



SEZÓNNE OČISŤOVANIE EKONOMICKÝCH ČASOVÝCH RADOV

SEASONAL ADJUSTMENT OF ECONOMIC TIME SERIES

Eva OSTERTAGOVÁ

Abstract

The paper deals with analysis of seasonal influences in time series. This analysis includes a construction of seasonally adjusted series. The construction of seasonally adjusted series helps to detect what long-term tendency of development has the analyzed index. Seasonal component is described with the model of proportional seasonality with linear trend. Seasonal adjustment is done by the help of seasonal factor.

Key words

Decomposition, seasonal component, model of proportional seasonality, seasonal factor, prediction.

Úvod

Ak analyzujeme vývoj časových radov, zaujíma nás nielen hlavná vývojová tendencia ich ukazovateľov, ale aj priebeh a intenzita prípadných periodických výkyvov, ktoré sa v daných časových radoch prejavujú. Periodické výkyvy (sezónne alebo cyklické), s ktorými sa pri mnohých hospodárskych javoch stretávame, sú v podstate nežiaduce. Preto sa snažíme príčiny týchto výkyvov (pokiaľ je to možné) odstraňovať alebo ich aspoň zmiernovať.

Krátkodobé výkyvy, ktoré sa prejavujú napr. pri výkonoch mestskej dopravy v jednotlivých hodinách dňa sú spôsobené okrem iného nástupom do práce a odchodom z práce. Tieto výkyvy je možné zmierniť organizačným zásahom do vymedzenia pracovnej doby jednotlivých inštitúcií. Sú však aj prípady, keď nie je možné odstrániť príčiny periodických výkyvov a ani zmierniť ich pôsobenie. To platí pre väčšinu sezónnych výkyvov napr. v poľnohospodárstve, v odvetviach priemyslu spracovávajúcich poľnohospodárske produkty, v stavebníctve, v cestovnom ruchu a pod., kde je sezónnosť daná predovšetkým charakterom týchto hospodárskych odvetví.

Základné koncepcie modelovania časových radov

Najjednoduchšou koncepciou modelovania časových radov je jednorozmerný model, ktorý je založený na predpoklade, že jediným faktorom dynamiky ukazovateľa v časovom rade je čas. Je ho možné zapísať v tvare $y = f(t) + \varepsilon$, kde y je hodnota ukazovateľa, t je čas, ε je náhodná (reziduálna) zložka (používa sa aj názov náhodná chyba) a funkcia f je regresná funkcia.

O náhodných chybách ε_i sa predpokladá, že sú to náhodné veličiny, ktoré majú normálne rozdelenie s nulovou strednou hodnotou a so smerodajnou odchýlkou $\sigma > 0$. Teda platí $\varepsilon_i \sim \text{norm}(0, \sigma)$, pre každé $i = 1, 2, \dots, n$. Podmienka, že stredná hodnota $E(\varepsilon_i) = 0$ pre $\forall i = 1, 2, \dots, n$ znamená, že náhodná zložka nepôsobí systematickým spôsobom na hodnoty vysvetľovanej premennej Y . Rozptyl $D(\varepsilon_i) = \sigma^2$ pre každé $i = 1, 2, \dots, n$; t.j.



rozptyl náhodnej zložky je konštantný. Táto podmienka vyjadruje, že variabilita náhodnej zložky nezávisí na hodnotách vysvetľujúcej premennej, t.j. na čase t .

Ďalej sa predpokladá, že $K(\varepsilon_i, \varepsilon_j) = 0$ pre $\forall i \neq j, i, j = 1, 2, \dots, n$; t.j. kovariancia náhodnej zložky je nulová. Teda hodnoty náhodnej zložky sú nekorelované a z toho vyplýva aj nekorelovanosť rôznych dvojíc pozorovaní vysvetľovanej premennej Y .

Nech $\hat{y}_i$ je teoretická (vyrovnaná) hodnota premennej Y . Rozdiel empirickej (pozorovanej) a teoretickej hodnoty $e_i = y_i - \hat{y}_i$ sa nazýva rezíduum. Rezíduum predstavuje bodový odhad náhodnej chyby ε_i . Rozptyl $D(\varepsilon_i) = \sigma^2, i = 1, 2, \dots, n$; obvykle nepoznáme a musí byť tiež odhadnutý z dát, ktoré sú k dispozícii. Pomocou reziduálneho súčtu štvorcov $SS_R = \sum_{i=1}^n e_i^2$ je možné stanoviť výberový reziduálny rozptyl $s_R^2 = \frac{SS_R}{n-k}$ (k je počet odhadovaných parametrov regresnej funkcie), ktorý predstavuje bodový odhad rozptylu σ^2 . Bodovým odhadom smerodajnej odchýlky σ je výberová reziduálna smerodajná odchýlka $s_R = \sqrt{s_R^2}$. Reziduálny rozptyl vyjadruje, ako veľmi sa v priemere teoretická hodnota $\hat{y}$ veličiny Y líši od empirickej hodnoty y . Udáva teda, ako sú hodnoty veličiny Y rozptýlené okolo regresnej funkcie.

