FMS 2021-13
ISBN 978-9935-522-10-8

Ferðamálastofa
Icelandic Tourist Board

SPÁLÍKAN UM FERÐAÞJÓNUSTU
TÖLFRÆÐILEG GRUNDVALLARATRIÐI

MAÍ 2021

INTELLECON

Útgefandi: Ferðamálastofa - Geirsgötu 9, 101 Reykjavík / Hafnarstræti 91, 600 Akureyri
Netfang: upplysingar@ferdamalastofa.is
Veffang: www.ferdamalastofa.is
Titill: Spálíkan um ferðaþjónustu - Tölfræðileg grundvallaratriði
Númer: FMS 2021-13
ISBN: 978-9935-522-10-8

Efnisyfirlit

Inngangur

Hér er fjallað um tölfræðileg grundvallaratriði spáa um ferðamennsku og umsvifa í henni.

Umfjöllunin er í þremur hlutum. Fyrst er almenn umfjöllun um tölfræðilegar spár og eiginleika þeirra. Því næst er fjallað um ólík líkön, kosti þeirra, galla sem og helstu niðurstöður rannsókna. Í lokin eru dregnar fram helstu niðurstöður, lærdómar og ábendingar.

I. Tölfræðilegar spár og eiginleikar þeirra

Tölfræði hagspáa er afar vítt fræðasvið. Viðfangsefnið er tæknilega flókið og ekki unnt að komast mikið áfram með það nema með því að beita tiltölulega sérhæfðum tölfræðilegum og stærðfræðilegum aðferðum. Þessi greinargerð er ekki vettvangur fyrir slíka nálgun. Því verður látið nægja að drepa á nokkur almenn meginatriði og útskýra þau í tiltölulega einföldu samhengi.

Allar tölfræðilegar spár byggjast á forsendu um að samband sé á milli spástærðarinnar, þ.e. þeirrar stærðar sem spá skal og safns annarra stærða. Þetta samband má (í vissum tilfellum) rita sem fallið:

(1) $y=F(x)+u$,

þar sem y er spástærðin, x hinar stærðirnar og u tilviljunarkenndur slembiliður. Spástærðin, y, er oft kölluð háða breytan og x-in óháðar breytur eða útskýringarbreytur. Í kerfislægu samhengi er y talin vera innri (e. endogenous) breyta og x-in ytri breytur (e. exogeneous). Bæði y og x eru vektorar. Fallið $F(.)$ er jafnvíður vektor og y. Ef y hefur fleiri en eina vídd, er t.d. $(N\times1)$ vektor, lýsir (1) jöfnukerfi. Slembiliðurinn, u og líkindadreifing hans gegnir lykilhlutverki í tölfræðilegu mati á sambandinu milli x og y , þ.e. fallinu $F(.)$

Til að geta spáð fyrir um gildi y þarf tvö meginskref:

1. Meta sambandið milli y og x, þ.e. fallið $F(.)$. Þetta er *matsvandamálið*.
2. Framkvæma spána og finna tölfræðilega eiginleika hennar. Þetta er *spávandamálið*.

Þægilegt er að fjalla um þessi tvö vandamál sitt í hvoru lagi.

I.1 Matsvandamálið

Til að meta óþekkta sambandið $y=F(x)$ þarf upplýsingar (gögn) um breyturnar í sambandinu, þ.e. (y,x). Í reynd er það oft hægara sagt en gert. Því er gagnaöflun veigamikill þáttur í því að leysa matsvandamálið.

Að gefnum gögnum er sjálft matið í eðli sínu tölfræðilegt viðfangsefni. Það er vegna þess að við þekkjum ekki hið sanna samband og því er sérhvert mat á því háð óvissu og hefur líkindadreifingu. Þar við bætist að stærðirnar y og x eru yfirleitt ekki fullkomlega mældar og hafa því einnig líkindadreifingu. Í þriðja lagi er hugsanlegt að sambandið sjálft, þ.e. fallið $F(.)$ sé ekki klippt og skorið (e. determinate) heldur háð líkindadreifingu.

Innan tölfræðinar hafa margar aðferðir verið þróaðar til að meta samband eins og (1). Með vissum einföldunum má þó segja að allar þessar aðferðir leitist við að vera jafngildar aðferð hámarkslíkinda (e. method of maximum likelihood).

Aðferð hámarkslíkinda byggir á þeirri grunnforsendu að fallsambandið hljóti að vera þannig að það sem gerðist, þ.e. gögnin, hafi verið það sem var líklegast að gerast myndi. Augljóslega er erfitt að draga þetta í efa.

Gallinn er hins vegar sá að til þess að átta sig á hvað er líklegast þarf að þekkja viðkomandi líkindadreifingu. Aðferð hámarkslíkinda krefst þess því að rannsakandinn tilgreini þá líkindadreifingu sem framleiðir gögnin. Sú líkindadreifing er oft kölluð gagnasköpunarferlið (e. data generating process).

Svokallað líkindafall (e. likelihood function) er meginverkfærið í aðferð hámarkslíkinda.

Ritum þetta fall svona (Silvey 1975):

(2) $$P(y,x;\theta),$$

þar sem *y* og *x* eru gögnin og θ óþekktir stuðlar (e. parameters) í gagnasköpunarferlinu. Það má t.d. sjá θ fyrir sér sem stuðla í fallinu *F*(.) í líkingu (1) hér að ofan. Þannig felur val á θ í sér ákvörðun um fallsambandið *F*(.). Athugið að fallformið endurspeglast í líkindafallinu.

Aðferð hámarkslíkinda velur θ þannig að líkur á þeim gögnum sem fyrir liggja séu í hámarki. M.ö.o.:

(3) $$P(y,x;\hat{\theta}) = \max_{\theta} P(y,x;\theta),$$

þar sem $\hat{\theta}$ eru metnu gildin á stuðlinum θ.

Mat á óþekktum stuðlum líkans með aðferð hámarkslíkinda hefur marga æskilega tölfræðilega eiginleika, sérstaklega þegar gangaúrtakið stækkar. Mestu máli skiptir að þegar úrtakið stækkar þá verður matið það sem kallast samkvæmt (þ.e. óhneigt og með minnsta mögulega breytileika) og metnu gildin normaldreifð (Silvey 1975).

Þótt aðferð hámarkslíkinda, eins og henni hefur nú verið lýst, sé skýr og einföld getur hún verið mjög margbrotin í reynd. Í fyrsta lagi getur verið erfitt að skýrgreina líkindafallið $P(y,x;\theta)$ og það getur verið mjög flókið. Í öðru lagi er ekki víst að hámarkið sem leitað er að í (3) sé til. Í þriðja lagi getur verið mjög erfitt (þ.e. krefst mikilla og flókinna reikninga) að finna hámark þess jafnvel þótt það sé til.

Til að varpa ljósi á notkun aðferðar hámarkslíkinda getur verið gagnlegt að skoða einfalt dæmi.

