



ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE

FAKULTA DOPRAVNÍ

**Predikce vývoje cargo přepravy na vybraných
trasách**

Diplomová práce

Bc. Lenka Mráčková

Vedoucí práce: Ing. Eva Endrizalová, Ph.D.

Praha, 2023



K621.....Ústav letecké dopravy

ZADÁNÍ DIPLOMOVÉ PRÁCE

(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení studenta (včetně titulů):

Bc. Lenka Mráčková

Studijní program (obor/specializace) studenta:

navazující magisterský – PL – Provoz a řízení letecké dopravy

Název tématu (česky): **Predikce vývoje cargo přepravy na vybraných trasách**

Název tématu (anglicky): Cargo Traffic Forecast on Selected Routes

Zásady pro vypracování

Při zpracování diplomové práce se řiďte následujícími pokyny:

- Cílem práce je pomocí vhodné metody provést předpověď poptávky po přepravě zboží letecky a kvantifikovat nárůst vytižení provozovaných typů letadel a počet dodatečných letů na vybraných trasách, případně pokles vytižení a snížení počtu letů.
- Provedte rešerši odborných zdrojů, které se zabývají předpovědí poptávky po leteckém cargo, provedte analýzu metod a přístupu.
- Vytvořte datovou sadu pro jednotlivé letištní páry, pomocí které bude vyčíslen objem a charakter provozu (eurostat), vytvořte skupiny spojení podle určitého typu růstu/poklesu výkonů.
- Vyberte letištní páry a provedte predikci vývoje na těchto trasách.
- Kvantifikujte nárůst nebo pokles vytižení pomocí parametrů Payload (ACTK dle typů provozovaných letadel, počet skutečně přepravených tun).
- Popište metodiku, diskutujte výsledky a vyvodte závěry provedeného výzkumu.



- Rozsah grafických prací: dle pokynů vedoucího diplomové práce
- Rozsah průvodní zprávy: minimálně 55 stran textu (včetně obrázků, grafů a tabulek, které jsou součástí průvodní zprávy)
- Seznam odborné literatury: R. Totamane a kol.: Air cargo demand modeling and prediction, 2014
M. A. F. Loaiza a kol.: Modeling demand for air cargo in the colombian context, 2017
Eurostat: Detailed freight and mail air transport by reporting countries and routes, online database

Vedoucí diplomové práce: **Ing. Eva Endrizalová, Ph.D.**

Datum zadání diplomové práce: **15. července 2022**
(datum prvního zadání této práce, které musí být nejpozději 10 měsíců před datem prvního předpokládaného odevzdání této práce vyplývajícího ze standardní doby studia)

Datum odevzdání diplomové práce: **30. listopadu 2023**
a) datum prvního předpokládaného odevzdání práce vyplývající ze standardní doby studia a z doporučeného časového plánu studia
b) v případě odkladu odevzdání práce následující datum odevzdání práce vyplývající z doporučeného časového plánu studia

doc. Ing. Jakub Kraus, Ph.D.
vedoucí
Ústavu letecké dopravy



prof. Ing. Ondřej Příbyl, Ph.D.
děkan fakulty

Potvrzuji převzetí zadání diplomové práce.

Bc. Lenka Mráčková
jméno a podpis studenta

V Praze dne..... 16. května 2023

Poděkování

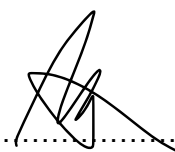
Tímto bych ráda poděkovala Ing. Evě Endrizalové, Ph.D. za její odborné rady, ochotu a konzultace, které mi v rámci mé diplomové práce poskytla. Dále bych velmi ráda poděkovala své rodině, příteli a přátelům za cennou podporu při celém studiu. Na závěr děkuji sobě, že jsem to nevzdala.

Čestné prohlášení

Prohlašuji, že jsem diplomovou práci s názvem Predikce vývoje cargo přepravy na vybraných trasách vypracovala samostatně a použila k tomu úplný výčet citací použitých pramenů, které uvádím v seznamu přiloženém k diplomové práci.

Nemám závažný důvod proti užití tohoto školního díla ve smyslu §60 Zákona č.121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon).

V Praze dne 30. listopadu 2023

A handwritten signature in black ink, consisting of several overlapping loops and a long horizontal stroke extending to the right, positioned above a dotted line.

Podpis

Abstrakt

Cílem diplomové práce je pomocí vhodné metody provést předpověď poptávky po přepravě zboží letecky a kvantifikovat nárůst vytížení provozovaných typů letadel a počet dodatečných letů na vybraných trasách, případně pokles vytížení a snížení počtu letů. Byly zvoleny trasy mezi Německem a Čínou. Po rešerši prediktivních metod byl jako hlavní nástroj zvolen statistický model ARIMA, který byl následně rozšířen o model GLM, vytvářející tak hybridní model ARIMA-GLM. Tento model byl použit pro predikci přepravených tun nákladu, s přihlédnutím k socioekonomickým proměnným pro roky 2023, 2024

Výsledky ukázaly, že kvůli současné stagnaci a pouze mírnému nárůstu v letecké dopravě prováděná kvantifikace nárůstu vytížení a počtu letů není žádoucí. Je předpokládáno, že mírný nárůst vytížení lze zvládnout efektivnějším využitím kapacity stávajících letů. Vzhledem k tomu, že predikce se týkají budoucích období let 2023 a 2024, nebylo možné je ověřit, což představuje omezení této práce.

Klíčová slova: predikce, ARIMA, letecký náklad

Abstract

The aim of the thesis is to forecast the demand for air freight transport using a suitable method and to quantify the increase in the utilisation of the aircraft types operated and the number of additional flights on selected routes, or the decrease in utilisation and the reduction in the number of flights. Routes between Germany and China were selected. After a search of predictive methods, the ARIMA statistical model was chosen as the main tool, which was then extended with a GLM model, thus creating a hybrid ARIMA-GLM model. This model was used to predict tons of cargo transported, taking into account socio-economic variables for the years 2023, 2024

The results showed that due to the current stagnation and only slight growth in air traffic, the quantification of the increase in load and number of flights carried out is not desirable. It is assumed that a moderate increase in load can be managed by a more efficient use of the capacity of existing flights. As the forecasts are for the future periods 2023 and 2024, it was not possible to verify them, which is a limitation of this work.

Keywords: prediction, ARIMA, air cargo

Seznam zkratek

Zkratka	Význam
ACF	Autokorelační funkce
ADF	Dickey-Fullerův test
AMS	Letiště Amsterdam Schiphol
ANOVA	Analýza rozptylu
AR	Autoregresní model
ARDL	Autoregresní model rozložených časových zpoždění
ARIMAX	Autoregresní integrovaný klouzavý průměr
ATK	Tunokilometry
BIC	Bayesovo informační kritérium
BPS	Statistický úřad Indonésie
CAN	Mezinárodní letiště Kanton Paj-jün
CGN	Letiště Kolín/Bonn
DANE	Správní oddělení národní statistiky
EMA	Letiště East Midlands
EUR	Euro
FRA	Letiště Frankfurt nad Mohanem
FTK	Nákladní tunokilometry
GBRT	Gradient Boosted Regression Trees
GLM	Generalizované lineární modely
HDP	Hrubý domácí produkt
HKG	Mezinárodní letiště Hongkong
IATA	Mezinárodní letecká dopravní asociace
ICAO	Mezinárodní organizace civilního letectví
IoT	Internet věcí
LEJ	Letiště Lipsko/Halle
LGG	Letiště Liege
LF	Faktor zatížení
LUX	Mezinárodní letiště Lucemburk
MAD	Střední absolutní odchylka
MAE	Střední absolutní chyba odhadu
MAPE	Střední absolutní procentní chyba odhadu
MA	Klouzavý průměr
MLAT	Multilaterace
MLR	Vícenásobná lineární regrese
MSE	Střední kvadratická chyba
NN	Neurální síť
PACF	Parciální autokorelační funkce
PEK	Mezinárodní letiště Peking
PSO	Optimalizace rojem částic
PVG	Mezinárodní letiště Šanghaj Pchu-tung
PZI	Přímé zahraniční investice
RMSE	Kořen průměrné kvadratické chyby

RNN	Rekurentní neuronové sítě
RTK	Příjmové tunokilometry
SDF	Letiště Louisville
SRM	Minimalizace strukturálního rizika
SVM	Stroj vektorové podpory
SVR	Podpůrná vektorová regrese
SZX	Mezinárodní letiště Šen-čen Pao-an
TSR	Regresní model časových řad
USA	Spojené státy americké
USD	Americký dolar
WIMA	Vážený klouzavý průměr
WMA	Algoritmus vážené většiny

Seznam obrázků

- Obrázek 1 - Rozdělení metod predikce, vlastní zpracování [2]
- Obrázek 2 - Složky časových řad, vlastní zpracování [2]
- Obrázek 3 - Ukázka původních nezpracovaných dat převáženého nákladu mezi Německem a partnerskými letišti v Číně [27]
- Obrázek 4 - Skript z programu R pro ARIMA model
- Obrázek 5 - Dekompozice časové řady FRA-PEK
- Obrázek 6 - ACF na trase FRA-PEK
- Obrázek 7 - Rozšířený Dickey-Fullerův test pro trasu FRA-PEK
- Obrázek 8 - ARIMA model pro trasu FRA-PEK
- Obrázek 9 - Dekompozice časové řady FRA-CAN.
- Obrázek 10 - ACF na trase FRA-CAN
- Obrázek 11 - ARIMA model pro trasu FRA-PEK
- Obrázek 12 - Dekompozice časové řady FRA-PVG
- Obrázek 13 - ACF na trase FRA-PVG
- Obrázek 14 - ARIMA model pro trasu FRA-PVG
- Obrázek 15 – Dekompozice časové řady CGN-CAN
- Obrázek 16 - ACF na trase CGN-CAN
- Obrázek 17 - - Rozšířený Dickey-Fullerův test pro trasu CGN-CAN
- Obrázek 18- ARIMA model pro trasu CGN-CAN
- Obrázek 19 - Dekompozice časové řady CGN-SZX
- Obrázek 20 - ACF na trase CGN-SZX
- Obrázek 21 - ARIMA model pro trasu CGN-SZX
- Obrázek 22 - Dekompozice časové řady CGN-PVG
- Obrázek 23 - ACF na trase CGN-PVG
- Obrázek 24 - ARIMA model pro trasu CGN-PVG
- Obrázek 25 - Dekompozice časové řady všech tras Německo-Čína
- Obrázek 26 - ACF na trasách Německo-Čína
- Obrázek 27 - ARIMA model pro trasy Německo-Čína
- Obrázek 28 - Dekompozice časové řady HDP
- Obrázek 29 - ACF u časové řady HDP
- Obrázek 30 - ARIMA model pro HDP
- Obrázek 31 - Dekompozice časové řady směnného kurzu.
- Obrázek 32 - ACF u časové řady směnného kurzu

Obrázek 33 - ARIMA model pro směnný kurz
Obrázek 34 - Dekompozice časové řady salda
Obrázek 35 - ACF u časové řady salda
Obrázek 36 - ARIMA model pro saldo zahraničního obchodu
Obrázek 37 - Skript z programu R pro GLM model
Obrázek 38 - Model GLM pro FRA-PEK
Obrázek 39 - Model GLM pro FRA-CAN
Obrázek 40 - GLM model pro FRA-PVG
Obrázek 41 - GLM model pro CGN-CAN
Obrázek 42 - GLM model pro CGN-SZX
Obrázek 43 - GLM model pro CGN-PVG
Obrázek 44 - Predikce kombinací modelů ARIMA+GLM pro FRA-PEK
Obrázek 45 - Proměnné pro ARIMA+GLM pro FRA-PEK
Obrázek 46 - Hodnoty predikce ARIMA+GLM pro FRA-PEK
Obrázek 47 - Predikce kombinací modelů ARIMA+GLM pro FRA-CAN
Obrázek 48 - Proměnné pro ARIMA+GLM pro FRA-CAN
Obrázek 49 - Hodnoty predikce ARIMA+GLM pro FRA-CAN
Obrázek 50 - Predikce kombinací modelů ARIMA+GLM pro FRA-PVG
Obrázek 51 - Proměnné pro ARIMA+GLM pro FRA-PVG
Obrázek 52 - Hodnoty predikce ARIMA+GLM pro FRA-PVG
Obrázek 53 - Predikce kombinací modelů ARIMA+GLM pro CGN-CAN
Obrázek 54 - Proměnné pro ARIMA+GLM pro CGN-CAN
Obrázek 55 - Hodnoty predikce ARIMA+GLM pro CGN-CAN
Obrázek 56 - Predikce kombinací modelů ARIMA+GLM pro CGN-SZX
Obrázek 57 - Proměnné pro ARIMA+GLM pro CGN-SZX
Obrázek 58 - Hodnoty predikce ARIMA+GLM pro CGN-SZX
Obrázek 59 - Predikce kombinací modelů ARIMA+GLM pro CGN-PVG
Obrázek 60 - Proměnné pro ARIMA+GLM pro CGN-PVG
Obrázek 61 - Hodnoty predikce ARIMA+GLM pro CGN-PVG

Seznam tabulek

Tabulka 1 - Shrnutí rešerše článků na téma predikce (vlastní zpracování)

Tabulka 2 - Letištní páry vyselektované pro predikci

Tabulka 3 - Predikované hodnoty HDP

Tabulka 4 - Historické hodnoty HDP

Tabulka 5 - Historické hodnoty směnného kurzu

Tabulka 6 - Predikované hodnoty směnného kurzu

Tabulka 7 - Predikované hodnoty salda zahraničního obchodu

Tabulka 8 - Rozptyl pro trasu FRA-PEK

Tabulka 9 - Počet letů na trase FRA-PEK

Tabulka 10 - Výsledné navýšení/snížení letů pro trasu FRA-PEK

Tabulka 11 - Rozptyl pro trasu FRA-CAN

Tabulka 12 - Výsledné navýšení/snížení letů pro trasu FRA-CAN

Tabulka 13 - Rozptyl pro trasu FRA-PVG

Tabulka 14 - Výsledné navýšení/snížení letů pro trasu FRA-PVG

Tabulka 15 - Rozptyl pro trasu CGN-CAN

Tabulka 16 - Výsledné navýšení/snížení letů pro trasu CGN-CAN

Obsah

ÚVOD	13
1. TEORIE METOD PREDIKCE	14
1.1 KVALITATIVNÍ METODY	18
1.2 KVANTITATIVNÍ METODY	19
2. REŠERŠE ODBORNÝCH PRACÍ ZABÝVAJÍCÍCH SE PREDIKCÍ POPTÁVKY PO LETECKÉ NÁKLADNÍ DOPRAVĚ	27
3. METODIKA A VÝSLEDKY	45
3.1 LETECKÝ TRH ZVOLENÝ PRO PREDIKCI NÁKLADU	45
3.2 DATA UŽITÁ PŘI PREDIKCI NÁKLADU	45
3.3 VOLBA LETIŠTNÍCH PÁRŮ	47
3.4 PREDIKCE VÝVOJE TRAS NA ZÁKLADĚ STATISTICKÉHO MODELU ARIMA	48
3.5 GLM MODEL HODNOTÍCÍ RELEVANTNOST PROMĚNNÝCH PRO PREDIKCI	68
3.6 PREDIKCE VÝVOJE TRAS NA ZÁKLADĚ KOMBINACE MODELŮ ARIMA + GLM	73
3.7 KVANTIFIKACE NÁRŮSTU A POKLESU VYTIŽENÍ PROVOZOVANÝCH TYPŮ LETADEL	82
DISKUZE	86
ZÁVĚR	87
SEZNAM POUŽITÉ LITERATURY	88
PŘÍLOHY	93

Úvod

Letecké cargo představuje klíčovou součástí světového obchodu. S rostoucím globálním obchodem a potřebou rychlé a efektivní přepravy zboží hraje letecká nákladní doprava nezastupitelnou roli v mezinárodním obchodním systému. Je to sektor charakterizovaný vysokou dynamikou a schopností rychle reagovat na měnící se tržní požadavky.

Vzhledem k rostoucí globalizaci a neustálým změnám v mezinárodním obchodě je důležité předvídat vývoj poptávky po letecké přepravě zboží. Tato diplomová práce se zaměřuje na analýzu a predikci cargo provozu na vybraných leteckých trasách, konkrétně mezi Německem a Čínou. Hlavním cílem je pomocí vhodných metod předpovědět budoucí poptávku po přepravě zboží letecky a kvantifikovat možný nárůst nebo pokles využití provozovaných letadel.

Covidová krize výrazně ovlivnila letecké cargo, přičemž způsobila významné změny v poptávce a v provozních kapacitách. Přerušování dodavatelských řetězců, omezení mezinárodního cestování a uzavření hranic vedly k významným výkyvům v nabídce a poptávce. V některých případech došlo k prudkému nárůstu poptávky po přepravě zdravotnického materiálu a jiných nezbytných zboží. Během pandemie došlo také k výrazným problémům s kapacitami. Mnoho leteckých společností muselo redukovat nebo úplně zastavit své pravidelné osobní lety, což vedlo k poklesu dostupné kapacity pro náklad. Tato situace zdůraznila potřebu pružné a adaptabilní infrastruktury v leteckém cargo sektoru.

Pokud jde o aktuální vývoj letecké nákladní dopravy ve světě, celkově nákladní lety mírně vzrostly, ale zůstávají pod předpandemickou úrovní, přičemž poměr nákladu přepravovaného nákladními a osobními letadly se vrátil na úroveň podobnou té před pandemií

Letecké cargo se vyznačuje několika klíčovými vlastnostmi, včetně schopnosti přepravovat zboží rychle, bezpečně a na velké vzdálenosti. Je rozděleno do několika kategorií, včetně standardního cargo, nebezpečného nákladu, chlazeného zboží a zvláštního cargo, jako jsou živé zvířata nebo cenné předměty.

1. Teorie metod predikce

Teoretická část o metodách predikce je primárně sepsána z dvou zdrojů a tím je kniha *Flying off Course* od autora Rigase Doganise [1] a kniha *Introduction to Air Transport Economics: From Theory to Applications* od Bijana Vasigha. [2] Zdrojů na téma metod predikce v cargo odvětví není velké množství a tyto knihy poskytují kvalitní základní přehled o problematice.

Predikce poptávky v tak dynamickém a na vnější faktory citlivém odvětví, jakým je letectví, je nesmírně obtížný úkol. [3] Pro pochopení predikce poptávky je nejprve nutné vysvětlit samotný pojem poptávka. Jedná se o množství zboží, které je kupující ochoten koupit za určitou cenu, v určitý čas na určitém místě. Dělíme ji na individuální a tržní. Individuální poptávka je poptávkou jednoho spotřebitele. Tržní poptávka je poptávka všech spotřebitelů po daném výrobku nebo službě. [4] Zákon poptávky znamená, že když cena produktu roste, spotřebitelé budou za stejných podmínek kupovat v určitém časovém období méně, a naopak kupovat více, když cena klesá. Důležitým ekonomickým principem je také elasticita. Formální definicí je procentní změna závislé proměnné vyplývající ze změny ceny o jedno procento neboli jak se změní spotřebiteli poptávané množství, když se cena změní o jedno procento. Můžeme měřit buď bodovou nebo obloukovou elasticitu. Bodová měří elasticitu funkce při určité hodnotě, zatímco oblouková měří elasticitu funkce v rozsahu hodnot. Podle výsledné hodnoty považujeme elasticitu za neelastickou, elastickou nebo jednotkově elastickou. Možným příkladem nepružné elasticity je poptávka po odeslání zboží letecky na poslední chvíli. Zvýšení ceny není tak podstatné ve chvíli, kdy je potřeba zboží převést, tudíž je zákazník ochoten zaplatit vyšší cenu. Poptávka po letecké dopravě má své specifické vlastnosti, které je třeba brát v potaz. Můžeme zmínit například cykličnost, sezónnost, ovlivnitelnost svátky či jednorázové směrové toky. [2] S poptávkou je potřeba vysvětlit i nabídku. Nabídka je služba nebo produkt, které výrobci plánují prodat v určitém časovém období při určité ceně. Na trhu se schází i nabídka i poptávka a v situaci, kdy na trhu není ani nedostatek ani přebytek statku, dochází k tržní rovnováze. Jedná se o situaci, kdy se tedy nabídka rovná poptávce. [4]

Predikce patří k nejkritičtějším oblastem řízení letecké společnosti. Předpověď poptávky je nutná k přizpůsobení nabídky služeb potřebné k uspokojení zákaznických potřeb. [1] Vzhledem k tomu, že poptávka po letecké dopravě není monolitická a pro každý let se liší, je třeba používat sofistikované predikční nástroje, které pomohou předpovědět velikost a charakter poptávky. Tento proces často zahrnuje stanovení míry rezervací pro každý let a také určení kombinace mezi zákazníky citlivými na cenu a zákazníky citlivými na čas. [2] Krátkodobá předpověď se pohybuje přibližně od 6 do 18 měsíců a je zahrnuta do provozního plánu a rozpočtu letecké společnosti na nynější a nadcházející finanční rok. Zároveň z ní

vychází taktická či operativní rozhodnutí, jako například plánování letadel, údržby, marketingové a prodejní kampaně atd. Z dlouhodobých predikcí vychází strategická rozhodnutí spojená s dlouhodobými plány a cíli společnosti, např. rozhodnutí o rozšíření flotily společnosti, otevírání nových tras či vstupy na nové trhy. Dlouhodobé predikce mohou být na 5 let, popřípadě i déle. [1] Krátkodobé predikce bývají přesnější než dlouhodobější predikce. [3]

Predikce se však neomezuje pouze na poptávku. Plánování lidských zdrojů, finančních zdrojů, rozvoje tras, rotace letadel a rozšiřování infrastruktury je založeno na určitém očekávání budoucích událostí. [3] Prognózy týkající se množství letů, požadavků na posádky, plánů výcviku, absencí a míry fluktuace zaměstnanců jsou pro letecké společnosti také velmi důležité a vzhledem k tomu, že projekty jsou analyzované po celou dobu své životnosti, musí prognózy zahrnovat i budoucí očekávané peněžní toky. [2] Pokud bude možné přesně předvídat objem nákladní dopravy v civilním letectví, sníží se také obtížnost plánování letového provozu a zlepší se efektivita přepravy nákladu. [5] Letecké společnosti mají různé obchodní strategie pro plánování a optimalizaci poptávky po leteckém nákladu, pro příklad můžeme uvést efektivní využití kapacity letecké sítě pomocí nových logistických koncepcí, posílení spolupráce s jinými leteckými nákladními společnostmi, kamionovými/vlakovými společnostmi a speditéry, či zavedení časově omezených služeb (např. garantované včasné dodání) atd. [6]

Je potřeba brát v potaz, že je predikce oblastí, ve které se nejčastěji chybí a kde je velká nejistota. Letecké společnosti využívají různé techniky predikce, z nichž má každá své výhody a nevýhody a žádná z nich nemůže zajistit konzistentní přesnost. Je však potřeba predikovat z již dříve zmíněných důvodů. Při predikci je ideálním postupem nejprve předpovědět celkovou poptávku na trhu a až následně predikovat vlastní podíl na této poptávce. Obecně letecké společnosti předpokládají, že přeprava nákladu poroste, pokud zůstanou ostatní proměnné konstantní. V realitě je však poptávka po leteckém nákladu ovlivněna velkým množstvím faktorů, které můžeme rozdělit na externí a interní. Externí faktory jsou velmi často nepředvídatelné, tudíž se podílí právě na již zmiňované nepřesnosti v predikci. Patří mezi ně například ekonomické klima, směnné kurzy či nenadálé události jako teroristické útoky atd. Mezi interní faktory patří například změna ve flotile a přechod z úzko na širokotrupé letouny, změna ve frekvencích letů či změna odletových časů. Ať jde o externí či interní faktory, letecká společnost musí zvládnout na tyto změny reagovat. [1] Pochopení povahy krátkodobého a dlouhodobého vztahu mezi poptávkou po leteckém nákladu a jejími determinanty je však velmi důležité, obzvláště pro vlády, letecké společnosti, letiště, logistické společnosti, plánovače dopravy a další zúčastněné strany. [7]

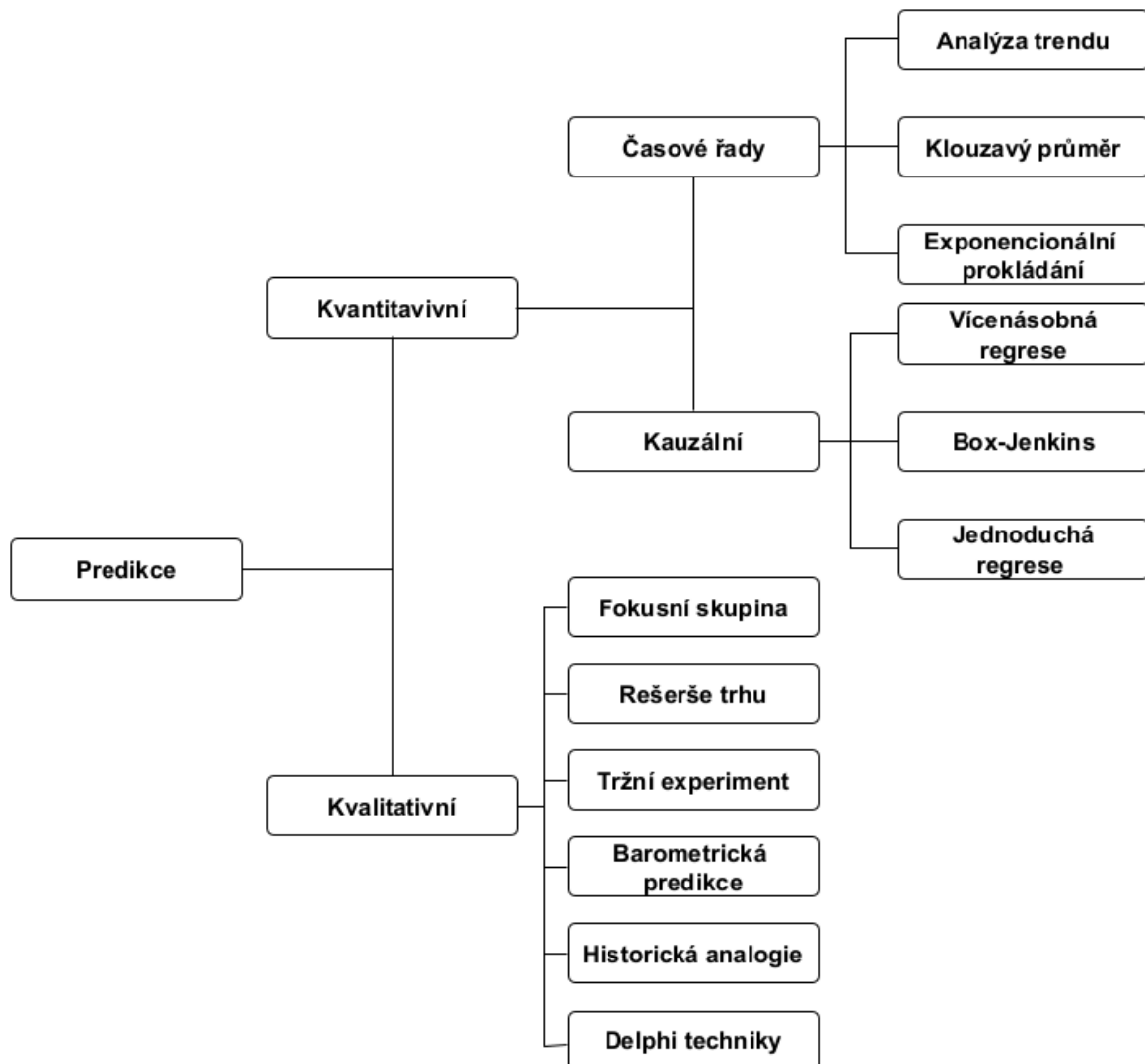
Obecně se dá říci, že předpověď cargo dopravy je méně úspěšná než předpověď poptávky cestujících v letecké dopravě. Faktory ovlivňující růst letecké nákladní dopravy jsou složitější a proměnlivější, než jsou faktory osobní přepravy. Také je potřeba brát v úvahu fakt, že přeprava nákladů bývá jednosměrná na rozdíl od přepravy cestujících, kteří mají tendenci vracet se do míst, odkud odlétají. Každá trasa disponuje množstvím komoditních sazeb nákladní dopravy, které bývají zpravidla méně stabilní než sazby pro cestující. [1] Část kapacity nákladní dopravy také vzniká jako vedlejší produkt kapacity osobní dopravy. Nákladní letadla však hrají v přepravě carga klíčovou roli, protože osobní letadla kvůli svým pevně stanoveným letovým řádům pro cestující, nejsou schopna úplně uspokojit poptávku po přepravě nákladu. V důsledku přepravy carga jak osobními, tak nákladními letadly často dochází k nadměrné nabídce kapacity, což vede k tlaku na snižování sazeb cargo dopravy. [6]

Kvůli těmto ovlivňujícím a proměnlivým faktorům se letecké společnosti při krátkodobé predikci cargo dopravy při výběru metody predikce obvykle spoléhají na projekci minulých trendů, kdy je následně upravují na základě průzkumů trhu a úsudku managementu. [1] Výjimkou, kterou je potřeba zmínit, byla pandemie Covid-19, neboť se letecké společnosti musely vypořádat s novou situací na trhu. Nákladní doprava tehdy měla problém s kapacitou, protože většina osobních letů byla pozastavena. Hlavním problémem byl tedy nedostatek letadel určených pro nákladní dopravu včetně břišní kapacity. [8] Jak bylo již řečeno, poptávka po pandemii se již vrací na své původní hodnoty, tudíž se i oblast predikce vrací k metodám, na které byla zvyklá před pandemií, samozřejmě pokud bereme v potaz přirozený vývoj predikčních technik v čase. [1]

Volba metody predikce se také odvíjí od toho, zda je zaváděna nová linka nebo jestli má společnost již s danou linkou zkušenosti. U linky na nové trase je potřebné znát, zda na lince operuje již konkurenční aerolinka či se zavádí úplně nová, dosud neobsluhovaná trasa. V takovém případě je predikce velmi náročná a je snadné dojít k chybě. Analýzy časových řad a regresní modely zde nejsou příliš užitečné, proto se občas tento problém řeší kombinací průzkumu trhu a úsudku manažerů anebo gravitačním modelem, který je jedním z kauzálních modelů vyvinutých právě pro predikci dopravy. Pokud se připojujeme na trasu, kde již operují jiné letecké společnosti, je možné použít k predikci historická data právě oněch společností, aby byla naše predikce přesnější. [1]

Klíčové dovednosti k predikci jsou především znalost leteckého průmyslu a schopnost aplikovat statistiku a ekonomické principy. Predikce se většinou provádí prostřednictvím softwaru a různých aplikací, pro příklad můžeme uvést například programy Microsoft Excel či SPSS. [2]

Metody predikce využívané leteckými společnostmi se primárně dělí do tří hlavní skupin: kvalitativních metod, projekcí časových řad a kauzálních či ekonometrických metod, přičemž projekce časových řad a ekonometrické či kauzální metody se dají zařadit pod kategorii kvantitativních metod. [1] Jedno z možných rozdělení je znázorněno na obrázku 1. Na závěr budou krátce představeny i poměrně nové metody strojového učení, které se už nyní čím dál tím více podílejí na predikci nákladní letecké dopravy. [3]



Obrázek 1 - Rozdělení metod predikce, vlastní zpracování [2]

1.1 Kvalitativní metody

Kvalitativní metody predikce jsou založeny především na názorech, průzkumech nebo přesvědčeních. Využívají statistické vztahy, které pomáhají predikovat budoucí události. Ačkoli by se mohlo zdát, že oproti kvantitativním prognózám, které jsou naopak založené na matematických vzorcích a statistice, budou predikce méně přesné, není tomu vždy tak.

