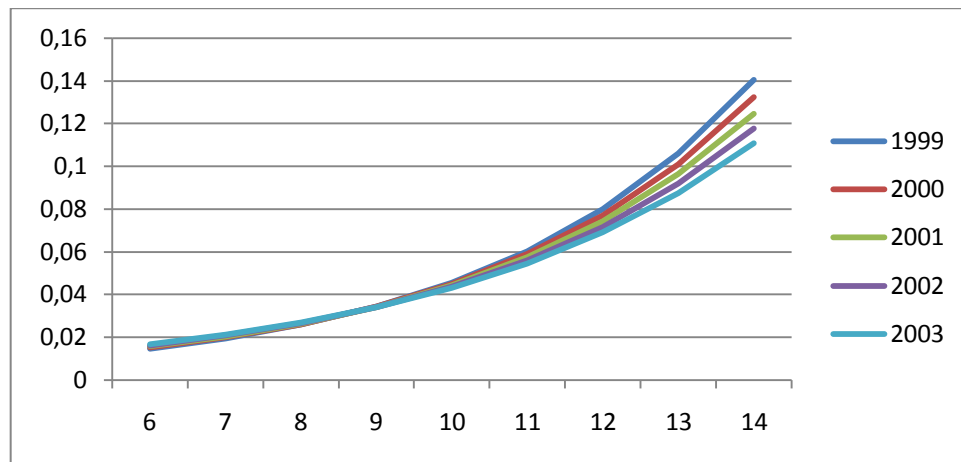
*Tabuľka 9 : Odhady pre rovnicu (7) pre roky 1999 – 2003*

Na záver, po odhadnutí všetkých potrebných regresii a koeficientov v nich, sme mohli vypočítať dáta pre roky, pre ktoré nie sú voľne prístupné údaje. Samozrejme ide o prvú časť, a teda staršie autá, a nemôžeme taktiež zabudnúť, že pracujeme s preškálovanými dátami. Aby sme si mohli počas výpočtov overiť náš postup, rozhodli sme sa, že budeme odhadovať zistené roky, a síce 1999, 2000 a 2001.

	<b>1999</b>	<b>2000</b>	<b>2001</b>	<b>2002</b>	<b>2003</b>
<b>6</b>	-4,21606972	-4,18444098	-4,15281224	-4,1211835	-4,08955476
<b>7</b>	-3,93452437	-3,91422578	-3,89392719	-3,8736286	-3,85333001
<b>8</b>	-3,65297902	-3,64401058	-3,63504214	-3,6260737	-3,61710526
<b>9</b>	-3,37143367	-3,37379538	-3,37615709	-3,3785188	-3,38088051
<b>10</b>	-3,08988832	-3,10358018	-3,11727204	-3,1309639	-3,14465576
<b>11</b>	-2,80834297	-2,83336498	-2,85838699	-2,883409	-2,90843101
<b>12</b>	-2,52679762	-2,56314978	-2,59950194	-2,6358541	-2,67220626
<b>13</b>	-2,24525227	-2,29293458	-2,34061689	-2,3882992	-2,43598151
<b>14</b>	-1,96370692	-2,02271938	-2,08173184	-2,1407443	-2,19975676

*Tabuľka 10 : Hodnoty logaritmov pravdepodobností vyradovania starších automobilov v rokoch 1999 - 2003 odhadnuté našimi rovnicami*

Posledným krokom, ktorý sme museli spraviť v tejto časti, bolo spätné preškáľovanie nami získaných dát. Keďže preškáľovanie znamenalo pre nás pretransformovať os-y na  $\ln(y)$ , spätné preškáľovanie znamená vypočítať exponenciálne hodnoty Tabuľky 10. Zobrazené ich vidíme na Obrázku 8 a môžeme vidieť, že naše očakávanie je splnené, a teda, čím je automobil novší, tým sa pravdepodobnosť zošrotovania znižuje.



*Obrázok 9 : Získané pravdepodobnosti vyradenia starších automobilov z dopravnej prevádzky od roku 1999 po rok 2003*

Staršie automobily (6 – 14 ročné) sme odhadli pomocou lineárnej regresie, keďže ich tvar bol veľmi blízky priamke. Pomocou metódy najmenších štvorcov sme zistili koeficienty v tejto regresii iba pre štyri roky, pre ktoré sme získali dáta. Aby sme boli schopní predpovedať pravdepodobnosť vyradenia staršieho automobilu do ďalších rokov, museli sme samotné koeficienty istým spôsobom odhadnúť. Znova sme použili MNŠ, ktorá nám dala odhadnuté hodnoty koeficientov lineárnej rovnice, podľa ktorej sme mohli vypočítať koeficienty rovnice priamky pre staršie automobily. Takto sme boli schopní zistiť hodnoty pravdepodobnosti vyradenia staršieho automobilu (Obrázok 9) pre iné roky, ako sme získali údaje. V Tabuľke 11 vidíme porovnanie hodnôt z rokov, pre ktoré sme údaje získali, ale zároveň sme ich aj namodelovali.