Dæmi: Aðferð minnstu kvaðrata sem aðferð hámarkslíkinda

Ímyndum okkur að samhengi *y* og *x* í líkingu (1) megi lýsa með staðlaða aðfallslíkaninu:

$$y = X\cdot\beta + u,\ u \sim N(0,\sigma^2\cdot I),$$

þar sem *X* er ($N\times M$) fylki athugana fyrir *M* útskýringarbreytur og allir hinir vektorarnir eru samsvarandi. *u* er slembiliður með hina mjög svo þægilegu líkindadreifingu $N(0,\sigma^2\cdot I)$, þ.e. vongildi núll og hvorki sjálffylgni né misdreifni, sem kallast gjarnan normal-dreifður hvítur hávaði. Óþekktir stuðlar í þessu líkani eru β og σ.

Samkvæmt líkindadreifingu slembiliðarins er líkindafallið fyrir ofangreint líkan:

$$P(y,X;\beta,\sigma^2)=\frac{1}{(2\pi\sigma^2)^{N/2}}\cdot e^{\left(-\frac{1}{2\sigma^2}\cdot(y-X\beta)^T(y-X\beta)\right)}.$$

Lógaritminn (logrinn) af líkindafallinu er:

$$\ln P(y,X;\beta,\sigma^2)=-\frac{N}{2}\cdot\ln(2\pi)-\frac{N}{2}\cdot\ln(\sigma^2)-\frac{1}{2\sigma^2}\cdot(y-X\beta)^T(y-X\beta).$$

Hámörkun á $\ln P(y,X;\beta,\sigma^2)$ með tilliti til β og σ^2 leiðir til metlana (e. parameter estimate):

$$\hat{\beta}=(X^TX)^{-1}\cdot X^Ty,$$

$$\hat{\sigma}^2=\frac{(y-X\hat{\beta})^T(y-X)}{N},$$

Sem eru hinu velþekktu metlar venjulegrar aðferðar minnstu kvaðrata[1] (sjá t.d. Theil 1971). M.ö.o. aðferð hámarkslíkinda felur í sér tölfræðilega réttlætingu á aðferð minnstu kvaðrata.

Rétt er að taka það fram að fremur sjaldgæft er að unnt sé að leiða út líkingar fyrir hámarkslíkindametla og þegar það er hægt eru þeir oftast miklu flóknari en hér. Oftast verður mat samkvæmt aðferð hámarkslíkinda að byggjast á tölulegum aðferðum.

I.2 Spávandamálið

Gerum ráð fyrir að sambandið milli y og x, þ.e. fallið $F(.)$, hafi verið metið. Þar sem spár um umsvif í ferðaþjónustu munu vera fram í tímann munum við tiltaka þann tíma sem þessar stærðir verða á og rita hið metna samband sem:

(4) $y(t)=F^*(x(t))$,

þar sem F^* táknar að hér er um að ræða metna sambandið.

Gerum jafnframt ráð fyrir að hið sanna líkan og líkindadreifing slembiliðarins sé:

(4‘) $y(t)=F^*(x(t))+u(t)$, $u(t)\sim N(0,\sigma^2)$,

Þótt það sé ekki meginatriði á þessu stigi málsins er rétt að ítreka það að vektorinn x getur innihaldið bæði tafðar ytri sem og innri breytur (samheiti þeirra er fyrirfram ákveðnar breytur). Almennt má rita vektorinn $x(t)=(x(t), x(t\text{-}1),\ldots; y(t\text{-}1),y(t\text{-}2),\ldots)$. Sérhvert tiltekið líkan (samband) í (1) hefur vissa útgáfu af vektornum $x(t)$. Tímaraðalíkön myndu t.d. aðeins hafa tafðar innri breytur meðal útskýringarbreyta, o.s.frv.

---

[1] Að undanteknu því að VAMK metillinn fyrir σ^2 hefur N-M í nefnara sem skiptir máli í litlum úrtökum.

Gerum nú ráð fyrir að óskað sé eftir að spá y fram í tímann, þ.e. finna „gott" gildi fyrir $y(t+\Delta)$, þar sem $\Delta>0$. Samkvæmt (1) er það gildi:

(5) $$y(t+\Delta)=F(x(t+\Delta))+u(t+\Delta).$$

Þar með blasir við að til að spá $y(t+\Delta)$ þarf að þekkja útskýringarbreyturnar $x(t+\Delta)$ og slembiliðinn $u(t+\Delta)$. Þar sem hvort tveggja er í framtíðinni þarf einnig að spá þessum stærðum.[2] Þar sem slembiliðurinn, $u(t+\Delta)$, getur ekki verið þekktur og vongildi hans er núll er besta spáin:

(6) $$y(t+\Delta)=F(x(t+\Delta))$$

Spáin, $y(t+\Delta)$, er þannig fall af tilviljunarkenndum stærðum, þ.e. tölfræðilegu mati á fallsambandinu $F(\ )$ og væntanlegu óvissumati á útskýringarbreytunum. Hún er því sjálf tilviljunarkennd og hefur því líkindadreifingu og óvissubil.

Spáskekkjan er munurinn á spánni og raunverulega gildinu. Þennan mun getum við ritað sem:

(7) $$e(t+\Delta)=F(x(t+\Delta))+u(t+\Delta)-F(x(t+\Delta).$$

Eins og sjálf spáin er spáskekkjan tilviljunarkennd stærð og hefur líkindadreifingu og óvissubil. Mikilvægt er að átta sig á að líkindadreifing spárinnar og líkindadreifing spáskekkjunnar eru ekki eins.

Það getur verið erfitt að finna þessar líkindadreifingu og þar með óvissubilin. Í vissum tilvikum er þó unnt að finna stærðfræðilegar líkingar fyrir þessar stærðir. Í öðrum er nauðsynlegt að grípa til tölulegra lausna, t.d. með hjálp svokallaðra Monte Carlo hermana.[3]

Dæmi: Línuleg aðfallsjafna

Til að varpa skýrara ljósi á tölfræði hagspáa munum við nú skoða dæmi um línulega aðfallsjöfnu og mat á grundvelli venjulegrar aðferðar minnstu kvaðrata (VAMK).

Látum aðfallsjöfnuna vera

$$y=X\cdot\beta+u,\ u\sim N(0,\sigma^2\cdot I)$$

og VAMK metilinn fyrir β

$$\hat{\beta}=(X^TX)^{-1}\cdot X^Ty=\beta+(X^TX)^{-1}\cdot X^Tu\cdot$$

[2] Nema auðvitað í því sértilfelli þegar allar útskýringarbreyturnar eru nægilega tafðar til þess að $x(t+\Delta)$ hafi þegar gerst.

