

KUOLEVUUDEN ENNUSTAMINEN ARIMA-MALLEILLA

Iiro Marttila

Suppea SHV-työ
15. maaliskuuta 2024

Abstract

In this study, we analyse univariate (zero-factor) time series models, specifically ARIMA(p,d,q) models, and their ability to predict Finnish age- and gender-specific mortality rates and period life expectancies. Recent academic studies have found that time series mortality models can perform better in out-of-sample back tests than Lee-Carter models. Therefore, we aim to further study the differences between (zero-factor) ARIMA models and more traditional Lee-Carter and Lee-Miller mortality models.

Based on time series analysis, we find that the simplest ARIMA model specifications, such as ARIMA(0,1,1) and ARIMA(1,1,1), usually have the best fit to Finnish log-mortality rates. However, we also find that the predicted mortality rates and life expectancies (expected values) are not very sensitive to the choice of model specification (at least when either $p \geq 1$ or $q \geq 1$). Instead, mortality rate predictions are more sensitive to the choice of time interval used for the model fitting as the observed historical mortality trend largely drives the predicted mortality rates too.

We compare ARIMA(0,1,1) and Lee-Miller predictions using models fitted with Finnish age-specific mortality rates from 1955 to 2009 which leaves years 2010—2022 for out-of-sample performance analysis. We find that ARIMA models produce more accurate mean predictions for females, but Lee-Miller model produces more accurate mean predictions for males. In addition, ARIMA models produce wider prediction intervals for both genders (especially at the beginning of the forecast horizon). Therefore, qualitative analysis indicates that ARIMA models capture the realized out-of-sample rates more often than Lee-Miller models. On the other hand, some ARIMA prediction intervals appear to be unreasonably wide especially when compared against the data from the COVID-19 pandemic.

We compare ARIMA(1,1,1), Lee-Miller and Lee-Carter predictions using models fitted with Finnish age-specific mortality rates from 1960 to 2022. We observe that ARIMA models predict somewhat lower mortality rates (especially for older age groups) and therefore longer life expectancies than Lee-Carter or Lee-Miller models. Furthermore, we observe that ARIMA models have wider prediction intervals than Lee-Miller models but narrower than Lee-Carter models.

We conclude that ARIMA models seem to be well suited for predicting Finnish mortality rates and life expectancies. However, we note that there is still much room for improvement in the (zero-factor) time series modelling framework. For example, by adding dependency modelling between independent ARIMA models we believe that it is possible to improve the overall performance of the models. In addition, more out-of-sample statistical testing is required before the prediction performance of ARIMA models can be properly evaluated.

Sisältö

1 Johdanto	1
2 ARIMA	3
2.1 ARIMA-prosessit	3
2.2 ARIMA-mallin valinta ja sovittaminen	5
2.3 ARIMA-malleilla ennustaminen	7
3 Kuolevuuden ja elinajanodotteiden ennustaminen ARIMA-malleilla	9
3.1 ARIMA-spesifikaation valinta	9
3.2 Ennusteet eri ARIMA-spesifikaatioilla	11
3.3 Vertailu Lee-Carter ja Lee-Miller -malleihin	14
4 Yhteenveto	21
Lähteet	22
A Kuolevuusaineisto ja elinajanodotteiden laskeminen	24
B ARIMA-mallien ennusteiden luottamusvälit	25
C Lee-Carter ja Lee-Miller -mallit	27
D Elinajanodotteiden ennusteiden luottamusvälit	30
E Rullaava ARIMA-ennuste	31
F Kuvakokoelma	32

Taulukot

1	Miesten ennustetut elinajanodotteet eri ARIMA-spesifikaatioilla.	14
2	Naisten ennustetut elinajanodotteet eri ARIMA-spesifikaatioilla.	14
3	ARIMA(1,1,1)-, Lee-Miller ja Lee-Carter -mallien ennustamat elinajanodotteet, miehet.	20
4	ARIMA(1,1,1)-, Lee-Miller ja Lee-Carter -mallien ennustamat elinajanodotteet, naiset.	20

Kuvaajat

1	50- ja 60-vuotiaiden miesten sekä 80- ja 90-vuotiaiden naisten log-kuolevuuden ja log-kuolevuuden muutosten aikasarjat 1950-luvulta alkaen.	10
2	ARIMA-mallin sovituksen aloitusvuoden vaikutus ennusteeseen.	12
3	ARIMA-mallin spesifikaation vaikutus ennusteeseen.	13
4	ARIMA(0,1,1)- ja Lee-Miller -mallien ennustamat kuolevuudet toteumaan verrattuna, 75-vuotiaat.	15
5	ARIMA(0,1,1)- ja Lee-Miller -mallien ennustamat elinajanodotteet toteumaan verrattuna.	17
6	ARIMA(0,1,1)- ja Lee-Miller -mallien MAPE-arvojen vertailu.	18
7	ARIMA(1,1,1)- ja ARMA(1,1)-malleilla toteutetut log-kuolevuuden ennusteet.	26
8	Log-kuolevuuden muutosten ennusteet ja 95 prosentin luottamusvälit Lee-Miller -mallilla.	29
9	Rullaava ARIMA-ennuste 80-vuotiaiden miesten ja naisten kuolevuudelle ja elinajanodotteelle.	31
10	Simuloituja esimerkkejä ARIMA(1,d,1)-prosesseista.	32
11	Esimerkkejä simuloitujen MA(1)-, AR(3)- ja ARMA(3,1)-prosessien estimoiduista otosautokorrelaatio- ja osittaisautokorrelaatiofunktioista luottamusväleihin.	33
12	ARMA-mallin spesifikaation vaikutus ennusteeseen.	34
13	Log-kuolevuuden ADF-testien p-arvot.	35
14	Log-kuolevuuden muutosten LB-testien p-arvot.	35
15	Log-kuolevuuden keskimääräiset otosautokorrelaatio- ja osittaisautokorrelaatiofunktiot.	36
16	Sopivimmat ARIMA-mallit auto.arima -funktion ja AICc-informaatiokriteerin perusteella.	37
17	ARIMA(0,1,1)-mallin residuaalien LB-testit.	38
18	ARIMA(0,1,1)-mallin residuaalien keskimääräiset otosautokorrelaatio- ja osittaisautokorrelaatiofunktiot.	39
19	Log-kuolevuuden muutosten ennusteet ja 95 prosentin luottamusvälit ARIMA-malleilla.	40
20	Elinajanodotteen ennusteet miehille eri ARIMA-spesifikaatiolla.	41
21	Elinajanodotteen ennusteet naisille eri ARIMA-spesifikaatiolla.	41
22	ARIMA(0,1,1)- ja Lee-Miller -mallien ennustamat kuolevuudet toteumaan verrattuna, 30-vuotiaat.	42
23	ARIMA(0,1,1)- ja Lee-Miller -mallien ennustamat kuolevuudet toteumaan verrattuna, 50-vuotiaat.	42
24	ARIMA(0,1,1)- ja Lee-Miller -mallien ennustamat kuolevuudet toteumaan verrattuna, 90-vuotiaat.	43
25	Elinajanodotteen ennusteiden vertailu, ARIMA(0,1,1) ja Lee-Miller.	43
26	ARIMA(1,1,1)-, Lee-Carter ja Lee-Miller -mallien kuolevuus, 90-vuotiaat.	44
27	ARIMA(1,1,1)-, Lee-Carter ja Lee-Miller -mallien elinajanodotteet.	44
28	Log-kuolevuuden muutosten korrelaatiomatriisi 20—110+ -vuotiaille miehille.	45
29	Log-kuolevuuden muutosten korrelaatiomatriisi 20—110+ -vuotiaille naisille.	46

1 Johdanto

Tässä SHV-työssä tarkastellaan suomalaisten miesten ja naisten ikäkohtaisten kuolevuuksien mallintamista ARIMA eli *Autoregressive Integrated Moving Average* -malleilla. Lisäksi tässä työssä tarkastellaan kuolevuusaikasarjojen tilastollisia ominaisuuksia aikasarja-analyysin näkökulmasta. Luotettavia kuolevuusennusteita tarvitaan vakuutuslalla esimerkiksi henkivakuutusvastuiden ja eläkemutoisten vastuiden arvioimiseen. Haastetta kuolevuuden ennustamiseen on perinteisesti tuonut kuolevuuden laskeva trendi, joka on suomalaisten miesten ja naisten kuolevuusaikasarjoissa jatkunut vaihtelevalla vauhdilla käytännössä yli sadan vuoden ajan (pl. sotavuodet) aina 2020-luvun alkuun asti. Toisaalta aivan viime vuosina eli koronapandemian aikana kuolevuus on merkittävästi kasvanut erityisesti yli 80-vuotiaissa, mikä tuo tällä hetkellä uudenlaista epävarmuutta kuolevuuden ennustamiseen.¹

Kuolevuuden ennustamista varten on luotu useita erilaisia tilastollisia malleja. Booth ja Tickle (2008) [2] toteavat, että kuolevuuden mallintaminen suoraan ARIMA-malleilla, eli ns. nollan faktorin ARIMA-malleilla, on kuitenkin ollut suhteellisen harvinaisia. Nollan faktorin ARIMA-mallien haasteena on mallikokonaisuuden valtava koko, sillä jokainen ikäkohtainen malli vaatii omat parametrinsa. Lisäksi ikäkohtaisten kuolevuusaikasarjojen keskinäinen riippuvuus ei tule yksiulotteisissa ARIMA-malleissa huomioiduksi, jolloin kuolevuus voi kehittyä epäuskottavasti erityisesti pidemmissä ennusteissa. Kuolevuusaikasarjojen keskinäinen riippuvuus voitaisiin huomioida moniulotteisia aikasarjamalleja käyttämällä, mutta tällöin mallin koko (dimensio) olisi suuri. Edellä mainittujen ongelmien takia ARIMA-malleja onkin yleensä käytetty vain osana kuolevuuden faktorimalleja. Faktorimalleissa ikäkohtaisten kuolevuuksien kehitys jaetaan yhteen tai useampaan faktoriin, jolloin ajassa muuttuvien faktoreiden kehitystä voidaan mallintaa ARIMA-malleilla. Näin ikäkohtaisten kuolevuusennusteiden keskinäinen riippuvuus saadaan mallinnettua ilman, että mallin koko paisuu liialliseksi.

Tällä hetkellä yleisimpiä kuolevuuden tilastollisia mallinnusmenetelmiä ovat Lee-Carter -malli (1992) [8] ja tästä pienillä muutoksilla saatava Lee-Miller -malli (2001) [9]. Malleissa kuolevuuden kehitys jaetaan kahteen faktoriin, ikä- ja aikafaktoriin, jolloin kuolevuuden ikäkohtaiset ennusteet saadaan muodostettua ennustamalla aikafaktorin kehitystä tulevaisuuteen satunnaiskulkuprosessilla (ks. luku 2.1). Suomessa Lee-Miller -malli on käytössä esimerkiksi Finanssivalvonnan ylläpitämässä henkivakuutuksen K2012-referenssikuolevuusmallissa (Sirén (2013) [11]) sekä Tapaturma-, Liikenne- ja Potilasvakuutuskeskuksen ylläpitämässä työtapa-turma, liikenne- ja potilasvakuutuksen K2021-referenssikuolevuusmallissa. Lee-Carter ja Lee-Miller -mallien tarkemmat tekniset kuvaukset esitetään liitteessä C.