	1999	1999	2000	2000	2001	2001
6	0,015969	0,014757	0,018402	0,015231	0,013833	0,01572
7	0,015847	0,019555	0,022489	0,019956	0,020582	0,020365
8	0,024151	0,025914	0,025403	0,026147	0,030065	0,026383
9	0,030577	0,03434	0,033315	0,034259	0,03048	0,034179
10	0,047636	0,045507	0,042778	0,044888	0,042059	0,044278
11	0,066558	0,060305	0,058869	0,058815	0,053599	0,057361
12	0,088587	0,079915	0,080134	0,077062	0,081221	0,074311
13	0,113683	0,105901	0,102655	0,10097	0,099913	0,096268
14	0,136301	0,140337	0,123994	0,132295	0,121466	0,124714

Tabuľka 11: Porovnanie nami získaných údajov (zvýraznené) a údajov z *Transportation energy data book*

### 3.3.1.2. Druhý prístup

Ako sme spomínali, na odhadnutie prechodových koeficientov sme zvolili dva prístupy. Prvý je podrobne rozobratý v predošlej kapitole a druhý uvedieme v nasledujúcej. Postup riešenia v tomto druhom prípade je veľmi podobný tomu prvému. Pomocou metódy najmenších štvorcov sme rovnako ako predtým odhadli hodnoty prechodového koeficientu  $\alpha_i$  pre roky, pre ktoré sme vedeli získať údaje (tabuľka 6). Rozdiel týchto dvoch prístupov bude vo fixovaní koeficientu  $\alpha_i$ , ktorý nahradíme priemerom z odhadnutých hodnôt (Tabuľka 12).

<b><math>\alpha</math> fixná</b>	<b>0,27588</b>
----------------------------------	----------------

Tabuľka 12 : Priemer odhadnutých  $\alpha_i$  z tabuľky 6

S takto fixne zvolenou  $\alpha$ , dopočítame hodnoty  $\beta_i$  pomocou rovnice (7). Získame tak hodnoty  $\beta_i$  pre štyri roky. Priemerom pre každý rok (1998-2001) dostaneme štyri hodnoty konštantného člena z rovnice (7). Avšak rovnako ako v prvom prístupe musíme hodnoty  $\beta_i$  určiť aj na nasledujúce roky. Tento výpočet bude rovnaký ako v prvom prístupe a teda pomocou metódy najmenších štvorcov odhadneme koeficienty lineárnej regresie, ktorá nám následne bude slúžiť na výpočet hodnôt  $\beta_i$  pre ďalšie roky. Posledným krokom bude pomocou oboch koeficientov (fixná  $\alpha$ ,  $\beta$ ) dopočítať hodnoty pravdepodobnosti vyradenia automobilu pre dlhšie obdobie (Tabuľka 13).

