

URČENÍ FAKTORŮ OVLIVŇUJÍCÍCH PŘESNOST PROGNOSTICKÝCH MODELŮ

IDENTIFICATION OF FACTORS AFFECTING THE ACCURACY OF FORECASTING MODELS

Bohumil Kába

Anotace:

Příspěvek pojednává o rozhodujících faktorech, které ovlivňují přesnost prognostických modelů ekonomických ukazatelů. Určení a kvantifikace těchto faktorů jsou velmi důležité pro výběr adekvátní prognostické metody. Výklad je doplněn výsledky empirické srovnávací analýzy.

Klíčová slova:

prognostický model, časová řada, předpověď, přesnost předpovědi

Abstract:

The paper deals with the crucial factors influencing the predictive performance of various forecasting models of economic indicators. Determining and measuring those factors can be of great help in finding and selecting an appropriate forecasting method. The results of a comparative empirical analysis are given.

Keywords:

forecasting model, time series, forecast, accuracy of forecasts

ÚVOD

Kvalita extrapolacních předpovědí z modelů jednorozměrných časových řad je velmi závislá na vlastnostech analyzovaných řad a na vlastnostech zvolených modelů. Vzhledem k tomu, že každá časová řada má své charakteristické a jedinečné vlastnosti, je velmi obtížné předem usuzovat na přesnost výsledků prognostického procesu. Obzvláště náročné je to v případě časových řad ukazatelů agrárního sektoru, pro něž je obvykle typický složitý a nepravidelný průběh.

Daný příspěvek prezentuje dílčí výsledky statistické studie, která byla věnována srovnávání vybraných statistických postupů při modelování a predikci vývoje ročních časových řad ukazatelů agrárního sektoru a potravinové vertikály České republiky. V rámci této analýzy byla též věnována pozornost variabilitě výsledků, získaných různými prognostickými postupy a jejím možným zdrojům.

CÍL A METODIKA

Daná studie se orientovala na prozkoumání možností identifikace a kvantifikace faktorů, které ovlivňují přesnost předpovědí z modelů časových řad. Analýza se týkala 160 ročních časových řad ukazatelů agrárního sektoru a potravinové vertikály České republiky a zahrnovala různé úseky období 1948 – 2002. Empirické testování různých tříd statistických modelů naznačilo, že modelování a tvorbu předpovědí (uvažován byl tříletý horizont předpovědi) těchto řad je účelné zakládat na některých adaptivních modelech a modelu náhodné procházky s posunem. Z adaptivních přístupů se nejvíce osvědčily Brownův model

dvojitého exponenciálního vyrovnávání, model exponenciálního vyrovnávání s tlumeným trendem a Holtův model exponenciálního vyrovnávání.

Zkoumání potenciálních faktorů ovlivňujících přesnost prognostických postupů se tedy zaměřilo na výsledky zmíněných čtyř modelů. Metodologicky vycházelo z technik regresní analýzy. Vysvětlovanou proměnnou byly předpovědní chyby jednotlivých modelů, charakterizované střední absolutní procentuální chybou MAPE [2]. Za vysvětlující proměnné byly zvoleny některé základní deskriptivní charakteristiky disponibilních řad a dále interpolační chyby (charakterizované opět metrikou MAPE, počítanou pro údaje z období interpolace).

VÝSLEDKY

Střední absolutní procentuální chyby předpovědí, získaných aplikací všech čtyř zmíněných prognostických postupů, byly popsány pomocí dvou regresních modelů. V prvním z nich – model I – byly jako vysvětlující proměnné uvažovány: délka analyzované řady N , variační koeficient hodnot řady CV , koeficient šikmosti S a koeficient špičatosti K hodnot řady. Odhadnuté regresní koeficienty tohoto vícenásobného regresního modelu a příslušný koeficient mnohonásobné determinace R^2 jsou uvedeny v tabulce 1. V závorce pod odhady koeficientů jsou uvedeny tzv. p – hodnoty (vypočtené hladiny významnosti), jež umožňují posoudit statistickou významnost těchto koeficientů.

Tabulka 1. Regresní koeficienty – model I

Model	Konstanta	N	CV	S	K	R^2
B2	3,665 ($<0,01$)	-0,043 (0,11)	0,205 ($<0,01$)	1,766 ($<0,01$)	0,921 (0,05)	38,74 % ($<0,01$)
HE	5,312 ($<0,01$)	-0,087 (0,06)	0,321 (0,01)	2,484 (0,03)	1,706 (0,03)	32,44 % ($<0,01$)
EDT	5,901 ($<0,01$)	-0,105 ($<0,01$)	0,288 ($<0,01$)	2,520 (0,01)	1,727 ($<0,01$)	45,22 % ($<0,01$)
RW	3,933 ($<0,01$)	-0,074 ($<0,01$)	0,255 ($<0,01$)	1,958 ($<0,01$)	1,575 ($<0,01$)	50,68 % ($<0,01$)

B2 – Brownův model dvojitého exponenciálního vyrovnávání

HE – Holtův model exponenciálního vyrovnávání

EDT – exponenciální vyrovnávání s tlumeným trendem

RW – model náhodné procházky s posunem

V další fázi analýzy byl konstruován model II. Množina jeho vysvětlujících proměnných vznikla rozšířením odpovídající množiny modelu I o proměnnou $MAPE_i$ (tzn. ukazatel interpolačních chyb). Odhadnuté regresní koeficienty modelu II jsou soustředěny v tabulce 2.

