

**M A S A R Y K O V A
U N I V E R Z I T A**

EKONOMICKO-SPRÁVNÍ FAKULTA

Porovnání in-sample a out-of-sample predikcí indexu průmyslové produkce

Diplomová práce

MICHAL STRNAD

Vedoucí práce: doc. Ing. Jan Čapek, Ph.D.

Katedra ekonomie
Program Matematické a statistické metody v ekonomii

Brno 2026

MUNI
ECON

POROVNÁNÍ IN-SAMPLE A OUT-OF-SAMPLE PREDIKCÍ INDEXU PRŮMYSLOVÉ PRODUKCE

MUNI
ECON

MASARYKOVA UNIVERZITA
EKONOMICKO-SPRÁVNÍ FAKULTA
LIPOVÁ 41A, 602 00 BRNO
IČ: 00216224
DIČ: CZ00216224

ZADÁNÍ DIPLOMOVÉ PRÁCE

Akademický rok: 2025/2026

Student:	Bc. Michal Strnad
Program:	Matematické a statistické metody v ekonomii
Název práce:	Porovnání in-sample a out-of-sample predikcí indexu průmyslové produkce
Název práce anglicky:	A comparison of in-sample and out-of-sample forecasts of industrial production index
Cíl práce, postup a použité metody:	Diplomová práce se zabývá porovnáním in-sample a out-of-sample predikcí indexu průmyslové výroby. Cílem práce je porovnat možnosti predikce vývoje indexu průmyslové produkce ve vybraných evropských zemích. K naplnění cíle budou vytvořeny ARIMA a VAR modely, které mezi sebou budou porovnávány na základě RMSE, případně dalších vybraných ukazatelů predikčních chyb.
Rozsah grafických prací:	Podle pokynů vedoucího práce
Rozsah práce bez příloh:	60 – 80 stran
Literatura:	BIANCHI, Carluccio; Alessandro CARTA; Dean FANTAZZINI; Maria Elena DE GIULI a Mario A. MAGGI. A copula-VAR-X approach for industrial production modelling and forecasting. <i>Applied Economics</i> . ENGLAND: ROUTLEDGE JOURNALS, TAYLOR & FRANCIS LTD, 2009. ISSN 0003-6846. Dostupné z: https://doi.org/10.1080/00036840802112349 . BODO, Giorgi; Roberto GOLINELLI a Giuseppe PARIGI. Forecasting industrial production in the Euro area. <i>EMPIRICAL ECONOMICS</i> . AUSTRIA: PHYSICA-VERLAG GMBH & CO, 2000. ISSN 0377-7332. Dostupné z: https://doi.org/10.1007/s001810000032 . FRANSES, Philip Hans a Dick VAN DIJK. The forecasting performance of various models for seasonality and nonlinearity for quarterly industrial production. <i>INTERNATIONAL JOURNAL OF FORECASTING</i> . NETHERLANDS: ELSEVIER, 2005. ISSN 0169-2070. Dostupné z: https://doi.org/10.1016/j.ijforecast.2004.05.005 .
Vedoucí práce:	doc. Ing. Jan Čapek, Ph.D.
Pracoviště vedoucího práce:	Katedra ekonomie

Datum zadání práce: 9. 2. 2024

Termín odevzdání diplomové práce a vložení do IS je uveden v platném harmonogramu akademického roku.

Zadání bylo schváleno prostřednictvím IS MU.

doc. Ing. Ondřej Krčál, Ph.D., 10. 11. 2025

V Brně dne: 4. 5. 2026

Bibliografický záznam

Autor:	Michal Strnad Ekonomicko-správní fakulta Masarykova univerzita Katedra ekonomie
Název práce:	Porovnání in-sample a out-of-sample predikcí indexu průmyslové produkce
Studijní program:	Matematické a statistické metody v ekonomii
Vedoucí práce:	doc. Ing. Jan Čapek, Ph.D.
Rok:	2026
Počet stran:	106
Klíčová slova:	in-sample, out-of-sample, predikce, ARIMA, SARIMA, index průmyslové produkce

Bibliographic record

Author: Michal Strnad
Faculty of Economics and Administration
Masaryk University
Department of Economics

Title of Thesis: A comparison of in-sample and out-of-sample
forecasts of industrial production index

Degree Programme: Mathematical and Statistical Methods in
Economics

Supervisor: doc. Ing. Jan Čapek, Ph.D.

Year: 2026

Number of Pages: 106

Keywords: in-sample, out-of-sample, forecast, ARIMA,
SARIMA, industrial production index

Anotace

Diplomová práce porovnává in-sample a out-of-sample predikce indexu průmyslové produkce České republiky a Slovenska v období 2000–2023. K modelování jsou použity ARIMA modely lišící se přístupem k zachycení sezónnosti a predikováním v rámci posuvného nebo rozšiřujícího se okna. Data jsou rozdělena do pěti období pro srovnání v různých fázích hospodářského cyklu. Výsledky ukazují, že v klidných obdobích je rozdíl mezi in-sample a out-of-sample predikcemi malý a způsoben především změnou parametrů. V krizových obdobích se rozdíl zvětšuje a roste i význam odlišné volby řádů modelu.

Abstract

The thesis compares in-sample and out-of-sample forecasts of the industrial production index for the Czech Republic and Slovakia from 2000 to 2023. ARIMA models are estimated in several variants differing in their approach to seasonality and in the use of expanding or rolling windows. The data are split into five periods to enable comparison across different phases of the business cycle. The results show that in calm periods the difference between in-sample and out-of-sample forecasts is small and driven mainly by changes in parameter estimates. In crisis periods the difference grows and the role of differences in selected model orders also increases.

Čestné prohlášení

Prohlašuji, že jsem diplomovou práci na téma Porovnání in-sample a out-of-sample predikcí indexu průmyslové produkce vypracoval samostatně pod vedením doc. Ing. Jana Čapka, Ph.D. a uvedl v ní všechny použité literární a jiné odborné zdroje v souladu s právními předpisy, vnitřními předpisy Masarykovy univerzity a vnitřními akty řízení Masarykovy univerzity a Ekonomicko-správní fakulty MU.

Při přípravě této práce jsem použil Claude od společnosti Anthropic pro zlepšení stylistické úrovně práce a asistenci při psaní kódu. Všechny výstupy generované AI jsem zkontroloval a přebírám plnou zodpovědnost za veškerý text uvedený v této práci.

V Brně 2. května 2026

.....
Michal Strnad

Poděkování

Na prvním místě bych chtěl poděkovat doc. Ing. Janu Čapkovi, Ph.D. za podnětné rady k psaní diplomové práce a produktivní konzultace. Rovněž bych rád poděkoval své rodině za podporu při studiu a pomoc při jazykové korektuře diplomové práce.

Obsah

Seznam obrázků	15
Seznam tabulek	17
Seznam pojmů a zkratek	19
1 Úvod	21
2 Rešerše literatury	23
2.1 In-sample a out-of-sample predikce	23
2.2 Volba transformace a přístup k sezónnosti.....	23
2.3 Využití jednorozměrných a vícerozměrných modelů pro predikce indexu průmyslové produkce	25
2.4 Alternativní přístupy k predikci indexu průmyslové produkce	28
2.5 Shrnutí rešerše	28
3 Data	30
4 Metodologie	36
4.1 Volba mezi ARIMA a VAR modely	36
4.2 Zachycení sezónnosti	37
4.3 ARIMA modely	37
4.4 Volba řádů modelu	39
4.5 Konstrukce in-sample a out-of-sample	41
4.6 Hodnocení predikcí.....	44
5 Srovnání predikcí při rozšiřujícím se okně	47
5.1 Modely auto.arima.....	47
5.2 Fixní modely	55
5.3 Shrnutí pro rozšiřující se okno	63
6 Srovnání predikcí při posuvném okně	65

6.1	Citlivostní analýza	65
6.2	Srovnání predikcí auto.arima mezi posuvným a rozšiřujícím se oknem	68
6.3	Srovnání predikcí fixního modelu mezi posuvným a rozšiřujícím se oknem	70
6.4	Srovnání predikcí auto.arima a fixního modelu posuvného okna	72
6.5	Dekompozice rozdílu in-sample a out-of-sample predikcí	74
6.6	Shrnutí pro posuvné okno.....	76
7	Srovnání predikcí při manuálním výběru modelu	77
7.1	Predikce finanční krize.....	77
7.2	Predikce roku 2018	82
7.3	Shrnutí pro manuální výběr	85
8	Diskuse	87
9	Závěr	90
	Použité zdroje	92
Příloha A	Alternativní transformace a modely	97
Příloha B	MAE predikcí	98
Příloha C	Analýza in-sample vzorku	101

Seznam obrázků

Obr. 1: Vývoj indexu průmyslové produkce České republiky a Slovenska	31
Obr. 2: Sezónní odchylky indexu průmyslové produkce od ročního průměru	33
Obr. 3: Konstrukce in-sample a out-of-sample predikcí pro rozšiřující se okno	43
Obr. 4: Konstrukce in-sample a out-of-sample predikcí pro posuvné okno	44
Obr. 5: Dekompozice rozdílu IS-OOS – deterministická sezónnost, rozšiřující se okno	53
Obr. 6: Dekompozice rozdílu IS-OOS – stochastická sezónnost, rozšiřující se okno	54
Obr. 7: Rozdíl RMSE mezi fixním a auto.arima modelem – ČR, deterministická sezónnost, rozšiřující se okno	56
Obr. 8: Rozdíl RMSE mezi fixním a auto.arima modelem – SK, deterministická sezónnost, rozšiřující se okno	57
Obr. 9: Rozdíl RMSE mezi fixním a auto.arima modelem – ČR, stochastická sezónnost, rozšiřující se okno	59
Obr. 10: Rozdíl RMSE mezi fixním a auto.arima modelem – SK, stochastická sezónnost, rozšiřující se okno	60
Obr. 11: Citlivostní analýza pro out-of-sample RMSE – auto.arima, deterministická sezónnost	66
Obr. 12: Citlivostní analýza pro out-of-sample RMSE – auto.arima, stochastická sezónnost	67
Obr. 13: Rozdíl RMSE mezi fixním a auto.arima modelem – ČR, deterministická sezónnost, posuvné okno	73
Obr. 14: Rozdíl RMSE mezi fixním a auto.arima modelem – SK, deterministická sezónnost, posuvné okno	74
Obr. 15: Dekompozice rozdílu IS-OOS – deterministická sezónnost, posuvné okno	75
Obr. 16: ACF a PACF IPP po první diferenci a odstranění deterministické sezónní složky, 01/2000–09/2008	79
Obr. 17: ACF a PACF IPP po první i sezónní diferenci 2010–2017	83
Obr. 18: Citlivostní analýza pro IS RMSE – auto.arima, deterministická sezónnost	101
Obr. 19: Citlivostní analýza pro IS RMSE – auto.arima, stochastická sezónnost	102
Obr. 20: ACF a PACF prvních diferencí IPP po odstranění deterministické sezónní složky, IS vzorek 2000–2009	104

SEZNAM OBRÁZKŮ

Obr. 21: ACF a PACF IPP po první i sezónní diferenci, IS vzorek 2010–2018	104
---	-----

Seznam tabulek

Tab. 1: Deskriptivní statistiky indexu průmyslové produkce	31
Tab. 2: HEGY test jednotkových kořenů	34
Tab. 3: Vymezení analyzovaných období	46
Tab. 4: RMSE po obdobích – auto.arima, deterministická sezónnost, rozšiřující se okno.....	49
Tab. 5: RMSE po obdobích – auto.arima, stochastická sezónnost, rozšiřující se okno.....	50
Tab. 6: Podíl predikcí, u nichž out-of-sample přesnější než in-sample – rozšiřující se okno.....	51
Tab. 7: Nejčastěji vybrané auto.arima modely a volba driftu – rozšiřující se okno	52
Tab. 8: RMSE po obdobích – ARIMA (0,1,1)+dummy, rozšiřující se okno	61
Tab. 9: RMSE po obdobích – SARIMA (0,1,1)(0,1,1), rozšiřující se okno	62
Tab. 10: Rozdíly RMSE mezi rozšiřujícím se a posuvným oknem po obdobích – auto.arima, deterministická sezónnost.....	69
Tab. 11: Nejčastěji vybrané modely – posuvné a rozšiřující se okno	70
Tab. 12: Rozdíly RMSE mezi rozšiřujícím se a posuvným oknem po obdobích – fixní model ARIMA(0,1,1)+dummy.....	71
Tab. 13: Testy stacionarity – out-of-sample vzorek pro finanční krizi..	78
Tab. 14: Výběr modelu pro finanční krizi na základě BIC.....	80
Tab. 15: RMSE predikcí – finanční krize, manuální výběr	81
Tab. 16: Testy stacionarity – out-of-sample vzorek pro rok 2018	82
Tab. 17: Výběr modelu pro rok 2018 na základě BIC.....	84
Tab. 18: RMSE predikcí – rok 2018, manuální výběr	85
Tab. 19: Rozdíly RMSE mezi log a non-log transformací – ČR, auto.arima, deterministická sezónnost, rozšiřující se okno.....	97
Tab. 20: RMSE ARIMA a VAR modelů – ČR, auto.arima, deterministická sezónnost, rozšiřující se okno.....	97
Tab. 21: MAE po obdobích – auto.arima, deterministická sezónnost, rozšiřující se okno.....	98
Tab. 22: MAE po obdobích – auto.arima, stochastická sezónnost, rozšiřující se okno.....	99
Tab. 23: MAE po obdobích – ARIMA(0,1,1)+dummy, rozšiřující se okno	99
Tab. 24: MAE po obdobích – SARIMA(0,1,1)(0,1,1), rozšiřující se okno	100
Tab. 25: Testy stacionarity – in-sample vzorek finanční krize	103

SEZNAM TABULEK

Tab. 26: Testy stacionarity – in-sample vzorek pro rok 2018	103
---	-----

Seznam pojmů a zkratek

ACF	– autokorelační funkce
ADF test	– rozšířený Dickey-Fuller test
AIC	– Akaikeho informační kritérium
AR	– autoregresní model
ARIMA	– autoregresní integrovaný model s klouzavým průměrem
BIC	– Bayesovské informační kritérium
ČR	– Česká republika
HEGY test	– Hylleberg-Engle-Granger-Yoo test
IPP	– index průmyslové produkce
IS	– in-sample
KPSS test	– Kwiatkowski-Phillips-Schmidt-Shin test
MA	– model klouzavého průměru
MAE	– střední absolutní chyba
OCSB test	– Osborn-Chui-Smith-Birchenhall test
OOS	– out-of-sample
PACF	– parciální autokorelační funkce
RMSE	– odmocnina ze střední kvadratické chyby
SARIMA	– sezónní autoregresní integrovaný model s klouzavým průměrem
SK	– Slovensko
VAR	– vektorový autoregresní model

1 Úvod

V literatuře o predikci časových řad se standardně rozlišují in-sample a out-of-sample predikce. In-sample predikce hodnotí, jak dobře model popisuje data, ze kterých byl odhadnut. Out-of-sample predikce naopak měří jeho schopnost predikovat data, která při odhadu neviděl. Oba typy predikcí se v literatuře běžně uvádějí vedle sebe, ovšem přímé srovnání in-sample a out-of-sample predikční přesnosti zůstává stranou hlavního zájmu. Většina existujících prací se vztahem mezi in-sample a out-of-sample metrikami zabývá ve formě statistických testů predikční síly jednotlivých prediktorů, nikoli kvantifikací samotného rozdílu. Přitom velikost tohoto rozdílu vypovídá o tom, nakolik lze in-sample přesnost považovat za spolehlivý ukazatel očekávané out-of-sample predikční přesnosti.

Pro analýzu volíme index průmyslové produkce, který patří mezi klíčové krátkodobé makroekonomické ukazatele. Průmysl tvoří významnou část přidané hodnoty v otevřených ekonomikách střední Evropy a vyznačuje se výraznou cyklickou citlivostí. Index je zveřejňován s typicky dvouměsíčním zpožděním, což je jedno z nejkratších mezi makroekonomickými ukazateli. Využívá se proto pro nowcasting hrubého domácího produktu, jehož čtvrtletní frekvence a delší publikační zpoždění omezují využitelnost v reálném čase. Centrální banky index sledují jako jeden z hlavních vstupů do krátkodobých predikčních modelů a pro průběžné hodnocení ekonomické situace. Predikci indexu průmyslové produkce se v literatuře dlouhodobě věnuje pozornost a standardním benchmarkem zůstávají ARIMA modely.

Cílem této diplomové práce je posoudit velikost rozdílu mezi in-sample a out-of-sample predikcemi indexu průmyslové produkce a metodologické faktory, které tento rozdíl ovlivňují. Analýza je provedena pro Českou republiku a Slovensko na měsíčních datech za období leden 2000 až prosinec 2023. Pro modelování jsou využity ARIMA modely v několika metodologických variantách zahrnujících různé přístupy k sezónnosti, konstrukci tréninkového vzorku i volbě řádu modelu. Vzorek je rozdělen na pět období podle makroekonomického kontextu, což umožňuje posoudit výkon predikcí jak ve stabilních obdobích, tak v obdobích strukturálních zlomů (finanční krize 2008–2009, pandemie COVID-19).

Práce je kromě úvodu členěna do osmi kapitol. Druhá kapitola obsahuje rešerši literatury věnující se vztahu mezi in-sample a out-of-sample

predikcemi, predikci průmyslové produkce a volbě modelů. Třetí kapitola představuje data a základní deskriptivní statistiky pro oba sledované státy. Čtvrtá kapitola zavádí metodologický rámec, konkrétně formální zápis ARIMA a SARIMA modelů, přístupy k zachycení sezónnosti, konstrukci in-sample a out-of-sample predikcí a dekompozici jejich rozdílů. Pátá kapitola prezentuje predikce rozšiřujícího se okna pro obě země a oba přístupy k sezónnosti. Šestá kapitola srovnává predikce rozšiřujícího se a posuvného okna. Sedmá kapitola se zaměřuje na manuální výběr modelu pro dvě kontrastní období, finanční krizi 2008–2009 a stabilní rok 2018. Osmá kapitola diskutuje získané výsledky, jejich interpretaci a limity práce. Devátá, závěrečná kapitola shrnuje hlavní zjištění a jejich praktické implikace.

2 Rešerše literatury

2.1 In-sample a out-of-sample predikce

Klíčovou otázkou práce je vztah mezi in-sample (IS) a out-of-sample (OOS) predikcemi. IS predikce hodnotí, jak dobře model popisuje data, ze kterých byl odhadnut. OOS predikce naopak měří jeho schopnost predikovat data, která při odhadu neviděl.

Literatura se problematice věnuje převážně prostřednictvím testů toho, zda konkrétní vysvětlující proměnná má predikční sílu pro cílovou veličinu (například zda úrokový spread predikuje inflaci nebo zda dividend-price ratio predikuje akciové výnosy). IS testy tuto otázku hodnotí t-testem nebo F-testem významnosti prediktoru v regresi na celém vzorku, OOS testy ji hodnotí srovnáním predikčních chyb modelu s benchmarkem. Inoue a Kilian (2005) teoreticky ukazují, že IS testy mají v mnoha situacích vyšší sílu než OOS testy. Clark (2004) naopak obhajuje OOS srovnání jako prevenci overfittingu. V makroekonomickém kontextu Clark a McCracken (2006) na příkladu Phillipsovy křivky ukazují, že rozpor mezi IS a OOS testy lze z velké části vysvětlit nízkou silou OOS testů doplněnou nestabilitou koeficientů. V oblasti predikce akciových výnosů Goyal a Welch (2008) dokumentují, že většina prediktorů equity premium selhává jak v IS, tak v OOS, zatímco Campbell a Thompson (2008) v reakci ukazují, že slabé restriktce na koeficienty umožňují predikčním modelům překonat historický průměr. Rapach a Wohar (2006) argumentují, že rozpor mezi IS a OOS testy se při použití silných OOS testů výrazně zmenšuje.

V literatuře o predikci časových řad však přímé srovnání IS a OOS přesnosti predikcí obvykle není hlavním tématem. Autoři zpravidla uvádějí IS diagnostické metriky (R^2 , informační kritéria, RMSE či MAE na tréninkovém vzorku) i OOS predikční metriky samostatně, bez jejich explicitního srovnání na shodných pozorováních.

2.2 Volba transformace a přístup k sezónnosti

Důležitým metodologickým rozhodnutím při predikci průmyslové produkce je zachycení sezónnosti, která vysvětluje značnou část variability dat. Literatura nabízí dva hlavní přístupy: deterministické sezónní

dummy proměnné v prvních diferencích a stochastickou sezónnost zachycenou sezónním diferencováním, případně v rámci SARIMA. Volba transformace ovlivňuje predikční přesnost a závisí na horizontu predikce i vlastnostech dat.

Franses a van Dijk (2005) srovnávají devět modelů s různými přístupy k modelování sezónnosti pro čtvrtletní data průmyslové produkce 18 zemí OECD na datech z období 1960–2002. Predikují tempo růstu, jelikož data logaritmují a diferencují. Autoři u všech modelů předpokládají jednotkový kořen na nulové frekvenci a provádí tedy první diferenci. Sezónnost modelují pomocí ARIMA modelů s deterministickými dummy sezónními proměnnými, SARIMA modelů se sezónními diferencemi, ale i složitějších modelů s nelineární specifikací a proměnlivou sezónností. Řády ARIMA modelů jsou voleny podle BIC kritéria za podmínky nepřítomnosti autokorelace reziduí. Out-of-sample predikce jsou generovány pro období 1988–2002 na horizontech 1 až 12 čtvrtletí. Autoři používají přístup rozšiřujícího se okna a pro každý krok znovu respecifikují model a reestimují parametry. Na krátkých horizontech (1–3 čtvrtletí) dosahují přesnějších predikcí jednoduché modely ARIMA s deterministickou dummy složkou a SARIMA. Na delších horizontech je naopak překonávají modely s časově proměnlivou sezónností. Přesnost predikcí se rovněž liší podle fází hospodářského cyklu. Zatímco jednoduché modely predikují lépe v expanzích, složitější modely lépe zachytí recese.

Osborn, Heravi a Birchenhall (1999) srovnávají čtyři AR modely lišící se přístupem k sezónnosti pro měsíční data průmyslové výroby osmi průmyslových odvětví v Německu, Francii a Velké Británii na datech z období 1978–1995. Predikují tempo růstu, jelikož data logaritmují a diferencují. Sezónnost vysvětluje u většiny řad přes 80 % variability prvních diferencí, u francouzských řad přes 90 %. Autoři testují sezónní jednotkové kořeny pomocí OCSB a HEGY testů. Testy zamítají přítomnost jednotkových kořenů u většiny řad, což podporuje modelování pomocí sezónních dummy proměnných s prvními diferencemi. Autoři nicméně používají modely s různými specifikacemi. Srovnávané modely zahrnují AR v prvních a sezónních diferencích, AR v sezónních diferencích, AR v prvních diferencích se sezónními dummy proměnnými a AR v úrovních se sezónními dummy a trendem. Řád autoregrese je volen podle HQ kritéria. Out-of-sample predikce na horizontech 1, 3, 6 a 12 měsíců jsou generovány pro období 1994–1995 přístupem rozšiřujícího se okna. Autoři porovnávají tři strategie: model s fixními parametry, rozšiřující se okno s reestimací parametrů a rozšiřující se okno s reestimací parametrů

i respecifikací řádu AR. Reestimace parametrů přináší největší přínosy u modelů se sezónními dummy proměnnými, kde může dosáhnout snížení RMSE až o 30 %, zatímco u modelů se sezónními diferencemi jsou zisky z reestimace podstatně menší. Dodatečná respecifikace řádu AR přináší jen marginální zlepšení oproti samotné reestimaci. Navzdory výsledkům OCSB a HEGY testů často specifikace se sezónními diferencemi produkuje přesnější out-of-sample predikce, zejména na delších horizontech. Model v prvních diferencích se sezónními dummy dosahuje nejlepších výsledků na krátkých horizontech u řad se silnou sezónností.

K obdobnému závěru jako Osborn, Heravi a Birchenhall (1999) docházejí Clements a Hendry (1997). Na čtvrtletních datech ukazují, že sezónní diferencování zlepšuje predikční přesnost i v případech, kdy HEGY test hovoří ve prospěch deterministické sezónnosti. Tento paradox vysvětlují robustností sezónního diferencování vůči strukturálním zlomům v sezónním chování. Diferencování automaticky přizpůsobuje predikce novému sezónnímu vzorci, zatímco deterministické dummy proměnné fixují sezónnost na základě historických odhadů.

Bruno a Lupi (2004) volí pro predikci italské průmyslové výroby z období 1988–2001 roční difference jako transformaci. Volbu zdůvodňují třemi argumenty: zájem o meziroční tempa růstu z hlediska krátkodobé ekonomické analýzy, empirická evidence naznačující, že filtrování pouze správných jednotkových kořenů nepřináší lepší predikce, a robustnost ročních diferencí vůči náhlým strukturálním změnám.

Bodo, Golinelli a Parigi (2000) pracují s logaritmovanými meziměsíčními diferencemi dat průmyslové produkce eurozóny a sezónnost modelují pomocí deterministických dummy proměnných doplněných o sezónní autoregresní složky. Volbu deterministické sezónnosti zdůvodňují OCSB testem, který zamítá přítomnost sezónních jednotkových kořenů.

2.3 Využití jednorozměrných a vícerozměrných modelů pro predikce indexu průmyslové produkce

Dalším rozhodnutím při predikci průmyslové produkce je volba mezi jednorozměrnými ARIMA modely a vícerozměrnými specifikacemi zahrnujícími exogenní proměnné. Literatura testuje různé přístupy. Od tradičních VAR modelů přes kointegrační ECM specifikace po faktorové modely redukující rozsáhlé datové soubory.

Bianchi et al. (2010) predikují průmyslovou výrobu Německa, Francie a Itálie na měsíčních sezónně očištěných datech z období 1990–2005. Data jsou transformována do logaritmovaných meziměsíčních diferencí. Trace test zamítá přítomnost kointegrace mezi průmyslovými výrobami tří zemí, a autoři proto volí VAR model v diferencích namísto VECM. Srovnávají copula-VAR model, standardní VAR, VAR s vícerozměrným studentovým t-rozdělením a ARMA(1,1) model. Vícerozměrné modely zahrnují exogenní proměnné (index průmyslové důvěry, úrokový spread, reálné úrokové míry a výnosy akciových trhů), zatímco ARMA(1,1) pracuje bez nich. Modely jsou odhadnuty na období 1990–2002 a out-of-sample predikce na horizontech jednoho a tří měsíců jsou generovány pro období 2003–2005 s fixními parametry bez reestimace. Index průmyslové důvěry se ukazuje jako významný předstihový ukazatel ve všech třech zemích, zatímco reálná úroková míra a úrokový spread jsou marginálně signifikantní pouze pro Itálii. Pro Itálii ARMA(1,1) mírně překonává všechny vícerozměrné modely navzdory absenci exogenních proměnných. Hansenův SPA test potvrzuje, že pro Francii je ARMA(1,1) signifikantně horší, avšak pro Německo a Itálii nejsou rozdíly mezi modely statisticky signifikantní.