[3] Sjá t.d. Davidson og McKinnon (1993)

Íhugum nú að spá gildum fyrir *y* þegar útskýringarbreyturnar eru *X**.[4] Samkvæmt líkaninu myndi þá *y* vera:

$$y^* = X^* \cdot \hat{\beta} + u^* = X^* \cdot \left(\beta + (X^T X)^{-1} \cdot X^T u\right) + u^* \cdot$$

Ef $E(X^*)=X$, þar sem $E(.)$ táknar vongildi af svigastærðinni og *X** er óháð slembiliðnum *u** er greinilegt að vongildi spárinnar er vongildi sönnu gildanna. Nánar tiltekið:

$$E(y^*) = E\left(X^* \cdot \left(\beta + (X^T X)^{-1} \cdot X^T u\right) + u^*\right) = X\beta + E(X^*) \cdot (X^T X)^{-1} \cdot X^T \cdot E(u) + E(u^*) = E(y)$$

Breytileikafylki (e. variance-covariance matrix) spárinnar er:

$$Var(y^*) = \sigma^2 \cdot X^* (X^T X)^{-1} X^{*T}.$$

Spáskekkjan er hins vegar

$$e = y - y^* = X^* \cdot \beta + u - X^* \hat{\beta} = X^* \cdot (\beta - \hat{\beta}) + u$$

Augljóst er að vongildi spáskekkjunnar er núll, þ.e. $E(e)=0$.

Breytileikafylkið fyrir spáskekkjuna er hins vegar:

$$\begin{aligned} & E(X^* \cdot (\beta - \hat{\beta}) + u^*) \cdot (X^* \cdot (\beta - \hat{\beta}) + u^*)^T \\ (8) \qquad & = E(X^* \cdot (X^T X)^{-1} X^T u + u^*) \cdot (X^* \cdot (X^T X)^{-1} X^T u + u^*)^T \\ & = \sigma^2 \cdot X^* \cdot (X^T X)^{-1} X^{*T} + \sigma^2 \cdot I \end{aligned}$$

þar sem *u* og *u** eru óháðir slembiliðir.

Þannig sjáum við að breytileikinn í spáskekkjunni er summa af tveimur liðum. Sá fyrri er vegna skekkjunnar í matinu á líkaninu. Sá síðari er vegna tilviljunarkenndra frávika frá líkaninu á spátímanum.

Af (8) er unnt að draga þá ályktun að spáskekkjan vex eftir því sem ytri stærðirnar *X** víkja lengra frá gögnunum, *X*, sem notuð voru til að meta sambandið.[5]

[4] Takið eftir að almennt er verið er að spá mörgum gildum fyrir *y*.

[5] Sjá t.d. Theil (1971).

II. Hin ýmsu líkön

Munum við nú gera stuttlega grein fyrir hinum ýmsu ólíku líkönum. Ólík líkön má flokka á ýmsan hátt. Gróf skipting væri t.d. annars vegar í megindleg líkön (e. quantitative) og hins vegar eigindleg (e. quantitative) líkön. Síðan má flokka ólík líkön í undirflokka og jafnvel líkön þar sem eigindlegar og megindlegar aðferðir skarast. Hér er farin sú leið að fylgja Song og fél. (2019) og flokka líkönin í fjóra flokka, þ.e. tímaraðalíkön (e. time series models), hagrannsóknarleg líkön (e. econometrics), gervigreindarlíkön (e. artificial intelligence) og upplýst skoðunarlíkön/dómgreindarlíkön (e. judgemental models).[6]

Mörg líkön skarast milli ofangreindra flokka.

II.1 Tímaraðalíkön (e. time series models)

Tímaraðalíkön byggja á þeirri hugmynd að hægt sé að láta þróun í fortíð segja til um gildi í framtíð. Þannig er reynt að finna leitni og annað mynstur í tímaröðinni.

Tímaröð (e. time series) er mæling á breytu yfir tíma, t.d. ef við höfum *n* margar mælingar á stærðinni

$$y_1, y_2, y_3, \dots, y_t, y_{t+1}, y_{t+2}, \dots y_{n-1}, y_n$$

Tölfræðilega er slík runa slembiferli (e. stochastic process).

Tímaraðalíkön geta verið allt frá því að vera mjög einföld í það að vera mjög flókin.

II.2 Barnalega líkanið (e. naive model)

Einfaldasta hreina tímaraðalíkanið byggir á því að spá síðasta gildi. Þetta líkan hefur verið kalla barnalega (e. naive) líkanið, sökum þess hve einfalt það er.

Barnalega líkanið má setja fram þannig,

$$y_t^{\,s} = y_{t-1}$$

þar sem $y_t^{\,s}$ stendur fyrir spáð gildi á y.

Þrátt fyrir einfaldleikann getur það komið vel út í samanburði við önnur líkön, sérstaklega þegar verið er að spá til mjög skammst tíma, enda best að spá óbreyttu ástandi ef óvissa um framtíðina er alger (e. random walk).[7]

---

[6] Þessi skipting er ekki ólík því sem aðrir aðilar nota s.s. ICAO (2006), Jiao og Chen (2019) og European Travel Commission (2021).
[7] Sjá Martin og Witt (1989).

II.3 Einföld sjálffylgnilíkön (e. autoregressive models /AR)

Sjálffylgnilíkön byggja á því að spá innri breytunni með því að notast við línulega samsetningu fyrri gilda þeirrar sömu breytu. Nafngift aðferðarinnar vísar til þess að breytan fylgi sjálfri sér.

Sjálffylgnilíkan af fyrstu gráðu má rita þannig,

$$y_t = \alpha\, y_{t-1} + \varepsilon_t$$

þar sem gildi nú ræðst af gildi síðasta tímabils að viðbættum villulið.

Annað einfalt tímaraðalíkan er svokallað *hlaupandi meðaltalslíkan* (e. moving average/MA) en einfalda útgáfu af því mætti rita þannig,

$$y_t = \frac{y_{t-1} + y_{t-2} + y_{t-3} + \cdots + y_{t-m}}{m} + \varepsilon_t$$

þ.e. að gildi nú ræðst af meðaltali ákveðið margra gilda sem á undan koma (hér m-margra).

Síðan má hugsa sér ýmsar fleiri útgáfur og sambland ólíkra líkana.

II.4 Flóknari tímaraðalíkön

Flóknari tímaraðalíkön taka tillit til þess að oft er um að ræða einhvers konar kerfislæga þætti í gögnunum, t.d. langtímaleitni og árstíðasveiflur. Slík atriði koma einkum fram í villuliðum hinna tölfræðilegu líkana.

Algengustu líkön af þessum meiði eru svokölluð ARIMA (e. autoregressive-integrated moving-average) líkön. Slík líkön eru leidd út frá einföldu AR og MA líkönunum sem minnst var á hér að ofan.

Hér gefst ekki tóm til að fjalla ítarlega um slík líkön en rétt er að gera nokkra grein fyrir hugmyndunum sem að baki liggja, enda eru slík líkön mjög algeng og vinsæl sem spálíkön.