Jednou z hlavních výhod kvalitativních predikcí je flexibilita. Predikce jsou snadno upravitelné tak, aby bylo možné včas rozpoznat neočekávané změny v ekonomice a datech. Na druhé straně mohou být kvalitativní prognózy obtížněji sledovatelné, a ne vždy je možné izolovat primární proměnnou, která způsobuje změny závislé proměnné. Dalším omezením je přesnost, neboť není možné spolehlivě určit, jak kvalitní predikce je. Kvalitativních metod je velké množství, pro lepší vysvětlení podstaty kvalitativních metod jsou zde zmíněny fokus skupiny, exekutivní posudek, rešerše trhu, tržní experiment a delphi techniky. [2]

Fokusní skupina je relativně neformální postup sběru dat, ale poměrně běžný při průzkumu trhu. Obvykle se jí účastní 8 až 12 osob, které diskutují o daném tématu za pomoci moderátora, který se snaží diskusi skupiny posouvat dopředu. Výzkumníci pozorují účastníky a jejich odpovědi, což jim umožňuje rychlý a relativně levný vhled do konkrétního výzkumného problému. Focus groups mohou být poměrně efektivní při hodnocení nových možností produktů a pomáhají určit cílové skupiny spotřebitelů, které jsou potřeba především pro plánování a marketing. [2]

Exekutivní posudek znamená predikci většinou od klíčových manažerů na základě jejich hluboké znalosti trhu a přehledu o ekonomické situaci, která by potenciálně mohla ovlivnit vývoj trasy v budoucnu. Výhody této metody jsou, že je poměrně rychlá a manažeři si mohou být lépe vědomi faktorů, které mohou ovlivnit budoucí poptávku na trase dříve, než bude možné tyto faktory zjistit pomocí číselných metod. [1]

Rešerše trhu jsou postavené především na postojových a behaviorálních dotaznících cestujících. Na rozdíl od fokusních skupin se dotazníky pro rešerše trhu vyplňují spíše individuálně. [1] Jednou z hlavních výhod průzkumů trhu je, že jsou snadno použitelné a k interpretaci výsledků nevyžadují pokročilou teorii nebo ekonometrickou analýzu. Potenciálním nedostatkem průzkumů trhu je, že přesnost závisí na velikosti a reaktivitě vzorku. [2] Tato metoda bývá u predikce carga využívána poměrně často, především na tratích, kde není převáženo takové množství nákladu, protože zde není snadné využít ekonomické metody či projekce časových řad i kvůli nedostatečným či nepřesným číselným údajům. [1]

Jednou z nejnákladnějších metod kvalitativního prognózování je tržní experiment. Ten zahrnuje testování nových výrobních faktorů na jednom nebo více vybraných testovacích trzích. Tržní experimenty využívají reálné regiony, což může být riskantní, pokud spotřebitelé změnu nepřijmou, neboť je změna může trvale odradit od využívání produktů či služeb společnosti. Vzhledem k nákladům, které jsou s tím spojeny, se tržní experimenty zřídka používají k predikci poptávky. [2]

Delphi techniky jsou založeny na shromáždění predikcí a názorů nezávislého panelu odborných osob, které jsou schopné předpovídat budoucí trendy. Jedná se o iterativní proces, který zahrnuje zpětnou vazbu prostřednictvím řady kol, ze které nakonec vznikne konsenzuální predikce, která je právě i největším nedostatkem této metody. Technika je vhodná spíše pro souhrnnou předpověď růstu na hlavních trzích či v regionech než pro predikci jednotlivých tras.

Ačkoliv jsou kvalitativní metody v některých situacích vhodné k užití, obecně se v letectví pro predikci poptávky po letecké dopravě spíše využívají kvantitativní metody. Od roku 2014 přešla i IATA k využívání spíše ekonometrických modelů pro své predikce, stejně jako ICAO. [1]

1.2 Kvantitativní metody

Kvantitativní metody můžeme, jak již bylo řečeno, rozdělit na projekce časových řad a kauzální či ekonometrické metody. Tyto metody využívají statistické údaje k analýze a predikci budoucího chování konkrétních proměnných. Jednou z hlavních výhod kvantitativních predikcí je možnost provádět testy spolehlivosti pro určení přesnosti prognózy. Při analýze časových řad lze na základě testu spolehlivosti zvolit nejpřesnější metodu predikce. V regresní analýze jsou prognostici schopni poskytnout nejen pravděpodobnost přesnosti celkové prognózy, ale také spolehlivost jednotlivých proměnných v predikci. Mezi hlavní nevýhody naopak patří fakt, že historie není vždy správným předpovědním faktorem pro budoucnost a že predikce obvykle vyžadují rozsáhlý sběr a zpracování dat. Kvalita kvantitativní prognózy je zcela závislá na kvalitě použitých dat. V závislosti na tom, jak přesná a úplná jsou data, vychází buď zkreslující predikce nebo naopak taková, která dokonale modeluje realitu. Zavedení pokročilých statistických programů však sběr a zpracování dat postupem času zjednodušuje. [2]

Projekce časových řad:
Projekce časových řad jsou nejčastěji používané techniky předpovědi leteckých společností. [1] Analýza časových řad hledá v datech zákonitosti, zatímco regresní analýza předpokládá příčinný vztah mezi dvěma nebo více proměnnými. Údaje používané pro predikci představují pozorování určitých proměnných ve stejných časových intervalech a ve více časových bodech.

[2] Předpokládají, že cokoliv ovlivnilo projekci v minulosti, ji bude ovlivňovat stejným způsobem i v budoucnu. Nezávislou proměnnou je zde tedy čas, závislou proměnnou je provoz (traffic). V praxi používají letecké společnosti obvykle exponenciální nebo lineární časové řady, často se ve vývoji také vystřídají obě tyto varianty. Exponenciální bývají využívané častěji jak pro svou jednoduchost, tak proto, že historické trendy v leteckém provozu se často zdají být exponenciálními. Rozdíl mezi nimi je takový, že u lineární časové řady je změna objemu carga za každou časovou jednotku konstantní, ale celkový provoz roste, tudíž roční procentuální růst postupně klesá. U exponenciální časové řady je nárůst objemu carga při každé časové jednotce větší než v té předchozí. Každý další růst je konstantní procentuální hodnota, ale vždy z většího předchozího součtu. Ideálním počtem uplynulých let je 7-10 let předchozích dat a někteří prediktoři tvrdí, že by se nemělo předvídat na více let, než je polovina ročních dat, která jsou k dispozici. Někteří prediktoři věří, že by měla být větší váha přiřazena datům z nedávné doby oproti starším datům. Mezi modely, které tuto teorii využívají, patří například Brownův double exponential smoothing model či Holt-Wintersův model. Tato teorie je zatím více využívaná v oborech mimo letectví, začíná se na ni však klást čím dál tím větší důraz i zde. Predikce časových řad také nemohou předvídat a umožnit náhlé a významné změny podmínek nabídky nebo poptávky, jako je například připojení nízkonákladového dopravce na obsluhu trasy. Často je následně potřeba upravit predikce na základě úsudku vedoucích pracovníků nebo výsledků průzkumu trhu. [1]

Letecké společnosti v praxi na základě vlastních zkušeností většinou preferují jednu nebo dvě z již zmíněných metod, kterých se drží. Většinou využívají projekci časových řad jako počáteční bod pro prognózu především i proto, že pokud je dostatek statistických informací, jsou snadno využitelné. Jak již bylo zmíněno, u dlouhodobější predikcí se riziko chyby u predikce zvyšuje, protože na poptávku začíná působit více okolních faktorů. Problémem časových řad, které jsou zde rozebírány, je fakt, že jsou pouze funkcemi času = jak se mění doba, mění se poptávka. Vztahovat poptávku jen na čisté změny v čase je pro přesnější dlouhodobější predikci nedostatečné, protože nejsou zahrnuty právě ostatní faktory. Letecké společnosti tedy mohou po projekci časových řad upravit poptávku na základě průzkumu trhu či úsudku manažerů nebo mohou použít ekonometrické či kauzální techniky predikce. [1] Pozorování u časových řad lze provádět ročně, čtvrtletně, měsíčně, týdně, denně nebo každou hodinu. Údaje časové řady jsou složeny ze čtyř složek, jak je znázorněno na obrázku 2. [2]



Obrázek 2 - Složky časových řad, vlastní zpracování [2]

Složka trendu znázorňuje pohyby časových řad v dlouhém časovém období. Jakékoli pravidelné vzorce hodnot nad a pod trendovou složkou lze pravděpodobně přičíst cyklické složce časové řady. V případě letecké dopravy se trendy většinou objevují kvůli faktorům jako je: liberalizace, změna disponibilního příjmu, zavedení nové technologie nebo růst počtu obyvatel. [2]

Sezónní složka zohledňuje pravidelné vzorce variability v rámci určitého časového období, třeba v průběhu roku. Poptávka v leteckém průmyslu je vysoce sezónní, ale je nutné zmínit, že především v přepravě cestujících. Tam je možné sledovat především silné letní měsíce, kdy aerolinky zvyšují četnost letů, na rozdíl od zimních měsíců, kde je snižují. [2] U přepravy nákladu bývá výrazný především pokles v prvním čtvrtletí roku. Vzhledem k velmi silné nákladní letecké dopravě na letištích v jihovýchodní Asii, je Čínský Nový rok svátkem, který množství nákladu velmi ovlivňuje. Letecké zásilky se odkládají na později, tudíž v únoru bývá nejnižší objem nákladu, a naopak v březnu nejvyšší, protože se opět začne odesílat ve velkém množství. Kromě únorového výkyvu jsou však měsíční objemy letecké nákladní dopravy mnohem stabilnější než objemy osobní přepravy. [9]

Cyklická složka se týká dlouhodobého kolísání statistik časových řad dle hospodářského cyklu, který zahrnuje čtyři fáze: expanzi, vrchol, pokles a dno. Předpovědět cyklické výkyvy je složitější než předpovědět sezónní výkyvy či trendy, ale pochopení cyklů může vedení letecké společnosti pomoci lépe se připravit na finanční překážky. [2]

Náhodné jevy jsou krátkodobé, neopakovatelné a neočekávané události, které ovlivňují hodnoty řad. Patří mezi ně například teroristický útok 11. září 2001 nebo pandemie Covid-19. Pokud se leteckým společností daří téměř bezchybně predikovat jiné složky, například cyklické operační náklady a výnosy, u náhodných jevů nic takového jako téměř bezchybná predikce neexistuje.

Trendová složka časových řad tedy analyzuje data za dlouhé časové období, cyklická složka se vyskytuje ve střednědobém horizontu a sezónní výkyvy se vyskytují v krátkodobém horizontu (do 1 roku). Náhodné jevy jsou jedinečné události. [2]

Metod k predikci časových řad existuje velké množství, zde jsou zmíněny čtyři základní. Jedná se o klouzavý průměr, vážený klouzavý průměr, exponenciální vyrovnávání a analýzu trendů. Klouzavý průměr je technika prokládání, která využívá průměr posledních hodnot dat k předpovědi na další období. Jedná se o jednoduchou techniku, která je založena na základním předpokladu, že nejnovější hodnoty nejlépe vystihují budoucnost. Přesnost predikce můžeme zjistit porovnáním rozdílu mezi skutečnou a predikovanou hodnotou. [2]

Vážený klouzavý průměr (WIMA) je podobný klouzavému průměru. Stále využívá historická data k poskytnutí predikované hodnoty, ale na rozdíl od klouzavého průměru mají jednotlivé hodnoty různé váhy. Přiřazené váhy jsou rozděleny dle úsudku prediktora, proto čím více zkušeností a odborných znalostí má, tím je pravděpodobnější, že přiřazené váhy budou přesné. Váhy lze přiřadit libovolnému počtu období, výsledný součet vah se však musí rovnat 100 % nebo 1. [10]

Třetí technikou je exponenciální vyrovnávání. Na rozdíl od klouzavého průměru, který k prognóze využívá více historických hodnot, k exponenciálnímu vyrovnávání jsou využívána data jen z předchozího období. Tím je vytvářena situace, kdy se váha hodnoty s postupem času exponenciálně zmenšuje. Váhu predikované hodnoty určuje vyrovnávací konstanta, která vychází ze skutečné hodnoty z předchozího období a z predikované hodnoty. Čím vyšší konstanta je, tím větší váhy skutečná hodnota dosáhne. Menší vyrovnávací konstanta zajišťuje větší kolísání předpovídané hodnoty, zatímco větší konstanta zajišťuje menší variabilitu predikce. Prognostik tyto hodnoty opět přiřazuje na základě svého úsudku. [2]

Poslední metodou je analýza trendu. Existenci trendu mezi nezávislými a závislými proměnnými v první fázi dobře identifikují diagramy rozptylu a využití spojnicové grafy, které znázorňují, jak těsně se body seskupují. Graficky však trendy není vždy jednoduché zobrazit, proto se k identifikaci trendové linie spíše využívá regresní analýza. Je nutné zmínit, že

analýza trendu předpovídá pouze trendovou část časové řady a cykličnost, sezónnost a náhodné faktory mohou způsobit zkreslení trendové linie. [6]

K ověření přesnosti predikce se využívají dvě hlavní metody, střední kvadratická chyba (MSE) a střední absolutní odchylka (MAD). Střední kvadratická chyba je průměrem čtvercového rozdílu mezi skutečnou hodnotou a predikovanou hodnotou. Hodnoty se kvadratickují, aby se eliminoval vliv záporných chyb, které ruší chyby kladné, a také aby se přiřkládala větší váha větším chybám. Střední absolutní odchylka zjišťuje průměr absolutních hodnot odchylek a je lépe interpretovatelná. Odchylky zde nejsou kvadratickované, tudíž se na rozdíl od MSE velkým odchylkám nepřikládá větší váha. Na základě měr přesnosti je snadnější zvolit vhodnou metodu predikce. [2]

Ekonometrické nebo kauzální metody

Ekonometrické nebo kauzální metody vztahují poptávku ke změnám faktorů, u nichž lze očekávat, že budou mít na poptávku vliv. Předpokládají tedy, že je poptávka po nákladu ovlivněna jedním či více ekonomickými, sociálními či nabídkovými faktory. Silnou stránkou kauzálních modelů je jejich logičnost. Dle ekonomické teorie poptávka závisí především na ceně, cenách konkurenčních produktů, původu produktu, míře, do jaké je pro zákazníka produkt důležitý, příjmu spotřebitele a spotřebitelově vkusu. Pokud se některá z těchto proměnných změní, změní se i finální poptávka. Ekonometrické modely s těmito proměnnými pracují i při predikci cargo dopravy. Postup je takový, že je nejprve třeba identifikovat a zvolit faktory pro model, a následně se určí vztahy mezi závislými a nezávislými proměnnými neboli se specifikuje forma modelu, který bude využit k predikci. V dalším kroku se model kalibruje a testuje matematickými výrazy, aby se zjistilo, jak na sebe faktory v rámci vztahů reagují a závisí. Pokud je výsledek signifikantní a statisticky významný, přechází se k poslednímu kroku, čímž je predikce nezávislých proměnných či užití cizích predikcí pro odvození námi hledané predikce. [1]

Regresní modely

Do ekonometrických metod spadají regresní modely, na kterých bývá založena většina predikcí leteckých společností. [1] Regresní analýza je zároveň i jednou z nejstudovanějších a nejpoužívanějších ekonometrických technik. [11] Některé konkrétní modely zde zmiňované jsou detailněji popsány v další části teorie u řešerše odborných prací.

Závislá proměnná je proměnná, která závisí na jiných faktorech a proměnných, zatímco nezávislá proměnná má hodnotu, která nezávisí na žádném jiném faktoru. Většina regresních analýz obsahuje několik nezávislých proměnných, protože závislou proměnnou většinou nelze

vysvětlit jen jedním faktorem. [2] Dvěma nejčastěji užívanými proměnnými jsou průměrná sazba za službu a určitá míra příjmu na obyvatele. Při predikci by měly být úroveň sazby, úroveň příjmu a ostatní ekonomické proměnné očištěné o inflaci a měly by být vyjádřeny v reálných hodnotách. [1] Protože je občas obtížné určit, které příjmy je třeba měřit, HDP bývá často rozumným zástupným ukazatelem příjmů. [2]

Před provedením regresní analýzy by měl být vytvořen graf, který znázorňuje rozptyl, který nám může pomoci pochopit povahu bodů s daty, neboť reálné údaje spolu nejsou nikdy v dokonalé korelaci kvůli faktorům způsobujícím zkreslení vztahů. Na základě grafu je možné následně zvolit vhodný typ regresní analýzy (např. lineární, exponenciální atd.), která bude nejvíce vhodná k přesné predikci závislé proměnné. Zjednodušeně řečeno, regresní analýza vypočítá kvantifikovatelný lineární/exponenciální/kvadratický vztah pro různé datové body. Všechny typy regresní analýzy plní stejnou funkci, existuje však mnoho různých metod pro vytvoření trendové čáry. Nejvyužívanější metodou je metoda nejmenších čtverců, při které dochází k minimalizaci kvadratické hodnoty reziduí, z čehož vychází, že větší chyby mají následně v postupu větší váhu. Reziduum v podstatě představuje rozdíl mezi skutečnou hodnotou a predikovanou hodnotou podél přímky trendu. [2]

Po specifikaci regresního modelu a nezávislých proměnných se model kalibruje na předchozí úrovni provozu a změny v nezávislých proměnných, obvykle se používají data časových řad. Společnosti testují různé formace modelu ještě před tím, než zvolí nezávislé proměnné, které budou užité v predikci. [1] Predikce je obvykle přesnější při vyšším počtu pozorování nebo při nižším počtu nezávislých proměnných. [2] Jednodušší modely se spíše využívají pro prognózy konkrétních tras, ekonometrické modely se používají pro predikci vývoje dopravy na širších trzích mezi jednotlivými zeměmi nebo mezi regiony. Mohou být velmi podrobné a složité, zejména pokud prediktor rozděluje trh do většího množství jednotlivých segmentů. [1]

Při snaze zachytit vliv náhodné události nebo dopady sezónnosti je také možné využít fiktivní proměnné. Regresní analýza s fiktivní proměnnou se většinou využívá k určení skutečné povahy vztahu mezi závislou proměnnou a nezávislými proměnnými. Fiktivní proměnná jen určuje, zda přítomnost zkoumaného faktoru existuje nebo ne a nabývá jen hodnot 1 nebo 0. Například když bychom chtěli zachytit vliv náhodné události 11. září 2001, ohodnotíme všechna čtvrtletí následující po třetím čtvrtletí 2001 hodnotou 1, čímž vyloučíme proměnné čtvrtletí, které nebyly ovlivněny touto událostí. [2]

U modelu je zároveň potřeba ověřit, zda je statisticky spolehlivý i po predikci, což lze zajistit pomocí statistických testů. [1] Pravděpodobně nejjednodušší metodou je využití koeficientu

determinace R neboli R-kvadrát, který vyjadřuje, jaký podíl variability závislé proměnné na nezávislé proměnné model vysvětluje. Dokonalá predikce hodnot závislé proměnné je značena hodnotou 1, zatímco 0 značí naprosto nepřesnou předpověď. [2] Při použití dat časových řad je v ideálním případě očekáván koeficient minimálně 0,9, pokud bychom model chtěli použít k predikci. Zároveň se dá koeficient využít k volbě mezi modely s různými kombinacemi nezávislých proměnných. [1]

Zároveň se zde vyskytují některé metodologické překážky. Například to, že vysoké koeficienty vícenásobné determinace samo o sobě neznamenají záruku kauzality a vztahu mezi nezávislými a závislými proměnnými. [1] Problémy mohou vzniknout i autokorelací, která může zkreslit přesnost a významnost regresního modelu. Dochází k ní ve chvíli, kdy byla vynechána významná nezávislá proměnná nebo pokud dochází k výraznému cyklickému kolísání závislé proměnné, jinak řečeno, když má trend. Pro zjištění autokorelace se dají použít testy či vykreslení v závislosti na čase. [2] Další překážkou může být multikolinearita, která nastává, když jsou dvě nebo více nezávislých proměnných navzájem vysoce korelovány, což může vést k problémům při interpretaci regresních koeficientů. [1] Multikolinearita v modelu se obvykle řeší buď získáním většího množství dat nebo vyloučením jedné či více kolineárních nezávislých proměnných. Z těchto dvou možností se většinou volí eliminace. Při výstupech regrese je potřeba vnímat také ANOVA F-statistiku. Protože je potřeba nezávislé proměnné vždy statisticky otestovat, F-statistika nám poskytuje údaje o celkové signifikanci regresní rovnice se zaměřením na celou množinu proměnných. Na rozdíl od T-statistiky, která je zaměřená jen na jednu jednotlivou proměnnou. Ideálně by nezávislé proměnné měly projít předem stanoveným testem, pokud ne, mělo by se přemýšlet o jejím vyřazení z modelu. Po kontrole regrese dochází obvykle ke kvantifikaci funkce poptávky, aby bylo následně možné vytvořit predikce. [2]

I přes nemožnost zaručení stoprocentní přesnosti predikcí k nim musí docházet, protože jsou na ně navázána důležitá rozhodnutí. Volbu správné techniky určuje několik faktorů. Nejdůležitější je si stanovit cíl predikce. Dále je nutné hledět i na rychlost predikce, náklady s ní spojené a dostupnost dat. Většina technik je poměrně přesná pro krátkodobé predikce. Pro delší časové období existují určité pochybnosti, ale je pravděpodobné, že kvantitativní techniky budou poskytovat přesnější prognózy. Někteří odborníci ale upozorňují, že pro toto tvrzení neexistují jasné důkazy. Některé menší letecké společnosti jsou v reálném světě často závislé na výrobcích letadel či predikcích IATA, které následně kombinují s vlastním úsudkem. [2]

Modely strojového učení

Kromě statistických metod je nutné zmínit ještě modely strojového učení. Důležitým důsledkem dnešních technologií a dostupnosti mobilních technologií a internetu věcí (IoT), je zvýšený objem sběru dat, který se označuje jako „big data“. Tato data jsou známá pod pěti písmeny V, do češtiny přeložené jako objem, rychlost, rozmanitost, pravdivost a hodnota. K jejich přeměně na smysluplné informace je zapotřebí složitější datová analytika. [12] Jednou z převratných technik datové analytiky, u níž se předpokládá, že ovlivní a zpřesní budoucnost predikce a v posledních letech se rychle rozvíjí, je strojové učení. Podstatná je především schopnost zachytit nelineární vzor dat a schopnost poskytnout přesnější výsledky predikce než statistické modely predikcí pro nelineární soubory dat. [13] Strojové učení je podmožinou umělé inteligence, kde algoritmus strojového učení jedná nebo provádí úkol, aniž by byl explicitně naprogramován. Stroj se může automaticky učit z minulých nezpracovaných dat a vytvářet predikční modely na základě předem navržených algoritmů. Obecně existují dva typy algoritmů učení: učení pod dohledem a učení bez dohledu. Algoritmy strojového učení s dohledem se učí z označených dat: vstupních a výstupních. Algoritmus je zodpovědný za nalezení vztahu mezi vstupem a výstupem a přestane se učit, když dosáhne přijatelné úrovně výkonu. Naproti tomu v případě nekontrolovaného učení existují pouze vstupní data a žádná odpovídající výstupní data. Cílem algoritmu je najít vzory a strukturu, aby se o daných datech dozvěděl více. Pro strojové učení se používají různé algoritmy, například neuronové sítě, stroje s podpůrnými vektory, rozhodovací stromy, náhodné lesy a jiné. [12] Neuronové sítě patří zatím mezi nejvyužívanější metody strojového učení právě díky svým přednostem jako je masivní paralelní zpracování, vysoká schopnost odolávat chybám a filtrovat šum, což jsou schopnosti užitečné pro predikci časových řad. [13] Dá se obecně říct, že v některých případech neuronové zpracování přesahuje tradiční ekonometrický přístup a nabízí zobecnění v chování časových řad, i když nejde jen o malé vzorky. [11] Nelze však říct, že se některá z metod dá využít pro nejpresnější predikci za všech okolností a také je nutné mít na paměti, že metody strojového učení bývají nákladnější a mnohem složitější na interpretaci než statistické metody. Je to však směr, kterým se bude predikce do budoucna pravděpodobně ubírat. [13]

2. Rešerše odborných prací zabývajících se predikcí poptávky po letecké nákladní dopravě

V této kapitole bude představeno deset odborných článků, které se zabývají predikcí poptávky po leteckém cargu. Z metod, které budou analyzovány, bude následně vybrána jedna metoda, která vzejde z rešerše jako ta nejvhodnější pro následující predikci. Metoda bude v následující kapitole užita k predikci přepraveného nákladu na konkrétních trasách.

1) Air cargo demand modeling and prediction [6]

Predikce poptávky leteckého nákladu uvedená v článku Totamane – Dasgupta – Rao, je založena na modelu induktivního uvažování, které vychází z pozorování celkové nabídky leteckého nákladu v mnoha obdobích (dnech) v průběhu let a zároveň z pozorování průměrného tržního podílu leteckých společností působících v určitých regionech.

Model je založený na přístupu „Potluck Problem“ a na základě predikce poptávky má za cíl navrhnout flexibilní letový řád a kapacitní plán nákladu, který může maximalizovat load factor (faktor zatížení) při zachování stávající úspěšnosti doručení. Load factor je poměr průměrného nákladu k celkové nákladní kapacitě letadla. [14]

Potluck problém je abstraktní herně-teoretické paradigma opakované nekooperativní hry v multiagentním systému. Agenti jsou producenti a spotřebitelé. Nekooperativní znamená, že spolu jednotliví agenti nespolupracují a nedávají znát své úmysly dopředu, naopak jedná každý jen ve svém zájmu. Všichni mají ale přístup k historickým údajům o produkci a spotřebě v minulých případech „večeře“. „Večeří“ se rozumí časové období poptávky, většinou se hovoří o hodinách/dnech/atp., záleží na oblasti. Spotřebitelé užívají statky, přičemž v rámci „večeře“ převážně dochází ke dvěma situacím – „hladovění“ a „přebytku“. „Hladovění“ aneb situaci, kdy část poptávky zůstává neuspokojena a „přebytku“, aneb situaci, kdy je nabídka zdroje ze strany producentů větší než poptávka ze strany spotřebitelů. Potluckovým problémem pojmenováváme situaci, kdy musí poskytovatelé na základě dat o poptávce určit, jaké zdroje budou nabízet v následujících časových obdobích. Mění se tedy jak poptávka, tak nabídka. Dochází zde k oscilačnímu chování, neboť pokud v části večeře dojde k hladovění, producenti dodají více zdrojů pro další večeři, a tak může dojít při této druhé večeři k přebytku. Pokud je však u večeře přebytek, producenti sníží svou produkci a při následující večeři může dojít opět k hladovění. K oscilacím může takto docházet stále dokola.

V článku je k predikci obecně využit algoritmus učení vážené většiny s různými prediktory pro předpověď budoucí poptávky. Na základě předpovídané poptávky, dostupné nákladní kapacity a globálních/regionálních ekonomických ukazatelů (HDP, inflace, růst průmyslu atd.) a za

použití různých strategií, je navržen nový plán nákladní kapacity, čímž se zlepší nákladní LF i finanční výsledky.

Pro předpověď poptávky při večeři jsou užívány různé sady více prediktorů, z nichž vzniká větší množství předpovědí, z kterých je následně určena celková předpověď poptávky pomocí algoritmu vážené většiny neboli WMA. Jedná se o algoritmus pro predikci posloupnosti hodnot pomocí předpovědí, které vzejdou z daného počtu prediktorů. Podstatnou informací je, že každému z prediktorů je přiřazena váha, která vyjadřuje úroveň důvěry v přesnost prediktoru v tu danou chvíli. Během procesu se dává větší váha předpovědím spolehlivějších prediktorů (těm více přesným) oproti těm méně spolehlivým. Pokud jsou nadále předpovědi přesné, váha jim je zvýšena a naopak. V práci je zmíněno 13 prediktorů, z nichž si letecká společnost vybírá libovolný počet k predikci. Po přidělení vah a jejich aktualizacích po jednotlivých iteracích, je vypočítán vážený průměr predikcí a je rozhodnuto o očekávané poptávce po nákladu. Do algoritmu je možné přidávat další prediktory, na spolehlivosti to téměř nic nemění.

Snahou je dosáhnout stavu blízkého rovnováze, kdy se agregátní nabídka zboží téměř rovná agregátní poptávce. Agenti dodavatelé předpovídají poptávku po zdrojích na večeři na základě historických dat a následně rozhodují o množství, které dodají. Čím vyšší je přesnost poptávky, tím více se přibližujeme rovnováze mezi agregátní poptávkou a nabídkou zdrojů letecké nákladní dopravy. Nabídkou S zde rozumíme součet kapacit nákladních letů provozovaných agentem na konkrétní trase. Poptávkou D značíme poptávku pro konkrétní den, který predikujeme. Pokud se $S = D$, dochází k rovnovážnému stavu. Pokud $S < D$, dochází k přeplnění nákladní kapacity, kterou letecké společnosti nezvládnou odbavit, pokud $S > D$, dochází k nevyužití dostupné nákladní kapacity. Poptávka se zároveň v různé dny liší. V ideálním případě dochází k maximalizaci faktoru zatížení právě ve chvíli, kdy se $S = D$.