Bodo, Golinelli a Parigi (2000) se zaměřují na predikci indexu průmyslové výroby eurozóny jako celku i jednotlivých zemí (Německo, Francie, Itálie, Španělsko) na měsíčních datech z období 1987–1999. Autoři testují kointegraci ve VAR modelu eurozóny a USA, kterou nacházejí. Proto používají ECM model, který využívá dlouhodobého rovnovážného vztahu mezi průmyslovými výrobami obou oblastí, a navíc zahrnuje index průmyslové důvěry. Naproti tomu čtyřrozměrný VAR model zahrnující IPP Německa, Francie, Itálie a Španělska kointegraci nenachází, a autoři ho proto odhadují v prvních diferencích. Kromě společného VAR modelu predikují autoři také IPP jednotlivých zemí samostatně. Jako benchmarkové modely jsou specifikovány ARIMA(p,1,0), jelikož autoři nenacházejí signifikantní MA složku. Out-of-sample predikce na jeden měsíc dopředu jsou generovány pro období 1997–1999 posuvným oknem o velikosti 96 měsíců s reestimací parametrů při každém kroku. Nejlepších výsledků dosahuje ECM model zahrnující USA IPP a index průmyslové důvěry. Jednoduchý ARIMA model překonává modely jednotlivých zemí i VAR modely bez exogenních proměnných, což autoři vysvětlují tím, že VAR bez vhodných indikátorů trpí nedostatečným počtem stupňů volnosti vzhledem k velkému množství odhadovaných parametrů.

Bruno a Lupi (2004) predikují italskou průmyslovou výrobu na měsíčních datech z období 1988–2001. Srovnávají ARMA(2,2) benchmark vybraný podle BIC kritéria s VAR modelem zahrnujícím index průmyslové výroby, objem železniční přepravy zboží a index očekávání budoucí výroby z podnikových průzkumů. VAR specifikaci vybírají přístupem general-to-specific z počáteční specifikace se 14 lagy. Out-of-sample predikce na horizontech 1 až 6 měsíců jsou generovány pro období 1999–2001 rozšiřujícím se oknem s reestimací parametrů. VAR model dosahuje na horizontech 1 až 4 měsíce lepších výsledků než ARMA benchmark, přičemž na horizontech 3 a 4 měsíce je rozdíl statisticky signifikantní podle modifikovaného Diebold-Mariano testu. Na delších horizontech (5–6 měsíců) je naopak ARMA benchmark přesnější.

Bulligan, Golinelli a Parigi (2010) analyzují predikční schopnosti dvanácti modelů pro italský index průmyslové produkce na horizontech 1 až 3 měsíce na měsíčních neočištěných datech z období 1980–2007 v úrovních bez logaritmické transformace. Srovnávají ARIMA(3,1,0) benchmark se dvěma alternativními přístupy: jednorovnicovými modely využívajícími několik vybraných krátkodobých indikátorů (spotřeba elektřiny, teploty, indexy z podnikových průzkumů) a faktorovými modely, které shrnují datový soubor 110 kvantitativních i kvalitativních ukazatelů do několika hlavních faktorů. Sezónnost je v modelech zachycena kombinací zmíněných indikátorů (spotřeba elektřiny a teploty mají přirozený roční vzorec) a speciálních dummy proměnných. Out-of-sample predikce jsou generovány pro období 2003–2007 posuvným oknem s reestimací parametrů. Autoři testují vliv délky estimačního okna (48, 84 a 138 měsíců) a nacházejí, že optimální délka odpovídá přibližně jednomu hospodářskému cyklu (84 měsíců), přičemž kratší i delší okna dávají statisticky signifikantně horší predikce. Důvodem je, že kratší okno zvyšuje variabilitu odhadů parametrů, zatímco delší zahrnuje strukturálně odlišná období, která zkreslují odhady. ARIMA benchmark dosahuje nejhorších výsledků ze všech testovaných přístupů a modely s vybranými indikátory ho signifikantně překonávají na všech horizontech.

2.4 Alternativní přístupy k predikci indexu průmyslové produkce

Kromě standardních ARIMA a VAR modelů se literatura věnuje i alternativním přístupům k predikci průmyslové výroby.

Hassani, Heravi a Zhigljavsky (2009) aplikují neparametrickou metodu Singular Spectrum Analysis (SSA) na 24 měsíčních neočištěných řad průmyslové výroby v Německu, Francii a Velké Británii (osm odvětví pro každou zemi) v logaritmické formě. SSA rozkládá řadu na součet nezávislých komponent (trend, oscilace, šum) bez parametrických předpokladů a na rozdíl od ARIMA tak nevyžaduje explicitní volbu diferencování. Autoři srovnávají SSA s Holt-Wintersovou metodou a SARIMA modely, jejichž řády $(p, d, q)(P, D, Q)_{12}$ jsou automaticky voleny minimalizací BIC kritéria. Na krátkých horizontech dosahují všechny tři metody srovnatelných výsledků, na delších horizontech SSA výrazně překonává ARIMA. V návazném článku Hassani, Heravi a Zhigljavsky (2013) rozšiřují SSA na vícerozměrnou variantu (MSSA) aplikovanou na osm řad průmyslové výroby Velké Británie a přidávají VAR model jako další benchmark. MSSA překonává jednorozměrný SSA, ARIMA i VAR modely, přičemž zlepšení oproti ARIMA dosahuje na různých horizontech 18 až 43 %. Výhoda SSA je patrná zejména v obdobích strukturálních zlomů, kdy je metoda robustnější vůči šokům v řadě než parametrické modely.

Thury a Witt (1998) srovnávají základní strukturální model (BSM) s modelem SARIMA(0,1,1)(0,1,1)₁₂ pro predikci rakouské a německé průmyslové výroby v logaritmické formě. Výsledky ukazují, že BSM produkuje přesnější predikce než SARIMA jak podle absolutních chyb, tak podle schopnosti správně předvídat směr změny, přičemž rozdíly se zvětšují s rostoucím horizontem predikce.

Tyto metody nabízejí alternativy ke standardním ARIMA a VAR modelům zejména v situacích, kdy jsou relevantní strukturální zlomy nebo delší horizonty predikce.

2.5 Shrnutí rešerše

Literatura ke vztahu mezi in-sample a out-of-sample přesností se zaměřuje především na statistické testy prediktability. Diskutuje se relativní síla IS a OOS testů, kdy IS testy bývají v mnoha situacích silnější, ale OOS srovnání naopak chrání před nadměrným přizpůsobením. Rozpor mezi

IS a OOS výsledky lze často vysvětlit nízkou silou OOS testů a nestabilitou koeficientů. Přímé srovnání IS a OOS predikční přesnosti však obvykle nebývá hlavním tématem výzkumu.

Napříč studiemi se ukazuje, že sezónní diferencování je z hlediska predikční přesnosti konkurenceschopné nebo lepší než deterministické dummy proměnné. To platí i v případech, kdy testy sezónních jednotkových kořenů hovoří ve prospěch deterministické sezónnosti. Volba mezi těmito přístupy závisí na horizontu predikce i vlastnostech dat. Na krátkých horizontech a u řad se silnou sezónností často vítězí deterministické dummy proměnné, zatímco na delších horizontech je sezónní diferencování obvykle přesnější. Modely s časově proměnlivou sezónností pak na dlouhých horizontech překonávají i jednodušší specifikace se sezónním diferencováním. Testy sezónních jednotkových kořenů tedy nemusí být spolehlivým vodítkem pro volbu predikčního modelu.

Index průmyslové důvěry se opakovaně ukazuje jako významný předstihový ukazatel průmyslové výroby. Vícerozměrné modely však nemusí automaticky překonávat jednoduché ARIMA benchmarky. Mohou trpět nadměrnou parametrizací a ztrátou stupňů volnosti, zejména pokud jim chybí vhodné exogenní proměnné. Přínosy vícerozměrných modelů se navíc koncentrují spíše na kratší horizonty, zatímco na delších se výhoda obvykle vytrácí.

Literatura se liší i ve volbě estimační strategie. Převažujícím přístupem je reestimace parametrů při každém novém pozorování, a to buď v posuvném okně s fixní délkou, nebo v rozšiřujícím se okně s rostoucím vzorkem. Výjimečně jsou modely odhadnuty jednorázově bez reestimace. Vedle samotné reestimace parametrů může být v každém kroku respecifikována i struktura modelu, například řád autoregrese podle informačního kritéria. Reestimace parametrů přináší výrazné zlepšení predikční přesnosti, u modelů se sezónními dummy proměnnými až o 30 %, zatímco dodatečná respecifikace struktury modelu přináší jen marginální zlepšení. Optimální délka estimačního okna odpovídá přibližně jednomu hospodářskému cyklu, přičemž kratší i delší okna vedou ke statisticky signifikantně horším predikcím.

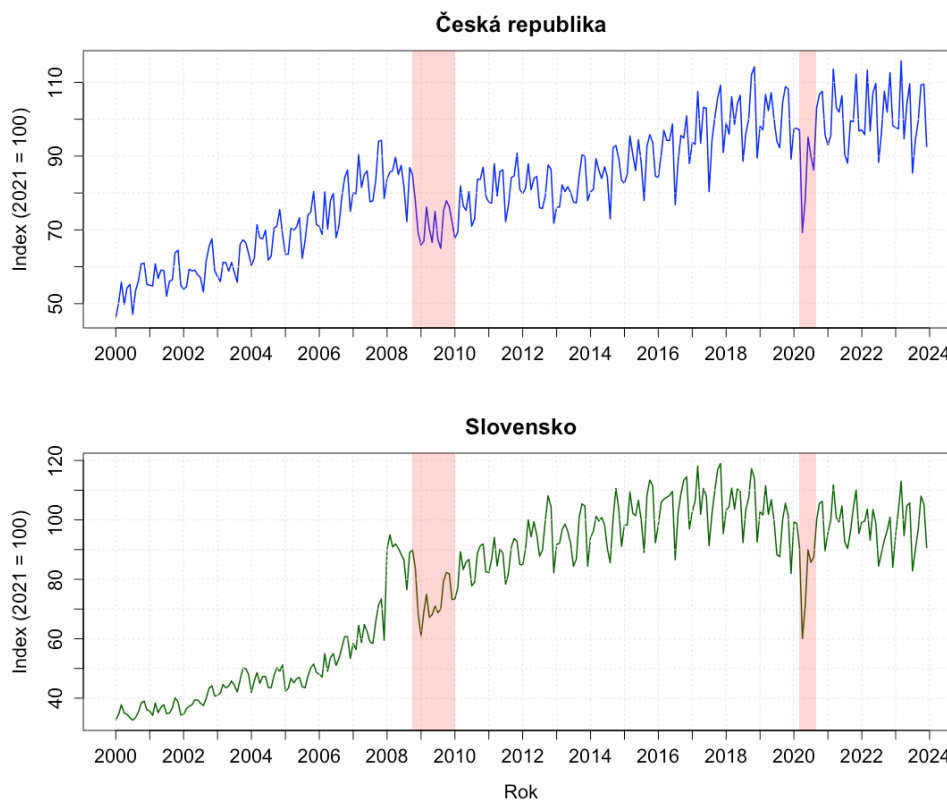
Existující studie predikce průmyslové výroby se zaměřují převážně na velké západoevropské ekonomiky s daty končícími před rokem 2010. Česká republika a Slovensko v této literatuře samostatně zkoumány nebyly. Tato práce tyto mezery v literatuře zaplňuje srovnáním in-sample a out-of-sample predikcí a aplikací na země a období, které v dosavadní literatuře predikovány nebyly.

3 Data

Data pro index průmyslové produkce pocházejí z databáze Eurostatu, datová sada Production in industry – monthly data (sts_inpr_m) (Eurostat, 2026). Pracujeme s agregátem sekcí B–D podle klasifikace NACE Rev. 2, který zahrnuje těžbu a dobývání, zpracovatelský průmysl a výrobu a rozvod elektřiny, plynu a tepla. Tento agregát odpovídá standardnímu souhrnnému indexu průmyslové produkce. Index je vyjádřen s referenčním rokem 2021 (2021 = 100). Zdrojová data shromažďují národní statistické úřady na základě měsíčních podnikových výkazů a následně je předávají Eurostatu. Data nejsou kalendářně ani sezónně očištěná, jelikož cílem je predikovat skutečný vývoj průmyslové produkce, nikoliv jeho sezónně očištěnou variantu.

Analyzované období zahrnuje leden 2000 až prosinec 2023, tedy 288 pozorování pro Česko a Slovensko. Vzorek obsahuje dostatečně dlouhý předkrizový úsek pro odhad ARIMA modelů a zároveň zahrnuje dvě období šoků, tedy finanční krizi 2008–2009 a pandemii COVID-19 v roce 2020. Data indexu průmyslové produkce jsou považována za finální s odstupem dvou let od publikace, do té doby mohou probíhat revize (Eurostat, 2026). Posledním pozorováním v analýze je proto prosinec 2023, který je v době vzniku práce nejpozdějším plně revidovaným měsícem.

Obr. 1 znázorňuje vývoj indexu průmyslové produkce obou zemí. Do roku 2008 vykazují obě řady rostoucí trend. Slovenská řada roste v tomto období rychleji než česká a v prvních měsících roku 2008 vykazuje obzvláště prudký nárůst. Finanční krize 2008–2009 přináší v obou zemích prudký propad, který začíná ve druhé polovině roku 2008 a dosahuje minima na počátku roku 2009. Po nejhlubším propadu následuje postupné zotavení, návrat na předkrizové úrovně však trvá několik let. Mezi lety 2010 a 2019 obě řady vykazují stabilní růst s pravidelnými sezónními výkyvy. Pandemie COVID-19 v březnu 2020 způsobuje výrazný, avšak krátký propad s minimem v dubnu 2020. Na rozdíl od finanční krize je šok COVID-19 krátkodobý a již během podzimu 2020 se obě řady vracejí k předpandemickému trendu.

Obr. 1: Vývoj indexu průmyslové produkce České republiky a Slovenska

Poznámka: červenými pruhy jsou označeny krize

Zdroj: vlastní zpracování

Tab. 1 zobrazuje deskriptivní statistiky IPP obou zemí. Průměr se mezi zeměmi výrazně neliší. Směrodatná odchylka je u slovenské řady výrazně vyšší, což je způsobeno nižší počáteční úrovní slovenského IPP a hlubším propadem v krizích.

Tab. 1: Deskriptivní statistiky indexu průmyslové produkce

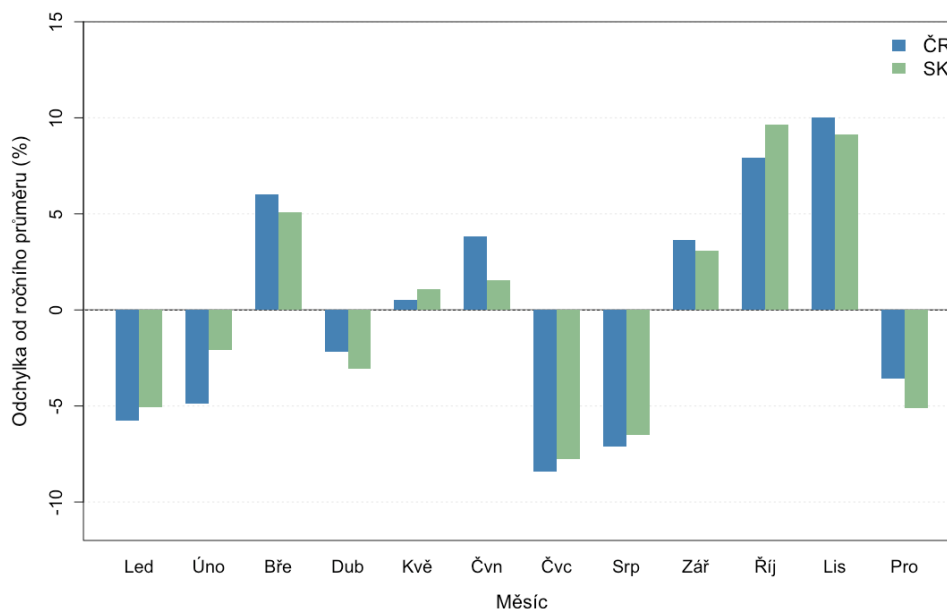
Statistika	ČR	SK
Počet pozorování (n)	288	288
Průměr	81,96	78,62
Směrodatná odchylka	16,01	25,77
Minimum	46,20	32,60
Maximum	115,80	119,00

Zdroj: vlastní zpracování

Obě řady vykazují výrazné pravidelné měsíční výkyvy, které jsou viditelné již v grafu vývoje indexu. Pro detailnější pohled na sezónní vzorec zobrazuje Obr. 2 průměrné procentní odchylky indexu od ročního průměru v jednotlivých kalendářních měsících. Sezónní profil obou zemí je velmi podobný. Nejnižších hodnot index dosahuje v červenci a srpnu, což odpovídá letnímu útlumu průmyslové aktivity spojenému s dovolenými a plánovanými odstávkami. Druhý podprůměrný úsek pokrývá zimní měsíce od prosince do února. Nejvyšší hodnoty index vykazuje v říjnu a listopadu, další lokální vrchol se objevuje v březnu. Výkyvy mezi sousedními měsíci jsou částečně ovlivněny různým počtem pracovních dnů. Například rozdíl mezi počtem pracovních dnů v únoru a březnu může přispívat k výraznému nárůstu v březnu, stejně jako Velikonoce snižují produkci v dubnu. Data nejsou kalendářně ani sezónně očištěná, takže tyto efekty v nich zůstávají zachovány.

Sílu sezónnosti kvantifikujeme po vzoru Osborn, Heravi a Birchenhall (1999) pomocí koeficientu determinace R^2 z regrese prvních diferencí indexu na měsíční dummy proměnné. Pro českou řadu dosahuje R^2 hodnoty 65 % a pro slovenskou 56 %. Oba podíly dokládají, že sezónní složka je v obou řadách dominantním zdrojem měsíční variability.

Obr. 2: Sezónní odchylky indexu průmyslové produkce od ročního průměru



Poznámka: Procentuální odchylka daného měsíce v daném roce od ročního průměru, zprůměrovaná přes všechny roky

Zdroj: vlastní zpracování

Pro modely ARIMA je klíčové ověřit vlastnost stacionarity. U sezónních měsíčních dat však standardní testy jako ADF nebo KPSS test zachytí pouze jednotkový kořen na nulové frekvenci a nedokážou odhalit potenciální jednotkové kořeny na sezónních frekvencích. Pro tato data je proto vhodnější test HEGY (Hylleberg, Engle, Granger a Yoo, 1990) ve verzi pro měsíční data podle Beaulieu a Miron (1993). Test rozkládá AR polynom na faktory odpovídající jednotlivým frekvencím a testuje, zda je na nich přítomen jednotkový kořen. Pro měsíční data se testuje sedm frekvencí: nulová (t_1), Nyquistova π (t_2) a pět komplexních párů na sezónních frekvencích ($F_{3:4}$ až $F_{11:12}$). Z pohledu modelování je klíčové rozlišit mezi jednotkovým kořenem na nulové frekvenci (implikuje potřebu první diference) a na sezónních frekvencích (implikuje potřebu sezónní diference). Stejnou metodiku pro IPP aplikují například Osborn, Heravi a Birchenhall (1999) a Bruno a Lupi (2004).

Výsledky HEGY testu uvádí Tab. 2. Pro obě země nemůžeme zamítnout nulovou hypotézu jednotkového kořene na nulové frekvenci, což naznačuje potřebu první diference. Společný F-test sezónních frekvencí naopak nulovou hypotézu zamítá pro obě země, což naznačuje, že není

nutné sezónní diferencování a sezónní složku lze modelovat deterministickými dummy proměnnými. Na některých individuálních sezónních frekvencích jednotkový kořen zamítnout nelze. Podobné výsledky dostávají i Bruno a Lupi (2004), kteří uvádějí, že přítomnost jednotkových kořenů na některých jednotlivých frekvencích nelze nikdy zcela vyloučit. Rozhodující je pro nás výsledek společného $F_{2:12}$ testu. Obě řady tedy vyžadují první diferenci ($d = 1$), ale nikoli sezónní diferenci ($D = 0$).

Tab. 2: HEGY test jednotkových kořenů

Test	ČR	SK
t_1	-2,83	-0,44
t_2	-3,03*	-2,99*
$F_{3:4}$	11,49**	13,13**
$F_{5:6}$	12,17**	4,62
$F_{7:8}$	5,51	5,11
$F_{9:10}$	22,64**	13,89**
$F_{11:12}$	8,00*	5,71
$F_{2:12}$	14,83**	11,25**
$F_{1:12}$	14,14**	10,32**

Poznámka: Hodnoty signifikantní na hladině 5 % označeny *, na hladině 1 % **.

Zdroj: vlastní zpracování

Při modelování časových řad je otázkou, zda řadu logaritmovat nebo pracovat s úrovněmi. Logaritmická transformace se aplikuje na řady s multiplikativní strukturou, kdy amplituda výkyvů roste úměrně s úrovní. Většina autorů při predikci IPP logaritmickou transformací aplikuje bez explicitního testu její vhodnosti. Jedním z mála příkladů, kdy je volba mezi lineární a logaritmickou specifikací formálně ověřena, jsou Marchetti a Parigi (2000), kteří na základě Box-Coxovy analýzy volí lineární specifikaci, tedy úroveň bez transformace. Pro posouzení vhodnosti transformace v našich datech používáme stejný přístup, tedy Box-Coxovu analýzu, která hledá optimální hodnotu parametru λ mocninné transformace. Hodnota λ blízko nuly signalizuje vhodnost logaritmické transformace. Hodnota blízko jedné naopak naznačuje, že úroveň jsou dostatečné. Metoda maximální věrohodnosti dává pro české IPP $\lambda \approx 1,05$ a pro slovenské $\lambda \approx 1,30$. Obě hodnoty jsou blízké jedné, z čehož vyplývá, že úroveň indexu lze modelovat bez transformace. Modelování přímo na úrovních je navíc vhodnější z praktického hlediska, protože predikujeme

IPP na původní škále indexu a logaritmická varianta by vyžadovala zpětný převod predikcí. V Příloze A uvádíme srovnání modelů odhadnutých na logaritmovaných datech s modely v úrovních, které ukazuje, že logaritmická transformace nepřináší zlepšení predikcí.

4 Metodologie

V této kapitole popisujeme metody použité v analytické části. Postupně zavádíme dva přístupy k zachycení sezónnosti v datech IPP, formální zápis ARIMA a SARIMA modelů, přístupy k volbě řádů modelu a konstrukci in-sample a out-of-sample predikcí. Kapitulu uzavírá popis metrik predikční přesnosti a rozdělení vzorku na období pro jejich hodnocení.

4.1 Volba mezi ARIMA a VAR modely

Před volbou konkrétní specifikace je třeba rozhodnout mezi jednorozměrnými a vícerozměrnými modely. Pro predikci indexu průmyslové produkce se v literatuře běžně používají ARIMA modely jako benchmark, zatímco VAR modely s exogenními indikátory mohou v některých případech přinášet zlepšení.

Prvním krokem analýzy tedy bylo srovnání ARIMA a VAR specifikace pro Českou republiku. Jde o dvourozměrné modely s Německem jakožto hlavním obchodním partnerem a se Slovenskem jakožto druhou ekonomikou zkoumanou v této práci, dále trojrozměrný VAR a varianty s indexem průmyslové důvěry jako endogenní proměnnou v duchu Bruno a Lupi (2004) nebo jako exogenní regresor v duchu Bianchi et al. (2010). Johansenův kointegrační test neprokázal dlouhodobý rovnovážný vztah mezi IPP sousedních ekonomik, a proto jsme VAR modely specifikovali v prvních diferencích. Detailní výsledky pro VAR modely s rozšiřujícím se oknem (kde dosahují nejlepších výsledků) jsou zobrazeny v Příloze A.

Žádná z testovaných VAR variant nepřekonává jednoduchý ARIMA+dummy model napříč horizonty. Tento výsledek je v souladu s Bianchi et al. (2010) a Bodo, Golinelli a Parigi (2000), kteří rovněž nenacházejí signifikantní zlepšení vícerozměrných modelů oproti jednoduchému ARIMA. Pro hlavní analýzu in-sample a out-of-sample predikcí proto volíme jednorozměrný ARIMA model. Místo srovnávání ARIMA modelů s VAR se zaměřujeme na srovnání ARIMA modelů s různými přístupy k zachycení sezónnosti.

4.2 Zachycení sezónnosti

Sezónnost je v datech IPP dominantní složkou variability a její specifikace je důležitým rozhodnutím. Literatura rozlišuje dva základní přístupy, které se liší předpokladem o povaze sezónního vzorce. Deterministická sezónnost předpokládá, že sezónní vzorec je v čase stabilní a každý rok se opakuje ve stejné podobě. Stochastická sezónnost naopak připouští, že se sezónní vzorec v čase postupně mění. (Ghysels a Osborn, 2001)

HEGY test provedený v kapitole Data dochází k závěru, že v datech není sezónní jednotkový kořen a sezónnost by se tak měla modelovat jako deterministická. Tyto výsledky je však třeba brát s rezervou, jelikož přítomnost strukturálních zlomů může výsledky HEGY testu zkreslovat (Clements a Hendry, 1997). Navíc literatura ukazuje, že sezónní diferencování může vést k přesnějším predikcím i tam, kde testy hovoří ve prospěch deterministické sezónnosti (Osborn, Heravi a Birchenhall, 1999). Z těchto důvodů v práci sezónnost modelujeme oběma přístupy.

4.3 ARIMA modely

Třída modelů ARIMA (Autoregressive Integrated Moving Average) patří mezi nejrozšířenější nástroje analýzy a predikce jednorozměrných časových řad. Podstatou přístupu je popsat chování sledované řady prostřednictvím její vlastní minulosti a posloupnosti náhodných šoků. U nestacionárních řad je tento přístup podmíněn předchozím diferencováním. Formální zápis rovnic v této podkapitole vychází z Hyndman a Athanasopoulos (2021).

Pro kompaktní zápis ARIMA modelu je výhodné použít zpoždovací operátor B , pro který platí:

$$By_t = y_{t-1} \quad (1)$$

Model ARIMA(p, d, q) lze potom zapsat jako:

$$\begin{aligned} (1 - \phi_1 B - \dots - \phi_p B^p)(1 - B)^d y_t \\ = c + (1 + \theta_1 B + \dots + \theta_q B^q) \varepsilon_t \end{aligned} \quad (2)$$

kde $\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$ je autoregresní polynom řádu p , člen $(1 - B)^d$ představuje d -násobnou diferenci zajišťující stacionaritu, $\theta(B) = 1 + \theta_1 B + \dots + \theta_q B^q$ je polynom klouzavého průměru řádu q . Konstanta c v modelu s první diferencí odpovídá lineárnímu trendu v úrovních y_t a označuje se jako drift. Symbol ε_t představuje bílý šum s nulovou střední hodnotou a konstantním rozptylem. Koeficienty ϕ kvantifikují sílu vazby mezi současnou hodnotou (po diferencování) a jejími zpožděními, zatímco koeficienty θ popisují, jak se minulé náhodné výkyvy promítají do současné hodnoty.

Sezónní strukturu v řadě lze do ARIMA modelu zahrnout přidáním dalších členů na zpožděních odpovídajících sezónní periodě. Výsledkem je SARIMA model, značený jako $\text{ARIMA}(p, d, q)(P, D, Q)_s$. V tomto modelu označují malá písmena nesezónní část, velká sezónní část, a s je délka sezónní periody. Sezónní část má analogickou strukturu jako nesezónní, ale používá zpožďovací operátor závislý na délce sezónní periody. Pro měsíční data, s nimiž pracujeme v této práci, je sezónní perioda $s = 12$. Například model $\text{ARIMA}(1, 1, 1)(1, 1, 1)_{12}$ bez konstanty lze zapsat jako:

$$\begin{aligned} (1 - \phi_1 B)(1 - \phi_1 B^{12})(1 - B)(1 - B^{12})y_t \\ = (1 + \theta_1 B)(1 + \theta_1 B^{12})\varepsilon_t \end{aligned} \quad (3)$$

Alternativním přístupem k modelování sezónnosti je její deterministické zachycení pomocí indikátorových (dummy) proměnných pro jednotlivé měsíce. Tento přístup předpokládá, že sezónní vzorec je v čase stabilní a lze jej zachytit konstantními koeficienty. Striktně vzato se nejedná o čistý ARIMA model, ale o regresi s ARIMA chybami, kde dummy proměnné tvoří systematickou regresní složku a reziduální proces u_t má ARIMA strukturu. V obecné formě lze tento model zapsat jako:

$$y_t = \beta_0 + \sum_{m=1}^{11} \beta_m D_{m,t} + u_t, u_t \sim \text{ARIMA}(p, d, q) \quad (4)$$

kde $D_{m,t}$ jsou binární dummy proměnné měsíců $m = 1, \dots, 11$ (prosinec slouží jako referenční kategorie), β_0 je konstanta (úroveň v referenčním měsíci), β_m jsou v čase konstantní sezónní odchylky jednotlivých měsíců od prosince, a reziduální složka u_t má $\text{ARIMA}(p, d, q)$ strukturu. (Hyndman a Athanasopoulos, 2021)

4.4 Volba řádů modelu

V této práci používáme několik přístupů k volbě modelu. Prvním je využití funkce `auto.arima`, která automaticky volí model. Dalším přístupem je využití fixní specifikace pro celý vzorek. Posledním přístupem je zvolení modelu ARIMA manuálně na základě ACF a PACF a informačního kritéria.