Villuliðurinn og greining á honum er grundvallaratriði í tímaraðagreiningu almennt. Ströng skilyrði þurfa að gilda um villuliðinn til að forðast ýmsa pytti í greiningunni sem geta leitt til villandi og jafnvel rangra niðurstaðna. Mikilvægasta skilyrðið er að tímaröðin sé það sem kallast sístæð (e. stationary) en almennt má segja að það þýði að að meðaltal, variáns og kóvaríans afgangsliða breytast ekki yfir tíma. Ástæðan er sú að forsendur venjulegrar aðferðar minnstu kvaðrata (e. ordinary least squares) halda ekki lengur ef röð er ósístæð. Þegar árstíðarsveiflur eða leitni er í tímaröðinni þá er hún ekki sístæð. Því er algengt í tímaraðagreiningu að kanna fyrst hvort röðin eða raðirnar sem verið er að skoða séu sístæðar eða ekki. Séu þær ekki sístæðar eru til aðferðir til að gera þær sístæðar.[8]

[8] Sjá t.d. Hamilton (1994), Judge og fél. 1988) eða Charemza og Deadman (1992).

Ef við blöndum saman einföldu AR og MA líkani kallast það ARMA líkan og rita einfalda útgáfu þess þannig[9],

$$y_t = \theta\, y_{t-1} + \varepsilon_t + \alpha\, \varepsilon_{t-1}$$

Hér ræðst síðasta gildi breytunnar sem við erum að skoða af gildi sínu á tímabilinu á undan að viðbættum villulið, en einnig af villulið tímabilsins þar áður. Þessi einfaldasta útgáfa kallast ARMA(1,1), því breytan skýrist af einu töfðu gildi af sjálfri sér og af einu töflu gildi af afgangsliðnum. Hægt er taka tillit til fleiri tafa og almennt má rita,

$$\begin{aligned} y_t = {} & \theta_1\, y_{t-1} + \theta_2\, y_{t-2} + \cdots + \theta_p\, y_{t-p} \\ & + \varepsilon_t + \alpha_1\, \varepsilon_{t-1} + \cdots + \alpha_q \varepsilon_{t-q} \end{aligned}$$

sem myndi þá kallast ARMA(p,q) þar sem síðasta gildi breytunnar ræðst af p-töfðum gildum af sjálfri sér og q-töfðum gildum af villuliðnum.

Ef það er leitni í gögnunum, sem er oft tilfellið, er tímaröðin ósístæð og þá þarf að leita leiða til að gera hana sístæða. Einföld og áhrifarík leið sem oft virkar er að taka mismun af röðinni sjálfri, þ.e. búa til nýja röð þannig,

$$x_t = y_t - y_{t-1}$$

Ef þessi nýja röð, x_t, er ARMA(p,q) þá kallast upphaflega röðin y_t ARIMA(p,1,q), með öðrum orðum þá þurfti *einn* mismun til að gera hana sístæða. Ef þarf að taka fleiri en einn mismun til að gera upphaflegu röðina sístæða, þá kallast hún ARIMA(p,d,q) þar sem d er fjöldi tafa sem nauðsynleg er til að gera hana síststæða.

II.5 Samþætting og skekkju-leiðrétting (e. Cointegration and error-correction)

Til þess að fá áreiðanlegar, þ.e. tölfræðilega samkvæmar (e. consistent) niðurstöður um aðfallsgreiningu með ósístæðum tímaröðum er nauðsynlegt að taka mismun af röðunum fyrst til þess að gera þær sístæðar eins og áður hefur verið nefnt. Séu raðirnar samþættar af 1° nægir að taka fyrsta mismun o.s.frv. Með dálitlum, nánast sjálfsögðum viðbótum, leiðir þetta til svokallaðra "error-correction" líkana (skekkju-leiðréttandi líkön).

Hugmyndin að baki skekkju-leiðréttandi líkönum er einfaldri mynd sem hér segir:

Ímyndum okkar að

$$y = a \cdot x\,.$$

Ef y og x eru samþættar af 1° mun aðfallsgeining ekki meta a tryggilega. Því er tekinn fyrsti mismunur[10]:

$$\Delta y = a\Delta x\,.$$

[9] Sjá t.d. Judge og fél. (1988).
[10] Raunar er yfirleitt tekinnlogri (lógaritmi) af breytunum áður, en það breytir ekki meginhugsuninni.

Gallinn við þessa líkingu er að hún flytur engar upplýsingar um jafnvægisvenslin $y = a \cdot x$. Því er þeim einfaldlega bætt við og líkanið ritað:

$$\Delta y = a \cdot \Delta x + b \cdot (y - a \cdot x),$$

þar sem gert er ráð fyrir að $b<0$, þannig að sé $y>a \cdot x$ hefur Δx tilhneiginu til að vera neikvætt og öfugt.

Líkingin hér að ofan er einföld útgáfa af skekkjuleiðréttandi líkani. Flóknari líkön eru þróuð með sama hætti. Einn af kostunum við skekkju-leiðréttandi líkön er að þau hafa bæði jafnvægislausn og aðlögunarferla.

Réttlæting þess að nota skekkjuleiðréttandi líkan til að lýsa samhengi milli samþættra raða er sett fram í svokallaðri *framsetningarsetningu Grangers*.[11] Samkvæmt setningunni er unnt að lýsa samhengi allra samþættra tímaraða með skekkjuleiðréttandi líkani og samþætting er nauðsynleg til þess að unnt sé að lýsa viðkomandi tímaröðum þannig.

Þróuð hefur verið talsvert viðamikil aðferðafræði til að kanna hvort skilyrðum framsetningarsetningar Grangers sé fullnægt.[12]

Augljóslega kemur til álita að nota skekkjuleiðréttingarlíkön til að útskýra ferðamennsku. Sennilegt virðist að viðkomandi tímaraðir, umfang ferðaþjónustu, landsframleiðsla, verð (þ.e. nafnvirði) og fjármunir í greininni hafi vaxandi leitni (þ.e. séu samþættar af 1°). Vandkvæðin kunna hins vegar að vera þau að hinar árlegu raðir eru nokkuð stuttar fyrir hefðbundna samþættingargreiningu. Ársfjórðungsgögnin mynda lengri raðir en veruleg árstíðasveifla er í þeim gögnum sem flækja málið nokkuð.

Slík líkön sem hér hefur verið lýst hafa sýnt sig að virka sæmilega til að spá um ferðamennsku. Má hér nefna líkan Gounopoulos og fél. (2012) er spáir fjölda ferðamanna til Grikklands, Lim og McAleer (2002) um komur ferðamanna til Ástralíu, Chu (2004) varðandi komur til Singapore, Du Preez og Witt um komur ferðamanna frá fjórum Evrópulöndum til Seychelle eyja og Kunendra og Witt (2001) þar sem borin voru saman líkön til að spá fyrir um ferðamennsku í fjórum Evrópulöndum og Bandaríkjunum.