Medzi základné metódy a postupy pre analýzu časových radov patrí dekompozičná metóda. Klasická dekompozičná metóda vychádza z predpokladu, že každý časový rad môže obsahovať tieto štyri zložky: trendovú zložku T , sezónnu zložku S , cyklickú zložku C a náhodnú zložku ε .

Trendová zložka (trend) vyjadruje dlhodobú tendenciu vo vývoji hodnôt analyzovaného ukazovateľa v čase. Trend môže byť rastúci, klesajúci alebo stagnujúci.

Sezónna zložka vyjadruje pravidelné kolísanie hodnôt okolo trendu s periódou, ktorá je kratšia alebo rovná jednému roku (napr. týždeň, mesiac, rok). Sezónne zmeny sú zapríčinené takými faktormi, akými sú napr. striedanie ročných období, vianočné a veľkonočné sviatky, prázdniny, dovolenky a pod.

Cyklická zložka vyjadruje kolísanie hodnôt okolo trendu spôsobené cyklickými vplyvmi (cyklus ekonomických kríz, výrobný cyklus, atď.) s periódou, ktorá je dlhšia ako jeden rok. Používa sa skôr v makroekonomických úvahách. Pre cykly je charakteristické, že nie je isté, kedy nastanú, aká bude ich intenzita a dĺžka ich trvania.

Náhodná (reziduálna) zložka je zložka, ktorá ostane v časovom rade po odstránení trendu, sezónnej aj cyklickej zložky. Vyjadruje náhodné výkyvy, ktoré nemajú systematický charakter. Toto náhodné kolísanie je vyvolané pôsobením faktorov náhodného charakteru, ktoré nie je možné predvídať (napr. zemetrasenie, štrajk, veľké suchá, rôzne epidémie a pod. alebo rôzne neidentifikovateľné vplyvy).

Súhrn sezónnej a cyklickej zložky predstavuje periodickú zložku časového radu $P = S + C$.

Klasický matematický model časového radu môže byť aditívny a multiplikatívny. Najčastejšie sa pri analýze časového radu predpokladá aditívny model opisu chovania časového radu. U tohto modelu je hodnota ukazovateľa v i -tom období daná súčtom:

$$y_i = T_i + P_i + \varepsilon_i, i = 1, 2, \dots, n.$$



Ak má časový rad tento tvar, tak hovoríme o periodickom časovom rade. V mnohých praktických úlohách je periodická zložka P_i reprezentovaná len sezónnou zložkou S_i , t.j. model môžeme zapísať v tvare $y_i = T_i + S_i + \varepsilon_i$ a príslušný časový rad nazývame sezónne zaťaženým časovým radom. Ak platí $P_i = 0$, hovoríme o neperiodickom časovom rade.

Analýza sezónnej zložky časového radu

Ďalej budeme analyzovať len sezónnu zložku časového radu. V prvom rade je potrebné zistiť či skúmaný časový rad reálne vykazuje sezónne výkyvy. Na identifikáciu sezónnej (ale aj cyklickej) zložky v časovom rade je možné použiť rôzne techniky a metódy, z ktorých najjednoduchšou je analýza grafu. Medzi ďalšie metódy patrí napr. analýza periodogramu, spektrálna analýza, atď.

Ďalší postup spočíva v riešení dvoch úloh:

- kvantifikácia sezónnych výkyvov,
- očistenie časového radu, t.j. vylúčenie sezónnej zložky.

Cieľom sezónneho očisťovania je odkrytie základnej dynamiky vývoja skúmaných javov a umožnenie bezprostredného porovnania vývoja v jednotlivých sezónach v rámci roka.

Pri vytváraní matematického modelu tu budeme uvažovať, že odhadujeme hodnotu ukazovateľa v priebehu rokov a v rámci roka ešte v čiastkových obdobiach (sezónach), ktoré sa každý rok opakujú, napr. mesiace, štvrťroky. Budeme teda používať dva indexy. Index $i = 1, 2, \dots, k$ bude označovať rok a index $j = 1, 2, \dots, m$ bude označovať sezónu.