4.4.1 Informační kritéria

Při volbě řádů ARIMA modelu je třeba rozhodnout, kolik autoregresních a klouzavých členů zahrnout. Přidání dalších zpoždění snižuje součet čtverců reziduí, zároveň však spotřebovává stupně volnosti a může zhoršit predikční výkonnost modelu. Informační kritéria umožňují tento kompromis kvantifikovat, jelikož penalizují složitější modely proti jednoduchým. (Enders, 2014)

V literatuře jsou dvě nejčastější informační kritéria: Akaikeho informační kritérium (AIC) a Bayesovské informační kritérium (BIC). Jejich základní tvary jsou:

$$AIC = T \cdot \ln(SSR) + 2n \quad (5)$$

$$BIC = T \cdot \ln(SSR) + n \cdot \ln(T) \quad (6)$$

kde SSR je součet čtverců reziduí, n počet odhadovaných parametrů ($p + q +$ případně konstanta) a T počet použitelných pozorování. Model s nižší hodnotou kritéria je preferován. (Enders, 2014)

Klíčový rozdíl mezi oběma kritérii spočívá v penalizaci dodatečného parametru. AIC ji váží koeficientem 2, zatímco BIC koeficientem $\ln(T)$, který pro typické velikosti vzorků výrazně převyšuje 2. Volba mezi AIC a BIC závisí na cíli analýzy. AIC cílí na model s nejlepší predikční přesností pro daný vzorek, zatímco BIC je konzistentní a s rostoucím počtem pozorování konverguje k výběru správného modelu (Burnham a Anderson, 2004). AIC má přitom tendenci volit složitější modely, jejichž přebytečné parametry mohou zhoršit predikce mimo vzorek.

V této práci je jako informační kritérium zvoleno BIC, které penalizuje počet parametrů přísněji než AIC a vede k úspornějším modelům. Volba je rovněž v souladu s články zabývajícími se predikcemi IPP, kde autoři modely volí na základě BIC.

4.4.2 ACF a PACF

Pro identifikaci řádů p a q v ARIMA modelu srovnáváme výběrovou autokorelační funkci (ACF) a parciální autokorelační funkci (PACF) s teoretickými vzorci různých ARMA procesů. ACF udává korelaci mezi pozorováními řady v různých zpožděních a PACF měří korelaci v daném zpoždění po eliminaci vlivu mezilehlých pozorování. Pro čistě autoregresní proces $AR(p)$ ACF postupně klesá (monotónně nebo oscilačně), zatímco PACF je nenulová pouze do lagu p a poté ostře klesá k nule. Pro proces klouzavého průměru $MA(q)$ je obrazec zrcadlově obrácený. ACF vykazuje nenulové hodnoty do lagu q a poté klesá k nule, zatímco PACF postupně klesá. U smíšeného modelu $ARMA(p, q)$ obě funkce klesají postupně, ACF od lagu q a PACF od lagu p . Pro ACF a PACF jsou definována pásma statistické významnosti $\pm 1,96/\sqrt{T}$ kolem nuly, která odpovídají 95 % intervalu spolehlivosti pro nulovou autokorelaci. Hodnoty přesahující toto pásmo jsou statisticky významné. (Enders, 2014)

4.4.3 Auto.arima

Řády modelů jsou v obou přístupech modelování sezónnosti voleny automaticky pomocí algoritmu Hyndman a Khandakar (2008). Tento algoritmus je implementován ve funkci `auto.arima()` balíčku `forecast` v programu R. Sezónní diference D se volí podle míry sezónní síly odvozené ze STL dekompozice. Pokud tato míra přesáhne hodnotu 0,64, algoritmus sezónní diferenci zařadí, v opačném případě ji vynechá. Následně se pro takto (případně) diferencovanou řadu volí nesezónní diference d postupnými KPSS testy. V každém kroku se testuje, zda je řada stacionární. Pokud ne, řada se diferencuje a test se provede znovu, dokud řada není stacionární. Poté se volí zbývající řády p, q, P, Q pomocí postupné (stepwise) procedury. Ve stochastickém přístupu sezónní diferenci D explicitně fixujeme na hodnotu 1, abychom zajistili konzistentní zařazení sezónní diference napříč všemi okny a zeměmi. STL kritérium D v hraničních případech vynechává, což by vedlo k modelům bez sezónní složky i tam, kde je sezónnost v datech přítomná. Stejně tak totiž sezónní dummy proměnné zahrnujeme ve všech případech.

V prvním kroku se odhadnou čtyři výchozí modely. V sezónním přístupu se jedná o modely $ARIMA(2,d,2)(1,D,1)_{12}$, $ARIMA(0,d,0)(0,D,0)_{12}$, $ARIMA(1,d,0)(1,D,0)_{12}$ a $ARIMA(0,d,1)(0,D,1)_{12}$. V přístupu s dummy proměnnými se odhadnou jejich nesezónní varianty, jelikož je sezónnost zachycena 11 měsíčními dummy proměnnými zahrnutými jako exogenní

regresory. Z těchto čtyř se jako aktuální vybere model s nejnižší hodnotou zvoleného informačního kritéria. Ve druhém kroku se zkouší všechny variace aktuálního modelu, kde se některý z řádů změní o ± 1 . Pokud má některá variace nižší hodnotu kritéria, stane se novým aktuálním modelem a druhý krok se opakuje. Procedura končí, když žádná variace nevede ke zlepšení.

Postupná procedura neprohledává celý prostor modelů, a proto nemusí nalézt model s globálně nejnižším informačním kritériem. Tato volba je učiněna s ohledem na velký počet iterací v rámci rozšiřujícího se okna, kde se `auto.arima()` spouští pro každé okno zvlášť. Pro řady delší než 150 pozorování algoritmus standardně používá při výběru modelu aproximaci věrohodnostní funkce pomocí podmíněného součtu čtverců (CSS). Tuto aproximaci explicitně vypínáme, aby byl výběr řádů modelu založen na přesných hodnotách informačního kritéria.

4.4.4 Fixní specifikace

Kromě funkce `auto.arima`, která pro každé okno znovu volí řády modelu, testujeme také fixní specifikace. Jde tedy o ARIMA modely s pevně stanovenými řády, které zůstávají stejné napříč všemi predikčními okny. Jednoduchou specifikací v tomto kontextu rozumíme model s maximálním řádem 1 pro všechny AR a MA složky, včetně sezónních. Ověřujeme, zda některá jednoduchá fixní specifikace nedosáhne srovnatelné nebo lepší predikční přesnosti než adaptivní `auto.arima`. Jednoduché modely mohou totiž out-of-sample překonávat složitější specifikace, protože jsou robustnější vůči strukturálním změnám v datech a méně náchylné k přeučení (Makridakis a Hibon, 2000).

4.5 Konstrukce in-sample a out-of-sample

V predikční literatuře je standardem rozdělit data na in-sample a out-of-sample. In-sample predikce jsou generovány modelem odhadnutým na těchto datech, kterým se snaží co nejlépe přizpůsobit. Out-of-sample predikce naopak model generuje pro pozorování, která během tréninku neviděl, čímž se lépe testuje jeho skutečná predikční schopnost. (Tashman, 2000)

Predikce vyhodnocujeme na třech horizontech $h = 1, 3$ a 6 měsíců. Volba reflektuje praxi v literatuře o predikci průmyslové produkce. Články s predikcemi měsíčního IPP obvykle testují horizonty 1, 3, 6

a případně 12 měsíců dopředu. Horizonty 1, 3 a 6 jsme tedy zvolili v souladu s literaturou, horizont 12 měsíců jsme vynechali z důvodu přehlednosti.

V literatuře se in-sample predikce obvykle hodnotí na základě jednokrokových predikcí v rámci celého tréninkového vzorku. Abychom mohli porovnávat rozdíl mezi in-sample a out-of-sample predikcemi napříč různými horizonty i obdobími, volíme odlišný přístup. V něm srovnáváme stejná konkrétní cílová pozorování a sledujeme rozdíl mezi predikcemi modelu, který má k dispozici pouze data dostupná v daném okamžiku, a modelu, který je odhadnut i z pozorování mezi tímto okamžikem a cílovým pozorováním.

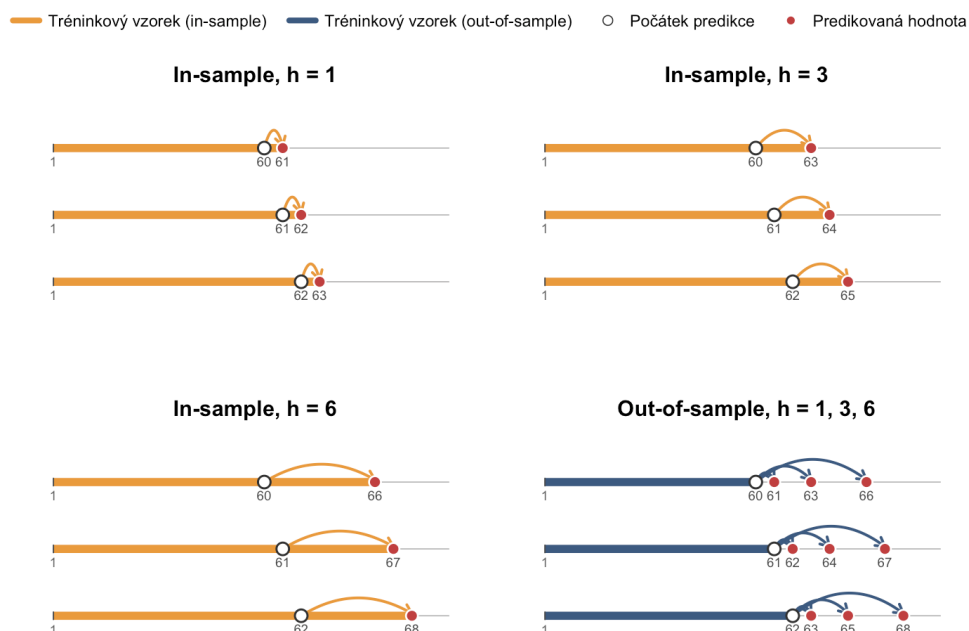
Out-of-sample model je odhadnut na vzorku bez cílového pozorování a bez všech pozorování mezi časem t a cílovým pozorováním. Pro horizont h tedy OOS nevidí h pozorování ($y_{t+1}, y_{t+2}, \dots, y_{t+h}$). To odpovídá situaci, kdy model v čase t generuje predikci budoucí hodnoty bez znalosti vývoje mezi časem t a cílovým pozorováním ($t+h$).

In-sample model je naopak odhadnut na rozšířeném vzorku, který těchto h pozorování zahrnuje, včetně cílového y_{t+h} . Pro každý horizont je tento rozšířený vzorek o jiný počet pozorování delší (o jedno pro $h = 1$, o tři pro $h = 3$ a o šest pro $h = 6$). Vlastní in-sample predikce hodnoty y_{t+h} je předpověď h kroků dopředu z pozice t , stejně jako u out-of-sample modelu. Samotný model (jeho specifikace i parametry) je přitom ale odhadnut na rozšířeném vzorku, tedy včetně hodnot y_{t+1} až y_{t+h} .

Tuto konstrukci in-sample a out-of-sample aplikujeme ve dvou různých režimech odhadování modelu, které se v literatuře běžně používají. V režimu rozšiřujícího se okna se s každou iterací přidá do tréninkového vzorku jedno nové pozorování a model je znovu odhadnut (velikost okna tedy postupně roste). V režimu posuvného okna naopak zůstává velikost tréninkového vzorku konstantní. S každou iterací se přidá nové pozorování a zároveň je odebráno nejstarší, takže se okno postupně posouvá časem (Rossi, 2013).

Konstrukci in-sample a out-of-sample predikcí pro rozšiřující se okno ilustruje Obr. 3, který zachycuje konstrukci predikcí pro první tři iterace. V každé z nich máme čtyři různé tréninkové vzorky: jeden pro out-of-sample, společný pro všechny horizonty, a tři pro in-sample (jeden pro každý horizont). In-sample vzorky se od out-of-sample liší o h pozorování přidaných napravo (cílová hodnota y_{t+h} a případná mezilehlá pozorování). Pro horizont $h = 1$ je tak in-sample vzorek o jedno pozorování delší, pro $h = 6$ o šest.

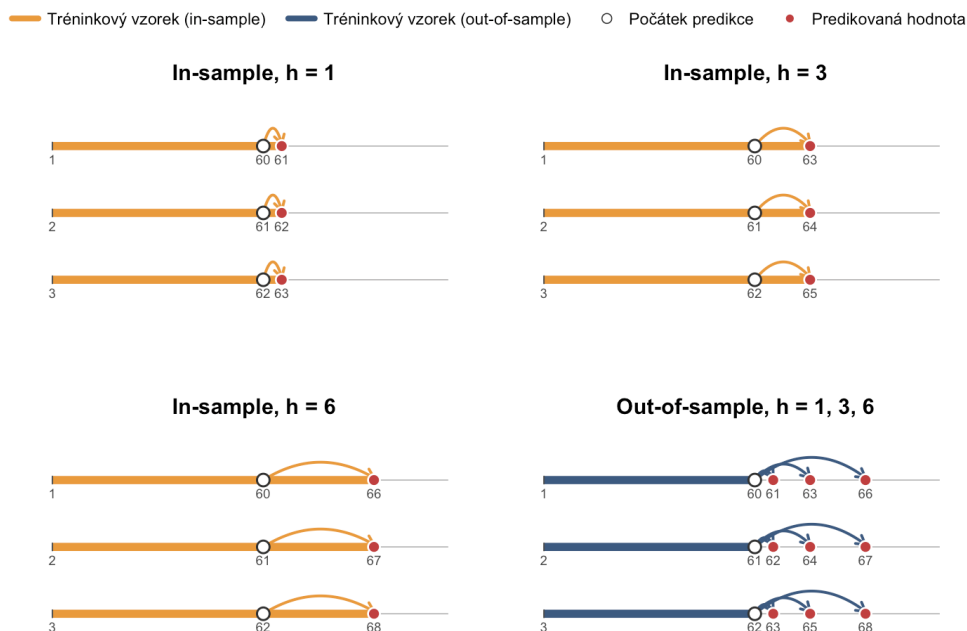
Obr. 3: Konstrukce in-sample a out-of-sample predikcí pro rozšiřující se okno



Zdroj: vlastní zpracování

Analogickou konstrukci pro režim posuvného okna ilustruje Obr. 4. Obrazek slouží ke znázornění in-sample a out-of-sample modelů a nereflektuje volbu velikosti posuvného okna, která je předmětem citlivostní analýzy.

Obr. 4: Konstrukce in-sample a out-of-sample predikcí pro posuvné okno



Zdroj: vlastní zpracování

4.6 Hodnocení predikcí

V této sekci popisujeme metriky pro hodnocení predikční přesnosti, dekompozici rozdílu mezi in-sample a out-of-sample predikcemi a rozdělení vzorku na analyzovaná období.

4.6.1 Metriky pro hodnocení predikcí

Pro hodnocení predikční přesnosti modelů jsou v literatuře standardně používány dvě metriky: odmocnina ze střední kvadratické chyby (RMSE) a střední absolutní chyba (MAE). Pro vzorek n pozorování y_i , $i = 1, 2, \dots, n$, a odpovídajících predikcí $\hat{y}_i$ jsou RMSE a MAE definovány jako:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (7)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (8)$$

Obě metriky vyjadřují průměrnou chybu ve stejných jednotkách jako data. Rozdíl mezi nimi spočívá v tom, že RMSE díky kvadratickému penalizování klade větší váhu na výrazné chyby, zatímco MAE váží všechny chyby lineárně. Modely s několika výraznými chybami jsou proto podle RMSE penalizovány silněji než podle MAE. Hodson (2022) srovnává obě metriky a dochází k závěru, že žádnou z nich nelze označit za univerzálně lepší. Je proto vhodné model hodnotit na základě obou.

4.6.2 Dekompozice rozdílu in-sample a out-of-sample predikcí

Rozdíl mezi in-sample a out-of-sample u auto.arima v práci rozkládáme na dva efekty. Prvním je parametrický efekt, tedy rozdíl mezi parametry odhadnutými na OOS a IS vzorku při stejné specifikaci modelu. Druhým je efekt výběru řádů, neboť auto.arima může na rozšířeném IS vzorku zvolit odlišné řády než na OOS vzorku. U modelu s fixní specifikací odpadá efekt výběru řádů (řády jsou pevně dané) a IS-OOS rozdíl je tak dán pouze parametrickým efektem.

Pro oddělení obou efektů zavádíme mezikrok IS s OOS řády (IS*). V tomto kroku zafixujeme specifikaci modelu na tu, kterou auto.arima vybrala na OOS vzorku (stejně řády (p, d, q), případně (P, D, Q), a stejnou indikaci driftu), a parametry odhadneme na rozšířeném IS vzorku. Predikci y_{t+h} získáme z tohoto modelu jako předpověď h kroků dopředu z pozice t, stejným způsobem jako u IS predikce.

Parametrický efekt pak vyjadřujeme jako rozdíl $\text{RMSE}(\text{OOS}) - \text{RMSE}(\text{IS}^*)$ a efekt výběru řádů jako rozdíl $\text{RMSE}(\text{IS}^*) - \text{RMSE}(\text{IS})$. Součet obou efektů odpovídá celkovému IS-OOS rozdílu $\text{RMSE}(\text{OOS}) - \text{RMSE}(\text{IS})$.

4.6.3 Rozdělení na období

Vedle celkových hodnot RMSE a MAE vyhodnocujeme predikční přesnost také v pěti obdobích, která jsou oddělena krizovými propady IPP. Hranice obou šokových období stanovujeme na základě meziročních změn IPP. Pro přehlednost je rozdělení zobrazeno v Tab. 3.

Finanční krizi vymezujeme od října 2008 do prosince 2009. Začátek krize stanovujeme podle české řady, kde se v říjnu 2008 meziroční změna IPP poprvé propadá do výrazně záporných hodnot (−9,7 %), což zároveň odpovídá měsíci po kolapsu Lehman Brothers. Slovenská řada stejnou interpretaci neumožňuje kvůli výjimečně silnému předkrizovému růstu IPP, který udržuje meziroční změnu v roce 2008 v kladných hodnotách navzdory reálnému propadu produkce. Společnou hranici říjen 2008 proto zakládáme na české řadě. Konec krize stanovujeme na prosinec 2009, kdy se meziroční změna poprvé současně vrací do kladných hodnot v obou zemích. Kritériem konce krize tedy není dno recese, ale návrat k růstu.

Pandemii COVID-19 vymezujeme březnem až srpnem 2020. V únoru 2020 byly meziroční změny IPP v obou zemích blízko standardních úrovní, v březnu se v důsledku lockdownů zavedených v polovině měsíce propadly do výrazného záporu (ČR −9,0 %, SK −18,7 %). Během léta došlo k rychlé obnově. V září 2020 se meziroční změny vracejí na hodnoty blízké předpandemickému stavu a od tohoto měsíce řadu zařazujeme do období po pandemii.

Počet predikcí n je pro jednotlivá období shodný napříč horizonty s výjimkou období před krizí, kde je n pro delší horizonty nižší ($n = 45$ pro $h = 1$, $n = 43$ pro $h = 3$ a $n = 40$ pro $h = 6$), neboť při delším horizontu predikce začínají až od pozdějšího pozorování. Stejně počty predikcí platí ve všech dalších tabulkách s analýzou po obdobích.

Tab. 3: Vymezení analyzovaných období

Období	n	Od	Do
Před krizí	45	01/2000	09/2008
Finanční krize	15	10/2008	12/2009
Mezikrizové	122	01/2010	02/2020
COVID	6	03/2020	08/2020
Po COVIDu	40	09/2020	12/2023

Zdroj: vlastní zpracování

5 Srovnání predikcí při rozšiřujícím se okně

V této kapitole prezentujeme predikce získané v rámci rozšiřujícího se okna. Počáteční velikost tréninkového vzorku je 60 měsíců a s každým krokem se zvětšuje o jeden měsíc. Predikce tedy pokrývají období od ledna 2005 do prosince 2023. Sezónnost modelujeme dvěma způsoby. Prvním je deterministický přístup s dummy proměnnými a druhým je stochastický přístup zahrnující sezónní složky (SARIMA model). Dále srovnáváme dvě varianty volby řádů modelu. V první podkapitole se zabýváme modely zvolenými pomocí funkce `auto.arima`, ve druhé modely s fixními řády. V textu uvádíme hodnoty RMSE pro porovnání predikční přesnosti. Analyzovali jsme však i hodnoty MAE, které uvádíme v Příloze B. Závěry se napříč metrikami nemění.

5.1 Modely `auto.arima`

Při tvorbě predikcí pomocí tohoto přístupu využíváme funkci `auto.arima` s výběrem řádů modelu na základě Bayesovského informačního kritéria. V této podkapitole se zabýváme srovnáním predikčních chyb napříč horizonty a obdobími. Rovněž analyzujeme, jakou část rozdílu mezi in-sample a out-of-sample predikcemi způsobuje změna parametrů a jakou změna řádů modelu.

5.1.1 Predikční výsledky

Tab. 4 prezentuje srovnání in-sample a out-of-sample RMSE pro Česko a Slovensko při modelování sezónnosti pomocí dummy proměnných. U obou zemí platí očekávaný růst RMSE se zvětšujícím se horizontem predikce. Celkové out-of-sample RMSE se u Česka zvyšuje z 4,77 při $h = 1$ na 6,45 při $h = 6$ a u Slovenska z 6,48 na 9,04. Slovenské predikce mají vyšší hodnoty RMSE, což odpovídá vyšší volatilitě slovenského IPP identifikované v deskriptivní analýze. Relativní rozdíl mezi in-sample a out-of-sample predikcemi se s horizontem výrazně nemění a u většiny případů se pohybuje v rozpětí 13–16 %. Výjimkou je $h = 1$ u Slovenska, kde dosahuje 24,5 %. Tento výsledek je způsoben především obdobím před krizí, kde in-sample RMSE je 4,07, zatímco out-of-sample RMSE je téměř dvojnásobné (7,96). Důvodem je skokový nárůst IPP Slovenska na začátku roku 2008, na který out-of-sample predikce nedokážou reagovat.

Přesnost predikcí se zároveň liší mezi jednotlivými obdobími. V mezikrizovém období dosahují modely nejnižších predikčních chyb, konkrétně se out-of-sample RMSE pohybuje mezi 3,81 a 4,29 u Česka a mezi 5,33 a 5,75 u Slovenska. V období před finanční krizí jsou chyby srovnatelně nízké pro Česko, ovšem u Slovenska jsou výrazně vyšší, což souvisí s odlišným vývojem slovenského IPP.

V období finanční krize dochází k výraznému nárůstu predikčních chyb (primárně na delších horizontech). Například u horizontu $h = 6$ je out-of-sample RMSE téměř třikrát větší než v mezikrizovém období. V období krize COVID jsou predikční chyby ještě vyšší, jelikož se jednalo o krátký výkyv a model se nedokázal přizpůsobit. Celkově se jedná o předpokládané výsledky, neboť krizová období představují strukturní zlomy, které model odhadnutý na předkrizových datech nemůže predikovat.

IS-OOS rozdíl se rovněž liší podle období. Nejmenší je v mezikrizovém období, kde u obou zemí napříč horizonty nepřesahuje 11 %. Ve finanční krizi se tento rozdíl zvětšuje. U Česka dosahuje 44 až 53 % a u Slovenska 18 až 24 %. V krizi COVID pozorujeme, že rozdíl klesá s rostoucím horizontem.

V klidných obdobích se parametry modelu na in-sample a out-of-sample vzorku liší jen málo a in-sample predikce je tak dobrým ukazatelem out-of-sample výkonu. V krizových obdobích je tomu naopak. Out-of-sample model nezachytí změnu dynamiky, zatímco in-sample model má výhodu, že vidí krizová data.

Tab. 4: RMSE po obdobích – auto.arima, deterministická sezónnost, rozšiřující se okno

Období	n	h=1		h=3		h=6	
		IS	OOS	IS	OOS	IS	OOS
Česko							
Před krizí	45	2,73	3,46	2,90	3,42	3,48	4,20
Fin. krize	15	3,54	5,14	4,48	6,85	8,37	12,05
Mezikrizové	122	3,44	3,81	3,48	3,84	3,87	4,29
COVID	6	12,76	14,75	15,85	17,54	17,00	18,43
Po COVIDu	40	5,18	5,42	5,59	5,82	6,66	7,28
Celkem	228	4,20	4,77	4,66	5,24	5,55	6,45
Slovensko							
Před krizí	45	4,07	7,96	6,87	8,63	9,36	10,88
Fin. krize	15	4,79	5,92	8,75	10,32	12,58	15,24
Mezikrizové	122	4,91	5,33	5,41	5,75	5,09	5,37
COVID	6	14,23	16,82	20,50	22,80	19,49	20,00
Po COVIDu	40	4,82	5,06	6,46	8,35	8,51	10,11
Celkem	228	5,20	6,48	6,96	8,08	7,98	9,04

Zdroj: vlastní zpracování

Tab. 5 zobrazuje výsledky pro modely se stochastickou sezónností, kde auto.arima vybírá model včetně sezónní složky. Celkové výsledky jsou srovnatelné s deterministickým přístupem. U Česka se celkové in-sample i out-of-sample RMSE mezi oběma přístupy liší maximálně o 2 %. U Slovenska jsou rozdíly větší, zejména v out-of-sample při $h = 6$ (10,81 oproti 9,04 u modelu s dummy proměnnými).

Výraznější rozdíly vidíme při srovnání jednotlivých období. V mezikrizovém období je SARIMA přesnější u obou zemí. U Česka při $h = 1$ dosahuje OOS RMSE 3,33 oproti 3,81 u modelu s dummy proměnnými, zatímco u Slovenska je výhoda ještě výraznější (např. při $h = 1$ hodnota 4,29 oproti 5,33 a na $h = 3$ dokonce 4,34 oproti 5,75). SARIMA tedy v klidném období zachycuje sezónní vzorec lépe než fixní dummy proměnné. Naopak v krizových obdobích je deterministický přístup výrazně lepší.

Nejvýraznější rozdíl vidíme u Slovenska ve finanční krizi při $h = 6$, kde deterministický přístup dosahuje OOS RMSE 15,24 oproti 28,22 u SARIMA. V období COVID je deterministický přístup lepší především na kratších horizontech. U Česka dosahuje OOS RMSE 14,75 při $h = 1$ oproti

17,26 u stochastického přístupu, kdežto pro Slovensko pozorujeme 22,80 při $h = 3$ oproti 25,84. Při $h = 6$ se rozdíly u obou zemí vyrovnávají.

Tyto výsledky naznačují, že modely se stochastickou sezónností jsou výhodnější v klidných obdobích, zatímco v obdobích strukturálních zlomů dosahují lepších výsledků modely s deterministickou sezónností.