II.6 Hagrannsóknarleg líkön (e. econometric models)

Hagrannsóknarleg líkön byggja á þeirri hugmynd að til séu sambönd orsaka og afleiðinga og að hægt sé að finna og nýta slík sambönd til að spá fyrir um framtíðina. Yfirleitt er gengið út frá því í byrjun að eitthvert orsakasamband sé til staðar og oftar en ekki stuðst við hagfræðilegar kenningar til að rökstyðja slík sambönd.[13] Síðan er beitt tölfræðilegum aðferðum til að meta umfang og eðli þessa sambands eða sambanda.

Þær tölfræðilegu aðferðir sem beitt er byggja yfirleitt á aðhvarfsgreiningum (e. regression). Yfirleitt þarf að aðlaga aðhvarsgreiningarnar að gögnunum og gera kröfu um sístæðni (e. stationarity) því annars er hætta á að tölfræðileg próf gefi misvísandi niðurstöður eins og áður

[11] Eagle og Granger (1987).
[12] Sjá t.d. Charemza og Deadman (1993) og Johansen (2006).
[13] Samböndin geta verið jafnt línuleg sem ólínuleg. Sjá t.d. Chen (2011).

hefur verið rætt. Það sem einkennir jafnan tölur um ferðaþjónustu eru árstíðarsveiflur.[14] Ástæðurnar eru vel þekktar enda ferðamannatímabil oft tengd jafnt árstíðarsveiflum í veðurfari, frítíma, (þ.m.t. sérstökum frídögum). Slíkar árstíðarsveiflur leiða alla jafna til ósístæðni sem leiðrétta þarf fyrir.[15]

Það er erfitt að segja fyrirfram hvaða breytur sé best að nota til að spá fyrir um framtíðargildi innri breytunnar. Fræðilegar niðurstöður um hvaða þættir það eru sem skipta máli, t.d. varðandi fjölda ferðamanna, geta veitt vísbendingar.[16] Tölfræðilega má prófa hvaða breytur auka spágetu líkansins, séu þær notaðar. Hér er þó að ýmsu að hyggja. Þannig geta ákveðnar skýribreytur reynst vel á ákveðnum tímum en ekki öðrum. Ástæður þessa geta m.a. verið þær að samspil skýribreyta og innri breytunnar breytist yfir tíma. Einnig getur verið að gögnin séu ekki nægilega góð, þ.e. að þau mæli ekki endilega það sem þeim er ætlað að mæla.

Oft og tíðum er blandað saman tímaraðagreiningum og hagrannsóknarlegum líkönum. Þar hefur gefist einna best að nýta hagrannsóknarleg sjálffylgnilíkön, þ.e. notast við sjálffylgnilíkön eins og lýst var hér að ofan, en bæta við fleiri skýristærðum. Fjölmargar rannsóknir af þessum meiði er að finna og er ekki að furða því með þeim má nýta spágetu tímaraðagreininga á sama tíma og hægt er að meta áhrif þekktra ytri breyta. Tiltölulega nýlegt dæmi um slíkt er rannsókn Li og fél. (2018) þar sem sjónum er beint að áhrifum breytinga í veðurfari á eftirspurn í Kína, sem og rannsókn González og Moral (1995) sem rannsaka eftirspurn eftir ferðum til Spánar og taka sérstakt tillit til breytinga í smekk og löngunum ferðamanna.

II.7 Líkön með breytilega stika (e. time varying parameter models)

Á síðustu áratugum hafa verið þróuð líkön þar sem stikar (e. parametrar) þeirra hafa ekki verið fastar, heldur geta breyst yfir tíma.

Einfalt dæmi um slíkt líkan væri:[17]

$$\Delta y_{it} = \Delta x'_{it}\beta_{it} + \epsilon_{it}$$

$$\beta_{it} = \beta_{it-1} + \nu_{it}$$

þar sem Δy_{it} er háða breytan, $\Delta x'_{it}$ er vektor skýribreyta, β_{it} er vektor stika sem geta breyst yfir tíma og ϵ_{it} og ν_{it} eru vektorar óháðra og jafndreifðra slembifrávika með meðaltal núll og fastan varíans. Slík líkön eru ekki síst áhugaverð þegar verið er að skoða langtímaþróun yfir tíma og ætla má að hagsambönd stærða breytist frá einum tíma til annars, t.d. vegna breytinga í framleiðsluháttum, neysluhegðun, o.s.frv.

[14] Sjá Chung (2009).
[15] Að fást við vanda ósístæðni er eitt helsta viðfangsefni spágerðar þar sem notast er við tímaraðir. Sjá t.d. Charemza og Deadman (1992), Hassler og Wolters (2005), Kulendran og King (1997) og Kulendran, og Witt (2001).
[16] Sjá D2 Hagfræðileg grundvallaratriði.
[17] Sjá Fildes og fél. (2011) en mun ítarlegri umræðu má t.a.m. finna hjá Judge og fél. (1985) og Tucci (1995).

II.8 Líkanakerfi – Jöfnukerfi

Þá má líka nefna að í stað þess að meta bara eina spájöfnu má einnig meta jöfnukerfi, þ.e. fleiri en eina jöfnu í einu. Er þar oft byggt á kenningum Deaton og Mullbauer (1980) um nánast fullkomið jöfnukerfi eftirspurnar (e. almost ideal demand system). Mjög gróflega má segja að þessar aðferðir geri mögulegt að meta eftirspurn eftir ákveðnum vörum eða þjónustu, mælt sem hlutfall þeirra í viðkomandi hagkerfi. Slík líkön má líka flækja með því að gera þau kvik (e. dynamic). Sem dæmi má nefna rannsóknir Li og fél. (2004, 2006) þar sem slík líkön eru notuð til að spá fyrir um eftirspurn breskra ferðamanna eftir áfangastaði á meginlandi Evrópu.

II.9 Gervigreindarlíkön (e. artificial intelligence models)

Gervigreindarlíkön nota reiknifræðilegar og tölfræðilegar aðferðir til að meta ólínuleg sambönd stærða án þess að fyrir liggi hugmyndir um samband stærðanna. Algengustu gervigreindarlíkön sem notuð hafa verið til að spá fyrir um fjölda ferðamanna eru svokölluð tauganet (e. neural networks).[18] Dæmi um rannsókn þar sem slíkum aðferðum hefur verið beitt er hjá Law og Au (1999) sem skoðuðu eftirspurn eftir ferðum Japana til Hong Kong og Chen og fél. (2012) sem rannsökuðu eftirpurn ferða til Tævan, frá Hong Kong, Makaó og Japan. Þá notuðu Tsaur og Kuo (2011) tauganetalíkan til að meta eftirspurn eftir ferðamennsku í Tævan.