Hodnotu ukazovateľa v i -tej perióde (rok) a v j -tom čiastkovom období vnútri periódy (mesiac, štvrťrok) môžeme v prípade aditívneho modelu vyjadriť v tvare $y_{ij} = T_{ij} + S_{ij} + \varepsilon_{ij}$, kde T_{ij} , S_{ij} , ε_{ij} sú postupne odhady trendu, sezónneho výkyvu a náhodnej zložky. Dĺžka časového radu je $n = m \cdot k$.

Matematické modely uvažujú dva druhy sezónnych vplyvov, a to konštantnú a proporcionálnu sezónnosť.

Model konštantnej sezónnosti

Pri tomto modeli je amplitúda (veľkosť výkyvu) sezónnej zložky stála, bez ohľadu na smer trendu, t.j. predpokladáme približne konštantnú veľkosť sezónnych výkyvov v rovnakej sezóne v rôznych rokoch.

Modelová rovnica prejde na tvar $y_{ij} = T_{ij} + b_j + \varepsilon_{ij}$, t.j. $S_{ij} = b_j$, pre $\forall i = 1, 2, \dots, k$; $j = 1, 2, \dots, m$; kde b_j sú neznáme sezónne parametre.

Teda na základe modelovej rovnice je zrejmé, že sa tu sezónnosť meria pomocou rozmernej sezónnej konštanty (vyjadrenej v rovnakých jednotkách, v akých sú vyjadrené hodnoty časového radu), ktorá sa k trendovej zložke pripočíta.

Ďalej predpokladáme, že pre sezónne parametre platí $\sum_{j=1}^m b_j = 0$ pre všetky roky $i = 1, 2, \dots, k$; t.j. pri regresii trendu uvažujeme, že sa kladné a záporné výkyvy od trendu v priebehu roka vzájomne vyrovnávajú.



Model proporcionálnej sezónnosti

Pri tomto modeli amplitúda sezónnej zložky korešponduje so smerom trendu. Teda pokiaľ v priebehu rokov napr. trend rastie, pre danú sezónu sa amplitúda s trendom zvyšuje. Hodnotu sezónneho výkyvu v i -tom roku a v j -tej sezóne môžeme teda vyjadriť v tvare $S_{ij} = c_j \cdot T_{ij}$, pre $\forall i = 1, 2, \dots, k; j = 1, 2, \dots, m$; kde c_j je konštanta úmernosti (sezónny parameter) nemenná pre danú sezónu v priebehu všetkých rokov.

Modelová rovnica má potom tvar $y_{ij} = T_{ij} + c_j \cdot T_{ij} + \varepsilon_{ij} = (1 + c_j) \cdot T_{ij} + \varepsilon_{ij}$. Celková konštanta úmernosti pre dané obdobie, t.j. výraz v zátvorke $(1 + c_j)$, sa nazýva sezónny faktor, ktorý označíme I_j . Teda na základe modelovej rovnice je zrejmé, že sa tu sezónnosť meria pomocou bezrozsmernej sezónnej konštanty, ktorou sa trendová zložka vynásobí.

Sezónny index I_{ij} pre sledovanú j -tú sezónu a uvažovaný i -tý rok je možné odhadnúť z nameraných hodnôt ukazovateľa y_{ij} a zo zodpovedajúcich vyrovnaných hodnôt trendu $\hat{T}_{ij}$ podľa vzorca $I_{ij} = \frac{y_{ij}}{\hat{T}_{ij}}$. Trend sa vyrovnáva niektorou zo známych metód, napr. lineárnou regresiou alebo kľzavými priermi či exponenciálnym vyrovnávaním.

Zo sezónnych indexov I_{ij} potom stanovíme pre určitú j -tú sezónu vo všetkých sledovaných rokoch (spolu k) spoločný odhad sezónneho faktora ako prostý aritmetický priemer $I_j = \frac{1}{k} \cdot \sum_{i=1}^k I_{ij}$, pre $j = 1, 2, \dots, m$. Približne platí $\sum_{j=1}^m I_j = m$.

Ďalej vypočítame vyrovnané hodnoty (systematickú zložku) časového radu ako súčin $\hat{y}_{ij} = \hat{T}_{ij} \cdot I_j$, prípadne môžeme stanoviť predpoveď na nasledujúci rok (roky). Rozdiel medzi skutočne nameranou hodnotou a takto vypočítanou hodnotou je odhad náhodnej zložky, t.j. rezíduum.