Jednotlivé kroky predikce leteckých společností jsou tedy následující:

- 1) Předpověď poptávky pomocí jednotlivých prediktorů
- 2) Předpověď poptávky pomocí všech prediktorů s využitím WMA
- 3) Rozhodnutí o nabídce S , kterou zvládnou obsloužit
- 4) Aktualizace vah všech prediktorů na základě aktuální poptávky a nabídky

Celková predikce poptávky po cargu pro konkrétní den roku je tedy vzata váženou většinou predikcí všech prediktorů. Po předpovědi poptávky po nákladu, musí agent rozhodnout o nabídce S . Ta je omezena na diskrétní hodnoty, které jsou kombinací různých nákladních kapacit letadel a počtů cest těchto letadel na dané trase a v daný den.

Jsou užitá data z let 2005 a 2006 (730 dní) různých nákladních leteckých společností, které působí na různých trasách v severoamerickém regionu. Data byla použita pro získání vah prediktorů ve WMA. Poté byly výsledky porovnány se skutečnými výsledky těchto leteckých společností v roce 2007. Ve všech čtyřech simulacích byly vážené předpovědi algoritmu přesné téměř v 75-81 % případů (275-295 dní v roce) s úrovní tolerance ± 25 % skutečné poptávky po nákladu letecké společnosti. Úroveň tolerance ± 25 % vedla k fixní hodnotě parametru β 0,87. β měří, jak drastické jsou změny předpovědí v průběhu iterací. Čím menší je β , tím drastičtější jsou změny. Algoritmus dosáhl ve všech případech zlepšení nákladního LF o více než 9 % (což prakticky znamená desítky milionů dolarů).

Změna load factor (LF) má významný vliv na letecké společnosti. Ovlivňuje efektivitu využití kapacity letadel a zohledňuje faktory jako vyvážení nákladu a těžiště. Pro zvýšení LF se letecké společnosti uchylují k různým strategiím. Jednou z nich je rušení letů, což je reálná a proveditelná možnost. Je však důležité nezvyšovat LF na maximum, protože to by znesnadnilo řešení neočekávaných situací, jako je špatné počasí, poruchy v letovém řádu nebo technické problémy. Úspěšní nákladní dopravci udržují LF na úrovni nejvýše 70 %, což umožňuje flexibilitu při řešení nepředvídatelných událostí. Jiné strategie zahrnují rychlé přidání nebo přesměrování letů, což je však složitější a závisí na mnoha omezeních, včetně schvalování letových plánů, finančních investic a dostupnosti posádky.

Z výsledné váhy prediktorů lze vyčíst, že nejzásadnější roli v predikci hrají average time varying function, což značí průměrnou poptávku v sobotu a prediktor zaměřující se na prázdninové dny. Ačkoli prediktory HDP a inflace mají vliv na určení budoucí čtvrtletní nebo roční poptávky, ve výsledku mají relativně menší vliv na denní nebo týdenní trendy poptávky.

Pro další účely predikce by mohly být přidány faktory jako jsou schopnosti, výkonnost a dostupnost letadel, regionální úroveň sídla každé letecké společnosti, dostupnost letištních slotů, zohlednění proměnných, jako je typ a stáří jednotlivých letadel, velikost letadel, objem nákladu, vyřazení letadel, využití, strategie a jiné.

2) Hybrid forecasting model to predict air passenger and cargo in Indonesia [15]

Článek Sulistyowati – Astuti užívá hybridní metody predikce cargo přepravy, konkrétněji kombinaci dvou lineárních a nelineárních modelů. Je zaměřen jak na predikci počtu pasažérů, tak na predikci objemu carga. Pro účely této práce zmiňují výsledky spojené s nákladem, pokud je možné tyto informace oddělit. Článek predikuje objem carga na třech mezinárodních letištích v Indonésii: Soekarno Hatta, I Gusti Ngurah Rai a Juanda Airport. Tato tři letiště se na celkovém objemu nákladu vyváženého z Indonésie podílí z 68,50 %.

První fází je lineární modelace pomocí regresního modelu časových řad (TSR – Time series regression model) a autoregresního integrovaného klouzavého průměru s exogenním faktorem (ARIMAX – Autoregressive Integrated Moving Average with Exogenous Factor). TSR je modelem spíše na dlouhodobější predikci. Nachází se zde 3 prediktory: trendy, sezónnost a kalendářní variace. Sezónností je myšlen opakující se vzorec se stejným obdobím, například 12 měsíců v roce, zatímco kalendářní variace jsou sezónní vzorce s různými obdobími. V této práci jde především o období kolem svátku Eid al-Fitr. Model ARIMAX je rozšíření standardního modelu ARIMA o regresní (exogenní) proměnné. ARIMA model je lineární kombinace prognostického modelu složeného z autoregresního modelu (AR) a modelu klouzavého průměru (MA). Model ARIMAX použitý v tomto článku je model s trendem, sezónními a kalendářními odchylkami.

Ve druhé fázi je chyba lineárního modelu analyzována pomocí metod strojového učení, jako je neuronová síť (NN – Neural Network) a podpurná vektorová regrese (SVR – Support Vector Regression). NN je metoda, která se používá k modelování vztahu mezi vícerozměrnou vstupní proměnnou x a vícerozměrnou výstupní proměnnou y a je složena z řady vzájemně propojených jednoduchých prvků zvaných uzly či neurony (původně model vyvinutý k napodobení fungování lidského mozku). V článku je při modelování časových řad využívána neuronová síť s dopředným posuvem. [16]

Hlavní myšlenka SVR je minimalizovat chybu a definovat optimální hyperrovinu pomocí minimalizace strukturálního rizika (SRM), rozdělení dat a minimalizace vzdálenosti mezi hyperrovinou a daty. [17]

Byly použity a porovnány čtyři hybridní modely: TSR-NN, TSR-SVR, ARIMAX-NN a ARIMAX-SVR. Nejvhodnější model byl vybrán na základě vyhodnocení přesnosti predikce. Kritériem pro výběr nejvhodnějšího modelu je střední absolutní procentní chyba (MAPE) v testovacích datech. Výhodou MAPE je, že je nezávislá na měřítku, a proto se často používá k porovnání výkonnosti předpovědi mezi různými soubory dat. Hodnocení založená na MAPE jsou však nevýhodná, pokud je některé z pozorování nekonečné či neurčité nebo má extrémní hodnotu blízkou nule. MAPE také více znevýhodňuje záporné chyby oproti chybám kladným. V této studii však nebyla hodnocení s extrémními hodnotami blízkými nule a chyby byly symetrické.

Data v článku jsou měsíční údaje, které jsou pravidelně zveřejňovány statistickým úřadem Indonésie (BPS). Časové období údajů je v rozmezí od ledna 2001 do srpna 2017. Jak již bylo řečeno, data jsou sbírána ze tří nejpřínosnějších indonéských mezinárodních letišť. Kalendářové odchylky dat jsou navázány především na Eid al-Fitr, což je muslimský svátek, kterým je zakončen měsíc Ramadánu. Časová řada celkového objemu nakládky na třech

mezinárodních letišťích značí rostoucí trend a sezónní průběh s rostoucím objemem nákladu v červenci a prosinci. Zároveň vychází, že pokud se Eid al-Fitr konal v prvním, druhém, třetím a čtvrtém týdnu, pak je průměrný náklad vysoký měsíc před svátkem a měsíc po svátku.

Postup užitý v článku je následující: V první fázi se modeluje osobní a nákladní letecká doprava na jednotlivých letišťích, tj. Soekarno Hatta, I Gusti Ngurah Rai a Juanda Airport, s využitím trendových, sezónních a kalendářních změn pomocí modelů TSR a ARIMAX. V modelu TSR je prvním krokem kontrola reziduálních předpokladů, které zahrnují bílý šum a normální rozdělení. Jak již bylo zmíněno, rezidua jsou formou variability, která se nám nedaří vysvětlit pomocí proměnných zařazených do modelu. Jedná se tedy o formu chyby predikce a lze je použít pro hodnocení celkové vhodnosti modelu. Pokud není splněn předpoklad bílého šumu reziduí, budou do modelu zahrnuta významná zpoždění. Pokud jsou všechny předpoklady splněny, provede se odhad parametrů a testování významnosti parametrů. Mezitím se při modelování ARIMAX se složkou trendu, sezónní a kalendářní variability data nejprve modelují regresním modelem časové řady a poté reziduum modeluje ARIMA na základě ACF (autocorrelation function) a PACF (partial autocorrelation function) modelu.

Druhou fází je modelování hybridní metodou, tj. TSR-NN, TSR-SVR, ARIMAX-NN a ARIMAX-SVR. Vstupem použitým v druhé fázi modelování jsou data s významným zpožděním z první fáze modelování. Cílem je zachytit zanedbanou nelineární složku modelování TSR nebo ARIMAX pomocí pokročilého modelování s využitím chyby. Určení parametrů v hybridním modelování TSR-SVR a ARIMAX-SVR probíhalo pomocí metody prohledávání mřížky. Vstupy použité v modelech TSR-SVR a ARIMAX-SVR byly podobné vstupům, které byly použity v TSR-NN a ARIMAX-NN. Model SVR má tři parametry, tj. 2σ , epsilon a C. Pro dosažení optimálních parametrů je nutné kombinovat hodnoty těchto tří parametrů tak, aby se dosáhlo nejmenší hodnoty MAPE.

Z výsledků je zřejmé, že užití hybridních modelů poskytuje přesnější predikci, obzvláště v případě tréninkových dat. Článek zaznamenal růst objemu carga na všech třech letišťích. Kromě toho má pohyb a nárůst nákladu na těchto třech letišťích periodický průběh, který se projevuje během červencových prázdnin akademického roku a prosincových vánočních a novoročních svátků. Objem se také zvyšuje o svátku Eid al-Fitr, což je důsledek kalendářních změn. Na základě srovnání prognostické výkonnosti výsledky ukazují, že metody ARIMAX-NN a TSR-NN poskytly lepší prognostickou výkonnost než TSR-SVR a ARIMAX-SVR. Použití NN v hybridních modelech má tedy lepší výkonnost ve srovnání s SVR pro predikci carga na třech největších mezinárodních letišťích v Indonésii. Článek také doplňuje, že kombinace metod predikce zajišťuje přesnější předpověď.

3) Modeling demand for air cargo in the colombian context [11]

Článek Loaiza-Solano-Ojito-Simancas využívá modelů lineární regrese a modelu neuronových sítí pro predikci letecké nákladní dopravy v Kolumbii. Původním záměrem bylo použít data přepravy mezi Kolumbií a 19 největšími letišti světa a jejich přepraveného nákladu, pro nedostatek informací však bylo několik světových letišť vyřazeno. Data jsou shromážděna za roky 2005-2014, jedná se o měsíční hodnoty. Proměnné v obou modelech jsou socioekonomické informace země, ve kterých se letiště nacházejí, získané většinou Správním oddělením národní statistiky (DANE) dané země.

V článku byl využit ekonometrický přístup s logaritmickou transformací dat a aplikován lineární regresní model, který sloužil k určení modelu pro předpověď celkové poptávky po leteckém nákladu z Číny. Model provádí prognózy na základě příčinné souvislosti mezi leteckou nákladní dopravou z Číny a ekonomickými aktivitami, související informace jsou získávány z časových řad údajů a ekonomických prognóz pro tři různé scénáře: pesimistický, nejpravděpodobnější a optimistický. Relevantními celosvětově užívanými proměnnými zahrnutými do modelu byly HDP, mezinárodní obchod země a tok přímých zahraničních investic.

Lineární regrese v článku využívá různé kombinace socioekonomických proměnných a produkce. Byly využity proměnné, které jsou relevantní světově, plus proměnné, které jsou relevantní pro analyzovanou oblast, jako například produkce květin. Užití proměnné jsou tedy následující: náklady na odeslání tuny nákladu z jednoho letiště na libovolné místo určení = odhaduje se, že čím vyšší je sazba, tím nižší je generování nákladu z letiště. Pro článek je určena cena za 1 kg mezi páry míst odletu-destinace a pro zjednodušení byl použit jeden tarif pro celou zemi. Dále produkce květin (1, když je produkce květin, 0 když ne), zemědělská plocha (plocha v ha určená pro zemědělskou produkci, předpokládaná záporná hodnota = není přepravována letecky), lesní plocha (plocha v ha určená pro dřevo přepravované letecky) a HDP (hrubý domácí produkt na obyvatele, kladná proměnná = vyšší HDP = větší zatížení ekonomické kapacity).

Jsou zde využity tři regresní modely, z výsledků vychází s nejvyšším R^2 model číslo 3, neboť má statisticky signifikantní všechny proměnné a jedná o nejlepší výsledek ze tří modelů.

V neurálních sítích byly využity stejné proměnné jako v případě lineární regrese: náklady, produkce květin, zemědělská plocha, HDP a lesní plocha. Opět jsou zde vypočítány tři modely se čtyřmi vstupními proměnnými a dvěma výstupními propojujícími neurony. V algoritmu jsou

pro predikci použity různé hodnoty koeficientů. Použitá skrytá aktivační funkce byla hyperbolická tangenciální aktivační funkce a výstup byl lineární.

V rámci lineární regrese vyšlo najevo, že sazba účtovaná za přepravu nákladu, HDP na obyvatele a zohlednění toho, zda se v daném departementu (zóně vlivu) nacházejí či nenacházejí květiny, mají velký vliv na celkovou poptávku po leteckém nákladu. Zvýšení sazeb má na poptávku negativní vliv, růst HDP a produkce květin má vliv pozitivní, což generuje zvýšení poptávky po letecké dopravě. Vliv má i rozloha lesů a zemědělská produkce, ale v porovnání s předchozími proměnnými výrazně nižší. V případě neuronových sítí byly použity stejné proměnné jako v lineární regresi, aby se ověřilo, že i při použití jiného modelu odhadu je jejich vliv zřejmý. Modely neuronových sítí poskytly přesnější výsledky než modely lineární regrese, ale byly odpovídající. V budoucích predikcích je možné přidat další proměnné, například i výrobu konkrétních produktů sledovaného trhu, jako například zde zmíněných květin.

4) Air Cargo Demand Prediction on Selected Routes by Holt-Winter Algorithm [17]

Článek Kaščeev-Endrizalová-Vittek je zaměřen na předpověď poptávky po leteckém nákladu na pěti trasách: dvou v rámci Evropy, dvou Evropa-Asie a jedné Evropa-S. Amerika. Trasy byly vybrány vícestupňovým výběrovým systémem a autoři se zaměřili výběrem pouze na trasy s neustále rostoucí tendencí co se týká vykazovaných objemů dopravy. Hlavním zdrojem dat byla databáze Eurostatu a jednalo se o měsíční údaje 1/2015-12/2020, tudíž data za 6 let. Jako metrika byla použita hmotnost naloženého a vyloženého nákladu mezi vybranými městy v tunách. Statistická významnost dat byla posuzována pomocí lineární a kvadratické regresní analýzy. Predikce byla provedena pomocí Holt-Wintersova algoritmu pro období 1/2021-4/2022. V metodě nejsou zahrnuty ekonomické parametry regionů.

Postup byl následující. Nejprve bylo zvoleno 15 největších letišť podle největšího objemu nákladu v roce 2019, v dalším kroku byla ponechána letiště s kompletními záznamy. Letiště, která zůstala byla: FRA, AMS, LEJ, LUX, LGG, CGN a EMA. Následně byly voleny trasy pro uvedená letiště. Po vyčištění dat, kdy došlo k zúžení množství tras z důvodu nekompletních záznamů, byly vybrány trasy podle průměrného měsíčního objemu nákladu s minimem 2000 tun za měsíc, čímž zůstalo pro následující regresní analýzu 90 tras. Lineární a kvadratická regresní analýza užitá pro každou „dvojici měst“, vyřadila další 43 tras, které hodnotově nebyly statisticky významné, o čemž bylo rozhodnuto na základě ANOVA F-statistiky a hodnoty P.

Cílem užití Holt-Wintersova algoritmu bylo určit budoucí růst nákladní dopravy na vybraných trasách. Algoritmus je AAA verzí exponenciálního vyrovnavání, založený na rovnici, která zohledňuje aditivní chyby, aditivní trendy a aditivní sezónnost v časových údajích. Jedná se o využití exponenciálních časových řad, kdy je nárůst objemu carga při každé časové jednotce větší než v té předchozí. V rámci predikce bylo potřeba finální hodnoty porovnat s historickými údaji, aby bylo možné získat přehled o tom, o kolik jednotlivé trasy porostou. Z předpovědi vzešlo 36 tras s rostoucím průměrným měsíčním nákladem a 11 s klesajícím. Z 36 tras bylo nakonec zvoleno 5, které byly vybrány i vzhledem k diverzifikaci typů letadel, které jsou na trasách provozovány. Finální trasy byly LEJ-CDG, CGN-SDF, LEJ-EMA, FRA-PVG a LUX-HKG. Konečná predikce nárůstu objemu nákladu naloženého a vyloženého byla na čtyřech trasách mírná, jen na trase LEJ-EMA velmi vysoká. Vznikly vždy tři predikce na každé jednotlivé trase, predikční linie a dolní a horní hranice spolehlivosti. Při verifikaci byl porovnáván rozdíl mezi těmito třemi údaji s reálnými údaji z Eurostatu.

Článek shrnuje, že použitá metoda není zcela přesná a vykazuje velmi široké rozpětí rozdílů. Důvody, které rozpětí mohly způsobit mohou být chybějící další proměnné (např. HDP, PZI, indexy spotřebitelský cen zemí atd.), nezahrnuté politické vlivy a regulace či úzké soustředění na konkrétní dvojice měst. Autoři zmiňují, že rozdíly mezi predikcí a skutečným vývojem byly u jednotlivých linek rozdílné. Algoritmická předpověď poptávky se ukázala jako nepříliš vhodná s průměrnou odchylkou 20 % u průměrných hodnot predikce. Tato metoda má však potenciál při zkoumání širšího trhu, kde se dají definovat spojení se specifickým růstem či poklesem proti obecnému trendu trhu.

5) Comparative analysis of forecasting for air cargo volume: Statistical techniques vs. machine learning [19]

Článek Liu-Ding-Guan-Gui-Xu představuje empirické hodnocení dvou statistických technik a tří modelů strojového učení pro predikci objemu leteckého carga. Ve studii jsou užity: vícenásobná lineární regrese (MLR), autoregrese s integrovaným klouzavým průměrem (ARIMA), regrese s podpůrnými vektory (SVR), neuronová síť (NN) a gradient boosting regression tree (GBRT). K testování modelů jsou shromážděny údaje o objemu letecké přepravy zboží v Číně za období jednoho roku a jsou navrženy tři predikční horizonty a dvě predikční strategie (statická a dynamická) pro jednodušší porovnání jednotlivých metod. Pro analýzu byl použit program Python. Pro hledání optimální kombinace je využito mřížkové vyhledávání.

Modely v článku zmíněné můžeme rozdělit do dvou kategorií. Jednou jsou statistické techniky, např. multilineární regrese, model autoregresního integrovaného klouzavého průměru (ARIMA), šedý prognostický model atd. S průlomem umělé inteligence identifikujeme druhou kategorii, metody strojového učení, jako jsou neuronové, rozhodovací stromy, support vector machine a jiné. Vícenásobná lineární regrese se užívá k modelování vztahu mezi spojitou závislou proměnnou a několika relativními nezávislými proměnnými. Na rozdíl od ní automatická regrese s integrovaným klouzavým průměrem (ARIMA) počítá s korelací mezi předpovídanou hodnotou proměnné a jejími historickými daty. Model ARIMA vykazuje v mnoha případech s velkým a lineárním souborem dat vysokou prognostickou výkonnost. Neuronová síť má za své silné stránky vysokou schopnost odolávat chybám a schopnost filtrovat šum, výkonnost však není vždy uspokojivá a mezi nevýhody spadá například upadnutí do lokálního optimálního řešení a přílišné přizpůsobení. Jedna z novějších metod strojového učení, support vector machine (SVM) má dobré extrapoláčnické vlastnosti a optimální nastavení parametrů se hledá za pomoci optimalizace pomocí roje částic (PSO). Poslední metodou je Gradient boosting, který jako ansámblová metoda integruje více slabších predikčních modelů. V článku je využita konkrétně Gradient boosted regression tree metoda, která strategicky integruje slabé rozhodovací stromy s cílem minimalizovat reziduum. Cílem každé iterace je zkonstruovat optimální slabý rozhodovací strom tak, aby se minimalizovala chyba ztráty. Protože obecně neexistuje jediný prognostický model, který by poskytoval nejlepší predikci pro všechny situace za všech okolností, zaměřuje se článek s těmito pěti metodami na leteckou nákladní dopravu, kde pro některé ze zmíněných metod zatím nejsou ještě dohledatelné zpracované studie.

Prvním krokem je analýza stability nelineárních dat časových řad pomocí Dickey-Fullerova testu, z níž vyplynulo, že nelze existenci jednotkového kořene zamítnout, tudíž je řada stacionární. Zároveň je potřeba pro ověření podmínek použitelnosti modelu ARIMA data rozložit na 3 vlivy, trend, sezónnost a rezidua, aby se zjistily charakteristiky časové řady. Zjistilo se, že rezidua splňují požadavky procesu bílého šumu a vliv týdenní sezónnosti je silný.

V článku jsou použity údaje jednoho roku, které zahrnují 365 pozorování denního objemu leteckého carga od ledna 2018 do prosince 2018. V datech je pozorován pravidelný průběh, prudší kolísání přichází v období Čínského Nového roku. Aktuální objem leteckého carga je závislý na objemu z předchozích dnů, tudíž jsou vstupy generovány na základě historických informací časových řad. Vzhledem k tomu, že jsou bližší historická data považována za relevantnější, je ve studii jako interval použito 7 dní, které předchází predikovanému dni. Následně jsou vybrána pozorování z několika týdnů s určitým intervalem. Jako interval je aplikováno 7 dní, které odpovídají předchozím třem týdnům, což generuje 10 vstupů

odpovídající jednomu měsíci s téměř 30 dny. Soubor tedy obsahuje informace o časové řadě za jeden měsíc. Pro testování výkonnosti modelu je soubor dat rozdělen na dvě části, tréninková data pro učení (265 dní pozorování) a testovací data pro ověření (100 dní). Je počítáno s třemi predikčními časovými horizonty – údaje získané 1, 5 a 10 dní před aktuálním dnem predikce. Pro predikční pozorování jsou stanoveny 3 horizonty, 1, 5 a 10 kroků dopředu (např. $t+1$). Jak již bylo řečeno, strategicky je využité statické a dynamické predikování pro lepší vyhodnocení modelů. Statické pozorování znamená pevně stanovenou strukturu a parametry modelů, zatímco dynamické znamená rekalibraci struktury a parametrů ve chvíli, kdy přichází nová pozorování. Tréninková data se tedy aktualizují přidáním jednoho pozorování po každé predikci.

K měření predikční výkonnosti se využívá pět statistických ukazatelů, střední kvadratická chyba, střední absolutní chyba a střední absolutní procentní chyba, které vyjadřují tři aspekty chyb predikce, index shody, což je standardizovaná míra chyby predikce a směrová symetrie, která měří výkonnost směru trendu.

Z výsledků predikce je znatelné výrazné zlepšení výkonnosti modelu ARIMA díky dynamické prognostické strategii, která převyšuje i predikční modely MLR. I u modelů strojového učení je viditelné zlepšení výkonnosti dynamickou strategií. Statistické modely mají lepší výsledky u předpovědí na 1 a 5 kroků, naopak modely strojového učení mají lepší výsledky u predikce na 10 kroku. Z toho plyne, že jsou statistické techniky spíše využitelnější pro krátkodobé predikční úlohy, zatímco modely strojového učení pro dlouhodobější, především z důvodu lepší schopnosti objevit nelineární vztahy v datech časových řad. U metod strojového učení je nejvýkonnější regrese s podpůrnými vektory, následně gradient boosting regression tree a poté neurální síť.

Výsledky experimentu ukázaly, že regrese s podpůrnými vektory převyšuje ostatní předpovědní modely. ARIMA je však také konkurenceschopným modelem, zejména v krátkodobých předpovědních úlohách, kde dokáže zachytit sezónnost. Proto mají ARIMA a SVR mírné výhody oproti ostatním modelům. Z výsledků také plyne, že statistické modely dosáhly lepších výsledků indexu shody a směrové symetrie a modely strojového učení dosáhly lepších výsledků RMSE, MAE a MAPE. Modely strojového učení mají stabilnější predikce. Objemy denního leteckého nákladu jsou ovlivněny mnoha faktory a dlouhodobými výkyvy. Pro budoucí predikci je možné zahrnout data z více zdrojů.

6) Air cargo demand in Africa: Application of cointegration and error correction modelling techniques [7]

Článek Akinyemi zkoumá krátkodobé a dlouhodobé dopady reálného důchodu, vývozu zboží, přímých zahraničních investic, reálné úrokové míry a cen leteckého paliva na poptávku po leteckém nákladu ve čtyřech afrických zemích v období 1980-2019. Jedná se o Egypt, Nigerii, Keňu a Jihoafrickou republiku, což jsou i přední trhy letecké nákladní dopravy v Africe. Země byly vybrány tak, aby byly ve studii reprezentovány všechny části kontinentu dle světových stran. Krátkodobá i dlouhodobá dynamika jsou integrovány do jednoho modelu. Růst letecké dopravy v Africe je spojen s hospodářským růstem, prováděním regulačních reforem, které zvýšily přístup na trh, toky investic a snížily náklady.

V článku jsou zkoumána roční panelová data za období 1980-2019. Jedná se o proměnné: údaje o objemu letecké cargo dopravy vyjádřené v milionech tunokilometrů (ATK); HDP na obyvatele značící reálný příjem; vývoz zboží, což je ukazatelem mezinárodního obchodu (v USD); PZI znamenají investice zahraničních investorů do domácí ekonomiky; reálná úroková míra v %, což je procentní nárůst hodnoty kupní síly úroků placených z investic a údaje o cenách leteckého paliva typu kerosin v USA (USD za galon), které znázorňují variabilní složku nákladů na provoz leteckých společností. Proměnné jsou převedeny na přirozené logaritmy za účelem snížení směrodatné odchylky ve vzorku a možnosti interpretace odhadů jako elasticit.

V metodice se jedná o ekonometrický model ke zkoumání krátkodobých a dlouhodobých vztahů mezi poptávkou po leteckém cargu a proměnnými. V prvním kroku se provádí test stacionarity, aby byl určen řád integrace a stacionarity proměnných, aby bylo zabráněno falešným regresím. Je použit Dickey-Fullerův zobecněný test nejmenšího kořene už dříve modifikovaný pomocí zobecněného přístupu nejmenších čtverců.

Jako model poptávky je využit kointegrační přístup autoregresního modelu rozložených časových zpoždění ARDL, který je výhodný v tom, že jej lze využít bez ohledu na řád integrace regresorů neboli nezávislých proměnných. Model funguje lépe v případě menších vzorků. Pro odhad elasticity poptávky je zahrnut náhodný chybový člen. Znamínka všech proměnných kromě leteckého paliva jsou kladná, neboť se předpokládá, že zvýšení všech proměnných povede ke zvýšení objemu letecké přepravy. U růstu cen leteckého paliva se předpokládá, že povede k poklesu objemu letecké cargo dopravy, tudíž koeficient bude záporný. V postupu je využit také model korekce chyb rovnice za účelem určení krátkodobé dynamiky mezi proměnnými.

Pokud mezi proměnnými existuje dlouhodobý vztah, vybere se vhodná délka zpoždění pro každou proměnnou, aby se zabránilo sériové autokorelaci a nenormálnosti. Stabilita odhadovaných parametrů je stanovována především proto, že nestabilní parametry mohou vést k chybné specifikaci modelu a zkresleným výsledkům. Po odhadu modelu korekce chyb se rezidua testovala pomocí testů založených na reziduích rekurzivní regrese. Díky těmto testům je zahrnuta krátkodobá dynamika do dlouhodobých koeficientů právě prostřednictvím reziduí. Nakonec jsou provedeny diagnostické testovací statistiky pro sériovou korelaci, heteroskedasticitu (když podmíněný rozptyl náhodná veličiny není konstantní) a normalitu, aby se určila dobrá shoda modelů.

Výsledky ukazují, že mezi proměnnými v modelu existuje dlouhodobý vztah pro všechny země zařazené do vzorku, pokud je objem letecké přepravy zboží závislou proměnnou. Z diagnostických testů pro sériovou korelaci, heteroskedasticitu a normalitu vychází, že jsou modely stabilní a dobře definované, protože koeficienty nejsou statisticky významné.

Tento článek je pro nedostatek údajů omezen pouze na čtyři země a vliv dalších proměnných může být zkoumán v návaznosti na tuto práci. Autoři se na závěr shodují, že je nutné zavést opatření, která posílí hospodářský rozvoj, obchod a lepší infrastrukturu letišť.

7) Freight traffic of civil aviation volume forecast based on hybrid ARIMA-LR model [20]

V článku Chen-Liu-Ruan-Yue-Long-Yao je na základě modelu autoregresního integrovaného klouzavého průměru (ARIMA) a lineárního regresního modelu (LR) navržen hybridní model ARIMA-LR pomocí vylepšeného Bayesovského kombinovaného modelu. Výhody kombinace obou modelů jsou především ve schopnosti se lépe přizpůsobit změnám způsobeným mimořádnými událostmi a vyšší celkovou přesností, než kdyby byly aplikovány oba modely zvlášť.

Mezi nejpoužívanější algoritmy dle předchozích studií patří neuronová síť, šedý GM, regresní analýza, ARIMA a gravitační model. Každý model však může mít dobrý predikční účinek pouze na data za určité období nebo s určitými charakteristikami. Proto se začalo používat kombinování modelů = vícenásobné modely = hybridní modely. Nedostatky jsou především v přiřazování stejné váhy různým časovým historickým datům a statické váze při řadách s různými fluktuacemi, kvůli čemuž nemohou modely trefně reagovat na mimořádné události. Z toho důvodu je užít zdokonalený Bayesovský kombinovaný model. Zdokonalený tak, aby diferencovaně využíval historickou přesnost předpovědi a věnoval větší pozornost informacím obsaženým v nedávné přesnosti předpovědi. Se změnou časové řady může dynamicky měnit

důraz na přesnost nedávných předpovědí. Model zohledňuje vztah mezi dobou vzniku chyby a referenční hodnotou a vliv kolísání dat na referenční hodnotu chyby historické předpovědi. Hybridní model je vyhodnocen pomocí MSE, MAPE a MSE.