Lee-Carter -malliin liittyvässä varhaisemmassa kirjallisuudessa nollan faktorin satunnaiskulkumalleja on jonkun verran hyödynnetty vertailukohtana Lee-Carter -mallille. Esimerkiksi Carter ja Lee (1992) [4] vertasivat nollan faktorin satunnaiskulkuprosessilla muodostettuja yhdysvaltalaisten miesten ja naisten elinajanodotteiden ennusteita Lee-Carter -mallilla tuotettuihin ennusteisiin. Heidän tulosten perusteella satunnaiskulkuprosessilla muodostettu ennuste pärjäsikin Lee-Carter -mallin ennustetta huonommin, kun ennustetta verrattiin toteutuneisiin elinajanodotteisiin. Sen sijaan Bell (1997) [1] mallinsi yhdysvaltalaisten miesten ja naisten ikäkohtaisten kuolevuuksien logaritmeja (ns. log-kuolevuutta) satunnaiskulkuprosessilla. Bellin tulosten perusteella yksinkertainen satunnaiskulkumalli pärjäsikin käytännössä yhtä hyvin kuolevuuden lyhyen aikavälin (1–5 vuotta) ennustamisessa kuin Lee-Carter -malli. Bellin tulos antaaakin tukea sille, että yksinkertaisillakin aikasarjamalleilla voidaan saada vertailukelpoisia

¹ks. <https://www.stat.fi/uutinen/elinajanodote-laski-kuolleisuuden-selvan-kasvun-myota-vuonna-2022>

kuolevuuden ennusteita aikaiseksi.

Kiinnostus nollan faktorin ARIMA-malleihin vaikuttaa hieman lisääntyneen 2010-luvun aikana. Esimerkiksi Giacometti ym. (2012) [5] esittivät mallin, jossa ikäkohtaisia kuolevuuksia mallinnetaan AR-ARCH -mallilla, jossa ARCH-termi (*Autoregressive Conditional Heteroskedasticity*) viittaa kuolevuusaikasarjan volatilitettiin ehdolliseen mallintamiseen. He sovittivat AR-ARCH -mallin Italian kuolevuusaikasarjoihin ja vertasivat mallin tuottamaa ennustetta Lee-Carter -mallin tuottamaan ennusteeseen kolmen peräkkäisen vuoden osalta. Vertailun perusteella AR-ARCH -mallin ennusteet olivat Lee-Carter -mallin ennusteita huomattavasti parempia. Hieman vastaavasti Syuhada ja Hakim (2021) [12] testasivat yhdysvaltalaisien miesten ja naisten kuolevuusaikasarjojen tilastollisia ominaisuuksia ja päätyivät tämän analyysin perusteella mallintamaan log-kuolevuuden muutoksia AR-ARCH ja AR-SVAR (*Stochastic Volatility Autoregressive*) -malleilla. Tutkimuksessa malleilla toteutettiin kuolevuudelle VaR-riskimitan tyylliset yhden vuoden riskimitat eli ns. "*Mortality-at-Risk*" -mitat.

Lin ym. (2015) [10] vertasivat ikäkohtaisten log-kuolevuuksien AR, AR-ARCH ja AR-GARCH -mallien (*General Autoregressive Conditional Heteroskedasticity*) ennustekykä Lee-Carter -malliin Japanin, Yhdysvaltojen ja Ison-Britannian kuolevuusaineistoilla vuosien 2001—2009 ajalta. Kaikki mallit oli sovitettu vuosien 1950—2000 aikasarjoihin. Heidän tulosten perusteella kaikkien aikasarjamallien ennustekykä oli Lee-Carter -mallia parempi. Lisäksi aikasarjamallit pääsääntöisesti tuottivat Lee-Carter -mallia leveämpiä ja paremmin toteutuneen kuolevuuden kattaneita ennusteiden luottamusvälejä. Valitettavasti tutkimuksessa ei kuitenkaan vertailtu aikasarjamalleja Lee-Miller -malliin, joka olisi voinut odotetun kuolevuuden osalta pärjätä Lee-Carter -mallia paremmin (ks. liite C). Lisäksi tutkimuksessa ei vertailtu malleja elinajanodotteiden ennusteiden näkökulmasta. Tutkimuksen tulokset antavat kuitenkin erittäin positiivisia viitteitä ARIMA-mallien potentiaalista.

Lyhyen kirjallisuuskatsauksen perusteella nollan faktorin ARIMA-mallit vaikuttavat olevan potentiaalisia työkaluja kuolevuuden mallintamiseen, mutta tutkimusta aiheesta on vielä verrattain rajallisesti. Tämän SHV-työn tarkoituksena onkin tarkastella, miten yksinkertaiset ARIMA-mallit soveltuisivat suomalaisten miesten ja naisten kuolevuuksien ja elinajanodotteiden ennustamiseen. Tämän työn luvussa 2 käydään läpi yksiulotteisten aikasarjojen ja ARIMA-prosessien olennainen teoria sekä kuvataan ARIMA-mallien valintaan, sovittamiseen ja ennusteisiin liittyvää teoriaa. Luvussa 3 kuvataan suomalaisten miesten ja naisten log-kuolevuuden aikasarjojen tilastollisia ominaisuuksia, valitaan ja sovitetaan ARIMA-mallit ja toteutetaan eri ARIMA-malleilla kuolevuuden ja elinajanodotteiden ennusteet. Lisäksi luvussa 3 vertaillaan eri ARIMA-mallien ennusteita sekä keskenään että Lee-Carter ja Lee-Miller -mallien ennusteisiin. Luvussa 4 tiivistetään tämän työn päätulokset.

2 ARIMA

ARIMA-mallien teoria kuvataan olennaisilta osin tässä luvussa. Aikasarjojen ja ARIMA-mallien teoria esitetään Rob J. Hyndmanin ja George Athanasopoulosin (2021) [6] aikasarjamallien käytännön soveltamista R:llä käsittelevän oppikirjan pohjalta sekä Peter J. Brockwellin ja Richard A. Davisin (2002) [3] aikasarjamalleja teoreettisemmin käsittelevän kirjan pohjalta.

2.1 ARIMA-prosessit

Aikasarjamalleja voidaan hyödyntää tilanteissa, joissa stokastista prosessia $\{y_t : t = 0, \pm 1, \pm 2, \dots\}$ (lyhyesti y_t) mallinnetaan tilastollisesti prosessin havaittujen arvojen eli aikasarjan (y_1, y_2, \dots, y_T) avulla, missä $t = 1, 2, \dots, T$ kuvaa aikasarjan havaintohetkiä. Stokastisen prosessin voidaan sanoa olevan heikosti stationaarinen, mikäli prosessin odotusarvo μ on ajanhetkestä t riippumaton, eli $E[y_t] = \mu$, ja prosessin autokovarianssifunktio $\gamma_{t,s}$ ei riipu ajanhetkistä s ja t vaan ainoastaan näiden erotuksesta $h = s - t$, eli $Cov(y_t, y_{t+h}) = \gamma_{t,t+h} = \gamma_{0,h} = \gamma_h$. Koska stokastisen prosessin varianssi on $Var[y_t] = \gamma_0$, saadaan prosessin autokorrelaatio puolestaan kaavalla $\rho_h = Cor[y_t, y_{t+h}] = \gamma_h/\gamma_0$. Yksinkertainen mutta hyödyllinen esimerkki heikosti stationaarisesta prosessista on valkoinen kohina ϵ_t , jolle pätee $E[\epsilon_t] = 0$, $Var[\epsilon_t] = \sigma^2$ ja $Cov[\epsilon_s, \epsilon_t] = 0$.

Odotusarvolle, autokovarianssille ja autokorrelaatioille voidaan luonnollisesti laskea otosestimaatit. Tällöin otoskeskiarvo \bar{y} saadaan kaavalla

$$\bar{y} = \frac{1}{T} \sum_{t=1}^T y_t, \quad (1)$$

otosautokovarianssi c_h kaavalla

$$c_h = \frac{1}{T-h} \sum_{t=1}^{T-h} (y_t - \bar{y})(y_{t+h} - \bar{y}), 0 \leq h < T \quad (2)$$

ja otosautokorrelaatio r_h kaavalla $r_h = c_h/c_0$, jota kutsutaan viipeen h suhteen esitettynä myös otosautokorrelaatiofunktioiksi. Otosautokorrelaation $(1 - \alpha)$ -luottamusvälille pätee ehto $\pm N_{1-\alpha/2}/\sqrt{T}$, jossa $N_{1-\alpha/2}$ on normaalijakauman $(1 - \alpha/2)$ -kvantiili. Tällöin otosautokorrelaation 95%-luottamusvälin tapauksessa $\alpha = 5\%$ ja itse luottamusväli on likimain $\pm 1,96/\sqrt{T}$.

Lineaarinen MA(q)-prosessi voidaan määrittellä kaavalla

$$y_t = \epsilon_t + \sum_{j=1}^q \theta_j \epsilon_{t-j}, \epsilon_t \sim iid(0, \sigma^2), \quad (3)$$

jossa termi "iid" tarkoittaa riippumattomia ja samoin jakautuneita satunnaismuuttujia (engl. *independent and identically distributed*). Toisin sanoen, MA(q)-prosessi muodostuu valkoisen kohinan ϵ_t realisaatiosta hetkellä t sekä q -määrästä prosessin menneitä havaintoja, joiden vaikutus prosessin y_t arvoon määräytyy kertoimien $\theta_j, j = 1, \dots, q$ perusteella. Vastaavasti AR(p)-prosessi voidaan määrittellä kaavalla

$$y_t = \sum_{i=1}^p \phi_i y_{t-i} + \epsilon_t, \epsilon_t \sim iid(0, \sigma^2), \quad (4)$$

eli prosessi muodostuu valkoisen kohinan realisaatiosta sekä p -määrästä prosessin y_t menneitä realisaatioita, joiden vaikutus arvoon y_t määräytyy kertoimien $\phi_i, i = 1, \dots, p$ perusteella. Luonnollisesti ARMA(p,q)-prosessi määritellään edellä kuvattujen prosessien yhdistelmänä, jolloin ARMA(p,q)-prosessia kuvaava yhtälö on muotoa

$$y_t = \sum_{i=1}^p \phi_i y_{t-i} + \sum_{j=1}^q \theta_j \epsilon_{t-j} + \epsilon_t, \epsilon_t \sim iid(0, \sigma^2). \quad (5)$$

ARMA-prosesseja voidaan esittää myös niin sanotun viivästysoperaattorin B avulla. Viivästysoperaattori määritellään prosessille x_t yhtälöllä $Bx_t = x_{t-1}$, jolloin induktiivisesti pätee myös $B^k x_t = x_{t-k}$. Viivästysoperaattorin avulla esimerkiksi ARMA(1,1)-prosessi, ts. $y_t = \phi y_{t-1} + \theta \epsilon_{t-1} + \epsilon_t$, voidaan kuvata yhtälöllä $(1 - B\phi)y_t = (1 + B\theta)\epsilon_t$. Yleisemmin ARMA(p,q)-prosessi voidaan viivästysoperaattorin avulla kuvata yhtälöllä

$$\left(1 - \sum_{i=1}^p \phi_i B^i\right) y_t = \left(1 + \sum_{j=1}^q \theta_j B^j\right) \epsilon_t. \quad (6)$$

Edellä kuvatut AR-, MA-, ja ARMA-prosessit ovat heikosti stationaarisia prosesseja, eli niillä ei voida mallintaa esimerkiksi trendin sisältäviä aikasarjoja. Trendin omaavia, tai yleisemmin epästationaarisia aikasarjoja voidaan kuitenkin käsitellä stationaaristen aikasarjojen menetelmillä poistamalla muunnoksilla alkuperäisen aikasarjan epästationaarisuus. Käytännössä yleisin sovellettava muunnos on alkuperäisen aikasarjan muuttaminen muutosaikasarjaksi, ts. $y_t - y_{t-1} = (1 - B)y_t$. Alkuperäisestä aikasarjasta voidaan ottaa muutos useamman kuin yhden kerran, jos yhden kerran otettu muutos ei riitä muuntamaan alkuperäistä aikasarjaa stationaariseksi. Useasti otettu muutos voidaan esittää notaatiolla $(1 - B)^d y_t$, jossa muutos on otettu d -kertaa. Käytännössä stokastista prosessia y_t sanotaan astetta d olevaksi integroituneeksi prosessiksi eli I(d)-prosessiksi, jos d -kertaa otetun muutoksen jälkeen prosessi on (asymptoottisesti) stationaarinen, mutta $(d - 1)$ -kertaa otetun muutoksen jälkeen ei.