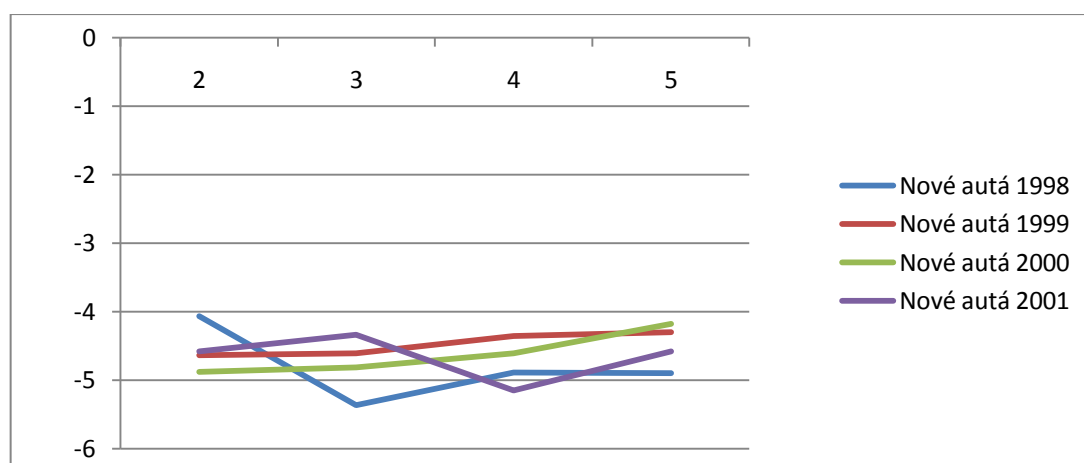
	1998	1999	2000	2001	2002	2003
6	-4,1797189	-4,19341	-4,2071	-4,2208	-4,23449	-4,24818
7	-3,903838625	-3,91753	-3,93122	-3,94491	-3,95861	-3,9723
8	-3,62795835	-3,64165	-3,65534	-3,66903	-3,68273	-3,69642
9	-3,352078075	-3,36577	-3,37946	-3,39315	-3,40685	-3,42054
10	-3,0761978	-3,08989	-3,10358	-3,11727	-3,13097	-3,14466
11	-2,800317525	-2,81401	-2,8277	-2,84139	-2,85509	-2,86878
12	-2,52443725	-2,53813	-2,55182	-2,56551	-2,57921	-2,5929
13	-2,248556975	-2,26225	-2,27594	-2,28963	-2,30333	-2,31702
14	-1,9726767	-1,98637	-2,00006	-2,01375	-2,02744	-2,04114

Tabuľka 13 : Hodnoty pravdepodobnosti vyradenia automobilu pre dlhšie obdobie

### 3.3.2. Novšie automobily

#### 3.3.2.1. Prvý prístup

Druhá časť, ktorú sme museli namodelovať, bola zložená z motorových vozidiel, ktoré majú dva až päť rokov (Obrázok 10). Rozhodli sme sa pre takéto delenie z dôvodu, že hodnoty pravdepodobnosti vyradenia automobilov v tomto veku nepripomínajú priamku, a preto aj prístup k namodelovaniu bude iný.



Obrázok 10 : Preškáľované hodnoty pravdepodobnosti vyradenia automobilov vo veku od dvoch do piatich rokov

Po analýze dát z grafu z Obrázka 10 vidíme, že krivky majú dosť nepredvídateľný charakter. Na druhej strane však môžeme pozorovať, že za najbližší tvar možno považovať parabolu. Z tohto dôvodu sme sa rozhodli interpolovať daný graf kvadratickým splajnom. K tomu sme potrebovali tri ľubovoľné body z grafu z Obrázka 10. K dispozícii máme štyri,

a preto sme vybrali také body, ktoré nám, na jednej strane, čo najlepšie resp. najpresnejšie dokážu opísať parabolický tvar, a na druhej strane sme ich nepovažovali za „outlierov“<sup>5</sup>. Pre roky 1998 a 2001 to boli body 2, 4 a 5 a pre ostatné roky, teda 1999 a 2000, to boli body 2, 3 a 5. Vybraté body sme následne mohli interpolovať kvadratickým splajnom. Princíp interpolovania nie je náročný. Pointou je jednoduchými výpočtami zistiť hodnoty koeficientov  $a$ ,  $b$  a  $c$  z interpolačného polynómu, v tomto prípade druhého rádu (9).

$$ax^2 + bx + c \quad (9)$$

	1998	1999	2000	2001
<b>a</b>	0,271572022	0,038179522	0,08498719	0,28390379
<b>b</b>	-2,177278113	-0,154725389	-0,3608592	-1,987991
<b>c</b>	-0,797385933	-4,479861624	-4,491485	-1,737897

*Tabuľka 14 : Hodnoty koeficientov  $a$ ,  $b$  a  $c$  odhadnuté pomocou interpolácie kvadratickým splajnom*

Takto nám vznikli štyri kvadratické funkcie, ktoré opisujú tvar kriviek vykreslených na Obrázku 10. Avšak, aby sme mohli povedať, ako budú tieto funkcie vyzerat' v nasledujúcich rokoch, musíme ich tak ako pri starších autách odhadnúť pomocou metódy najmenších štvorcov. Postup je analogický ako v kapitole 3.3.1, a preto uvediem iba výsledné koeficienty lineárnych regresí (Tabuľka 14) a koeficienty interpolačných polynómov pre roky 1999 – 2003 (Tabuľka 15).

	konštanta	koeficient
<b>a</b>	-16,586746	0,008380298
<b>b</b>	-73,49762621	0,03617275
<b>c</b>	563,612994	-0,2833157