Metodika se skládá nejprve z modelu ARIMA. Klíčový bod pro stanovení modelu ARIMA (p,d,q) spočívá ve výběru tří parametrů (p,d,q) . d je stupeň diferenciacce. Účelem diferenciacce je změnit původní řadu pozorování na stacionární časovou řadu. V tomto článku se k výběru p a q používá Bayesovo informační kritérium (BIC). Dále se metodika skládá z lineárně regresního modelu, aby byly co nejlépe zohledněny změny trendu posledních údajů. V článku jsou pro regresi použity dva datové body z předchozího období a pro další období predikce pouze aktuální časový bod. Následně je užit hybridní ARIMA-LR model jako tradiční Bayesovský kombinovaný model. Kombinací všech metod lze vypočítat pravděpodobnost, že časová řada predikovaná jediným modelem se bude rovnat skutečně pozorované hodnotě. Kombinace vede ke zpoždění v úpravě kreditních hodnot jednotlivých modelů a kombinovaný model je tak díky tomu méně citlivý na změny způsobené neočekávanými situacemi. V článku jsou údajům z předchozích období přiřazovány různé váhy na základě myšlenky "čím kratší je časový úsek od bodu predikce, tím vyšší je referenční hodnota". Podobně jako u myšlenky exponenciálního vyrovnávání se referenční hodnota každého období exponenciálně zvyšuje. Při velkých změnách údajů časové řady bude mít nedávná chyba předpovědi větší váhu při výpočtu kumulativní váhové chyby, čímž se zvýší přizpůsobivost modelu změnám trendu. Při velkém kolísání řady naopak.

Jako data pro predikci se v článku užívá objem cargo dopravy za posledních 150 měsíců za období 1/2010-6/2022, který zveřejňuje Národní statistický úřad v Číně. Prvních 50 měsíců je použito jako historická data, posledních 100 měsíců jako data pro predikci a iteraci. V datech je patrná periodicitata cargo dopravy, která má větší amplitudu výkyvů. Jak bylo již řečeno v předchozích pracích, minimální hodnota se objevuje v únoru, na což má vliv Čínský Nový rok. Od roku 2020 má nákladní doprava v důsledku pandemie Covid-19 mírně klesající trend.

Pro zjištění stacionarity modelu se využil rozšířený Dickey-Fullerův test.

Po vykreslení predikce modelu ARIMA je znatelné, že jen zlehka popisuje trend a některé výkyvy objemu ve sledovaném období. Stále však existují výkyvy, které se v predikci nezohlednily. LR model je citlivý na kolísání dat, takže predikovaná hodnota může odrážet směr nárůstu nebo poklesu skutečně pozorované hodnoty, ale amplituda je často nepřesná.

Po využití postupné predikce dostáváme vykreslení hybridního modelu ARIMA-LR. Model lépe odráží trend sekvence, získaný z modelu ARIMA. Zároveň zvládá lépe kontrolovat

rozsah kolísání predikované hodnoty. Konečné zpracování prokazuje, že pokud má původní sekvence velké výkyvy, hybridní model může lépe předpovídat výskyt výkyvů dat.

Hodnocení přesnosti predikce pomocí tří ukazatelů značí, že čím nižší hodnota, tím přesnější predikce modelu. Oproti samostatným modelům vychází hybridní model ve třech ukazatelech (MAE, MSE, MAPE) lépe (niž) o 1,06,29,02,0,03 než u modelu ARIMA a ve srovnání s modelem LR o 3,00,92,00,0,06. Finální výsledek je tedy takový, že hybridní model má nejpresnější predikci.

8) Econometric Forecasting Models for Air Freight in Indonesia (And How Will It be Affected by Covid-19?) [20]

Článek Suryan – Sari – Amalia – Habillah predikuje nadcházející počet přeprav v letecké dopravě do roku 2030, kdy předpovídá průměrný roční růst za deset let. Také zmiňuje Covid-19, který ovlivnil ekonomiku v zemi, dle autorů se snížila kupní síla, vzrostla nejistota, která způsobila nízké investice a vývoz komodit z Indonésie do světa klesl v důsledku nízké poptávky. Pro predikci je využita jednoduchá lineární regrese.

Z výsledků vychází, že modely prognózování pomocí ekonometrie pro leteckou nákladní dopravu v Indonésii souvisejí s HDP a hospodářskou krizí. Počet obyvatel není při předpovídání zohledněn. Do roku 2030 vzroste celkový počet letecké nákladní dopravy o -14 % (nízký), 0 % (střední) a 14 % (vysoký). Do příštích studií autoři doporučují rozšířit metody a zvýšit množství zkoumaných údajů. Dále tento článek navrhuje vedení letiště rozšířit budování kapacit lidských zdrojů pro efektivní odbavování nákladu.

9) Market and Infrastructure Analysis of Future Air Cargo Demand in China [21]

Článek Jiang – Ren – Hansman analyzuje budoucí poptávku po letecké nákladní dopravě v Číně do roku 2020 a její důsledky pro systémovou infrastrukturu. Využívá k tomu ekonometrické metody predikce. Data jsou vzata z Čínského národního statistického úřadu pro ekonomický/demografický vývoj, Čínského ministerstva zahraničního obchodu a ekonomické spolupráce pro mezinárodní obchod a Čínské generální správy civilního letectví pro leteckou dopravu.

Dle článku je hlavní hnací silou poptávky po leteckém nákladu hospodářský růst, tudíž se dá říct, že souhrnná predikce budoucího leteckého nákladu v Číně je založena na vztahu mezi poptávkou po leteckém nákladu a hospodářským rozvojem. Právě ke stanovení tohoto vztahu je užita ekonometrická metrika. Je provedena predikce budoucí poptávky po leteckém nákladu

na pevninské Číně a konkrétní predikce odbavení nákladu na některých hlavních letištích včetně Hongkongu a Tchaj-wanu. Je stanovena predikce HDP získaná analýzou trendů z historických údajů a prognózami čínské vlády a uznávaných institucí. Pro zohlednění nejistoty hospodářského vývoje jsou uvedeny tři prognózy HDP, vysoká, nízká a základní. Na základě historických údajů a HDP je predikován budoucí objem převáženého nákladu v roce 2020 na konkrétních 21 letištích. Zjištěná míra růstu je u některých letišť dlouhodobě optimistická, protože je predikováno především na základě historických dat. Proto pokud jsou tato data rostoucí, bude rostoucí i predikce. Nakonec byly navrženy změny v infrastruktuře letišť.

10) Air cargo transport demand forecasting using ConvLSTM2D, an artificial neural network architecture approach [3]

Článek Anguita – Olariaga provádí krátkodobou predikci (5 let) v Kolumbii, přičemž bylo zohledněno nejzávažnější období Covidu-19, čímž byl rok 2020. Jedná se o predikci na základě dat z období 1979-2020. Predikce je založena na metodě strojového učení/hlubokého učení (ML/DL), které zahrnuje hybrid konvolučních neuronových sítí a neuronových sítí s rekurentní pamětí, které umožňují časoprostorovou nelineární analýzu, jako jsou prostory s více proměnnými a časové více kroky. Metodika umožňuje identifikovat nejvýznamnější socioekonomické rysy při predikci poptávky po leteckém nákladu a interpretovat výsledky ML/DL získané prostřednictvím variační analýzy různých kombinací rysů. Konkrétně se jedná o model odvozený z umělých neuronových sítí typu ConvLSTM2D („Conv“ pro konvoluční a „LSTM2D“ pro dvojrozměrnou krátkodobou a dlouhodobou paměť). Jde o hybrid mezi konvolučními neuronovými sítěmi (CNN), které jsou vhodné pro extrakci vzorů neměnných prostorových souvislostí, a mezi rekurentními neuronovými sítěmi (RNN), které jsou vhodné pro extrakci vzorů rysů časových souvislostí. Predikční metody na bázi ML/DL jsou výhodné v tom, že mohou zahrnovat více prvků analýzy a mohou podporovat vícerozměrné zpracování (několik vstupních proměnných) a nelineární analýzu, tudíž mají potenciál být přesnější než čistě statistické metody.

Soubor dat je složený z jedenácti vstupních proměnných – dvou leteckých (vnitrostátní a mezinárodní letecké cargo) a devíti socioekonomických (HDP, HDP na obyvatele, počet obyvatel, CPI, saldo zahraničního obchodu, PZI a další) Proměnné byly vybrány za účelem vysvětlení příčin vývoje poptávky po letecké cargo dopravě.

Prvním krokem v postupu je příprava dat, která zahrnuje normalizaci neboli škálování proměnných na stupnicích mezi 0 (minimální hodnota prvku) a 1 (maximální hodnota prvku),

aby byly srovnatelné jako vstupy pro ConvLSTM2D model. Následně byla aplikována další normalizace, která spočívala v přiřazení standardního rozdělení řadám se střední hodnotou 0 a směrodatnou odchylkou 1. Data byla předem zpracována detekcí a odstraněním odlehlých hodnot a poté normalizována tak, aby bylo zajištěno, že všechny proměnné jsou na stejné stupnici. Druhým krokem je rozdělení historické datové sady na strukturu 5D tenzorů X a 3D tenzorů Y, aby byly přizpůsobeny vstupní a výstupní vrstvě neuronového modelu ConvLSTM2D. Data X a Y byla dále rozdělena na tréninková a testovací data. Bylo zvoleno 90 % z celkových dat pro tréninková data a 10 % pro testovací. Pro zjednodušení byl uvažován scénář predikce s jednou výstupní proměnnou, ačkoli bylo uvažováno s více vstupními proměnnými. Predikce Y se skládá z vektoru časových složek. Vstupní proměnná X zahrnuje všechny vstupní proměnné uvažované v jednotlivých scénářích, 42 historických let od roku 1979 do roku 2020. Model ConvLSTM2D byl testován na trénovací množině a jako hodnotící metriky byly použity střední absolutní procentní chyba (MAPE) a střední kvadratická chyba (RMSE).

K určení nejdůležitějších socioekonomických charakteristik byla provedena variační analýza. K určení optimální délky cyklů a podcyklů, jakož i optimálního určení funkcí a jejich kombinací, právě za účelem nej přesnější predikce. Na množině tréninkových dat byla vycvičena hybridní architektura neuronové sítě ConvLSTM2D. Model se skládá z konvolučních a LSTM vrstev, které se používají k extrakci časoprostorových rysů ze vstupních dat. Časoprostorové rysy jsou extrahovány pomocí hybridní architektury neuronové sítě ConvLSTM2D. Konvoluční vrstvy se používají k extrakci prostorových rysů ze vstupních dat, zatímco vrstvy LSTM se používají k zachycení časových závislostí a vzorů. Kombinací těchto dvou typů vrstev je model schopen extrahovat časoprostorové rysy, které jsou relevantní pro predikci. Model je zkoušen pomocí zpětného šíření a stochastického gradientního sestupu. Zpětné šíření je proces používaný v neuronových sítích k aktualizaci vah neuronů na základě chyby mezi předpovědaným výstupem a skutečným výstupem. Zahrnuje výpočet gradientu nákladové funkce vzhledem k vahám neuronů a následně použití tohoto gradientu k aktualizaci vah v opačném směru gradientu. Tento proces se opakuje iterativně, dokud není chyba minimalizována. Stochastický sestup po gradientu je optimalizační algoritmus používaný ve strojovém učení k minimalizaci nákladové funkce. Zahrnuje aktualizaci vah neuronů na základě gradientu nákladové funkce vzhledem k náhodně vybrané podmnožině trénovacích dat. Po dokončení dávek tréninkových párů vstupu X a výstupu Y se proces opakoval ve 400 cyklech.

Z výsledků vyplývá, že výkonnost modelu se v průběhu času zlepšuje, přičemž s rostoucím počtem cyklů se snižují hodnoty MSE a RMSE pro testovací soubor. Zároveň je výkonnost

lepší na trénovací množině než na testovací množině, což je běžný problém ve strojovém učení a zdůrazňuje potřebu pečlivého vyhodnocení a validace modelu. Z výsledků také vyplývá, že delší tříleté subcykly a zahrnutí socioekonomických proměnných do modelu zvyšuje přesnost predikce poptávky po leteckém nákladu. Nejdůležitějšími proměnnými jsou HDP na obyvatele, vývoz a HDP. Konkrétně poptávka po mezinárodní letecké nákladní dopravě je silně závislá na HDP a HDP na obyvatele, zatímco poptávka po vnitrostátní letecké nákladní dopravě je významně závislá na HDP na obyvatele. Výsledky také ukazují velmi rychlé oživení poptávky po leteckém nákladu při prepandemických sazbách. Výkonnost modelu je citlivá na volbu délky subcyklu a kombinaci prvků a je třeba provést další výzkum k optimalizaci těchto parametrů, zanesení proměnných do modelu však zvyšuje přesnost predikce.

Výsledky rešerše článků byly vloženy do tabulky 1, aby byly přehledně viditelné užití metodiky a závěry jednotlivých článků, jak z pohledu přesnosti jednotlivých metod, tak vlivu proměnných na predikci nákladní letecké dopravy. Pokud bylo uvedeno doporučení přínosné pro tuto práci, je do tabulky zaneseno také.

Číslo článku	Metodika pro predikci	Shrnutí výsledků, vliv proměnných na predikci
1	Potluck problem, algoritmus určení vážené většiny s různými prediktory	Ekonomické modely růstu průmyslu a HDP jsou do jisté míry užitečné, ale sezónní a cyklická složka mají na časovou řadu větší vliv.
2	Hybridní modely, kombinace 2 lineárních a 2 nelineárních modelů – TSR a ARIMAX + NN a SVR	Hybridní modely poskytují přesnější predikci než jednotlivé modely zvláště, NN přesnější než SVR
3	Model lineární regrese a neuronové sítě	Neuronové sítě poskytly přesnější výsledky, HDP na obyvatele a saldo zahraničního obchodu mají vliv na predikci poptávky
4	Lineární a kvadratická regresní analýza, Holt-Winters algoritmus	Metoda není příliš přesná, chybí další proměnné, politické vlivy a regulace
5	5 metod – MLR a ARIMA statistické, NN, GBRT a SVR modely strojového učení	ARIMA je ze statistických metod nejpřesnější, MLR je u strojového učení nejpřesnější, statistické metody jsou přesnější na krátkodobější predikci, strojové učení na dlouhodobou predikci
6	Ekonometrický kointegrační přístup autoregresního modelu rozložených časových zpoždění	Nedostatek údajů pro sledované země v Africe, doporučení na přidání více proměnných
7	Hybridní model ARIMA-LR	Výhody hybridu: schopnost se lépe přizpůsobit změnám způsobeným mimořádnými událostmi a vyšší přesnost než u jednotlivých modelů, lépe odráží trend sekvence, zároveň lépe kontroluje rozsah kolísání predikované hodnoty.
8	Lineární regrese	HDP a hospodářská krize mají vliv na predikci, nedostatečný počet proměnných, autoři doporučují "rozšířit metody"
9	Ekonometrická predikce – analýza trendů z historických dat ve spojení s predikovaným HDP	Hospodářský růst posiluje poptávku po leteckém cargu, HDP je nejdůležitější ukazatel
10	Model strojového/hlubokého učení s hybridem neuronových sítí a neuronových sítí s rekurentní pamětí	Model je výhodný v tom, že může zahrnovat nelineární i vícerozměrnou analýzu. Souhlasí, že zahrnutí socioekonomických proměnných zvyšuje přesnost predikce poptávky po leteckém nákladu. Nejdůležitějšími proměnnými jsou HDP na obyvatele, saldo zahraničního obchodu a HDP.

Tabulka 1 - Shrnutí rešerše článků na téma predikce (vlastní zpracování)

3. Metodika a výsledky

V této části jsou nejprve zvolené letištní páry, pro které bude následně provedena predikce. Jsou zvolené i socioekonomické proměnné, které budou později zanesené do modelů. Zároveň jsou zde uvedené zdroje, odkud byly data čerpány. Metodika je rozdělena na čtyři části. V první je na základě statistického modelu ARIMA predikován vývoj převáženého objemu na sledovaných trasách. V druhé části je zvolen model GLM pro stanovení, které z proměnných jsou pro predikci převáženého objemu nákladu relevantní a které ne. Ve třetí části dochází k nové predikci již se zanesenými socioekonomickými proměnnými na základě kombinace modelů ARIMA a GLM. Ve čtvrté části dochází k odhadu navýšení či snížení počtu letů na základě predikce z předchozího kroku.

3.1 Letecký trh zvolený pro predikci nákladu

Nejprve je potřeba vyselektovat letištní páry, pro které bude následně vytvořena predikce. Konkrétním trhem, na který je metodická část zaměřena, je Německo. Jedná se o jeden z největších trhů v Evropě s největším cargo uzlem ve Frankfurtu, takže je pravděpodobné, že bude dostatečný objem informací a zároveň jsou data pro trh dobře dostupná. [23] Pro pozdější zapojení socioekonomických proměnných je vhodné i to, že se jedná o největší ekonomiku v Evropě dle výsledků HDP za rok 2022. [24] Díky velikosti leteckého cargo trhu a ve spojitosti s tím i velikostí ekonomiky země je možné predikovat větší počet tras, než by bylo možné v případě menších trhů. Druhým trhem, se kterým budou trasy z německých letišť tvořit letištní páry, je Čína. Jedná se o největší asijský trh letecké nákladní dopravy, tudíž bude množství dostupných dat dostatečné. [25] Zároveň budou dobře vysvětlitelné ekonomické či popřípadě politické vlivy působící na dosavadní vývoj linek.

3.2 Data užitá při predikci nákladu

Veškerá data spojená s leteckým nákladem, která se v práci objevují, jsou vzata z Eurostat databáze. Eurostat poskytuje snadný přístup ke statistikám, zároveň širokou škálu důvěryhodných dat. Konkrétně pro jednotlivé trasy se jedná o data z datasetu Freight and mail air transport between the main airports of Germany and their main partner airports (routes data), avia_gor. Výřez z nezpracovaných dat z toho datasetu je na obrázku 3. [26] Pro predikci celého trhu Německo-Čína byla využita data z datasetu International extra-EU freight and mail air transport by reporting country and partner world regions and countries, avia_goe. [27]

Freight and mail air transport between the main airports of Germany and their main partner airports (routes data)		Settings: Default							
(online data code: avia_gor_de)									
Source of data: Eurostat									
AIRP_PR		TIME	2021-Q2	2021-Q3	2021-Q4	2022-Q1	2022-Q2	2022-Q3	2022-Q4
FRANKFURT/MAIN airport - UNKNOWN airport - UHINA	:	:	:	:	:	:	:	:	:
FRANKFURT/MAIN airport - BEIJING/CAPITAL airport	.7	23 922.9	25 025.2	31 223.3	24 567.9	26 960.7	24 321.5	18 945.8	:
FRANKFURT/MAIN airport - TIANJIN/BINHAI airport	:	:	:	:	:	:	:	:	:
FRANKFURT/MAIN airport - GUANGZHOU/BAIYUN air...	.8	13 073.8	10 794.3	10 132.2	10 673.9	15 492	10 965.4	13 520.5	:
FRANKFURT/MAIN airport - SHENZHEN/BAOAN airport	.9	3 695.8	3 067.8	3 563.3	3 395.7	4 244.4	3 195	4 371	:
FRANKFURT/MAIN airport - ZHENGZHOU/XINZHENG8	1 037.9	2 466.5	1 324.7	1 204.4	2 303.4	117	:	:
FRANKFURT/MAIN airport - WUHAN/TIANHE airport	.4	6 500.5	5 444.7	4 687.9	3 757.9	5 528.5	5 098.1	6 334	:
FRANKFURT/MAIN airport - airport SYSTEM - SHANG...	:	:	:	:	:	:	:	:	:
FRANKFURT/MAIN airport - FUZHOU/CHANGLE airport	.1	2 218.1	661.2	:	:	:	:	:	:
FRANKFURT/MAIN AIRPORT - HANGZHOU/XIAOSHA...	:	:	:	:	:	:	:	:	:
FRANKFURT/MAIN airport - NINGBO/LISHE airport	.9	2 233.4	548	:	:	:	:	:	:
FRANKFURT/MAIN airport - NANJING/LUKOU airport	:	:	:	:	:	:	:	:	:
FRANKFURT/MAIN airport - SHANGHAI/PUDONG airp...	.5	59 969	43 618.2	47 945.8	50 095.4	39 128.8	54 163.8	57 031.6	:
FRANKFURT/MAIN airport - QINGDAO/LIUTING airport	:	:	:	:	:	:	:	:	:
FRANKFURT/MAIN airport - SHANGHAI CITY airport	:	:	:	:	:	:	:	:	:
FRANKFURT/MAIN airport - CHONGQING/JIANGBEI ai...	27	2 796.9	2 783.4	3 566.9	3 248.7	5 246.8	3 449.9	2 448.2	:
FRANKFURT/MAIN airport - CHENGDU/SHUANGLIU ai...	.4	3 507.8	2 915.4	3 605.3	1 917.4	1 780.6	1 727.9	1 796.4	:
FRANKFURT/MAIN airport - SHENYANG/TAOXIAN airp...	:	:	:	:	:	:	:	:	:
HAMBURG airport - BEIJING/CAPITAL airport	:	116	:	:	:	:	:	:	:
HAMBURG airport - WUHAN/TIANHE airport	:	112.5	:	:	:	:	:	:	:
HAMBURG airport - NANJING/LUKOU airport	29	229.1	:	:	:	:	:	:	:
HAMBURG airport - SHANGHAI/PUDONG airport	.8	:	:	2 289.6	2 140	:	:	:	:
HAMBURG airport - URUMQI/DIWOPU airport	:	:	:	:	:	:	:	:	:
KOELN/BONN airport - GUANGZHOU/BAIYUN airport	.5	3 525.8	3 509.8	4 872.3	3 316.6	2 262.6	2 465.3	3 349.6	:
KOELN/BONN airport - SHENZHEN/BAOAN airport	.8	4 600.2	4 534.7	4 499	3 538.4	4 817	6 023	9 526.4	:
KOELN/BONN airport - ZHENGZHOU/XINZHENG airport	:	:	:	:	519.3	1 222.3	:	:	:
KOELN/BONN airport - SHANGHAI/PUDONG airport	.3	606.8	526.1	1 583.6	2 166.2	715.8	745.4	3 009.3	:
KOELN/BONN airport - SHANGHAI CITY airport	:	:	:	:	:	:	:	:	:

Obrázek 3 - Ukázka původních nezpracovaných dat převáženého nákladu mezi Německem a partnerskými letišti v Číně [27]

Jednotky, ve kterých jsou data z Eurostatu vzata, jsou tuny a jedná se konkrétně o „Freight and mail loaded and unloaded“. Pro přehled, hlavními ukazateli leteckého nákladu jsou nákladní tunokilometry (FTK), dostupné tunokilometry (ATK), příjmové tunokilometry (RTK), koeficient zatížení a objem naloženého a vyloženého leteckého nákladu, který se obvykle měří v metrických tunách. První tři parametry lze převést z kilometrů na míle podle původu údajů. FTK, ATK, RTK a koeficient zatížení poskytují informace o efektivitě využívání nákladních letadel a o jejich příjmovém generování. Objem naloženého a vyloženého leteckého nákladu pak ilustruje tok nákladu na letištích a po jednotlivých trasách.

Data ohledně socioekonomických proměnných, která budou využita v predikci, jsou brána z vícero zdrojů. HDP Německa je vzato ze Spolkového statistického úřadu ze serveru Genesis-online. Soubor má název „Volkswirtschaftliche Gesamtrechnungen“ a jde o záložku 81000-001. Jedná se o kvartální data v miliardách eur, která jsou cenově, sezónně i kalendářně očištěna, což je pro model podstatné. [28] Druhou proměnnou je Reálný efektivní směnný kurz na základě indexu spotřebitelských cen. Data jsou cenově očištěna, proto se jedná o reálný kurz. Jedná se o index a data jsou vzata z databáze Mezinárodního měnového fondu. Jedná se o kvartální data pro Německo. [29] Třetí proměnnou je Saldo zahraničního obchodu. Data jsou ze souboru číslo 51000-0020, „Aus – und Einfuhr (kalandre- und saisonbereinigt) (Außenhandel): Deutschland, Monate“, jsou uvedena

v tisících EUR a jsou kalendářně a sezónně očištěna. [30] Jedná se o data exportu a importu, tudíž saldo vznikne odečtením importu od exportu. [31] Hodnoty v datasetu jsou měsíční, tudíž bylo následně nutné vytvořit průměr ze tří měsíců, aby vznikly kvartální hodnoty, které jsou následně vloženy do modelu.

Z časového hlediska bylo pro trasy vyselektováno 32 hodnot. Byly zvoleny čtvrtletní časové úseky od počátku roku 2015 do konce roku 2022, jedná se tedy o 32 čtvrtletí. Kvartální data jsou oproti měsíčním či ročním vhodná především kvůli možnosti zapojení socioekonomických proměnných, které nejsou v měsíčních či ročních hodnotách všechny dohledatelné v takovém stavu, v jakém je třeba je využít. Zároveň je potřeba dodat, že pokud by se jednalo o měsíční hodnoty, časová řada, která by vznikla, by kvůli množství údajů měla ideálně začínat v období aktivního Covidu-19, což by mohlo negativním způsobem ovlivnit stabilitu a přesnost predikce. V případě ročních dat může zase v predikci uniknout zachycení krátkodobých výkyvů nebo sezónních vzorců vývoje.

Pro predikci vývoje spojení trhů byly zvolené stejné parametry, jen jiný dataset, jak již bylo zmíněno dříve. Také se jedná o kvartální hodnoty 2015-2022, parametr Freight and mail loaded and unloaded a jednotky jsou tuny. Geopolitická reportující entita je Německo, partnerská entita Čína včetně Hong Kongu.

3.3 Volba letištních párů

Z datasetu byla vyselektována data potřebná pro další predikci. Pro predikci konkrétních tras byla ze sekce airport pairs (routes) vybrána jen letiště na území Německa. Na druhé straně letištního páru byla zvolena letiště v Číně (vč. Hongkongu). Po aplikaci tohoto kritéria zbylo z 1883 tras 84, spojující jen letiště mezi Německem a Čínou. Letiště v Německu jsou Berlin – Brandenburg, Erfurt – Weimar, Frankfurt am Main, Frankfurt – Hahn, Hamburg, Köln Bonn, Düsseldorf, München, Nürnberg, Leipzig Halle, Stuttgart, Berlin – Tegel a Hannover. Čínská letiště jsou Beijing/Capital, Shanghai/Pudong, Dalian/Zhoushuizi, Guangzhou/Baiyun, Shenzhen/Baoan, Wuhan/Tianhe, Nanjing/Lukou, Tianjin/Binhai, Zhengzhou/Xinzheng, Fuzhou/Changle, Hangzhou/Xiaoshan, Ningbo/Lishe, Qingdao/Liuting, Shanghai City, Chongqing/Jiangbei, Chengdu/Shuangliu, Shenyang/Taoxian, Urumqi/Diwopu, Xi'an City, Nantong/Xingdong, Macao/Intl, Dongsheng, Wuxi a Xi'an/Xianyang.

Data 84 letištních párů, které byly vyselektovány v datasetu, byla následně očištěna. Cílem tohoto procesu je zajistit, že data, která jsou dále používána pro modelování, jsou kompletní, spolehlivá a vhodná pro stanovené účely. Limitujícím faktorem zde byl nedostatek záznamů i

delší časové období, kdy byly některé z linek ukončeny nebo během tohoto času započaty. Na to měla pravděpodobně vliv i pandemie Covid-19. Po odstranění tras, kde nebyla data kompletní, zbylo 6 letištních párů, které obsahují údaje od Q1 2015 po Q4 2022 bez přerušení. Jedná se o páry vypsané v tabulce 2. Z těchto kompletních časových řad je následně možné vytvořit jednotlivé predikce tak, aby byly všechny na stejné úrovni a mohly tak být i analyzovány.

Letiště Frankfurt nad Mohanem – Pekingské mezinárodní letiště
Letiště Frankfurt nad Mohanem – Mezinárodní letiště Kanton Paj-jün
Letiště Frankfurt nad Mohanem – Mezinárodní letiště Šanghaj Pchu-tung
Letiště Kolín/Bonn – Mezinárodní letiště Kanton Paj-jün
Letiště Kolín/Bonn – Mezinárodní letiště Šen-čen Pao-an
Letiště Kolín/Bonn – Mezinárodní letiště Šanghaj Pchu-tung

Tabulka 2 - Letištní páry vyselektované pro predikci

Většina praktické části byla vytvořena v programu R, což je softwarové prostředí pro statistické výpočty a grafiku. U šesti zvolených tras je v následující části predikován objem naložených a vyložených tun nákladu pro roky 2023 a 2024. Jedná se o krátkodobou predikci na dva roky i z důvodu objemu množství dat, které jsou v práci využité. Ve statistické praxi je obecným pravidlem, že by se mělo ideálně predikovat na období zhruba 10 % z množství dat, které máme k dispozici. [32] Z 32 kvartálů, které jsou shromážděné, to vychází na 3,2 kvartálů, tudíž téměř celý rok 2023. Aby byl však naznačen i následující vývoj, zvolila jsem i následující čtyři kvartály, aby byla predikce uzavřena posledním kvartálem roku 2024. Je to i z toho důvodu, že k proměnným, které budou užité v další části, jsou brána aktuální data z prvních tří kvartálů 2023 a bylo by vhodné mít větší prostor na následující predikci.

3.4 Predikce vývoje tras na základě statistického modelu ARIMA

Ze statistických metod byl zvolen model ARIMA. Model pro predikci byl stanoven na základě teoretické části a proběhlé rešerše. Ačkoliv jsou modely strojového učení považovány v některých případech za přesnější predikční nástroje oproti statistickým metodám, není to tak vždy a stále se jedná o poměrně nové, ne tak prozkoušené metody, které jsou náročné na vypracování jak časově, tak finančně. Jak již bylo řečeno v teoretické části, pravděpodobně však budou důležitou součástí budoucí predikce. V rešerši je také zmíněno, že jsou statistické techniky využitelnější spíše pro krátkodobé predikční úlohy, zatímco modely strojového učení pro dlouhodobější. V této práci je prováděna krátkodobá predikce a je zvolena statistická metoda. V rámci rešerše je model ARIMA několikrát zmíněn, v jedné z prací je i označen v porovnání s ostatními statistickými modely jako nejpřesněji predikující. Zároveň se jedná o

jeden z nejužívanějších modelů. Ve dvou pracích jsou též zmíněny hybridní modely, které poskytují přesnější predikce než jednotlivé modely zvlášť. I z toho důvodu je model ARIMA ve finálním kroku spojen s modelem GLM, aby vznikl hybridní model ARIMA-GLM.

V roce 2019 vznikla na Ústavu letecké dopravy na FD ČVUT bakalářská práce na téma Konverze osobních letadel na nákladní v souvislosti s přepravní kapacitou, která je rovněž jedním z odborných článků rozebíraných v teoretické části. Cíl práce je rozdílný, práce je však také založena na predikci časových řad letecké nákladní dopravy na vybraných trasách a je k tomu užit Holt-Wintersův algoritmus založený na exponenciálním vyrovnávání. V závěrečném zhodnocení bylo zmíněno, že navazující práce by mohly využít dataset, který nebude zaměřený jen na evropská cargo letiště, ale také na ta zahraniční. Zároveň je v závěru zmíněno, že by další práce mohly zařadit hlubší analýzu s dalšími faktory vloženými do modelování, které budou vést k přesnějším výsledkům predikce vývoje carga. [8] I z toho důvodu jsou zvoleny socioekonomické proměnné a zahraniční trh.

