

FMS 2022-22
ISBN 978-9935-522-24-5



Ferðamálastofa
Icelandic Tourist Board

SPÁLÍKAN UM FERÐAÞJÓNUSTU ÁFANGASKÝRSLA: PRÓFUN Á SPÁJÖFNUM

ÁGÚST 2022


INTELLECON

D6 Prófun á spájöfnum

Drög – 30/08/2022

Spálíkan um ferðaþjónustu

© Ferðamálastofa 2022

Útgefandi: Ferðamálastofa - Geirsgötu 9, 101 Reykjavík / Hafnarstræti 91, 600 Akureyri
Netfang: upplýsingar@ferdamalastofa.is
Veffang: www.ferdamalastofa.is
Titill: Spálíkan um ferðaþjónustu - Áfangaskýrsla: Prófun á spájöfnum
Númer: FMS 2022-22
ISBN: 978-9935-522-24-5

Öll réttindi áskilin. Skýrsluna má ekki afrita með neinum hætti, svo sem með ljósmyndun, prentun, hljóðritun eða á annan sambærilegan hátt, að hluta eða í heild, án skriflegs leyfis útgefanda.

EFNISYFIRLIT

Inngangur	3
Spár til millilangs og langs tíma - Heilbrigðisgreining á spájöfnum	4
1. Árlega spájafnan	4
1.1 Fallform.....	5
1.2 Stöðugleiki stuðla.....	5
1.3 Brot í fallinu.....	6
1.4 Viðbótarútskýringarbreytur.....	7
1.5 Lengra gagnatímabil	7
1.6 Raðaðar innri spár	8
1.7 Niðurstöður: Samantekt varðandi spájöfnur til millilangs og langs tíma	9
Spár til skamms tíma - greining á spájöfnum	10
2. Líkön til skamms tíma	10
2.1.1 Um gögnin og spáskekkjur – fjöldi ferðamanna.....	11
2.1.2 Marktækni stuðla – fjöldi ferðamanna.....	12
2.1.3 Greining á afgangslíðum ólíkra líkana – fjöldi ferðamanna	13
2.2. Fjöldi gistinátta	16
2.2.1 Um gögnin og spáskekkjur– fjöldi gistnátta	16
2.2.2 Marktækni stuðla – fjöldi gistinátta	17
2.2.3 Greining á afgangslíðum ólíkra líkana – fjöldi gistinátta	18
2.3 Eyðsla ferðamanna.....	21
2.3.1 Um gögnin og spáskekkjur – eyðsla ferðamanna	21
2.3.2 Marktækni stuðla – eyðsla ferðamanna	22
2.3.3 Greining á afgangslíðum ólíkra líkana – eyðsla ferðamanna.....	22
2.4. Niðurstöður: Samantekt varðandi spájöfnur til skamms tíma.....	26
Tilvísanir.....	27
Viðauki A.....	28

INNGANGUR

Hér má finna ítarlega prófun á spájöfnum, próf fyrir stöðugleika og áreiðanleika þeirra líkana sem unnið er með.

Rannsóknarverkefnið snýst um að búa til spálíkan fyrir ferðapjónustu þar sem megin árherslan er lögð á að spá fjölda ferðamanna til Íslands. Verkefnið felur í sér ýmsa verkþætti sem tengjast innbyrðis.

Á fyrri stigum verkefnisins var sjónum beint að fræðilegum grunni spágerðarinnar, ólíkum líkönum og gagnasöfnun. Hér birtast fyrstu niðurstöður ólíkra líkana og spáa. Um er að ræða skref í átt að fullgerðri umgjörð um slíkar spár. Eðli málsins samkvæmt er um að ræða bráðabirgðaniðurstöður. Hér er einungis um að ræða lýsingu á hluta þeirrar vinnu sem farið hefur fram við prófun á spájöfnum en tilgangurinn er að undirbyggja fyrstu spár sem verða birtar fljótlega.

Umfjölluninni er skipt í tvo hluta. Annars vegar spár með ársögnum til millilangs og langs tíma. Hins vegar spár til skemmri tíma með mánaðargögnum.

Um er að ræða afurð verkefnis sem er í vinnslu og því má gera ráð fyrir að afurðin muni taka einhverjum breytingum í verkferlinu.

SPÁR TIL MILLILANGS OG LANGS TÍMA - HEILBRIGÐISGREINING Á SPÁJÖFNUM

Í fyrri minnisgreinum var gerð grein fyrir mati á árlegum spájöfnum til langs nokkurra (1-3) ára í senn (minnisgreinar 6 og 7) og til ársfjórðungslegum spájöfnum til nokkurra (4-8) ársfjórðunga í senn. (minnisgrein 8). Í þessari minnisgrein verður framkvæmd heilbrigðisgreining (e. diagnostic checks) á þessum spájöfnum. Slík greining felur í sér upplýsingar um hversu traust hin metnu sambönd eru og gefur vísbendingu um æskilega tíðni á að uppfæra matið með nýjum gögnum.

1. Árlega spájafnan

Í minnisgrein 6 er gerð grein fyrir því sem nefnt er sémileg spálíking. Tölfræðilegt form þeirrar jöfnu er:

$$(1) \quad \frac{\Delta x(t)}{x(t-1)} = a_0 + a_1 \cdot gdp(t) + a_2 \cdot gdp(t-1) + b \cdot x(t-1) + u(t),$$

þar sem $x(t)$ er fjöldi erlendra farþega til landsins, gdp vísitala fyrir verga landsframleiðslu í OECD löndum og $u(t)$ er tilviljunarkenndur slembiliður. Háða breytan í jöfnunni, þ.e. sú stærð sem verið er að útskýra, er hlutfallsleg breyting í fjölda ferðamanna frá einu ári til annars.

Líking (1) var metin með venjulegri aðferð minnstu kvaðrata (VAMK). Sú matsaðferð er jafngild aðferð hámarkslíkinda (e. maximum likelihood) ef slembiliðurinn er hvítur hávaði og fallformið rétt. Stöðluð tölfræðipróf fyrir misdreifni, sjálffylgni og normaldreifingu afganga gaf til kynna að afgangar væru hvítur hávaði (e. white noise) og renndu þar með stoðum undir að VAMK væri jafngild aðferð hámarkslíkinda. Frekari heilbrigðisgreining m.a. á (i) misdreifni aðganga og (ii) stöðugleika hinna metna stuðla staðfesti þetta enn frekar.

Með umritun á (1) er hin metna spájafna:

$$(2) \quad x(t) = x(t-1) + (-5.5857 + 13.978 \cdot gdp(t) - 8.1639 \cdot gdp(t-1) - 0.000638 \cdot x(t-1)) \cdot x(t-1).$$

1.1 Fallform

Fyrir mat sem byggist á tímaröðum byggjast flest próf á réttmæti fallformsins á því að kanna sjálffylgni afgangsstæðra. Eftirfarandi próf fyrir fallformið voru framkvæmd:

Tafla 1 Sjálffylgnipróf á afganga				
Tafnir sjálfylnistuðar	t-próf á metna stuðla	Lagrange-margfaldarapróf	DW-próf	Box-Pierce-Ljung próf
m	$t(n-m)$	Normal-dreifing	Durbin-Watson gildi	$\chi^2(m)$
1	0.247	0.50	1.84	0.08
2	-1.19	1.87	2.45	2.21

Eins og ráða má af töflu 1 er H_0 tilgátunni um að viðkomandi sjálfylnistuðlar séu núll aldrei hafnað miðað við venjuleg mörk (t.d. 5% eða 1%).

Sé hins vegar horft til afgangslíðanna í heild er meiri óvissa. Pinkse-próf (Pinkse 1998) á hvort engin tengsl séu milli afgangslíða gefur gildið 8.49 sem hefur dreifinguna $\chi^2(1)$. Samkvæmt þessu prófi er miðað við venjuleg mörk unnt að hafna núll-tilgátunni um að engin tengsl séu milli afgangslíða.

Niðurstaðan er því sú að þar sem flest prófin gefa slíka niðurstöðu og ekkert þeirra er mjög öflugt séu litlar vísbendingar séu um að fallformið sé ekki viðeigandi.

1.2 Stöðugleiki stuðla

Til að kanna stöðugleika hinna metnu stuðla voru framkvæmd Hansen (1992) og svokölluð CUSUM og CUSUMSQ próf (sjá t.d. Brown, Durbin og Evans 1975, Harvey og Collier 1977 og Harvey 1990).

Niðurstöður í Hansen (1992) prófi fyrir stöðugleika stuðla og varians afgangslíða voru:

Tafla 2 Hansen próf fyrir stöðugleika stuðla			
Stuðlar	Hansen gildi	Höfnunargildi (10%)	Niðurstaða
a_0	0.088	0.353	Stöðugt
a_1	0.088	0.353	Stöðugt
a_2	0.087	0.353	Stöðugt
b	0.097	0.353	Stöðugt

Varians	0.204	0.353	Stöðugt
Allt í einu	0.790	1.280	Stöðugt

Niðurstöður Harvey-Collier t -prófs á afganga úr endurteknu (e. recursive) mati voru:

Endurtekið mat áfram: $t(5)$ -1.69

Endurtekið mat afturábak: $t(5)$: -1.66

Það er með öðrum orðum ekki forsendur til að hafna H_0 tilgátunni um stöðuga stuðla.

Niðurstöður úr CUSUMSQ prófi (Brown, Durbin og Evans 1975, Harvey 1990) voru:

Tafla 3 CUSUMSQ-próf			
Endurtekið (recursive) mat: Áfram			
Athuganir	Lægri mörk (5%)	CUSUMSQ	Efri mörk (5%)
6	-0.342	0.003	0.675
7	-0.175	0.003	0.841
8	-0.09	0.105	1.009
9	0.158	0.295	1.175
10	0.325	0.994	1.342
11	0.491	1.000	1.509
Endurtekið (recursive) mat: Afturábak			
Athuganir	Lægri mörk (5%)	CUSUMSQ	Efri mörk (5%)
7	-0.342	0.014	0.675
6	-0.175	0.623	0.841
5	-0.009	0.973	1.009
4	0.158	0.982	1.175
3	0.325	0.982	1.342
2	0.491	1.000	1.509

Eins og sjá má í töflu 3, eru CUSUMSQ ætíð innan 5% tölfræðilegra vikmarka. Því er ekki heldur unnt að hafna því samkvæmt CUSUMSQ-prófinu að stuðlar metnu líkingarinnar séu stöðugir.

1.3 Brot í fallinu

Vel-þekkt próf á því hvort fallið (þ.e. stuðlar þess) taki breytingum á einhverjum tímapunkti á matstímabilinu eru Chow og Goldfeld-Quandt (G-Q) próf. Niðurstöður þessara prófa fyrir brotapunkt á miðju tímabilinu (gagnaröðin er stutt, aðeins 11 athuganir og því ekki forsendur til prófa marga slíka brotapunkta) eru:

Chow-próf (F(4,2) dreifing): 28.77

G-Q-próf (F(1,1) dreifing): 0.20

Samkvæmt þessum niðurstöðum virðist ekki vera unnt að hafna þeirri tilgátu að ekki sé brot í fallinu.

1.4 Viðbótarútskýringarbreytur

Líking (1) útskýrir erlendan ferðamannafjölda til Íslands einungis með (i) ferðamannafjöldanum sjálfum á liðnum árum samkvæmt ólínulegri mismunajöfnu (S-bugðu kenningin sjá Arnason 2021 minningu grein 2) og (ii) landsframleiðslu í OECD-ríkjum. S-bugðu kenningin hefur mjög verulegan og stöðugan útskýringarmátt og viðist því grundvallaratriði. Ördú máli kann að gegna um landsframleiðslu í OECD löndum.

Auk landsframleiðslu í OECD-ríkjum hafa verið könnuð hafa verið áhrif tveggja annarra útskýringarbreyta, (iii) verðlags á Íslandi á mælikvarða raunvirðis erlendra myntar og (iv) framboðs á flugsætum til Íslands. Það að bæta þessum breytum, sitt í hvoru lagi, við líkingu (1) bætti útskýringarmátt líkingarinnar umtalsvert og gildi metnu stuðlanna var marktækt frá núlli. Ýmsir aðrir tölfræðilegir eiginleikar jöfnunnar urðu hins vegar lakari.

Notkun þessara útskýringarbreyta í stað landsframleiðslu í OECD gaf einnig bæði minni útskýringu og lakari tölfræðilega eiginleika.

Niðurstaða var því sú að landsframleiðsla í OECD löndum væri tölfræðilega séð vænlegri útskýringarbreyta en bæði verðlag á Íslandi og framboð flugsæta. Jafnframt væri ekki ástæða til að bæta annarri hvorri þessara útskýringarbreyta við.