Yleinen esimerkki I(1)-prosessista on satunnaiskulku, jota voidaan kuvata yhtälöllä $y_t - y_{t-1} = c + \epsilon_t$, jossa $\epsilon_t \sim iid(0, \sigma^2)$. Peräkkäisillä sijoituksilla voidaan näyttää, että prosessi y_t seuraa tällöin yhtälöä

$$y_t = y_0 + ct + \sum_{j=1}^t \epsilon_j, t = 1, 2, \dots, \quad (7)$$

jossa vakio c kuvaa prosessin y_t deterministisen trendin kulmakerrointa ja $\sum_{j=1}^t \epsilon_j$ stokastista heiluntaa deterministisen trendin ympärillä. Prosessin y_t odotusarvoksi saadaan $E[y_t] = y_0 + ct$, varianssiksi $Var[y_t] = Var[\sum_{j=1}^t \epsilon_j] = \sigma^2 t$ ja autokorrelaatioksi $Cor[y_t, y_{t+h}] = 1/(\sqrt{(1 + h/t)})$, eli prosessin arvot riippuvat ajanhetkestä t (ts. satunnaiskulkuprosessi on selvästi epästationaarinen prosessi).

Yhdistämällä I(d)-prosessi ARMA(p,q)-prosessiin saadaan muodostettua ARIMA(p,d,q)-prosessi, jota voidaan kuvata yhtälöllä

$$\left(1 - \sum_{i=1}^p \phi_i B^i\right) (1 - B)^d y_t = \left(1 + \sum_{j=1}^q \theta_j B^j\right) \epsilon_t + c. \quad (8)$$

Esimerkiksi ARIMA(1,1,1)-prosessin yhtälö on tällöin muotoa

$$y_t - y_{t-1} = \phi(y_{t-1} - y_{t-2}) + \theta \epsilon_{t-1} + \epsilon_t + c. \quad (9)$$

Vakiotermin c arvolla ja muutosten määrällä eli d :n arvolla on ARIMA-prosessin kululle merkittävä vaikutus. Seuraavassa listassa esitellään ARIMA(p,d,q)-prosessien kulkuja d :n vaihdellessa nollan ja yhden välillä ja c :n ollessa nolla tai nollasta poikkeava.

- Jos $c = 0$ ja $d = 0$, niin stationaarinen prosessi heiluu nollan ympärillä.
- Jos $c \neq 0$ ja $d = 0$, niin stationaarinen prosessi heiluu nollasta poikkeavan keskiarvon c ympärillä.
- Jos $c = 0$ ja $d = 1$, niin prosessi on epästationaarinen, mutta ei osoita selvää trendiä.
- Jos $c \neq 0$ ja $d = 1$, niin prosessilla on lineaarinen trendi ct , jonka ympärillä prosessi heiluu satunnaisesti.

Liitteen F kuvassa 10 esitetään edellä kuvattujen eri ARIMA-prosessien simuloituja kulkuja.

2.2 ARIMA-mallin valinta ja sovittaminen

Sopivaa ARIMA-mallin spesifikaatiota (eli p :n, d :n ja q :n arvoja) valittaessa ensimmäisenä haasteena on yleensä löytää I(d)-prosessin d :n arvo. Käytännössä tässä työssä rajoitutaan tilanteisiin, joissa d on joko 1 tai 0. Prosessin I(d) testaamisessa olennaisessa osassa ovat tilastolliset yksikköjuuritestit. Yksikköjuurien ideasta voidaan lyhyesti todeta, että mikäli $y_t \sim AR(p)$ ja prosessilla y_t ei ole yksikköjuuria, on y_t tällöin asymptoottisesti stationaarinen prosessi ja $d = 0$. Sen sijaan yksikköjuuren olemassaolo viittaa siihen, että prosessi y_t tarvitsee muuntaa muutosprosessiksi $(1 - B)y_t = \Delta y_t$, eli $d = 1$.

Yksikköjuuren olemassaoloa voidaan testata useiden erilaisten tilastollisten testien avulla. Näistä yleisin on täydennetty Dickey-Fuller -testi (engl. Augmented Dickey-Fuller, lyhyesti ADF).² Testin nollahypoteesina on yksikköjuuren olemassaolo ja vastahypoteesina aikasarjan stationaarisuus tai stationaarisuus trendin suhteen. Toisin sanoen, jos testin nollahypoteesi jää voimaan, testattu aikasarja sisältää yksikköjuuren ja AR-malli tulee perustaa aikasarjan muutoksille. Toinen yleinen yksikköjuuritesti on Kwiatkowski-Phillips-Schmidt-Shin-testi (lyhyesti KPSS) [7], jossa nollahypoteesina on testatun aikasarjan stationaarisuus. Toisin kuin ADF-testissä, pienet p-arvot KPSS-testissä viittaavat tarpeeseen muuntaa testattu aikasarja muutosaikasarjaksi.³

MA- ja AR-prosessien sopivuutta ja spesifikaatioita (ts. q :n ja p :n arvoja) voidaan arvioida autokorrelaatiofunktioiden avulla, kun d :n arvo tiedetään. Otosautokorrelaatioiden lisäksi aikasarjoista voidaan estimoida osittaisautokorrelaatiot, jotka mittaavat arvon y_t ja arvon y_{t-h} välistä riippuvuutta puhdistettuna muiden viipeiden (t.s. 1, 2, ..., $h - 1$) tuomista vaikutuksista. AR(p)-prosessin osittaisautokorrelaatiolle voidaan näyttää pätevän, että jos $y_t \sim AR(p)$, niin osittaisautokorrelaatio viipeelle p on AR(p)-prosessin parametrin ϕ_p suuruinen ja $\phi_i = 0$, kun $i > p$. Otosautokorrelaatioiden tapaan myös estimoitujen osittaisautokorrelaatioiden luottamusvälit ovat muotoa $\pm N_{1-\alpha/2}/\sqrt{T}$. Autokorrelaatiofunktioiden avulla tehtävää analyysia voidaan tiivistää seuraavien ehtojen avulla:

- Mikäli otosautokorrelaatiofunktiossa on selvä katkos viipeen q kohdalla ja osittaisautokorrelaatiofunktio on vaimeneva tai siniaaltainen, viittaavat autokorrelaatiofunktio MA(q)-prosessiin.

²Ks. <https://www.rdocumentation.org/packages/aTSA/versions/3.1.2/topics/adf.test>.

³Ks. <https://www.rdocumentation.org/packages/tseries/versions/0.10-54/topics/kpss.test>

- Mikäli osittaisautokorrelaatiofunktiossa on selvä katkos viipeen p kohdalla ja otosautokorrelaatiofunktio on vaimeneva tai siniaaltainen, viittaavat autokorrelaatiofunktio AR(p)-prosessiin.

Liitteen F kuvassa 11 esitellään simuloidun MA(1)-, AR(3)- ja ARMA(3,1)-prosessin otosautokorrelaatio- ja osittaisautokorrelaatiofunktioita.

Estimoidut autokorrelaatiofunktioita voivat olla hyödyllisiä työkaluja sopivan AR- ja MA-spesifikaation etsimisessä, mutta eivät kuitenkaan aina riittäviä. Erityisesti ARMA-prosessi tuottaa haasteita, sillä autokorrelaatiofunktioita perusteella ei ARMA-prosessin tapauksessa aina voida suoraan päätellä p :n ja q :n arvoja (ks. liitteen F kuva 11). Tämän takia eräs hyödyllinen keino eri ARMA-spesifikaatioita vertailussa onkin informaatiokriteerien käyttäminen. Tunnettuja informaatiokriteereitä ovat esimerkiksi Akaiken informaatiokriteeri (lyhyesti AIC) ja korjattu AIC (lyhyesti AICc), joista jälkimmäinen rankaisee liian usean parametrin käyttämisestä erityisesti, kun mallinnettava aikasarja on lyhyt. ARIMA-malleille AIC:n määritelmä on

$$AIC = -2\log(L) + 2(p + q + k + 1), \quad (10)$$

ja AICc:n määritelmä on

$$AICc = AIC + \frac{2(p + q + k + 1)(p + q + k + 2)}{T - p - q - k - 2}, \quad (11)$$

jossa $k = 0$ kun $c = 0$, tai $k = 1$ kun $c \neq 0$, L on ARIMA-mallin uskottavuusfunktio ja $\log(\cdot)$ viittaa luonnolliseen logaritmiin.

Virhetermin ϵ_t normaalisuusoletuksen alla (ts. $\epsilon_t \sim \text{nid}(0, \sigma^2)$) voidaan näyttää, että jos $y_t \sim \text{ARMA}(p, q)$, niin tällöin vektorille $\mathbf{y} = (y_1, \dots, y_T)'$ pätee, että $\mathbf{y} \sim N(\mathbf{0}, \sigma^2 \mathbf{\Sigma})$, jossa $T \times T$ kovarianssimatriisi $\mathbf{\Sigma}$ riippuu parametrivektorista $\boldsymbol{\beta} = (\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)'$. Tällöin ARMA-mallin uskottavuusfunktio saadaan multinormaalijakauman tiheysfunktion perusteella muotoon

$$L(\boldsymbol{\beta}, \sigma^2; \mathbf{y}) = \frac{\exp(-\frac{1}{2\sigma^2} \mathbf{y}' \mathbf{\Sigma}(\boldsymbol{\beta})^{-1} \mathbf{y})}{\sqrt{(2\pi)^T \sigma^2 |\mathbf{\Sigma}(\boldsymbol{\beta})|}}. \quad (12)$$

ARMA-mallin parametrit $\boldsymbol{\beta}$ (ja σ) voidaan täten estimoida käyttäen suurimman uskottavuuden menetelmää (engl. *Maximum likelihood method*) eli etsimällä parametriavaruudesta se parametrivektori $\boldsymbol{\beta}^*$, joka maksimoi uskottavuusfunktion havaitulla aikasarjalla. Täten on ilmeistä, että mitä pienempi informaatiokriteerin AIC tai AICc arvo on, sitä paremmin kyseinen ARMA(p, q)-mallin spesifikaatio sopii dataan. ARIMA-mallien parametrien tekniseen estimointiin ei tässä työssä paneuduta tämän enempää, vaan lukijaa kehoitetaan tutustumaan esimerkiksi Brockwellin ja Davisin (2002) [3] kirjan lukuun 5.2 aiheen osalta.