*Tabuľka 15 : Koeficienty a konštanty z lineárnej regresie, ktorou sme odhadovali koeficienty kvadratického splajnu*

<sup>5</sup> outlier – dátový bod, ktorý ma veľmi rozdielnú hodnotu od ostatných; štatisticky je nevýznamný

	a	b	c
<b>1999</b>	0,165469699	-1,18829896	-2,7350903
<b>2000</b>	0,173849997	-1,15212621	-3,018406
<b>2001</b>	0,182230295	-1,11595346	-3,3017217
<b>2002</b>	0,190610593	-1,07978071	-3,5850374
<b>2003</b>	0,198990891	-1,04360796	-3,8683531

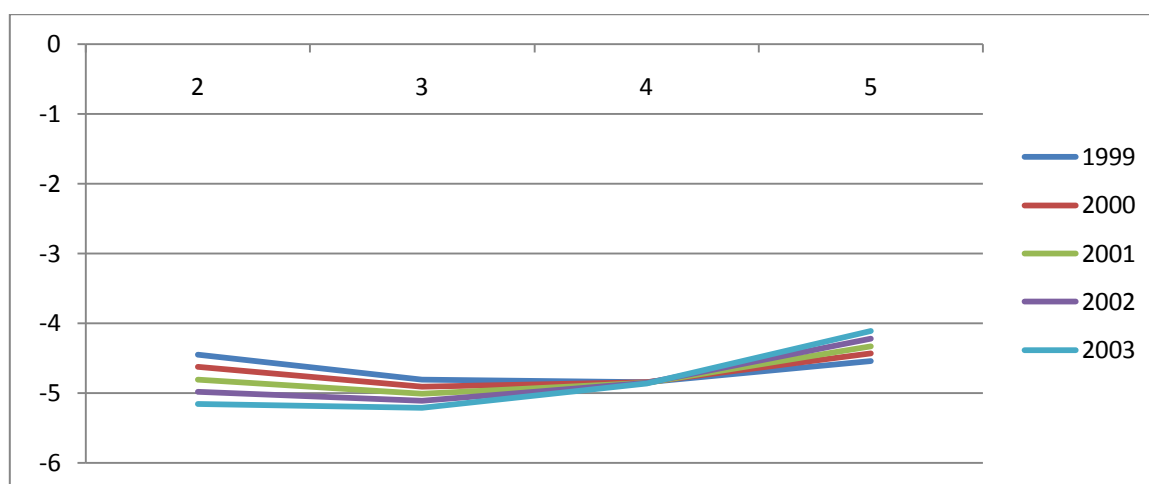
*Tabuľka 16 : Hodnoty koeficientov interpolačných polynómov pre roky 1999 – 2003*

Koeficienty  $a$ ,  $b$  a  $c$  nám následne poslúžili na vygenerovanie dát pre roky 1999 až 2003, ktoré sme si zvolili. V Tabuľke 16 vidíme porovnanie údajov pre roky 1999, 2000 a 2001. Na Obrázku 11 sme vykreslili grafy, ktoré nami vygenerované dáta zobrazujú.

	1999	1999	2000	2000	2001	2001
<b>2</b>	-4,63659	-4,44981	-4,87325	-4,62726	-4,57826	-4,80471
<b>3</b>	-4,60042	-4,81076	-4,80918	-4,91013	-4,33794	-5,00951
<b>4</b>	-4,35179	-4,84077	-4,60357	-4,84531	-5,14674	-4,84985
<b>5</b>	-4,299	-4,53984	-4,1711	-4,43279	-4,58026	-4,32573

*Tabuľka 17 : Porovnanie dát našich a reálnych*

V Tabuľke 17 môžeme pozorovať, že rozdiely medzi nami získanými dátami a dátami reálnymi sú väčšie ako to bolo v prípade starších automobilov. Je to spôsobené najmä nepravidelnosťou reálnych údajov a väčším množstvom outlierov.