Samotný model ARIMA je využíván pro předpovídání vývoje v časových řadách, kde datová řada reprezentující předpovídané objekty v čase je chápána jako náhodná posloupnost. Matematické modelování tohoto vztahu umožňuje efektivní predikci budoucího vývoje. ARIMA (p, d, q) je základní formou modelu, který může být rozšířen o sezónní složku (P, D, Q) S pro data s periodickými vlastnostmi. Klíčovým aspektem je zpracování nestacionárních informací v časových řadách pomocí diferencování. Nejlepším lineárním predikčním modelem pro krátkodobé předpovědi je model ARIMA s reziduální posloupností připomínající bílý šum. ARIMA je dále rozšířením auto regresivního klouzavého průměru, který překonává omezení na stacionární řady a umožňuje aplikaci přímo na neupravené datové řady. Stupeň difference 'd' zajišťuje, že řada se stane stacionární a umožňuje analýzu a využití statistických metod na datech, která jsou silně korelována v časové posloupnosti, kde konvenční metody selhávají. [33] [34]

Skript z programu R popisující části předcházející predikci modelem ARIMA a model ARIMA je vypsán na obrázku 4.

```

## decomposition ###
decomposedRes <- decompose(ts, type="additive") #"additive" for additive components
plot(decomposedRes) # see plot below
stlRes <- stl(ts, s.window = "periodic")

## autocorrelation ###
acfRes <- acf(ts) # autocorrelation
pacfRes <- pacf(ts) # partial autocorrelation

## stationary test ###
#Use Augmented Dickey-Fuller Test (adf test). A p-Value of less than 0.05 in adf.test() indicates that it is stationary.
adf_test <- adf.test(ts) # p-value < 0.05 indicates the TS is stationary

## arima ###
autoarima <- auto.arima(ts, trace=F, stepwise = FALSE)
plot(forecast(autoarima, h=8))
write.csv2(forecast(autoarima, h=8), file=paste0("./results/ARIMA_forecast_", meta$Short, ".csv", sep=""))

```

Obrázek 4 - Skript z programu R pro ARIMA model

První analyzovanou trasou je Letiště Frankfurt nad Mohanem (FRA)– Pekingské mezinárodní letiště (PEK). U každé trasy je nejprve znázorněn graf dekompozice časové řady dat. (obrázek 5) Dekompozice se obvykle používá ke znázornění různých složek, které tvoří časovou řadu, což může být užitečné pro pochopení základních vzorců a trendů v datech. Na ose X je časové období, kvartály 2015-2022, na ose Y jsou přepravené tuny.

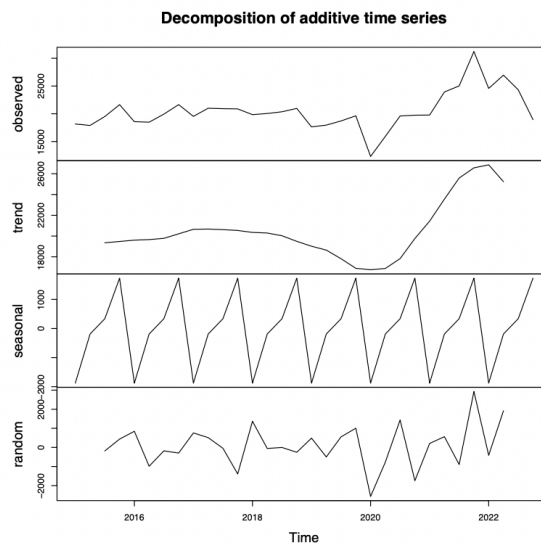
První část grafu představuje původní data, časovou řadu, která je následně dekomponována na trend, sezónnost a po očištění zůstane náhodná složka. Zobrazuje celkovou tonáž přepravenou čtvrtletně na trase Frankfurt-Peking za roky 2015-2022. Čára v čase kolísá, což naznačuje proměnlivost údajů, která může být způsobena sezónními vlivy, trendy nebo nepravidelnými složkami.

Druhou částí grafu je trend, který zachycuje dlouhodobou tendenci časové řady růst nebo klesat. [35] Získává se vyhlazením pozorovaných údajů, aby se zvýraznil dlouhodobý průběh nebo zákonitosti. Na ose Y jsou znázorněné přepravené tuny. U první části grafu je rozpětí tun větší než u druhé. Je to z toho důvodu, že jsou obě části znázorněné na stejné úrovni, ale druhý graf je na nižším rozpětí, protože trend rozpětí zplošťuje a zároveň tam nejsou znázorněny výkyvy jako u původních dat, takže stačí ke znázornění nižší rozpětí. V tomto případě se trend jeví jako relativně plochý v první polovině pozorovaného období. Následně dochází k propadu z důvodu pandemie Covid-19, po níž dochází k prudkému nárůstu následovanému lehkým poklesem.

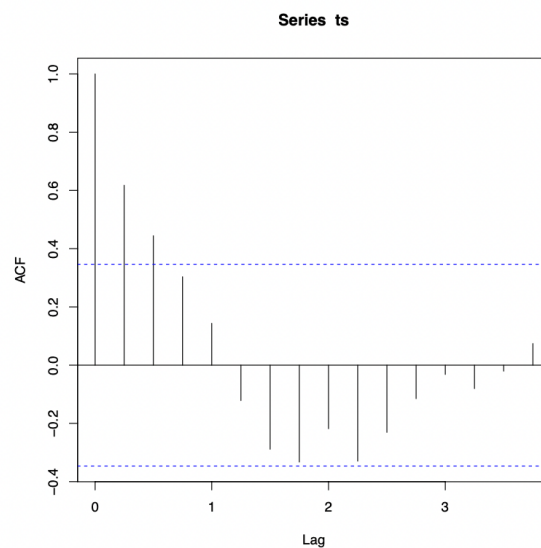
Třetí částí prvního grafu je znázorněna sezónnost, což je pravidelně se opakující fluktuaace související se střídáním ročních období, svátky, množství denního světla atd., které mohou ovlivnit objem přepravy. [35] Sezónnost v podstatě zachycuje zprůměrovaná jednotlivá čtvrtletí a jejich odchýlení od trendové složky. Sezónní složka čtyř kvartálu zde kolísá.

Vzorec se zde zdá být poměrně konzistentní, pravidelně se objevují vrcholy a poklesy, což naznačuje silný sezónní vliv na přepravu nákladu.

Poslední částí je nepravidelná/náhodná/zbytková složka, která představuje výkyvy, které jsou dynamické, nepředvídatelné, nemají systematický charakter a jsou velmi těžce vysvětlitelné. Představuje šum v datech po odstranění trendové a sezónní složky. [35]



Obrázek 5 - Dekompozice časové řady FRA-PEK



Obrázek 6 - ACF na trase FRA-PEK

Druhým grafem, který je na obrázku 6, je ACF = autokorelační funkce, která počítá, jaké je v časové řadě zpoždění neboli jaká je cyklická část/sezónnost. Osa y představuje hodnoty autokorelačního koeficientu (ACF), který měří korelaci časové řady s jejími vlastními zpožděnými hodnotami. Osa x zobrazuje zpoždění z hlediska časové jednotky, což je v tomto

případě rok, který je rozdělen na čtyři části. Modré přerušované čáry představují hranice významnosti, které se vypočítávají podle toho, kolik máme hodnot v časové řadě. Pokud by byla časová řada delší, byla by hodnota významnosti nižší. I z toho důvodu se upřednostňují delší časové řady pro predikce. Spodní hladina významnosti je symetrická k vrchní. V kladných hodnotách je kladná korelace, v záporných je záporná korelace, která může být také statisticky významná, což v tomto případě není. Pokud sloupce přesahují tyto hranice, naznačuje to statisticky významnou korelaci v daném zpoždění kvartálů. Existuje silná autokorelace při zpoždění 0, které je vždy 1, protože řada je dokonale korelovaná sama se sebou při zpoždění 0. Pokud časovou řadu zpozdím o jeden kvartál, tak porovnáme aktuální kvartální hodnoty proti předchozímu období, předchozímu kvartálu. Autokorelace při zpoždění 1 je kladná a zdá se být poměrně významná, což naznačuje, že hodnota řady v jednom čtvrtletí je pozitivně korelována s její hodnotou v předchozím čtvrtletí. S rostoucím zpožděním se koeficienty autokorelace snižují a stávají se nevýznamnými, protože spadají pod hranici významnosti. To naznačuje, že čtvrtletní hodnoty časových řad nejsou významně korelovány s jejich hodnotami ve čtvrtletích dále v minulosti. Pokud je cyklická či sezónní složka zřetelná, vystoupá hodnota až nad hladinu významnosti. Zde je zřetelná cykličnost jen v prvních dvou kvartálech, u nichž se dá říct, že spolu korelují, zatímco další čtvrtletí již ne. Půlroční hodnoty spolu korelují lépe než roční. Časová řada tedy nemá silnou sezónní složku.

Následně je proveden rozšířený Dickey-Fullerův test (ADF), což je běžný statistický test používaný k určení stacionarity časové řady. Je znázorněn na obrázku 7.

Augmented Dickey-Fuller Test

```
data: ts
Dickey-Fuller = -2.1119, Lag order = 3, p-value = 0.5297
alternative hypothesis: stationary
```

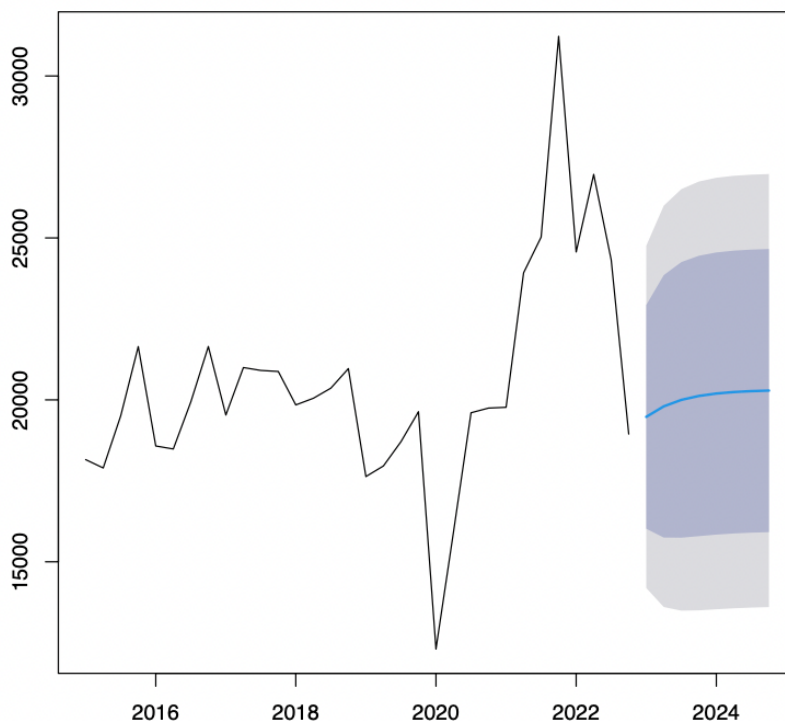
Obrázek 7 - Rozšířený Dickey-Fullerův test pro trasu FRA-PEK

Testovací statistika je -2,1119. Tato hodnota se porovnává s kritickými hodnotami pro Dickey-Fullerovo rozdělení, aby se určilo, zda se má zamítnout nulová hypotéza o nestacionaritě. Nulová hypotéza zní, že časová řada není stacionární, alternativní hypotéza tvrdí, že je stacionární. Vzhledem k p hodnotě vyšší než 0,05 se nezamítá nulová hypotéza, tudíž platí, že časová řada není stacionární.

Pokud je p hodnota nižší než 0,05, časová řada je stacionární. V tomto případě není, což by mělo znamenat, že je v časové řadě přítomný trend či cyklická složka či je trend statisticky významný.

Posledním grafem ke každé trase je predikce pomocí modelu ARIMA. (obrázek 8) Zde se jedná o predikci pomocí modelu ARIMA (1,0,0) s nenulovým průměrem. Plná černá linie představuje historicky pozorovaná data. Z parametrů (1,0,0) vyplývá, že se jedná o autoregresní model (AR) řádu jedna bez diferenciaci a bez složek klouzavého průměru (MA). Nenulový průměr naznačuje, že model předpokládá, že řada má průměr odlišný od nuly, což je logické vzhledem k umístění tun nákladu na ose Y. Predikované údaje jsou znázorněny modrou linií. Kolem prognózy se nachází stínovaná oblast představující interval spolehlivosti pro predikci. Jedná se o rozsah, v němž se s určitou pravděpodobností očekává, že se skutečné hodnoty budou pohybovat. Tmavěji stínovaná oblast představuje užší interval spolehlivosti 80 % a světleji stínovaná oblast představuje širší interval spolehlivosti 95 %. Z predikce vyplývá, že zůstane tonáž nákladní dopravy na této trase během let 2023 a 2024 relativně stabilní s lehkým potenciálem k růstu. Nejsou zde viditelné žádné výkyvy ani sezónnost. Dá se říci, že má na predikci velký podíl i nevysvětlitelná složka časové řady.

Forecasts from ARIMA(1,0,0) with non-zero mean



Obrázek 8 - ARIMA model pro trasu FRA-PEK

Pokud jde o kombinaci parametrů, program R užitý k predikci hledá různé kombinace členů AR (p), I (d) a MA (q) a počítá významnost kombinace parametrů, kdy nakonec vybírá tu nejlepší shodu. Parametr "p" představuje počet autoregresních členů, "d" představuje počet nesezónních diferencí potřebných pro stacionaritu a "q" představuje počet zpožděných chyb předpovědi v rovnici předpovědi. Pro predikci si sám navolí parametry, za kterých bude model nejpřesnější. Ve skriptu na obrázku 4 se jedná o funkci „autoarima“.

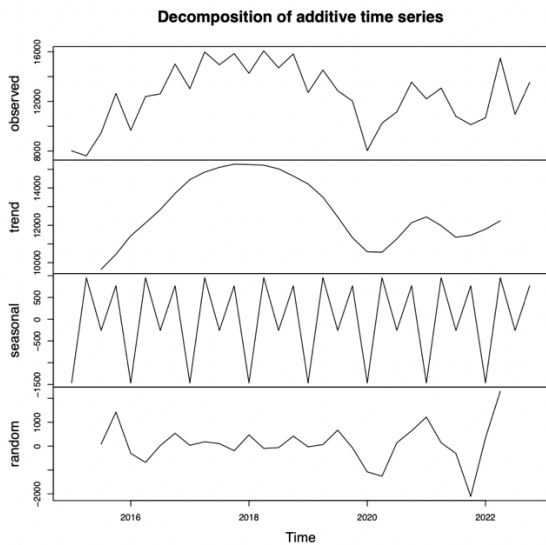
Druhou trasou je Letiště Frankfurt nad Mohanem (FRA)– Mezinárodní letiště Kanton Paj-jün (CAN).

Dekompozice probíhá stejným způsobem jako u první trasy. (obrázek 9) Zde je již vidět viditelný trend, který je nejprve rostoucí, následně zasažený poklesem v období pandemie Covid-19, tudíž klesající, poté opět stoupá, klesá a následně stoupá. Sezónnost znázorňuje jasný vzorec s pravidelnými vrcholy a poklesy, které se každoročně opakují, což naznačuje sezónnost objemu nákladu.

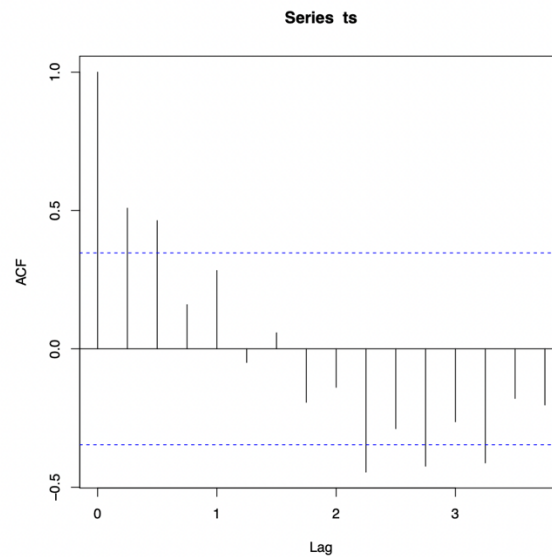
V grafu znázorňujícím autokorelační funkci (obrázek 10) je vývoj podobný první trase, kdy spolu korelují první a druhé čtvrtletí, zatímco další zpoždění se dostávají pod hladinu významnosti. Na druhé straně se zde objevují záporné korelace až po období tří let a jednoho čtvrtletí.

U Dickey-Fullerova testu opět neexistuje s p-hodnotou 0,6385 dostatečný statistický důkaz pro zamítnutí nulové hypotézy, což naznačuje, že časová řada není stacionární.

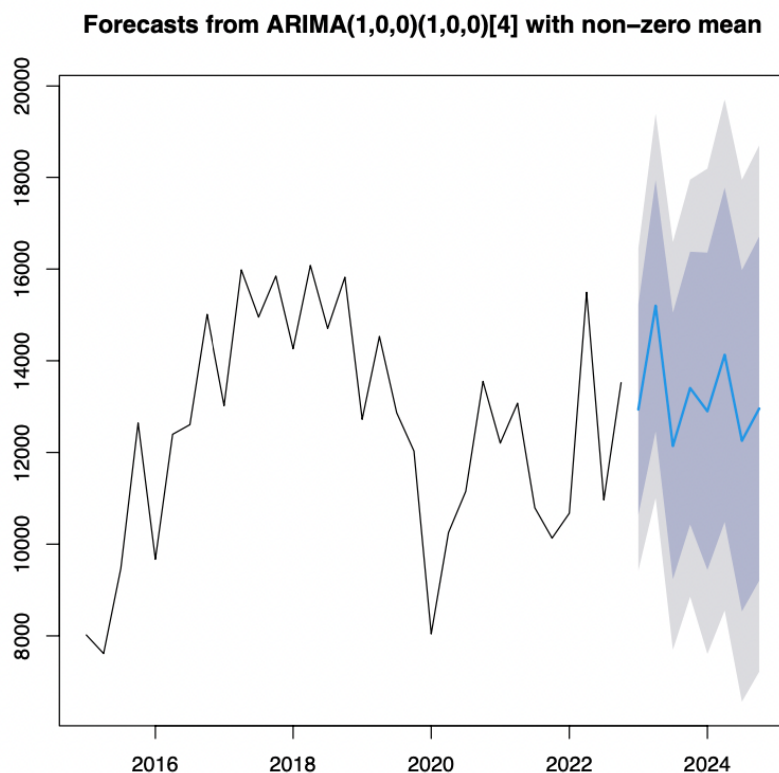
Predikce pro tuto trasu je s parametry (1,0,0)(1,0,0)[4]. (obrázek 11) Jedná se o predikci se zpožděním čtyři, což značí čtyři čtvrtletí, tudíž cyklus jednoho roku. První parametry znázorňují původní parametry, druhé parametry jsou parametry pousnuté časové řady o 4. Jak již bylo řečeno, model může mít v některých případech nejpřesnější predikci právě při zpoždění. Z grafu je v predikci viditelný cyklus, který vzniká právě i na základě tohoto parametru. Predikce je zároveň mírně klesající. Graf naznačuje vyrovnávání objemu nákladu, přičemž intervaly spolehlivosti se v průběhu času rozšiřují, což odráží rostoucí nejistotu v predikci s tím, jak se promítá dále do budoucnosti.



Obrázek 9 - Dekompozice časové řady FRA-CAN.



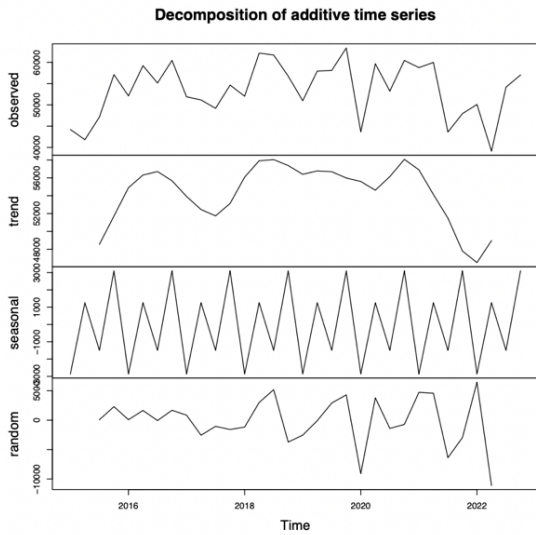
Obrázek 10 - ACF na trase FRA-CAN



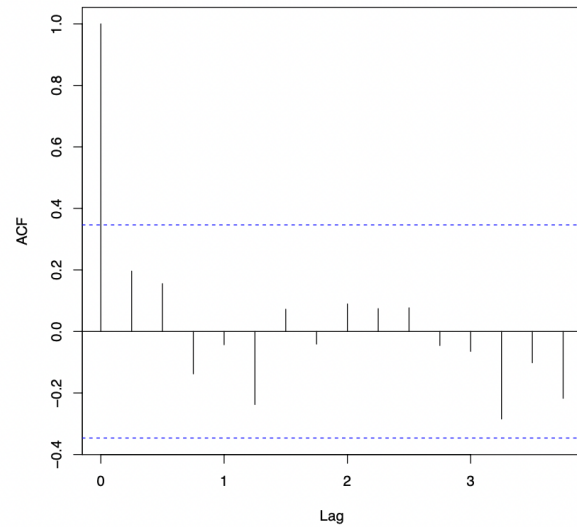
Obrázek 11 - ARIMA model pro trasu FRA-PEK

Třetí trasou je Letiště Frankfurt nad Mohanem (FRA)– Mezinárodní letiště Šanghaj Pchutung (PVG). Složka trendu (obrázek 12) odhaluje základní pohyb v řadě, který se zdá být v čase proměnlivý. Neexistuje zde však žádný jasný dlouhodobý vzestupný nebo sestupný trend, což naznačuje, že i když mohou nastat období růstu nebo poklesu, objemy nákladu se časem vracejí do rovnováhy. Sezónní složka vykazuje konzistentní vzorec vrcholů a poklesů,

kteře se objevují v pravidelných intervalech, což svědčí o cyklické povaze poptávky po nákladu na této trase. Naopak je viditelný velký vliv náhodné složky. V případě autokorelační funkce (obrázek 13) zde není viditelná žádná korelace. Časová řada s p-hodnotou 0,2665 není stacionární. Predikce ukazuje (obrázek 14), že předpokládaný objem nákladu je do budoucna relativně rovnoměrný. Na predikci nemá vliv téměř nic i vzhledem k parametrům (0,0,0).

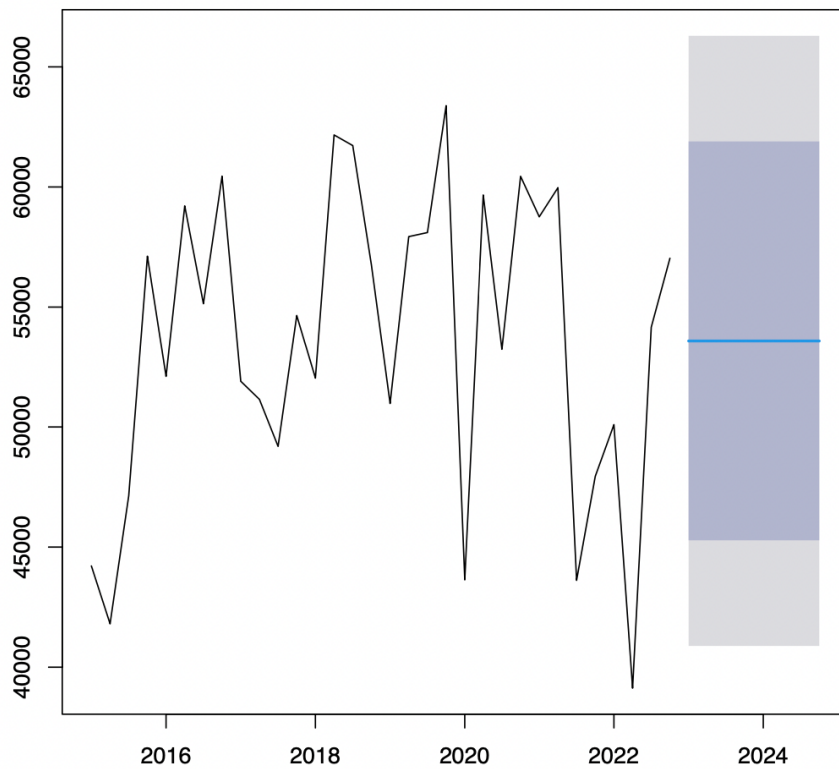


Obrázek 12 - Dekompozice časové řady FRA-PVG



Obrázek 13 - ACF na trase FRA-PVG

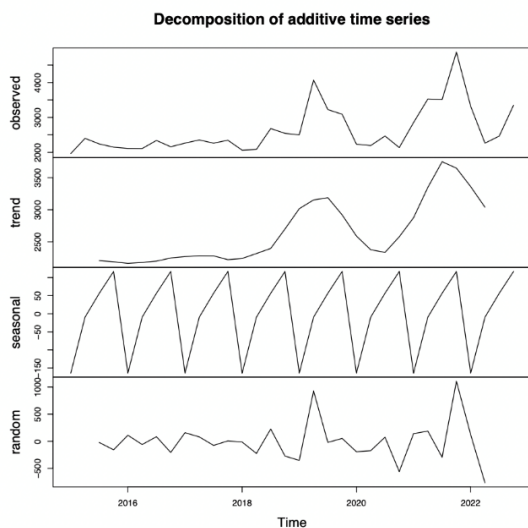
Forecasts from ARIMA(0,0,0) with non-zero mean



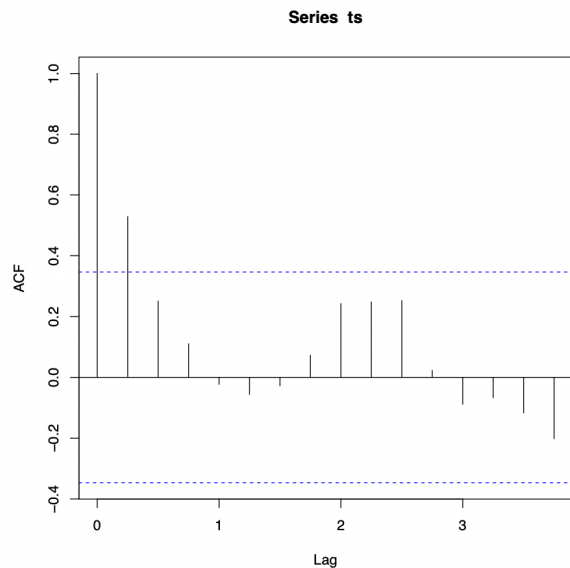
Obrázek 14 - ARIMA model pro trasu FRA-PVG

Čtvrtou trasou je Letiště Kolín/Bonn (CGN)– Mezinárodní letiště Kanton Paj-jün (CAN). Výrazný trend je zde viditelný v druhé polovině časové řady, ale není dlouhodobý, jedná se o pravidelné střídání. (obrázek 15) Je zde i výrazná sezónní složka, která naznačuje silný cyklický vzorec. Autokorelační funkce naznačuje korelaci jen jednoho kvartálu. (obrázek 16) Dickey-Fullerův test odhalil (obrázek 17), že s p-hodnotou 0,01277 lze nulovou hypotézu zamítnout, což naznačuje, že je časová řada stacionární a nemá dlouhodobý trend, přičemž by měla stagnovat kolem průměru.

U predikce (obrázek 18) má model ARIMA parametry (0,1,2). Jak již bylo řečeno, řada naznačuje cykly, z nichž je viditelné, že se ve fázi posledního kvartálu 2022 nacházela časová řada právě v cyklu, v rostoucí fázi. Predikce naznačuje, že bude přeprava pokračovat v ve vzorci pozorovaném v historických údajích, tudíž že klesne, ale následně se ustálí na jedné hodnotě, což by odpovídalo výsledku stacionarity modelu.



