



Vývoj obchodní letecké přepravy cestujících v ČR

Časové řady

Seminář výpočetní statistiky

Vypracoval:	Zdeněk Styblík
Zařazení:	1. ročník INFONK, PEF ČZU Praha
Cvičení:	kombinované

Obsah

Vývoj obchodní letecké přepravy cestujících v ČR.....	1
Úvod.....	3
Řešení.....	4
Závěr.....	10

Úvod

Tématem mé práce je vývoj obchodní letecké dopravy cestujících v ČR v časovém období 1995-2012 s predikcí na 4 roky, tzn. predikcí pro období 2013-2016.

Data byla získána z webových stránek ČSÚ. Rozsah dat je od roku 1995 do roku 2012 včetně, a reprezentují sumu pasažérů přepravených ve vnitrostátní a mezinárodní letecké dopravě za dané období, resp. za daný rok.

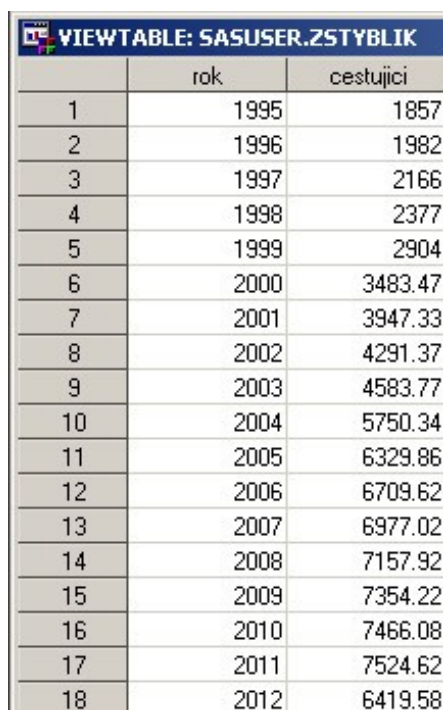
Rok	Cestující v tis.
1995	1857.00
1996	1982.00
1997	2166.00
1998	2377.00
1999	2904.00
2000	3483.47
2001	3947.33
2002	4291.37
2003	4583.77
2004	5750.34
2005	6329.86
2006	6709.62
2007	6977.02
2008	7157.92
2009	7354.22
2010	7466.08
2011	7524.62
2012	6419.58

Tabulka 1: počet cestujících v tisících

Data z roku 2013 byla pro neúplnost a nepřesnost vynechána.

Řešení

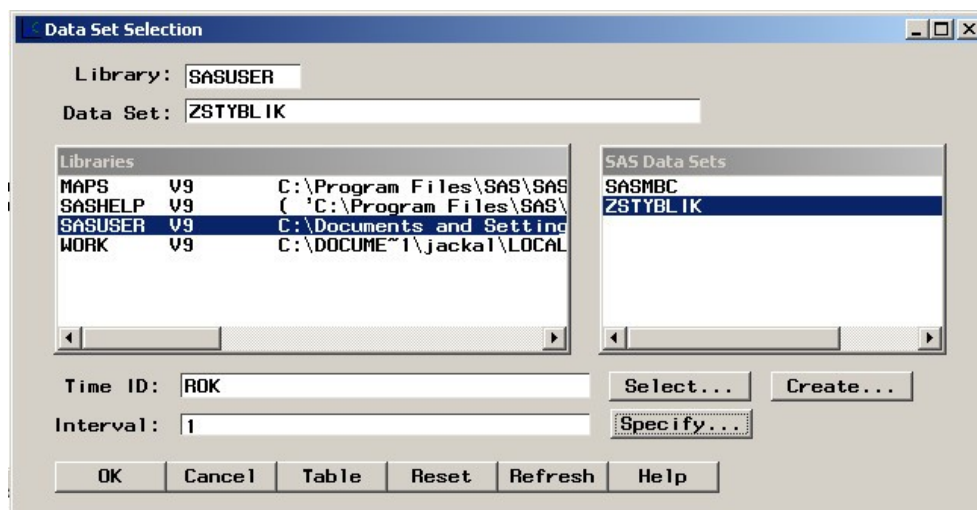
Pro řešení daného problému jsem použil aplikaci SAS v9.2. Data do aplikace jsem nainportoval ze souboru ve formátu CSV, protože SAS v9.2 nepodporuje formát OpenOffice. Alternativním řešením bylo zadat data do aplikace ručně.