*Obrázok 11 : Grafy hodnôt pre roky 1999 až 2003 vygenerované našimi odhadmi pre novšie autá*

### 3.3.2.2. Druhý prístup

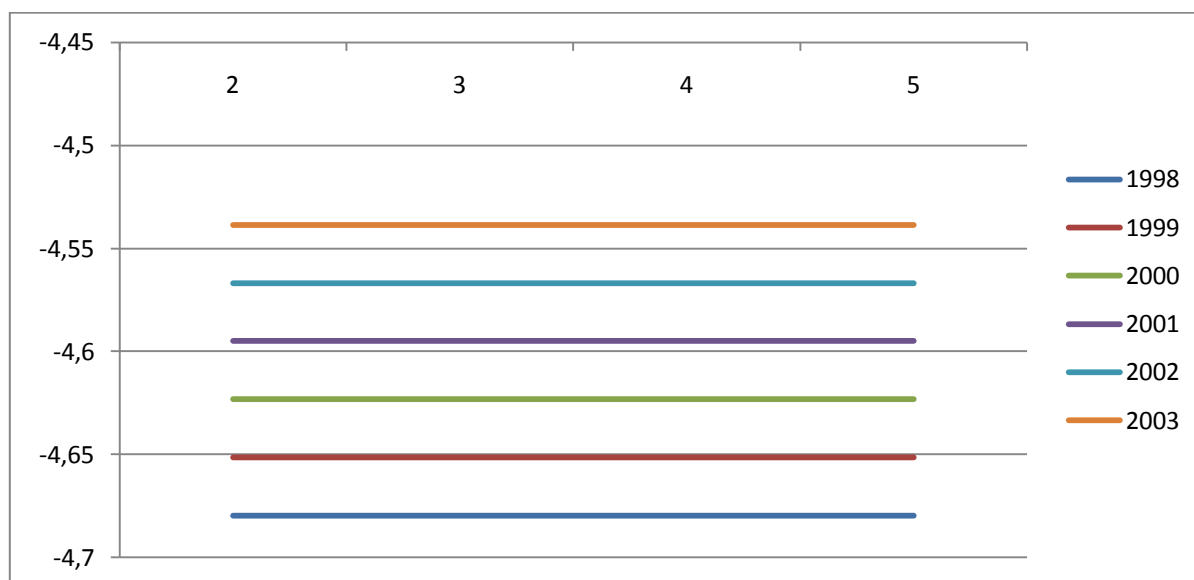
Druhý prístup, ktorý sme zvolili pri výpočte novších automobilov, vychádza z Greenspanovej práce, v ktorej môžeme nájsť graf pravdepodobnosti zotrvania automobilu v prevádzke (Obrázok 2). Na tomto grafe môžeme vidieť, že novšie autá zostávajú v dopravnej prevádzke s veľkou konštantnou pravdepodobnosťou. Z tohto dôvodu sme sa rozhodli určiť novšie automobily ako *konštantné funkcie*, ktorých tvar bude hodnota získaná ako priemer dát, ktoré máme k dispozícii.

1998	1999	2000	2001
-4,80224	-4,47195	-4,61428	-4,6608

Tabuľka 18 : Hodnoty konštantných funkcií pre novšie automobily získané ako priemer dostupných údajov

Pomocou MNŠ sme znova zistili hodnoty lineárnej regresie a prostredníctvom nej odhadneme tieto konštantné funkcie do budúcnosti.

Obrázok 12 : Novšie automobily - druhý prístup od roku 1998-2003



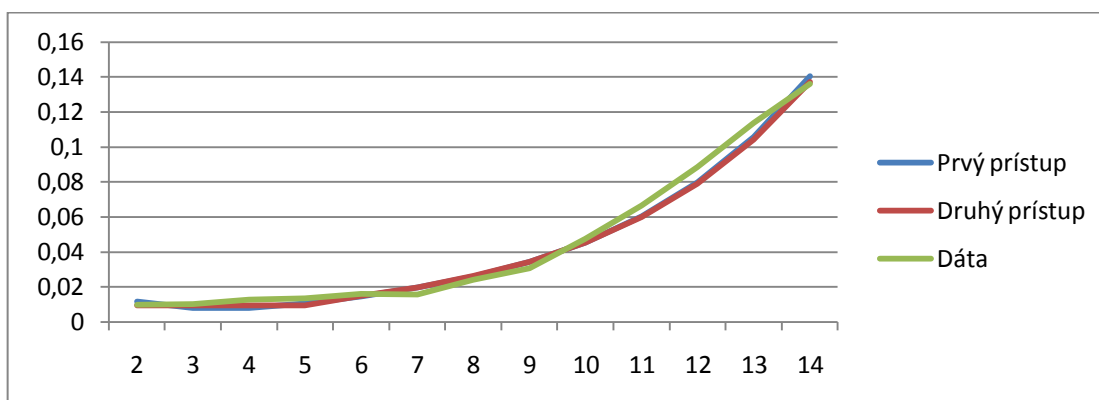
### 3.4. Výsledky a ich interpretácia

#### 3.4.1. Porovnanie oboch prístupov s reálnymi dátami

Po konečnom odhadnutí a namodelovaní všetkých vekových kategórií automobilov v oboch prístupoch môžeme zhrnúť všetky výsledky, ku ktorým sme prišli a porovnať ich s dátami, ktoré sme mali k dispozícii. Keďže údaje sme mali iba k rokom 1998-2001, rozhodli sme sa porovnať všetky hodnoty v roku 1999.