1.5 Lengra gagnatímabil

Aðfallslíking (1) var upphaflega metin á grundvelli gagna frá 2010-20 og alls 10 athugunum (gagnapunktum) þar sem ein athugun glataðist vegna tímatafa. Nú hefur verið aflað gagna frá árinu 1999 sem gefur kost á að meta sömu jöfnu með 21 gagnapunktum.

Markverðustu niðurstöður úr því mati eru að:

- 1) Lengra gagnatímabil felur enn í sér verulegan stuðning við grunnkenninguna um S-bugðu aðlögun.
- 2) Útskýringarmáttur aðfallslíkinganna á mælikvarða R^2 er talsvert minni en áður.
- 3) Það virðist hafa orðið kerfishnykkur (structural break) á árunum 2009-11.

Niðurstaða 1) gefur ástæðu til að ætla að S-bugðu kenningin og sú mismunajafna sem lýsir henni sé eitt af grunnlögmálum þróunarinnar allt tímabilið. Niðurstöður 2) og 3) benda hins vegar til að (i) breytingar hafi orðið á stuðlum mismunajöfnunnar á tímabilinu og/eða fleiri útskýringarbreytur þurfi (þ.e. hin útskýringarbreytan, landsframleiðsla í OECD, dugi ekki til að

lýsa þróuninni). Vísbendingin um kerfishnykk 2009-11 gefur til kynna að á þeim árum hafi staða Íslands sem áfangastaðar fyrir ferðmenn breyst. Jafnframt er hugsanlegt að fjármálakreppan 2008-10 eigi hér hlut að máli.

Fyrir ársspár um fjölda ferðamanna er lærdómurinn eftirfarandi:

- S-bugðu kenningin virðist virka
- Ekki byggja spájöfnu á of löngu gagnatímabili
- Endurmeta stuðla spájöfnunnar á hverju ári þegar ný gögn liggja fyrir.

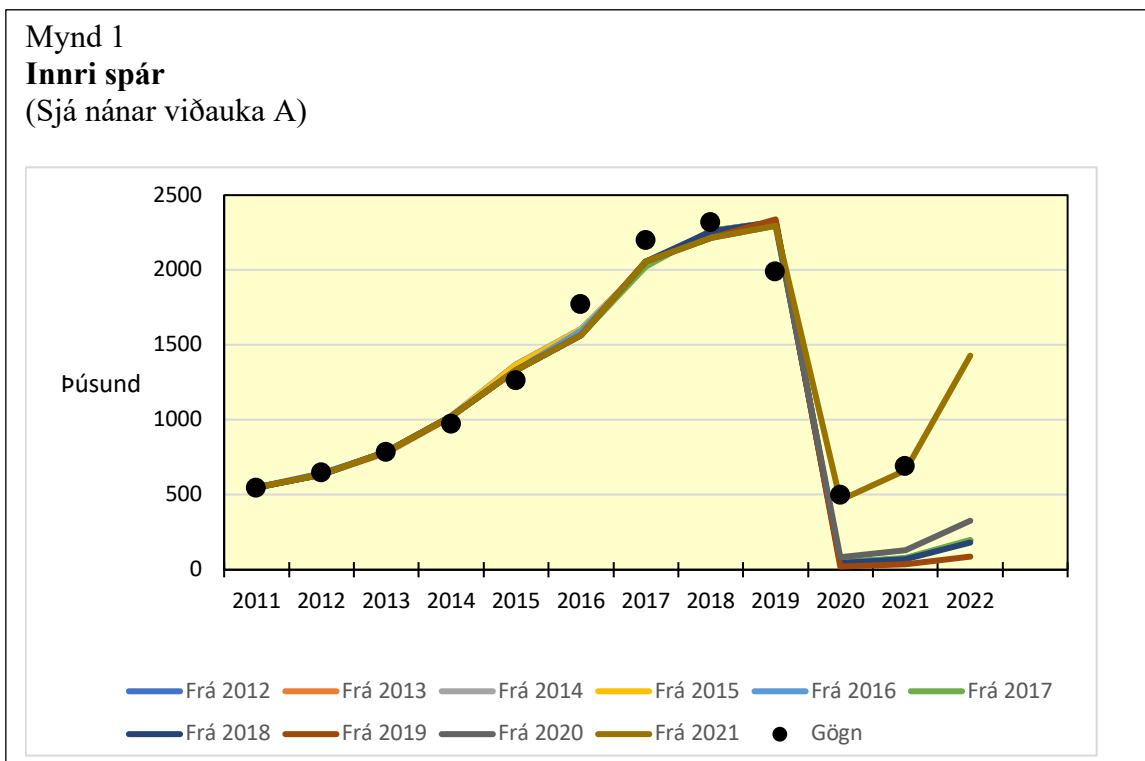
1.6 Raðaðar innri spár

Þær spár sem nýta sé mæld gildi á útskýringarbreytum eru kallaðar ytri (e. exogenous) spár. Hinar sem nota spád gildi á töfðum útskýringarbreytum (og öðrum útskýringarbreytum) eru kallaðar innri (e. endogenous) spár. Innri spár eru eðli málsins samkvæmt jafnan lakari en ytri spár. Í fyrri minnisgreinum (minnisgreinum 5 og 6, Arnason 2021b og 2021c) hefur verið gerð grein fyrir ytri spám. Eins og þar er rakið falla þær vel að þeim gögnum sem til eru. Í þessum kafla verður fjallað um innri spár.

Framkvæmdar hafa verið innri spár sem hér segir:

- Notuð er hin metna spájafna, jafna (2).
- Spád er út gagnatímabilið 2012-2021 og fyrir árið 2022.
- Mæld gildi á landsframleiðslu í OECD eru notuð.
- Spád gildi á fjölda erlendra ferðamanna eru notuð frá upphafsári innri spánna.
- Spád er fyrir mismunandi upphafsár frá og með 2012 (fyrsta mögulega spáár) til 2021. Því er fjöldi innri spáa tíu. Fram að spáári er spáin ytri spá

Helstu niðurstöður eru raktar í mynd 1. Þar sem um 10 innri spár er að ræða er á vissum árum ekki auðvelt að greina á milli spánna. Til frekari glöggvunar eru því spárnar á töluforni settar fram í töfluformi í viðauka A.



Eins og mynd 1 gefur til kynna og geggra er í töflunni í viðauka A eru innri spár á grundvelli hinnar metnu spájöfnu (2) góðar og næstum eins góðar og ytri spárnar þar til kemur að árinu 2020. Þá eru innri spár sem byggja aðeins á mældum fjölda erlendra ferðamanna fyrir 2020 talsvert fjarri lagi. Innri spáin fyrir 2021 sem nýtir mældan ferðamannafjölda á árinu 2020 er hins vegar allgóð. Því má ljóst vera að á árinu 2020 gerist eitthvað sem ekki er innifalið í metnu spájöfnunni þótt hún standi sig síðan vel á árunum 2020 þegar hún hefur réttar upplýsingar um ferðamannafjöldann 2019, sem og á árinu 2021 sem er innri spá.

Böndin berast að sjálfsögðu að áhrifum Covid árunum 2020 og 2021 sem dró mjög úr ferðamennsku um heiminn. Það nægir þó ekki því innri spárnar sem fram til 2020 vanmeta en ofmeta ekki ferðamannafjöldann 2020 og 2021. Allar spárnar ofmeta hins vegar ferðamannafjöldann árið 2019, en þá er talið að fall Wow hafi fækkað ferðamönnum til landsins verulega. Því virðist líklegt að það sé bæði fall Wow og Covid sem útskýri hina miklu skekkju innri spánna árin 2020 til 2022. Þriðja möglega skýringin er að áhrif landsframleiðslu í OECD ríkjum séu ofmetin í metnu aðfallsjöfnunni, en miklu lækkun varð í þessari breytni á árinu 2020 en áður á gagnatímabilinu.

Kjarni málsins er hins vegar sá að svo virðist sem innri spár samkvæmt jöfnum (1) og (2) séu góðar, jafnvelmörg ár fram í tímann svo framalega sem ekki verða stór ytri áföll sem eðlima'sins geta ekki verið hluti af spájöfnunni. Því virðist viðeigandi af hafa slíkan fyrirvara um spár nokkur ár fram í tímann.

1.7 Niðurstöður: Samantekt varðandi spájöfnur til millilangs og langs tíma

Heilbrigðisgreining á árlegu spájöfnu gefur til kynna að hún sé tölfræðilega sæmilega traust. Sú staðreynd að hún verður lakari ef gagnaraðirnar eru lengdar og innri spárnar geta brugðist verulega ef óvænt áföll gerast, eru rök fyrir því að

- (1) Endurmeta stuðla spájöfnunnar sæmilega oft, helst árlega
- (2) Hafa þann fyrirvara um spár meira en eitt ár fram í tímann að þær geti brugðist ef óvænt áföll eins og Covid eða hrun stórs flutningsaðila verða.

SPÁR TIL SKAMMS TÍMA - GREINING Á SPÁJÖFNUM

2. Líkön til skamms tíma

Í D5 var fjallað um þau spálíkön sem notuð eru til að spá til skamms tíma. Þar var m.a. útskýrt að slík líkön henti ekki vel til að spá lengur en nokkra mánuði fram í tímann.

Sérstaklega hefur verið unnið með eftirfarandi líkön (sjá D4 og D5) og hafa þau verið skammstöfuð til hæðarauka.

- Sjálffylgnilíkan (ARIMA)
- Veldisjöfnun (ETS)
- Veldisjöfnun með reki (THETA)
- Tauganet (NN)

Eins og lýst var í D5 er notast við KPSS próf til að ákvarða fjölda mismuna (e. differences) í Hyndman-Khandakar algrími og p, d og q í ARIMA(p, d, q) líkani eru valin þannig að AIC sé lágmarkað, þar sem notast er við skrefaða leit (e. stepwise search). Ítrun er notuð til að finna það líkan sem sýndi besta tölfræðieiginleika.¹

Spálíkön til skamms tíma geta verið af ýmsum toga. Algengast er þó að nota tímaraðalíkön sem eru sambland af sjálffylgni og hreyfanlegum meðaltölum. Algengt er að setja þau fram sem ARIMA(p, d, q) þar sem p er fjöldi tafðra breyta, d er fjöldi mismunaliða til að gera gögnin sístæð og q er fjöldi tafðra gilda afgangslíða.² Þannig eru til ótal útgáfur af slíkum líkönum, eftir ólíkum gildum á p, d og q .

Hér er sjónum beint að því að meta tölfræðilega eiginleika hinna ólíku líkana.³ Það er gert með því í fyrsta lagi að keyra þau á ólík tímabil og kanna spágildi þeirra á fyrst sjö mánuði ársins 2022.⁴ Þannig sést hve vel líkönin spá miðað við það sem raunverulega gerðist. Í öðru lagi eru útlistaðir tölfræðilegir eiginleikar ólíkra líkana, með því að skoða marktækni stuðla, dreifingu afgangslíða og fjallað um það hvort ætla megi að brot hafi átt sér stað í föllunum.

1. Nákvæmni spáa miðað við raunveruleika (MAPE, mean average percentage error).
3. Marktækni stuðla.
3. Greining á afgangslíðum.
4. Breytingar í spájöfnum yfir tíma.

Covid-19 faraldurinn hafði gríðarlega mikil og ófyrirséð áhrif á ferðaþjónustu á Íslandi sem annars staðar í heiminum eins og fjallað hefur verið um hér að framan. Af þessum sökum er

¹ Sjá D5: Mat á spájöfnum. Verkefni: Spálíkön í ferðaþjónustu, sem og Hyndman og Khandakar (2008) og Wang, Smith og Hyndman (2006).

² ARIMA stendur fyrir Autoregressive Integrated Moving Averages.

³ Tauganet eru ekki skoðuð sérstaklega enda eru þau í eðli sínu 'svartur kassi' og erfitt að átta sig á hvernig meta eigi tölfræðilega eiginleika þeirra.

⁴ Tekinn er lógaríþmi af öllum gögnum.

áhugavert að prófa tölfræðilega eiginleika hinna ýmsu líkana með því að skipta tímabilinu upp í tvennt, eins og hér segir:

- 2002-loka árs 2021 (Allt gagnasettið)
- 2002 – loka árs 2019 (allt gagnasettið til Covid-19)

Þessi gagnatímabil eru síðan notuð til að bera spár líkananna saman við það sem raungerðist fyrstu sjö mánuði ársins 2022. Er þar horft til fjölda ferðamanna, fjölda gistinguátta og eyðslu ferðamanna samkvæmt kortaveltu.