	rok	cestujici
1	1995	1857
2	1996	1982
3	1997	2166
4	1998	2377
5	1999	2904
6	2000	3483.47
7	2001	3947.33
8	2002	4291.37
9	2003	4583.77
10	2004	5750.34
11	2005	6329.86
12	2006	6709.62
13	2007	6977.02
14	2008	7157.92
15	2009	7354.22
16	2010	7466.08
17	2011	7524.62
18	2012	6419.58

Obrázek 1: data v programu SAS

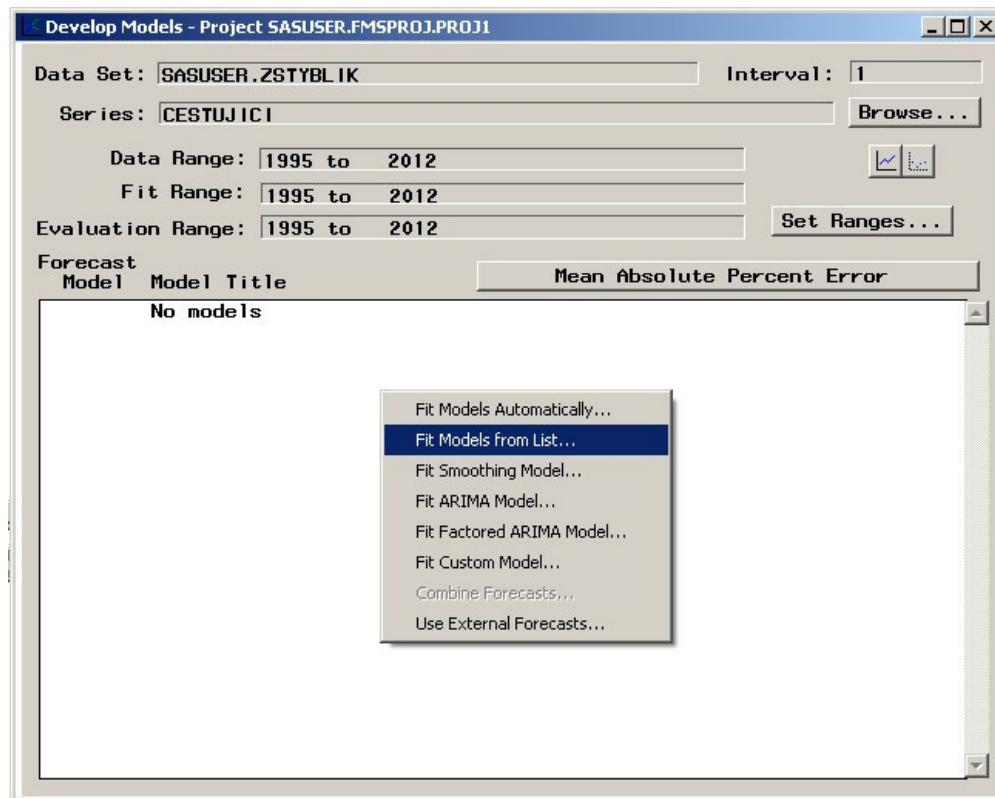
Po zadání dat spustíme modul pro výpočet předpovědi časových řad(*Time Series Forecasting System*). V modulu(Obrázek 2) nejdříve vybereme dataset do kterého jsme uložili data. Potom vybereme sloupec, který reprezentuje časovou proměnnou(sloupec pojmenovaný *rok*), a v jakém intervalu. V našem případě zvolíme interval 1.