1999			
	Prvý prístup	Druhý prístup	Dáta
2	0,01168079	0,009548077	0,009690645
3	0,00814167	0,009548077	0,010047594
4	0,00790096	0,009548077	0,012883711
5	0,01067509	0,009548077	0,013582127
6	0,01475653	0,01509471	0,015968817
7	0,019555	0,019890149	0,015846995
8	0,02591382	0,026209053	0,024150743
9	0,03434037	0,034535409	0,030577428
10	0,04550704	0,045506966	0,047636497
11	0,06030484	0,059964079	0,066558442
12	0,07991453	0,079014072	0,088587176
13	0,10590082	0,104116061	0,113682678
14	0,14033724	0,137192702	0,136300658

Tabuľka 19 : Pravdepodobnosť vyradenie automobilu v roku 1999 prvým a druhým prístupom v porovnaní s reálnymi dátami



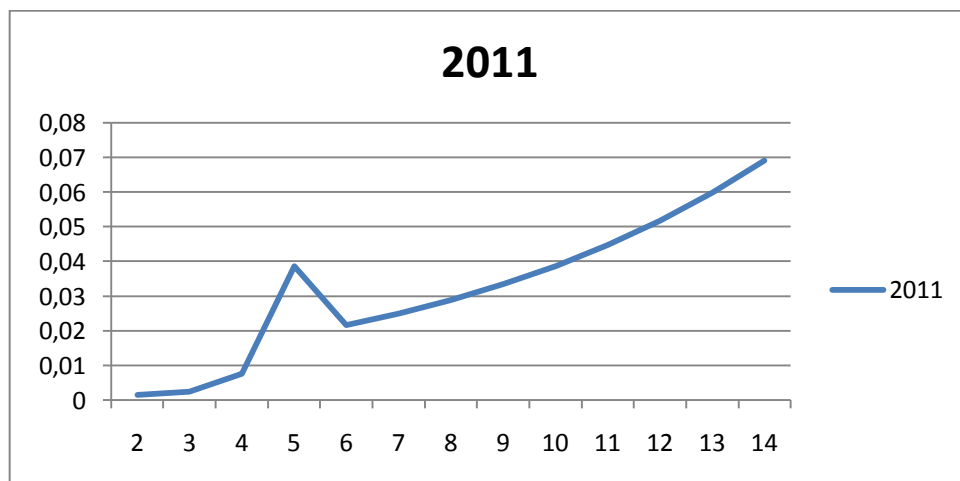
Obrázok 13 : Graf k tabuľke 18



Ak by sme mali za úlohu opísať dané krivky a hodnoty, povedali by sme, že oba prístupy veľmi dobre popisujú vekové rozloženie, resp. pravdepodobnosť vekového rozloženia automobilov v roku 1999 v USA. Vidíme, že odchýlky od skutočných hodnôt sú veľmi malé. Je zaujímavé, že oba prístupy, ktoré sme zvolili nám dali podobné hodnoty a môžeme pozorovať, že zo začiatku voľba konštantnej funkcie z druhého prístupu bola správna. Neskôr, t.j. v okolí štvrtého, piateho roku života, sa odchýlky zväčšujú, a v tomto prípade sa začína zdať prvý prístup lepším. Neskôr a ku koncu dát sú hodnoty z oboch prístupov veľmi blízke, a ak by sme si mali zvoliť práve jeden prístup, tak z hľadiska jednoduchosti výpočtov by to bol práve druhý.

### 3.4.2. Predikcia vekového rozloženia do ďalších rokov

V našich modelovaniach sme sa rozhodli pre dva prístupy, ako pozerat' na pravdepodobnosť vyradovania automobilov. Líšia sa najmä vo výpočtovej časti. Prvý prístup poskytuje výpočtovo zložitejší postup, ktorý má však jeden vážny nedostatok. Chybu môžeme pozorovať pri vykreslení grafu pre neskorší rok, t.j. 2011, kedy nám pre päť ročné automobily vzniká veľký odskok dát (Obrázok 14).