Obrázek 15 – Dekompozice časové řady CGN-CAN

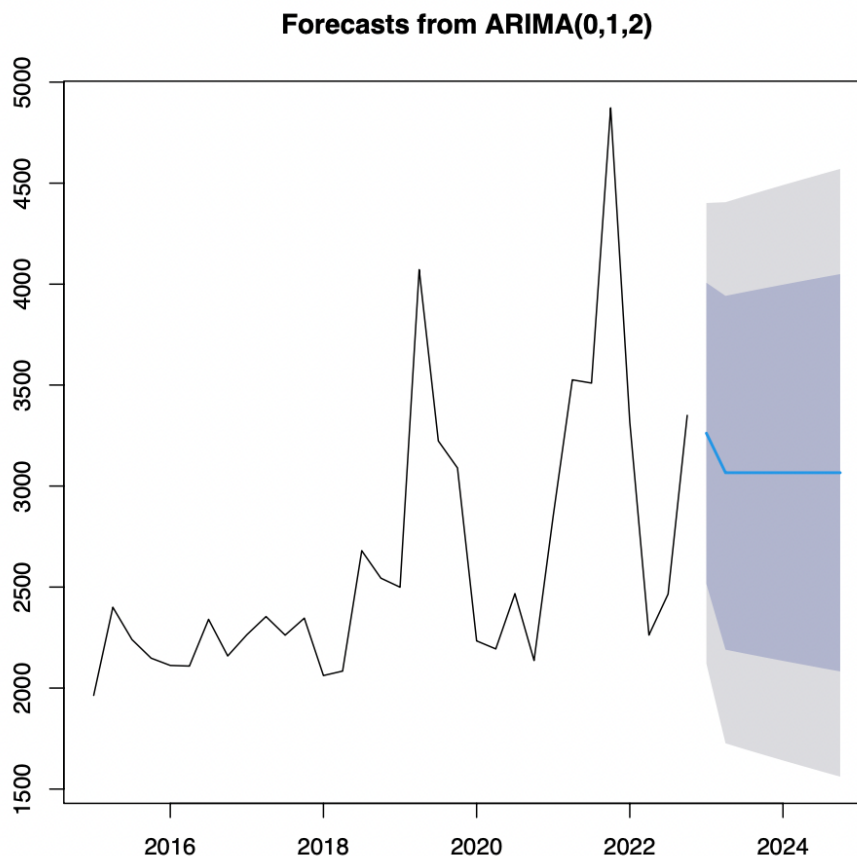


Obrázek 16 - ACF na trase CGN-CAN

Augmented Dickey-Fuller Test

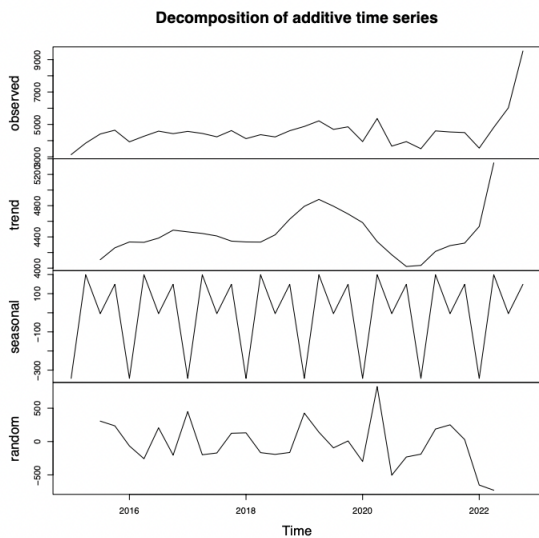
```
data: ts
Dickey-Fuller = -4.2491, Lag order = 3, p-value = 0.01277
alternative hypothesis: stationary
```

Obrázek 17 - - Rozšířený Dickey-Fullerův test pro trasu CGN-CAN

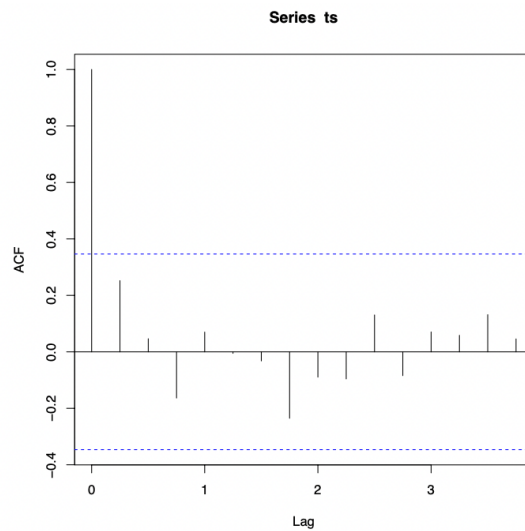


Obrázek 18- ARIMA model pro trasu CGN-CAN

Pátá trasa je Letiště Kolín/Bonn (CGN) – Mezinárodní letiště Šen-čen Pao-an (SZX). Není zde viditelný výrazný trend, až výrazný růst v posledním roce, jinak je velmi proměnlivý. (obrázek 19) Časová řada dle Dickey-Fullerova testu s p-hodnotou 0,9897 není stacionární. Mezi čtvrtletími neexistuje korelace (obrázek 20). Predikce (obrázek 21) s parametry (1,0,2) předpokládá na základě historických hodnot pokles převážených tun téměř zpět na původní hodnoty. Predikce je však s nabývajícím kvartály čím dál tím méně přesná.

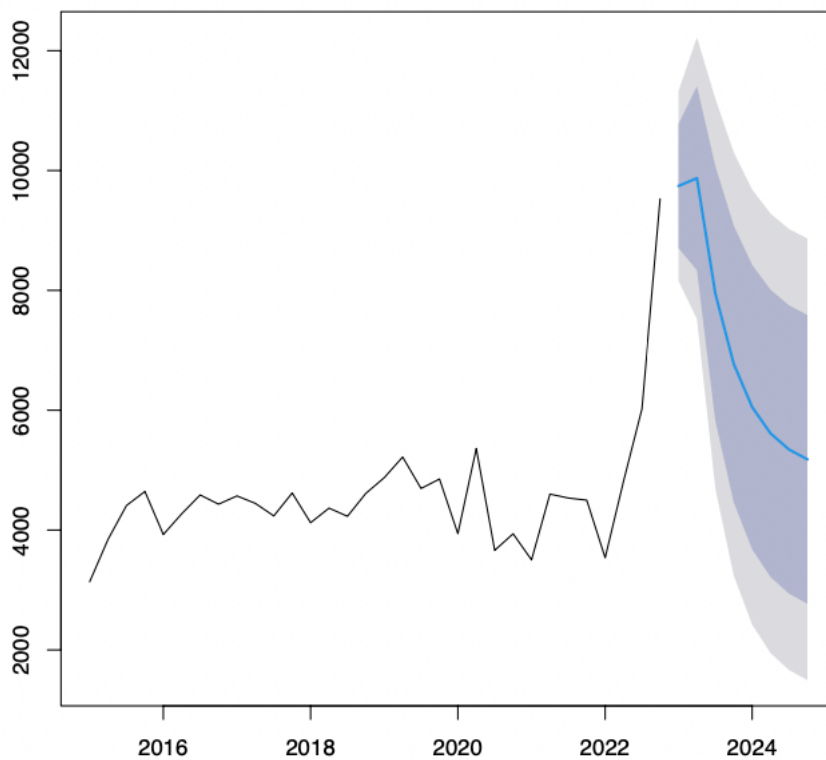


Obrázek 19 - Dekompozice časové řady CGN-SZX



Obrázek 20 - ACF na trase CGN-SZX

Forecasts from ARIMA(1,0,2) with non-zero mean

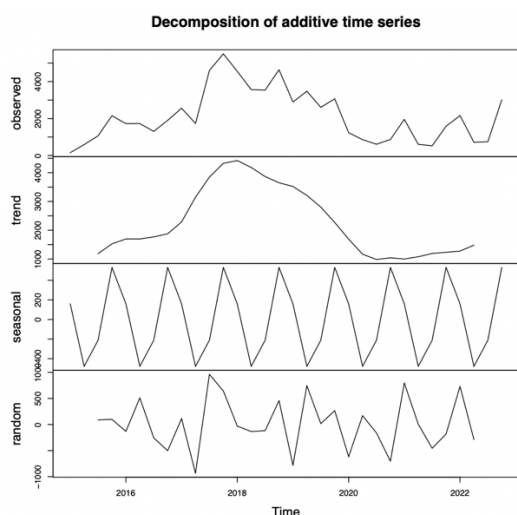


Obrázek 21 - ARIMA model pro trasu CGN-SZX

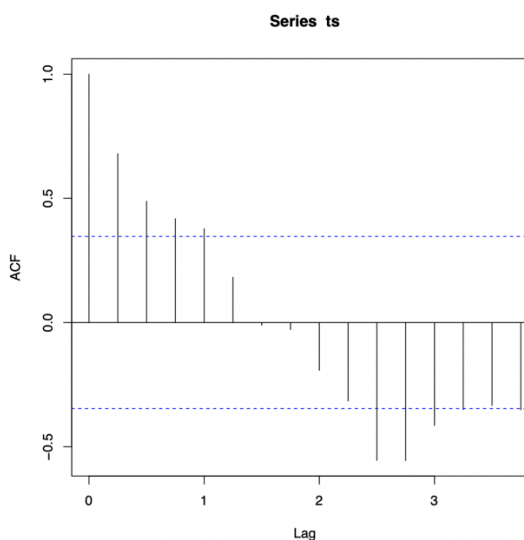
Poslední predikovanou z šesti tras je spojení Letiště Kolín/Bon (CGN) – Mezinárodní letiště Šanghaj Pchu-tung (PVG). V trendu (obrázek 22) je zde vidět poměrně prudký nárůst do roku 2018, po němž byl dlouhodobý klesající trend, který vystřídal nynější stav, což je

několikaleté obnovování po Covidu-19 a velmi mírné stoupání. Je zde viditelná jak sezónnost, tak nejvýraznější korelace ze všech tras, které jsou analyzovány. (obrázek 23) Zpoždění do jednoho roku se nachází nad hladinou významnosti. Dle Dickey-Fullerova testu není časová řada s p-hodnotou 0,7433 stacionární.

Pro predikci (obrázek 24) jsou zvoleny parametry $(1,0,0)(0,0,1)[4]$. Opět se jedná o zpoždění čtyř kvartálů, což je propojené i s korelací ve čtyřech čtvrtletích, což se následně poměrně výrazně projevuje i v predikci. Jak již bylo řečeno, model si generuje parametry, které budou vycházet nejpřesněji a se zpožděním čtyři jsme se již setkali. Pro vysvětlení, datová řada je původně definovaná s frekvencí čtyři, takže i model se při vyhledávání zpoždění pro predikci snaží hledat primárně zpoždění čtyři a násobky. Predikce tedy detailněji popisuje pokles nyní rostoucího převáženého nákladu, který však neklesne na hodnoty, které jsou vidět v historických cyklech, ale opět začne růst, přičemž poté opět poklesne, a to v poměrně krátkém časovém intervalu několika kvartálů.

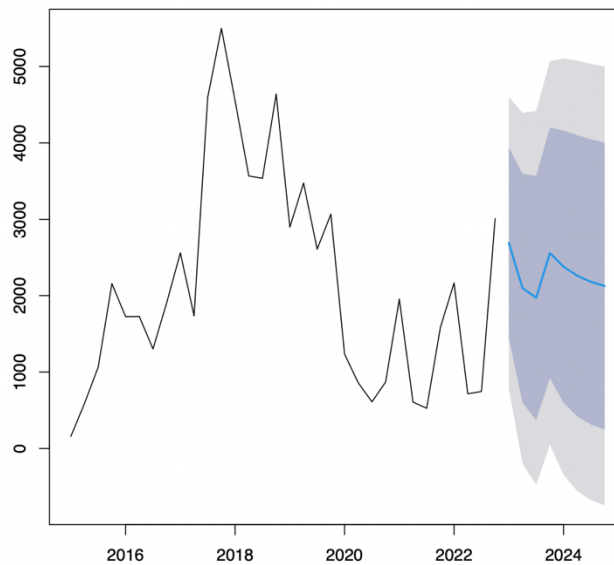


Obrázek 22 - Dekompozice časové řady CGN-PVG



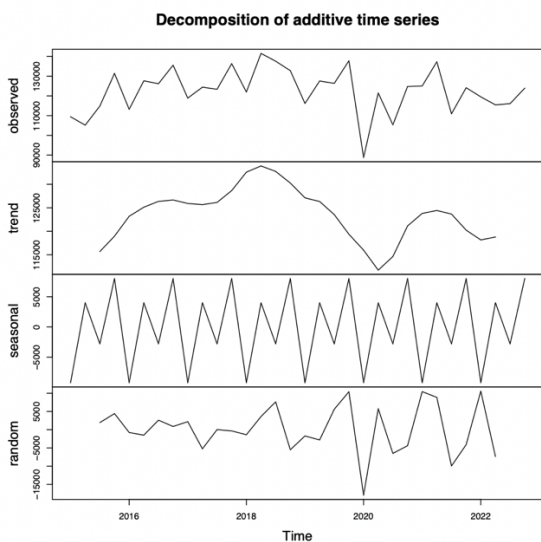
Obrázek 23 - ACF na trase CGN-PVG

Forecasts from ARIMA(1,0,0)(0,0,1)[4] with non-zero mean

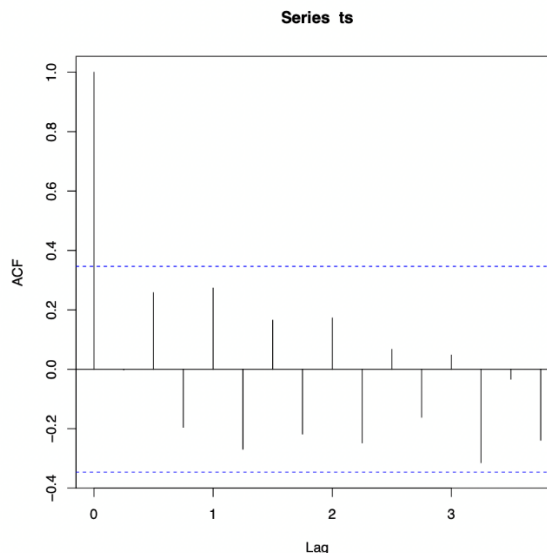


Obrázek 24 - ARIMA model pro trasu CGN-PVG

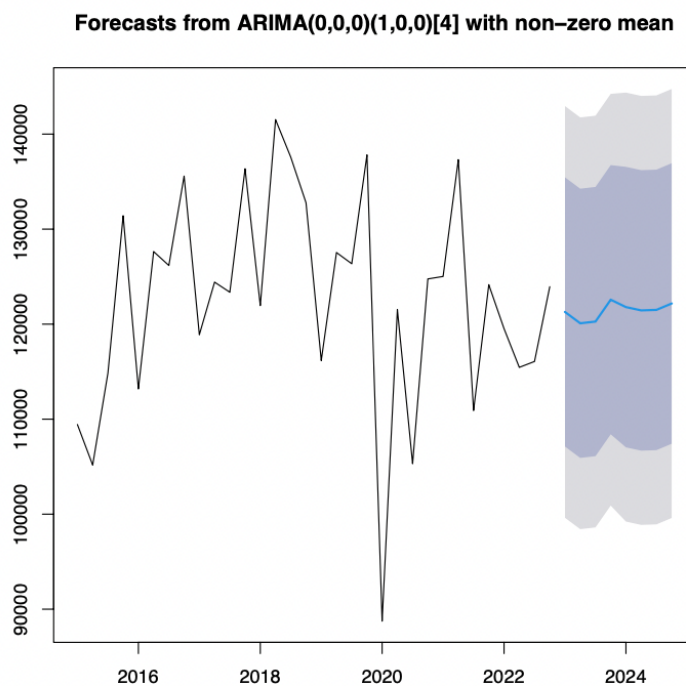
Poslední ARIMA predikcí je vývoj carga na všech trasách spojujících německá a čínská letiště. Trend (obrázek 25) naznačuje výrazné klesání od roku 2018, následované výrazným růstem v roce 2020, které opět přechází v pokles. Není tam tedy pozorovatelný dlouhodobý trend. V posledním kvartálu před predikovaným obdobím došlo k mírnému nárůstu. Náhodný šum naznačuje výrazný vliv na poslední řadu. V případě autokorelační funkce (obrázek 26) zde není viditelná žádná korelace. Časová řada dle Dickey-Fullerova testu s p-hodnotou 0,494 není stacionární. ARIMA (obrázek 27) je opět se zpožděním čtyři a celkovými parametry (0,0,0)(1,0,0)[4] a předpokládá pokračování cykličnosti viditelné v historických datech, ovšem s tendencí lehkého růstu a budoucího ustálení. Predikce spojení těchto trhů v podstatě odpovídá predikci celosvětového vývoje leteckého carga. Zatím se nachází spíše ve stagnaci, ale do budoucna bude pravděpodobně možné pozorovat růst.



Obrázek 25 - Dekompozice časové řady všech tras Německo-Čína



Obrázek 26 - ACF na trasách Německo-Čína



Obrázek 27 - ARIMA model pro trasy Německo-Čína

Predikce vývoje socioekonomických proměnných na základě statistického modelu ARIMA

Jak již bylo zmíněno, do finální predikce vývoje přepravy nákladu budou vloženy tři socioekonomické proměnné. Vzhledem k tomu, že se jedná stejně jako u tras o data jen do roku 2022, bylo potřeba predikovat i jejich vývoj pro rok 2023 a 2024. Po predikci, která byla provedená však zůstaly všechny proměnné na konstantních hodnotách pro všech 8 čtvrtletí. Proto byl nakonec upraven postup tak, že byly dohledány aktuální hodnoty za rok 2023, přičemž byly k dispozici data za Q1, Q2 a Q3. Pro přesnější predikci a možnost získat větší

množství informací, se tedy zkrátil interval predikce jen na 5 kvartálů – Q4 roku 2023 a čtyři čtvrtletí pro rok 2024.

Proměnnými jsou HDP, směnný kurz a saldo zahraničního obchodu. Dle článku Air cargo demand modeling and prediction má ekonomická aktivita hlavní vliv na rozvoj světové letecké nákladní dopravy. [6] Především HDP a zahraniční obchod, což jsou hodnoty, které jsou v zájemném závislém vztahu a ovlivňují poptávku po letecké nákladní dopravě. [11] HDP je označován jako nejužívanější ekonomický faktor pro predikci leteckého carga. Dále mezi ně patří například zahraniční obchod či směnný kurz. [7] Dle článku Air cargo transport demand forecasting using ConvLSTM2D, an artificial neural network architecture má HDP i větší vliv na poptávku po nákladní letecké dopravě než po osobní přepravě. Také potvrzuje, že dalším socioekonomickým faktorem, který ovlivňuje leteckou nákladní dopravu, je směnný kurz. [3]

První proměnnou je tedy HDP. Zásadní informací je, že hodnoty pro 35 čtvrtletí od roku 2015 do třetího kvartálu roku 2023 užitá k predikci, jsou cenově, sezónně a kalendářně očištěné. Úprava o ceny odstraňuje vliv inflace nebo deflace, takže údaje o HDP odrážejí reálné (ne nominální) hodnoty. Sezónní úpravy odstraňují účinky sezónních výkyvy, což jsou opakující se vlivy souvisejících s ročním obdobím, jako jsou prázdniny nebo počasí. Kalendářní úpravy zohledňují výkyvy v počtu pracovních dnů nebo načasování svátků, které mohou ovlivnit ekonomickou aktivitu.

HDP

Dekompozice časové řady (obrázek 28) je opět rozdělena na trend, sezónnost a náhodnou složku. Jak je z grafu patrné, HDP s mírným kolísáním rostlo, dokud nepřišel signifikantní propad kvůli pandemii Covid-19. Poté došlo opět k prudkému nárůstu, kdy se hodnoty nákladu dostaly až nad hranici původních hodnot. Následně dochází ke stabilizaci. Autokorelační funkce (obrázek 29) značí korelaci tří kvartálů, poté jsou všechny hodnoty pod hladinou významnosti. Časová řada dle Dickey-Fullerova testu s p-hodnotou 0,4353 není stacionární.

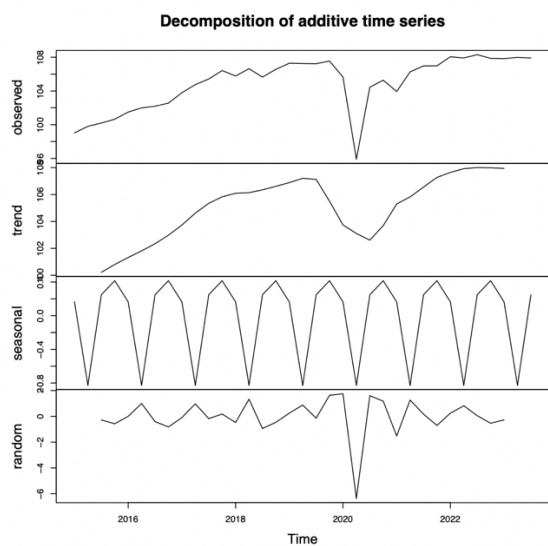
Třetí graf (obrázek 30), který představuje předpovědi z modelu ARIMA (0,1,1), naznačuje pokračující stagnaci, což je viditelné i na hodnotách predikce (tab. 3, první sloupec; poslední čtyři sloupce značí predikci v intervalu spolehlivosti). Na základě historického vývoje HDP v posledních čtvrtletích (tab. 4, hodnoty v miliardách EUR) je viditelné, že nedochází ani k růstu ani k poklesu proměnné, tudíž predikce jen odráží pokračování této časové řady ve stejném duchu.

	Point.Forecast	Lo.80	Hi.80	Lo.95	Hi.95
2023 Q4	107,915543203006	105,063942814475	110,767143591537	103,554396753587	112,276689652424
2024 Q1	107,915543203006	104,626354944689	111,204731461322	102,885163858158	112,945922547853
2024 Q2	107,915543203006	104,240506601634	111,590579804378	102,29505974045	113,536026665562
2024 Q3	107,915543203006	103,891486960332	111,93959944568	101,761280261543	114,069806144468
2024 Q4	107,915543203006	103,570412219137	112,260674186875	101,270238819273	114,560847586738

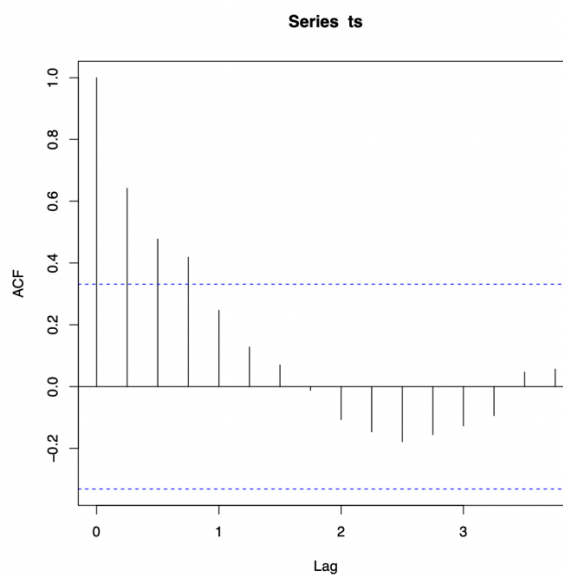
Tabulka 3 - Predikované hodnoty HDP

2022-Q1	2022-Q2	2022-Q3	2022-Q4	2023-Q1	2023-Q2	2023-Q3
108,05	107,91	108,30	107,86	107,83	107,98	107,90

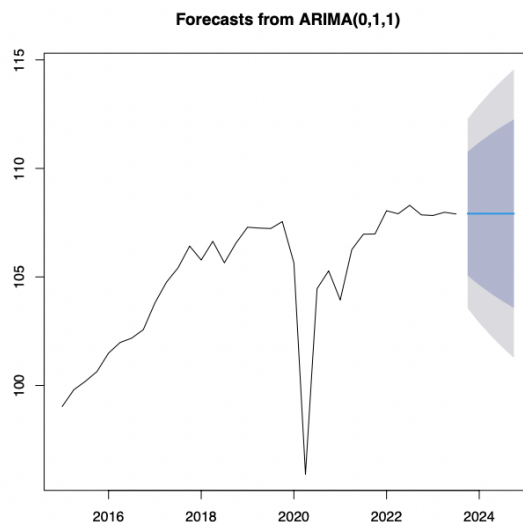
Tabulka 4 - Historické hodnoty HDP



Obrázek 28 - Dekompozice časové řady HDP



Obrázek 29 - ACF u časové řady HDP

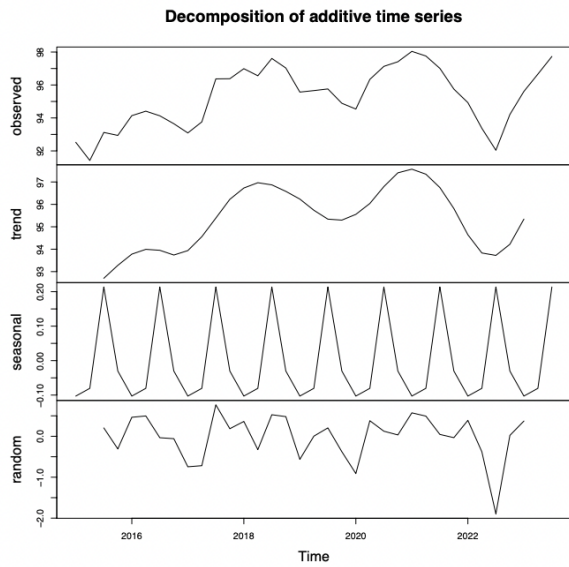


Obrázek 30 - ARIMA model pro HDP

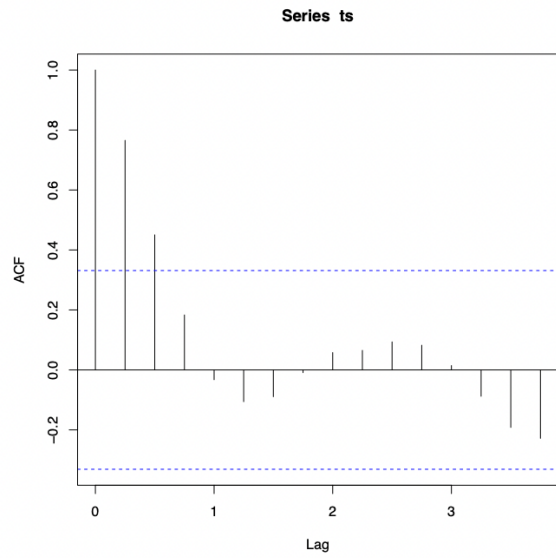
Reálná predikce HDP v Německu světovými organizacemi, jako například Organizace pro hospodářskou spolupráci a rozvoj (OECD) očekává na konci roku 2023 stagnaci ekonomiky, po níž bude následovat mírný pokles a v roce 2024 mírný růst. [36] Evropská komise předpovídá pro rok 2023 podobný vývoj, pro rok 2024 mírný růst. [37]

Reálný efektivní směnný kurz na základě indexu spotřebitelských cen

Reálný efektivní směnný kurz je, jak již z názvu vyplývá, cenově očištěn. V dekompozici (obrázek 31) je zde viditelný trend, který nebyl ani příliš významně ovlivněn pandemií, až v roce 2022 došlo k prudkému propadu, což mohlo být způsobené především válkou na Ukrajině. Po propadu však následoval opět prudký růst. Autokorelační funkce (obrázek 32) naznačuje znázorňuje korelaci dvou čtvrtletí, kde dochází k prudkému poklesu v korelaci. Časová řada dle Dickey-Fullerova testu s p-hodnotou 0,1159 není stacionární. Predikce ARIMA (0,1,0) (obrázek 33) je stejně jako u HDP predikuje stagnaci. Ačkoliv jsou hodnoty před predikcí výrazně rostoucí (tabulka 4), model predikuje na základě historických nižších hodnot ustálení kurzu na kvartální hodnotě 97,73, (tabulka 5) která odpovídá hodnotě třetího kvartálu roku 2023.

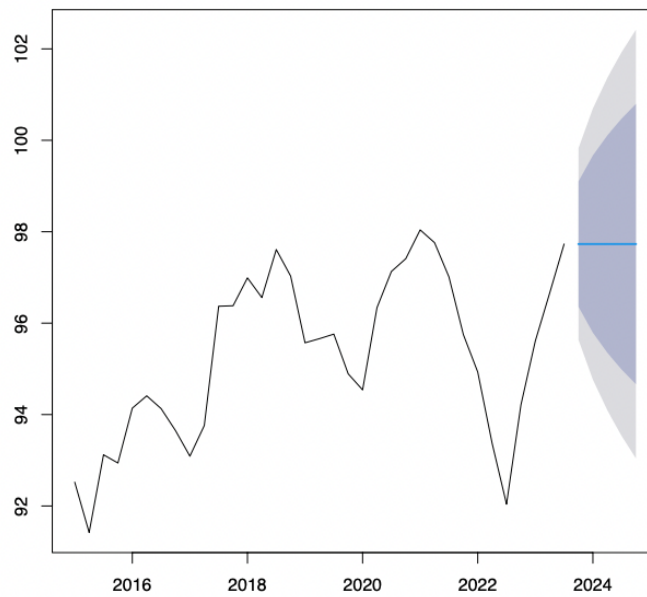


Obrázek 31 - Dekompozice časové řady směnného kurzu.



Obrázek 32 - ACF u časové řady směnného kurzu

Forecasts from ARIMA(0,1,0)



Obrázek 33 - ARIMA model pro směnný kurz

2022-Q1	2022-Q2	2022-Q3	2022-Q4	2023-Q1	2023-Q2	2023-Q3
94,93	93,37	92,04	94,21	95,61	96,66	97,73

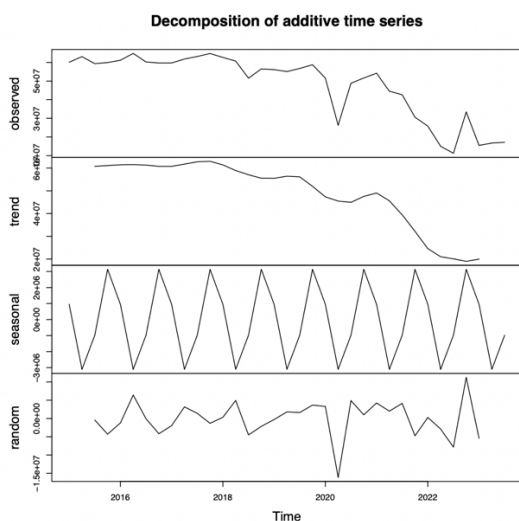
Tabulka 5 - Historické hodnoty směnného kurzu

	Point.Forecast	Lo.80	Hi.80	Lo.95	Hi.95
2023 Q4	97,73	96,3588303617974	99,1011696382026	95,632977000828	99,827022999172
2024 Q1	97,73	95,7908733013397	99,6691266986603	94,7643616339627	100,695638366037
2024 Q2	97,73	95,3550645208373	100,104935479163	94,0978496207936	101,362150379206
2024 Q3	97,73	94,9876607235948	100,472339276405	93,535954001656	101,924045998344
2024 Q4	97,73	94,6639714802952	100,796028519705	93,0409140234709	102,419085976529

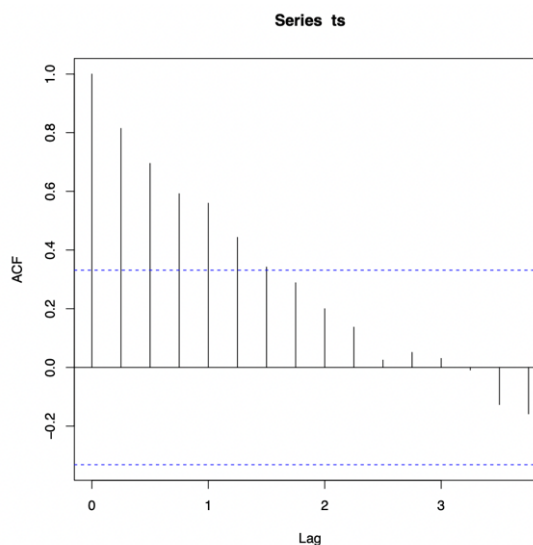
Tabulka 6 - Predikované hodnoty směnného kurzu

Saldo zahraničního obchodu

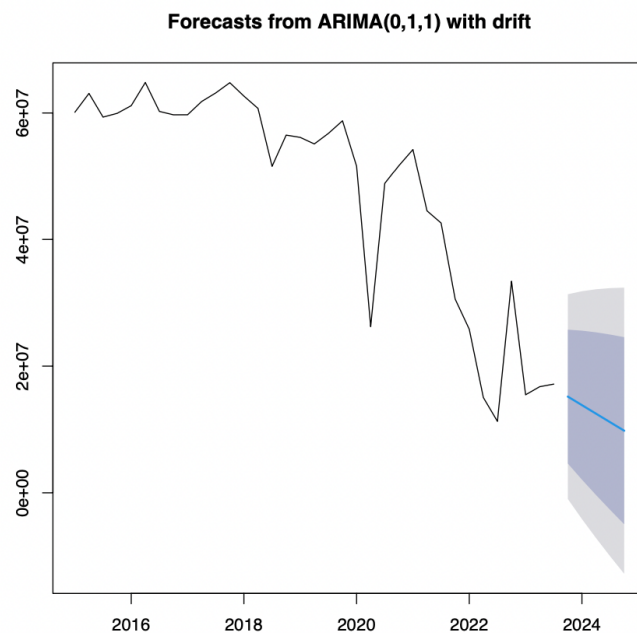
Hodnoty salda jsou zde uvedené a do grafů zanesené v tisících EUR a jsou sezónně a kalendářně očištěné. Nejsou však očištěné o inflaci, tudíž i tento fakt může ovlivňovat kolísání predikce. V grafu dekompozice není viditelný klesající trend (obrázek 34). Autokorelační funkce (obrázek 35) odhaluje silnou korelaci na šest čtvrtletí, i přes to, že je postupně klesající. To značí sezónnost či cykličnost.