Obrázek 2: výběr dat a časové proměnné

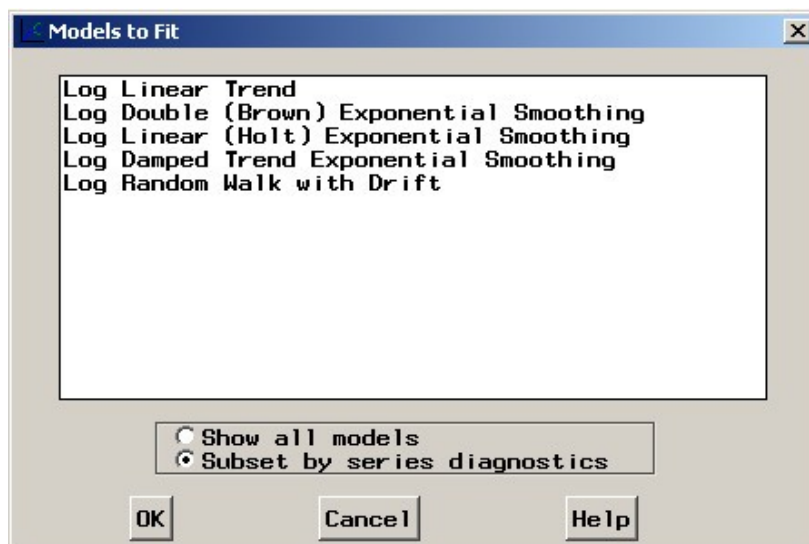
Jako další krok máme možnost provést výběr vhodných modelů ručně (*Develop Models*), nebo přenechat volbu na aplikaci (*Fit Models Automatically*). Zvolíme první možnost a v následujícím dialogovém okně vybereme sloupec, ve kterém se nachází naměřené hodnoty (sloupec pojmenovaný *cestujici*).

Dále je vhodné navolit parametry modelů a rozsahu předpovědi. Rozsah předpovědi nastavíme na 4 roky, i když při rozsahu dat by bylo možné nastavit delší časové období. Dalším důležitým parametrem je kritérium, podle kterého budeme vybírat vhodnost modelu. Zde jsem zvolil *Mean Absolute Percent Error* (MAPE).



Obrázek 3: nastavení parametrů modelů

Po navolení parametrů provedeme diagnostiku hodnot pomocí *Diagnose Series* a vybereme modely, které chceme aplikovat. Vzhledem k malému počtu modelů (Obrázek 4), které byly doporučeny, jsem vybral všechny.