Obrázok 14 : Pravdepodobnosť vyradenia automobilu odhadnuté prvým prístupom

Z tohto dôvodu sme sa rozhodli, že predikciu počtu áut v dopravnej prevádzke budeme robiť pomocou druhého prístupu. K tomu, aby sme boli schopní kompletne predikovať vekové rozloženie v dopravnej prevádzke na roky po 2002, museli sme zistiť aký bol predaj nových áut pre čo najväčšie obdobie. V edíciách, ktoré spomíname v časti 3.3, sa

nachádzajú údaje aj o predajnosti automobilov v USA. Pre našu predikciu sme si vybrali Edíciu 29, z ktorej sme boli schopní získať dáta na obdobie 1980 – 2009 [9]. Keď už sme mali dáta k dispozícii, mohli sme zistiť hodnoty vekového rozloženia automobilov pre roky 1998 – 2011. Z dôvodov nedostatku miesta pracovného listu uvádzame našu predikciu iba pre roky 2002 – 2009, ktorá je avšak postačujúca.

	2002	2003	2004	2005	2006	2007	2008	2009
<b>1</b>	8,423	8,103	7,610	7,545	7,720	7,821	7,618	6,813
<b>2</b>	8,790	8,333	8,014	7,524	7,457	7,628	7,725	7,522
<b>3</b>	8,009	8,696	8,241	7,923	7,436	7,368	7,534	7,627
<b>4</b>	7,476	7,923	8,600	8,148	7,831	7,347	7,277	7,438
<b>5</b>	7,778	7,397	7,836	8,503	8,053	7,737	7,257	7,185
<b>6</b>	7,231	7,667	7,292	7,727	8,386	7,944	7,634	7,162
<b>7</b>	8,392	7,095	7,525	7,159	7,588	8,237	7,805	7,502
<b>8</b>	7,283	8,183	6,921	7,343	6,988	7,409	8,046	7,626
<b>9</b>	7,174	7,045	7,920	6,701	7,113	6,772	7,183	7,804
<b>10</b>	6,510	6,865	6,746	7,588	6,424	6,823	6,499	6,898
<b>11</b>	6,418	6,140	6,481	6,373	7,174	6,078	6,461	6,160
<b>12</b>	6,185	5,938	5,687	6,009	5,915	6,666	5,654	6,016
<b>13</b>	6,068	5,575	5,361	5,142	5,440	5,363	6,053	5,141
<b>14</b>	5,374	5,280	4,861	4,683	4,501	4,772	4,713	5,328
<b>Suma</b>	101,111	100,240	99,094	98,367	98,028	97,966	97,459	96,221

*Tabuľka 20 : Predikcia druhým prístupom vekového rozloženia automobilov v dopravnej prevádzke*

Výpočet hodnôt v Tabuľke 20 ilustrujeme na príklade počtu trojročného automobilu v roku 2001. Z Tabuľky 4 vieme, že tento počet bol 7,555 miliónov motorových vozidiel. Ďalej vieme, že tieto automobily budú v roku 2002 štvorročné, a preto pomocou predikovanej hodnoty pravdepodobnosti vyradenia automobilu z prevádzky vieme povedať, že ich bude  $7,555 - (7,555 * (\text{pravdepodobnosť vyradenia trojročného automobilu v roku 2001}))$ . Analogický postup sme spravili do roku 2011. Je samozrejmé, že počty áut budú klesať, čo môžeme pozorovať v našej Tabuľke 20. Čo však nie je očakávané, je klesajúci trend celkového počtu automobilov. Tento fakt je spôsobený nezahrnutím počtu automobilov, ktoré mali viacej ako 15, resp. 14 rokov, ktorých počet stúpa rýchlejšie ako celkové počty klesajú, a preto by aj celkový počet v konečnom dôsledku stúpал.

## Záver

Cieľom bakalárskej práce bolo poskytnúť nový pohľad na modelovanie vekového rozloženia v dopravnej prevádzke. Vo všeobecnosti neexistuje veľa štúdií, ktoré by pojednávali na túto tému, pretože väčšina dát je spolplatnených. Avšak v tejto práci sme sa pokúsili ukázať dva prístupy, ako možno z malého množstva údajov vygenerovať matematický model. Pomocou metódy najmenších štvorcov sme odhadli všetky potrebné koeficienty v lineárnych regresióch, ktoré sme na odhady použili a vytvorili tak dva rozličné postupy, ktorými sme vygenerovali počty automobilov, v rôznych rokoch, v rôznom veku.