Þá hefur verið skoðað sérstaklega hvort það bæti spágetu að bæta við öðrum skýristærðum en tölum gildum. Ýmislegt var prófað í því sambandi, s.s. gengi krónunnar og breytingar í verðlagi. Með hliðsjón af reynslu af tímaraðalíkönum, ætti ekki að koma á óvart að niðurstöður slíkra greininga voru allar á þann veg að slíkt bætti ekki matið. Ástæða þessa er sú að breytingar í slíkum þáttum geta haft í för með sér áhrif á eftirspurn eftir ferðum en á móti kemur að slíkar breytingar skila sér jafnan seint í eftirspurnina og yfirleitt eru ýmis önnur áhrif jafnvel sterkari, til skamms tíma litið.

2.1.1 Um gögnin og spáskekkjur – fjöldi ferðamanna

Gögnin eru ósístæð án umbreytinga. Þetta kemur skýrt fram með útvíkuðu Dickey-Fuller prófi eins og sjá má í meðfylgjandi töflu:

Tafla 4:

Tímabil	ADF-prófstærð	fjöldi tafa	p-gildi
2002-2019	-4,4991	5	0,01
2002-2021	-2,473	6	0,377

Í gögnunum er bæði leitni og augljósar árstíðarsveiflur.

Verður nú gerð grein fyrir tölfræðilegum eiginleikum ofangreindra líkana á þessum tímabilum.

Til að bera saman óvissu ólíkra líkana er reiknað meðaltal prósentuskekkju (e. mean average percentage error) í spám fyrir hvert og eitt líkan á fyrstu sjö mánuðum ársins 2022 fyrir gagnatímabilin þrjú. Við reiknum líkönin með mismunandi gagnatímabilum. Þetta er gert til að sjá hve næm ólík líkön eru fyrir ólíkum gagnatímabilum.

Í eftirfarandi töflu má sjá meðaltalprósentuskekkju ólíkra líkana þegar metinn er heildarfjöldi ferðamanna og spáð 7 mánuði fram í tímann.

Tafla 5:

Líkön spá 7 mánuðum	2002-2021	2002-2019
ARIMA	745,58	20,55
ETS	15,28	17,21
THETA	17,28	42,10

Af þessu má glögglega sjá hve líkönin eru næm fyrir því gagnatímabili sem notað er til að spá. Sé tímabil Covid-19 tekið með eykur það mjög spáskekkju ARIMA en ekki ETS og THETA líkananna. Ástæðan er trúlega sú að leitnin og rekið (e. drift) í líkönunum tveimur nær að einhverju leyti að fanga samdráttinn sem var farið að gæta áður en Covid-19 reið yfir.

Annað sem vert er að hafa í huga er að hér er skoðuð spáskekkja, mæld á þennan hátt, þegar spáð er 7 mánuði fram í tímann. Sé spátímabilið styttra, þ.e. spáð fyrir apríl – júlí 2022, þá minnkar spáskekkjan verulega ef notað er gagnatímabilið án Covid-19. Þetta sést glögggt í eftirfarandi töflu þar sem sýnd er niðurstaða sömu spáa, en einungis spá 4 mánuði fram í tímann.

Tafla 6:

Líkön spá 4 mánuðum	2002-2021	2002-2019
	ARIMA	1236,46
ETS	20,21	2,15
THETA	19,37	38,42

Þessi staðreynd endurspeglar að líkön sem þessi eru heppilegri til að spá stutt fram í tímann en langt. Einnig sýna þessar niðurstöður glögglega hve óheppilegt er að nota gögn yfir Covid-19 tímabilið til að spá, jafnvel þótt til mjög skamms tíma sé. Sé því gagnatímabili sleppt er spáskekkja líkananna mun minni en ella.

2.1.2 Marktækni stuðla – fjöldi ferðamanna

Meðfylgjandi töflur sýna marktækni stuðla ARIMA líkansins fyrir ólík gagnatímabil.

2002-2021 ARIMA – fjöldi ferðamanna

stuðull	mat	std. Error	t-gildi	p-gildi
ma1	0,273139473	0,065166989	4,191377842	3,92E-05
ma2	-0,219389665	0,065308764	-3,359268383	9,11E-04
ma3	-0,271156238	0,073956513	-3,666428114	3,04E-04
ma4	-0,32321143	0,076531393	-4,223252918	3,43E-05
sma1	0,151642851	0,071512073	2,120520994	3,50E-02
sma2	0,45897142	0,103246111	4,445411208	1,35E-05

2002-2019 ARIMA – fjöldi ferðamanna

stuðull	mat	std. Error	t-gildi	p-gildi
ar1	0,72266777	0,069538918	10,392278	1,54E-20

ar2	0,24260668	0,069269219	3,50237357	5,68E-04
sar1	-1,2191809	0,084407444	-14,443998	5,97E-33
sar2	-0,7658037	0,073992589	-10,349735	2,06E-20
sma1	0,67695813	0,125157779	5,4088378	1,78E-07
sma2	0,29100773	0,1200731	2,42358802	1,62E-02

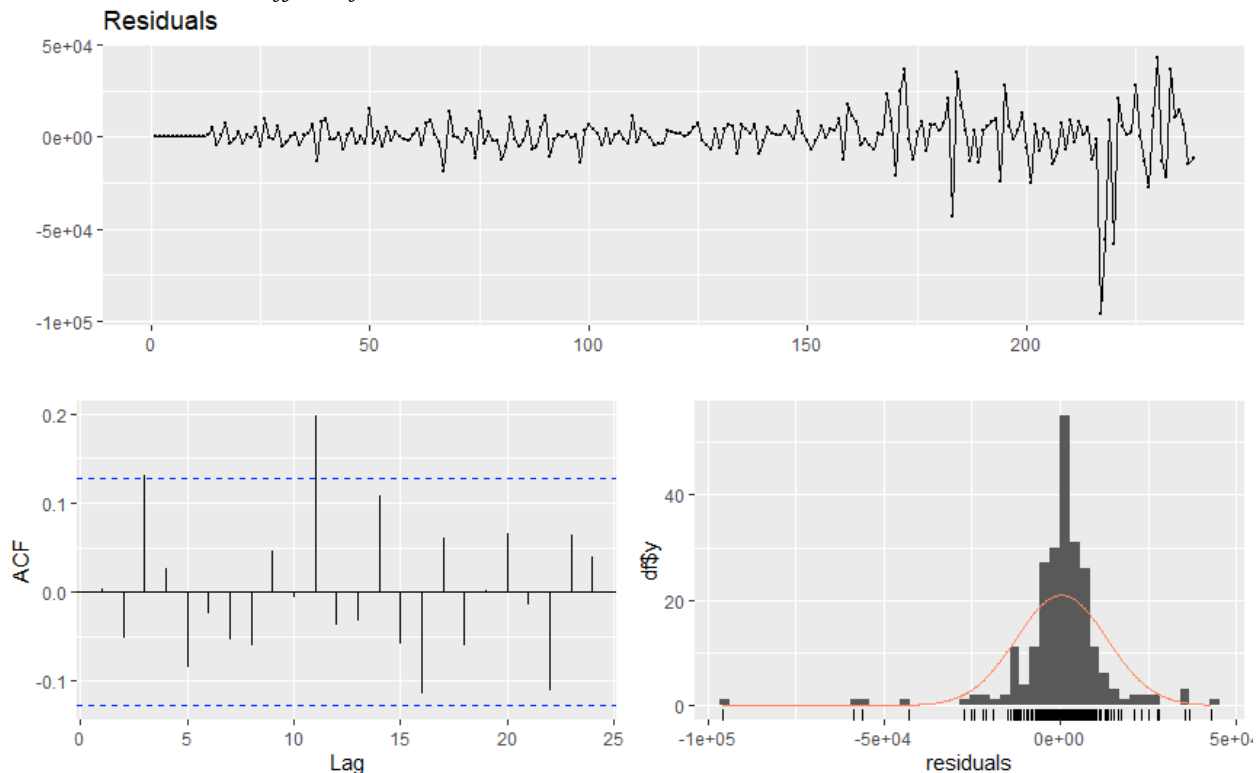
Eins og sjá má eru tölfræðilegir eiginleikar stuðlamats góðir í þeim skilningi að metnir stuðlar mælast líklegir til að vera tölfræðilega frábrugðnir núlli. Þá kemur í ljós, eins og búast mátti við, að ólík líkön eiga við ólík gagnatímabil.

2.1.3 Greining á afgangslíðum ólíkra líkana – fjöldi ferðamanna

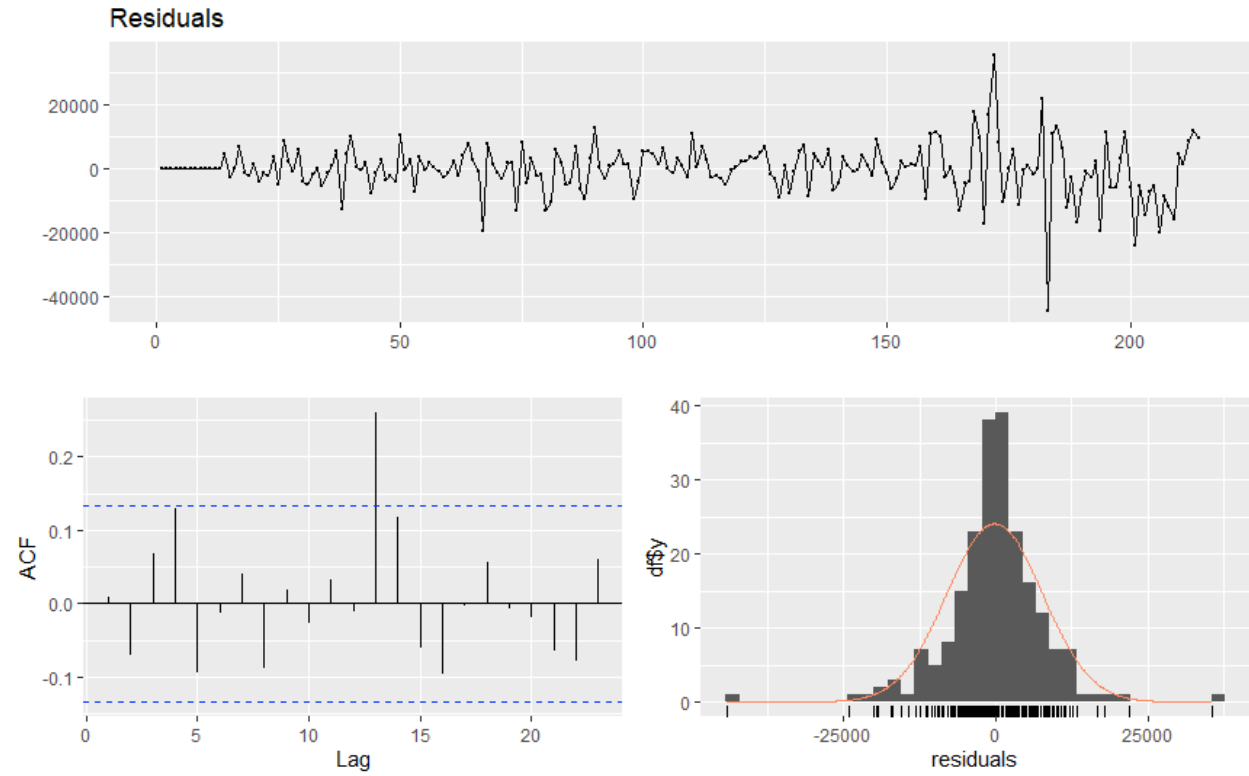
Þær aðferðir sem hér hefur verið beitt byggja á því að fortíð sem endurspeglast í gögnunum, nýtist til að spá fyrir um framtíðina. Hér hefur verið tryggt að gagnaraðirnar séu sístæðar.

Á eftirfarandi myndum má sjá afgangslíði og dreifingu hinna ólíku metnu líkana fyrir ólík tímabil.