Tab. 5: RMSE po obdobích – auto.arima, stochastická sezónnost, rozšiřující se okno

Období	n	h=1		h=3		h=6	
		IS	OOS	IS	OOS	IS	OOS
Česko							
Před krizí	45	3,14	3,53	3,36	3,76	3,93	4,21
Fin. krize	15	4,26	5,65	5,81	7,36	9,96	12,12
Mezikrizové	122	3,08	3,33	3,20	3,29	3,79	4,04
COVID	6	15,14	17,26	18,17	20,51	17,95	18,56
Po COVIDu	40	4,44	5,21	5,25	5,46	5,69	8,13
Celkem	228	4,21	4,81	4,85	5,35	5,64	6,57
Slovensko							
Před krizí	45	4,30	6,33	6,80	7,80	9,51	10,21
Fin. krize	15	5,61	10,93	10,95	18,58	14,95	28,22
Mezikrizové	122	4,13	4,29	4,00	4,34	4,53	5,08
COVID	6	16,48	17,77	23,93	25,84	18,43	19,28
Po COVIDu	40	4,58	4,74	7,01	7,36	10,21	10,70
Celkem	228	5,07	6,15	7,02	8,48	8,38	10,81

Zdroj: vlastní zpracování

Tab. 6 prezentuje podíl predikcí, v nichž jsou out-of-sample predikce přesnější než in-sample. U deterministického přístupu jsou in-sample predikce téměř vždy přesnější. Česko při $h = 1$ nemá jediný případ, kdy by OOS predikce překonala IS predikci, a i při delších horizontech zůstává podíl nízký (13,9 % při $h = 6$). U Slovenska jsou out-of-sample predikce přesnější v ještě menším počtu případů (maximálně 6,3 % u $h = 6$). V případě stochastické sezónnosti je obraz odlišný. Již při $h = 1$ je OOS přesnější v 10,5 % případů pro Česko a při $h = 3$ dokonce ve 35,8 %. Tento výsledek může souviset s vyšší složitostí modelů se sezónní

ARIMA složkou, která vede k větší variabilitě odhadů parametrů a vyššímu riziku přeučení na in-sample datech.

Tab. 6: Podíl predikcí, u nichž out-of-sample přesnější než in-sample – rozšiřující se okno

	h = 1	h = 3	h = 6
ČR deterministická	0,0 %	7,5 %	13,9 %
ČR stochastická	10,5 %	35,8 %	35,4 %
SK deterministická	0,4 %	4,4 %	6,3 %
SK stochastická	3,5 %	19,0 %	20,2 %

Zdroj: vlastní zpracování

Tab. 7 shrnuje nejčastěji vybrané modely funkcí auto.arima v rámci rozšiřujícího se okna pro out-of-sample predikce. U deterministické sezónnosti vidíme, že pro Česko je výběr modelu rozptýlenější než pro Slovensko. Nejčastější model (2,1,0) pokrývá necelých 39 % oken a nejčastější 3 modely dohromady 81 %. U Slovenska naopak výrazně dominuje model (0,1,1), který je vybrán v 62 % oken, přičemž 3 nejčastější modely pokrývají 95 % případů.

U stochastické sezónnosti je výběr modelu pro Česko méně stabilní než u deterministické varianty. Nejčastější model pokrývá jen 25,4 % oken a nejčastější 3 modely dohromady 43 %. U Slovenska naopak dominuje specifikace (1,0,1)(0,1,1) s podílem 40,8 % oken, přičemž první tři modely pokrývají 69 %. Zároveň je patrné, že stochastická sezónnost vede u Česka k výrazně složitějším modelům s vyšším počtem parametrů (například (3,0,0)(2,1,2) oproti (2,1,0) u deterministické sezónnosti). Slovenské modely zůstávají poměrně jednoduché i při volbě stochastické sezónnosti. Obě země preferují odlišné modely. U deterministické sezónnosti Česka jsou vybrány AR složky, zatímco Slovensko preferuje MA složky. Tento rozdíl naznačuje odlišnou dynamiku průmyslové produkce obou zemí.

Podíl oken, v nichž auto.arima zahrnuje drift, se výrazně liší mezi přístupy i zeměmi. U deterministické sezónnosti volí drift u Česka 40,8 % oken a u Slovenska jen 18,0 %. Pro Česko tedy dlouhodobý lineární trend přispívá modelu výrazně častěji. U stochastické sezónnosti podíl driftu pro obě země roste (ČR 51,8 %, SK 55,7 %). U Slovenska to znamená trojnásobný nárůst oproti deterministické variantě.

Tab. 7: Nejčastěji vybrané auto.arima modely a volba driftu – rozšiřující se okno

Pořadí	ČR model	ČR po- díl	SK model	SK po- díl
<i>Deterministická</i>				
1.	(2,1,0)	38,6 %	(0,1,1)	61,8 %
2.	(3,1,0)	22,8 %	(0,1,2)	20,6 %
3.	(0,1,1)	19,3 %	(0,1,0)	12,3 %
Drift		40,8 %		18,0 %
<i>Stochastická</i>				
1.	(3,0,0)(2,1,2)	25,4 %	(1,0,1)(0,1,1)	40,8 %
2.	(3,0,0)(0,1,1)	8,8 %	(1,0,0)(0,1,1)	15,8 %
3.	(4,0,0)(0,1,1)	8,8 %	(0,1,1)(0,1,0)	12,7 %
Drift		51,8 %		55,7 %

Zdroj: vlastní zpracování

5.1.2 Dekompozice rozdílu in-sample a out-of-sample predikcí

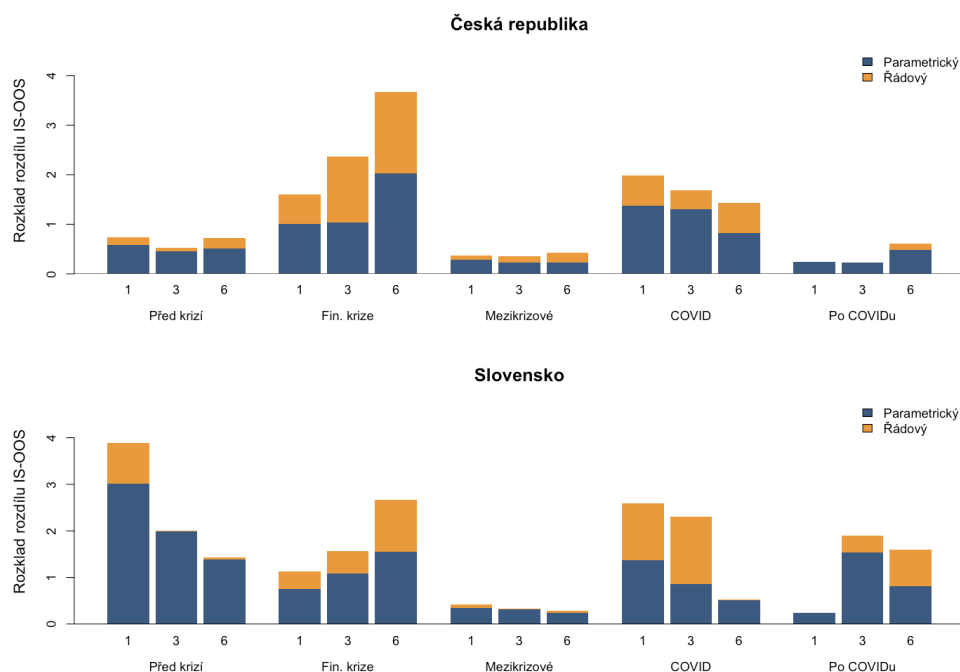
Rozdíl mezi in-sample a out-of-sample predikcemi při výběru modelu pomocí funkce auto.arima má dvě příčiny. První je parametrický efekt, tedy model při in-sample odhaduje jiné hodnoty parametrů než out-of-sample model, jelikož vidí dodatečná pozorování. Druhou je řádový efekt, tedy in-sample model může volit jiné řády modelu na základě dodatečných pozorování. V této podkapitole se zabýváme oběma efekty zvlášť. Jejich součet dává dohromady rozdíl mezi in-sample a out-of-sample RMSE.

Na Obr. 5 je pro jednotlivá období zobrazena dekompozice pro deterministickou sezónnost. U České republiky dominuje ve všech obdobích změna parametrů. Řádový efekt je zejména po COVIDu na kratších horizontech prakticky nulový, v mezikrizovém období tvoří přibližně čtvrtinu až polovinu celkového rozdílu, ale v absolutních hodnotách zůstává malý (nepřesahuje 0,2). V krizových obdobích podíl řádového efektu roste a na delších horizontech tvoří 30 až 55 % celkového rozdílu. Během krizí se mění dynamika vývoje, tudíž dodatečná pozorování vzorku vedou k volbě odlišných řádů modelu.

U Slovenska pozorujeme obdobný vzorec. Zvláštní pozornost si zaslouží období před krizí, kde celkový IS-OOS rozdíl klesá s rostoucím horizontem, což je opačný průběh než v ostatních obdobích. Tento pokles vysvětluje parametrický efekt, který za ním stojí. Skokový nárůst

slovenského IPP na začátku roku 2008 je součástí in-sample vzorku a výrazně mění odhady parametrů oproti out-of-sample modelu trénovanému ještě na datech před neobvyklým růstem. S rostoucím horizontem se vliv tohoto šoku v out-of-sample parametrech postupně vyrovnává, a IS-OOS rozdíl se tak zmenšuje.

Obr. 5: Dekompozice rozdílu IS-OOS – deterministická sezónnost, rozšiřující se okno



Poznámka: Čísla pod sloupci (1, 3 a 6) označují horizonty predikce. Kladné hodnoty rozdílu znamenají, že IS predikce jsou přesnější než OOS, záporné naopak.

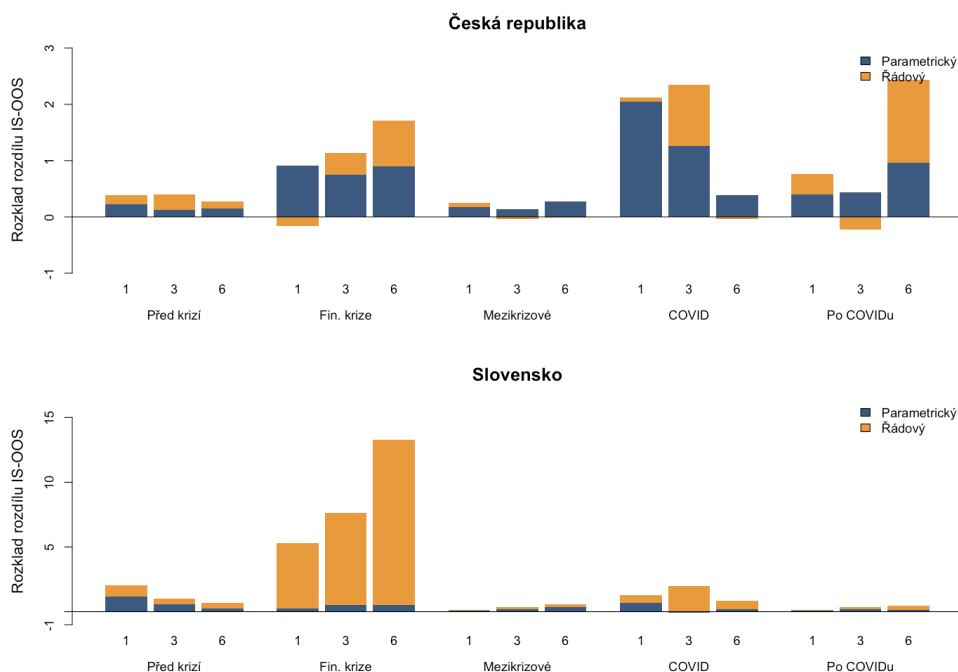
Zdroj: vlastní zpracování

Na Obr. 6 je dekompozice provedena pro stochastickou sezónnost. Oproti deterministickému přístupu se výsledky mění ve dvou směrech. Zaprvé je IS-OOS rozdíl ve finanční krizi výrazně větší. Zadruhé se v některých obdobích prosazuje řádový efekt, který u deterministického přístupu zůstával sekundární. U Slovenska dosahuje ve finanční krizi na $h = 6$ hodnoty 13,27, což je téměř pětinasobek oproti deterministickému přístupu (2,67). Tento rozdíl je navíc téměř zcela tvořen řádovým efektem (na $h = 6$ činí řádový efekt RMSE 12,77, tedy 96 % celkového rozdílu, zatímco parametrický efekt tvoří jen RMSE 0,50). Auto.arima tedy na in-

sample vzorku obsahujícím krizová pozorování volí velmi odlišný řád než na out-of-sample vzorku, kterému část krizových pozorování chybí, a parametrické rozdíly jsou oproti volbě řádu zanedbatelné.

V období COVID pozorujeme u stochastické sezónnosti nekonzistenci napříč horizonty. Při $h = 1$ dominuje parametrický efekt. Při $h = 3$ význam řadového efektu roste a u Slovenska dokonce převažuje nad parametrickým. Při $h = 6$ jsou celkové IS-OOS rozdíly u obou zemí malé ve srovnání s kratšími horizonty. Tento pokles celkového rozdílu s rostoucím horizontem vysvětluje úbytek parametrického efektu. U Česka klesá z 2,04 při $h = 1$ na 0,39 při $h = 6$, což odráží postupné vyrovnávání odhadů parametrů v delších horizontech, kdy predikce závisejí více na dlouhodobé úrovni než na krátkodobé dynamice.

Obr. 6: Dekompozice rozdílu IS-OOS – stochastická sezónnost, rozšiřující se okno



Poznámka: Čísla pod sloupci (1, 3 a 6) označují horizonty predikce. Kladné hodnoty rozdílu znamenají, že IS predikce jsou přesnější než OOS, záporné naopak.

Zdroj: vlastní zpracování

5.2 Fixní modely

V předchozí sekci jsme porovnávali predikce pro modely `auto.arima`. To jsou modely, u kterých očekáváme nejlepší in-sample predikce, jelikož je řád modelu pro každé okno zvolen znovu. Navíc na IS-OOS rozdíl se podílí jak změna parametrů, tak změna řádů modelu. Otázkou je, zda existuje model s fixními řády, který by dokázal lépe predikovat out-of-sample. V rámci takového modelu bychom mohli přímo porovnat rozdíly mezi in-sample a out-of-sample predikcemi způsobené pouze změnou parametrů.

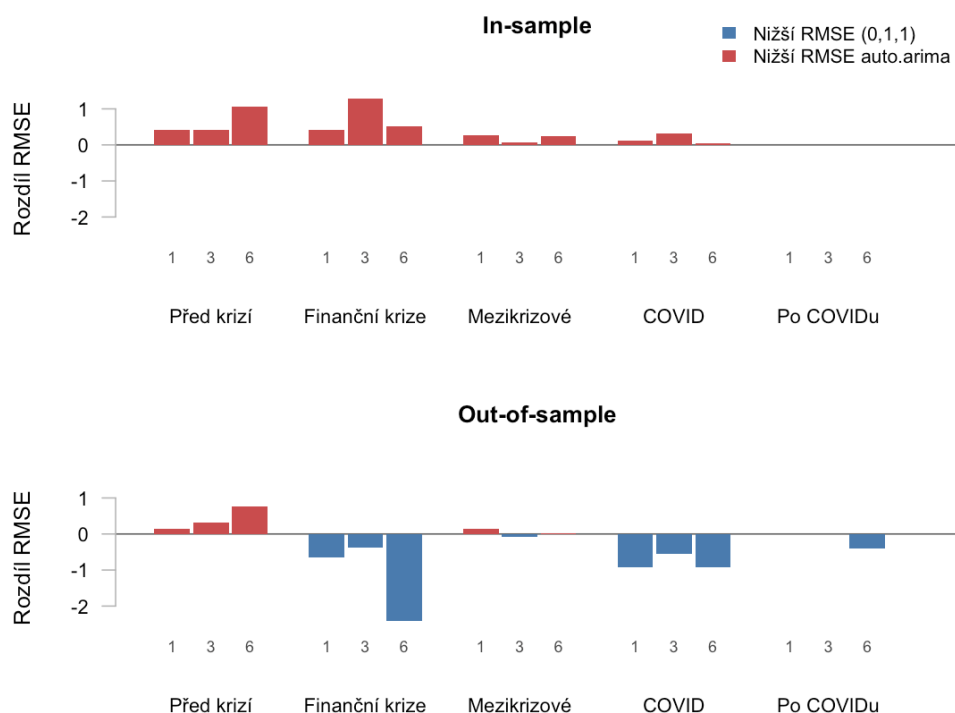
Pro modely s deterministickou sezónností jsme testovali jednoduché specifikace s nejvýše jednou AR nebo MA složkou: $ARIMA(1,1,0)$ a $ARIMA(0,1,1)$. Model s $MA(1)$ složkou měl výrazně lepší predikce než model s $AR(1)$ složkou. To souhlasí s Tab. 7, v níž jsou zobrazeny nejčastěji vybrané modely funkcí `auto.arima`. Ve třech nejčastěji vybraných modelech se $ARIMA(0,1,1)$ objevuje u obou zemí.

Otázkou pro fixní model bylo, zda přidat drift. Pro obě země jsme testovali obě varianty a celkové OOS RMSE bylo nižší u modelu bez driftu. Z tohoto důvodu jsme zvolili fixní model bez driftu. Výsledný fixní model pro deterministickou sezónnost je tedy $ARIMA(0,1,1)$ s dummy proměnnými a bez driftu.

Na Obr. 7 je znázorněn rozdíl mezi RMSE modelu s fixními řády a modelu `auto.arima` pro Českou republiku. Kladná hodnota značí vyšší RMSE modelu s fixními řády, a tedy `auto.arima` dosahuje přesnějších predikcí. `Auto.arima` má přesnější in-sample predikce napříč horizonty i obdobími, jelikož přizpůsobuje řády modelu konkrétním tréninkovým datům. V období před krizí jsou predikce `auto.arima` přesnější i v out-of-sample, což je způsobeno hlavně tím, že `auto.arima` pro toto období přidává drift (v 91 % oken), kdežto fixní model je bez driftu. V mezikrizovém období jsou rozdíly mezi modely minimální. V krizových obdobích je naopak fixní model výrazně přesnější. V období finanční krize dosahuje nižšího OOS RMSE na všech horizontech. Na $h = 6$ je rozdíl dokonce 2,42. V období COVIDu je fixní model rovněž přesnější (RMSE nižší o 0,5 až 0,9). V období po COVIDu vidíme jediný rozdíl na horizontu 6. Tento rozdíl je způsoben predikcí pro jediné pozorování (říjen 2020), které je druhým pozorováním po COVIDu. Tréninkový vzorek pro tuto predikci končí v dubnu 2020, tedy na dně krizového propadu. `Auto.arima` nevolí jednoduchý model $ARIMA(0,1,1)$ jako v ostatních případech, ale $ARIMA(0,1,3)$, jelikož je ovlivněna propadem. Ostatní rozdíly v tomto

období jsou nulové, jelikož `auto.arima` ve všech dalších oknech volí právě model $ARIMA(0,1,1)$.

Obr. 7: Rozdíl RMSE mezi fixním a `auto.arima` modelem – ČR, deterministická sezónnost, rozšiřující se okno



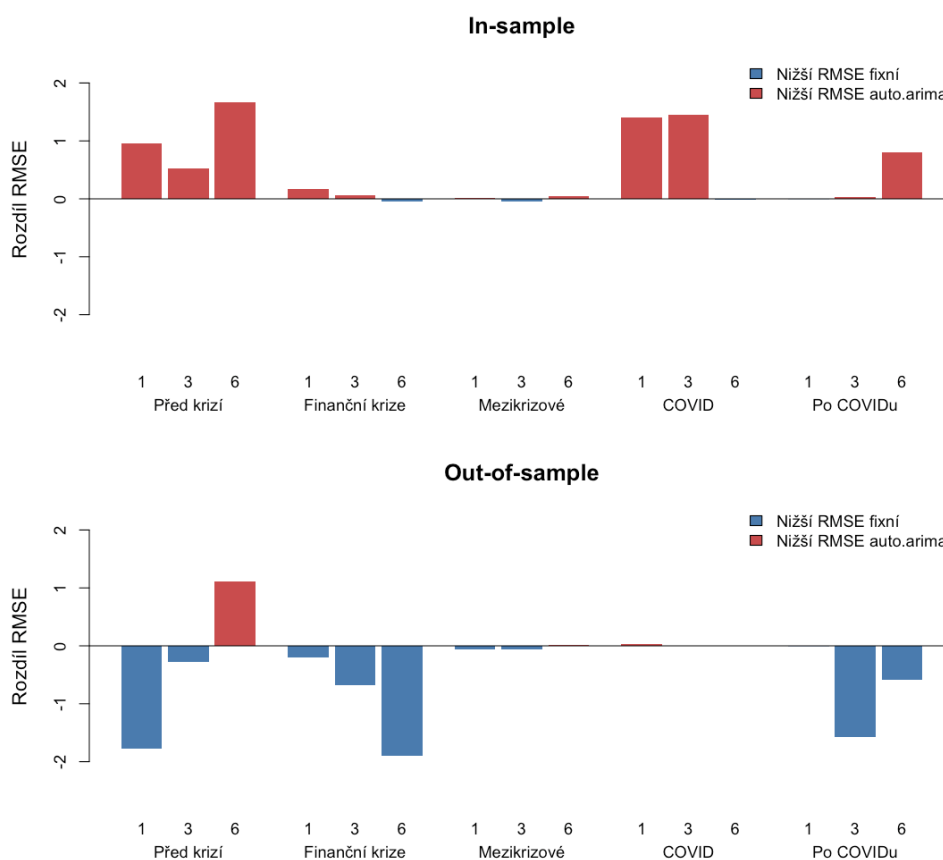
Poznámka: Čísla pod sloupci (1, 3 a 6) označují horizonty predikce.

Zdroj: vlastní zpracování

Obr. 8 zobrazuje stejné srovnání pro Slovensko. Celkový obraz je podobný, ovšem existují dílčí rozdíly. In-sample predikce jsou opět ve většině případů lepší u modelu `auto.arima`. Na rozdíl od Česka však v období finanční krize a mezikrizovém období jsou rozdíly mezi in-sample predikcemi obou modelů menší a fixní model dokonce ve čtyřech případech dosahuje lepších predikcí (Fin. krize $h = 6$, mezikrizové $h = 3$, COVID $h = 6$, Po COVIDu $h = 1$, ale všechny rozdíly pod 0,05). Pro out-of-sample predikce vidíme rozdíl v období před krizí, kdy fixní model na kratších horizontech dosahuje přesnějších predikcí, ale na horizontu 6 naopak lépe predikuje `auto.arima`. V období finanční krize je fixní model lepší

napříč horizonty, přičemž na $h = 6$ je rozdíl v RMSE 1,90 ve prospěch fixního modelu. Stejně tak v období po COVIDu dosahuje fixní model nižšího OOS RMSE, a to zejména na $h = 3$ (rozdíl 1,58) a $h = 6$ (0,59). Pro out-of-sample predikce v rámci krize COVID je rozdíl prakticky nulový na rozdíl od ČR, kde fixní model překonával auto.arima napříč horizonty.

Obr. 8: Rozdíl RMSE mezi fixním a auto.arima modelem – SK, deterministická sezónnost, rozšiřující se okno



Poznámka: Čísla pod sloupci (1, 3 a 6) označují horizonty predikce.

Zdroj: vlastní zpracování

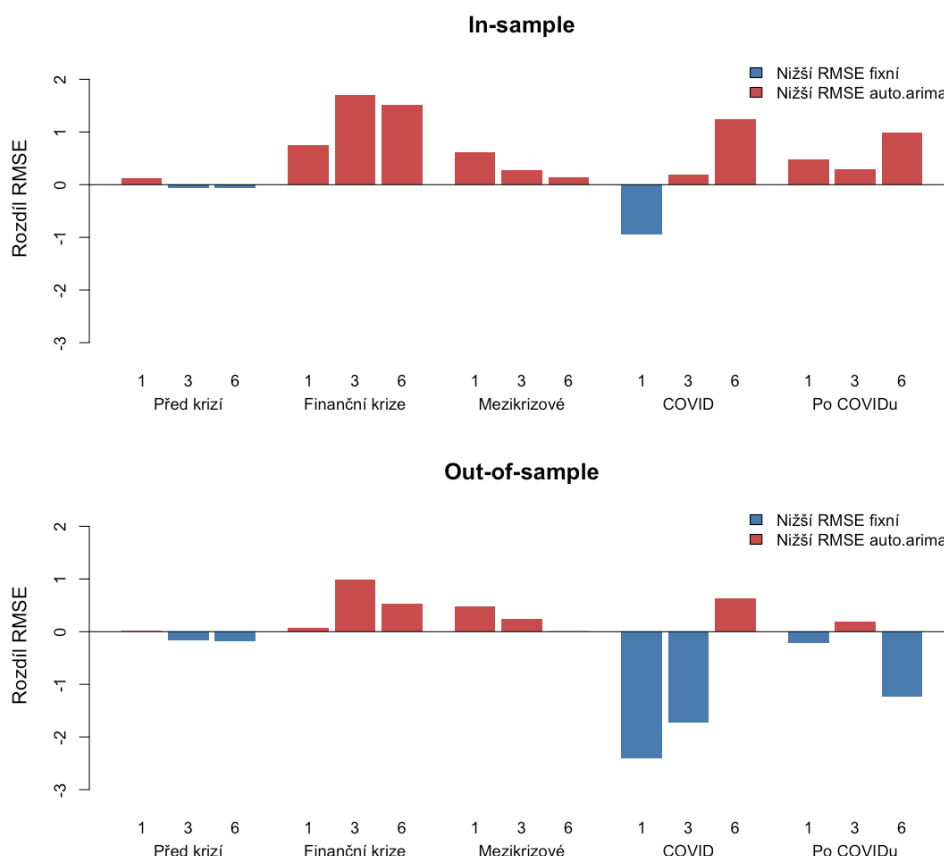
Výběr jednoduchého SARIMA modelu je méně přímočarý než u deterministického přístupu. Zatímco u deterministické sezónnosti stačilo srovnat dvě specifikace s AR(1) nebo MA(1) složkou, u SARIMA přibývají sezónní složky (P, D, Q) a volba mezi nesezónní diferencí, sezónní diferencí, nebo jejich kombinací. Proto jsme pro stochastickou sezónnost testovali

33 SARIMA specifikací s omezením $d + D \in \{1, 2\}$, každou složkou AR a MA v $\{0, 1\}$ a maximálně dvěma nenulovými AR a MA řády.

Nejnižšího out-of-sample RMSE dosahoval napříč zeměmi a horizonty model $(0,1,1)(0,1,1)$. Tento model je v literatuře o modelování měsíční průmyslové produkce používaným benchmarkem. Thury a Witt (1998) jej srovnávají s dalšími modely pro Rakousko a Německo. Hassani, Heravi a Zhigljavsky (2009) jej podle BIC vybírají jako optimální ARIMA specifikaci u 5 z 8 britských sérií. Drift u této specifikace nezahrnujeme, neboť již v modelu máme nesezónní i sezónní diferenci.