Gervigreindarlíkön hafa verið gagnrýnd fyrir að erfitt sé að átta sig á hvernig líkönin virka og þannig tengja gögnin við niðurstöður. Þá krefjast slík líkön yfirleitt mikilla gagna. Með tilkomu Internetsins hefur aðgangur að gríðarlegu magni upplýsinga (e. big data) aukist og um leið opnast nýir möguleikar við líkanagerð af þessu tagi. Á síðustu árum hefur einmitt færst í vöxt að nota það sem kalla mætti óhefðbundin gögn til að spá, t.d. fjölda leitarorða í leitarvélum á Internetinu eða reikigögn farsímafyrirtækja.[19]

Law og fél. (2019) nota gervigreindarlíkön og ýmis konar gögn, þ.á.m. um fjölda leitaðra orða víða að á Internetinu, til að spá eftirspurn eftir ferðamennsku í Macao. Þeir telja niðurstöður rannsóknarinnar lofa góðu þótt ýmislegt megi enn bæta bæði hvað varðar aðferðirnar og nýtingu gagna.

II.10 Kannanalíkön/dómgreindarlíkön (e. judgemental methods)

Það hefur lengi tíðkast að safna saman upplýsingum frá ýmsum aðilum og þá oftast frá þeim sem til þekkja í greininni, með skipulögðum hætti og búa þannig til spá um framtíðina.

Algengasta aðferðin af þessum toga er svonefnt Delphi líkan. Almennt má segja að aðferðin byggist á því að nota spurningakönnun til að fá fram skoðanir sérfræðinga og haghafa á framtíðarhorfum. Þessum skoðunum eða viðhorfum er safnað saman, þau metin og dregnar af þeim ályktanir. Þá er aftur haft samband við umrædda aðila, þeim kynntar niðurstöður og/eða sviðsmyndir og þeir beðnir um að endurmeta spá sína. Þannig má fara nokkra hringi með að uppfæra spárnar þar til endanleg spá liggur fyrir.

[18] Almenna umfjöllun um tauganet má t.d. finna hjá Kosko (1992).
[19] Sjá t.d. Bangway-Skeete og fél. (2015), Choi ogVarian (2012), Yang og fél. (2015) og Dergiades og fél. (2018).

Aðferðir sem þessar hafa verið gagnrýndar af ýmsum ástæðum og snýr mikilvægasta gagnrýnin að því að erfitt getur verið að meta vissu eða óvissu niðurstaðna. Oft og tíðum er slíkum aðferðum þó beitt samfara öðrum þeim aðferðum sem hér hafa verið ræddar. Þá hefur verið töluvert um að megindlegum aðferðum og skoðunarlíkönum sé steypt saman. Þannig telja Song, Gao og Lin (2013) að bæta megi spágetu tímaragreiningarlíkana með því að bera nðurstöður þeirra undir sérfræðinga og aðlaga þau að þeim.

III. Umræða

Það er ómögulegt að segja hvaða líkan eða tölfræðilega aðferð sé ávallt best til að spá. Rannsóknir sýna, svo ekki verður um villst, að spágeta ólíkra líkana er mismunandi eftir aðstæðum og tímabilum. Einnig getur verið að ákveðið líkan spái betur en önnur fyrir hreina heppni. Ef nógu oft er spáð er líklegt að í einhverri tilrauninni rambi einhver spáin á að spá nokkurn veginn rétt og það fyrir hreina tilviljun. Þá getur verið erfitt að meta hvaða líkön séu best, þar sem hægt er að draga í efa hagnýtt gildi birtra spáa. Kemur það ekki síst til af því að áhersla þeirra er oftar en ekki á fræðilega æskilega eiginleika spánna frekar en hagnýtt gildi niðurstaðnanna.[20] Þá er langt í frá sjálfgefið að flóknari líkön spái betur en einfaldari og jafnvel einföldustu líkön.[21]

Tölfræðilegar spár eru vélrænn framreikningur og niðurstöðurnar byggja á gögnum og þeim forsendum sem gefnar eru. Breytist þær forsendur að verulegu leyti eru allar líkur til að skynsamlegt sé að breyta spánni til samræmis. Af þessu leiðir að tölfræðilegar spár eru yfirleitt ekki þess burðugar að spá fyrir um meiriháttar breytingar í framtíðinni.

Í viðamikilli rannsókn Fildes og fél. (2011) voru borin saman fjölmörg tölfræðileg líkön sem byggðu öll á tímaraðagreiningu með sjálffylgni. Hvað varðar þessi líkön virtist vera sem það skipti jafnvel meira máli að velja réttu skýristærðirnar frekar en hver er nákvæm útfærsla (e. specification) líkananna sjálfra. Þá virðist sem líkön þar sem stikarnir geta breyst yfir tíma komi ekki betur út en önnur.

Varðandi það hve margar tímatafir er best að nota, þá eru til tölfræðilegar aðferðir til að ákveða slíkt á sama hátt og lýst hefur verið hér að ofan. Meðal breyta sem rannsóknir hafa sýnt að geta skipt máli eru tekjur ferðamanna, hlutfallslegt verðlag á ákvörðunarstað (miðað við heimaland) og verð á helstu neysluvörum og þjónustu ferðamanna á stöðum í samkeppni (Li og fél, 2005, Song og Li, 2008). Aðrir þættir sem nefndir hafa verið er stjórnmálalegur stöðugleiki (Saha og Yap, 2014), einstakir viðburðir s.s. sjúkdómsfaraldrar (Page og fél. 2011), hryðjuverk (Bonham og fél, 2006) og fjármálakreppur (Son og fél., 2011).

Hvað varðar áföll í ferðaiðnaði þá sýna rannsóknir að einstakir stóratburðir, s.s. hryðjuverkaárásirnar 11. september 2001 höfðu einungis tímabundin áhrif og að ekki er nauðsynlegt að taka tillit til slíka atburða þegar verið er að spá til lengri tíma (Lai og Lu, 2005, Njegovan, 2006). Þó má nefna að sumir fræðimenn telja að innanlandsflug í Bandaríkjunum hafi tekið töluverðan tíma að ná fyrra stigi hvað varðar farþegafjölda í kjölfar þessa viðburðar (Blunk, Clark og McGibany, 2006). Nú á tímum kórónaveirufaraldurs er áhugavert að skðoa niðurstöður slíkra rannsókna og hvort og hvernig heimfæra megi þær upp á núverandi ástand.

Þá má benda á að flestar birtar spár fjalla um horfur í fjölda ferðamanna almennt, til ákveðinna landa. Það eru ekki margar spár sem fjalla um innanlandsmarkaði eða innanlandsflug en dæmi um slíkt eru þó Athanasopulos og Hyndman (2008), Blunk og fél. (2006). Dæmi um spár varðandi eftirspurn eftir gistingu eru Pan og fél. (2012), Yang og fél. (2014) og Rivera (2016). Þá hafa Chen og fél. (2003) og Ellis og Doren (1965) reynt að spá fyrir um eftirspurn eftir sérstökum ferðamannastöðum frekar en löndum.