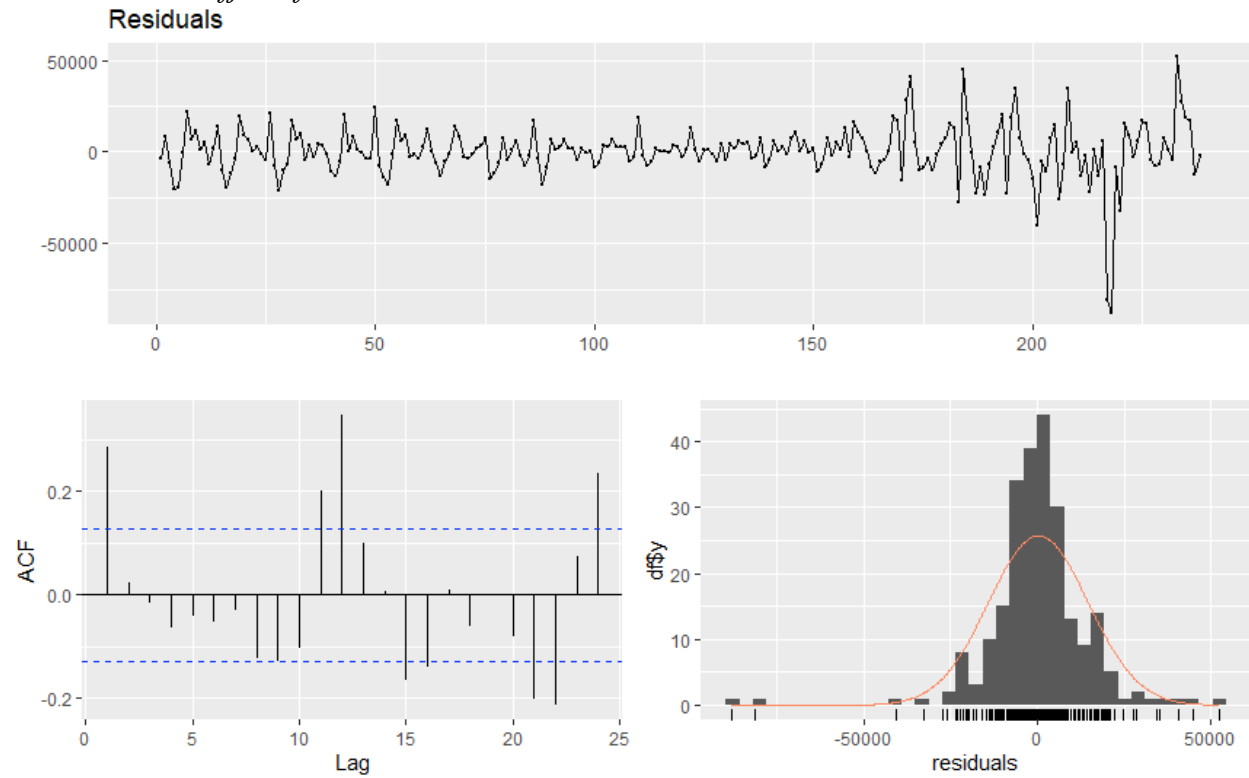
ARIMA 2002-2021 – fjöldi ferðamanna



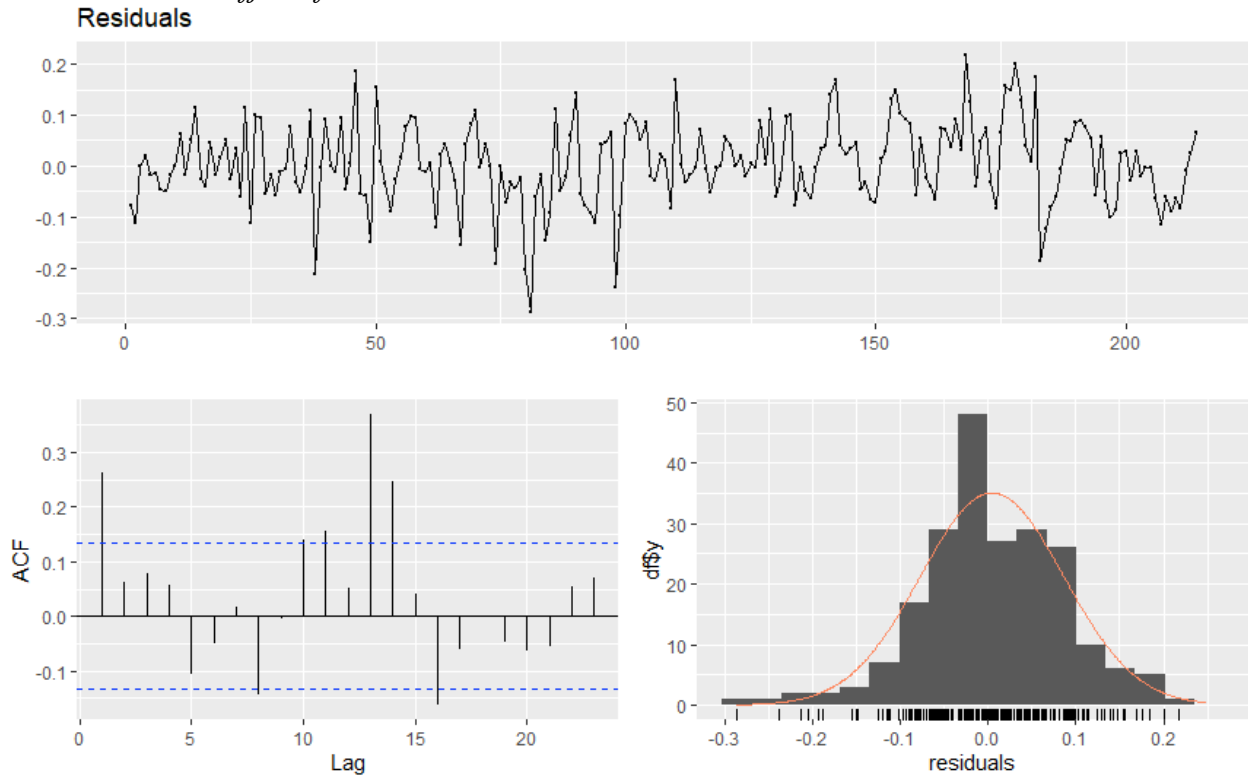
ARIMA 2002-2019 – fjöldi ferðamanna



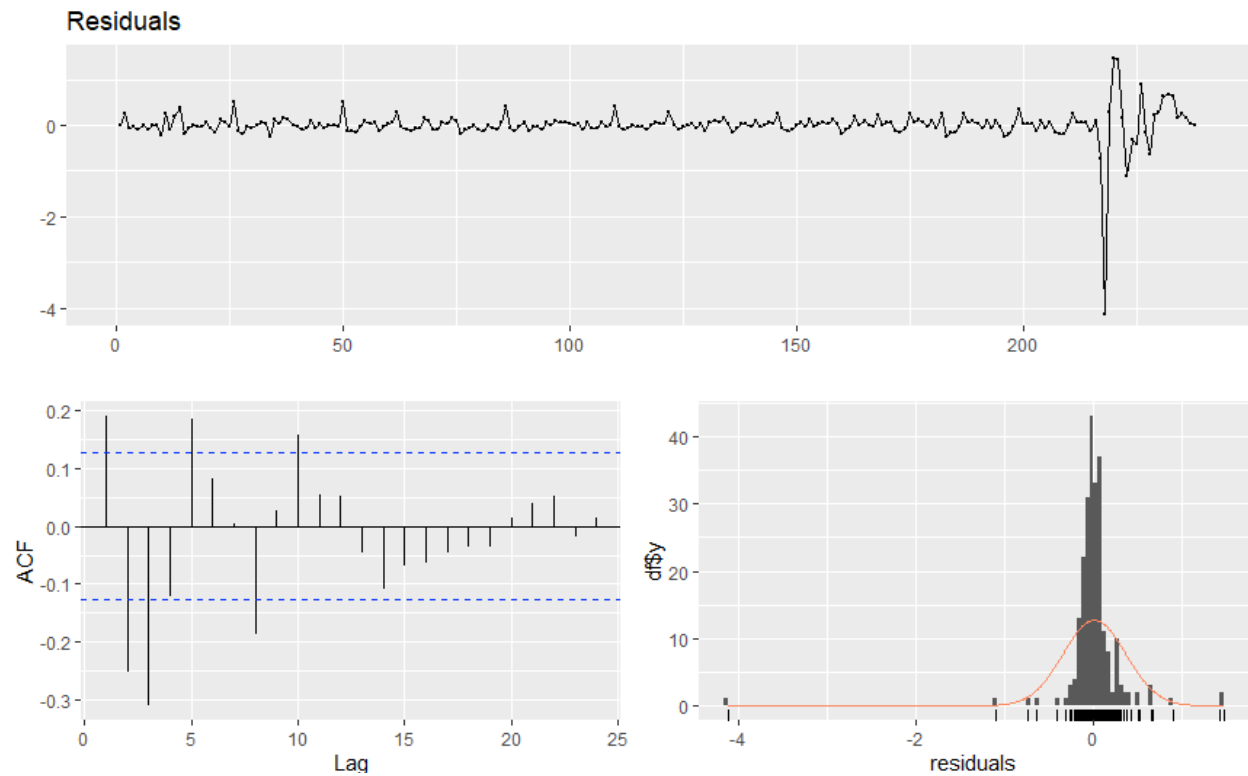
ETS 2002-2021-fjöldi ferðamanna



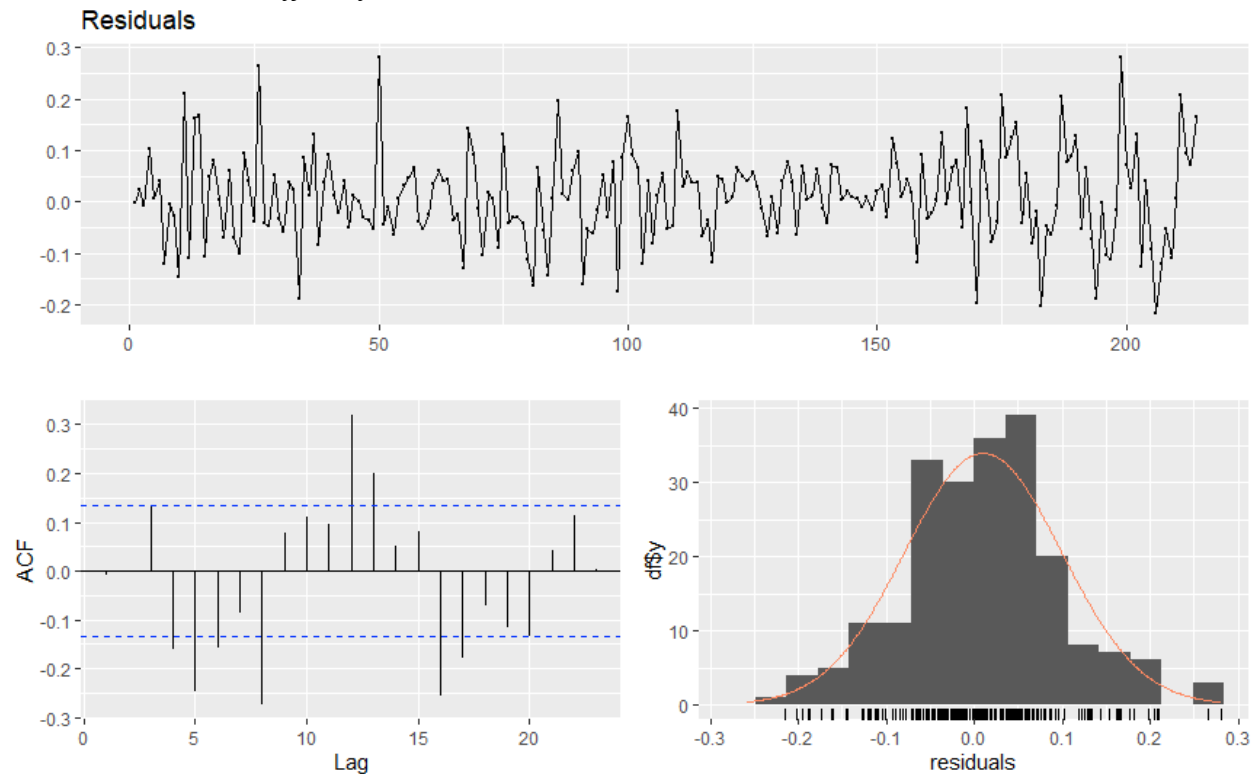
ETS 2002-2019 – fjöldi ferðamanna



THETA 2002-2021 – fjöldi ferðamanna



THETA 2002-2019 – fjöldi ferðamanna



Þessi samanburður á afgangslíðum ólíkra líkana keyrð á ólík gagnatímabil er áhugaverður. Hér sést enn og aftur hve varhugavert það er að nota gögn frá Covid-19 tímanum. Það má orða það sem svo að skýringarmáttur þeirra talna sé minni en enginn.

2.2. Fjöldi gistinátta

2.2.1 Um gögnin og spáskekkjur– fjöldi gistnátta

Í eftirfarandi töflu má sjá meðaltalprósentskekkju ólíkra líkana þegar metinn er heildarfjöldi gistinátta og spáð 7 mánuði fram í tímann.