Na Obr. 9 je znázorněn rozdíl mezi RMSE fixního modelu $SARIMA(0,1,1)(0,1,1)$ a modelu `auto.arima` pro Českou republiku. In-sample predikce jsou opět přesnější u `auto.arima` s výjimkou období před krizí, kde jsou rozdíly minimální, a období COVIDu, kde je na $h = 1$ naopak fixní model přesnější o 0,94. V out-of-sample predikcích vidíme odlišnosti oproti deterministické sezónnosti. V období před krizí jsou rozdíly zanedbatelné a v mezikrizovém období `auto.arima` dosahuje mírně přesnějších predikcí. Ve finanční krizi dokonce `auto.arima` konzistentně poráží fixní model (+0,99 na $h = 3$), což je opačný vzorec než u deterministického přístupu. Výrazná převaha fixního modelu se však projevuje v období pandemie COVID, kdy na kratších horizontech dosahuje podstatně přesnějších predikcí (-2,41 na $h = 1$ a -1,73 na $h = 3$). `Auto.arima` se v těchto oknech snaží adaptovat na propad složitějšími specifikacemi, což vede k horším krátkodobým predikcím. Na $h = 6$ má naopak `auto.arima` o 0,64 nižší RMSE. V období po pandemii je fixní model přesnější na horizontu $h = 6$ (-1,24).

Obr. 9: Rozdíl RMSE mezi fixním a auto.arima modelem – ČR, stochastická sezónnost, rozšiřující se okno



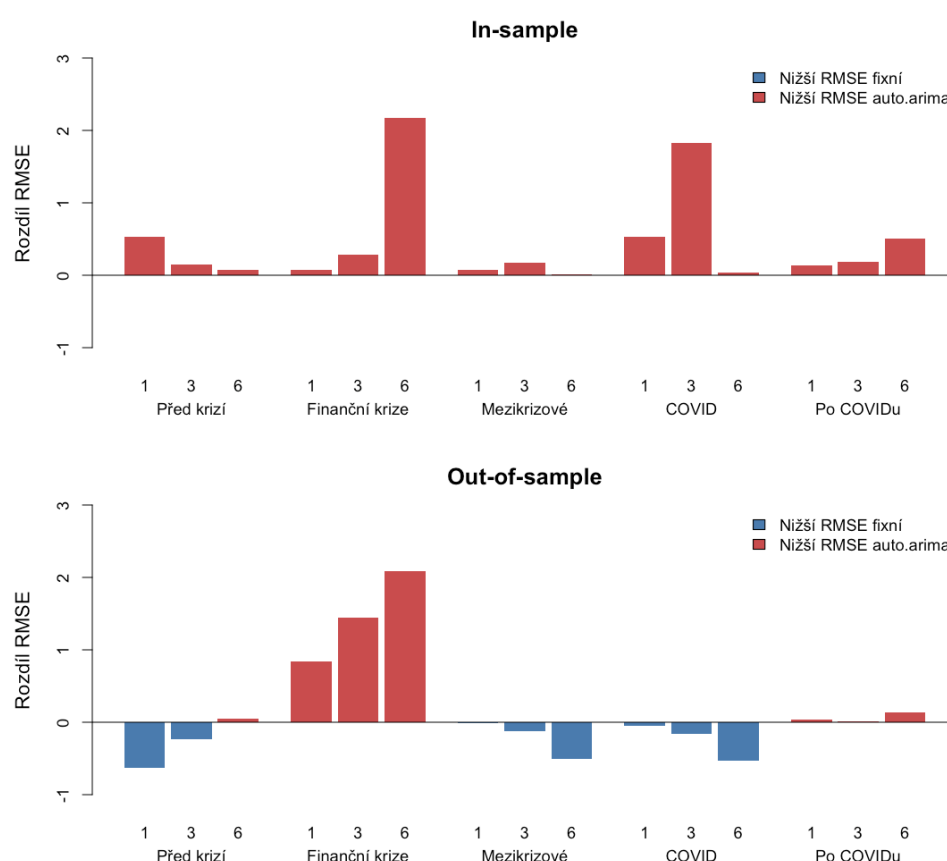
Poznámka: Čísla pod sloupci (1, 3 a 6) označují horizonty predikce.

Zdroj: vlastní zpracování

Obr. 10 zobrazuje stejné srovnání pro Slovensko. In-sample predikce jsou napříč všemi obdobími přesnější u auto.arima, přičemž největší rozdíly jsou ve finanční krizi na $h = 6$ (+2,17) a u COVIDu na $h = 3$ (+1,83). V out-of-sample predikcích auto.arima vítězí pouze v období finanční krize (+0,84 na $h = 1$, +1,44 na $h = 3$, +2,09 na $h = 6$). Tento vzorec je opačný než u deterministické varianty SK, kde v krizi vítězil fixní ARIMA(0,1,1)+dummy. Fixní SARIMA(0,1,1)(0,1,1) s dvojnásobnou diferencí v krizi trpí nadměrným rozdílováním, zatímco auto.arima má větší flexibilitu při výběru řádu. V období před krizí je na krátkých horizontech přesnější fixní model (-0,63 na $h = 1$, -0,23 na $h = 3$), v mezikrizovém období dosahuje fixní model nižšího OOS RMSE zejména na $h = 6$ (-0,51),

v období COVIDu jsou rozdíly malé a převážně ve prospěch fixního modelu ($-0,05$ až $-0,53$). V období po pandemii jsou rozdíly prakticky nulové.

Obr. 10: Rozdíl RMSE mezi fixním a auto.arima modelem – SK, stochastická sezónnost, rozšiřující se okno



Poznámka: Čísla pod sloupce (1, 3 a 6) označují horizonty predikce.

Zdroj: vlastní zpracování

Tab. 8 popisuje IS a OOS RMSE pro fixní model ARIMA(0,1,1)+dummy po jednotlivých obdobích. Rozdíl mezi IS a OOS v této tabulce odpovídá parametrickému efektu, tedy změně odhadu parametrů po přidání dodatečných pozorování do tréninkového vzorku.

V klidných obdobích je parametrický efekt malý. U deterministického modelu se u České republiky rozdíl mezi IS a OOS RMSE pohybuje mezi 0,19 a 0,24, u Slovenska mezi 0,21 a 0,35. Dodatečná pozorování

v těchto obdobích neposkytují výrazně odlišné informace o dynamice řady, a parametry se tedy odhadují podobně jako z kratšího vzorku.

V krizových obdobích se parametrický efekt znatelně zvyšuje. Pro Českou republiku dosahuje u deterministického modelu 0,54 až 0,74 ve finanční krizi a 0,48 až 0,94 v období COVIDu. U Slovenska je efekt silnější a ve finanční krizi dosahuje až 0,82 na $h = 6$, v období COVIDu až 1,22 na $h = 1$. Tato zvýšená citlivost odpovídá tomu, že šoková pozorování v tréninkovém vzorku výrazně mění odhady MA parametru a sezónních dummy proměnných.

Specifický vzorec pozorujeme u IS-OOS rozdílu v období COVID na přích horizonty. Zatímco v ostatních obdobích je parametrický efekt na přích horizonty přibližně konstantní nebo mírně klesá, v COVIDu je naopak nejvyšší při $h = 1$ a klesá s horizontem (u Česka z 0,94 na 0,48). Příčinou je, že pro delší horizonty končí tréninkový vzorek OOS modelu ještě před pandemií, a OOS je tak bez šokových pozorování. Přestože IS model obsahuje krizová pozorování, jejich malý počet (6 pozorování na vzorku obsahujícím přes 240 pozorování) vede k podobným odhadům parametrů jako u OOS. Parametrický efekt zůstává nízký navzdory nárůstu RMSE.

Tab. 8: RMSE po obdobích – ARIMA (0,1,1)+dummy, rozšiřující se okno

Období	n	h = 1		h = 3		h = 6	
		IS	OOS	IS	OOS	IS	OOS
Česko							
Před krizí	45	3,14	3,60	3,31	3,73	4,54	4,98
Fin. krize	15	3,96	4,50	5,77	6,48	8,88	9,62
Mezikrizové	122	3,72	3,96	3,56	3,76	4,12	4,31
COVID	6	12,87	13,81	16,18	16,98	17,04	17,52
Po COVIDu	40	5,18	5,42	5,59	5,82	6,66	6,88
Celkem	228	4,41	4,74	4,86	5,17	5,84	6,13
Slovensko							
Před krizí	45	5,03	6,19	7,39	8,36	11,03	12,00
Fin. krize	15	4,95	5,73	8,81	9,64	12,53	13,35
Mezikrizové	122	4,92	5,27	5,37	5,69	5,14	5,38
COVID	6	15,62	16,84	21,94	22,80	19,48	20,00
Po COVIDu	40	4,81	5,05	6,49	6,78	9,31	9,52
Celkem	228	5,48	6,05	7,17	7,67	8,52	8,98

Zdroj: vlastní zpracování

Tab. 9 ukazuje stejný detail pro model $(0,1,1)(0,1,1)$. Ten vykazuje oproti deterministické variantě odlišnosti hlavně u Slovenska. V mezikrizovém období je pro obě země srovnatelný nebo mírně lepší. U Slovenska dosahuje OOS RMSE 4,27 oproti 5,27 u deterministického modelu na $h = 1$. Zásadní slabinou je však jeho chování ve finanční krizi slovenské řady, kde OOS RMSE prudce roste až na 30,31 na $h = 6$, tedy více než dvojnásobek deterministické varianty (13,35). Tento extrémní výsledek dokumentuje, že stochastický přístup s dvojitou diferenciací u krátkých krizových oken selhává, protože kombinace nesezónní a sezónní difference nedokáže dostatečně rychle absorbovat strukturní šok. Deterministická sezónnost je v tomto ohledu robustnější, jelikož dummy proměnné zachycují sezónní hladinu stabilně bez ohledu na krizové odchylky.

U České republiky je parametrický efekt u $(0,1,1)(0,1,1)$ celkově menší než u deterministické varianty (například v mezikrizovém období 0,11 oproti 0,24 na $h = 1$). To naznačuje, že dvojitá diferenciacie u české řady vede ke stabilnějším odhadům parametrů v klidných obdobích, i když celkové OOS RMSE je u obou modelů podobné.

Tab. 9: RMSE po obdobích – SARIMA $(0,1,1)(0,1,1)$, rozšiřující se okno

Období	n	h = 1		h = 3		h = 6	
		IS	OOS	IS	OOS	IS	OOS
Česko							
Před krizí	45	3,27	3,54	3,30	3,60	3,86	4,02
Fin. krize	15	5,01	5,73	7,51	8,35	11,48	12,64
Mezikrizové	122	3,70	3,81	3,48	3,54	3,93	4,05
COVID	6	14,19	14,85	18,36	18,78	19,19	19,20
Po COVIDu	40	4,92	5,00	5,55	5,65	6,69	6,89
Celkem	228	4,54	4,75	5,17	5,37	6,16	6,41
Slovensko							
Před krizí	45	4,83	5,70	6,94	7,57	9,58	10,26
Fin. krize	15	5,68	11,76	11,23	20,02	17,12	30,31
Mezikrizové	122	4,20	4,27	4,17	4,21	4,55	4,57
COVID	6	17,01	17,72	25,76	25,68	18,46	18,75
Po COVIDu	40	4,71	4,78	7,20	7,38	10,72	10,83
Celkem	228	5,26	6,13	7,33	8,62	8,79	11,07

Zdroj: vlastní zpracování

5.3 Shrnutí pro rozšiřující se okno

Predikční přesnost klesá s rostoucím horizontem pro obě země i oba přístupy k sezónnosti. Slovenské predikce jsou konzistentně méně přesné než české, což reflektuje vyšší volatilitu slovenského IPP.

Nejlepších OOS predikcí obě země dosahují v klidných obdobích (mezikrizovém, po COVIDu a u Česka také před krizí), kde modely fungují podle očekávání na stabilních datech. Nejhorší výsledky přicházejí v obou krizových obdobích, zejména v COVIDu, kde krátké trvání pandemie v kombinaci se strukturálním šokem znemožnilo modelům adekvátně reagovat. Výjimkou je slovenská řada se stochastickou sezónností, kde nejhorších výsledků model dosahuje ve finanční krizi na $h = 6$ (OOS RMSE 28,22), protože kombinace nesezónní a sezónní difference nedokáže absorbovat prudký propad. V období před krizí se obě země liší. U Česka jsou predikce přesné, u Slovenska výrazně horší kvůli skokovému nárůstu IPP v roce 2008.

Dva přístupy k sezónnosti dávají odlišné výsledky. V klidných obdobích je stochastická sezónnost mírně přesnější, protože flexibilní sezónní složka lépe zachytí postupně se měnící sezónní vzorec. V krizových obdobích je naopak deterministický přístup výrazně robustnější, protože dummy proměnné zachycují sezónní hladinu stabilně bez ohledu na krizové odchylky.

In-sample predikce jsou u obou zemí téměř vždy přesnější než out-of-sample. Rozdíl mezi nimi má dva zdroje. Parametrický efekt dominuje v klidných obdobích, kde řádový efekt je zanedbatelný. V krizových obdobích roste význam řádového efektu, který se stává druhou významnou složkou celkového rozdílu. V období COVID u stochastické sezónnosti klesá parametrický efekt s rostoucím horizontem (u Česka z 2,04 na $h = 1$ na 0,39 na $h = 6$), což vede k překvapivě malému celkovému IS-OOS rozdílu na $h = 6$ navzdory vysoké úrovni RMSE.

Fixní modely tyto závěry doplňují a v některých případech zlepšují. Auto.arima má díky flexibilnímu výběru řádu zpravidla lepší IS predikce, ale v OOS ji ve většině období překonávají fixní modely. Deterministický fixní model ARIMA(0,1,1) s dummy proměnnými je lepší nebo srovnatelný s auto.arima u obou zemí téměř ve všech obdobích, zejména ve finanční krizi a v pandemii COVID. Hlavní výjimkou je období před krizí u Česka, kde auto.arima díky zahrnutí driftu predikuje přesněji. Stochastický fixní model SARIMA(0,1,1)(0,1,1) je naopak méně konzistentní. U Slovenska je v klidných obdobích lepší než auto.arima, u Česka jsou

výsledky smíšené. Ve finanční krizi slovenské řady fixní model selhává, protože dvojitá diferenciace není schopna absorbovat strukturní šok. Výsledky potvrzují závěr, že deterministická sezónnost je v krizových obdobích spolehlivější volbou.

6 Srovnání predikcí při posuvném okně

V předchozí kapitole jsme prezentovali predikce v rámci rozšiřujícího se okna, kde se tréninkový vzorek s každým krokem rozšiřuje o jeden měsíc. Alternativním přístupem je posuvné okno s pevnou velikostí tréninkového vzorku, které se s každým krokem posouvá vpřed a pracuje pouze s nejnovějšími pozorováními. V této kapitole nejprve určujeme optimální velikost posuvného okna pomocí citlivostní analýzy a následně srovnáváme jeho predikce s rozšiřujícím se oknem pro `auto.arima` i fixní modely.

6.1 Citlivostní analýza

Volba velikosti posuvného okna je zásadní metodologické rozhodnutí. Příliš malé okno vede k nestabilnímu odhadu parametrů s vysokou variancí predikcí, naopak příliš velké okno se svým chováním přibližuje rozšiřujícímu se oknu a ztrácí výhodu rychlejší adaptace na aktuální data. V této podkapitole testujeme délky posuvného okna od 36 do 120 měsíců pro obě země a oba přístupy k sezónnosti. Všechny hodnoty RMSE jsou počítány na společných pozorováních od ledna 2010, aby byly srovnatelné napříč velikostmi okna.

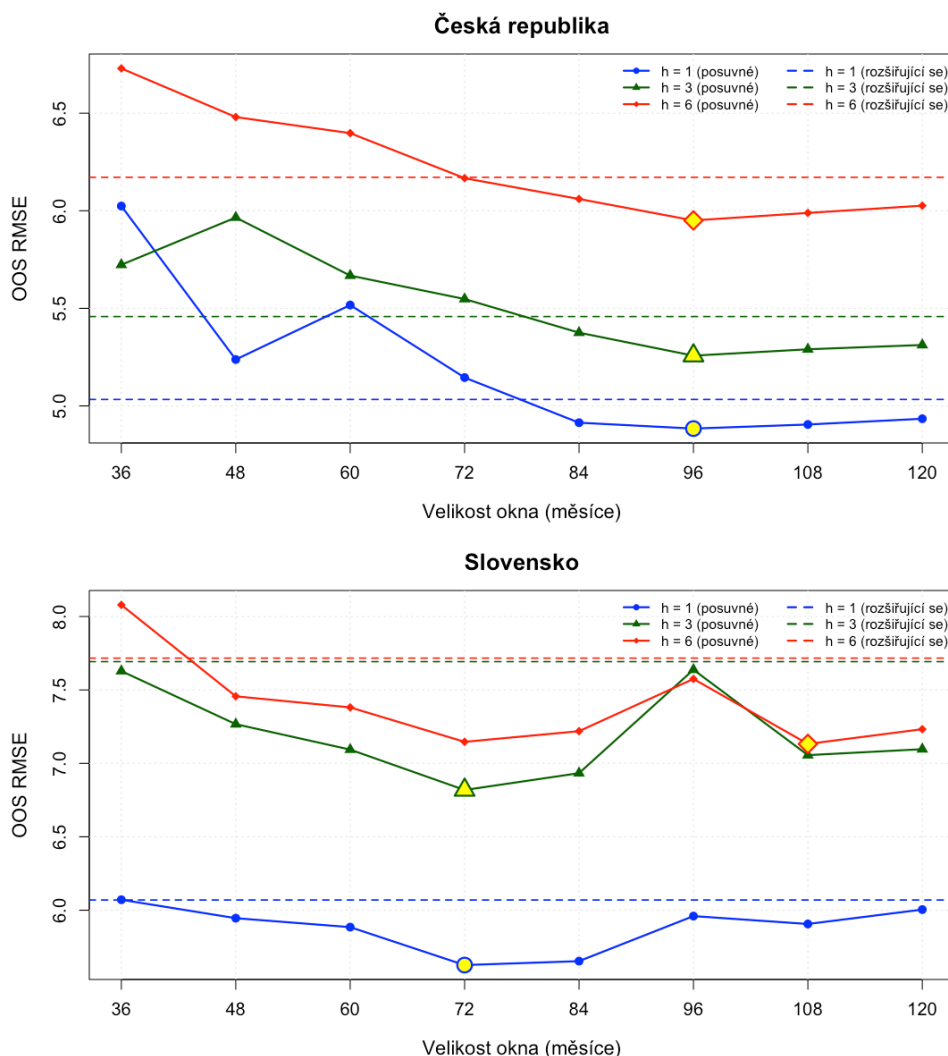
Pro volbu optimální délky posuvného okna se zaměřujeme na out-of-sample RMSE. In-sample RMSE se u obou přístupů chová opačně. U deterministického přístupu se s rostoucí velikostí posuvného okna zvyšuje (sezónní dummy proměnné lépe popisují kratší vzorek), zatímco u stochastického klesá (SARIMA potřebuje delší vzorek pro stabilní odhad sezónních složek). Grafy pro citlivostní analýzu IS RMSE uvádíme v Příloze C.

Na Obr. 11 je znázorněna citlivostní analýza pro deterministický přístup modelování sezónnosti. V grafu vidíme celkové out-of-sample RMSE v závislosti na délce okna. V grafu jsou rovněž přerušovanými čarami znázorněny RMSE na jednotlivých horizontech pro rozšiřující se okno. Můžeme tak srovnat, zda má na společných predikcích posuvné okno nižší predikční chyby než rozšiřující se okno.

Pro Českou republiku je patrné, že optimální délkou posuvného okna je 96, protože na této délce je RMSE minimální pro všechny horizonty. Pro Slovensko není shoda na všech horizontech, nicméně se jako nejlepší jeví délka okna 72. Při této délce je RMSE nejmenší při $h = 1$

a $h = 3$. RMSE pro $h = 6$ je nejnižší při délce okna 108, ale toto zlepšení je malé oproti oknu o délce 72 (7,13 oproti 7,15). Jako optimální délku posuvného okna tedy volíme 96 pro Českou republiku a 72 pro Slovensko.

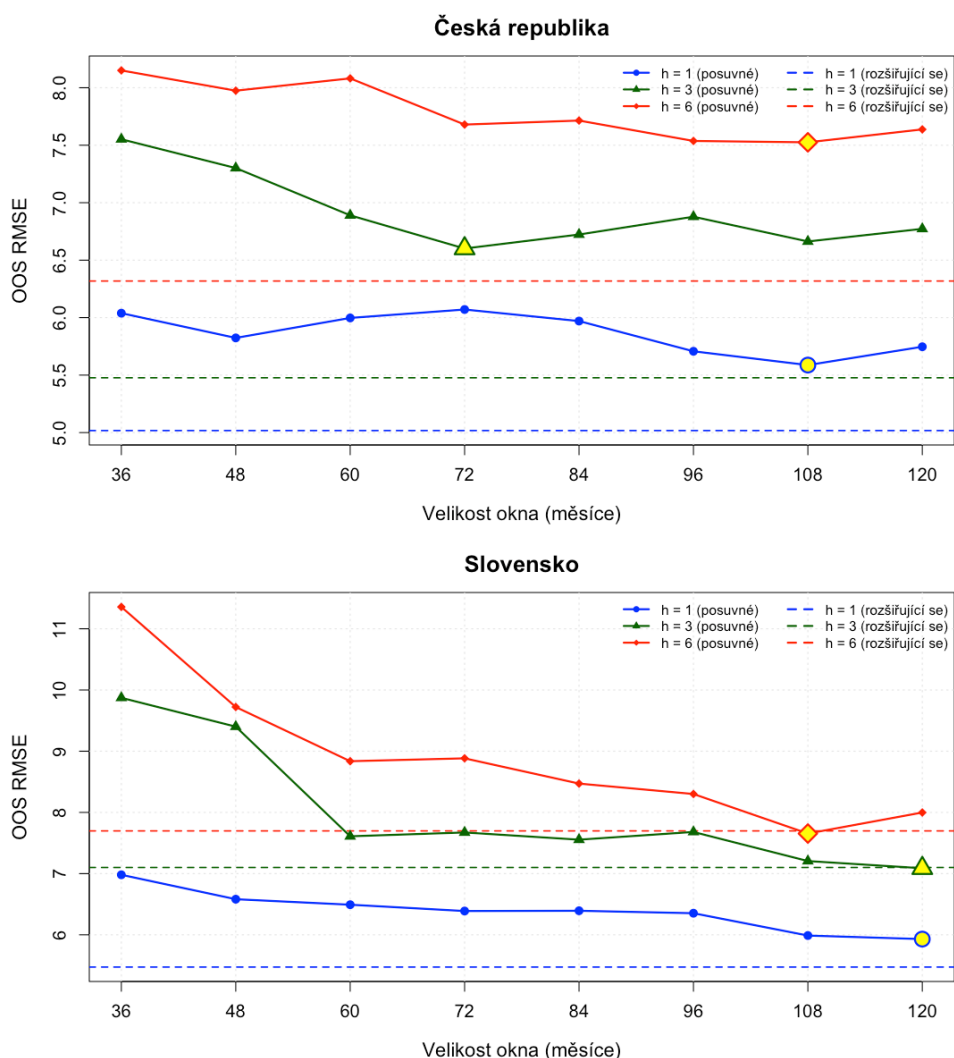
Obr. 11: Citlivostní analýza pro out-of-sample RMSE – auto.arima, deterministická sezónnost



Poznámka: Plné čáry zobrazují posuvná okna o různých délkách, přerušované čáry zobrazují referenční hodnoty pro rozšiřující se okno. Všechny hodnoty jsou vypočteny na společných pozorováních (od ledna 2010). Žlutě zvýrazněné body označují minimum pro daný horizont.

Zdroj: vlastní zpracování

Obr. 12: Citlivostní analýza pro out-of-sample RMSE – auto.arima, stochastická sezónnost



Poznámka: Plné čáry zobrazují posuvná okna o různých délkách, přerušované čáry zobrazují referenční hodnoty pro rozšiřující se okno. Všechny hodnoty jsou vypočteny na společných pozorováních (od ledna 2010). Žlutě zvýrazněné body označují minimum pro daný horizont.

Zdroj: vlastní zpracování

Na Obr. 12 je zobrazen vývoj celkového out-of-sample RMSE v závislosti na velikosti posuvného okna pro model se stochastickou sezónností. Na rozdíl od modelů s deterministickou sezónností nevidíme u modelů se stochastickou sezónností výrazné zlepšení oproti rozšiřujícímu se oknu. Pro Českou republiku je nejnižší RMSE pro každý horizont na jiné délce

okna, nicméně se neblíží RMSE pro rozšiřující se okno. U Slovenska je situace odlišná. Posuvné okno dosahuje u horizontu 3 a 6 při vhodné délce okna mírně nižšího RMSE než rozšiřující se okno, ale zlepšení je zanedbatelné (do 0,05). U horizontu 1 zůstává rozšiřující se okno přesnější oproti všem délkám posuvného okna.

Z těchto poznatků docházíme k závěru, že pro SARIMA modely je nejvhodnější metodou rozšiřující se okno a dále se jim v této kapitole nebudeme věnovat.

6.2 Srovnání predikcí auto.arima mezi posuvným a rozšiřujícím se oknem

Tab. 10 ukazuje rozdíly mezi rozšiřujícím se a posuvným oknem po jednotlivých obdobích pro IS i OOS predikce. Kladné hodnoty znamenají, že posuvné okno má nižší RMSE, záporné hodnoty, že je lepší rozšiřující se okno.

Celkově je u obou zemí posuvné okno přesnější než rozšiřující se okno. U Česka je výhoda mírná (OOS o 0,18 až 0,34 RMSE), u Slovenska výrazně větší (o 0,46 až 0,82). Vzhledem k vyšší úrovni RMSE (přes 6 u celkového OOS) jsou tyto rozdíly v relativním vyjádření malé, ale u Slovenska jsou viditelnější než u Česka.

Při pohledu na jednotlivá období je posuvné okno přesnější především v mezikrizovém období a po COVIDu. U Slovenska je výhoda posuvného okna v mezikrizovém období výraznější (až +1,41 při $h = 3$) než u Česka (+0,25 při $h = 1$). Slovenské predikce tedy více profitují z toho, že posuvné okno zapomíná starší data, která už nereflektují aktuální dynamiku.

Výjimku představuje období COVID, kde u obou zemí při $h = 1$ dosahuje lepších výsledků rozšiřující se okno. Nejmarkantnější rozdíl je u Slovenska (-3,81 při $h = 1$). Rozšiřující se okno obsahuje ve svém tréninkovém vzorku finanční krizi 2008, což může vést k lepším odhadům při novém šoku. Posuvné okno naopak pokrývá pouze klidné mezikrizové období, a proto na šok reaguje méně přesně. Tento výsledek je ale třeba interpretovat opatrně kvůli malému počtu pozorování ($n = 6$).

IS predikce vykazují podobný vzorec jako OOS. Posuvné okno je obecně přesnější i v IS (u Česka o 0,21 až 0,48, u Slovenska o 0,20 až 1,28) a rozdíly jsou přitom často ještě větší než u OOS, zejména u Slovenska v mezikrizovém období (IS rozdíl +1,73 oproti +1,36 u OOS).

Srovnání IS a OOS rozdílů odhaluje odlišnou povahu obou přístupů v klidných a krizových obdobích. V klidných obdobích posuvné okno zlepšuje IS více než OOS, tedy má větší IS-OOS rozdíl než rozšiřující se okno. Kratší vzorek posuvného okna se lépe přizpůsobí konkrétním datům v IS, ale tato výhoda se jen zčásti přenáší do OOS predikcí. V krizových obdobích pozorujeme opačný vzorec. U Česka ve finanční krizi má rozšiřující se okno nižší IS RMSE ($-0,20$ až $-0,62$), ale vyšší OOS RMSE ($+0,41$ až $+1,12$). Rozšiřující se okno se tedy lépe přizpůsobuje krizovým datům, ale jeho delší tréninkový vzorek vede k zahrnutí krizových pozorování, která se v dalších oknech neopakují.