ARIMA-mallin sovittamisen jälkeen mallin sopivuutta voidaan tarkastella tutkimalla empiiristen residuaalien eli todellisten aikasarjan havaintojen y_t ja sovitetulla ARIMA-mallilla estimoitujen havaintojen \hat{y}_t arvojen erotusta, ts. $\hat{\epsilon}_t = \hat{y}_t - y_t$. Käytännössä havaittujen residuaalien tulisi vasta teoreettisten residuaalien oletuksia, eli olla normaalijakautuneita ja riippumattomia odotusarvolla 0 ja varianssilla σ^2 . Jos sovitetun mallin residuaalit eivät noudata residuaalien teoreettisia ominaisuuksia, on ARIMA-mallin sopivuus dataan kyseenalainen.

Tässä työssä keskitytään erityisesti residuaalien autokorrelaatioita eli riippumattomuuden toteutumisen tarkasteluun. Estimoitujen autokorrelaatioita merkittävyttä voidaan arvioida

Ljung-Box -testillä (lyhyesti LB-testi). Testi nojaa oletukseen, jonka mukaan otosautokorrelaatiovektori (r_1, \dots, r_H) on asympotoottisesti multinormaalijakautunut $N(\mathbf{0}, \mathbf{I}/T)$, jossa \mathbf{I} on $H \times H$ yksikkömatriisi. Tämän nojalla voidaan muodostaa LB-testisuure

$$Q_{LB} = T(T+2) \sum_{h=1}^H \frac{r_h^2}{T-h} \stackrel{as}{\sim} \chi_H^2. \quad (13)$$

Ljung-Boxin testin nollahypoteesina on, että $\rho_1 = \dots = \rho_H = 0$, eli aikasarja on autokorreloitamaton, kun taas vastahypoteesina on, että aikasarja on autokorreloitunut.

Käytännössä LB-testillä testataan havaitun aikasarjan (y_1, \dots, y_T) lineaarista riippuvuutta. Epälineaarista riippuvuutta voidaan sen sijaan testata ottamalla LB-testi neliöityjen havaintojen aikasarjasta (y_1^2, \dots, y_T^2) . ARIMA-mallin sopivuutta eli residuaalien $(\hat{\epsilon}_1, \dots, \hat{\epsilon}_T)$ lineaarista riippumattomuutta testattaessa LB-testisuureta verrataan χ_{H-p-q}^2 jakaumaan, kun taas neliöityjen residuaalien tapauksessa verrattava jakauma on edelleen χ_H^2 . Mikäli testien perusteella mallin residuaalien todetaan olevan lineaarisesti riippumattomia mutta epälineaarisesti kuitenkin riippuvia, tarkoittaa tämä sitä, etteivät residuaalit myöskään toteuta normaalijakautuneisuuden oletusta.

2.3 ARIMA-malleilla ennustaminen

Seuraavassa kuvataan prosessin y_t ennustamisen yleiset periaatteet ARIMA(p,1,q)-mallille. Käytännössä ennuste tehdään ensin muutosprosessille Δy_t , jonka perusteella lasketaan puolestaan alkuperäisen y_t -prosessin ennuste. ARIMA(p,0,q)-mallin ennusteprosessi noudattaa käytännössä samoja periaatteita, mutta suoraan y_t -prosessille ennustettuna. Ennustamisen tavoitteena on ennustee prosessin y_t tulevaa arvoa periodilla $T+h$, eli prosessin y_t periodin $T+h$ ehdollista odotusarvoa hetken T tiedoilla. Ennustetta merkitään lyhyesti notaatiolla $\hat{y}_{T+h|T} = E[y_{T+h}|T]$.

Ennustamista varten ARIMA(p,1,q)-mallin $(p, q \geq 1)$ yhtälö kirjoitetaan muutosprosessille Δy_t ARMA(p,q)-mallin muodossa, jolloin

$$\Delta y_t = \sum_{i=1}^p \phi_i \Delta y_{t-i} + \sum_{j=1}^q \theta_j \epsilon_{t-j} + \epsilon_t + c, \quad (14)$$

jossa $\epsilon_t \sim \text{nid}(0, \sigma^2)$ eli oletuksena on residuaalien riippumattomuus ja normaalijakautuneisuus. Tällöin muutosprosessin yhden periodin ennuste saadaan yhtälöstä

$$\Delta \hat{y}_{T+1|T} = \sum_{i=1}^p \phi_i \Delta y_{T-i+1} + \sum_{j=1}^q \theta_j \hat{\epsilon}_{T-j+1} + c. \quad (15)$$

Ennusteyhtälö saadaan tähän muotoon, koska prosessin y_t ja täten prosessin Δy_t arvot ovat periodiin T asti tunnettuja, residuaaliprosessin ϵ_t ehdollinen odotusarvo hetkellä $T+1$ on nolla (ts. $E[\epsilon_{T+1}|T] = 0$) ja residuaalien menneet arvot ovat havaittujen residuaalien $\hat{\epsilon}_t$ mukaisia. Luonnollisesti prosessin y_t hetken $T+1$ mukainen ennuste saadaan tällöin yhtälöstä $\hat{y}_{T+1|T} = y_T + \Delta \hat{y}_{T+1|T}$ eli viimeisen havaitun arvon ja ennustetun muutoksen välisenä summana. Esimerkiksi ARIMA(1,1,1)-prosessin ensimmäisen periodin ennusteen yhtälö on $\hat{y}_{T+1|T} = \phi \Delta y_T + \theta \hat{\epsilon}_T + c + y_T$.

Jatkettaessa ennusteprosessia eteenpäin periodille $T + 2$ muutosprosessin ennusteyhtälö saadaan muotoon

$$\Delta\hat{y}_{T+2|T} = \phi_1\Delta\hat{y}_{T+1|T} + \sum_{i=2}^p \phi_i\Delta y_{T-i+2} + \sum_{j=2}^q \theta_j\hat{\epsilon}_{T-j+2} + c, \quad (16)$$

jossa $p \geq 2$ ja $q \geq 2$. Periodin $T + 2$ ennusteyhtälössä MA-prosessin summa saadaan esitettyyn muotoonsa, sillä $E[\epsilon_{T+2}|T] = E[\epsilon_{T+1}|T] = 0$. Toisaalta AR-prosessin osassa periodin $T + 1$ ennuste $\Delta\hat{y}_{T+1|T}$ otetaan periodin $T + 2$ ennusteen muodostamiseen mukaan. Näin jatkaen ennusteprosessia h -periodin verran eteenpäin, päädytään lopulta tilanteeseen, jossa $p < h$ ja $q < h$. Tällöin ennusteyhtälö muodostuu pelkästään prosessin aiempien periodien ennusteista sekä vakiotermistä c , jolloin ennusteyhtälö saadaan muotoon

$$\Delta\hat{y}_{T+h|T} = \sum_{i=1}^p \phi_i\Delta\hat{y}_{T+h-i|T} + c. \quad (17)$$

Luonnollisesti prosessin y_t ennusteelle pätee yleisesti, että $\hat{y}_{T+h|T} = \hat{y}_{T+h-1|T} + \Delta\hat{y}_{T+h|T}$, kun $h > 1$. Esimerkiksi ARIMA(1,1,1)-mallin toisen periodin ennusteyhtälö saadaan muotoon $\hat{y}_{T+2|T} = \phi\Delta\hat{y}_{T+1|T} + c + \hat{y}_{T+1|T}$ ja kolmannen periodin $\hat{y}_{T+3|T} = \phi\Delta\hat{y}_{T+2|T} + c + \hat{y}_{T+2|T}$.⁴

Edellä esitettyjä ennusteyhtälöitä voidaan luonnollisesti laajentaa siten, että periodin $T + h$ ennusteyhtälöön sijoitetaan iteratiivisesti edeltävien periodin ennusteyhtälöt. Tästä esitystavasta on hyötyä erityisesti mallin ARMA-osan käsittelyssä. Esimerkiksi ARMA(1,1)-mallin ennusteelle pätee (kun $h \geq 2$)

$$\Delta\hat{y}_{T+h|T} = \phi\Delta\hat{y}_{T+h-1|T} + c = \phi^h\Delta y_T + \phi^{h-1}\theta\epsilon_T + \phi^{h-1}c + c. \quad (18)$$

Mikäli parametrille ϕ pätee $|\phi| < 1$ (ts. prosessi on stationaarinen), niin kerroin ϕ^h (ja ϕ^{h-1}) lähestyy nollaa ennustehorisontin h kasvaessa. Toisin sanoen, ARMA(1,1)-mallin ennuste muuttuu vakioksi c ennustehorisontin h kasvaessa tarpeeksi suureksi, ts. $\Delta\hat{y}_{T+h|T} \approx c$, jolloin ARIMA(1,1,1)-mallin ennusteyhtälölle pätee $\hat{y}_{T+h|T} \approx c + \hat{y}_{T+h-1|T}$. Käytännössä voidaan näyttää, että ARMA-prosessin ollessa stationaarinen ennusteen asettuminen vakiouralle yleistyy kaikille ARMA(p,q)-spesifikaatioille.⁵ Karkeasti voidaan sanoa, että ennusteen vakiouran saavuttamisessa kestää sitä pidempään, mitä suurempia p :n ja q :n arvot ovat (toki riippuen myös itse parametrien numeerisista arvoista). Tätä ilmiötä kuvataan esimerkiksi liitteen F kuvassa 12.

Kaikkien ARIMA-mallien tapauksessa ennusteen ensimmäisen periodin luottamusväli on vielä yksinkertaisesti määriteltävissä ja muotoa $\hat{y}_{T+1|T} \pm \hat{\sigma}N_{1-\alpha/2}$, jossa $N_{1-\alpha/2}$ on normaalijakauman $(1 - \alpha/2)$ -kvantiili ja $\hat{\sigma}$ on havaittujen residuaalien $\hat{\epsilon}_t$ keskihajonta. Sen sijaan pidempien ($h \geq 2$) ennusteiden luottamusvälien kaavat riippuvat ARIMA-mallin spesifikaatiosta eli p :n ja q :n arvoista, joten tarkkojen kaavojen esittäminen ei ole yhtä suoraviivaista. Käytännössä ARIMA-mallien ennusteiden luottamusvälit levenevät ennustehorisontin h kasvaessa, mikäli $d \geq 1$. Jos sen sijaan $d = 0$, niin luottamusvälit konvergoituvat saman suuruisiksi ennustehorisontin h kasvaessa tarpeeksi suureksi. Ennusteiden luottamusvälejä tarkastellaan tarkemmin liitteessä B.

⁴ARIMA-malli voidaan määritellä myös siten, että joko $p = 0$ tai $q = 0$, jolloin myös ennusteyhtälöt ovat yksinkertaisempia. Esimerkiksi ARIMA(0,1,1)-mallin ensimmäisen periodin ennusteyhtälö on $\hat{y}_{T+1|T} = \theta\hat{\epsilon}_T + c + y_T$ ja toisesta periodista ($h \geq 2$) alkaen ennusteyhtälö on muotoa $\hat{y}_{T+h|T} = c + \hat{y}_{T+h-1|T}$, eli edeltävään periodiin nähden ennuste muuttuu vain deterministisen lineaarisen trendin kulmakertoimen c verran.