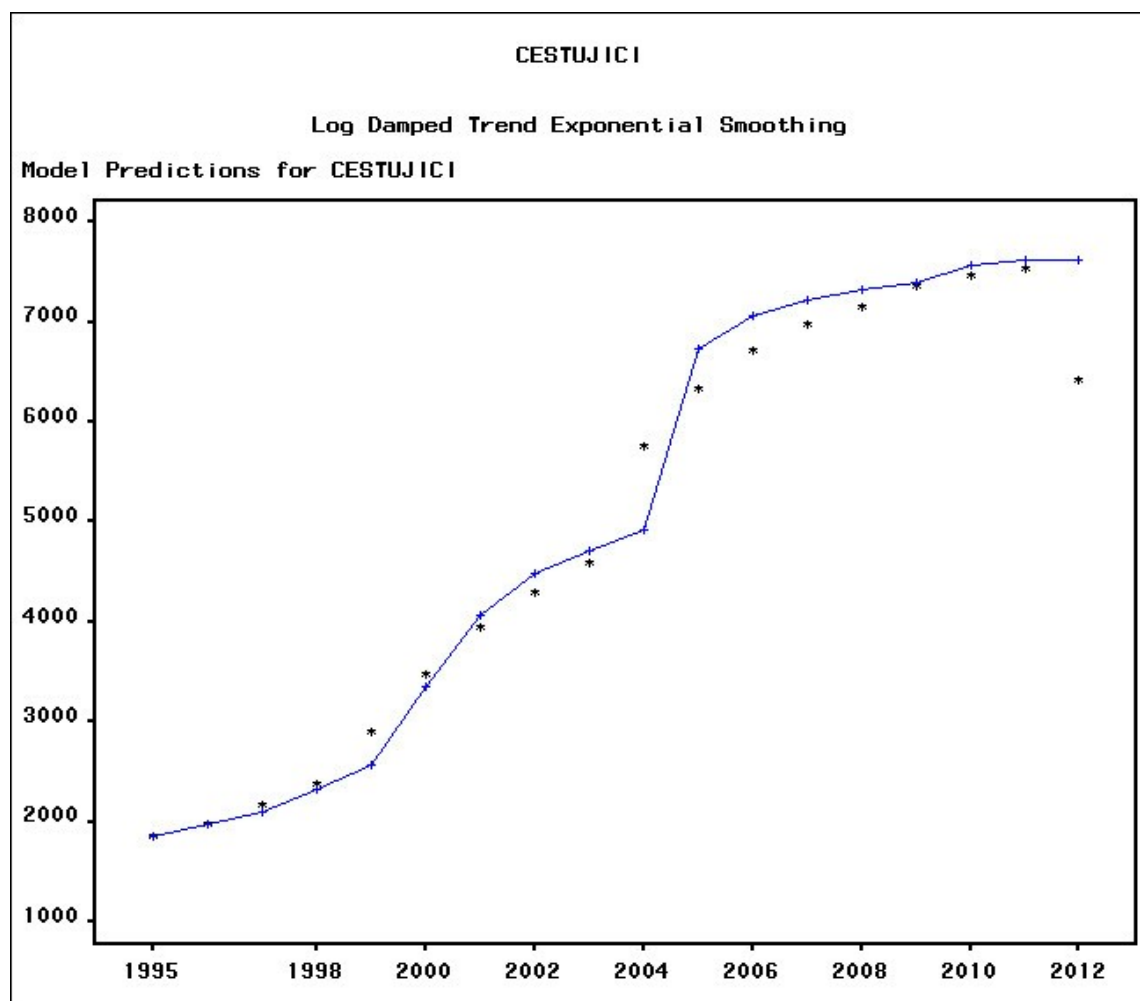
Obrázek 4: nabídnuté modely

Jako nejvhodnější se dle *MAPE* jeví trendová funkce logaritmická (*Log Damped Trend Exponential Smoothing*). Vypovídá o tom i hodnota *R-Square*, která je 0.964. Funkce, které mají *MAPE* větší než 5% můžeme vyřadit, protože jsou nevyhovující.

Develop Models - Project SASUSER.FMSPROJ.PROJ1		
Data Set: SASUSER.ZSTYBLIK		Interval: 1
Series: CESTUJICI		Browse...
Data Range: 1995 to 2012		✓ ✗
Fit Range: 1995 to 2012		
Evaluation Range: 1995 to 2012		Set Ranges...
Forecast Model	Model Title	Mean Absolute Percent Error
☐	Log Linear Trend	13.34482
☐	Log Double (Brown) Exponential Smoothing	5.07330
☐	Log Linear (Holt) Exponential Smoothing	5.19619
☒	Log Damped Trend Exponential Smoothing	4.68330
☐	Log Random Walk with Drift	6.11818

Obrázek 5: modely a jejich *MAPE* hodnoty

Z grafu je vidět, že proložení naměřených hodnot logaritmickou funkcí téměř odpovídá naměřeným hodnotám. Jedinou výjimku zde tvoří rok 2012 kde došlo k velkému poklesu přepravených cestujících.



Obrázek 6: graf logaritmické funkce

Statistics of Fit	
CESTUJICI	
Log Damped Trend Exponential Smoothing	
Statistic of Fit	Value
Mean Square Error	151441.5
Root Mean Square Error	389.15485
Mean Absolute Percent Error	4.68330
Mean Absolute Error	247.07405
R-Square	0.964