[20] Sjá umfjöllun hjá Gunther og fél. (2019).
[21] Sjá t.d. Witt og Witt (1995).

Nýrri aðferðir sem notast við gervigreind og tauganet eru áhugaverðar en bæði þarf að þróa þær frekar auk þess sem aðferðirnar sjálfar bjóða ekki uppá auðvelda túlkun niðurstaðna varðandi það hvað veldur framtíðarþróun, sem er oft mikilvægur hluti spágerðarinnar.

Hvaða eða hvers konar líkön eru best ræðst ekki síst af því til hvers nota á spána. Þar sem hér er ætlunin að sem flestir aðilar geti nýtt spárnar við stefnumótun og fjárfestingarákvarðanir virðist á þessum tímapunkti fýsilegast að beita tímaraðagreiningum með ytri skýristærðum. Slík líkön hafa gefist vel, þótt þau séu ekki gallalaus frekar en önnur. Sú leið hefur t.d. verið farin í Nýja Sjálandi, þar sem ráðuneyti ferðamála (Ministry of Business, Innovation and Employment) hefur skilgreint stefnu í ferðamálum með mörgum markmiðum. Sem hluti af framfylgd stefnunnar hefur verið búin til ákveðin umgjörð fyrir spár í ferðaþjónustu. Notast er við tímaraðagreiningu til að spá fyrir um fjölda ferðamanna frá lykilmörkuðum og löndum. Við gerð spálíkananna er tekið tillit til athugasemda og ábendinga nefndar í hverri sitja sérfræðingar og hagaðilar.[22]

Mikilvægt er að öflug umgjörð sé í kringum spágerðina, ekki einungis til að tryggja gæði spánna sjálfra, heldur einnig að spárnar séu yfirfarnar og gagnrýndar svo hægt sé að bæta þær yfir tíma.

Af því sem hér hefur verið ritað virðist vænlegt að notast við hagfræðilegar tímaraðagreiningar þar sem tekið er tillit til ytri breyta svo hægt sé að einangra áhrif stjórnstærða og aðgerða. Slíkar aðferðir eru vel þekktar og hafa gefist sæmilega til spágerðar. Auk þess gefa slíkar aðferðir færi á að meta áhrif ólíkra aðgerða, t.a.m. stjórnvalda eða ytri og innri áfalla eða búhnykkja. Slíka líkanagerð má styrkja með ítrun á verklagi, sem og með skipulagðri ytri rýni.

[22] New Zealand Government (2019) og Deloitte (2018).

Heimildir

Assimakopoulos, V. og Nikolopoulos, K. (2000). The theta-model: a decomposition approach to forecasting. *International Journal of Forecasting*, Vol. 16, 521-530.

Athanasopoulos, G. og Hyndman, R.J. (2008). Modelling and forecasting Australian domestic tourism. *Tourism Management*, 29, 19-31.

Athanasopoulos, G., Hyndman, R.J., Song, H., og Wu, D.C. (2010). The tourism forecasting competition. www.robjhyndman.com

Bangway-Skeete, P.F., og Skeete, R.W. (2015). Can Google data improve the forecasting performance of tourist arrivals? Mixed-data sampling approach. *Tourism Management*, vol. 46, 454-464.

Blunk, S. S., Clark, D. E. og McGibany, J. M. (2006). Evaluating the long-run impacts of the 9/11 terrorist attacks on US domestic airline travel. *Applied Economics*, 38, 363–370.

Charemza, W.W. og Deadman, D.F. (1992). *New Directions in Econometric Practice*. Edwin Elgar, London.

Chen, K-Y. (2011). Combining linear and nonlinear model in forecasting tourism demand. *Expert Systems with Applications*, Vol. 38, 10368-10376.

Chen, R.J., Bloomfield, P. og Fu, J.S. (2003). An evaluation of alternative forecasting methods to recreation visitation. *Journal of Leisure Research*, 35(4), 441-454.

Chen, C-F., Lai, M-C., og Yeh, C-C. (2012). Forecasting tourism demand based on empirical mode decomposition and neural network. *Knowledge-Based Systems*, Vol. 26, 281-287.

Choi, H. og Varian, H. (2012). Predicting the Present with Google Trends. *Economic Record*, Vol. 8 (Special Issue, June), 2-9.

Chu, F-L. (2003). Forecasting tourism demand. a cubic polynomial approach. *Tourism Management*, Vol. 25, 209-218.

Chung, J.Y. (2009). Seasonality in Tourism: A Review. *e-Review of Tourism Research (eRTR)*, Vol. 7, No. 5.

Davidson, R og J. MacKinnon. 1993. *Estimation and Inference in Econometrics*. Oxford University Press. Oxford.

Deaton, A. og Muellbauer, J. (1980). An almost ideal demand system. *The American Economic Review*, 70(3), 312-326.

Deloitte (2018). *2018 travel and hospitality industry outlook*. Deloitte,

Dergiades, T., Mavragani, E. og Pan, B. (2018). Google Trends and tourists‘ arrivals: Emerging biases and proposed corrections. *Tourism Management*, Vol. 66, 108-120.

Du Preez, J. og Witt, S.F. (2003). Univariate versus multivariate time series forecasting: an application to international tourism demand. *International Journal of Forecasting*, 19, 435-451.

Eagle, R.F. og Granger, C.W.J. (1987). Co-integration and error correction: Representation, estimation and testing, *Econometrica*, 55(2), 251-276.

European Travel Commission (2021). *Handbook on Tourism Forecasting Methodologies*. A report produced for the European Travel Commission by Inzights & Silverbullet Research, Brussels (February).

Fildes, R., Wei, Y, og Ismail, S. (2009). Evaluating the forecasting performance of econometric models of air passenger traffic flows using multiple error measures. *International Journal of Forecasting*, Vol. 27), 902-922.

González, P., og Moral, P. (1995). An analysis of the international tourism demand in Spain. *International Journal of Forecasting*, Vol. 11, Issue, 2 (June), 233-251.

Gounopoulos, D., Petmezas, D. og Santamaria, D. (2011). Forecasting tourist arrivals in Greece and the impact of macroeconomic shocks from the countries of Tourists' origin. *Annals of Tourism Research*, vol. 39, No.2, 641-666.

Gunther, U., Önder, I. og Smeral, E. (2019). Scientific value of econometric tourism demand studies. *Annals of Tourism Research*, Vol. 78, 102738.

Gurthrie, H.W. (1960). Demand for Tourists' Goods and Services in a World Market. *Cowles Foundation Discussion Paper No. 93*. Cowles Foundation for Research in Economics. Yale University, New Haven.

Hamilton, J.D. (1994). *Time Series Analysis*. Princeton University Press, Princeton, NJ.