⁵Ks. Brockwell ja Davis (2002) [3] luku 3.1 ARMA-prosessien stationaarisuuden ehtoihin liittyen.

3 Kuolevuuden ja elinajanodotteiden ennustaminen ARIMA-malleilla

Tässä työssä suomalaisten miesten ja naisten ikäkohtaisia kuolevuuksia ennustetaan ARIMA-malleilla. Mallien pohjalla käytetään *Human Mortality Database* -tietokannan (lyhyesti HMD) tasoitettua periodikuolevuusaineistoa (ns. *Life tables*).⁶ Aineistoa kuvataan tarkemmin liitteessä A ja HMD:n metodiprotokollassa (Wilmoth ym. (2021) [13]). Tilastollisessa aineiston analysoimisessa sekä varsinaisessa mallintamisessa käytetään R-ohjelmistoa ja erityisesti Rob J. Hyndmanin forecast-pakettia ARIMA-mallien osalta.⁷

Väestökuolevuuden ennustaminen toteutetaan siten, että HMD-tietokannasta kerättyihin tasoitettuihin log-kuolevuusaikasarjoihin sovitetaan omat ARIMA-mallit ikäkohtaisesti erikseen sekä miehillä että naisille.⁸ Toisin sanoen, $\log(m_{x,t}) \sim ARIMA(p, d, q)$, jossa $m_{x,t}$ on x -ikäisen miehen tai naisen keskikuolevuus vuonna t , $\log(\cdot)$ on luonnollinen logaritmi ja $x = 0, 1, 2, \dots, 110+$, jossa ikä 110+ sisältää kaikki yli 109-vuotiaat. Aikasarjamallit sovitetaan siten, että ARIMA-mallien spesifikaatiot (ts. p, d ja q) ovat samoja kaikille kuolevuusaikasarjoille, mutta parametrien (ϕ, θ, σ ja c) arvot vaihtelevat aikasarjakohtaisesti. Elinajanodotteiden ennusteet lasketaan puolestaan kuolevuuden ennusteiden pohjalta periodiaineistosta liitteiden A ja D mukaisesti. Kuolevuuden ja elinajanodotteiden ennusteet toteutetaan mallien soviteaikavälin lopusta lähtien useaksi vuodeksi eteenpäin (ts. $h \geq 2$).

3.1 ARIMA-spesifikaation valinta

Kuten luvussa 2.2 kuvattiin, aikasarjamallintamisen alussa on olennaista hahmottaa ARIMA-mallin spesifikaatio eli $p:n$, $d:n$ ja $q:n$ arvot. Ensimmäisessä vaiheessa ilmeinen tapa tutkia sopivaa $d:n$ arvoa on tarkastella itse mallinnettavien aikasarjojen kuvaajia. Aikasarjojen tarkasteluväliksi valitaan vuodet 1950—2022 (muutokset 1951—2022)⁹. Koska aikasarjoja on tässä tapauksessa yli kaksisataa, kuvassa 1 esitetään esimerkinomaisesti log-kuolevuuden sekä log-kuolevuuden muutosten aikasarjat vain muutamille eri ikäisille miehille ja naisille. Kuvan perusteella havaitaan, että log-kuolevuuden esimerkkiaikasarjat eivät selvästikään ole stationaarisia, sillä molempien aikasarjojen käyttäytyminen muistuttaa yksikköjuuren sisältävää prosessia. Sen sijaan muutosaikasarjat vaikuttavat silmämääräisesti tarkasteltuina stationaarisilta. Nämä havainnot viittaavat siihen, että ainakin esimerkkiaikasarjojen osalta $d:n$ arvon tulisi olla 1.

Aikasarjojen yksikköjuuren olemassaoloa testataan tilastollisesti käyttämällä ADF-testiä. ADF-testien aikasarjakohtaisten tulosten perusteella havaitaan, että log-kuolevuuden ADF-testin nollahypoteesi eli yksikköjuuri jää pääsääntöisesti voimaan viiden prosentin luottamustasolla, vaikka satunnaisia poikkeuksia esiintyykin erityisesti alle 30-vuotiaissa. Sen sijaan muutosaikasarjojen ADF-testien nollahypoteesit pääsääntöisesti kumoutuvat, joten testien tulokset antavat selvää tukea sille, että yleisessä ARIMA-mallin spesifikaatiossa $d:n$ tulisi olla 1. Liitteen F kuvassa 13 esitellään ADF-testien tuloksia (p-arvoja) tarkemmin.

ADF-testien tulosten perusteella lineaarisen ja epälineaarisen riippuvuuden LB-testit teh-

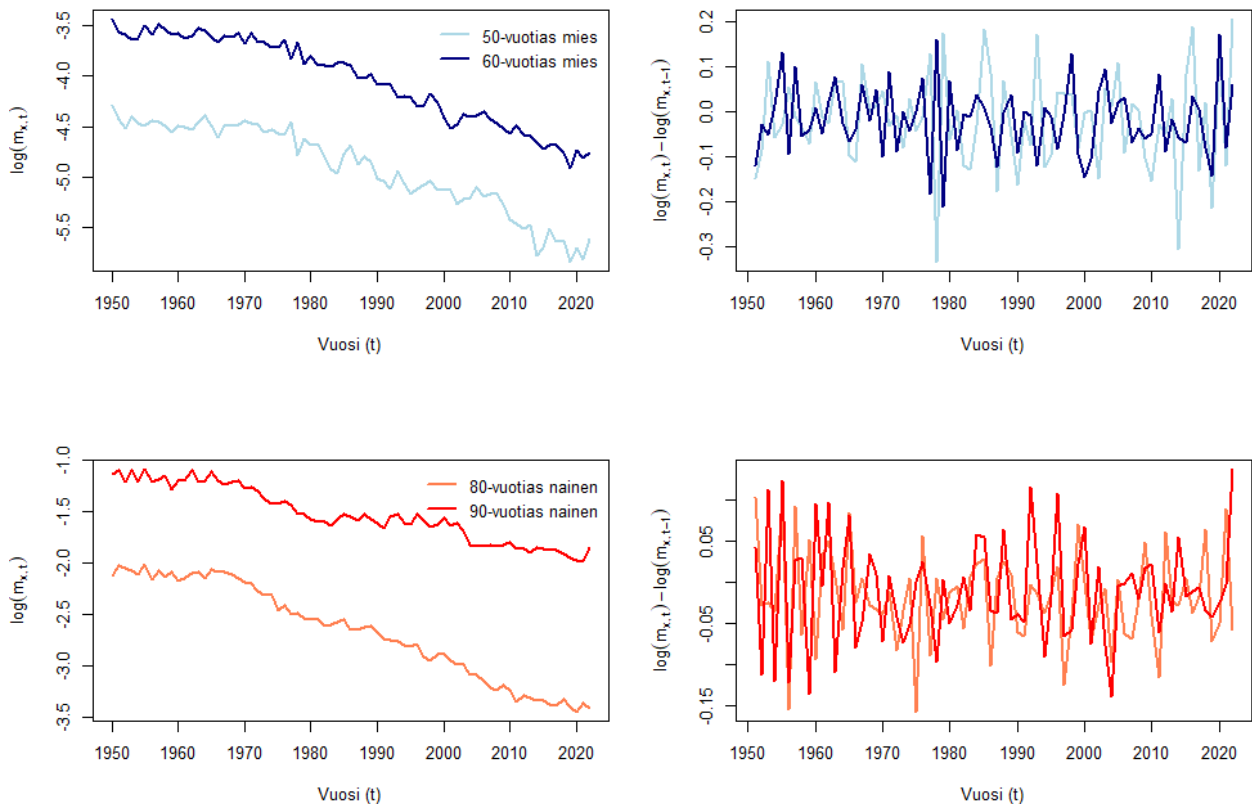
⁶<https://www.mortality.org/Country/Country?cntr=FIN>

⁷<https://github.com/robjhyndman/forecast>

⁸Analysoitavia aikasarjoja sekä sovitettavia ARIMA-malleja on yhteensä 222 kappaletta.

⁹Tilastolliseen tarkasteluun on valittu käytännössä sotien jälkeiset vuodet. Aikasarjojen tilastolliset ominaisuudet eivät kuitenkaan muutu merkittävästi, vaikka aineiston alku- tai loppupäästä jätettäisiin muutamia vuosia pois.

dään log-kuolevuuden muutosajasarjoille. Testien perusteella suurimmassa osassa aikasarjoja havaitaan tilastollisesti merkitsevää lineaarista riippuvuutta viiden prosentin luottamustasolla, joskin joitakin poikkeuksia myös löytyy. Sen sijaan neliöidyissä aikasarjoissa tilastollisesti merkitsevää autokorreloituneisuutta havaitaan harvemmin, joten epälineaarinen riippuvuus ei vaikuta olevan sarjoissa yhtä vahvaa kuin lineaarinen.¹⁰ Testien tulokset antavatkin tukea sille, että log-kuolevuuden muutoksia voidaan mallintaa ARMA-malleilla. LB-testien tuloksia esitellään tarkemmin liitteen F kuvassa 14.



Kuva 1: 50- ja 60-vuotiaiden miesten sekä 80- ja 90-vuotiaiden naisten log-kuolevuuden ja log-kuolevuuden muutosten aikasarjat 1950-luvulta alkaen.

Tilastollisesti merkitsevän autokorreloituneisuuden olemassaolo aikasarjoissa ei riitä vielä ratkaisemaan ARMA-mallin spesifikaatiota eli $p:n$ ja $q:n$ arvoja. Tätä varten tarkastellaan log-kuolevuuden muutosten autokorrelaatiofunktioita. Koska mallinnettavia aikasarjoja on useita ja tavoitteena on löytää kaikille sarjoille yhteinen ARMA-spesifikaatio, esitetään liitteen F kuvassa 15 ikäkohtaisista autokorrelaatiofunktioista lasketut keskimääräiset otos- ja osittaisautokorrelaatiofunktioit miehille ja naisille. Keskimääräisten autokorrelaatiofunktioiden perusteella havaitaan, että yksinkertainen ARMA(0,1)-spesifikaatio saa vahvaa tukea, sillä sekä miesten

¹⁰Epälineaarinen eli neliöityjen log-kuolevuuden muutosten tilastollinen riippumattomuus useiden aikasarjojen tapauksessa tarkoittaa sitä, että ehdollisen varianssin GARCH-mallin käyttö ei saa varauksetonta tukea (pl. n. 90-vuotiaat ja vanhemmat naiset, joilla havaitaan tilastollisesti merkitsevää epälineaarista riippuvuutta). Toisaalta GARCH-malli vaikuttaisikin kuolevuusennusteissa lähinnä ennusteiden luottamusväleihin.

että naisten otosautokorrelaatiofunktioissa on selvät katkokset viipeen 1 kohdalla ja molemmat osittaisautokorrelaatiofunktioit ovat vaimenevia.

Sopivan ARIMA-mallin spesifikaation valinta eri aikasarjoille voidaan automatisoida käyttämällä R:n forecast paketin auto.arima -funktioita, jonka algoritmi kuvataan tarkemmin Hyndmanin ja Athanasopoulosin (2021) [6] kirjan luvussa 9.7.¹¹ Lyhyesti funktion käyttämää algoritmia voidaan kuvata siten, että ensin valitaan sopivin d :n arvo (0, 1 tai 2) toistetuilla KPSS-testeillä, jonka jälkeen p :n ja q :n arvot valitaan eri ARIMA-spesifikaatioiden informaatiokriteereitä vertailemalla.