Získané sezónne faktory môžeme použiť aj na opačné výpočty. Vydelením jednotlivých nameraných hodnôt sezónnym faktorom ich zbavíme vplyvu sezónnosti a získame tak trendové zložky len s náhodnými výkyvmi, tzv. sezónne očistené hodnoty $\frac{y_{ij}}{I_j}$.

Aplikácia modelu proporcionálnej sezónnosti v ekonomickej praxi

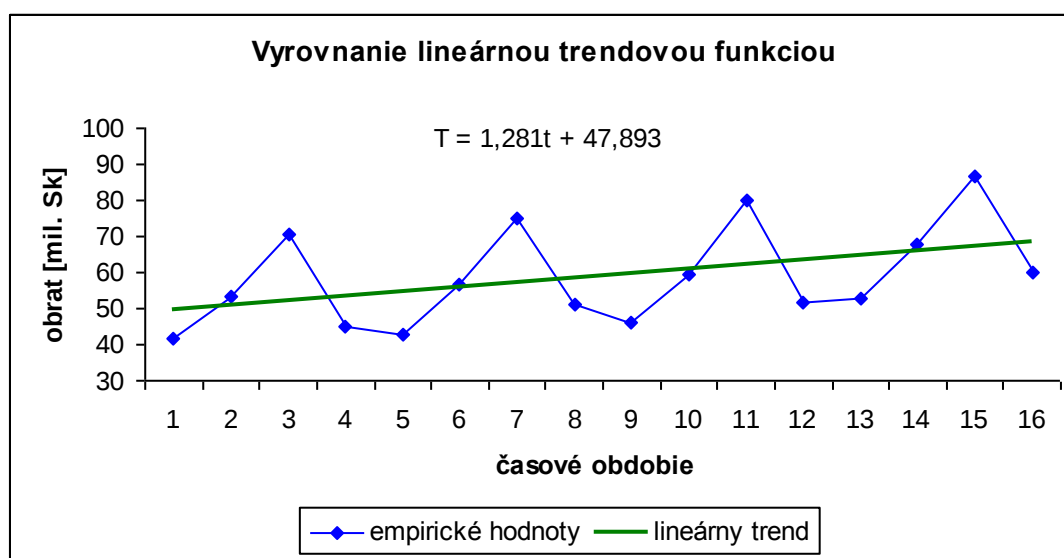
V tab.1 sú údaje o vývoji tržieb (y_{ij}) v mil. Sk vybraného podniku za roky 2005 – 2008. Budeme analyzovať vývoj tržieb v minulých obdobiach a urobíme prognózu na rok 2009 podľa štvrtí rokov.



Tab. 1 Údaje o vývoji tržieb podniku

Rok	2005				2006				2007			
Štvrťrok	I.	II.	III.	IV.	I.	II.	III.	IV.	I.	II.	III.	IV.
y_{ij}	41,5	53,6	70,7	45,2	42,6	56,7	75,1	51,1	45,9	59,4	79,8	51,6
Rok	2008											
Štvrťrok	I.	II.	III.	IV.								
y_{ij}	52,7	67,7	86,9	60,0								

Najskôr budeme analyzovať trend vyrovnaním časového radu vhodnou trendovou funkciou. Použitím programu EXCEL dostaneme pre daný časový rad spojnicový graf uvedený na obrázku 1. Z grafického zobrazenia možno usúdiť, že postačí vyrovnanie priamkou.



Obr. 1 Vyrovnanie daného časového radu priamkou

Z grafickej analýzy daného časového radu je zrejmé, že vykazuje sezónnosť. Použijeme model proporcionálnej sezónnosti. Všetky ďalšie výpočty sú usporiadané v tabuľke 2.

Tab. 2 Tabuľka výsledkov pre model proporcionálnej sezónnosti

Číslo intervalu	Časové obdobie	Indexy $i; j$	Nameraná hodnota y_{ij}	Trend (lineárna regresia) $\hat{T}_{ij}$	Sezónny index $I_{ij} = \frac{y_{ij}}{\hat{T}_{ij}}$	Vyrovnaná hodnota $\hat{y}_{ij} = \hat{T}_{ij} \cdot I_{ij}$
1	I. štvrťrok 2005	1;1	41,5	49,17	0,84	39,34
2	II. štvrťrok 2005	1;2	53,6	50,45	1,06	51,46
3	III. štvrťrok 2005	1;3	70,7	51,74	1,37	68,30
4	IV. štvrťrok 2005	1;4	45,2	53,02	0,85	45,60
5	I. štvrťrok 2006	2;1	42,6	54,30	0,78	43,44
6	II. štvrťrok 2006	2;2	56,7	55,58	1,02	56,69
7	III. štvrťrok 2006	2;3	75,1	56,86	1,32	75,06