Tafla 7:

Líkön spá 7 mánuðum	2002-2021	2002-2019
ARIMA	126,54	22,39
ETS	25,07	24,21
THETA	19,99	52,73

Á sama hátt má reikna meðaltalsprósentskekkju hinna ólíku líkana þegar spáð er 4 mánuði fram í tímann.

Tafla 8:

Líkön spá 4 mánuðum	2002-2021	2002-2019
ARIMA	182,79	11,89
ETS	25,88	13,33
THETA	19,29	52,95

Líkt og áður minnkar spáskekkjan verulega fyrir ARIMA og ETS líkönin þegar Covid-19 tímabilinu er sleppt.

2.2.2 Marktækni stuðla – fjöldi gistinátta

Í eftirfarandi töflum má sjá marktækni stuðla fyrir ARIMA líkan sem metið er með fjölda gistinátta líkt og áður.

2002-2021 ARIMA – fjöldi gistinátta

stuðull	mat	std. error	t-gildi	p-gildi
ar1	0,880729445	0,033658479	26,16664389	1,29E-70
ma1	0,379355586	0,062224033	6,096608787	4,58E-09
	-	-	-	-
sma1	0,866788759	0,070444884	12,30449546	5,09E-27
sma2	0,179415627	0,086296836	2,079052213	3,87E-02

2002-2019 ARIMA – fjöldi gistinátta

stuðull	mat	std. error	t-gildi	p-gildi
ar1	1,61149908	0,065380947	24,64783941	6,04E-63
	-	-	-	-
ar2	-0,742224212	0,063826077	11,62885536	2,75E-24
	-	-	-	-
ma1	-1,867485684	0,046793684	39,90892607	4,78E-98
ma2	0,941284396	0,042995446	21,89265325	2,44E-55
	-	-	-	-
sar1	-0,440694054	0,081234845	5,424938664	1,64E-07
	-	-	-	-
sar2	-0,145890629	0,080498597	1,812337533	7,14E-02

Sama verður sagt um tölfræðilega eiginleika þessara metnu jafna og um þær er notaðar eru til að spá fyrir um fjölda ferðamanna. Metnir stuðlar eru tölfræðilega frábrugðnir núlli og nokkur munur á milli líkana eftir því hvort tímabilið er tekið.

Þessar niðurstöður sýna enn og aftur hve líkön sem byggja á skammtímagögnum eru næm fyrir því gagnatímabili sem liggur til grundvallar. Einnig koma greinilega fram áhrif Covid-19

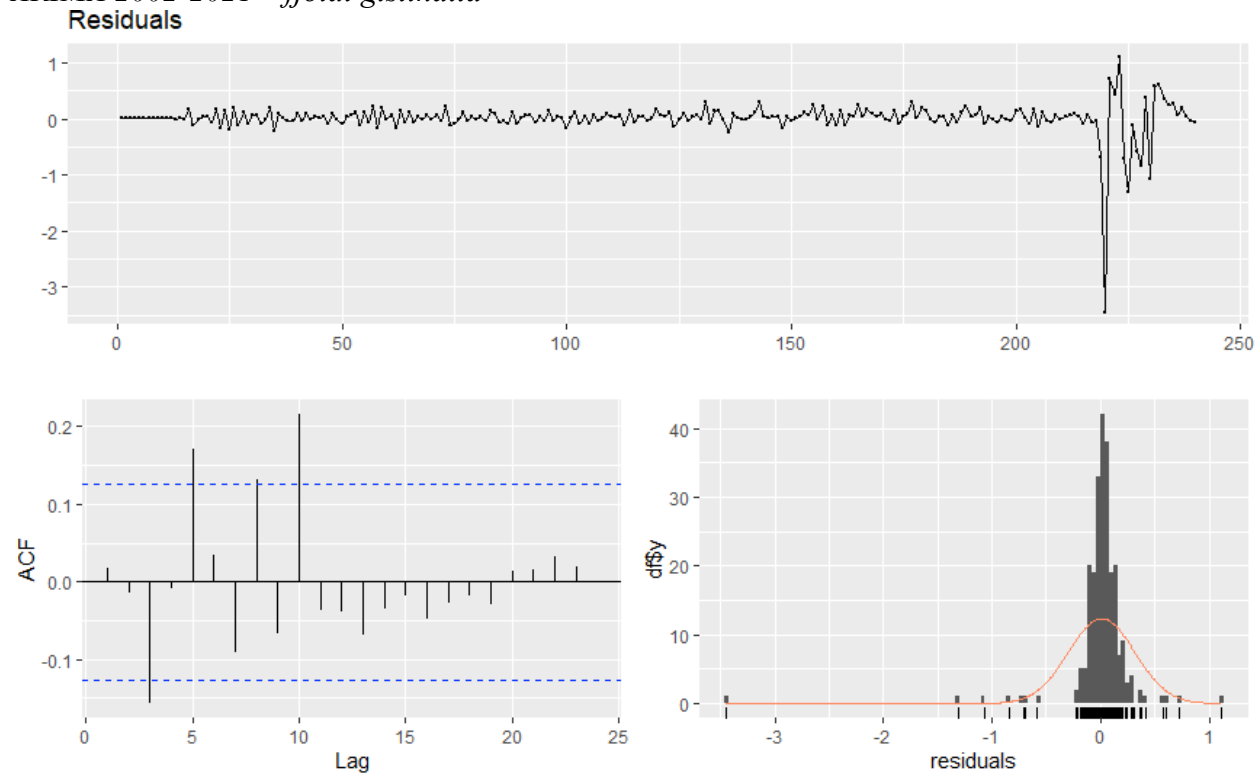
faraldursins og staðfesta þessar niðurstöður að varlega þarf að fara þegar notast á við gögn frá því tímabili til að spá fyrir um framtíðina.

Það er í eðli sjálffylgnilíkana sem þessara að stuðlar þeirra og jafnvel fallform (þ.e. p , d eða g) getur breyst yfir tíma, þar sem þau nýta upplýsingar úr fortíðinni sem skýristærðir um það sem ætla má að muni gerast í framtíðinni, að öðru óbreyttu. Stöðugleiki parametra getur þannig breyst eftir því sem tímar líða fram.

2.2.3 Greining á afgangslíðum ólíkra líkana – fjöldi gistinátta

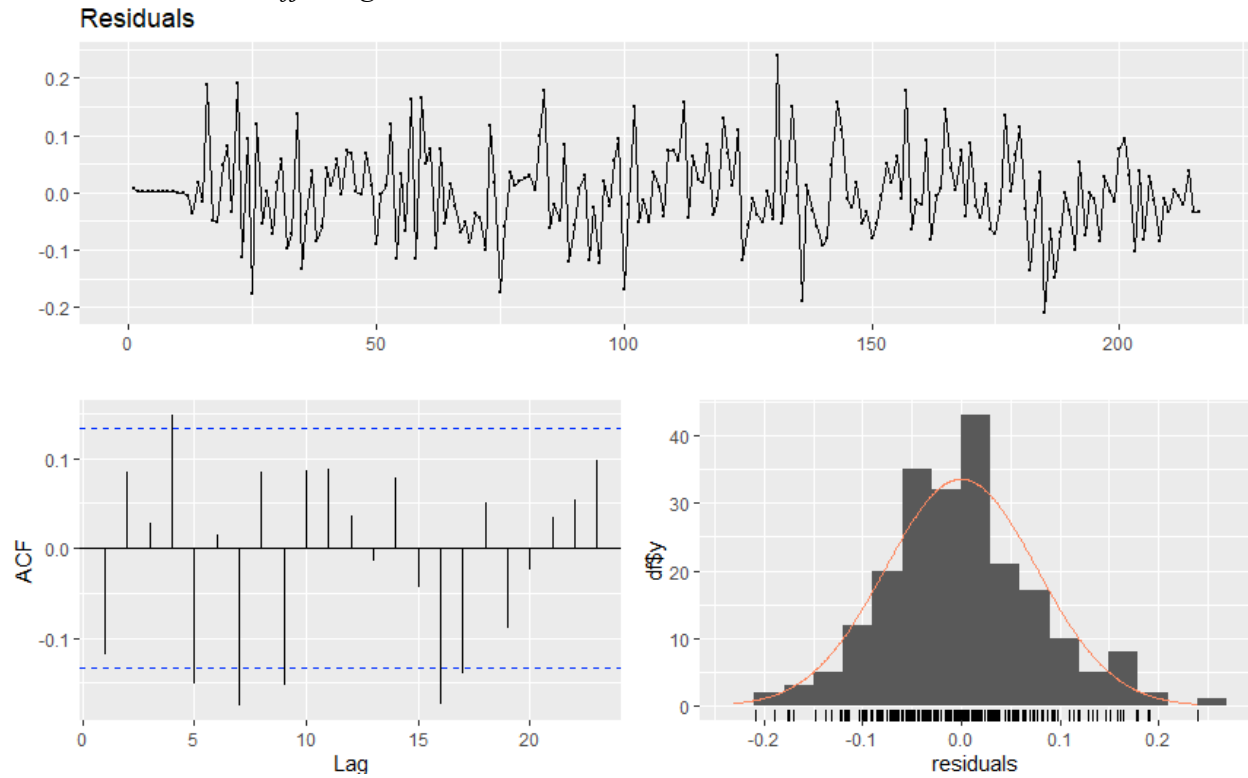
Á sama hátt má skoða afgangslíði og dreifingu þeirra fyrir fjölda gistinátta.

ARIMA 2002-2021 – fjöldi gistinátta



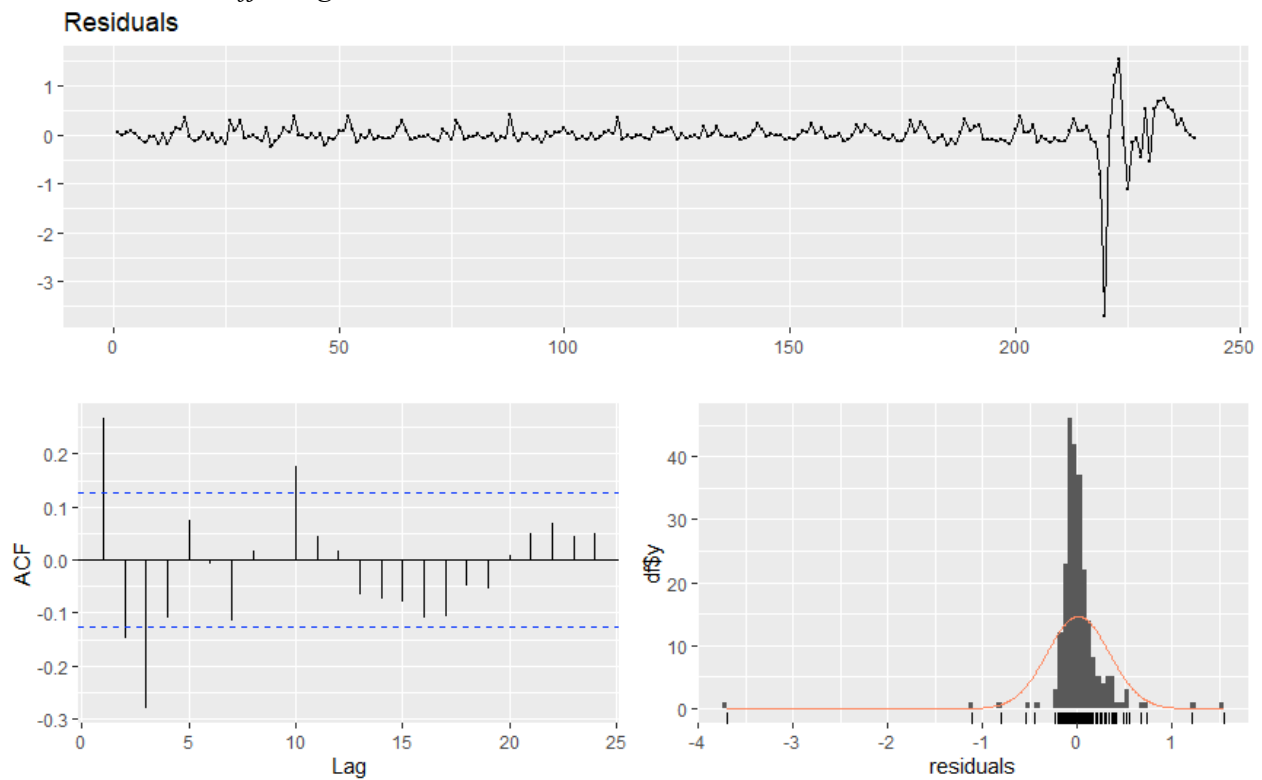
Skýrt sjást merki brots í gögnunum þegar Covid-19 faraldurinn reið yfir og nánast lokaði landinu sem og ferðaþjónustu í öðrum löndum.