Obrázek 34 - Dekompozice časové řady salda



Obrázek 35 - ACF u časové řady salda



Obrázek 36 - ARIMA model pro saldo zahraničního obchodu

Model ARIMA (0,1,1) (obrázek 36) predikuje kontinuální pokles navazující na historická data, bez výkyvů. Pojem "drift" v analýze časových řad označuje stálý dlouhodobý trend v datech. Zahrnutí driftu do modelu ARIMA znamená, že vedle zákonitostí zachycených samotným procesem ARIMA připouštíte i lineární trend v čase. To je často relevantní v případě, že data časové řady vykazují trvalý vzestupný nebo sestupný trend, což je případ této časové řady. Data predikce jsou v tabulce 7.

	Point.Forecast	Lo.80	Hi.80	Lo.95	Hi.95
2023 Q4	15185637,1347089	4619668,92128247	25751605,3481353	-973616,78599499	31344891,0554128
2024 Q1	13840035,1604661	2082363,20432508	25597707,1166071	-4141772,3575795	31821842,6785117
2024 Q2	12494433,1862233	-344804,149744604	25333670,5221913	-7141485,83882613	32130352,2112728
2024 Q3	11148831,2119806	-2687685,00093955	24985347,4249007	-10012294,2377813	32309956,6617424
2024 Q4	9803229,2377378	-4963366,39062181	24569824,8660974	-12780329,9291167	32386788,4045923

Tabulka 7 - Predikované hodnoty salda zahraničního obchodu

3.5 GLM model hodnotící relevantnost proměnných pro predikci

Po základní predikci tras pomocí modelu ARIMA přecházíme k GLM modelu, který zhodnotí, které z proměnných jsou pro predikci převáženého objemu nákladu relevantní a které ne. Generalizovaný lineární model (GLM) představuje flexibilní rozšíření klasické lineární regrese, které se přizpůsobuje různým typům dat a rozdělení chyb. Jeho esencí je využití

spojovací funkce, která propojuje lineární prediktor s proměnnou odezvy a umožňuje provádění různých druhů regresních analýz, jako jsou lineární, logistická či Poissonova regrese. GLM zahrnuje širokou třídu modelů, zahrnující nejenom lineární regresi, ale i analýzu rozptylu, logitové a probitové modely, stejně jako logaritmické a polynomické přístupy. Tyto modely sdílejí společné charakteristiky a metodologii odhadu parametrů, což umožňuje vědeckým pracovníkům po celém světě zkoumat široké spektrum statistických modelů s využitím GLM, přestože každý model je definován zvlášť. [38] [39]

Tato část nezahrnuje predikci, jedná se jen o GLM model, který nám stanovuje pět proměnných, u kterých pozorujeme, zda ovlivňují časové řady šesti zvolených tras za období 2015-2022 (32 čtvrtletí). Pět proměnných zahrnuje časovou proměnnou (časový trend), vliv čtvrtletí, HDP, směnný kurz a saldo zahraničního obchodu. U ARIMA modelu byla k dispozici jen časová proměnná a vliv čtvrtletí, kdy se model snažil stanovit predikci na základě historických hodnot časových řad. Nyní se do GLM modelu přidávají tři zbylé proměnné. GLM model samotný nevytváří predikci, protože zatím nemá údaje pro roky 2023 a 2024 u všech proměnných. Kombinace obou metod přichází až v dalším kroku. Skript z programu R pro GLM model je na obrázku 37.

```
# GLM ####
idx_flights <- c(4:9)
idx_covariates <- c(1, 3, 11, 13, 14)

data <- readRDS("./data/data.rds")
metadata <- read.csv2(file="./data/metadata.csv")

data <- data[c(1:32), ]

tsData <- ts(data, frequency = 4, start=2015)

for (i_f in idx_flights) {
  meta <- find_metadata(colnames(tsData)[i_f])
  ts <- tsData[, c(i_f, idx_covariates)]
  glm1 <- glm(ts[,1] ~ ts[,2:6])
  plot(glm1)
```

Obrázek 37 - Skript z programu R pro GLM model

První trasou je Letiště Frankfurt nad Mohanem – Pekingské mezinárodní letiště. Výsledky jsou vidět na obrázku 38. Každý řádek představuje vliv jedné predikční proměnné. Sloupec Estimate udává odhadovaný účinek na objem carga, Std. Error označuje standardní chybu odhadu, čím nižší chyba, tím je odhad přesnější. Jednotky jsou u standardní chyby stejné jako u odhadované proměnné. Hodnota t je testovací statistika pro test hypotézy a $Pr(>|t|)$ je p-hodnota pro tento test, což je pro nás nejpodstatnějším ukazatelem.

Intercept je konstantní člen modelu. Odhad -95120 značí, že pokud by všechny ostatní proměnné byly nulové, objem nákladu by měl tuto hodnotu. Signifikanční kódy pod proměnnými značí vysvětlivky kódů významnosti, které znázorňují rozsahy p-hodnot a jejich ohodnocení počtem hvězd či tečkou, které se nacházejí za sloupcem p-hodnoty. U interceptu značí, že je tento odhad statisticky významný na hladině 0,05 (p-hodnota = 0,04711). Zde je má jednu hvězdu, tudíž je právě v intervalu (0,01-0,05) a mohu říct, že tato proměnná s 5 % pravděpodobností tuto časovou řadu neovlivňuje. Dále je princip takový, že pokud se mi proměnná „čas“ zvýší o jednotku (jedno čtvrtletí, máme k dispozici 32 čtvrtletí), tak mi v průměru stoupne či klesne objem převážených tun o hodnotu ve sloupci „Estimate“. V tomto případě se mi s následujícím kvartálem sníží počet tun o 421 tun a s 2 % pravděpodobností tuto časovou řadu neovlivňuje. Čtvrtletí jsou hodnoty 1,2,3 a 4. Vliv čtvrtletí je důležitý, protože je spojen se sezónností. Množství nákladu se v různých částech sezóny mění, proto je důležité, které čtvrtletí aktuálně hodnotíme. Pokud se mi zvýší čtvrtletí o jednotku, zvýší se objem převážených tun o 841,7 a s 5 % pravděpodobností tuto časovou řadu neovlivňuje. Pokud se HDP zvýší o jednotku (miliarda EUR), objem přepravených tun se zvýší o 684. HDP tuto trasu s 0,7 % pravděpodobností neovlivňuje. Pokud se zvýší směnný kurz o jednotku, zvýší se objem přepravených tun o 666,1 a s 10 % pravděpodobností tuto časovou řadu neovlivňuje. Poslední proměnnou je saldo zahraničního obchodu. Pokud se zvýší jednotku (tisíc EUR), sníží se objem přepravených tun o 0,0002 a s 0,2 % pravděpodobností tuto časovou řadu neovlivňuje.