8	IV. štvrťrok 2006	2;4	51,1	58,14	0,88	50,00
9	I. štvrťrok 2007	3;1	45,9	59,42	0,77	47,54
10	II. štvrťrok 2007	3;2	59,4	60,70	0,98	61,91
11	III. štvrťrok 2007	3;3	79,8	61,98	1,29	81,81
12	IV. štvrťrok 2007	3;4	51,6	63,26	0,82	54,40
13	I. štvrťrok 2008	4;1	52,7	64,55	0,82	51,64
14	II. štvrťrok 2008	4;2	67,7	65,83	1,03	67,15
15	III. štvrťrok 2008	4;3	86,9	67,11	1,29	88,59
16	IV. štvrťrok 2008	4;4	60,0	68,39	0,88	58,82

Jednotlivé vyrovnané hodnoty $\hat{T}_{ij}$ pomocou lineárnej regresie dostaneme dosadením do $\hat{T} = 1,281 \cdot t + 47,893$ postupne za $t = 1, 2, \dots, 16$. Odhady sezónnych indexov I_{ij} vypočítame podľa vzorca $I_{ij} = \frac{y_{ij}}{\hat{T}_{ij}}$. Napr. pre sezónny index pre prvý rok a prvý štvrťrok

$$\text{dostaneme } I_{1;1} = \frac{y_{1;1}}{\hat{T}_{1;1}} = \frac{41,5}{49,17} = 0,84.$$

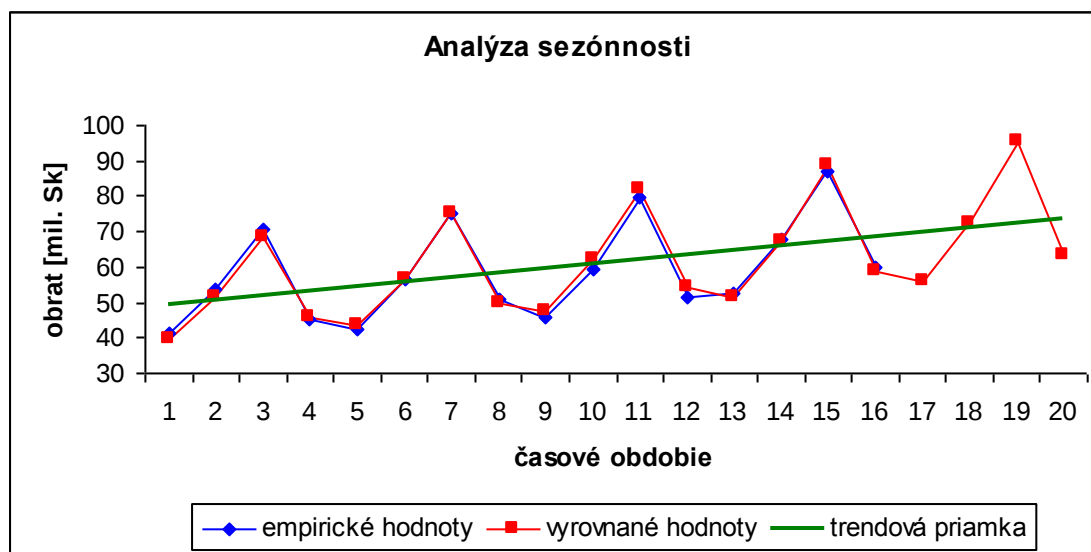
Potom vypočítame odhady sezónnych faktorov ako prosté aritmetické priemery sezónnych indexov. Napr. sezónny faktor I_1 pre 1. štvrťrok je prostý aritmetický priemer sezónnych indexov pre 1. štvrťrok za 4 roky. Teda je $I_1 = \frac{0,84 + 0,78 + 0,77 + 0,82}{4} = 0,80$. Analogicky vypočítame sezónne faktory pre ďalšie štvrťroky. Dostaneme $I_2 = 1,02$; $I_3 = 1,32$; $I_4 = 0,86$.