Tab. 10: Rozdíly RMSE mezi rozšiřujícím se a posuvným oknem po obdobích – auto.arima, deterministická sezónnost

Období	n	h=1		h=3		h=6	
		IS	OOS	IS	OOS	IS	OOS
Česko							
Fin. krize	15	−0,20	+0,52	−0,62	+0,41	−0,39	+1,12
Mezikrizové	122	+0,44	+0,25	+0,37	+0,25	+0,12	−0,05
COVID	6	+0,48	−1,17	+0,11	−0,22	+0,36	+0,37
Po COVIDu	40	+0,92	+0,58	+0,76	+0,41	+0,74	+0,77
Celkem	183	+0,48	+0,18	+0,31	+0,22	+0,21	+0,34
Slovensko							
Fin. krize	15	+0,54	+0,61	+0,06	+0,46	−0,03	+0,46
Mezikrizové	122	+1,73	+1,36	+1,72	+1,41	+0,88	+0,66
COVID	6	+0,87	−3,81	−0,44	−1,43	+1,59	+0,40
Po COVIDu	40	+0,94	+0,48	+0,51	+1,30	−1,10	+0,67
Celkem	183	+1,28	+0,46	+0,86	+0,82	+0,20	+0,54

Poznámka: Kladné hodnoty = posuvné okno má nižší RMSE. Záporné hodnoty = rozšiřující se okno má nižší RMSE. Česko $w = 96$, Slovensko $w = 72$. Společná pozorování.

Zdroj: vlastní zpracování

Tab. 11 znázorňuje, jaké modely auto.arima nejčastěji vybírala v rozšiřujícím se a posuvném okně na společných pozorováních pro out-of-sample predikce. U Česka jsou tři nejčastější modely stejné v obou přístupech, liší se ale jejich zastoupení. Rozšiřující se okno častěji vybírá AR modely, zejména (2,1,0) s 36,5 % oproti 24,5 % u posuvného okna. Posuvné okno naopak mírně více preferuje MA model (0,1,1). Výrazný je rozdíl ve

výběru driftu, který rozšiřující se okno na rozdíl od posuvného vybírá ve 31,8 % oken oproti 14,1 %.

U Slovenska jsou rozdíly mezi přístupy výraznější. Rozšiřující se okno preferuje model (0,1,1) s 59,7 %, zatímco posuvné okno nejčastěji vybírá (0,1,0) ve 47,2 % oken, tedy random walk se sezónními dummy proměnnými bez ARMA dynamiky. Na kratším vzorku o 72 pozorováních BIC častěji vyhodnotí, že přidané ARMA parametry nepřinášejí dostatečné zlepšení modelu. Zajímavý je také rozdíl u modelu (2,1,0), který rozšiřující se okno téměř nikdy nevybírá (0,5 %), zatímco posuvné okno jej volí v 16,2 % oken. Drift vykazuje oproti Česku opačný vzorec. Posuvné okno ho vybírá častěji (28,7 %) než rozšiřující se okno (15,7 %).

Tab. 11: Nejčastěji vybrané modely – posuvné a rozšiřující se okno

Model	Rozšiřující se okno	Posuvné okno
<i>Česko</i>		
(2,1,0)	36,5 %	24,5 %
(3,1,0)	27,1 %	17,2 %
(0,1,1)	22,4 %	22,9 %
Drift	31,8 %	14,1 %
<i>Slovensko</i>		
(0,1,1)	59,7 %	19,4 %
(0,1,2)	21,8 %	10,6 %
(0,1,0)	13,0 %	47,2 %
(2,1,0)	0,5 %	16,2 %
Drift	15,7 %	28,7 %

Zdroj: vlastní zpracování

6.3 Srovnání predikcí fixního modelu mezi posuvným a rozšiřujícím se oknem

Tab. 12 ukazuje rozdíly mezi posuvným a rozšiřujícím se oknem pro fixní model ARIMA(0,1,1) s dummy proměnnými. Protože fixní model má stejné řády v obou přístupech, rozdíly odrážejí pouze vliv odlišného tréninkového vzorku na odhad parametrů.

Posuvné okno je u fixního modelu celkově mírně přesnější v OOS (u Česka o 0,07 až 0,14, u Slovenska o 0,13 až 0,55). Výhoda se projevuje

především v mezikrizovém období, zejména u Slovenska (+1,28 až +1,45). Výhoda posuvného okna plyne z toho, že parametry jsou odhadnuté na aktuálnějších datech, která lépe odrážejí dynamiku klidného období. Ve finanční krizi jsou rozdíly minimální, jelikož se tréninkové vzorky obou přístupů z velké části překrývají.

V období COVID je naopak rozšiřující se okno na kratších horizontech výrazně přesnější (u Česka OOS -1,25 při $h = 1$, u Slovenska -1,91), což jsme viděli rovněž v Tab. 10.

Posuvné okno je v IS přesnější napříč všemi obdobími s výjimkou finanční krize, kde jsou rozdíly malé. Nejvýraznější je výhoda posuvného okna v mezikrizovém období, po COVIDu a zejména v COVIDu u Slovenska (+1,41 při $h = 6$). Kontrast s OOS je nejvíce zřetelný právě v COVIDu, kdy posuvné okno dosahuje nejnižšího IS RMSE, ale zároveň nejvyššího OOS RMSE.

Tab. 12: Rozdíly RMSE mezi rozšiřujícím se a posuvným oknem po obdobích – fixní model ARIMA(0,1,1)+dummy

Období	n	h=1		h=3		h=6	
		IS	OOS	IS	OOS	IS	OOS
Česko							
Fin. krize	15	+0,08	+0,03	−0,03	−0,09	+0,00	−0,06
Mezikrizové	122	+0,43	+0,24	+0,35	+0,20	+0,27	+0,14
COVID	6	+0,41	−1,25	+0,33	−0,98	+0,43	−0,35
Po COVIDu	40	+0,92	+0,58	+0,76	+0,41	+0,74	+0,39
Celkem	183	+0,48	+0,14	+0,36	+0,07	+0,33	+0,11
Slovensko							
Fin. krize	15	+0,26	−0,03	+0,19	−0,18	−0,03	−0,36
Mezikrizové	122	+1,55	+1,28	+1,68	+1,45	+0,85	+0,63
COVID	6	+1,28	−1,91	+1,11	−1,12	+1,41	−0,76
Po COVIDu	40	+0,89	+0,33	+0,27	−0,41	+0,34	−0,08
Celkem	183	+1,17	+0,55	+0,96	+0,44	+0,55	+0,13

Poznámka: Fixní model ARIMA(0,1,1) s dummy proměnnými bez driftu. Česko $w = 96$, Slovensko $w = 72$. Společná pozorování.

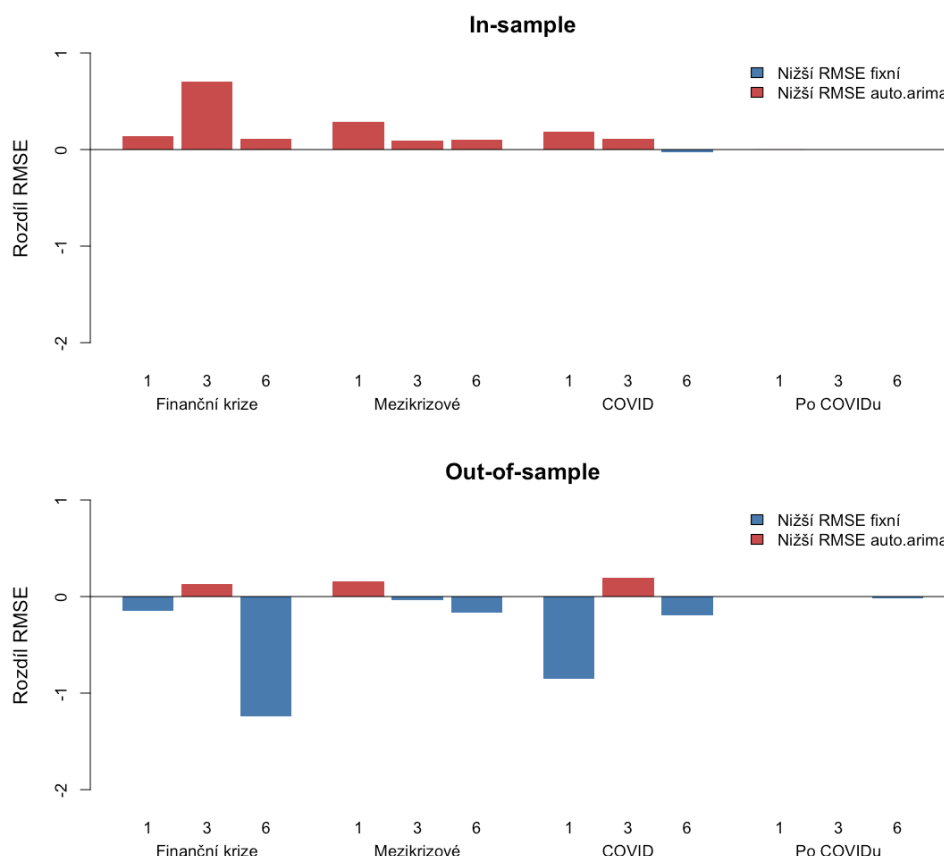
Zdroj: vlastní zpracování

6.4 Srovnání predikcí auto.arima a fixního modelu posuvného okna

V předchozí kapitole byly srovnávány fixní modely s auto.arima při použití rozšiřujícího se okna. Stejně srovnání provedeme také pro posuvné okno. Období Před krizí v analýze posuvného okna neuvádíme, jelikož posuvné okno má pro toto období malé množství pozorování.

Na Obr. 13 je znázorněn rozdíl mezi RMSE fixního modelu a auto.arima pro Českou republiku. Stejně jako u rozšiřujícího se okna jsou in-sample predikce přesnější u auto.arima téměř ve všech horizontech i obdobích, ačkoliv rozdíly jsou v případě posuvného okna menší. V out-of-sample predikcích se výhoda fixního modelu soustředí do krizových období. Ve finanční krizi při $h = 6$ dosahuje fixní model o 1,24 nižšího RMSE, v období COVID vítězí fixní model při $h = 1$ s rozdílem $-0,85$. V mezikrizovém období a po pandemii jsou rozdíly mezi oběma modely minimální. Ve srovnání s rozšiřujícím se oknem jsou rozdíly celkově menší, což odráží skutečnost, že oba modely mají ve verzi posuvného okna kratší vzorek pro odhad parametrů.

Obr. 13: Rozdíl RMSE mezi fixním a auto.arima modelem – ČR, deterministická sezónnost, posuvné okno

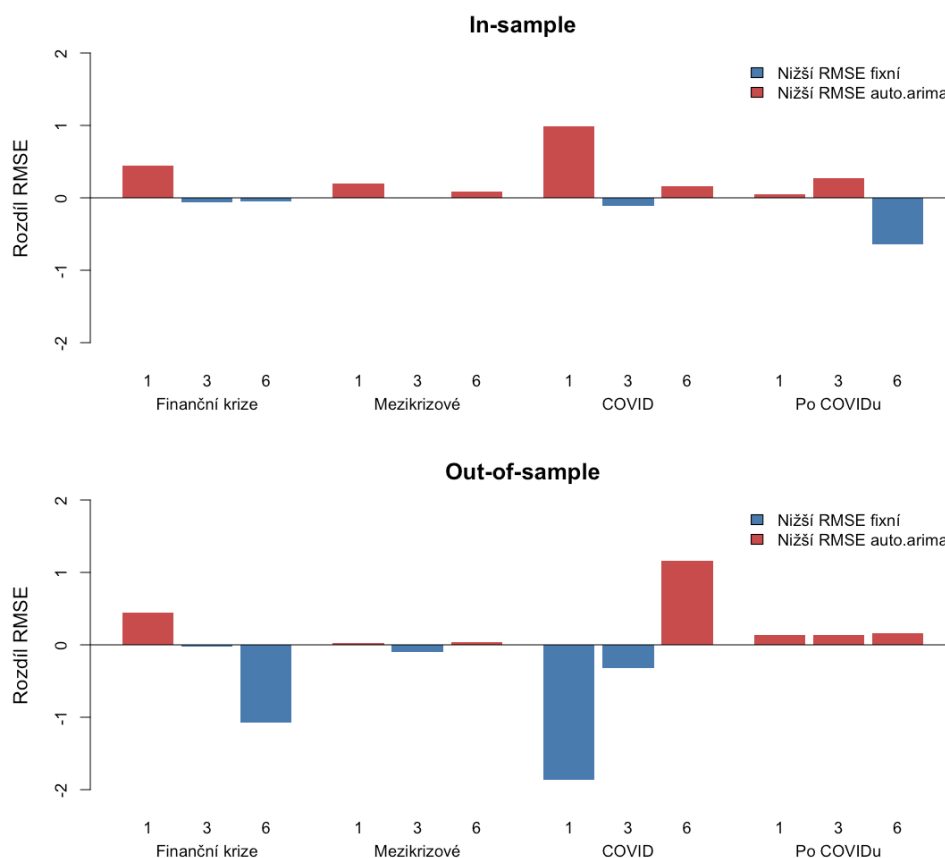


Poznámka: Čísla pod sloupci (1, 3 a 6) označují horizonty predikce.

Zdroj: vlastní zpracování

Obr. 14 zobrazuje stejné srovnání pro Slovensko. In-sample predikce jsou u většiny kombinací přesnější u auto.arima, nejvýrazněji v COVIDu a finanční krizi při $h = 1$. Výjimkou je horizont $h = 6$ po pandemii, kde fixní model dosahuje o 0,64 nižšího IS RMSE. V out-of-sample predikcích pozorujeme ve finanční krizi nekonzistenci napříč horizonty. Při $h = 1$ je auto.arima přesnější o 0,44, zatímco při $h = 6$ je fixní model přesnější o 1,07. V období COVID je vzorec ještě výraznější (při $h = 1$ fixní model vítězí o 1,87, při $h = 6$ naopak auto.arima vítězí o 1,16). V mezikrizovém období a po pandemii jsou rozdíly zanedbatelné. Oproti rozšiřujícímu se oknu je u Slovenska v období COVID při $h = 1$ převaha fixního modelu výrazně silnější.

Obr. 14: Rozdíl RMSE mezi fixním a auto.arima modelem – SK, deterministická sezónnost, posuvné okno



Poznámka: Čísla pod sloupci (1, 3 a 6) označují horizonty predikce.

Zdroj: vlastní zpracování

6.5 Dekompozice rozdílu in-sample a out-of-sample predikcí

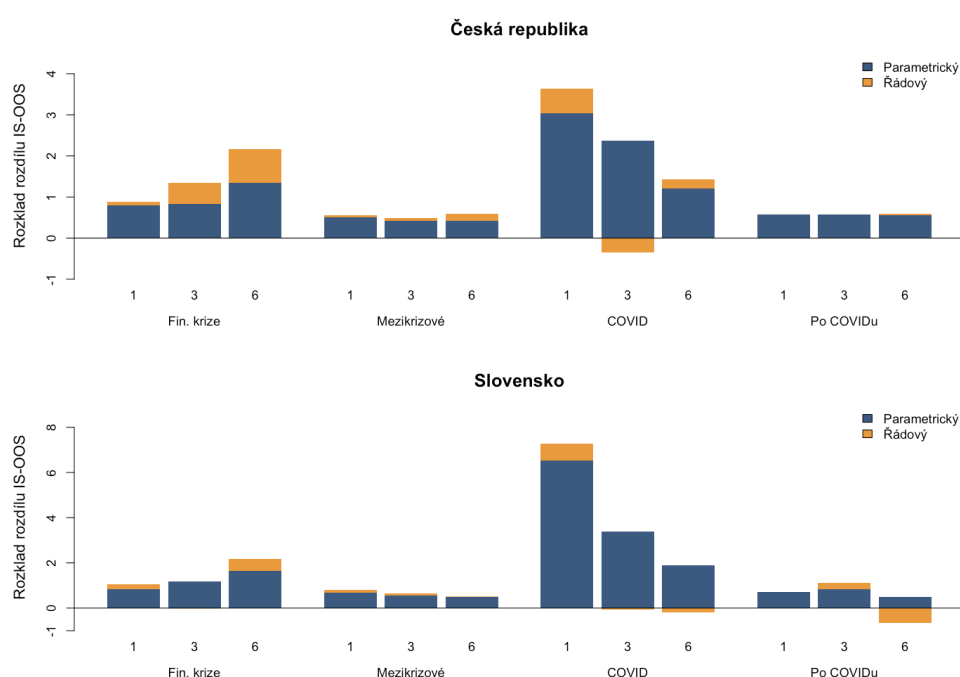
Stejně jako u rozšiřujícího se okna rozkládáme i u posuvného okna rozdíl mezi in-sample a out-of-sample predikcemi auto.arima na parametrický a řadový efekt. Vzhledem k závěru citlivostní analýzy se omezujeme pouze na deterministickou sezónnost.

Obr. 15 zobrazuje parametrický a řadový efekt po obdobích pro deterministickou sezónnost. Výsledky se v klidných obdobích téměř neliší

od výsledků pro rozšiřující se okno, tedy hlavní je parametrický efekt a řádový je prakticky nulový. Zásadní rozdíly se ale objevují v krizových obdobích. Parametrický efekt je u posuvného okna v krizích výrazně vyšší. Nejmarkantnější je to v období COVID. U Česka při $h = 1$ dosahuje +3,03 oproti +1,37 u rozšiřujícího se okna, u Slovenska při $h = 1$ dokonce +6,53 oproti +1,37. To je pravděpodobně způsobeno tím, že krátký tréninkový vzorek posuvného okna je k přidávání krizových dat citlivější.

Řádový efekt se u posuvného okna v některých krizových obdobích obrací do záporných hodnot, což u rozšiřujícího se okna prakticky nepozorujeme. Například u Česka v COVIDu při $h = 3$ je řádový efekt $-0,35$ a u Slovenska po COVIDu při $h = 6$ dokonce $-0,67$. Záporný řádový efekt znamená, že auto.arima na IS vzorku posuvného okna vybírá horší řády než na OOS vzorku. U posuvného okna se tak v některých obdobích parametrický a řádový efekt částečně neutralizují, takže celkový IS-OOS rozdíl je nižší, než by odpovídalo samotnému parametrickému efektu.

Obr. 15: Dekompozice rozdílu IS-OOS – deterministická sezónnost, posuvné okno



Poznámka: Čísla pod sloupci (1, 3 a 6) označují horizonty predikce. Kladné hodnoty rozdílu znamenají, že IS predikce jsou přesnější než OOS, záporné naopak.

Zdroj: vlastní zpracování

6.6 Shrnutí pro posuvné okno

Citlivostní analýza identifikovala optimální délku posuvného okna 96 měsíců pro Českou republiku a 72 měsíců pro Slovensko. U stochastické sezónnosti posuvné okno nedosahovalo konkurenceschopných výsledků vůči rozšiřujícímu se oknu, proto jsme se soustředili pouze na deterministickou sezónnost.

Posuvné okno je u obou zemí celkově mírně přesnější v OOS predikcích než rozšiřující se okno, zejména v klidných obdobích. Slovenské predikce z posuvného okna profitují více než české, protože starší data méně reflektují aktuální dynamiku. Výjimkou je období COVID při $h = 1$, kde je rozšiřující se okno přesnější u obou zemí. Posuvné okno totiž pokrývá pouze klidné mezikrizové období, a proto na pandemický šok reaguje méně přesně.

Rozdíl mezi oběma přístupy má v klidných a krizových obdobích odlišný původ. V klidných obdobích jej způsobují především aktuálnější odhady parametrů na kratším vzorku. Potvrzuje to i fixní model s konstantními řády, který vykazuje velmi podobné vzorce. V krizových obdobích se ale výhoda posuvného okna u fixního modelu oslabuje nebo obrací. Výhoda `auto.arima` v posuvném okně tedy pramení hlavně z odlišné volby řádů modelu. To dokládá i analýza výběru modelů. Posuvné okno preferuje u Slovenska `random walk (0,1,0)` ve 47 % oken, zatímco rozšiřující se okno silně preferuje `(0,1,1)`.

Dekompozice rozdílu IS-OOS u posuvného okna ukazuje, že v klidných obdobích dominuje parametrický efekt, shodně s rozšiřujícím se oknem. V COVIDu je parametrický efekt u posuvného okna mnohem silnější než u rozšiřujícího se okna, kvůli velké váze krizových pozorování v kratším vzorku.

Napříč všemi testovanými přístupy s deterministickou sezónností dosahuje nejnižšího IS RMSE `auto.arima` při posuvném okně u obou zemí a většiny horizontů. Důvodem je krátký tréninkový vzorek v kombinaci s flexibilní volbou modelu, který se nejlépe přizpůsobí tréninkovým datům. V OOS predikcích jsou výsledky odlišné. U Česka jsou v případě posuvného okna fixní a `auto.arima` modely téměř identické. U Slovenska je fixní model posuvného okna mírně přesnější na kratších horizontech. Při $h = 1$ má OOS RMSE 5,43 oproti 5,63 u `auto.arima` v posuvném okně, tedy rozdíl 0,2. Rozdíly jsou ale napříč horizonty velmi malé a u $h = 3$ i $h = 6$ jsou oba modely prakticky vyrovnané.

7 Srovnání predikcí při manuálním výběru modelu

V předchozích kapitolách jsme hodnotili predikce napříč celým dostupným vzorkem a po rozdělení do období. V této kapitole se zaměříme na dvě kontrastní období: finanční krizi 2008–2009 jako období strukturálního šoku a rok 2018 jako poslední klidný rok průmyslové expanze před cyklickým zpomalením v roce 2019. Cílem je porovnat manuálně volenou ARIMA specifikaci s automatickým postupem `auto.arima` v obou uvedených obdobích, a to jak v identifikaci modelu, tak v jeho predikční přesnosti.

V tomto případě používáme odlišný přístup k in-sample predikcím. Na rozdíl od konstrukce v předchozích kapitolách zde máme pouze jeden in-sample model zahrnující celé predikované období. Jinak se odhad predikcí neliší a cílová pozorování jsou stále h kroků vzdálená od počátku predikce. Out-of-sample predikce vznikají stejně jako u rozšiřujícího se okna.

Analýzu provádíme zvlášť pro Česko a Slovensko a pro každou zemi vybíráme model na základě ACF/PACF diagnostiky a hodnoty BIC. Tréninkový vzorek volíme v přístupu rozšiřujícího se okna s pevně stanoveným počátkem. Pro analýzu finanční krize 2008–2009 je počátkem leden 2000 (celá dostupná předkrizová historie). Pro analýzu roku 2018 je počátkem leden 2010, což z tréninkových dat vylučuje finanční krizi. Důvodem tohoto asymetrického výběru je, že krizová dynamika 2008–2009 je strukturálně odlišná od klidného období 2010–2018, a její zahrnutí do tréninku pro predikci klidného roku by snižovalo přesnost odhadů parametrů relevantních pro toto období.

ACF a PACF grafy zobrazujeme v hlavním textu pro OOS vzorek (tréninkový vzorek před cílovým obdobím). Pro úplnost jsou v Příloze C uvedeny i odpovídající ACF a PACF pro IS vzorek (OOS vzorek rozšířený o cílové období), který slouží jako kontrolní diagnostika. Totéž platí pro testy stacionarity.

7.1 Predikce finanční krize

Pro období finanční krize volíme ARIMA model s deterministickou sezónností (dummy proměnné), jelikož předchozí analýza ukázala, že je

v krizovém období přesnější než stochastický SARIMA. Cílovým obdobím predikcí je 15 měsíců od října 2008 do prosince 2009. Počáteční tréninkový vzorek je tedy pro out-of-sample 01/2000–09/2008. Pro in-sample je pouze jeden tréninkový vzorek 2000–2009.

7.1.1 Testování stacionarity

Stacionaritu řad IPP testujeme kombinací KPSS testu s nulovou hypotézou stacionarity a ADF testu s nulovou hypotézou jednotkového kořene. Jelikož sezónnost v následných modelech zachycujeme deterministicky pomocí dummy proměnných, testy aplikujeme na rezidua z OLS regrese y_t na sezónní dummy proměnné, jak doporučují Dickey, Bell a Miller (1986).

V Tab. 13 jsou zobrazeny výsledky ADF a KPSS testů. U úrovní KPSS zamítá stacionaritu na 1 % hladině významnosti u obou zemí. Po první diferenci KPSS nezamítá stacionaritu a ADF zamítá jednotkový kořen u obou zemí. Je proto vhodné použít ARIMA model s dummy proměnnými a první diferencí.

Tab. 13: Testy stacionarity – out-of-sample vzorek pro finanční krizi

Země	Období	Úrovně		Diference	
		KPSS	ADF	KPSS	ADF
ČR	2000–09/2008	0,010	0,610	> 0,100	< 0,010
SK	2000–09/2008	0,010	0,810	> 0,100	< 0,010

Poznámka: Hodnoty < 0,010 a > 0,100 reflektují hraniční hodnoty tabulky těchto testů v R

Zdroj: vlastní zpracování

7.1.2 Identifikace modelů

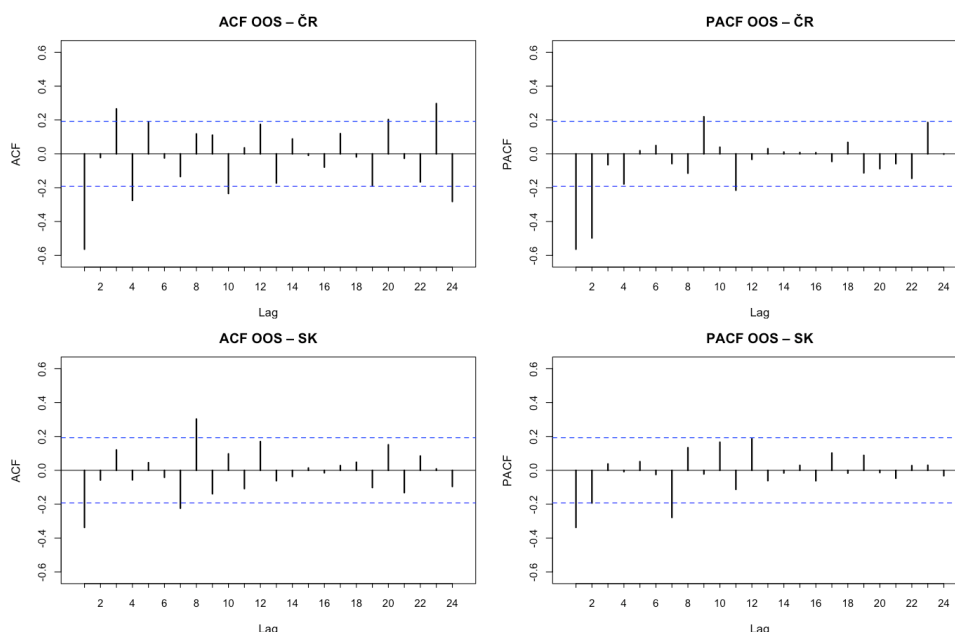
Modely identifikujeme pomocí ACF a PACF první difference IPP po odstranění deterministické sezónní složky (regresí na měsíční dummy proměnné). Obě funkce jsou znázorněny na Obr. 16.

U české řady vykazuje ACF záporný spike na lagu 1 (–0,56) následovaný oscilacemi se střídavým znaménkem. PACF má dva výrazné záporné spiky na lagu 1 a 2, přičemž další lagy již nejsou významné. Tento vzor naznačuje autoregresní strukturu pravděpodobně řádu 2, ale pro robustní výběr testujeme i alternativní specifikace.

U slovenské řady je vzor méně jednoznačný. ACF vykazuje jediný výrazný spike na lagu 1, ostatní počáteční lags jsou nevýznamné. PACF obsahuje rovněž pouze jeden výrazný spike na lagu 1. Vzor je kompatibilní jak s MA(1), tak s AR(1), případně s jejich kombinací.

Pro výběr modelu porovnáváme kandidátní specifikace pomocí BIC. Pro Česko testujeme ARIMA(p,1,q) s $p \in \{0, 1, 2\}$ a $q \in \{0, 1\}$, pro Slovensko s podobným nastavením, ale s omezením $p \in \{0, 1\}$ odpovídajícím užšímu prostoru kandidátních modelů naznačenému ACF a PACF. V obou variantách zvažujeme i zahrnutí driftu.