ARIMA 2002-2020 – fjöldi gistinátta

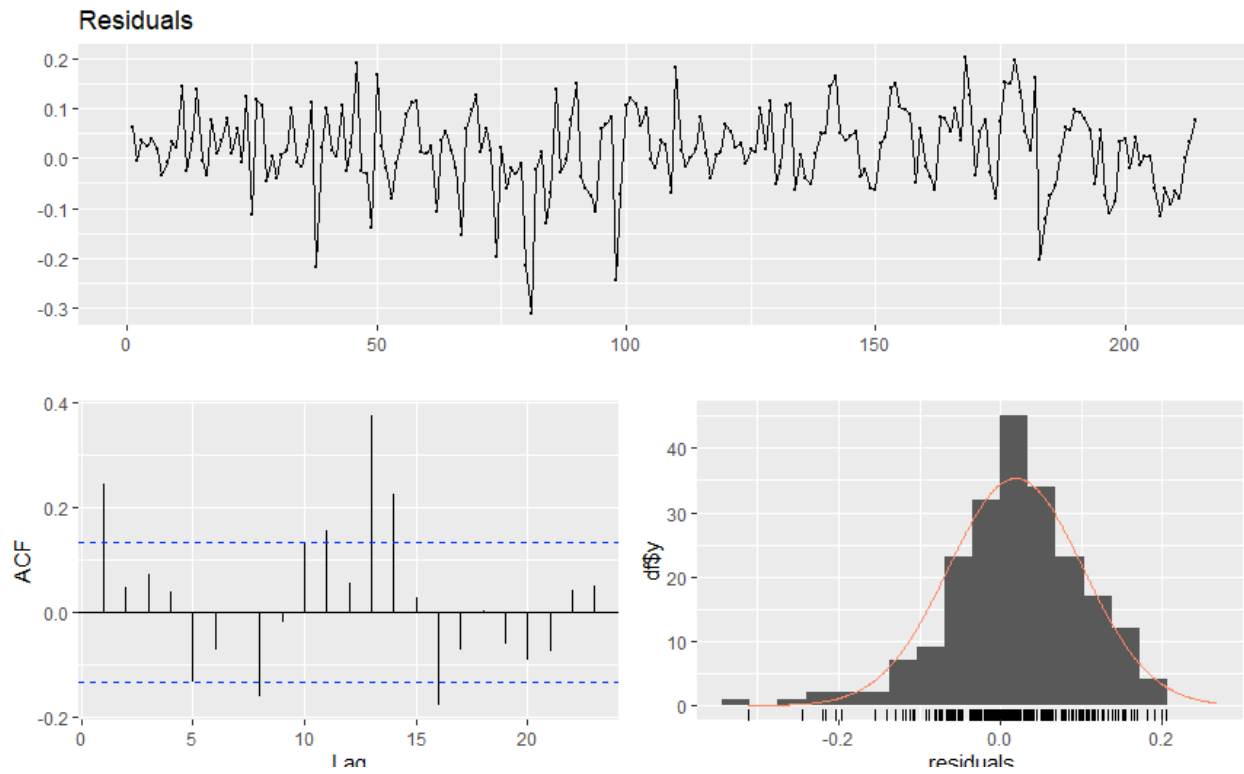


Ef Covid-19 tímabilið er tekið út sést að afgangslíðir hafa eðlilega dreifingu.

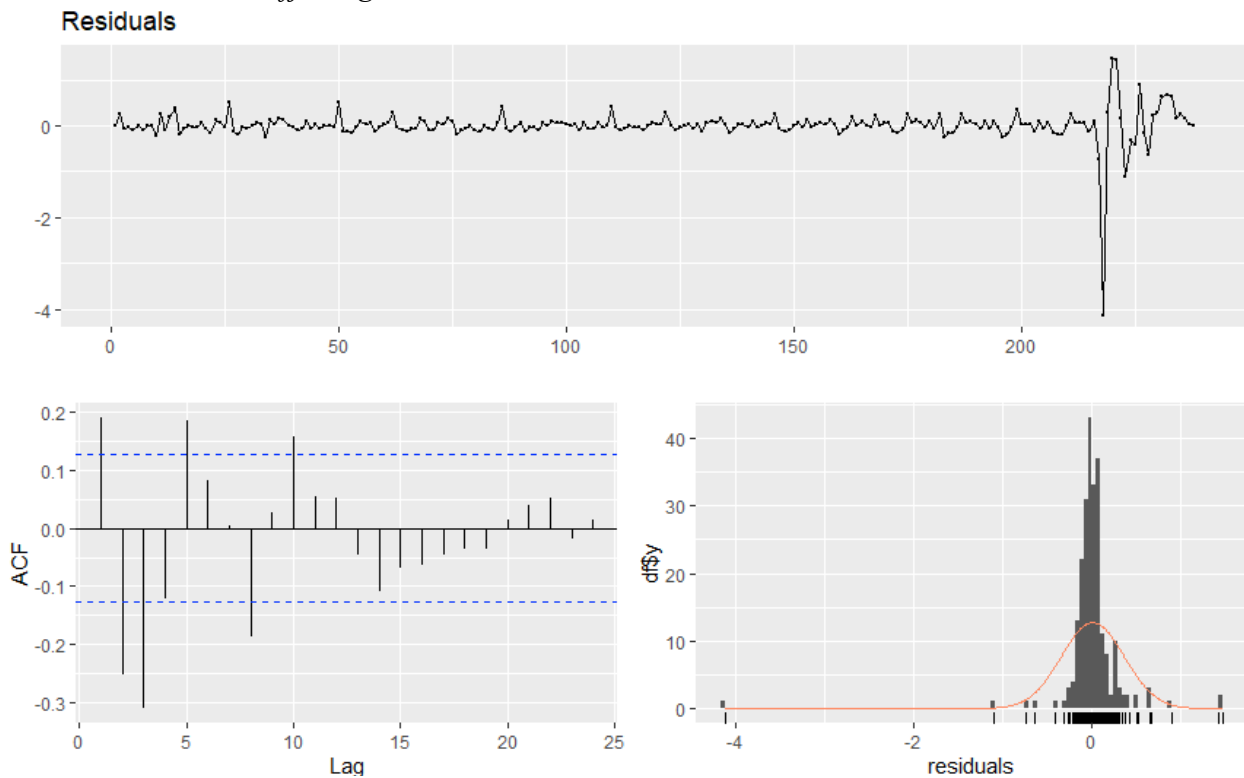
ETS 2002-2021 – fjöldi gistinátta



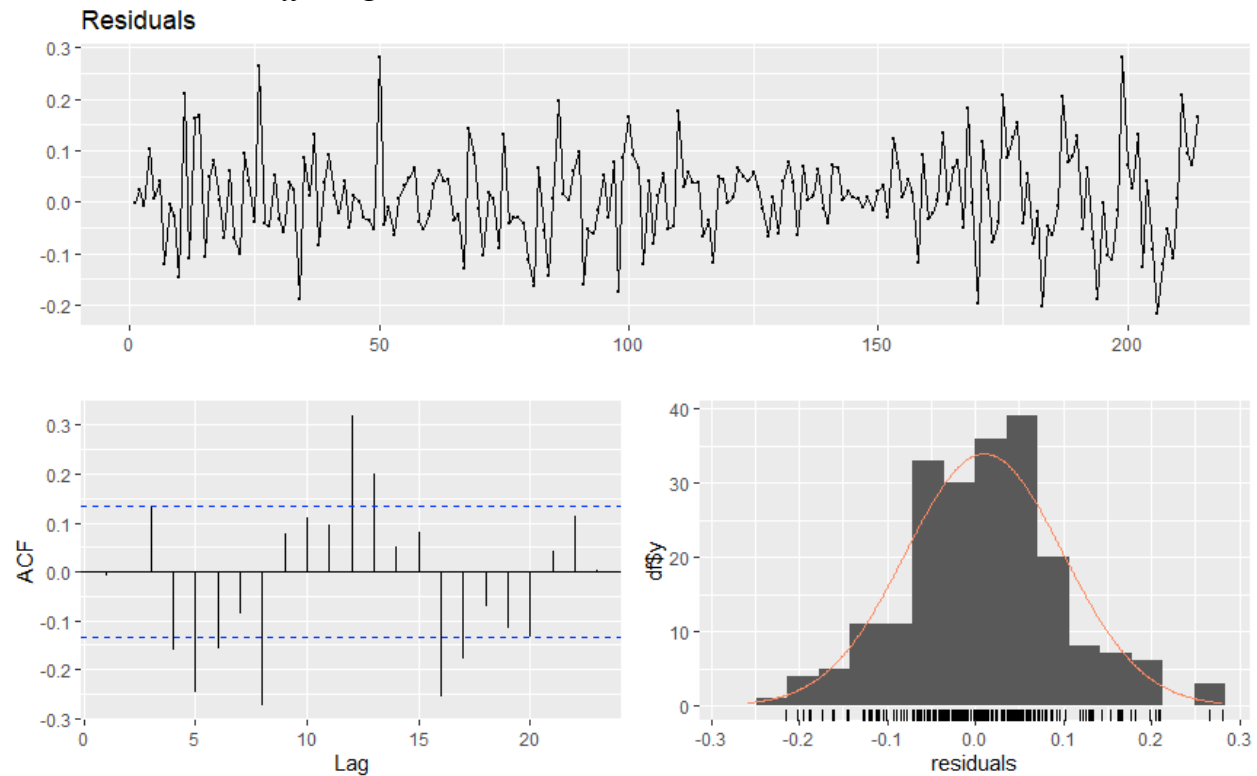
ETS 2002-2020 – fjöldi gistinátta



THETA 2002-2021 – fjöldi gistinátta



THETA 2002-2020 – fjöldi gistinátta



Greining á afgangslíðum leiðir til sömu niðurstöðu og áður, þ.e. að varasamt sé að nota upplýsingar um fjölda gistinátta á tíma Covid-19 faraldursins, þar sem slíkt skekkir spániðurstöður líkana.

2.3 Eyðsla ferðamanna

Fyrir liggja mánaðargögn um kreditkortaveltu erlendra ferðamanna á Íslandi en þau gögn ná skemmra aftur en aðrar gagnaraðir eða bara til janúar 2012.

2.3.1 Um gögnin og spáskekkjur – eyðsla ferðamanna

Í eftirfarandi töflu má sjá meðaltalprósentskekkju ólíkra líkana þegar metinn er eyðsla ferðamanna og spáð 7 mánuði fram í tímann.

Tafla 9:

Líkön spá 7 mánuðum	2012-2021	2012-2019
ARIMA	309,58	23,36
ETS	30,27	19,32
THETA	59,32	21,34

Sem áður má glögglega sjá hve líkönin eru næm fyrir því gagnatímabili sem notað er til að spá. Sé tímabil Covid-19 tekið með eykur það mjög spáskekkju ARIMA og ETS líkananna. Þetta er í samræmi við fyrri niðurstöður.

Líkt og áður er áhugavert að stytta spátímabilið. Sé spátímabilið styttra, þ.e. spáð fyrir apríl – júlí 2022, þá minnkar spáskekkjan verulega ef notað er gagnatímabilið án Covid-19. Þetta sést glögggt í eftirfarandi töflu sem sýnir prósentuspáskekkjuna fyrir ólík líkön og ólík gagnatímabil.

Tafla 10:

Líkön spá 4 mánuðum	2002-2021	2002-2019
ARIMA	501,91	19,95
ETS	42,01	10,31
THETA	90,00	12,94

Enn og aftur sýna þessar niðurstöður glögglega hve óheppilegt er að nota gögn yfir Covid-19 tímabilið til að spá, jafnvel þótt til mjög skamms tíma sé sem og að nota skammtímaspár til að spá langt fram í tímann.

2.3.2 Marktækni stuðla – eyðsla ferðamanna

Í eftirfarandi töflum má sjá marktækni stuðla fyrir ARIMA líkan sem metið er með gögnum um eyðslu ferðamanna eins og hér hefur verið lýst.