Teraz môžeme vyjadriť priebeh deterministických zložiek časového radu, a to jednak priebeh trendu podľa rovnice lineárnej regresie a jednak po vynásobení trendu sezónnym faktorom dostaneme „predpovedanú“ hodnotu so sezónnym kolísaním $\hat{y}_{ij} = \hat{T}_{ij} \cdot I_j$. Napr. pre 1. rok a 1. štvrťrok bude $\hat{y}_{1;1} = \hat{T}_{1;1} \cdot I_1 = 49,17 \cdot 0,80 = 39,34$ (viď Tab. 2). Skutočná hodnota bola 41,5. Teda rezíduum je $41,5 - 39,34 = 2,16$. Predpoveď na rok 2009 je v tabuľke 3.

Tab. 3 Tabuľka s prognózou na rok 2009

Číslo intervalu	Časové obdobie	Trend (lineárna regresia) $\hat{T}_{ij}$	Predpovedaná hodnota $\hat{y}_{ij} = \hat{T}_{ij} \cdot I_j$
17	I. štvrťrok 2009	69,67	55,74
18	II. štvrťrok 2009	70,95	72,37
19	III. štvrťrok 2009	72,23	95,34
20	IV. štvrťrok 2009	73,51	63,22

Na obr. 2 sú graficky znázornené už aj vyrovnané hodnoty sledovaného časového radu spolu s prognózovanými hodnotami na 1. – 4. štvrťrok roku 2009.



Obr. 2 Grafická analýza sezónnosti

Súhrn

Sezónne očisťovanie ekonomických časových radov predstavuje jednu z prakticky veľmi užitočných procedúr, ktorá so vzrastajúcim záujmom o meranie dynamiky ekonomických javov nadobúda stále väčší význam. Sezónne fluktuácie v ekonomických časových radoch predstavujú problém pre vytváranie modelu a neberúc ich do úvahy riskujeme, že neodstránené sezónne výkyvy spôsobia neadekvátne odhady.

Dôvodom pre sezónne očisťovanie je možnosť konštrukcie predpovedí krátkodobých časových radov, použitie očistených časových radov pri konštrukcii ekonometrických modelov, ako aj dosiahnutie porovnateľnosti hodnôt za jednotlivé štvrťroky, prípadne mesiace.

Kľúčové slová

Dekompozícia, sezónna zložka, model proporcionálnej sezónnosti, sezónny faktor, predikcia.

Príspevok bol pripravený v rámci riešenia grantového projektu VEGA č. 1/0102/11 Metódy a techniky experimentálneho modelovania vnútropodnikových výrobných a nevýrobných procesov.

Použitá literatúra

- [1] ARLT, J., ARLTOVÁ, M., RUBLÍKOVÁ, E.: *Analýza ekonomických časových řad s příklady*. VŠE Praha, 2004,
- [2] BELL, W.R., HILLMER, S.C.: *Issues Involved with the Seasonal Adjustment of Economic Time Series*. Journal of Bussines and Economic Statistics. 1984, p. 291-320.
- [3] CYHELSKÝ, L., KAŇOKOVÁ, J., NOVÁK, I.: *Základy teorie statistiky pro ekonomy*. SNTL/Alfa, Praha, 1979,
- [4] HANČLOVÁ, J., TVRDÝ, L.: *Úvod do analýzy časových řad*. VŠB-TU, Ostrava, 2003.
- [5] CHAJDIK, J., RUBLÍKOVÁ, E., GUBÁDA, M.: *Štatistické metódy v praxi*. Statis, Bratislava, 1994,



- [6] KŘIVÝ, I.: *Analýza časových řad*. Ostravská univerzita, 2006.
- [7] OSTERTAGOVÁ, E.: *Aplikovaná statistika*. Elfa, Košice, 2011, ISBN 978-80-8086-171-1,
- [8] OSTERTAGOVÁ, E.: *Time series modeling*. Modelling of Mechanical and Mechatronical Systems: proceedings of the 4th international conference. Herľany, 2011, p. 366-373,
- [9] POPELKA, J., SYNEK, V.: *Úvod do statistické analýzy dat*. UJEP, Ústí nad Labem, 2009,
- [10] SVATOŠOVÁ, L., KÁBA, B.: *Statistické metody 2*. PEF ČZU, Praha, 2008.

Kontaktná adresa

PhDr. Eva Ostertagová, PhD.,
Technická univerzita v Košiciach,
Fakulta elektrotechniky a informatiky,
Katedra matematiky a teoretickej informatiky,
Němcovej 32, 040 01 Košice,
e-mail: eva.ostertagova@tuke.sk.