Obr. 16: ACF a PACF IPP po první diferenci a odstranění deterministické sezónní složky, 01/2000–09/2008



Zdroj: vlastní zpracování

Tab. 14 uvádí čtyři nejlepší modely pro tréninkový vzorek OOS (2000–09/2008) a IS (2000–2009) na základě BIC. Řády modelu vybrané pomocí BIC jsou pro oba vzorky stejné. Pro Česko vychází ARIMA(2,1,0) a pro Slovensko ARIMA(0,1,1). Rozdíl je v zahrnutí driftu. Na OOS vzorku BIC drift preferuje, protože lineární trend dobře zachycuje monotónně rostoucí vývoj. Na IS vzorku, který již zahrnuje krizová pozorování, BIC drift naopak vyřazuje, jelikož lineární trend je v rozporu s prudkým

propadem roku 2008. Model ARIMA(2,1,0) je tudíž zvolen pro Česko a ARIMA(0,1,1) pro Slovensko.

Tab. 14: Výběr modelu pro finanční krizi na základě BIC

Model	BIC OOS (pořadí)	BIC IS (pořadí)
<i>Česko</i>		
ARIMA(2,1,0) + drift	543,48 (1)	647,02 (3)
ARIMA(2,1,1) + drift	547,90 (2)	648,34 (4)
ARIMA(2,1,0)	548,21 (3)	644,56 (1)
ARIMA(2,1,1)	552,17 (6)	645,43 (2)
<i>Slovensko</i>		
ARIMA(0,1,1) + drift	633,20 (1)	734,51 (4)
ARIMA(0,1,1)	633,45 (2)	731,43 (1)
ARIMA(1,1,0)	634,51 (3)	732,56 (2)
ARIMA(0,1,0)	641,31 (7)	733,51 (3)

Zdroj: vlastní zpracování

Před prezentací predikčních výsledků kontrolujeme rezidua zvolených modelů Ljung-Boxovým testem na lagu 24. Vzhledem k tomu, že práce reportuje pouze bodové predikce, neprovádíme dodatečné testy normality a homoskedasticity, které by měly význam především pro konstrukci konfidenčních intervalů. Na OOS vzorku test nezamítá nulovou hypotézu nekorelovanosti reziduů na 5 % hladině pro ČR ARIMA(2,1,0)+dummy ($p = 0,058$) ani pro SK ARIMA(0,1,1)+dummy ($p = 0,446$). Na IS vzorku se u modelu ČR zbytková autokorelace zhoršuje ($p = 0,012$), což odráží šokovou povahu krizových pozorování přidaných do tréninkového vzorku. Vybraný model přes přítomnou zbytkovou autokorelaci akceptujeme, neboť ani žádný z ostatních kandidátních modelů identifikovaných ACF a PACF analýzou LB(24) testem neprochází.

7.1.3 Výsledky predikcí

Uvádíme výsledky predikcí modelů vybraných v předchozí podkapitole. Pro in-sample predikce tedy volíme ARIMA(2,1,0) pro Česko a ARIMA(0,1,1) pro Slovensko. Pro out-of-sample predikce volíme stejné řády modelu, ale rozlišujeme model bez driftu a s driftem.

Tab. 15: RMSE predikcí – finanční krize, manuální výběr

	IS	OOS	OOS + drift
<i>Česko</i>			
h = 1	3,87	4,57	4,94
h = 3	5,42	5,86	6,46
h = 6	8,96	9,42	11,00
<i>Slovensko</i>			
h = 1	4,93	5,73	6,07
h = 3	8,72	9,64	10,38
h = 6	12,56	13,35	15,43

Zdroj: vlastní zpracování

V Tab. 15 prezentujeme RMSE predikcí manuálně vybraných modelů v období finanční krize. Procentuální IS-OOS rozdíl klesá s rostoucím horizontem (u ČR z 18 % u $h = 1$ na 5 % u $h = 6$, u SK z 16 % na 6 %). Klíčovým rozhodnutím je nakládání s driftem. BIC sice drift na OOS tréninkovém vzorku jednoznačně preferuje pro obě země, avšak jeho zahrnutí zhoršuje predikce krize (u ČR o 8 až 17 % a u SK o 6 až 16 % oproti variantě bez driftu).

Ve srovnání s auto.arima přístupem z kapitoly 5 (varianta s dummy proměnnými) dosahují manuálně vybrané OOS modely bez driftu výrazně nižší RMSE (u ČR o 11 až 22 % a u SK o 3 až 12 % napříč horizonty). Toto srovnání je ovšem nutné interpretovat opatrně, jelikož drift byl vyřazen na základě znalosti, že krize přijde, kterou auto.arima nemá. Férovějším srovnáním je manuálně vybraný model s driftem proti auto.arima, kde oba přístupy nechávají BIC drift vybrat. Při tomto přístupu dosahuje ARIMA(2,1,0) s driftem u Česka o 4 až 9 % nižší RMSE než auto.arima, u Slovenska jsou hodnoty prakticky shodné (manuál o 1 až 3 % horší). Hlavní výhoda manuální specifikace tedy nespočívá v lepším řádu modelu, ale v možnosti vědomě vyřadit drift v očekávání propadu. Pokud vybereme model pouze ze znalosti předkrizových dat, out-of-sample predikce se od auto.arima výrazně neliší.

U Česka není IS RMSE manuálně vybraného modelu nižší než u auto.arima, ačkoliv tréninkový vzorek obsahuje všechna krizová pozorování a mohl by se jim tak lépe přizpůsobit. Z toho plyne, že auto.arima dokáže v krizi nově zvolenou specifikaci dosáhnout lepšího odhadu na tréninkových datech, což se však nepromítá do přesnějších out-of-

sample predikcí. U Slovenska to neplatí a IS RMSE je mezi přístupy téměř stejný.

7.2 Predikce roku 2018

Pro klidné období roku 2018 volíme SARIMA model se stochastickou sezónností. Z předchozí analýzy vyplývá, že v mezikrizovém období dosahuje tento přístup nižších OOS RMSE než ARIMA s deterministickými dummy proměnnými. Cílovým obdobím predikcí je 12 měsíců od ledna do prosince 2018.

7.2.1 Testování stacionarity

Stejně jako u finanční krize testujeme stacionaritu tréninkových vzorků pomocí KPSS testu s nulovou hypotézou stacionarity a ADF testu s nulovou hypotézou jednotkového kořene. Na rozdíl od krize pracujeme se stochastickým modelem typu SARIMA se sezónní diferencí $D = 1$. Testy proto aplikujeme na řadu po této diferenci a ověřujeme, zda je potřeba doplnit také nesezónní diferencí $d = 1$.

Výsledky Tab. 16 jsou shodné pro obě země. Po samotné sezónní diferenci $D = 1$ ADF test nezamítá jednotkový kořen u žádné z řad a u Česka je navíc KPSS na samé hranici 5 % hladiny. Sezónní difference k zajištění stacionarity nestačí. Po doplnění nesezónní difference $d = 1$ oba testy stacionaritu jednoznačně potvrzují u obou zemí. Pro modelování proto v obou případech volíme specifikaci SARIMA($p, 1, q$)($P, 1, Q$).

Tab. 16: Testy stacionarity – out-of-sample vzorek pro rok 2018

Země	Období	Po $D = 1$		Po $d = 1 + D = 1$	
		KPSS	ADF	KPSS	ADF
ČR	2010–2017	0,053	0,258	> 0,100	< 0,010
SK	2010–2017	> 0,100	0,237	> 0,100	< 0,010

Poznámka: Hodnoty < 0,010 a > 0,100 reflektují hraniční hodnoty tabulky těchto testů v R

Zdroj: vlastní zpracování

7.2.2 Identifikace modelů

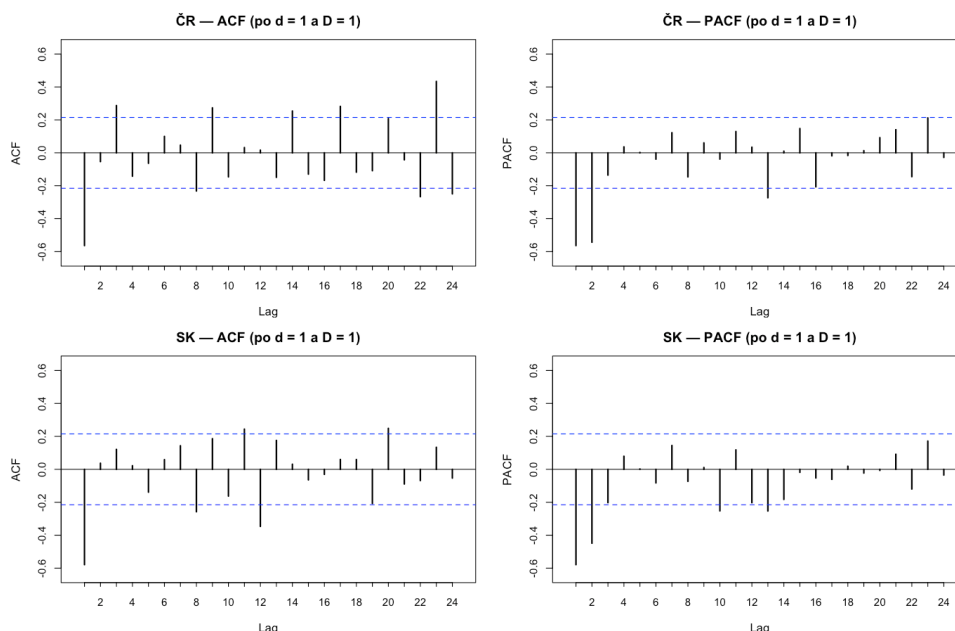
Modely identifikujeme pomocí ACF a PACF řady IPP po nesezónní i sezónní diferenci ($d = 1, D = 1$) v souladu s výsledky testů stacionarity. Obě funkce jsou znázorněny na Obr. 17.

U české řady vykazuje ACF záporný spike na lagu 1 následovaný oscilacemi se střídavým znaménkem. Sezónní lag 12 v ACF není významný, na lagu 24 však vykazuje výrazný záporný spike. PACF má dva výrazné záporné spiky na lagu 1 a 2, další lags včetně sezónních (12 a 24) již nejsou významné. Vzor odpovídá nesezónnímu AR(2) procesu se sezónní MA složkou.

U slovenské řady vykazuje ACF rovněž záporný spike na lagu 1. Sezónní lag 12 v ACF je výrazně záporný, lag 24 však významný není. PACF má stejnou strukturu jako u české řady. Vzor odpovídá nesezónnímu AR(2) procesu se sezónní MA složkou prvního řádu.

Pro výběr modelu porovnáváme kandidátní specifikace pomocí BIC. Pro Česko testujeme SARIMA($p, 1, q$)($0, 1, Q$) s $p \in \{0, 1, 2\}$, $q \in \{0, 1\}$, $Q \in \{0, 1, 2\}$. Pro Slovensko používáme stejný prostor parametrů, ale s $Q \in \{0, 1\}$, jelikož ACF slovenské řady na sezónním lagu 24 nevykazuje významný spike.

Obr. 17: ACF a PACF IPP po první i sezónní diferenci 2010–2017



Zdroj: vlastní zpracování

Tab. 17 uvádí tři nejlepší modely pro OOS vzorek (2010–2017) a IS vzorek (2010–2018) na základě BIC. Řády modelu vybrané pomocí BIC jsou pro oba vzorky stejné. Pro Česko vychází SARIMA(2,1,0)(0,1,2) a pro Slovensko SARIMA(2,1,0)(0,1,1). Oba modely sdílí stejnou nesezónní AR(2) strukturu a liší se pouze řádem sezónní MA složky, což odpovídá identifikaci z ACF a PACF.

Tab. 17: Výběr modelu pro rok 2018 na základě BIC

Model	BIC OOS (pořadí)	BIC IS (pořadí)
<i>Česko</i>		
SARIMA(2,1,0)(0,1,2)	459,79 (1)	521,89 (1)
SARIMA(2,1,1)(0,1,2)	462,35 (2)	524,38 (2)
SARIMA(2,1,0)(0,1,0)	465,46 (3)	529,36 (5)
<i>Slovensko</i>		
SARIMA(2,1,0)(0,1,1)	477,33 (1)	538,38 (1)
SARIMA(0,1,1)(0,1,1)	478,86 (2)	541,27 (2)
SARIMA(2,1,1)(0,1,1)	479,77 (3)	541,61 (3)

Zdroj: vlastní zpracování

Stejně jako u krizových modelů kontrolujeme rezidua Ljung-Boxovým testem na lagu 24. Na OOS vzorku test zamítá nulovou hypotézu nekorelovanosti reziduí na 5 % hladině pro ČR SARIMA(2,1,0)(0,1,2) ($p = 0,02$), zatímco pro SK SARIMA(2,1,0)(0,1,1) nezamítá ($p = 0,36$). Na IS vzorku se diagnostika ČR modelu zlepšuje a test již nezamítá nulovou hypotézu na 5 % hladině ($p = 0,09$). Pozorování z klidného roku 2018 přidaná do tréninkového vzorku tedy zbytkovou autokorelaci spíše snižují. SK model nezamítá ani na IS vzorku ($p = 0,39$). Vybraný ČR model přes zbytkovou autokorelaci na OOS akceptujeme, neboť LB(24) testem procházejí jen tři z osmnácti kandidátních modelů identifikovaných ACF a PACF analýzou, a to s p -hodnotami těsně nad 5 % hladinou. Všechny tyto modely navíc dosahují horších predikčních chyb na všech horizontech.

7.2.3 Výsledky predikcí

V Tab. 18 prezentujeme RMSE predikcí manuálně vybraných modelů v roce 2018. Pro Česko volíme SARIMA(2,1,0)(0,1,2) a pro Slovensko SARIMA(2,1,0)(0,1,1).

IS-OOS rozdíl je v roce 2018 řádově menší než během krize. U Česka mezera kolísá mezi 0,11 a 0,19 RMSE napříč horizonty, u Slovenska mezi 0,03 a 0,12. Toto pozorování je konzistentní s výsledky analýzy rozšiřujícího se okna z kapitoly 5, kde i auto.arima modely vykazovaly v mezikrizovém období jen malou IS-OOS mezeru. Klidné období bez strukturálních zlomů zajišťuje stabilní odhady parametrů, takže přidání cílového pozorování do tréninkového vzorku už predikce znatelně nemění.

Tab. 18: RMSE predikcí – rok 2018, manuální výběr

	IS	OOS
<i>Česko</i>		
h = 1	3,04	3,23
h = 3	3,15	3,26
h = 6	3,29	3,41
<i>Slovensko</i>		
h = 1	2,85	2,92
h = 3	3,43	3,46
h = 6	3,57	3,69

Zdroj: vlastní zpracování

7.3 Shrnutí pro manuální výběr

Manuální výběr modelu byl proveden ve dvou kontrastních obdobích, během finanční krize 2008–2009 a v klidném roce 2018. Pro krizové období jsme zvolili ARIMA s deterministickou sezónností (dummy proměnné), pro klidné období stochastickou SARIMA specifikaci. Identifikace vedla pro Česko napříč obdobími k nesezónní AR(2) struktuře. U Slovenska je v krizi vybrána MA(1) a v roce 2018 stejná AR(2) jako u Česka se sezónní MA(1) složkou. Konkrétně ARIMA(2,1,0) pro Česko a ARIMA(0,1,1) pro Slovensko v krizi, SARIMA(2,1,0)(0,1,2) pro Česko a SARIMA(2,1,0)(0,1,1) pro Slovensko v roce 2018.

Výsledky kapitoly potvrzují hlavní vzory pozorované v analýze rozšiřujícího se okna v kapitole 5. Vybrané řády jsou kompatibilní s tím, co typicky volí auto.arima (zejména AR(2) struktura napříč obdobími u ČR), a IS-OOS mezera je v roce 2018 stejně malá jako u auto.arima v mezikrizovém období. Při férovém srovnání s driftem v krizi, kdy oběma přístupům necháme BIC drift vybrat, je rozdíl v predikční přesnosti malý. Fixní

model pro Česko má jen mírnou převahu (4–9 %) a model pro Slovensko je s `auto.arima` srovnatelný. Volba řádu modelu sama o sobě tedy mezi manuálním a automatickým přístupem zásadní rozdíl nepřináší.

Zároveň ale analýza ukazuje, že `auto.arima` má omezení. Například ze samotných `auto.arima` výsledků nelze odlišit, jakou část rozdílu mezi in-sample a out-of-sample predikcemi v krizi způsobuje volba řádů modelu a jakou zahrnutí driftu, jelikož řádový efekt v dekompozici z kapitoly 5 obě složky slévá. Manuální analýza tuto otázku zodpovídá tím, že drift v predikčních variantách kontroluje explicitně. Obdobně volba délky tréninkového vzorku, která pro analýzu roku 2018 vylučovala krizová pozorování, není standardní součástí `auto.arima` procedury. Manuální výběr tak nepředstavuje alternativu k `auto.arima`, ale spíš její doplnění v situacích, kdy je k dispozici dodatečná informace o povaze sledovaného období.

8 Diskuse

V práci docházíme k několika zjištěním. Rozdíl mezi in-sample a out-of-sample predikcemi závisí především na charakteru období. V klidných obdobích je rozdíl malý a je primárně způsoben rozdíly v odhadech parametrů mezi in-sample a out-of-sample modely. V krizových obdobích je výrazně vyšší z důvodu jak změny odhadů parametrů, tak i volby odlišné specifikace in-sample modelem kvůli přítomnosti krizových pozorování.

Dalším zjištěním je, že jednoduchý fixní model s deterministickou sezónností poskytuje ve většině období přesnější out-of-sample predikce než auto.arima s respecifikací řádu v každém kroku. Hlavní výhoda se projevuje v krizích, kde se auto.arima při přidání krizových pozorování do vzorku přizpůsobuje jejich dynamice a volí složitější specifikace, které na delší horizonty predikují hůře. U stochastické sezónnosti fixní model tak konzistentní převahu nemá. Jeho výsledky jsou napříč zeměmi i obdobími smíšené, jelikož pevně daná dvojitá diferenciací omezuje schopnost modelu přizpůsobit se různým režimům, zatímco auto.arima volbou nových řádů pro každé okno si zachovává flexibilitu.

Volba mezi deterministickou a stochastickou sezónností vykazuje podobnou asymetrii. V klidných obdobích je SARIMA přesnější, jelikož flexibilní sezónní složka lépe zachytí postupně se měnící sezónní vzorec. V krizových obdobích je naopak deterministická ARIMA s dummy proměnnými robustnější. Sezónní difference $D = 1$ v SARIMA reaguje na krizový šok tak, jako by se sezónní vzorec změnil, zatímco deterministická dummy reprezentace zůstává zafixovaná a vůči jednorázovému šoku v úrovni řady je odolnější.

Volba mezi rozšiřujícím se a posuvným oknem závisí na přístupu k sezónnosti. Pro deterministické modelování s dummy proměnnými se jako přesnější ukazuje metoda posuvného okna, která snižuje in-sample a out-of-sample RMSE díky aktuálnějším odhadům parametrů. Pro SARIMA modely je naopak vhodnější rozšiřující se okno, jelikož stochastická sezónnost potřebuje delší tréninkový vzorek pro stabilní odhad sezónních složek.

Manuální výběr ARIMA modelu ve zkoumaných krizových i klidných obdobích sloužil jako kontrolní analýza. Pro Česko obě metody konvergují k AR(2) struktuře napříč obdobími. Manuální výběr modelu tedy je

v souladu s hlavním zjištěním práce a automatická volba řádu pomocí `auto.arima` se jeví jako věcně správná.

Přesnost predikcí jsme v práci hodnotili pomocí RMSE. Rovněž jsme ovšem analyzovali i hodnoty MAE (uvedené v Příloze B). U většiny období a horizontů se rozdíly mezi in-sample a out-of-sample predikcemi v obou metrikách významně neliší a docházíme tak ke stejným závěrům. Jediné významnější rozdíly pozorujeme u Slovenska v období před krizí a finanční krize, kdy v krátké době došlo k výraznému nárůstu a poklesu indexu průmyslové produkce, což vedlo k větším predikčním chybám, které RMSE penalizuje více.

Hlavní zjištění práce o rozdílech mezi in-sample a out-of-sample predikcemi by mohlo být vhodné podpořit testem statistické významnosti pozorovaných rozdílů. Standardním nástrojem pro tento typ srovnání je Diebold-Marianův test, který je však konstruován pro porovnání out-of-sample predikcí dvou modelů. V naší konstrukci je situace asymetrická, neboť in-sample model při odhadu cílovou hodnotu využívá, zatímco out-of-sample model ne. Tato asymetrie vede k porušení předpokladů testu (zejména nezávislosti a shodné distribuce chyb), což znemožňuje jeho korektní interpretaci. Z tohoto důvodu v práci Diebold-Marianův test nepoužíváme.

Vhodným rozšířením výzkumu by rovněž mohlo být srovnání in-sample a out-of-sample predikcí pouze na klidném období, kdy by cílové období i tréninkový vzorek sestávaly výhradně z klidných pozorování. Na takovém vzorku by se mohl ještě prohloubit rozdíl mezi deterministickou a stochastickou sezónností. SARIMA je v klidném období sice přesnější, ale jeho odhady parametrů mohou být v naší analýze negativně ovlivněny přítomností finanční krize v tréninkovém vzorku.

Dalším možným rozšířením je srovnání in-sample a out-of-sample predikcí na sezonně očištěné řadě IPP. Na takto upravené řadě by ARIMA modelovala pouze trendovou a cyklickou složku a sezónní vzor by byl odstraněn ještě před modelováním. Srovnání in-sample a out-of-sample rozdílů na očištěné a neočištěné řadě by ukázalo, do jaké míry pramení pozorovaný rozdíl ze samotného modelování sezónnosti a do jaké míry z nestability trendu.

Za úvahu stojí také použití jiných modelů než ARIMA. Musely by být vhodné pro reestimaci, aby bylo možné porovnat in-sample a out-of-sample na dlouhém vzorku. V naší práci jsme testovali VAR model, který ovšem neposkytl vyšší predikční přesnost než jednorozměrná `auto.arima`. VAR vyžaduje volbu vstupních proměnných a testy

kointegrace mezi nimi. Tyto vztahy se však mohou v čase měnit, což v rámci automatické a opakované respecifikace modelu není snadné průběžně ověřovat, a snižuje to praktickou použitelnost VAR v naší metodologii. Alternativními kandidáty jsou state space modely s dekompozicí řady na trend, sezónní a nepravidelnou složku, neparametrické vícerozměrné metody typu MSSA, nebo modely strojového učení schopné zachytit nelineární dynamiku.

9 Závěr

Tato diplomová práce se zabývala srovnáním in-sample a out-of-sample predikcí indexu průmyslové produkce České republiky a Slovenska s cílem posoudit, jak velký je rozdíl mezi oběma typy predikcí a jaké faktory tento rozdíl ovlivňují. Hlavním nástrojem bylo opakované generování predikcí pomocí ARIMA modelů v několika metodologických variantách.

Nejprve byla provedena rešerše literatury, na jejíž základě byla zvolena metodologie vycházející ze závěrů odborných článků zabývajících se predikcí indexu průmyslové produkce. Odborné články ovšem pokrývají pouze období do roku 2010 a práce tak přináší nové poznatky aplikací na novější data.

Na rešerši literatury navazuje analýza dat a volba metodologie. Analýza pro data z období 2000–2023 modelovala predikce 1, 3 a 6 měsíců dopředu, k čemuž byly využity dva standardní přístupy, rozšiřující se a posuvné okno. Řády modelů jsou voleny pomocí funkce `auto.arima` v každém kroku, případně je definován model s fixními řády. Vzorek byl rozdělen na pět období podle makroekonomického kontextu (před krizí, finanční krize 2008–2009, mezikrizové, COVID, po COVIDu).

V analytické části byl nejprve vyhodnocen výkon rozšiřujícího se okna napříč obdobími pro obě země. Výsledky ukázaly, že velikost rozdílu mezi in-sample a out-of-sample predikcemi závisí především na charakteru období. V klidných obdobích je rozdíl malý a je primárně způsoben rozdíly v odhadech parametrů mezi in-sample a out-of-sample modely. V krizích rozdíl výrazně roste a významnou roli v něm hraje řádivý efekt, kdy `auto.arima` na in-sample vzorku volí výrazně odlišné řády modelu. Ze dvou přístupů k modelování sezónnosti dosahuje SARIMA v klidných obdobích vyšší přesnosti, ale v krizích je deterministická ARIMA s dummy proměnnými robustnější.

Následně byla srovnána rozšiřující se a posuvná okna. Pro deterministickou sezónnost dosahuje posuvné okno s optimální velikostí okna (96 měsíců pro Česko a 72 měsíců pro Slovensko) nižších OOS predikčních chyb než rozšiřující se okno, zejména v klidných obdobích. Pro stochastickou sezónnost není posuvné okno konkurenceschopné v žádném horizontu, jelikož SARIMA potřebuje delší tréninkový vzorek pro stabilní odhad sezónních složek. Jednoduchý fixní model s deterministickou sezónností ve většině období překonává `auto.arima`, výjimkou je pouze předkrizové období, kdy `auto.arima` díky driftu predikuje přesněji.

V poslední části práce byl proveden manuální výběr ARIMA modelu pro dvě kontrastní období, finanční krizi 2008–2009 a rok 2018. Identifikace pomocí ACF, PACF a BIC kritéria vedla pro Česko ke shodné AR(2) struktuře napříč oběma obdobími. U Slovenska se vybrané struktury liší. V krizi je vybrána MA(1) a v roce 2018 stejná AR(2) jako u Česka. Manuální výběr ukazuje, že volba řádu modelu sama o sobě nepřináší zásadní rozdíl oproti automatickému postupu `auto.arima`. Hlavní přínos manuálního přístupu spočívá v možnosti využít dodatečnou informaci o povaze období, například ve vědomém vyřazení driftu v očekávání propadu.

Hlavním zjištěním práce je, že rozdíl mezi in-sample a out-of-sample predikcemi není konstantní vlastností modelu, ale silně závisí na charakteru predikovaného období. V klidných obdobích zůstává rozdíl malý a in-sample i out-of-sample přesnost se výrazně neliší. V krizových obdobích se rozdíl výrazně rozšiřuje a stejně tak roste citlivost na volbu modelu. Žádný z testovaných přístupů tak není univerzálně nejlepší. Pro období strukturálních zlomů jsou vhodnější jednodušší fixní specifikace s deterministickou sezónností, v klidných obdobích naopak SARIMA s rozšiřujícím se oknem. Výsledky práce tak ukazují, že charakter období hraje klíčovou roli při volbě modelu.