2002-2021 ARIMA – eyðsla ferðamanna

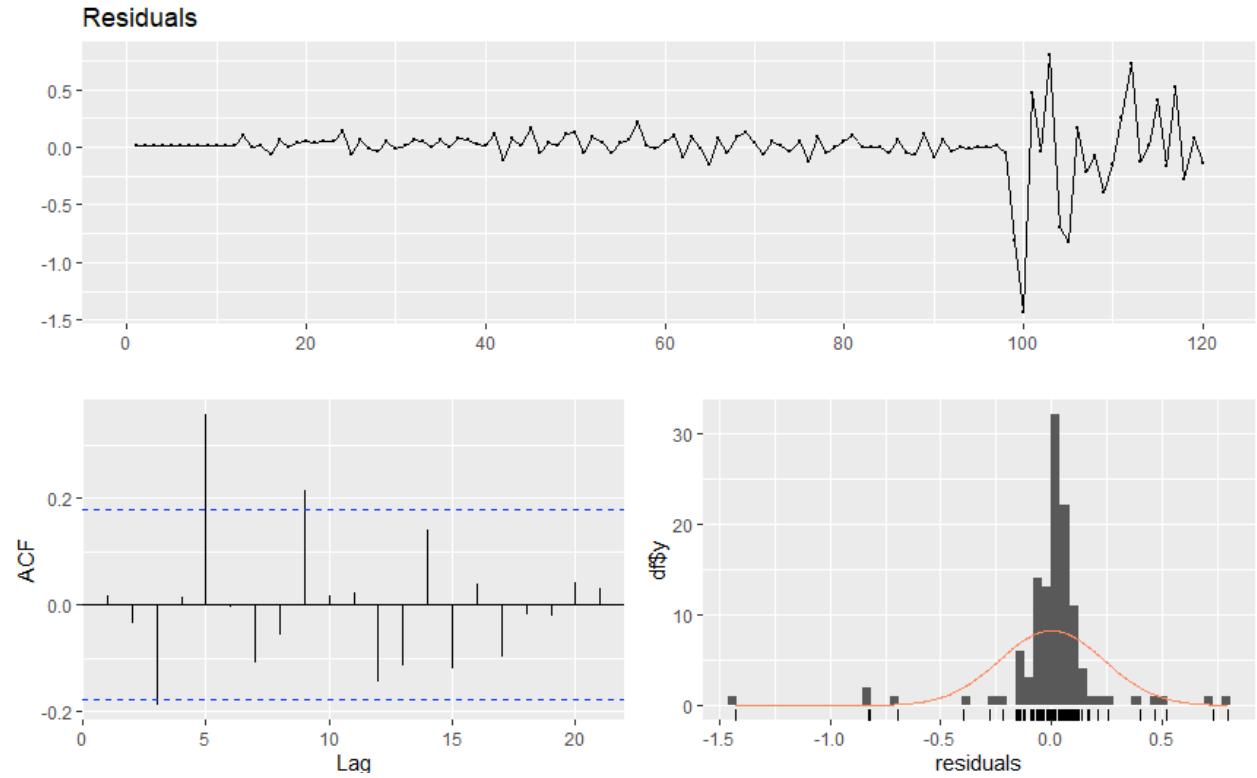
stuðull	mat	std. Error	t-gildi	p-gildi
ar1	0,877072212	0,047658858	18,40313123	4,45E-35
ma1	0,591777849	0,089505197	6,61165906	1,50E-09
sar1	-0,718110985	0,082684709	-8,684930901	4,46E-14

Sama verður sagt um tölfræðilega eiginleika þessarar metnu jafnu og um þær er notaðar eru til að spá fyrir um fjölda ferðamanna og gistinætur.

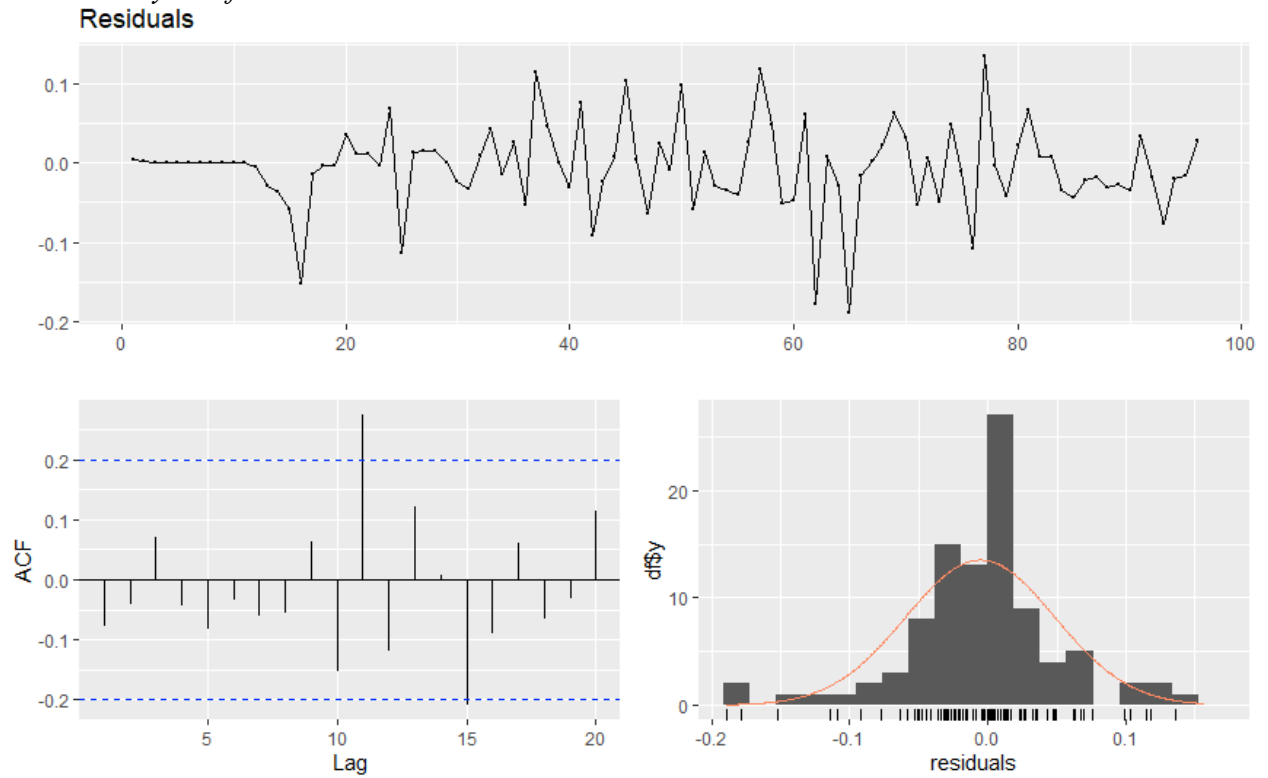
2.3.3 Greining á afgangslíðum ólíkra líkana – eyðsla ferðamanna

Hér má sjá samskonar myndir og áður af afgangslíðum og dreifingu þeirra fyrir ólík líkön og ólík gagnatímabil.

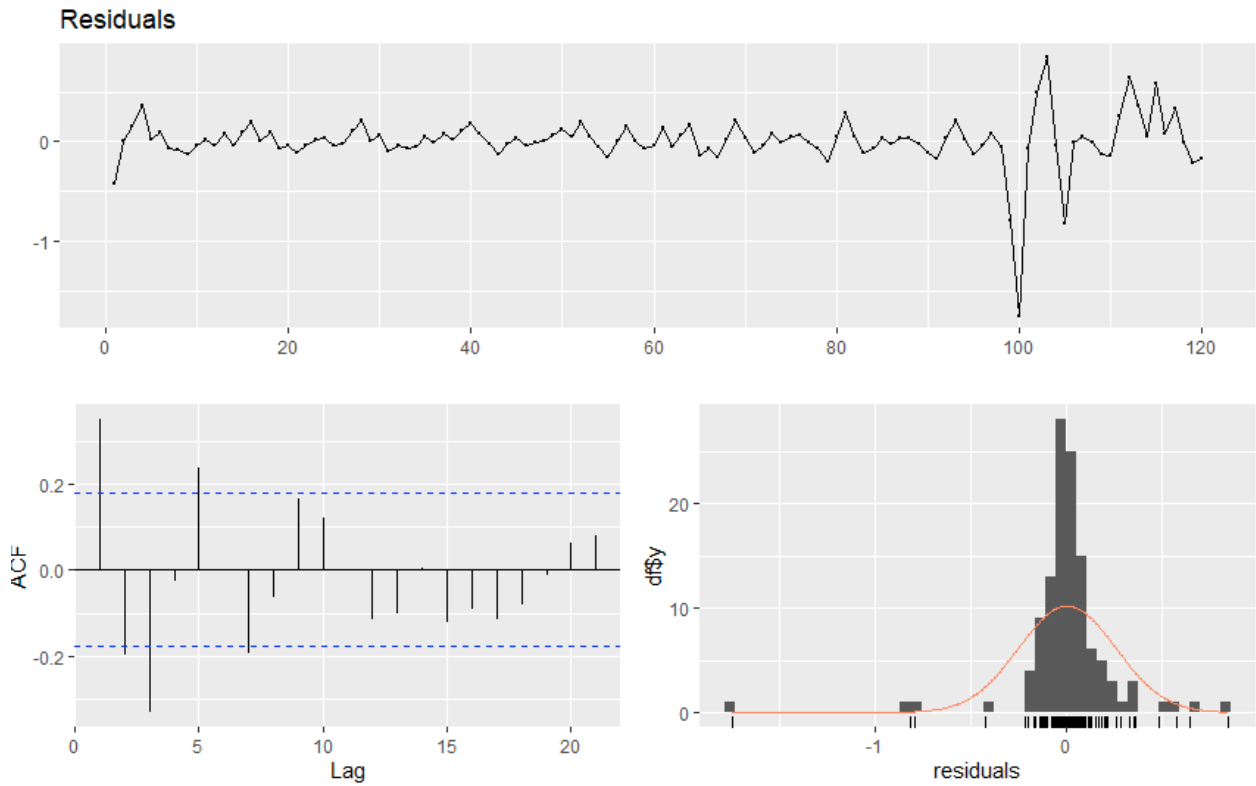
ARIMA – eyðsla ferðamanna 2018 – júlí 2022