Tässä työssä auto.arima -funktion avulla etsitään jokaiselle log-kuolevuusaikasarjalle sopivin ARIMA-spesifikaatio. Kaikkiaan miehille tulee valituksi seitsemäntoista ja naisilla kahdeksantoista erilaista spesifikaatiota. Useimmiten sopivimpia ovat suhteellisen yksinkertaiset spesifikaatiot, joissa $d = 1$. Selvästi yleisin spesifikaatio on ARIMA(0,1,1), joka sopii parhaiten noin puolelle aikasarjoista. Saatu tulos on täten linjassa keskimääräisten autokorrelaatiofunktioiden tulosten kanssa. Toiseksi yleisin valittu malli on ARIMA(1,1,1), joka on sopivin noin kolmellekymmenelle aikasarjalle. Tarkemmin auto.arima -funktiolla saatuja tuloksia kuvataan sukupuolittain eriteltynä liitteen F kuvassa 16.

Valitun ARIMA-spesifikaation sopivuutta tarkastellaan sovittamalla sama spesifikaatio kaikkiin log-kuolevuuden aikasarjoihin ja testaamalla sovitettujen mallien residuaaliaikasarjojen lineaarista ja epälineaarista riippuvuutta LB-testeillä. Testien tulokset ARIMA(0,1,1)-mallille esitetään liitteen F kuvassa 17 ja residuaalien keskimääräiset autokorrelaatiofunktioit esitetään liitteen F kuvassa 18. Tulosten perusteella ARIMA(0,1,1)-spesifikaation residuaaliaikasarjoissa ei pääsääntöisesti ole havaittavissa tilastollisesti merkitsevää lineaarista tai epälineaarista autokorreloituneisuutta, joten ARIMA(0,1,1)-malli vaikuttaa useimmissa tapauksissa olevan riittävä mallintamaan log-kuolevuuden aikasarjoja. Toisaalta myös poikkeavia testituloksia saadaan erityisesti noin 20-vuotiaissa ja sitä nuoremmista ikäluokissa, joille ARIMA(0,1,1)-spesifikaatio ei vaikuta täysin optimaaliselta mallilta.¹² Residuaalien LB-testien tulokset ovat olennaisesti samoja myös ARIMA(1,1,1)-mallille.

Edeltävä aikasarja-analyysi osoittaa, ettei kuolevuuden aikasarjoille löydy yhtä yhteistä ja optimaalista ARIMA-spesifikaatiota. Analyysien perusteella kuolevuuden ennusteet kannattaa kuitenkin toteuttaa ainakin yleisimmillä eli ARIMA(0,1,1)- ja ARIMA(1,1,1)-spesifikaatioilla. Vertailun vuoksi kuolevuuden ennusteet toteutetaan kuitenkin myös ARIMA(2,1,2)-spesifikaatiolla. Kyseinen malli tulee auto.arima -funktion perusteella valituksi vain muutamia kertoja sopivimmaksi ARIMA-spesifikaatioiksi, mutta liitteen F kuvan 16 (oikea puoli) perusteella kyseisen mallin AICc-arvot ovat pääsääntöisesti vielä suhteellisen hyvällä tasolla. Tästä syystä ARIMA(2,1,2)-spesifikaatiota käytetään esimerkkinä hieman isomman ARIMA-mallin ominaisuuksista.

3.2 Ennusteet eri ARIMA-spesifikaatioilla

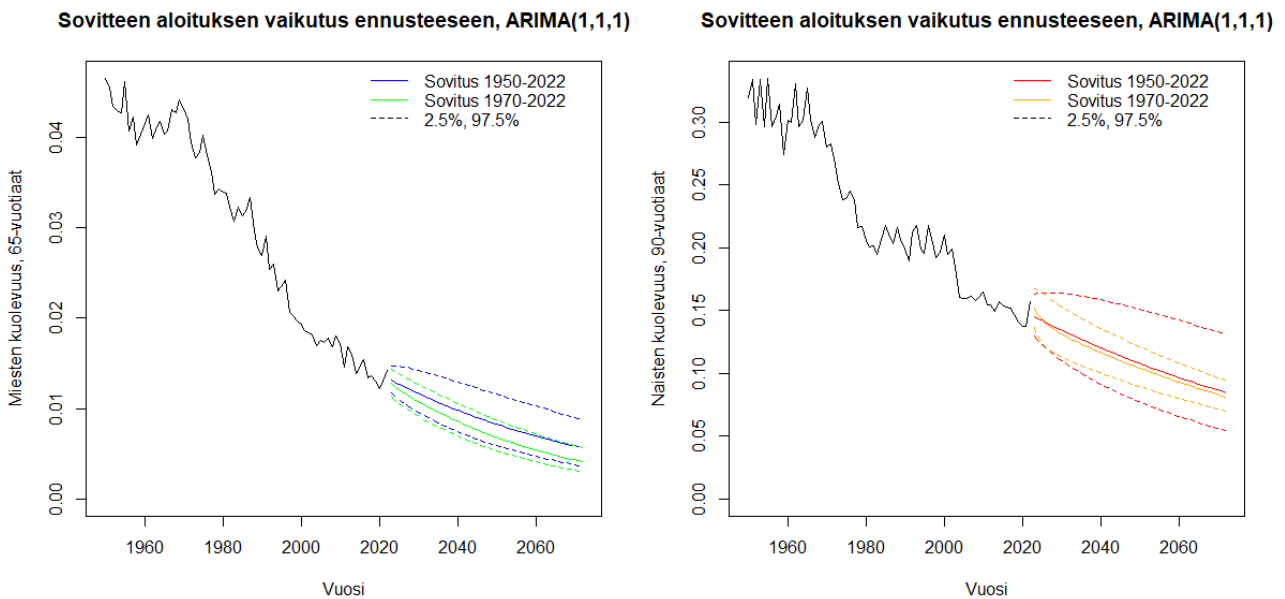
Tässä luvussa toteutetaan kuolevuuden ja elinajanodotteiden ennusteet luvussa 3.1 valituilla ARIMA(0,1,1)-, ARIMA(1,1,1)- ja ARIMA(2,1,2)-spesifikaatioilla. ARIMA-mallien spesifikaatioiden vertailun lisäksi tässä luvussa testataan mallien sovittamisessa käytettävän aikavälin vaikutusta kuolevuuden ennusteisiin. Käytännössä mallien sovittaminen toteutetaan R:n

¹¹<https://www.rdocumentation.org/packages/forecast/versions/8.21.1/topics/auto.arima>

¹²Yleisesti voidaan todeta, että ARIMA-mallit sopivat huonoiten juuri nuorempien (n. 30-vuotiaat ja nuoremmat) ikäluokkien kuolevuuksien mallintamiseen.

forecast-paketin Arima-funktiolla¹³ ja ennustaminen saman paketin forecast-funktiolla¹⁴.

Kuvassa 2 nähdään, miten ARIMA(1,1,1)-mallin sovittamisessa käytettävä aikaväli vaikuttaa kuolevuuden ennusteisiin 65-vuotiaiden miesten ja 90-vuotiaiden naisten tapauksissa. Erityisesti 65-vuotiaiden miesten tapauksessa havaitaan, että vuosille 1970—2022 sovitettu ARIMA-malli ennustaa alhaisempaa kuolevuutta kuin vuosille 1950—2022 sovitettu malli. Syy tähän eroon on vuosina 1970—1990 tapahtuvassa kuolevuuden jyrkässä pienenemisessä, jonka vaikutus korostuu 1970-vuodesta sovitetuissa malleissa (ts. trendiparametrissa c tulee voimakkaampi erityisesti miehillä). Toisaalta 1950-vuodesta sovitettujen mallien ennusteiden luottamusväleistä tulee leveämpiä kuin 1970-vuodesta sovitetuilla malleilla. Tämä johtuu siitä, että 50- ja 60-luvuilla kuolevuuden ARIMA-mallista poikkeava vaihtelu on ollut suurempaa kuin myöhempinä vuosikymmeninä, joten residuaalien keskihajontoihin pohjautuvista ennusteen luottamusväleistä tulee leveämpiä.



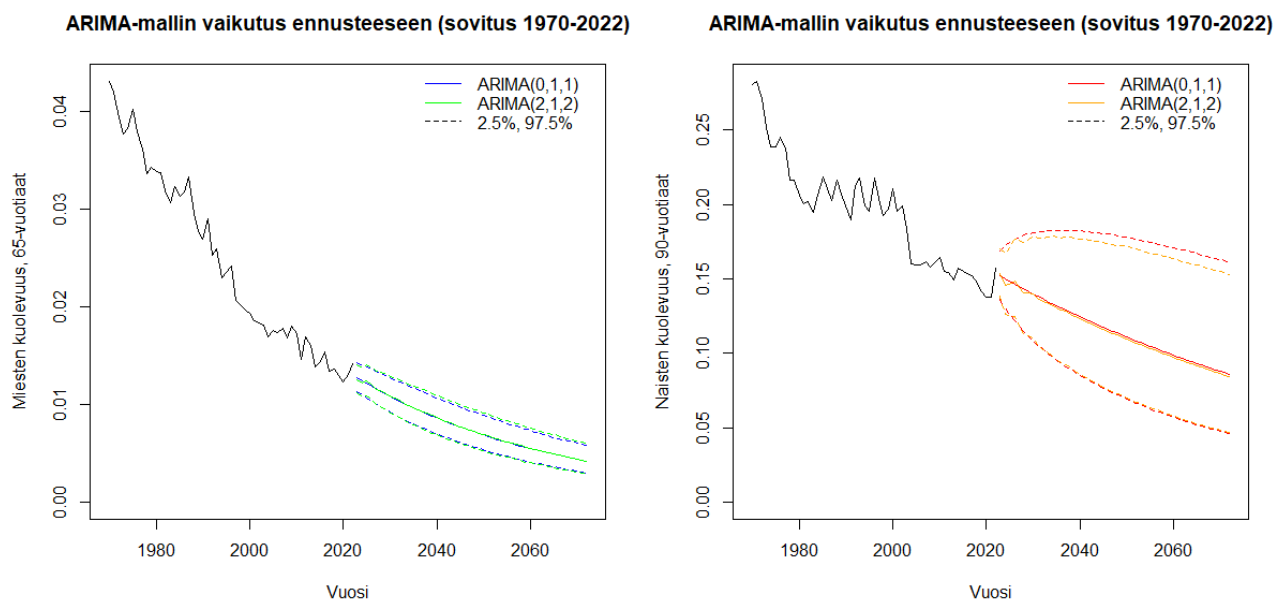
Kuva 2: ARIMA-mallin sovitteen aloitusvuoden vaikutus ennusteeseen, 65-vuotiaat miehet ja 90-vuotiaat naiset.