Použité zdroje

BEAULIEU, J. Joseph a Jeffrey A. MIRON, 1993. Seasonal unit roots in aggregate U.S. data. *Journal of Econometrics*, Vol. 55, Iss. 1-2, pp. 305-328. ISSN 0304-4076. DOI: [https://doi.org/10.1016/0304-4076\(93\)90018-Z](https://doi.org/10.1016/0304-4076(93)90018-Z)

BIANCHI, Carluccio, Alessandro CARTA, Dean FANTAZZINI, Maria Elena DE GIULI a Mario A. MAGGI, 2010. A copula-VAR-X approach for industrial production modelling and forecasting. *Applied Economics*, Vol. 42, Iss. 25, pp. 3267-3277. ISSN 0003-6846. DOI: <https://doi.org/10.1080/00036840802112349>

BODO, Giorgio, Roberto GOLINELLI a Giuseppe PARIGI, 2000. Forecasting industrial production in the euro area. *Empirical Economics*, Vol. 25, Iss. 4, pp. 541-561. ISSN 0377-7332. DOI: <https://doi.org/10.1007/s001810000032>

BRUNO, Giancarlo a Claudio LUPI, 2004. Forecasting industrial production and the early detection of turning points. *Empirical Economics*, Vol. 29, Iss. 3, pp. 647-671. ISSN 0377-7332. DOI: <https://doi.org/10.1007/s00181-004-0203-y>

BULLIGAN, Guido, Roberto GOLINELLI a Giuseppe PARIGI, 2010. Forecasting monthly industrial production in real-time: from single equations to factor-based models. *Empirical Economics*, Vol. 39, Iss. 2, pp. 303-336. ISSN 0377-7332. DOI: <https://doi.org/10.1007/s00181-009-0305-7>

BURNHAM, Kenneth P. a David R. ANDERSON, 2004. Multimodel Inference: Understanding AIC and BIC in Model Selection. *Sociological Methods & Research*, Vol. 33, Iss. 2, pp. 261-304. ISSN 0049-1241. DOI: <https://doi.org/10.1177/0049124104268644>

CAMPBELL, John Y. a Samuel B. THOMPSON, 2008. Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average?

Review of Financial Studies, Vol. 21, Iss. 4, pp. 1509-1531. ISSN 0893-9454. DOI: <https://doi.org/10.1093/rfs/hhm055>

CLARK, Todd E., 2004. Can Out-of-Sample Forecast Comparisons Help Prevent Overfitting? *Journal of Forecasting*, Vol. 23, Iss. 2, pp. 115-139. ISSN 0277-6693. DOI: <https://doi.org/10.1002/for.904>

CLARK, Todd E. a Michael W. MCCRACKEN, 2006. The Predictive Content of the Output Gap for Inflation: Resolving In-Sample and Out-of-Sample Evidence. *Journal of Money, Credit and Banking*, Vol. 38, Iss. 5, pp. 1127-1148. ISSN 0022-2879. DOI: <https://doi.org/10.1353/mcb.2006.0068>

CLEMENTS, Michael P. a David F. HENDRY, 1997. An empirical study of seasonal unit roots in forecasting. *International Journal of Forecasting*, Vol. 13, Iss. 3, pp. 341-355. ISSN 0169-2070. DOI: [https://doi.org/10.1016/S0169-2070\(97\)00022-8](https://doi.org/10.1016/S0169-2070(97)00022-8)

DICKEY, David A., William R. BELL a Robert B. MILLER, 1986. Unit Roots in Time Series Models: Tests and Implications. *The American Statistician*, Vol. 40, Iss. 1, pp. 12-26. ISSN 0003-1305. DOI: <https://doi.org/10.1080/00031305.1986.10475349>

ENDERS, Walter, 2014. *Applied Econometric Time Series*. 4th edition. Hoboken: Wiley. ISBN 978-1-118-80856-6.

EUROSTAT, 2026. Production in industry – monthly data (sts_inpr_m) [online]. [cit. 2026-04-21]. Dostupné z: https://ec.europa.eu/eurostat/databrowser/view/sts_inpr_m/default/table

FRANSES, Philip Hans a Dick VAN DIJK, 2005. The forecasting performance of various models for seasonality and nonlinearity for quarterly industrial production. *International Journal of Forecasting*, Vol. 21, Iss. 1, pp. 87-102. ISSN 0169-2070. DOI: <https://doi.org/10.1016/j.ijforecast.2004.05.005>

GHYSELS, Eric a Denise R. OSBORN, 2001. The Econometric Analysis of Seasonal Time Series. Cambridge: Cambridge University Press. ISBN 978-0-521-56515-3

GOYAL, Amit a Ivo WELCH, 2008. A Comprehensive Look at The Empirical Performance of Equity Premium Prediction. *Review of Financial Studies*, Vol. 21, Iss. 4, pp. 1455-1508. ISSN 0893-9454. DOI: <https://doi.org/10.1093/rfs/hhm014>

HASSANI, Hossein, Saeed HERAVI a Anatoly ZHIGLJAVSKY, 2009. Forecasting European industrial production with singular spectrum analysis. *International Journal of Forecasting*, Vol. 25, Iss. 1, pp. 103-118. ISSN 0169-2070. DOI: <https://doi.org/10.1016/j.ijforecast.2008.09.007>

HASSANI, Hossein, Saeed HERAVI a Anatoly ZHIGLJAVSKY, 2013. Forecasting UK Industrial Production with Multivariate Singular Spectrum Analysis. *Journal of Forecasting*, Vol. 32, Iss. 5, pp. 395-408. ISSN 0277-6693. DOI: <https://doi.org/10.1002/for.2244>

HODSON, Timothy O., 2022. Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not. *Geoscientific Model Development*, Vol. 15, Iss. 14, pp. 5481-5487. ISSN 1991-959X. DOI: <https://doi.org/10.5194/gmd-15-5481-2022>

HYLLEBERG, Svend, Robert F. ENGLE, Clive W. J. GRANGER a Byung Sam YOO, 1990. Seasonal integration and cointegration. *Journal of Econometrics*, Vol. 44, Iss. 1-2, pp. 215-238. ISSN 0304-4076. DOI: [https://doi.org/10.1016/0304-4076\(90\)90080-D](https://doi.org/10.1016/0304-4076(90)90080-D)

HYNDMAN, Rob J. a George ATHANASOPOULOS, 2021. Forecasting: Principles and Practice [online]. 3rd edition. Melbourne: OTexts. [cit. 2026-04-21]. Dostupné z: <https://otexts.com/fpp3/>

HYNDMAN, Rob J. a Yeasmin KHANDAKAR, 2008. Automatic Time Series Forecasting: The forecast Package for R. *Journal of Statistical*

Software, Vol. 27, Iss. 3, pp. 1-22. ISSN 1548-7660. DOI:
<https://doi.org/10.18637/jss.v027.i03>

INOUE, Atsushi a Lutz KILIAN, 2005. In-Sample or Out-of-Sample Tests of Predictability: Which One Should We Use? *Econometric Reviews*, Vol. 23, Iss. 4, pp. 371-402. ISSN 0747-4938. DOI:
<https://doi.org/10.1081/ETC-200040785>

MAKRIDAKIS, Spyros a Michèle HIBON, 2000. The M3-Competition: results, conclusions and implications. *International Journal of Forecasting*, Vol. 16, Iss. 4, pp. 451-476. ISSN 0169-2070. DOI:
[https://doi.org/10.1016/S0169-2070\(00\)00057-1](https://doi.org/10.1016/S0169-2070(00)00057-1)

MARCHETTI, Domenico J. a Giuseppe PARIGI, 2000. Energy consumption, survey data and the prediction of industrial production in Italy: a comparison and combination of different models. *Journal of Forecasting*, Vol. 19, Iss. 5, pp. 419-440. ISSN 0277-6693. DOI:
[https://doi.org/10.1002/1099-131X\(200009\)19:5<419::AID-FOR749>3.0.CO;2-J](https://doi.org/10.1002/1099-131X(200009)19:5<419::AID-FOR749>3.0.CO;2-J)

OSBORN, Denise R., Saeed HERAVI a C. R. BIRCHENHALL, 1999. Seasonal unit roots and forecasts of two-digit European industrial production. *International Journal of Forecasting*, Vol. 15, Iss. 1, pp. 27-47. ISSN 0169-2070. DOI: [https://doi.org/10.1016/S0169-2070\(98\)00055-7](https://doi.org/10.1016/S0169-2070(98)00055-7)

RAPACH, David E. a Mark E. WOHR, 2006. In-sample vs. out-of-sample tests of stock return predictability in the context of data mining. *Journal of Empirical Finance*, Vol. 13, Iss. 2, pp. 231-247. ISSN 0927-5398. DOI:
<https://doi.org/10.1016/j.jempfin.2005.08.001>

ROSSI, Barbara, 2013. Advances in Forecasting under Instability. In: Graham ELLIOTT a Allan TIMMERMAN. *Handbook of Economic Forecasting*, Vol. 2, Part B. Amsterdam: Elsevier, pp. 1203-1324. ISBN 978-0-444-62731-5. DOI: <https://doi.org/10.1016/B978-0-444-62731-5.00021-X>

POUŽITÉ ZDROJE

TASHMAN, Leonard J., 2000. Out-of-sample tests of forecasting accuracy: an analysis and review. *International Journal of Forecasting*, Vol. 16, Iss. 4, pp. 437-450. ISSN 0169-2070. DOI: [https://doi.org/10.1016/S0169-2070\(00\)00065-0](https://doi.org/10.1016/S0169-2070(00)00065-0)

THURY, Gerhard a Stephen F. WITT, 1998. Forecasting industrial production using structural time series models. *Omega*, Vol. 26, Iss. 6, pp. 751-767. ISSN 0305-0483. DOI: [https://doi.org/10.1016/S0305-0483\(98\)00024-3](https://doi.org/10.1016/S0305-0483(98)00024-3)

Příloha A Alternativní transformace a modely

Tab. 19: Rozdíly RMSE mezi log a non-log transformací – ČR, auto.arima, deterministická sezónnost, rozšiřující se okno

Metoda	h = 1		h = 3		h = 6	
	IS	OOS	IS	OOS	IS	OOS
Rozšiřující se okno	-0,13	-0,12	-0,02	-0,05	+0,11	+0,01
Posuvné okno	-0,03	+0,03	-0,04	+0,06	+0,04	+0,20

Poznámka: Hodnoty v tabulce představují rozdíl mezi log a non-log RMSE na původní škále (po zpětné exponenciaci predikcí z log škály). Záporné hodnoty znamenají, že log transformace zlepšuje predikce, kladné že zhoršuje.

Zdroj: vlastní zpracování

Tab. 20: RMSE ARIMA a VAR modelů – ČR, auto.arima, deterministická sezónnost, rozšiřující se okno

Model	h = 1	h = 3	h = 6	Průměr
Auto.arima+dummy	4,77	5,24	6,45	5,49
VAR(IPP_CR, IPP_GER)	4,87	5,30	6,31	5,49
VAR(IPP_CR, IPP_SK)	4,75	5,27	6,34	5,46
VAR(IPP_CR, IPP_GER, IC_CR) endogenní	4,75	5,25	6,31	5,44
VARX(IPP_CR, IPP_GER) + IC_CR exogenní	4,91	5,25	6,27	5,48

Zdroj: vlastní zpracování

Příloha B MAE predikcí

Tab. 21: MAE po obdobích – auto.arima, deterministická sezónnost, rozšiřující se okno

Období	n	h = 1		h = 3		h = 6	
		IS	OOS	IS	OOS	IS	OOS
Česko							
Před krizí	45	2,20	2,86	2,31	2,85	2,74	3,33
Fin. krize	15	2,75	3,97	3,27	5,33	6,91	10,62
Mezikrizové	122	2,53	2,84	2,76	3,06	3,11	3,43
COVID	6	10,01	11,77	13,03	14,69	13,44	14,09
Po COVIDu	40	4,16	4,35	4,57	4,76	5,12	5,46
Celkem	228	2,96	3,42	3,30	3,78	3,94	4,54
Slovensko							
Před krizí	45	2,54	3,96	3,86	4,95	5,84	7,20
Fin. krize	15	3,60	4,31	6,45	7,55	10,91	13,25
Mezikrizové	122	3,79	4,12	4,37	4,66	4,27	4,53
COVID	6	11,12	14,24	16,78	19,41	14,80	15,20
Po COVIDu	40	4,00	4,20	5,24	6,07	6,74	7,59
Celkem	228	3,76	4,38	4,90	5,55	5,73	6,43

Zdroj: vlastní zpracování

Tab. 22: MAE po obdobích – auto.arima, stochastická sezónnost, rozšiřující se okno

	h = 1			h = 3		h = 6	
Období	n	IS	OOS	IS	OOS	IS	OOS
Česko							
Před krizí	45	2,48	2,90	2,52	2,88	3,01	3,35
Fin. krize	15	3,43	4,30	4,71	5,64	8,63	10,12
Mezikrizové	122	2,28	2,44	2,51	2,60	3,11	3,30
COVID	6	12,56	14,81	14,96	17,71	14,06	14,42
Po COVIDu	40	3,69	4,31	4,22	4,25	4,21	5,59
Celkem	228	2,91	3,31	3,29	3,55	3,95	4,48
Slovensko							
Před krizí	45	2,69	3,70	3,95	4,93	5,53	6,89
Fin. krize	15	3,85	5,47	7,62	11,48	11,55	21,81
Mezikrizové	122	3,09	3,21	3,15	3,44	3,62	4,05
COVID	6	14,09	15,40	20,43	22,55	13,30	14,23
Po COVIDu	40	3,85	4,00	5,92	6,21	7,77	8,13
Celkem	228	3,48	3,92	4,55	5,25	5,50	6,76

Zdroj: vlastní zpracování

Tab. 23: MAE po obdobích – ARIMA(0,1,1)+dummy, rozšiřující se okno

Období	n	h = 1		h = 3		h = 6	
		IS	OOS	IS	OOS	IS	OOS
Česko							
Před krizí	45	2,58	2,98	2,57	2,90	3,75	4,12
Fin. krize	15	3,21	3,64	4,23	4,77	6,99	7,60
Mezikrizové	122	2,83	3,02	2,79	2,96	3,36	3,51
COVID	6	10,37	11,29	13,44	14,24	13,47	13,80
Po COVIDu	40	4,16	4,35	4,57	4,76	5,12	5,29
Celkem	228	3,24	3,51	3,44	3,68	4,26	4,49
Slovensko							
Před krizí	45	2,90	3,53	4,34	4,99	7,39	8,16
Fin. krize	15	3,76	4,29	6,12	6,75	10,92	11,68
Mezikrizové	122	3,78	4,06	4,33	4,60	4,32	4,53
COVID	6	13,06	14,23	18,59	19,41	14,92	15,20
Po COVIDu	40	3,86	4,05	5,17	5,40	7,17	7,35
Celkem	228	3,86	4,24	4,98	5,35	6,11	6,45

Zdroj: vlastní zpracování

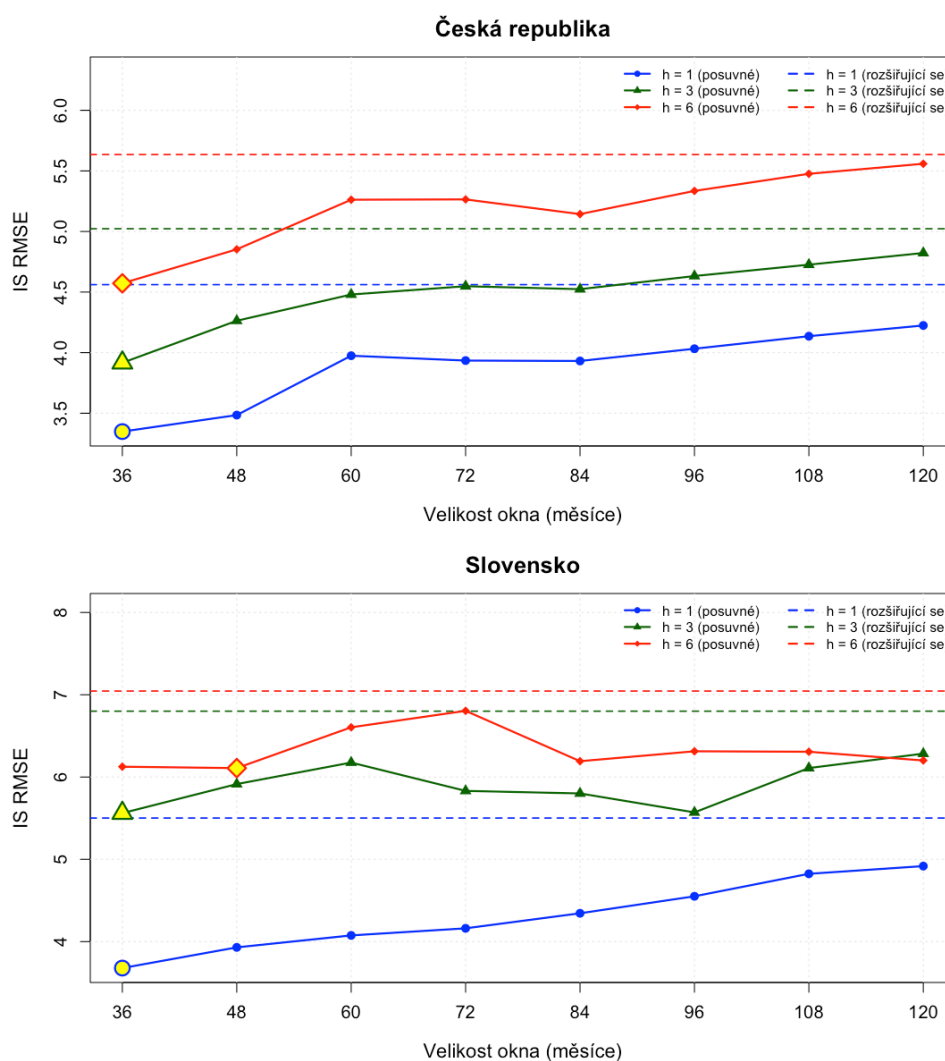
Tab. 24: MAE po obdobích – SARIMA(0,1,1)(0,1,1), rozšiřující se okno

Období	n	h = 1		h = 3		h = 6	
		IS	OOS	IS	OOS	IS	OOS
Česko							
Před krizí	45	2,61	2,82	2,55	2,73	3,00	3,14
Fin. krize	15	3,74	4,29	5,59	6,22	9,09	10,12
Mezikrizové	122	2,74	2,82	2,76	2,81	3,20	3,29
COVID	6	11,86	12,60	15,68	16,37	15,57	15,56
Po COVIDu	40	4,15	4,22	4,41	4,49	5,12	5,23
Celkem	228	3,27	3,42	3,54	3,68	4,24	4,40
Slovensko							
Před krizí	45	2,93	3,42	4,12	4,64	6,13	6,87
Fin. krize	15	4,08	6,01	8,18	12,86	14,71	23,86
Mezikrizové	122	3,11	3,16	3,35	3,39	3,65	3,68
COVID	6	14,57	15,27	22,32	22,36	13,50	13,68
Po COVIDu	40	3,92	3,97	5,98	6,11	8,24	8,34
Celkem	228	3,58	3,86	4,79	5,24	5,93	6,71

Zdroj: vlastní zpracování

Příloha C Analýza in-sample vzorku

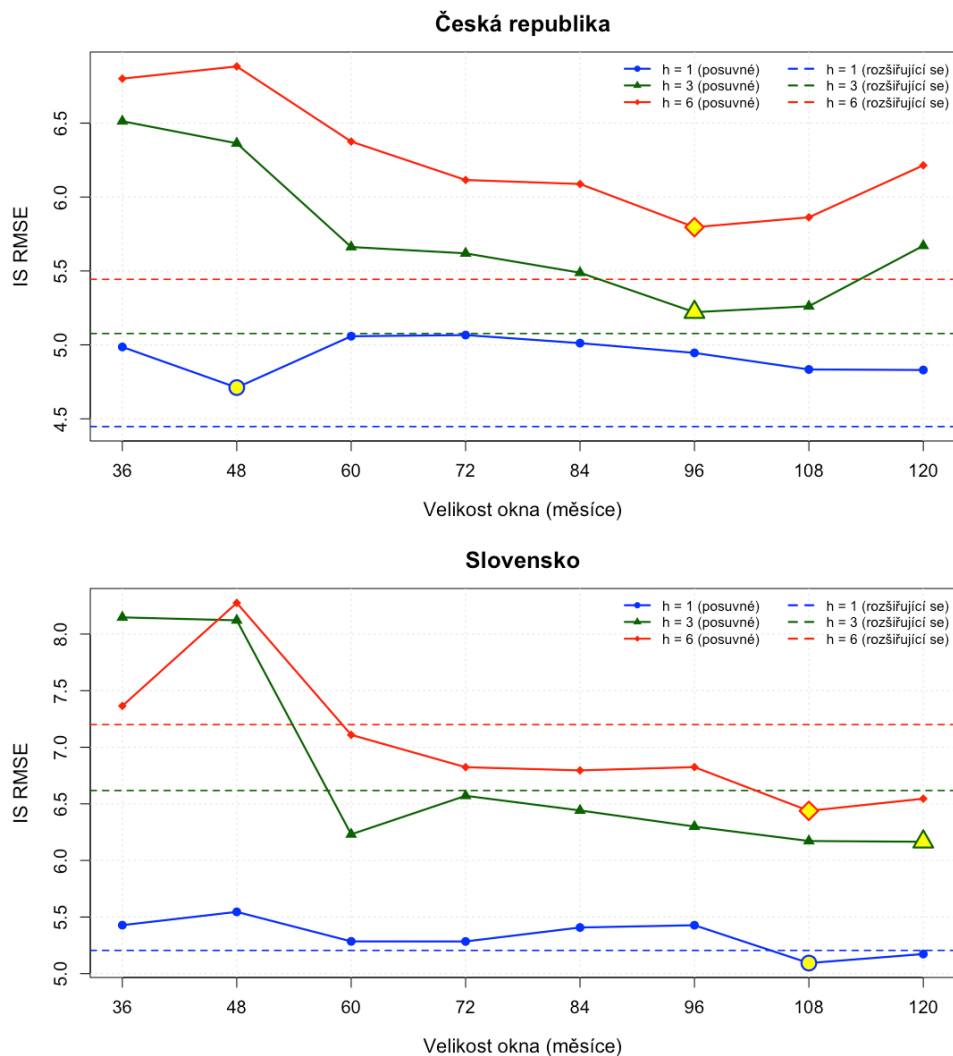
Obr. 18: Citlivostní analýza pro IS RMSE – auto.arima, deterministická sezónnost



Poznámka: Plné čáry zobrazují posuvná okna o různých délkách, přerušované čáry zobrazují referenční hodnoty pro rozšiřující se okno. Všechny hodnoty jsou vypočteny na společných pozorováních (od ledna 2010). Žlutě zvýrazněné body označují minimum pro daný horizont.

Zdroj: vlastní zpracování

Obr. 19: Citlivostní analýza pro IS RMSE – auto.arima, stochastická sezónnost



Poznámka: Plné čáry zobrazují posuvná okna o různých délkách, přerušované čáry zobrazují referenční hodnoty pro rozšiřující se okno. Všechny hodnoty jsou vypočteny na společných pozorováních (od ledna 2010). Žlutě zvýrazněné body označují minimum pro daný horizont.

Zdroj: vlastní zpracování

Tab. 25: Testy stacionarity – in-sample vzorek finanční krize

Země	Období	Úrovně		Diference	
		KPSS	ADF	KPSS	ADF
ČR	2000–2009	0,010	0,957	> 0,100	< 0,010
SK	2000–2009	0,010	0,358	> 0,100	< 0,010

Poznámka: IS vzorek zahrnuje tréninkový vzorek rozšířený o cílové krizové období (říjen 2008 – prosinec 2009). $n = 120$ pozorování. Hodnoty < 0,010 a > 0,100 reflektují hraniční hodnoty tabulky testů v R.

Zdroj: vlastní zpracování

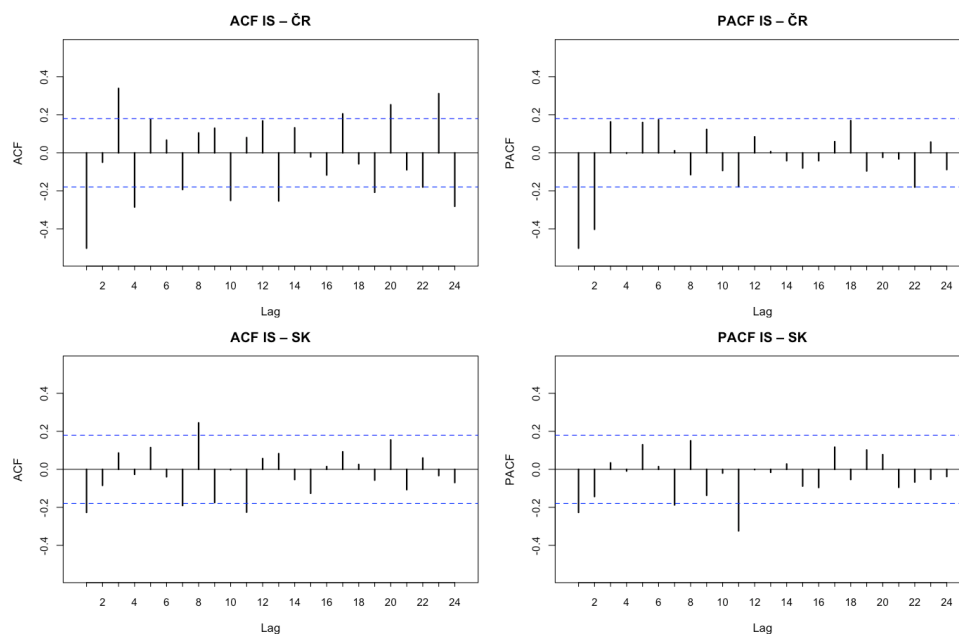
Tab. 26: Testy stacionarity – in-sample vzorek pro rok 2018

Země	Období	Po $D = 1$		Po $d = 1 + D = 1$	
		KPSS	ADF	KPSS	ADF
ČR	2010–2018	0,079	0,130	> 0,100	< 0,010
SK	2010–2018	0,032	0,139	> 0,100	< 0,010

Poznámka: IS vzorek zahrnuje tréninkový vzorek rozšířený o cílové období roku 2018. $n = 108$ pozorování. Hodnoty < 0,010 a > 0,100 reflektují hraniční hodnoty tabulky testů v R.

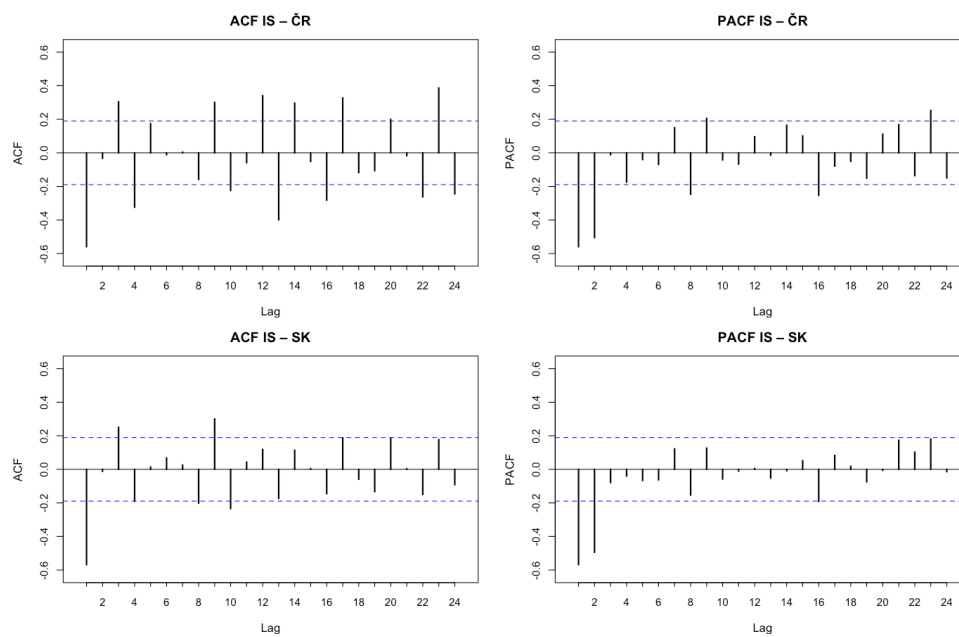
Zdroj: vlastní zpracování

Obr. 20: ACF a PACF prvních diferencí IPP po odstranění deterministické sezónní složky, IS vzorek 2000–2009



Zdroj: vlastní zpracování

Obr. 21: ACF a PACF IPP po první i sezónní diferenci, IS vzorek 2010–2018



Zdroj: vlastní zpracování