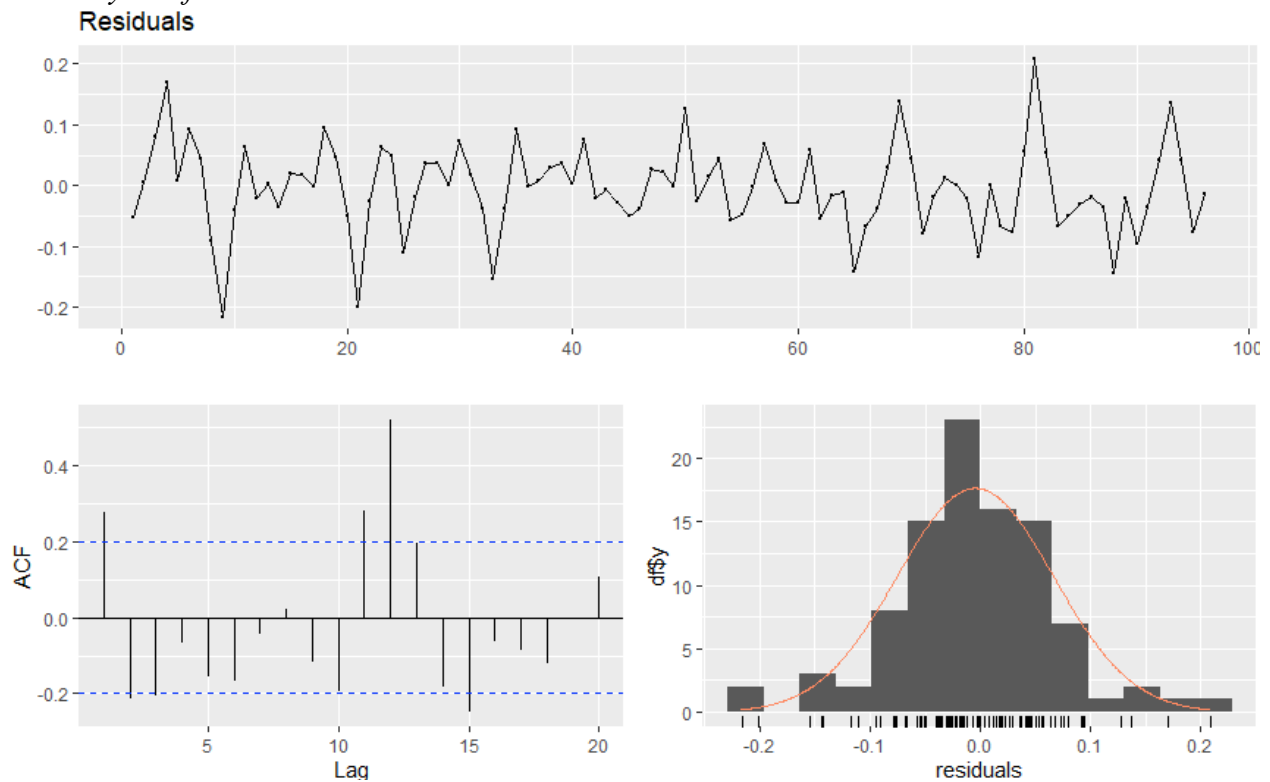
ARIMA – eyðsla ferðamanna 2018 – 2020



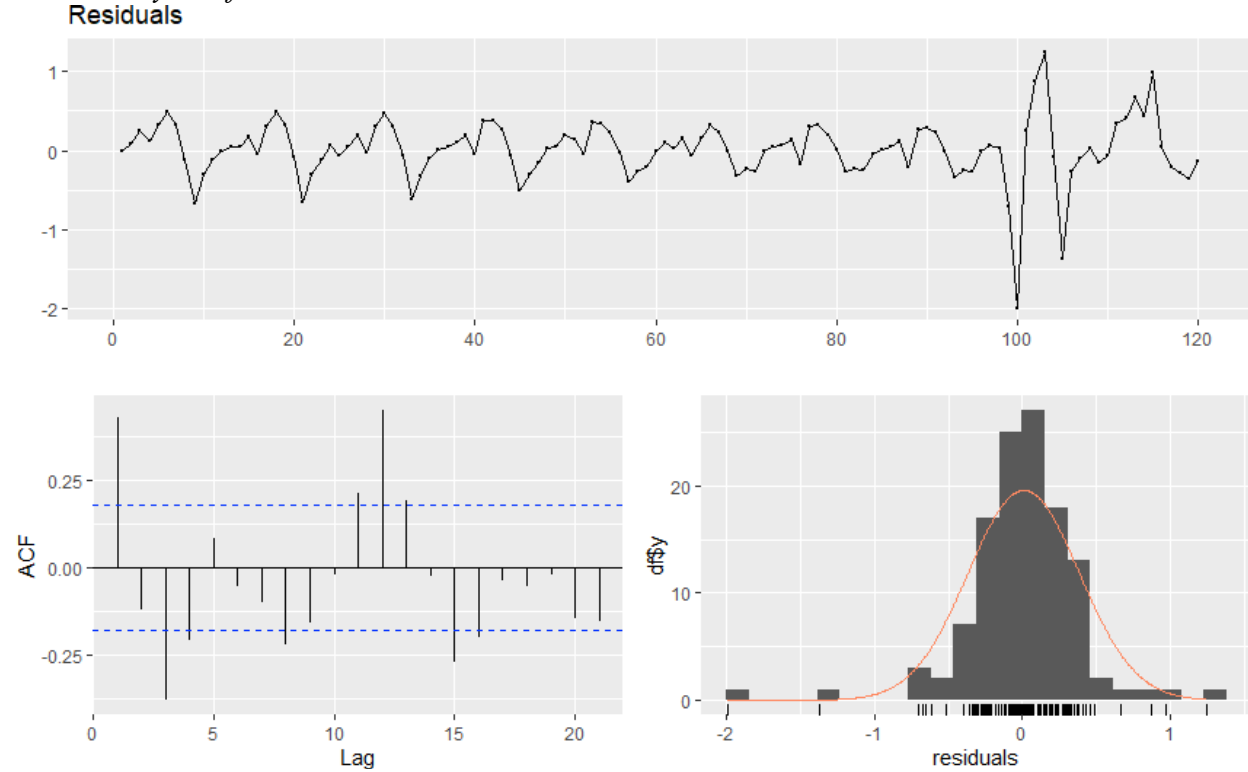
ETS – eyðsla ferðamanna 2018 – júlí 2022



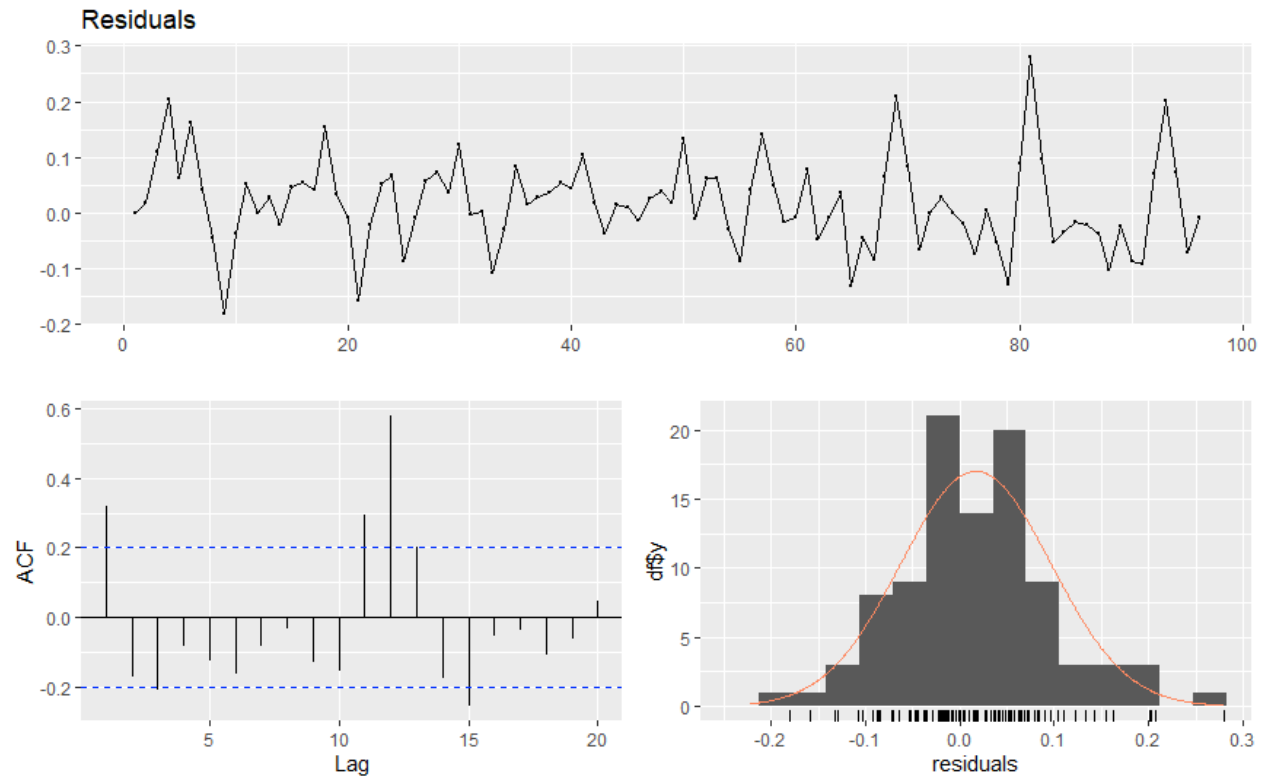
ETS – eyðsla ferðamanna 2018 – 2020



THETA – eyðsla ferðamanna 2018 – 2020



THETA – eyðsla ferðamanna 2018 – 2020



Líkt og áður sjást áhrif Covid-19 greinilega á tölfræðilega eiginleika spálíkananna. Þessar niðurstöður gefa til kynna að varlega þurfi að fara í að nota skammtímalíkön til að spá fyrir um jafn slembikennda breytur sem eyðslu erlendra ferðamanna.

2.4. Niðurstöður: Samantekt varðandi spájöfnur til skamms tíma

Sú greining sem hér hefur verið lýst á spájöfnum sem spá til skamms tíma með mánaðargögnum leiðir í ljós að líkönin eru mjög næm fyrir því hvaða gagnatímabil er valið. Spáskekkjur aukast einnig verulega eftir því sem reynt er að spá lengra fram í tímann. Þessar niðurstöður eru mjög í samræmi við niðurstöður varðandi spájöfnur til meðallangs og langs tíma sem fjallað er um í kafla 1.7 hér að ofan. Samkvæmt greiningunni er vert að huga að þrennu varðandi spár til skemmri tíma til að nýta þær með sem bestum hætti, þ.e.;

- (1) Að endurmeta þurfi spájöfnur reglulega og sem oftast.
- (2) Endurskoða þarf spárnar þegar stórir óvenjulegir atburði verða
- (3) Beita þarf jafnt niðurstöðum tölfræðilegra prófa sem byggjuviti við spár á hverjum tíma.

TILVÍSANIR

- Arnason, R. 2021. Minnisgrein 2. Líkan fyrir ferðamennsku. Verkefni: Spálíkön í ferðaþjónustu.
- Arnason, R. 2021. Minnisgrein 5. Spár til millilangs og langs tíma: Tilraun 3. Verkefni: Spálíkön í ferðaþjónustu.
- Arnason, R. 2021. Minnisgrein 6. Sæmileg spálíking til millilangs og langs tíma. Verkefni: Spálíkön í ferðaþjónustu.
- Brown, R. J. Durbin og J. Evans 1975. Techniques for Testing the Constancy of Regression Relationships over Time. *Journal of the Royal Statistical Society. Series B* (Methodological). 37:149-192.
- Davidson, R. og MacKinnon, J. 1993. *Estimation and Inference in Econometrics*. Oxford University Press. UK
- Hansen, B. 1992. Testing for Parameter Stability in Linear models. *Journal of Policy Modelling* 14:517-33.
- Harvey A. og P. Collier 1977. Testing for Functional Misspecification in Regression Analysis. *Journal of Econometrics* 6:103–119.
- Harvey, A. 1990. *The Econometric Analysis of Time Series*. 2nd ed. MIT Press.
- Hyndman, R.J. og Khandakar, Y. 2008. Automatic time series forecasting: The forecast package for R, *Journal of Statistical Software*, 26(3).
- Pinkse, J. 1998. A Consistent Nonparametric Characteristic Function Based Test for Serial Independence. *Journal of Econometrics* 84:205-31.
- Wang, X., Smith, K.A., Hyndman, R.J. (2006). Characteristic-based clustering for time series data, *Data Mining and Knowledge Discovery*, 13(3).

VIÐAUKI A

Innri og ytri spár

Spálíking: $x(t) = x(t-1) + (-5.5857 + 13.978 \cdot gdp(t) - 8.1639 \cdot gdp(t-1) - 0.000638 \cdot x(t-1)) \cdot x(t-1)$,

þar sem $x(t)$ er fjöldi erlendra ferðega til landsins, gdp vísitala fyrir verga landsframleiðslu í OECD löndum og t tími (hér ár). Í ytri spám eru mældu gildin á útskýringarbreytunum (á hægri hlið spálíkingar) notuð. Í innri spánum eru spáð gildi á $x(t-1)$ notuð frá og með upphafsári innri spárinnar. Fram að því ári er spáin ytri spá.

Fjöldi erlendra ferðamanna: Innri og ytri spár											
		Upphafsár innri spáa									
Ár	Ytri spá	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
2011	548.613	548.613	548.613	548.613	548.613	548.613	548.613	548.613	548.613	548.613	548.613
2012	634.709	641.124	634.709	634.709	634.709	634.709	634.709	634.709	634.709	634.709	634.709
2013	787.429	782.744	777.510	787.429	787.429	787.429	787.429	787.429	787.429	787.429	787.429
2014	1021.000	1022.396	1018.155	1026.162	1021.000	1021.000	1021.000	1021.000	1021.000	1021.000	1021.000
2015	1329.518	1367.808	1364.889	1370.381	1366.849	1329.518	1329.518	1329.518	1329.518	1329.518	1329.518
2016	1566.914	1605.984	1605.099	1606.756	1605.695	1593.505	1566.914	1566.914	1566.914	1566.914	1566.914
2017	2055.562	2033.202	2032.988	2033.388	2033.132	2030.090	2022.796	2055.562	2055.562	2055.562	2055.562
2018	2214.418	2261.162	2261.201	2261.127	2261.175	2261.731	2263.018	2256.706	2214.418	2214.418	2214.418
2019	2295.619	2320.336	2320.319	2320.350	2320.330	2320.098	2319.560	2322.179	2338.406	2295.619	2295.619
2020	464.519	47.978	48.002	47.957	47.986	48.324	49.109	45.286	21.393	83.666	464.519
2021	705.955	76.427	76.465	76.395	76.440	76.973	78.208	72.186	34.296	132.136	665.903
2022	1464.259	192.516	192.610	192.435	192.547	193.864	196.913	182.027	87.311	328.146	1426.942

ÚTGEFIÐ Í ÁGÚST 2022



Ferðamálastofa
Icelandic Tourist Board

Geirsgata 9 • 101 Reykjavík • Iceland • Hafnarstræti 91 • 600 Akureyri • Iceland
Sími/Tel +354 535 5500 • upplýsingar@ferdamalastofa.is

www.ferdamalastofa.is