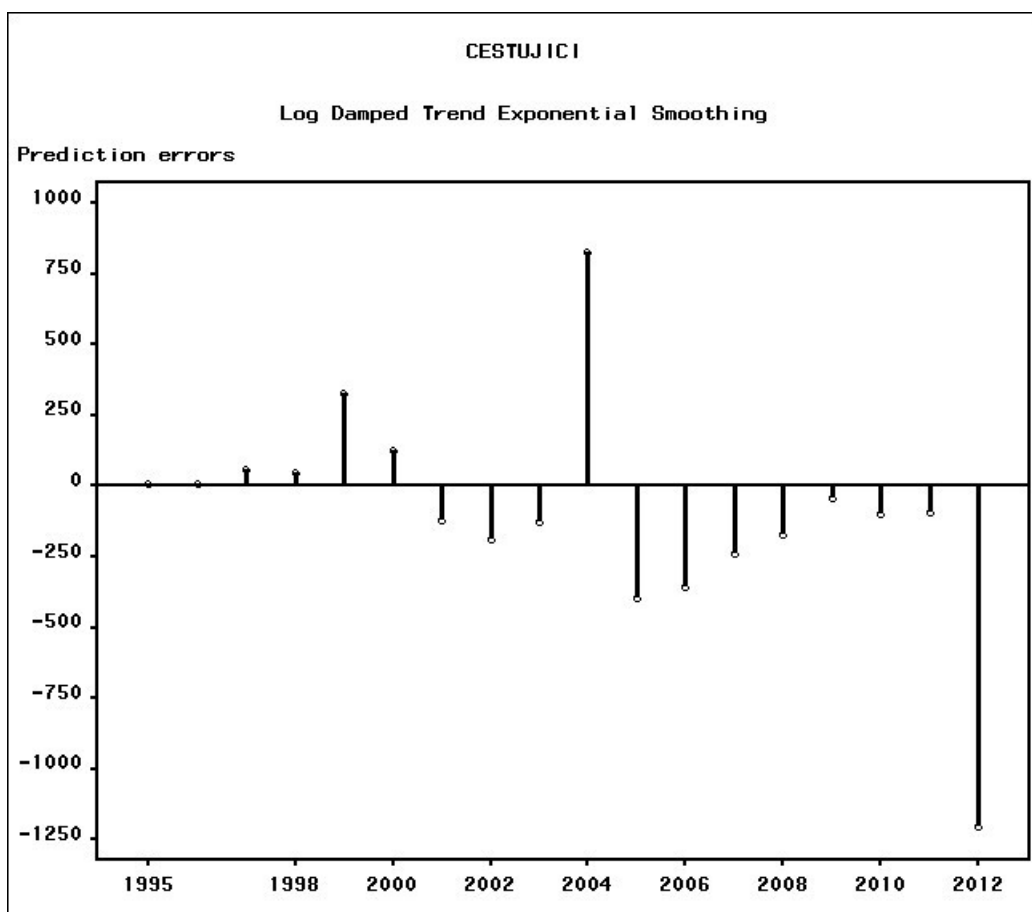
Evaluation Range: 1995 to 2012

Obrázek 7: charakteristika kvality logaritmického trendu

Parameter Estimates				
CESTUJICI				
Log Damped Trend Exponential Smoothing				
Model Parameter	Estimate	Std. Error	T	P
LEVEL Smoothing Weight	0.99900	0.2540	3.9327	*
TREND Smoothing Weight	0.63143	0.4555	1.3863	*
DAMPING Smoothing Weight	0.91543	0.0911	10.0463	*
Residual Variance (sigma squared)	0.00548	.	.	.
Smoothed Level	8.76728	.	.	.
Smoothed Trend	-0.09635	.	.	.