Tabulka 2. Regresní koeficienty – model II

Model	Konstanta	N	$MAPE_i$	CV	S	K	R^2
B2	2,385 (0,03)	-0,022 (0,43)	0,298 (0,02)	0,125 (0,03)	1,319 (0,05)	0,731 (0,11)	45,17 % ($<0,01$)
HE	1,855 (0,26)	-0,032 (0,44)	0,922 ($<0,01$)	0,101 (0,24)	1,169 (0,23)	1,053 (0,13)	51,83 % ($<0,01$)
EDT	3,329 ($<0,01$)	-0,066 (0,02)	0,705 ($<0,01$)	0,122 (0,03)	1,504 (0,02)	1,246 (0,01)	65,07 % ($<0,01$)
RW	3,181 ($<0,01$)	-0,063 (0,01)	0,182 (0,09)	0,209 ($<0,01$)	1,716 ($<0,01$)	1,459 ($<0,01$)	53,93 % ($<0,01$)

DISKUSE

Z výsledků prezentovaných v tabulce 1 lze usuzovat na středně silnou korelaci mezi chybami předpovědí z uvažovaných prognostických modelů a některými základními deskriptivními charakteristikami analyzovaných časových řad. Nízké vypočtené hladiny významnosti pro většinu regresních koeficientů naznačují, že odhadnuté regresní modely lze považovat za vyhovující a mohou být využity pro získání základní představy o kvalitě extrapoláčních předpovědí a vysvětlení variability příslušných předpovědních chyb (zejména to platí pro exponenciální vyrovnávání s tlumeným trendem a model náhodné procházky).

Tabulka 2 ukazuje, že zařazení ukazatele interpolačních chyb do množiny vysvětlujících proměnných může v některých případech poměrně značně zvýšit schopnost regresního modelu vysvětlovat celkovou variabilitu předpovědních chyb. Regresní model pro exponenciální vyrovnávání s tlumeným trendem již signalizuje silnou korelaci mezi extrapoláčními chybami a danou množinou vysvětlujících proměnných a regresní modely pro náhodnou procházku a pro Holtův model rovněž vysvětlují značný podíl variability předpovědních chyb. Je zde tedy možnost využít regresního modelu II pro konstrukci dostatečně spolehlivých regresních odhadů chyb extrapoláčních předpovědí z příslušných prognostických modelů.

Podrobnější analýza všech dosažených výsledků a některé doplňující rozbor, uskutečněné technikami tzv. etapovité regrese (*stepwise regression*), orientované na identifikaci „optimální“ podmnožiny vysvětlujících proměnných ukázaly, že úroveň extrapoláčních chyb je závislá zejména na velikosti interpolačních chyb a variabilitě původních hodnot (charakterizované variačním koeficientem) analyzovaných časových řad. Vliv proměnné „délka analyzované řady“ na přesnost předpovědí se ukázal méně podstatný a v řadě případů byl statisticky nevýznamný.

ZÁVĚRY

Identifikace adekvátních prognostických modelů časových řad ukazatelů agrárního sektoru respektive ukazatelů potravinové vertikály je obvykle náročná a nevede k jednoznačnému výsledku, protože z teoretických vlastností potenciálních modelů nelze vyvozovat kvalifikované závěry o reálných možnostech jejich uplatnění. Daná studie naznačila, že užitím prezentovaných regresních modelů I a II lze v mnoha případech získat dostatečně spolehlivý odhad chyb předpovědí z modelů, které se osvědčily při analýze a extrapolaci ročních časových řad zmiňovaných ukazatelů. Lze tedy konstatovat, že důkladná průzkumová analýza hodnot disponibilních časových řad, případně doplněná informací o chování modelu ve fázi interpolace, může být velmi užitečným nástrojem při výběru prognostické metody.

Literatura:

1. Kába, B.: Interpolační a extrapoláční vlastnosti modelů časových řad zemědělských ukazatelů, In: Sborník z mezinárodní vědecké konference „Firma a konkurenční prostředí 2003“, PEF MZLU, Brno 2003, ISBN 80-7157-695-6, s. 119 – 124
2. SAS/ETS Software: Time Series Forecasting, Version 6, First Edition, Cary, NC, SAS Institute Inc., 1995

Kontaktní adresa:

Doc. RNDr. Bohumil Kába, CSc
katedra statistiky PEF ČZU
165 21 Praha 6 – Suchdol
tel.: 224 382 236
e-mail: kaba@pef.czu.cz