U této trasy jsou tedy statisticky významné proměnné čas, HDP a Saldo. Čtvrtletí je téměř významné. Jsou to proměnné, které s největší pravděpodobností nejvíce ovlivňují převážený náklad na trase mezi letišti Frankfurt nad Mohanem a Pekingské mezinárodní letiště.

```
Call:
glm(formula = ts[, 1] ~ ts[, 2:6])
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-9.512e+04	4.564e+04	-2.084	0.04711	*
ts[, 2:6]Time	-4.210e+02	1.659e+02	-2.539	0.01746	*
ts[, 2:6]Quarter	8.417e+02	4.154e+02	2.026	0.05312	.
ts[, 2:6]HDP	6.840e+02	2.353e+02	2.907	0.00737	**
ts[, 2:6]CPI	6.661e+02	3.948e+02	1.687	0.10350	
ts[, 2:6]Saldo	-2.847e-04	8.230e-05	-3.459	0.00188	**

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Obrázek 38 - Model GLM pro FRA-PEK

Druhou trasou je Letiště Frankfurt nad Mohanem – Mezinárodní letiště Kanton Paj-jün.

(obrázek 39) Pomiňme dále u proměnných intercept, protože se nejedná o proměnnou, která by ovlivňovala predikci v dalším kroku spojení modelů ARIMA a GLM. Ze zbylých proměnných je statisticky významný čas, HDP a Saldo. Čas neovlivňuje tuto časovou řadu jen s 4 % pravděpodobností, směnný kurz s 2 % pravděpodobností a HDP s 0,2 % pravděpodobností.

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-1.134e+05	3.429e+04	-3.306	0.00277	**
ts[, 2:6]Time	-2.703e+02	1.246e+02	-2.169	0.03942	*
ts[, 2:6]Quarter	4.986e+02	3.121e+02	1.598	0.12218	
ts[, 2:6]HDP	6.007e+02	1.768e+02	3.399	0.00219	**
ts[, 2:6]CPI	7.368e+02	2.966e+02	2.484	0.01974	*
ts[, 2:6]Saldo	-7.678e-05	6.182e-05	-1.242	0.22539	

Obrázek 39 - Model GLM pro FRA-CAN

Třetí trasou je Letiště Frankfurt nad Mohanem – Mezinárodní letiště Šanghaj Pchu-tung.

Jedinou proměnnou, která se blíží tomu, aby byla statisticky významná na hladině 0,05 je čtvrtletí. Z 6 % pravděpodobně neovlivňuje objem nákladu na této trase. Zbylé proměnné sice nejsou významné, ale i přesto značí, z kolika procent je pravděpodobné, že neovlivňují časovou řadu. (obrázek 40)

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-3.413e+04	1.029e+05	-0.332	0.7429
ts[, 2:6]Time	1.934e+02	3.740e+02	0.517	0.6094
ts[, 2:6]Quarter	1.819e+03	9.367e+02	1.942	0.0631 .
ts[, 2:6]HDP	-3.786e+02	5.305e+02	-0.714	0.4819
ts[, 2:6]CPI	1.175e+03	8.902e+02	1.320	0.1984
ts[, 2:6]Saldo	1.534e-04	1.856e-04	0.827	0.4160

Obrázek 40 - GLM model pro FRA-PVG

Čtvrtou trasou je **Letiště Kolín/Bonn – Mezinárodní letiště Kanton Paj-jün**. Zde není žádná z proměnných statisticky významná na hladině 0,05. Zbylé proměnné sice nejsou významné, ale i přesto značí, z kolika procent pravděpodobně neovlivňují časovou řadu. (obrázek 41)

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-9.867e+02	1.028e+04	-0.096	0.924
ts[, 2:6]Time	4.143e+01	3.737e+01	1.108	0.278
ts[, 2:6]Quarter	7.710e+01	9.361e+01	0.824	0.418
ts[, 2:6]HDP	3.233e+01	5.302e+01	0.610	0.547
ts[, 2:6]CPI	-1.063e+01	8.896e+01	-0.120	0.906
ts[, 2:6]Saldo	7.193e-06	1.854e-05	0.388	0.701

Obrázek 41 - GLM model pro CGN-CAN

Pátá trasa je Letiště Kolín/Bonn – Mezinárodní letiště Šen-čen Pao-an. Proměnnými, které jsou statisticky významné, je čtvrtletí a směnný kurz. Čtvrtletí s 4 % pravděpodobností a směnný kurz s 5 % pravděpodobností nejspíše neovlivňují časovou řadu. (obrázek 42)

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	3.080e+04	1.611e+04	1.912	0.0670 .
ts[, 2:6]Time	9.325e+01	5.854e+01	1.593	0.1233
ts[, 2:6]Quarter	3.185e+02	1.466e+02	2.172	0.0392 *
ts[, 2:6]HDP	-2.154e+01	8.304e+01	-0.259	0.7974
ts[, 2:6]CPI	-2.877e+02	1.393e+02	-2.065	0.0490 *
ts[, 2:6]Saldo	2.085e-05	2.905e-05	0.718	0.4794

Obrázek 42 - GLM model pro CGN-SZX

Poslední z šesti tras je spojení Letiště Kolín/Bonn – Mezinárodní letiště Šanghaj Pchutung. Statisticky významný je čas, směnný kurz a velmi významné je zde HDP, které neovlivňuje objem přepraveného carga jen s 0,07 % pravděpodobností. Čas potom s 3 % a směnný kurz s 2 % pravděpodobností. (obrázek 43)

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-7.038e+04	1.835e+04	-3.836	0.000715	***
ts[, 2:6]Time	-1.540e+02	6.667e+01	-2.310	0.029066	*
ts[, 2:6]Quarter	1.486e+02	1.670e+02	0.890	0.381681	
ts[, 2:6]HDP	3.628e+02	9.458e+01	3.836	0.000716	***
ts[, 2:6]CPI	3.968e+02	1.587e+02	2.501	0.019025	*
ts[, 2:6]Saldo	-2.050e-05	3.308e-05	-0.620	0.540920	

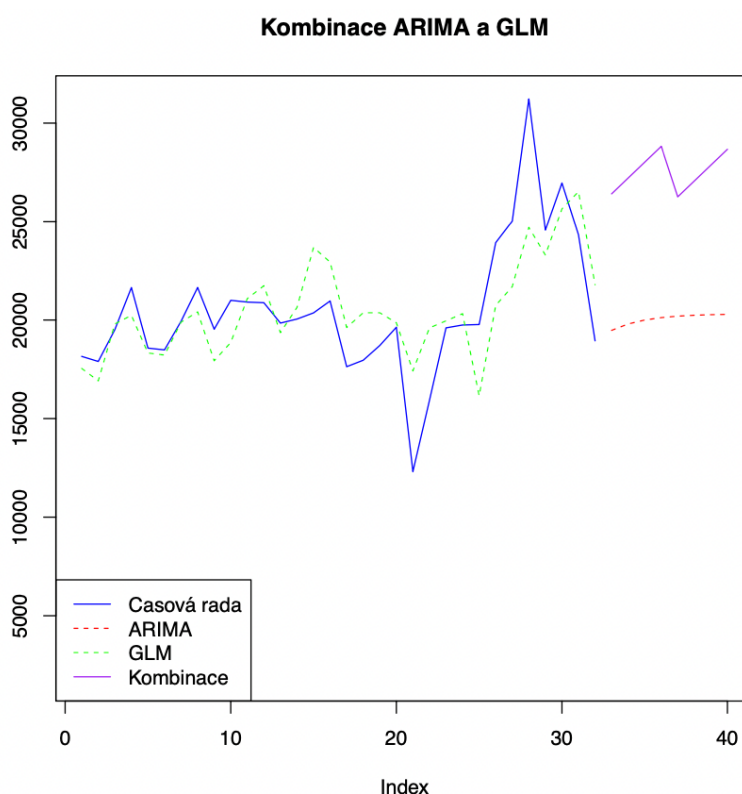
Obrázek 43 - GLM model pro CGN-PVG

3.6 Predikce vývoje tras na základě kombinace modelů ARIMA + GLM

V poslední části predikce se informace z předchozích dvou modelů zkombinují. Skript z programu R pro tuto kombinace je k nalezení v Příloze X kvůli velikosti souboru. **Popišme nyní metodu na grafu první trasy (obrázek 44), Letiště Frankfurt nad Mohanem – Pekingské mezinárodní letiště.** Na ose Y se nachází tuny náklady, na ose X indexy, které znázorňují jednotlivá čtvrtletí. 32 čtvrtletí je zde pro hodnoty, které jsou dostupné, zbylých 8 čtvrtletí je období pro predikci. Dohromady je na ose X 40 kvartálů, končících posledním kvartálem roku 2024. Modrá čára znázorňuje původní časovou řadu, jedná se jen o původní historická data. Červenou přerušovanou čárou je značena predikce modelem ARIMA, která vznikla v první části metodiky. Zelená přerušovaná čára je model GLM, který odhaduje reálně naměřené tuny na základě vstupních vypsání parametrů. Model si v podstatě zkusil vypredikovat historickou přepravu nákladu od roku 2015-2022. Na predikci pro roky 2023 a 2024 již neměl dostatečná data u proměnných. Proto jsou do něj dosazena data socioekonomický proměnných, v našem případě reálná data z Q1-Q3 2023 a vypredikovaná za Q4 2023 a Q1-Q4 2024 pomocí modelu ARIMA, jak bylo uvedeno v první části metodiky. Data byla dosazena do modelu GLM z části 2. Fialová čára je tedy predikcí modelu GLM se socioekonomickými parametry, dosazenými modelem ARIMA. Predikce je pro roky 2023 a 2024. Ačkoliv není znázorněna návaznost, navazuje na modrou čáru.

V seznamu proměnných, které ovlivňují finální předpověď je opět vidět, které proměnné měli jakým způsobem vliv na predikovaný objem. Zde je statisticky významné na hladině 0,05 především čtvrtletí, tedy vliv sezónnosti a následně směnný kurz a saldo zahraničního obchodu. Čtvrtletí s 0,6 % pravděpodobností, že neovlivňuje objem přepraveného carga, směnný kurz s 3 % a saldo s 2 % pravděpodobností. Čas se blíží hladině významnosti. (obrázek 45)

Finální predikce se zde se započítáním vlivu proměnných velmi liší od predikce modelem ARIMA. Stejně jako zde, tak i na následujících trasách jsou predikce kombinací obou modelů velmi ovlivněné sezónností, což je poznat i na jejich tvaru. Obecně se dá říci, že GLM zachycuje sezónnost lépe než ARIMA, která ji v některých případech vůbec nenašla. Predikce se zde pohybuje ve vysokých hodnotách. Po hlubokém poklesu, který nastal na konci roku 2022, kombinace modelů predikuje návrat do vysokých hodnot a pokračování v signifikantních výkyvech. Na obrázku 46 je predikce vypsána v konkrétních hodnotách v tunách pro jednotlivé kvartály od Q1 2023 do Q4 2024, považujeme za podstatnou hodnotu ve sloupci „fit“.



Obrázek 44 - Predikce kombinací modelů ARIMA+GLM pro FRA-PEK

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-7.757e+04	3.980e+04	-1.949	0.06216 .
Time	-2.754e+02	1.446e+02	-1.904	0.06802 .
Quarter	1.067e+03	3.622e+02	2.946	0.00671 **
HDP	1.139e+02	2.052e+02	0.555	0.58366
CPI	8.047e+02	3.442e+02	2.338	0.02736 *
Saldo	-1.762e-04	7.176e-05	-2.456	0.02106 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

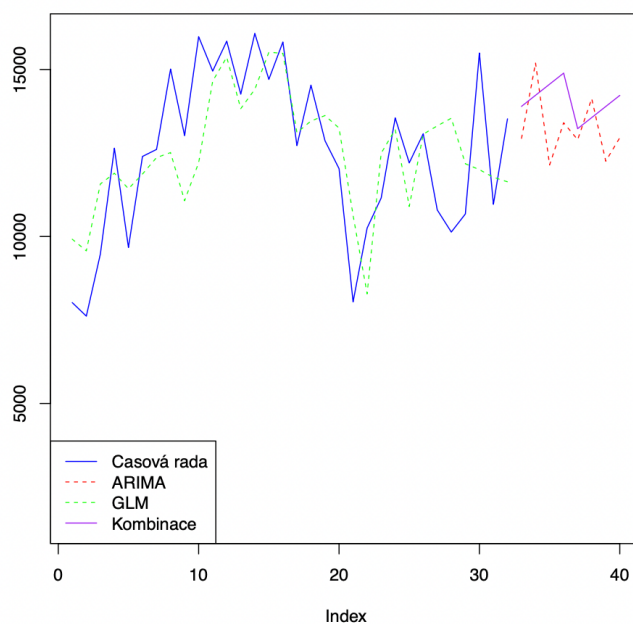
Obrázek 45 - Proměnné pro ARIMA+GLM pro FRA-PEK

	fit
1	26411,868813772
2	27215,6383527151
3	28019,4078916582
4	28823,1774306013
5	26260,0448168843
6	27063,8143558274
7	27867,5838947705
8	28671,3534337136

Obrázek 46 - Hodnoty predikce ARIMA+GLM pro FRA-PEK

Druhou trasou je Letiště Frankfurt nad Mohanem – Mezinárodní letiště Kanton Paj-jün. Predikované hodnoty (obrázek 49) se zde pohybují v podobném intervalu jako u predikce pomocí modelu ARIMA, na rozdíl od předchozí trasy. Dochází zde opět k sezónním výkyvům a dle predikce nedojde příští dva roky k velkému propadu, který můžeme pozorovat v historických datech. (graf, obrázek 47) Pokud jde o proměnné, žádná z nich není statisticky významná. I tato skutečnost naznačuje, proč jsou hodnoty predikce na podobné úrovni jako hodnoty predikce modelem ARIMA. O HDP jako nejdůležitější socioekonomické proměnné se dá říci, že s pravděpodobností 87 % neovlivňuje objem převáženého nákladu na této trase. (obrázek 48)

Kombinace ARIMA a GLM



Obrázek 47 - Predikce kombinací modelů ARIMA+GLM pro FRA-CAN

Coefficients:

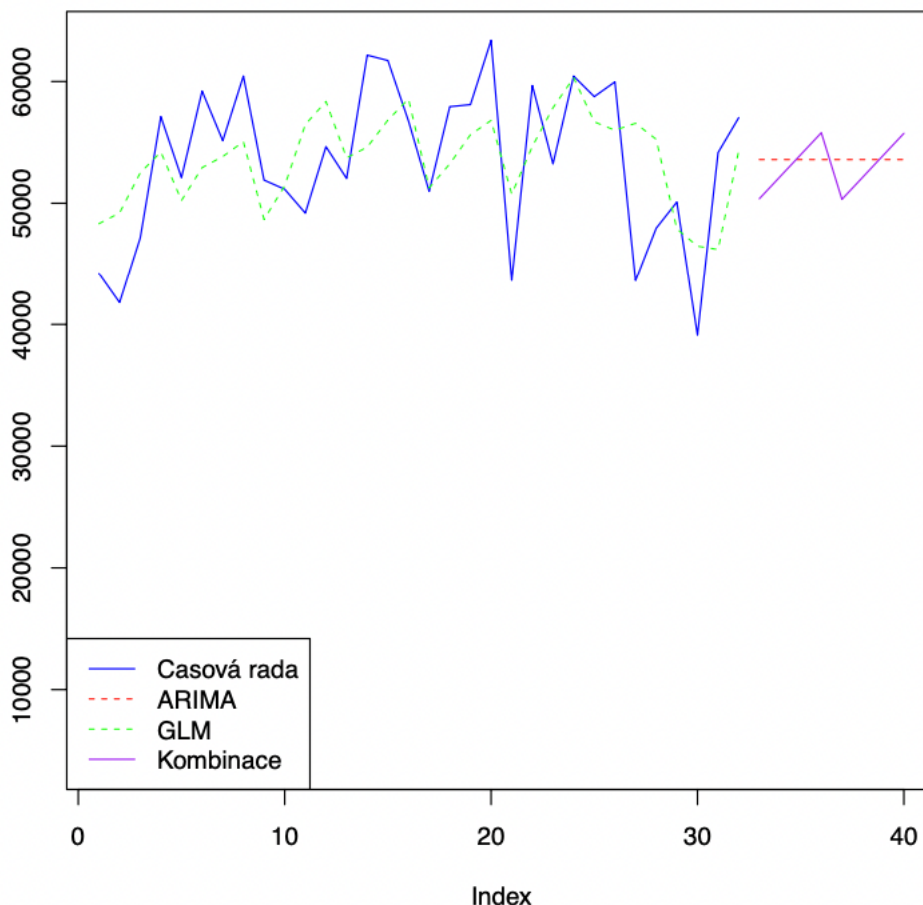
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-2.638e+03	3.252e+04	-0.081	0.936
Time	-4.589e+00	1.182e+02	-0.039	0.969
Quarter	2.744e+02	2.960e+02	0.927	0.362
HDP	-2.824e+01	1.677e+02	-0.168	0.868
CPI	5.395e+01	2.813e+02	0.192	0.849
Saldo	1.058e-06	5.865e-05	0.018	0.986

Obrázek 48 - Proměnné pro ARIMA+GLM pro FRA-CAN

	fit
1	13895,9688075042
2	14227,6658073267
3	14559,3628071491
4	14891,0598069715
5	13228,1945169888
6	13559,8915168113
7	13891,5885166337
8	14223,2855164561

Obrázek 49 - Hodnoty predikce ARIMA+GLM pro FRA-CAN

Třetí trasou je Letiště Frankfurt nad Mohanem – Mezinárodní letiště Šanghaj Pchu-tung. (obrázek 50) Žádná z proměnných není statisticky významná (obrázek 51), jen čtvrtletní vliv se tomu blíží s pravděpodobností 6 %, že neovlivňuje přepravní objem. Čtvrtletí značí sezónnost, o které jsme již hovořili, že je v predikci patrná. Predikce zde odpovídá predikci modelu ARIMA, která se drží zhruba v oblasti středu zprůměrovaných historických hodnot. Predikované hodnoty jsou na obrázku 52.



Obrázek 50 - Predikce kombinací modelů ARIMA+GLM pro FRA-PVG

Coefficients:

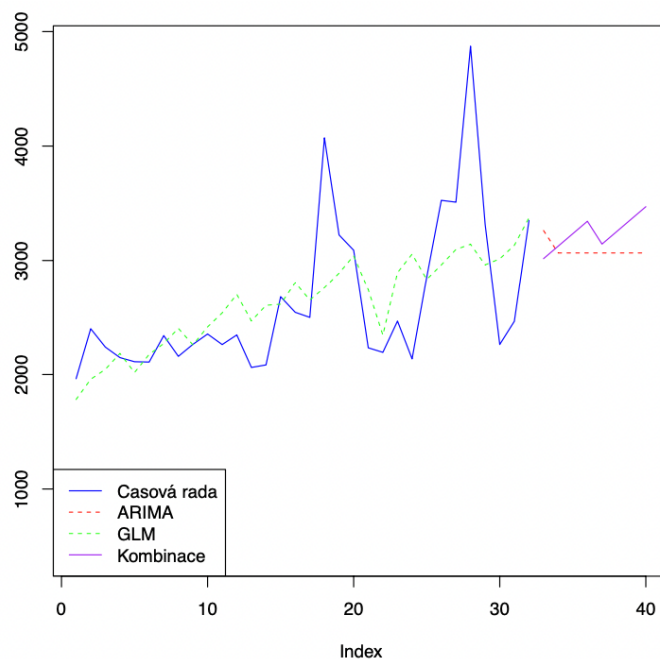
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-8.771e+04	1.029e+05	-0.852	0.4018
Time	1.934e+02	3.740e+02	0.517	0.6094
Quarter	1.819e+03	9.367e+02	1.942	0.0631
HDP	-3.786e+02	5.305e+02	-0.714	0.4819
CPI	1.175e+03	8.902e+02	1.320	0.1984
Saldo	1.534e-04	1.856e-04	0.827	0.4160

Obrázek 51 - Proměnné pro ARIMA+GLM pro FRA-PVG

fit	
1	50374,4336113234
2	52180,524038136
3	53986,6144649485
4	55792,7048917611
5	50322,4715246369
6	52128,5619514495
7	53934,652378262
8	55740,7428050746

Obrázek 52 - Hodnoty predikce ARIMA+GLM pro FRA-PVG

Čtvrtou trasou je **Letiště Kolín/Bonn – Mezinárodní letiště Kanton Paj-jün**. V tomto případě opět není žádná z proměnných statisticky významná a ani se tomu žádná z nich neblíží. Zbylé proměnné sice nejsou významné, ale i přesto značí, z kolika procent pravděpodobně neovlivňují časovou řadu. GLM model se zde příliš neshoduje s časovou řadou reálných dat. Predikce spojená se socioekonomickými proměnnými však opět přibližně odpovídá umístění v grafu vzhledem k predikci modelu ARIMA. Predikce předpovídá, že budou hodnoty pokračovat v rostoucím trendu, což je viditelné i z číselných hodnot predikce.



Obrázek 53 - Predikce kombinací modelů ARIMA+GLM pro CGN-CAN

Coefficients:

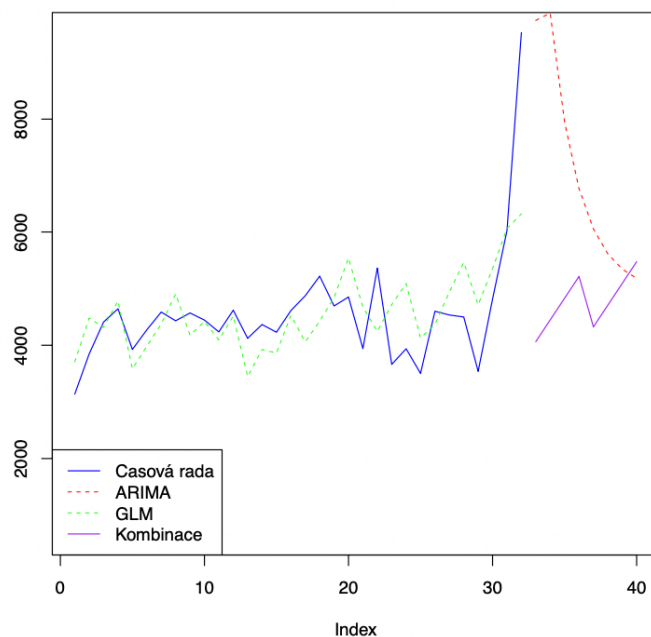
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1.693e+03	1.149e+04	0.147	0.884
Time	2.297e+01	4.174e+01	0.550	0.587
Quarter	1.107e+02	1.046e+02	1.059	0.299
HDP	-4.282e+01	5.921e+01	-0.723	0.476
CPI	1.597e+01	9.935e+01	0.161	0.874
Saldo	1.290e-05	2.071e-05	0.623	0.539

Obrázek 54 - Proměnné pro ARIMA+GLM pro CGN-CAN

	fit
1	3016,58102321062
2	3125,43268738664
3	3234,28435156267
4	3343,13601573869
5	3143,56820778648
6	3252,41987196251
7	3361,27153613853
8	3470,12320031456

Obrázek 55 - Hodnoty predikce ARIMA+GLM pro CGN-CAN

Pátá trasa je Letiště Kolín/Bonn – Mezinárodní letiště Šen-čen Pao-an. Žádná z proměnných není statisticky významná na hladině významnosti 0,05, jen čtvrtletní vliv se tomu blíží se pravděpodobností 5 %, že neovlivňuje přepravní objem. Zde je signifikantní rozdíl mezi predikcemi. Zatímco předpověď pomocí modelu ARIMA značí pozvolnější pokles, který by měl trvat po celé dva roky predikce, kombinace modelů předpokládá okamžitý pokles na hodnoty, před tím než došlo k rychlému nárůstu na konci roku 2022. Následně predikce pokračuje s mírným stoupáním s výkyvy, které byly pozorovatelné už před zmiňovaným nárůstem.



Obrázek 56 - Predikce kombinací modelů ARIMA+GLM pro CGN-SZX

Coefficients:

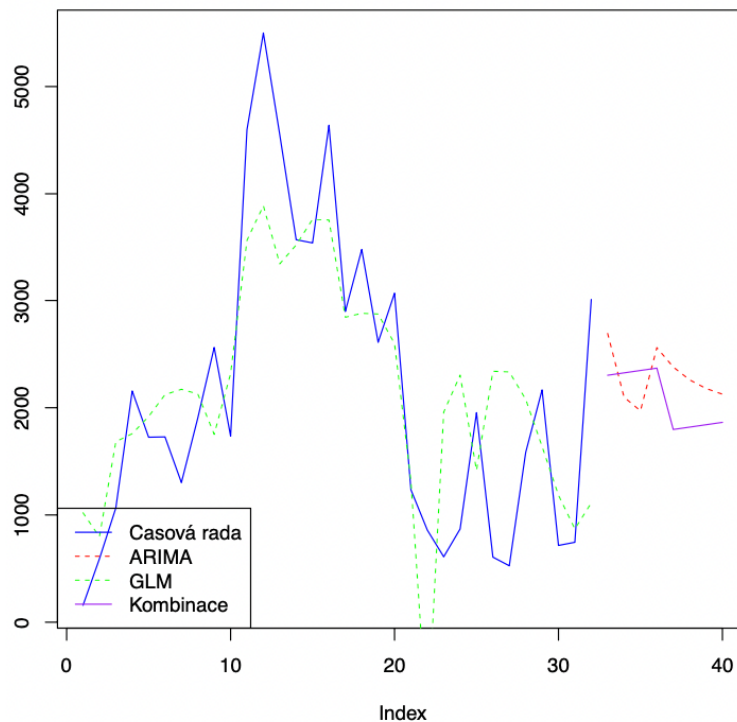
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	2.161e+04	1.462e+04	1.478	0.1514
Time	5.294e+01	5.313e+01	0.996	0.3283
Quarter	2.688e+02	1.331e+02	2.020	0.0538
HDP	-4.773e+01	7.537e+01	-0.633	0.5321
CPI	-1.907e+02	1.265e+02	-1.508	0.1435
Saldo	1.090e-06	2.636e-05	0.041	0.9673

Obrázek 57 - Proměnné pro ARIMA+GLM pro CGN-SZX

	fit
1	4063,57098083315
2	4447,24393392571
3	4830,91688701826
4	5214,58984011083
5	4324,36152155117
6	4708,03447464373
7	5091,70742773629
8	5475,38038082886

Obrázek 58 - Hodnoty predikce ARIMA+GLM pro CGN-SZX

Poslední z šesti tras je spojení Letiště Kolín/Bonn – Mezinárodní letiště Šanghaj Pchutung. (obrázek 59) Jedinou proměnnou, která je zde statisticky významná na hladině významnosti 0,05, je vliv čtvrtletí. (obrázek 60). Jedná se o 4 % pravděpodobnost, že proměnná neovlivní objem převáženého nákladu. Predikce je zde opět v podobném rozpětí nákladu jako je predikce pomocí modelu ARIMA. Model předpovídá zastavení růstu, ke kterému docházelo na konci roku 2022 a následný pokles přepravených tun, který se bude blížit hodnotám pozorovaných v letech 2021 a 2022. Číselné hodnoty predikce jsou na obrázku 61.



Obrázek 59 - Predikce kombinací modelů ARIMA+GLM pro CGN-PVG

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-9.512e+03	1.557e+04	-0.611	0.5467
Time	-1.631e+00	5.660e+01	-0.029	0.9772
Quarter	3.044e+02	1.418e+02	2.147	0.0413 *
HDP	6.840e+01	8.029e+01	0.852	0.4020
CPI	7.515e+00	1.347e+02	0.056	0.9559
Saldo	1.901e-05	2.808e-05	0.677	0.5045

Obrázek 60 - Proměnné pro ARIMA+GLM pro CGN-PVG

	fit
1	2303,00939794708
2	2325,16136004277
3	2347,31332213843
4	2369,46528423412
5	1797,22039214193
6	1819,3723542376
7	1841,52431633328
8	1863,67627842895

Obrázek 61 - Hodnoty predikce ARIMA+GLM pro CGN-PVG

Obecně se dá shrnout, že největší vliv má na predikci tras kvartální proměnná. Jak již bylo zmíněno, GLM model zachycuje v těchto případech sezónnost lépe než ARIMA, tudíž jsou pak viditelné projevy sezónnosti a pravděpodobně i cykličnosti v predikci. Socioekonomické proměnné nejsou většinu času v modelu statisticky významné. Procentuálně je však pravděpodobné, že modely nějakou měrou ovlivňují, ve většině případů to však není na hladině 0,05.

3.7 Kvantifikace nárůstu a poklesu vytížení provozovaných typů letadel

Poslední částí práce je kvantifikace nárůstu vytížení provozovaných typů letadel a počet dodatečných letů na vybraných trasách, případně pokles vytížení a snížení počtu letů. Postup je takový, že je osm kvartálních hodnot predikovaných kombinací modelů GLM + ARIMA zprůměrovaných do jedné hodnoty a ta je porovnána s jednou zprůměrovanou hodnotou vzniklou z osmi kvartálů historických údajů trasy. Z toho je získána informace, o kolik přepravovaný objem na trase poroste či klesne. K zprůměrování byl použit vzorec Averagea v Excelu. Protože predikujeme osm čtvrtletí, kdy se jedná o krátkodobou predikci, z historických dat bylo vzato také osm hodnot s ohledem na předpokládané krátkodobé změny v predikovaném počtu letadel u leteckých společností. Tyto průměry byly použity k výpočtu rozptylu pomocí vzorce pro rozptyl. Vzorec rozptylu vypočítává, jaký je průměrný kvartální rozdíl mezi predikovanými hodnotami a historickými hodnotami časové řady. Rozptyl lze uvést číselně nebo v procentech. [40] Existují dva způsoby výpočtu rozptylu:

$$\text{Variance [\%]} = \text{Forecast AVG} - \text{Data AVG Data AVG}$$

$$\text{Variance [tonnes]} = \text{Forecast AVG} - \text{Data AVG}$$

Data pro první trasu **Letiště Frankfurt nad Mohanem – Pekingské mezinárodní letiště** jsou následující (tabulka 8):

	Data	Predikce
Průměr [tuny]	24 342	27 541
Rozptyl [%]		13,00 %
Rozptyl [tuny]		3199

Tabulka 8 - Rozptyl pro trasu FRA-PEK

Pro další postup byla potřeba zjistit počet letů a typy letadel, které jsou na trase provozovány. Typy letadel byly zjištěny z historických dat na Flightradar24, což je globální nástroj pro sledování letů a veškerých informací o nich. Flightradar24 kombinuje data z několika zdrojů dat včetně ADS-B, MLAT, satelitních a radarových dat. Zároveň je možné vyhledávat data o letech v rozsahu dle zaplaceného předplatného. [41] Data o počtu letů byla vzata z Eurostatu, konkrétně stejného datasetu, který byl používán pro predikci převáženého nákladu (Freight and mail air transport between the main airports of Germany and their main partner airports (routes data), avia_gor), jen zde byl zaměněn náklad za počet letů a bylo změněno období. [42] Zároveň bylo nutné znát maximální payload každého letadla, které určuje maximální hmotnost nákladu, který lze naložit na palubu. V praxi se letadla téměř neprovozují s maximálním payloadem. Obvykle se množství přepravovaného nákladu proto vyjadřuje pomocí koeficientu zatížení. Load factor je poměr průměrného nákladu k celkové nákladní kapacitě letadla. IATA odhadla průměrné užitečné zatížení v celém odvětví v prosinci 2022 na 47,2 %, tudíž bylo toto číslo následně využito i pro výpočty. [43] Počet letů (tabulka 9) na první trase mezi lety 2015 a 2022 činil průměrně 579 letů za čtvrtletí. Na této trase operují dvě letecké společnosti, Lufthansa a Air China a nejběžnější nákladní letadla na této trase jsou Boeing 747-400F a Boeing 777F. Jejich maximální payload je pro B747-400F 113 tun a pro B777F 102 tun. [44] Na základě těchto údajů a výše popsané metody je počet dalších letadel, která pokryjí rostoucí objem nákladu do posledního čtvrtletí roku 2024, 60 v případě Boeingu 747-400F a 67 v případě Boeingu 777F.(tabulka 10)

FRA-PEK počet letů (čtvrtletí)	506	577	547	773	612	668	568	380
---------------------------------------	-----	-----	-----	-----	-----	-----	-----	-----

Tabulka 9 - Počet letů na trase FRA-PEK

A/C Typ	Max. Payload [tuny]	Upravený Payload [tuny]	Rozptyl [tuny]	Výsledné navýšení letů
B747-400 F	113	53,30	3199	60
B777F	102	48,10	3199	67

Tabulka 10 - Výsledné navýšení/snížení letů pro trasu FRA-PEK

Druhou trasou je Letiště Frankfurt nad Mohanem – Mezinárodní letiště Kanton Paj-jün.

Na této trase je nižší převážené množství nákladu než na první, tudíž i průměry historických dat a predikce se pohybují v nižších hodnotách. Rozptyl obou hodnot je 1953 tun. (tabulka 11) Na trase operuje jen společnost China Southern Airlines s Boeingem 777F. Upravený payload je 48,1 tun, po vydělení rozptylu payloadem nám vychází výsledné doporučení navýšení letů/letadel na této trase o 41 do konce roku 2024. (tabulka 12)

	Data	Predikce
Průměr [tuny]	12 107	14 060
Rozptyl [%]		16,00 %
Rozptyl [tuny]		1953

Tabulka 11 - Rozptyl pro trasu FRA-CAN

A/C Typ	Max. Payload [tuny]	Upravený Payload [tuny]	Rozptyl	Výsledné navýšení letů
B777F	102	48,10	1953	41

Tabulka 12 - Výsledné navýšení/snížení letů pro trasu FRA-CAN

Třetí trasou je Letiště Frankfurt nad Mohanem – Mezinárodní letiště Šanghaj Pchu-tung.

Rozptyl obou hodnot je 1719 tun. (tabulka 13) Na trase operují čtyři aerolinky: China Southern Airlines s B777F, China Eastern Airlines také s A330F, Air China s B777F a B747-400F a Lufthansa s Boeingem 777F. Upravený payload je viditelný v tabulce X a po vydělení rozptylu payloadem nám vychází výsledné doporučení navýšení letů/letadel na této trase u B777F o 41 letadel, u B747-400F by byl výsledný počet 32 letadel a u A330F 52 letadel do konce roku 2024. (tabulka 14)

	Data	Predikce
Průměr [tuny]	51 339	53 058
Rozptyl [%]		3 %
Rozptyl [tuny]		1719

Tabulka 13 - Rozptyl pro trasu FRA-PVG

A/C Typ	Max. Payload [tuny]	Upravený Payload [tuny]	Rozptyl [tuny]	Výsledné navýšení letů
B777F	102	48,10	1719	36
B747-400F	113	53,30	1719	32
A330F	70	33,00	1719	52

Tabulka 14 - Výsledné navýšení/snížení letů pro trasu FRA-PVG

Čtvrtou trasou je Letiště Kolín/Bonn – Mezinárodní letiště Kanton Paj-jün.

Na této trase je se opět pohybujeme v násobně nižších hodnotách než na předchozí trase, co se týká převáženého množství nákladu. Rozptyl obou hodnot je -26 tun. (tabulka 15) Na trase operuje jen společnost FedEx s Boeingem 777F. Upravený payload je 48,1 tun, po vydělení rozptylu payloadem nám vychází výsledné doporučení na snížení letů/letadel na této trase o 1 do konce roku 2024. (tabulka 16)

	Data	Predikce
Průměr [tuny]	3 269	3 243
Rozptyl [%]		-1 %
Rozptyl [tuny]		-26

Tabulka 15 - Rozptyl pro trasu CGN-CAN

A/C Typ	Max. Payload [tuny]	Upravený Payload [tuny]	Rozptyl [tuny]	Výsledné snížení letů
B777F	102	48,10	-26	1

Tabulka 16 - Výsledné navýšení/snížení letů pro trasu CGN-CAN

Pro pátou a šestou trasu kalkulace nebyla dokončena, protože nejsou k dispozici data. U páté se nepodařilo zjistit typy letadel provozovaných na lince přes flightrada24 a jinde nejsou tyto údaje dohledatelné a u šesté trasy nejsou k dispozici počty tras přes Eurostat databázi. Zároveň se u páté trasy dá říct, že nemá smysl kalkulaci tvořit, neboť je zde signifikantní nárůst v posledním čtvrtletí, zatímco predikce počítá s okamžitým poklesem, což je velmi nepravděpodobná varianta. Hodnoty tedy nebudou odpovídající.

Diskuze

Tato práce byla zadávána v první polovině loňského roku s předpokladem, že bude docházet k návratu k předpandemickým hodnotám a růstu nad ně. V tu chvíli by byla kalkulace dodatečných letadel smysluplné. Vzhledem k neočekávané stagnaci není v podstatě poslední část této práce příliš žádaná. Je viditelný vliv ekonomické situace v Německu i ve světě na leteckou nákladní dopravu. Vzhledem ke stagnaci ekonomiky a jen náznakům budoucího růstu byl zpomalen i růst leteckého nákladu. Dá se říci, že cargo mezi Německem a Čínou dle výsledků kopíruje vývoj světového leteckého carga a světové letecké cargo je zpomalováno stagnací světové ekonomiky oproti původně predikovaným hodnotám pro letošní a nadcházející rok. Toto odpovídá i ne příliš zřetelnému vlivu socioekonomických proměnných na trasy v této práci, zároveň se zde ale jemný vliv vyskytuje. Dá se také říci, že predikce všech tras by zprůměrovaně stagnovala či mírně rostla, což také odpovídá současnému stavu. Tento jev je také viditelný v predikci pomocí modelu ARIMA u všech tras mezi Německem a Čínou. Predikce většiny tras jen pokračuje v trendu, který je započatý. U tras, které zásadním způsobem rostou či klesají je nepravděpodobné, že to bude odpovídat reálnému vývoji, ale na to se může zaměřit až navazující práce. Navazující práce by se také mohla zaměřit na přesnější metodu kvantifikace zatížení. S nedostatečnými daty a získáváním dostupných dat přes server flightradar24 je snadné udělat chybu, bez ohledu na to, že se jedná i o velmi časově náročnou variantu. Spojením těchto faktů a také metody kalkulace mohou být odhady vzdálené od reálných čísel.

Závěr

V diplomové práci jsem se zaměřila na analýzu a predikci cargo provozu na vybraných trasách. Hlavním cílem bylo pomocí vhodné metody provést předpověď poptávky po přepravě zboží letecky a kvantifikovat nárůst vytížení provozovaných typů letadel a počet dodatečných letů na vybraných trasách, případně pokles vytížení a snížení počtu letů.

Po rešerši metod predikce byl pro predikci zvolen statistický model ARIMA. Vzhledem k výsledkům rešerše, kdy bylo několikrát zmíněno, že hybridní modely dosahují větší přesnosti predikce, byl k modelu ARIMA přidán model GLM. Po přidání socioekonomických proměnných vznikl model ARIMA-GLM, který predikoval vývoj přepravených tun na trasách mezi Německem a Čínou.

Na závěr práce měl být kvantifikován nárůst vytížení provozovaných typů letadel a počet dodatečných letů na vybraných trasách, případně pokles vytížení a snížení počtu letů. Vzhledem ke stagnaci letecké dopravy a jen mírného náznaku růstu v budoucí době není žádoucí kvantifikaci nyní provádět. Mírný nárůst vytížení by mělo být zvládnutelné navýšením load factoru jednotlivými aeroliniemi tak, aby nebylo potřebné navyšovat počet letadel/letů.

Celkové predikce nejsou zatím ověřitelné, neboť je predikováno budoucí období let 2023 a 2024, což je i považováno za omezení této práce. I z toho důvodu byla do práce přidána kombinace s druhým statistickým modelem a socioekonomickými proměnnými, aby byly výsledky co nejpřesnější pro případné budoucí porovnání v navazujících pracích. Pro budoucí výzkum bych doporučila predikci pomocí modelů strojového učení, neboť tyto modely dle teorie zvládají přesnější predikce nelineárních modelů a měly by být součástí budoucí predikce nejen v letecké dopravě.

Seznam použité literatury

- [1] DOGANIS, Rigas, 2019. 9 Forecasting demand. In: *Flying off Course: Airline Economics and Marketing*. Fifth edition. Milton, UNITED KINGDOM: Taylor & Francis Group, s. 188–212 [vid. 2023-08-21]. ISBN 978-1-315-40297-0.
- [2] VASIGH, Bijan, 2008. 9 Forecasting in the air transport industry. In: *Introduction to Air Transport Economics: From Theory to Applications*. Milton, UNITED KINGDOM: Taylor & Francis Group, s. 300–352 [vid. 2023-08-21]. ISBN 978-1-351-15535-9
- [3] GERARDO MUROS ANGUIA, Juan a Oscar DÍAZ OLARIAGA, 2023. Air cargo transport demand forecasting using ConvLSTM2D, an artificial neural network architecture approach. *Case Studies on Transport Policy* [online]. **12**, 101009. ISSN 2213-624X. Dostupné z: doi:10.1016/j.cstp.2023.101009
- [4] JUREČKA, Václav a KOLEKTIV. *Mikroekonomie* [online]. 3. B.m.: Grada, s. 55-70 [vid. 2023-11-04]. ISBN 978-80-271-0146-7. Dostupné z: [https://www.grada.cz/mikroekonomie-\(2\)-10151/?gad_source=1&gclid=Cj0KCQjwtJKqBhCaARIsAN_yS_k-1jCAaIW_zcKnsI2xsjpK8m8iiRVVpdQbf0fhje5K5IOx4QMogTQaAl1vEALw_wcB](https://www.grada.cz/mikroekonomie-(2)-10151/?gad_source=1&gclid=Cj0KCQjwtJKqBhCaARIsAN_yS_k-1jCAaIW_zcKnsI2xsjpK8m8iiRVVpdQbf0fhje5K5IOx4QMogTQaAl1vEALw_wcB)
- [5] CHEN, Bin, Jiacheng LIU, Zhouying RUAN, Ming YUE, Hansen LONG a Weiping YAO, 2022. Freight traffic of civil aviation volume forecast based on hybrid ARIMA-LR model. In: [online]. s. 69. Dostupné z: doi:10.1117/12.2657975
- [6] TOTAMANE, Raghavendra, Amit DASGUPTA a Shrisha RAO, 2014. Air Cargo Demand Modeling and Prediction. *IEEE Systems Journal* [online]. **8**, 52. Dostupné z: doi:10.1109/JSYST.2012.2218511
- [7] AKINYEMI, Yingigba Chioma, 2023. Air cargo demand in Africa: Application of cointegration and error correction modelling techniques. *Journal of Air Transport Management* [online]. **109**, 102399. ISSN 0969-6997. Dostupné z: doi:10.1016/j.jairtraman.2023.102399
- [8] KAŠČEEV, Artem, 2021. Konverze osobních letadel na nákladní v souvislosti s přepravní kapacitou [online]. [vid. 2023-11-27]. Dostupné z: <https://dspace.cvut.cz/handle/10467/97451>

- [9] LUCAS, Patrick, 2018. Airport markets and seasonal variations. *ACI World Insights* [online]. [vid. 2023-11-30]. Dostupné z: <https://blog.aci.aero/airport-markets-and-seasonal-variations/>
- [10] ŘEZANKOVÁ, Hana, Luboš MAREK a Michal VRABEC. Iostat – INTERAKTIVNÍ UČEBNICE STATISTIKY. *Časovky* [online] [vid. 2023-11-30]. Dostupné z: <https://iostat.vse.cz/>
- [11] LOAIZA, Miguel, Roberto SOLANO, Ricardo SIMANCAS a Víctor HIGUERA OJITO, 2017. Modeling Demand for Air Cargo in the Colombian Context. In: [online]. Dostupné z: doi:10.2991/amsce-17.2017.31
- [12] AAMER, Ammar, Luh Putu EKA YANI a I Made ALAN PRIYATNA, 2020. Data Analytics in the Supply Chain Management: Review of Machine Learning Applications in Demand Forecasting. *Operations and Supply Chain Management: An International Journal* [online]. 1–13. ISSN 2579-9363. Dostupné z: doi:10.31387/oscm0440281
- [13] LIU, Jiaming, Lina DING, Xiaoyu GUAN, Jiao GUI a Jianbin XU, 2020. Comparative analysis of forecasting for air cargo volume: Statistical techniques vs. machine learning. *Journal of Data, Information and Management* [online]. 2(4), 243–255. ISSN 2524-6364. Dostupné z: doi:10.1007/s42488-020-00031-1
- [14] PANDE, Pranjal a Jake HARDIMAN, 2021. What Is An Airline's Load Factor And Why Does It Matter? *Simple Flying* [online] [vid. 2023-11-30]. Dostupné z: <https://simpleflying.com/load-factor/>
- [15] SULISTYOWATI, Ratna, SUHARTONO, Heri KUSWANTO, SETIAWAN a Erni Tri ASTUTI, 2018. Hybrid forecasting model to predict air passenger and cargo in Indonesia. In: *2018 International Conference on Information and Communications Technology (ICOIACT)* [online]. Yogyakarta: IEEE [vid. 2023-08-21]. Dostupné z: doi:10.1109/ICOIACT.2018.8350816
- [16] TRILOBYTE. *Neuronová síť | QC-Expert* [online] [vid. 2023-11-30]. Dostupné z: <https://www.trilobyte.cz/QC-Expert/Neuronova-sit.html>

- [17] NOVOTNÝ, Dominik, 2019. Predikce nemocnosti zaměstnanců Škoda Auto pomocí strojového učení [online]. [vid. 2023-11-30]. Dostupné z: <https://dk.upce.cz/handle/10195/73031>
- [18] KAŠČEEV, Artem, Eva ENDRIZALOVÁ a Peter VITTEK, 2022. Air Cargo Demand Prediction on Selected Routes by Holt-Winter Algorithm. In: *2022 New Trends in Civil Aviation (NTCA)* [online]. s. 87–91 [vid. 2023-11-30]. ISSN 2694-7854. Dostupné z: [doi:10.23919/NTCA55899.2022.9934392](https://doi.org/10.23919/NTCA55899.2022.9934392)
- [19] LIU, Jiaming, Lina DING, Xiaoyu GUAN, Jiao GUI a Jianbin XU, 2020. Comparative analysis of forecasting for air cargo volume: Statistical techniques vs. machine learning. *Journal of Data, Information and Management* [online]. **2**(4), 243–255. ISSN 2524-6364. Dostupné z: [doi:10.1007/s42488-020-00031-1](https://doi.org/10.1007/s42488-020-00031-1)
- [20] CHEN, Bin, Jiacheng LIU, Zhouying RUAN, Ming YUE, Hansen LONG a Weiping YAO, 2022. Freight traffic of civil aviation volume forecast based on hybrid ARIMA-LR model. In: [online]. s. 69. Dostupné z: [doi:10.1117/12.2657975](https://doi.org/10.1117/12.2657975)
- [21] SURYAN, Viktor, Anggi SARI, Direstu AMALIA a M. HABILLAH, 2020. Econometric Forecasting models for Air Freight in Indonesia (And How Will It be Affected by COVID-19?). *Journal of Airport Engineering Technology (JAET)* [online]. **1**, 30–33. Dostupné z: [doi:10.52989/jaet.v1i1.5](https://doi.org/10.52989/jaet.v1i1.5)
- [22] JIANG, Hong, Liling REN a R. HANSMAN, 2003. Market and Infrastructure Analysis of Future Air Cargo Demand in China [online]. Dostupné z: [doi:10.2514/6.2003-6770](https://doi.org/10.2514/6.2003-6770)
- [23] STATISTA RESEARCH DEPARTMENT. Total cargo volume handled by major European airports 2022. *Statista* [online] [vid. 2023-11-30]. Dostupné z: <https://www.statista.com/statistics/1402888/total-cargo-volume-by-european-airport/>
- [24] Gross domestic product 2022. In: *World bank* [online]. 2023 [cit. 2023-11-30]. Dostupné z: https://databankfiles.worldbank.org/public/ddpext_download/GDP.pdf
- [25] BUSINESS MARKET INSIGHTS. *Asia Pacific Air Cargo Market to 2028 - By Size, Share, Growth by End User and Forecast* | *Business Market Insights* [online] [vid. 2023-11-30]. Dostupné z: <https://www.businessmarketinsights.com/reports/asia-pacific-air-cargo-market>

[26] EUROSTAT. Freight and mail air transport between the main airports of Germany and their main partner airports (routes data). *Statistics | Eurostat* [online] [vid. 2023-11-30].

Dostupné

z: https://ec.europa.eu/eurostat/databrowser/view/avia_gor_de__custom_8425420/default/table?lang=en

[27] EUROSTAT, [b.r.]. International extra-EU freight and mail air transport by reporting country and partner world regions and countries. *Statistics | Eurostat* [online] [vid. 2023-11-30]. Dostupné

Dostupné

z: https://ec.europa.eu/eurostat/databrowser/view/avia_goexcc__custom_8710039/default/table?lang=en

[28] Federal statistical office. Gross domestic product in the 3rd quarter of 2023 down 0.1% on the previous quarter. *Federal Statistical Office* [online] [vid. 2023-11-30]. Dostupné

z: https://www.destatis.de/EN/Press/2023/10/PE23_420_811.html

[29] IMF. *Exchange Rates incl. Effective Ex. Rates – IMF Data* [online] [vid. 2023-11-30].

Dostupné z: <https://data.imf.org/regular.aspx?key=61545850>

[30] Federal statistical office. *Aus- und Einfuhr (kalender- und saisonbereinigt)*

(Außenhandel): Deutschland, Monate [online] [vid. 2023-11-30]. Dostupné z: <https://www-genesis.destatis.de/datenbank/beta/statistic/51000/table/51000-0020>

[31] CZSO. Zahraniční obchod se zbožím – prosinec 2022. *Zahraniční obchod se zbožím – prosinec 2022* [online] [vid. 2023-11-30]. Dostupné

z: <https://www.czso.cz/csu/czso/cri/zahranicni-obchod-se-zbozim-prosinec-2022>

[32] Konzultace se statistikem Mgr. Petrem Gregorem

[33] WANG, Meng, Jinhua PAN, Xinghui LI, Mengying LI, Zhixi LIU, Zhao QI, Linyun LUO, Haiping CHEN, Sirui CHEN, Feng JIANG, Liping ZHANG, Weibing WANG a Ying WANG, 2022. ARIMA and ARIMA-ERNN models for prediction of pertussis incidence in mainland China from 2004 to 2021. *BMC Public Health* [online]. **22**. Dostupné z: doi:10.1186/s12889-022-13872-9

- [34] SHARMA, Anubhav, Pratik PUNJ, Niranjan KUMAR, Alok DAS a Ajit KUMAR, 2023. Lifetime Prediction of a Hydraulic Pump Using ARIMA Model. *Arabian Journal for Science and Engineering* [online]. Dostupné z: doi:10.1007/s13369-023-07976-6
- [35] CZSO. Sezónně očištěná data. *Sezónně očištěná data* [online] [vid. 2023-11-30]. Dostupné z: https://www.czso.cz/csu/czso/sezonne_ocistena_data
- [36] Germany projection note OECD Economic Outlook June 2023. *issuu* [online] [vid. 2023-11-28]. Dostupné z: <https://issuu.com/oecd.publishing/docs/germany-oecd-economic-outlook-june-2023>
- [37] EUROPEAN COMMISSION. *Economic forecast for Germany* [online] [vid. 2023-11-28]. Dostupné z: https://economy-finance.ec.europa.eu/economic-surveillance-eu-economies/germany/economic-forecast-germany_en
- [38] ALAQEELI, Omar a Raad ALTURKI, 2023. Evaluating the Performance of the Generalized Linear Model (glm) R Package Using Single-Cell RNA-Sequencing Data. *Applied Sciences* [online]. **13**, 11512. Dostupné z: doi:10.3390/app132011512
- [39] RAGAZOU, Konstantina, Ioannis PASSAS, Alexandros GAREFALAKIS, Eleni ZAFEIRIOU a Grigorios KYRIAKOPOULOS, 2022. The Determinants of the Environmental Performance of EU Financial Institutions: An Empirical Study with a GLM Model. *Energies* [online]. **15**, 15. Dostupné z: doi:10.3390/en15155325
- [40] BURNS, David. Forecast vs actual variance. ExcelJet [online]. ExcelJet [Accessed 2021-07-10]. Available at: <https://exceljet.net/formula/forecast-vs-actual-variance>.
- [41] FLIGHTRADAR24. Live Flight Tracker – Real-Time Flight Tracker Map. *Flightradar24* [online] [vid. 2023-11-30]. Dostupné z: <https://www.flightradar24.com/how-it-works>.
- [42] EUROSTAT. *Statistics | Eurostat* [online] [vid. 2023-11-30]. Dostupné z: https://ec.europa.eu/eurostat/databrowser/view/avia_gor_de__custom_8768044/default/table?lang=en
- [43] IATA, 2022. Air Cargo Market Analysis December 2022 [online]. Dostupné z: <https://www.iata.org/en/iata-repository/publications/economic-reports/air-cargo-market-analysis---december-2022/>
- [44] BOEING, [b.r.]. *Boeing Launches the New 777-8 Freighter* [online] [vid. 2023-11-30]. Dostupné z: <https://www.boeing.com/commercial/freighters>

Přílohy

Příloha 1: Dataskript z programu R pro model ARIMA+GLM

Příloha 2: Data tras, socioekonomický proměnných, kalkulace vytížení

Příloha 3: Výsledky predikce ARIMA, GLM, ARIMA+GLM

Příloha 4: Číselné výsledky predikce ARIMA, ARIMA+GLM