Prvá kapitola obsahuje podrobné rozobratie existujúcej štúdie, ktorá nám bola jedinou inšpiráciou. Pokúsili sme sa vysvetliť ako autori danej práce zmýšľali, a čo bolo ich motiváciou. V druhej kapitole sme sa zmienili o možných chybách, ktoré sme v existujúcej štúdií vyzozorovali, a k nim sme sa pokúsili uviesť aj vhodné riešenia. Na záver tejto kapitoly sme sa pokúsili o zreprodukovanie celého modelovania, ktoré autori spravili. Prišli sme k záveru, že autori štúdie museli mať k dispozícii relevantnejšie dáta, pretože pomocou tých, ktoré sme my získali, sme neboli schopní vypočítať ani prvé základné koeficienty. Tretia kapitola sa skladá z dvoch základných častí. Prvou je analýza dát, kde sme vysvetlili odkiaľ a v akej forme sme získali dáta. Druhou časťou je matematický model založený na diskretnom demografickom modeli. Pomocou dostupných údajov sme sa pokúsili zostaviť matematický model, ktorým by sme vedeli predikovať dáta do ďalších rokov. Zvolili sme dva prístupy odhadu parametrov a na koniec tejto záverečnej kapitoly sme ich porovnali a s lepším z nich sme vygenerovali dáta, vekové rozloženie dopravnej prevádzky, do roku 2011. týmto spôsobom sme zodpovedali základnú otázku, ktorú sme si položili v úvode tretej kapitoly.

Dôležitým prínosom predkladanej bakalárskej práce je vygenerovanie matematického modelu, ktorý sa ešte nikto na vedeckej úrovni nepokúsil vytvoriť. Získané výsledky naznačujú, že model je správny a s jeho pomocou sme schopní zistiť počty automobilov rôzneho vekového rozloženia. Nie je však dokonalý, pretože dáta ktoré sme mali k dispozícii nie sú rovnomerné a konzistentné.

## Zoznam použitej literatúry

- [1] GREENSPAN, A. 1996. Motor Vehicle Stocks, Scrappage, and Sales. [online]. 1996. Washington D.C. [cit. 2011.15.05.] Dostupné na internete : <<http://www.federalreserve.gov/pubs/feds/1996/199640/199640pap.pdf>>. (preklad autora)
- [2] R.L. Polk & Co. [online]. 2011. [cit. 2011.20.05.] Dostupné na internete:<[http://en.wikipedia.org/wiki/R.L.\\_Polk\\_%26\\_Co.](http://en.wikipedia.org/wiki/R.L._Polk_%26_Co.)>. (preklad autora)
- [3] United States Census Bureau. [online]. 2011 [cit. 2011.20.05.] Dostupné na internete : < [http://en.wikipedia.org/wiki/Census\\_Bureau](http://en.wikipedia.org/wiki/Census_Bureau) >. (preklad autora)
- [4] Oak Ridge National Laboratory. [online]. 2011 [cit. 2011.20.05.] Dostupné na internete : < [http://en.wikipedia.org/wiki/Oak\\_Ridge\\_National\\_Laboratory](http://en.wikipedia.org/wiki/Oak_Ridge_National_Laboratory) >. (preklad autora)
- [6] Alan Greenspan. [online]. 2011 [cit. 2011.21.05.] Dostupné na internete : <<http://www.nndb.com/people/164/000023095/>>.
- [7] STACY, C. – SUSAN, W. 2007. Transportation Energy Data Book - Edition 26. [online]. 2007. [cit. 2011.23.05.] Dostupné na internete : <[http://www.realworldengineering.org/download.php?proj=7&file=RWEP\\_Final6\\_20\\_Edition26\\_Full\\_Doc.pdf](http://www.realworldengineering.org/download.php?proj=7&file=RWEP_Final6_20_Edition26_Full_Doc.pdf)>.
- [8] FAIZ, A. – WEAVER, CH. – WALSH, M. 2006. Air Pollution from Motor Vehicles. [online]. 1996. Washington D.C. [cit. 2011.24.05.] Dostupné na internete : <[http://books.google.sk/books?id=Hqsyv\\_KD0lgC&printsec=frontcover#v=onepage&q&f=false](http://books.google.sk/books?id=Hqsyv_KD0lgC&printsec=frontcover#v=onepage&q&f=false)>.
- [9] STACY, C. – SUSAN, W. – BOUNDY, R. 2007. Transportation Energy Data Book - Edition 29. [online]. 2010. [cit. 2011.25.05.] Dostupné na internete : <[http://cta.ornl.gov/data/tedb29/Edition29\\_Chapter04.pdf](http://cta.ornl.gov/data/tedb29/Edition29_Chapter04.pdf)>.