Kuvassa 3 esitetään ennusteet samoille 65-vuotiaiden miesten ja 90-vuotiaiden naisten aikasarjoille, mutta toteutettuna ARIMA(0,1,1)- ja ARIMA(2,1,2)-spesifikaatioilla. Kuvasta havaitaan, että spesifikaatioiden tuottamat kuolevuuden ennusteet ovat suhteellisen lähellä toisiaan, joskin erityisesti 90-vuotiaiden naisten tapauksessa ARIMA(2,1,2)-ennuste heiluu ennustehorisontin alussa hieman enemmän kuin ARIMA(0,1,1)-ennuste.¹⁵ Ennusteen heilunta noudattaa luvussa 2.3 ja liitteen F kuvassa 12 kuvattua mallien käyttäytymistä, jonka mukaan useasta menneestä havainnosta riippuvan ARIMA-spesifikaation ennusteella kestää pidempään saavuttaa vakioura kuin pienemmän ARIMA-spesifikaation ennusteella. ARIMA(2,1,2)-ennusteen

¹³<https://www.rdocumentation.org/packages/forecast/versions/8.21.1/topics/Arima>

¹⁴<https://www.rdocumentation.org/packages/forecast/versions/8.4/topics/forecast>

¹⁵Eri ARIMA-spesifikaatioiden ennusteiden luottamusväleissä vaikuttaa olevan enemmän eroja kuin ennusteiden odotusarvoissa (vertaa erityisesti 90-vuotiaiden naisten ARIMA(1,1,1)- ja ARIMA(0,1,1)-ennusteita). Tämä johtuu osaltaan siitä, että spesifikaatioiden sopivuus ikäkohtaisiin kuolevuusajaksarjoihin vaihtelee, jolloin myös eri spesifikaatioiden residuaaliajakasarjojen keskihajonnat vaihtelevat.



Kuva 3: ARIMA-mallin spesifikaation vaikutus ennusteeseen, 65-vuotiaat miehet ja 90-vuotiaat naiset.

heilunta on vielä selvemmin nähtävillä liitteen F kuvasta 19, jossa esitetään kuvan 3 mukaiset kuolevuuden aikasarjat ja ennusteet log-muutoksina.

Taulukoissa 1 ja 2 esitetään edellä kuvatuilla ARIMA-malleilla ennustetut miesten ja naisten elinajanodotteet vastasyntyneille (0-vuotiaat), 40-vuotiaille ja 80-vuotiaille vuosina 2023, 2050 ja 2070. Lisäksi liitteen F kuvissa 20 ja 21 näytetään ARIMA(0,1,1)- ja ARIMA(2,1,2)-spesifikaatioilla toteutetut elinajanodotteiden ennusteet ja luottamusvälit aikasarjoina. Koska koronapandemian aikana eli vuosina 2021 ja 2022 kuolevuus kasvoi merkittävästi erityisesti vanhemmissa ikäluokissa, taulukoissa 1 ja 2 esitetään vertailun vuoksi myös ennusteet, joissa pahimpia koronavuosia ei ole huomioitu. ARIMA-mallit on tällöin sovitettu vuosien 1950—2020 datalle.

Taulukoiden 1 ja 2 ennustetut elinajanodotteet antavat tukea kuvassa 2 havaitulle ilmiölle, jonka mukaan 70-luvulta alkaen sovitetut mallit ennustavat pidempää elinaikaa kuin 50-luvulta alkaen sovitetut mallit. Suurimmat erot havaitaan vuonna 2070 syntyneillä miehillä, joille erot mallien antamissa elinajanodotteiden ennusteissa ovat runsaan vuoden luokkaa. Naisilla erot 50- ja 70-luvulta sovitettujen mallien välillä ovat kauttaaltaan miehiä pienempiä. Koronavuosien poistaminen (eli mallien sovitus vuosille 1950—2020) odotetusti pidentää ennustettuja elinajanodotteita, sillä esimerkiksi vuonna 2070 syntyneiden miesten ja naisten elinajanodotteet ovat vajaan vuoden verran pidempiä kuin vuosille 1950—2022 sovitettujen mallien tapauksessa.

Taulukoiden 1 ja 2 perusteella eri ARIMA-spesifikaatioilla vaikuttaa olevan vähemmän vaikutusta ennustettuihin elinajanodotteisiin kuin mallin sovittamisessa käytettävällä aikavälillä, sillä ennusteiden väliset erot ovat suurimmillaankin vajaan kolmen kuukauden luokkaa. Tämä toisaalta tukee kuvasta 3 saatuja havaintoja, jonka mukaan eri ARIMA-spesifikaatioiden ennusteiden odotusarvot ovat melko samanlaisia (sen sijaan ennusteiden luottamusväleissä voi olla enemmän eroja). Myös liitteen F kuvissa 20 ja 21 esitettyjen elinajanodotteiden ennusteiden aikasarjojen perusteella havaitaan, etteivät ennusteet vaikuta olevan kovinkaan herkkiä

Ikä	Vuosi	ARIMA(0,1,1)			ARIMA(1,1,1)			ARIMA(2,1,2)		
		1950 - 2022	1970 - 2022	1950 - 2020	1950 - 2022	1970 - 2022	1950 - 2020	1950 - 2022	1970 - 2022	1950 - 2020
0	2023	79.32	79.44	79.72	79.35	79.46	79.71	79.36	79.44	79.73
0	2050	83.36	84.22	83.9	83.33	84.23	83.88	83.15	84.05	83.85
0	2070	85.86	87	86.49	85.83	87.02	86.47	85.64	86.84	86.44
40	2023	40.52	40.65	40.95	40.54	40.66	40.96	40.58	40.65	40.99
40	2050	44.07	44.96	44.65	44.05	44.97	44.65	43.87	44.79	44.61
40	2070	46.36	47.52	47.01	46.34	47.55	47.02	46.15	47.36	46.98
80	2023	8.32	8.35	8.54	8.35	8.38	8.54	8.35	8.36	8.55
80	2050	9.72	9.96	10.04	9.72	10	10.04	9.71	9.98	10.04
80	2070	10.74	11.06	11.14	10.74	11.11	11.14	10.73	11.09	11.14

Taulukko 1: Miesten ennustetut elinajanodotteet eri ARIMA-spesifikaatioilla.

Ikä	Vuosi	ARIMA(0,1,1)			ARIMA(1,1,1)			ARIMA(2,1,2)		
		1950 - 2022	1970 - 2022	1950 - 2020	1950 - 2022	1970 - 2022	1950 - 2020	1950 - 2022	1970 - 2022	1950 - 2020
0	2023	84.59	84.58	85.18	84.63	84.59	85.18	84.63	84.56	85.15
0	2050	88.18	88.37	88.88	88.14	88.36	88.88	88	88.21	88.87
0	2070	90.36	90.66	91.11	90.31	90.65	91.11	90.17	90.51	91.1
40	2023	45.22	45.21	45.79	45.26	45.23	45.79	45.25	45.2	45.77
40	2050	48.51	48.78	49.19	48.47	48.77	49.19	48.33	48.62	49.18
40	2070	50.57	50.96	51.3	50.52	50.95	51.31	50.38	50.81	51.29
80	2023	10.01	10.01	10.37	10.04	10.01	10.38	10.06	10	10.38
80	2050	11.73	11.94	12.2	11.7	11.96	12.21	11.6	11.84	12.21
80	2070	12.92	13.23	13.46	12.89	13.24	13.48	12.78	13.14	13.47

Taulukko 2: Naisten ennustetut elinajanodotteet eri ARIMA-spesifikaatioilla.

eri ARIMA-spesifikaatioille. Esimerkiksi ARIMA(2,1,2)-mallin ominainen ennusteen heilunta ei juurikaan näy elinajanodotteiden ennusteiden aikasarjoissa.

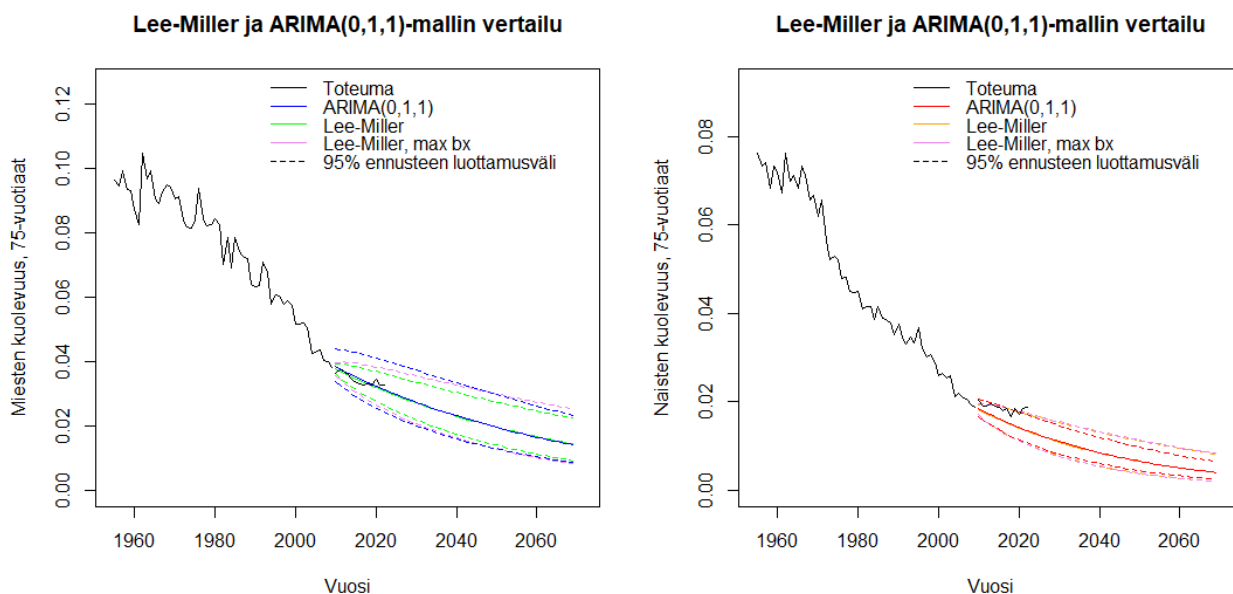
3.3 Vertailu Lee-Carter ja Lee-Miller -malleihin

Tässä työssä Lee-Carter ja Lee-Miller -malleja käytetään vertailukohtana ARIMA-malleille. Mallien vertaileminen toteutetaan siten, että mallit sovitetaan saman aikavälin aineistoihin (HMD-tietokannasta kerätty tasoitettu kuolevuusaineisto ja ikäväli 0—110+), jonka jälkeen mallien tuottamia kuolevuuden ja elinajanodotteiden ennusteita vertaillaan keskenään. Vertailua tehdään ennusteiden odotusarvojen ja 95 prosentin luottamusvälien osalta. Mallivertailujen tavoitteena on tarkastella eroja eri mallien tuottamien ennusteiden välillä. Lee-Carter ja Lee-Miller -mallien käytännön toteutus tehdään Sirénin (2013) [11] SHV-työn liitteessä H kuvattulla tavalla hyödyntäen Rob J. Hyndmanin demography-pakettia.¹⁶ Lee-Carter ja Lee-Miller -mallien lyhyet kuvaukset esitetään tämän työn liitteessä C.