Hassler, U. og Wolters, J. (2005). Autoregressive Distributed Lag Models and Cointegration. *Diskussionsbeitrage, No2005/22*, Freie Universitat Berlin, Facbbereich Wirtschaftswissenschaft, Berlin.

ICAO (2006). *Manual on Air Traffic Forecasting*. Doc 8991 AT/722/3. International Civil Aviation Organization.

Intervistas (2013). Transport Canada Forecasts. Kynning 10. apríl 2013 á Canadian Forecasting Workshop.
http://www.intervistas.com/downloads/presentations/InterVISTAS%20Airport%20Traffic%20Forecasting%20Workshop%20-%20Presentations,%2010April2013.pdf

Judge, G.G., Hill, R.C., Griffiths, W.E., Lutkepohl, M. og Lee, T-C. (1988). *Introduction to the Theory and Practice of Econometrics* (2. útg.). John Wiley and Sons.

Jiao, E.X., og Chen, J.L. (2019). Tourism forecasting: A review of methodological developments over the last decade. *Tourism Economics*, Vol. 25(3), 469-492.

Kosko, B. (1992). *Neural Networks and Fuzzy Systems*. Prentice-Hall, Englewood Cliffs, NJ.

Kulendran, N. og King, M.L. (1997). Forecasting international quarterly tourist flows using error-correction and time-series models. *International Journal of Forecasting*, Vol. 13, 319-327.

Kulendran, N. og Witt, S.F. (2001). Cointegration versus least squares regression. *Annals of Tourism Research*, Vol. 28, No. 2, 291-311.

Lai, S. L., og Lu, W. L. (2005). Impact analysis of September 11 on air travel demand in the USA. *Journal of Air Transport Management*, 11(6), 455-458.

Law, R. og Au, N. (1999). A neural network model to forecast Japanese demand for travel to Hong Kong. *Tourism Management*, Vol. 20, 89-97.

Law, R., Li, G., Fong, D.K.C. og Han, X. (2019). Tourism demand forecasting: A deep learning approach. *Annals of Tourism Research*, Vol 75, 410-423.

Li, H., Goh, C., Hung, K. og Chen, J.L. (2018). Relative Climate Index and Its Effect on Seasonal Tourism Demand. *Journal of Travel Research*, vol 57(2), 178-192.

Li, G., Song, H. og Witt, S.F. (2004). Modeling tourism demand: A dynamic linear AIDS approach. *Journal of Travel Research*, Vol. 43, Issue 2 (November), 141-150.

Li, G. Song, H. og Witt, S.F. (2006). Time varying parameter and fixed parameter linear AIDS: An application to tourism demand forecasting. *International Journal of Forecasting*, Vol. 22, Issue 1, 57-71.

Lim, C. og McAleer, M. (2002). Time series forecasts of international travel demand for Australia. *Tourism Management*, Vol, 23, Issue 4 (August), 389-396.

Menges, G. (1958). Die trouristische konsumfunktion der Schweiz 1929-1956. *SISS Journal of Economics and Statistics*, 94(3), 328-268.

Martin, C.A. og Witt, S.F. (1989). Accuracy of econometric forecasts of tourism. *Annals of Tourism Research*, Vol. 16, Issue 3, 407-428.

New Zealand Government (2019). *New Zealand Tourism Forecasts 2019-2025* (May). Ministry of Business, Innovation & Employment.

Njegovan, N. (2006). Are shocks to air passenger traffic permanent or transitory? Implications for long-term air passenger forecasts for the UK. *Journal of Transport Economics and Policy*, 40(2), 315–328.

Pan, B., Wu, D.C. og Song, H. (2012). Forecasting hotel room demand using search engine data. *Journal of Hospitality and Tourism Technology*, 3(3), 196-210.

Rivera, R. (2016). A dynamic linear model to forecast hotel registrations in Puerto Rico using Google Trends data. *Tourism Management*, Vol. 57, 12-20.

Silvey, S. 1975. *Statistical Inference*. Chapman and Hall. London.

Song, H., Gao, B.Z. og Lin, V.S. (2013). Combining statistical and judgmental forecasts via a we-based tourism demand forecasting system. *International Journal of Forecasting*, Vol. 29, 295-310.

Song, H. og Li, G. (2008). Tourism demand modelling and forecasting – A review of recent research. *Tourism Management*, Vol. 29, 203-230.

Song, H., Qiu, R.T. og Park, J. (2019). A review of research on tourism demand forecasting: Launching the *Annals of Tourism Research* Curated Collection on tourism demand forecasting. *Annals of Tourism Research*, Vol 75, 338-362.

Song, H., og Witt, S.F. (2006). Forecasting International Tourist Flows to Macau. *Tourism Management*, Vol. 27, 214-224.

Song, H., Witt, S.F. og Jensen, T.C. (2003). Tourism forecasting: accuracy of alternative econometric models. *International Journal of Forecasting*, Vol. 19, 123-141.

Song, H., Witt, S.F., og Li, G. (2003). Modelling and forecasting the demand for Thai tourism. *Tourism Economics*, 9(4), 363-387.

Soren, J. (2006). Cointegration: a survey, í Mills, T.C. og Patterson, K. (ritstj.) Palgrave *Handbook of Econometrics: Vol. I, Economic Theory*, Palgrave MacMillan, Basingstoke, UK og New York, USA, 540-577.

Stienmetz, J.L., Maxcy, J.G., og Fesenmaier, D.R. (2015). Evaluating Destination Advertising. *Journal of Travel Research*, Vol 54(1), 22-35.

Tsaur, C-C. og Kuo, T-C. (2011). The adaptive fuzzy system time series model with an application to Tawain‘s tourism demand. *Expert Systems with Applications*, Vol. 38, 9164-9171.

Theil, H. 1971. *Principles of Econometrics*. North Holland. Amsterdam.

Transport Canada (2003). Presentation and Discussion of the PODM-V2 Model, (June).

Tucci, M.P. (1995). Time-varying parameters: a critical introduction. *Structural Change and Economic Dynamics*, Vol. 6, 237-260.

Witt, S.F., og Witt, C.A. (1995). Forecasting tourism demand: A review of empirical research. *International Journal of Forecasting*, Vol. 11, 447-475.

Wong, K.K.F., Song, H.,, Witt, S.F., og Wu, D.C. (2007). Tourism forecasting: To combine or not to combine? *Tourism Management*, Vol. 28, 1068-1078.

Yang, X., Pan, B., Evans, J.A. og Lv, B. (2015). Forecasting Chinese tourist volume with search engine data. *Tourism Management*, Vol. 46, 386-397.

Yang, X., Pan, B. og Song, H. (2014). Predicting hotel demand using destination marketing organizations‘s web traffic data. *Journal of Travel Researchs*, 53(4), 433-447.

ÚTGEFIÐ Í MAí 2021

Ferðamálastofa
Icelandic Tourist Board

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

www.ferdamalastofa.is