Fit Range: 1995 to 2012

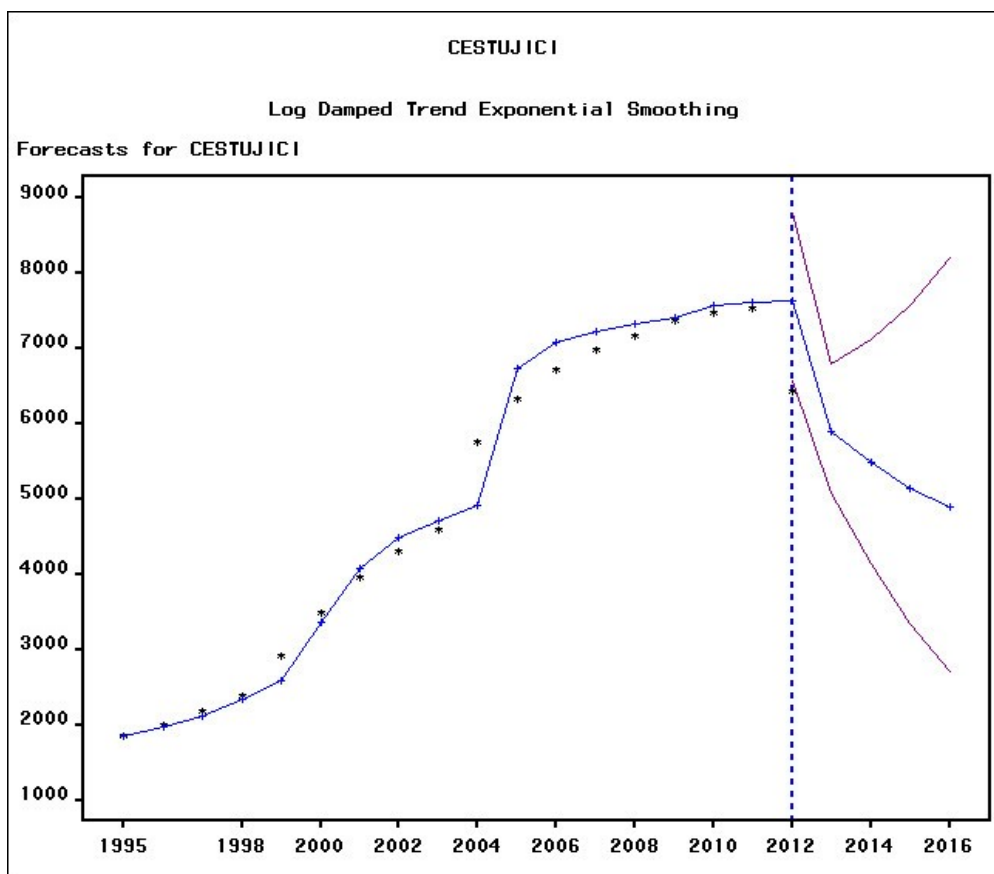
Obrázek 8: odhad parametrů logaritmického trendu



Obrázek 9: odchylky skutečných a predikovaných hodnot

Závěr

Na základě výstupu logaritmické funkce můžeme konstatovat, že v budoucích letech bude počet přepravených pasažérů nadále klesat. Avšak pro přesnější odhad by bylo potřeba modelovat nikoliv data souhrnná, ale minimálně data za kvartály jednotlivých let. I tak by predikce nebyla naprosto přesná, protože zde působí mnoho dalších faktorů, např. cena paliva, letištní poplatky, důchody občanů, nabídky cestovních kanceláří.



Obrázek 10: predikce na 2013-2016

Výrazný pokles v roce 2012 lze přisoudit prohlubující se ekonomické krizi a opatrnosti občanů v utrácení. Nicméně si nemyslím, že by byl predikovaný pokles realný. K poklesu oproti předchozím obdobím jistě dojde, ale ustálí se na nové hodnotě a pokles nebude tak strmý. Ostatně předběžná data pro rok 2013 tuto úvahu podporují.

Rok	Skutečná hodnota	Odhadnutá hodnota	Odchylka
1995	1857	1851.48	5.52
1996	1982	1976.68	5.32
1997	2166	2105.7	60.3
1998	2377	2331.23	45.77
1999	2904	2575.86	328.14
2000	3483.47	3355.47	128
2001	3947.33	4070.71	-123.38
2002	4291.37	4480.61	-189.24
2003	4583.77	4709.21	-125.44
2004	5750.34	4922.37	827.97
2005	6329.86	6725.54	-395.68
2006	6709.62	7068.78	-359.16
2007	6977.02	7216.1	-239.08
2008	7157.92	7327.14	-169.22
2009	7354.22	7399.09	-44.87
2010	7466.08	7567.75	-101.67
2011	7524.62	7618.53	-93.91
2012	6419.58	7624.27	-1204.69
2013		5894.73	
2014		5474.64	
2015		5146.71	
2016		4899.18	
2017		4722.45	
2018		4609.03	
2019		4553.42	
2020		4551.89	

Tabulka 2: skutečné a odhadnuté hodnoty

Za povšimnutí také stojí relativně veliký nárůst cestujících v letech 2004-2005. Příčinou byl s největší pravděpodobností vstup ČR do Evropské unie. Tím se otevřel trh pro další společnosti a byl umožněn volný pohyb osob po zemích EU.