Mallien tuottamien ennusteiden vertaileminen jaetaan kahteen osaan siten, että ensimmäisessä osassa vertaillaan vuosivälille 1955—2009 sovitettuja ARIMA- ja Lee-Miller -malleja, jolloin Lee-Miller -malli vastaa käytännössä henkivakuutuksen K2012-kuolevuustutkimuksen väestökuolevuusmallia (ks. Sirén (2013) [11]). Vuodesta 2010 alkavia ennusteita verrataan myös vuosien 2010—2022 toteutuneisiin kuolevuuksiin ja elinajanodotteisiin, jolloin voidaan arvioida

¹⁶<https://github.com/robjhyndman/demography>

ennusteiden tarkkuutta ensimmäisen kolmentoista vuoden ajalta.¹⁷ Jälkimmäisessä osassa verrataan vuosivälille 1960—2022 sovitettuja ARIMA-malleja vastaaviin Lee-Carter ja Lee-Miller -malleihin. Koska luvussa 3.2 nähtiin, että erot eri ARIMA-spesifikaatioiden ennusteissa ovat suhteellisen pieniä (luottamusvälejä lukuun ottamatta), käytetään ensimmäisessä vertailussa vain ARIMA(0,1,1)- ja jälkimmäisessä ARIMA(1,1,1)-spesifikaatiota.



Kuva 4: ARIMA(0,1,1)- ja Lee-Miller -mallien ennustamat kuolevuudet vuosivälin 2010—2022 toteumaan verrattuna, 75-vuotiaat naiset ja miehet. Mallit on sovitettu vuosivälille 1955—2009.

Kuvassa 4 esitetään ARIMA(0,1,1)-mallin ja Lee-Miller -mallin tuottamat kuolevuuden ennusteet sekä näiden 95 prosentin mukaiset luottamusvälit 75-vuotiaille miehille ja naisille. Lee-Miller -mallin luottamusvälit esitetään myös siten, että mallin ikäkohtaisen intensiteetin parametri b_x on vakioitu ikävälillä 20—110+ maksimiarvoon (ks. liite C tarkempien tietojen osalta). Silmämääräisesti kuvista nähdään, että odotetun kuolevuuden ennusteet ovat hyvin lähellä toisiaan, mutta ennusteiden luottamusväleissä on enemmän eroa. Erityisesti miesten tapauksessa ARIMA-ennusteen luottamusväli on Lee-Miller -mallin luottamusväliä leveämpi ennustehorisontin alkupäässä. Toisaalta ennustehorisontin kasvaessa Lee-Miller-ennusteen luottamusväli lähestyy ARIMA-ennusteen luottamusväliä. Naisten tapauksessa ARIMA-ennusteen luottamusväli on ennustehorisontin alussa hieman leveämpi, mutta jää nopeasti Lee-Miller -mallin luottamusväliä kapeammaksi horisontin kasvaessa. Sekä miesten että naisten tapauksessa Lee-Miller -mallin suurimpaan b_x -parametriin perustuva luottamusväli on lopulta levein pitkissä ennusteissa. Miesten tapauksessa vuosien 2010—2022 toteuma on ollut verrattain lähellä ennusteiden odotusarvoja, kun taas naisten tapauksessa kuolevuuden toteuma on ollut ennusteen luottamusvälin ylärajan tasolla.

¹⁷Tässä työssä esitettävät malli- ja toteumavertailut sisältävät rajoitteita. Vertailuja toteumaan suoritetaan vain ensimmäisen kolmentoista vuoden osalta ja yhdellä aikavälillä, joten tulokset eivät välttämättä sellaisenaan yleisty pidemmille ennusteille ja muille aikaväleille. Lisäksi ennusteiden luottamusvälejä arvioidaan vain kvalitatiivisesti, joten luottamusvälejä ei testata tarkemmin tilastollisesti. Sen sijaan ennusteiden odotusarvoja tarkastellaan tilastollisesti.

Liitteen F kuvissa 22, 23 ja 24 esitetään kuvaa 4 vastaavia vertailuja 30-, 50- ja 90-vuotiaiden miesten ja naisten kuolevuuksien ennusteille. Kuvien perusteella havaitaan, että näissä ikäluokissa myös ennusteiden odotusarvot eroavat hieman toisistaan. Lisäksi Lee-Miller -ennusteiden 95 prosentin luottamusvälien havaitaan edelleen olevan ARIMA-ennusteiden luottamusvälejä kapeampia erityisesti ennustehorisontin alkupäässä. Kuvien perusteella nähdään, että kuolevuuden toteuman suhteellinen vaihtelu on ollut huomattavan suurta erityisesti 30-vuotiaissa, jolloin toteutunut kuolevuus on ennusteen ensimmäisen kolmentoista vuoden aikana osunut hieman useammin ARIMA-ennusteiden leveämpien luottamusvälien sisään.¹⁸ Vastaavasti korona-aikana havaittu selvä nousu 90-vuotiaiden kuolevuudessa jää ARIMA-ennusteiden luottamusvälien sisään, joskin huomattavan kauas luottamusvälien ylärajoista, kun taas Lee-Miller -ennusteissa toteumat ylittävät niukasti luottamusvälien ylärajat. Erityisesti korona-ajan poikkeuksellisten havaintojen kohdalla on kuitenkin normaaliakin vaikeampaa arvioida, miten toteuman tulisi suhteutua ennusteisiin.¹⁹

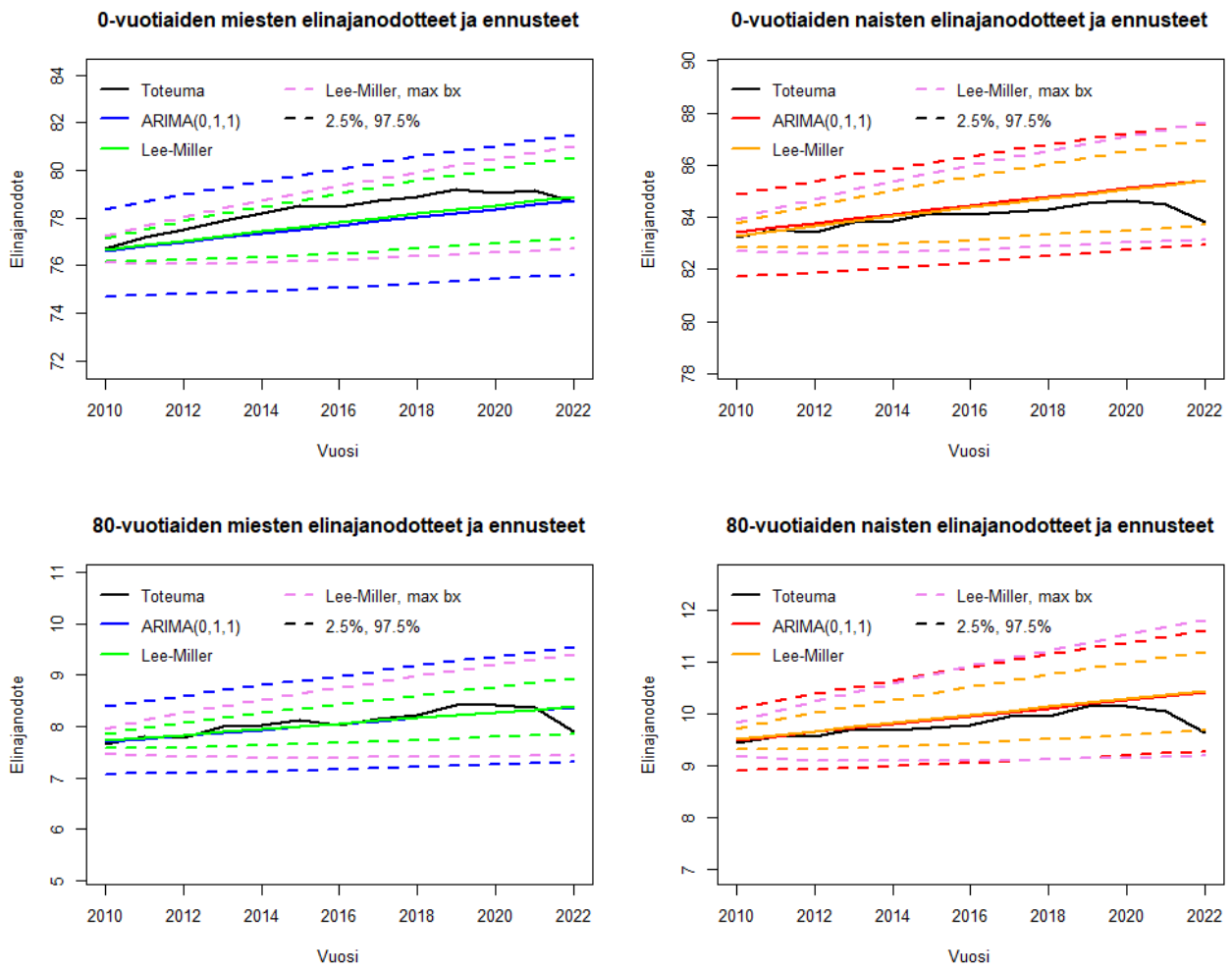
Tässä esitetyt luottamusvälien vertailut eivät sinällään vielä kerro mallien paremmuudesta mitään, mutta paljastavat kuitenkin mielenkiintoisia eroja mallien ennustejakaumissa. Sekä ARIMA- että Lee-Miller (ja Lee-Carter) -malleissa ikäkohtaisten log-kuolevuuksien ennusteiden luottamusvälit lasketaan normaalijakaumaan perustuen (ks. liite D), mutta mallit eroavat siinä, miten luottamusvälien laskennassa käytettävät ennustevirheet muodostetaan. Perinteisesti Lee-Miller (ja Lee-Carter) -mallin ennusteiden luottamusväleissä huomioidaan vain kaikille ikäluokille yhteisen aikafaktorin ennusteen epävarmuus (ks. liite C tarkempien tietojen osalta), sillä tämän vaikutuksen on näytetty dominoivan erityisesti pidemmissä ennusteissa.²⁰ Sen sijaan ARIMA-malleissa ennustevirheen keskihajonta perustuu suoraan ikäkohtaisten mallien ennustevirheisiin (ks. liite B). Nämä erot vaikuttavatkin johtavan siihen, että ARIMA-mallien luottamusväleistä muodostuu usein Lee-Miller -ennusteen luottamusvälejä leveämpiä etenkin ennustehorisontin alkupäässä.

Kuvassa 5 esitetään ARIMA(0,1,1)-mallin ja Lee-Miller -mallin ennustamat elinajanodotteet sekä näiden 95 prosentin mukaiset ennusteen luottamusvälit vastasyntyneille sekä 80-vuotiaille miehille ja naisille vuosilta 2010—2022. Kuvasta nähdään, että elinajanodotteiden ennusteiden odotusarvot ovat hyvin lähellä toisiaan, mutta ARIMA-ennusteiden luottamusvälit ovat tässäkin Lee-Miller -ennusteiden luottamusvälejä leveämpiä erityisesti ennustehorisontin alkupäässä. Naisten tapauksessa suurimman b_x -parametrin mukaiset luottamusvälit ovat ennustehorisontin alkupäässä ARIMA-mallia kapeampia, mutta levenevät muutaman vuoden aikana lähelle tai yli ARIMA-ennusteiden luottamusvälien. Miesten tapauksessa ARIMA-ennusteen luottamusvälit ovat leveimpiä koko esitetyn ennustehorisontin ($h = 13$) ajalta.

¹⁸Tässä 30-vuotiaiden tapauksessa ARIMA-ennusteiden luottamusvälien leveyteen vaikuttaa todennäköisesti myös se, ettei käytetty ARIMA-malli itsessään sovellu parhaalla tavalla näiden kuolevuuksien mallintamiseen (eli ennustevirhe kasvaa).

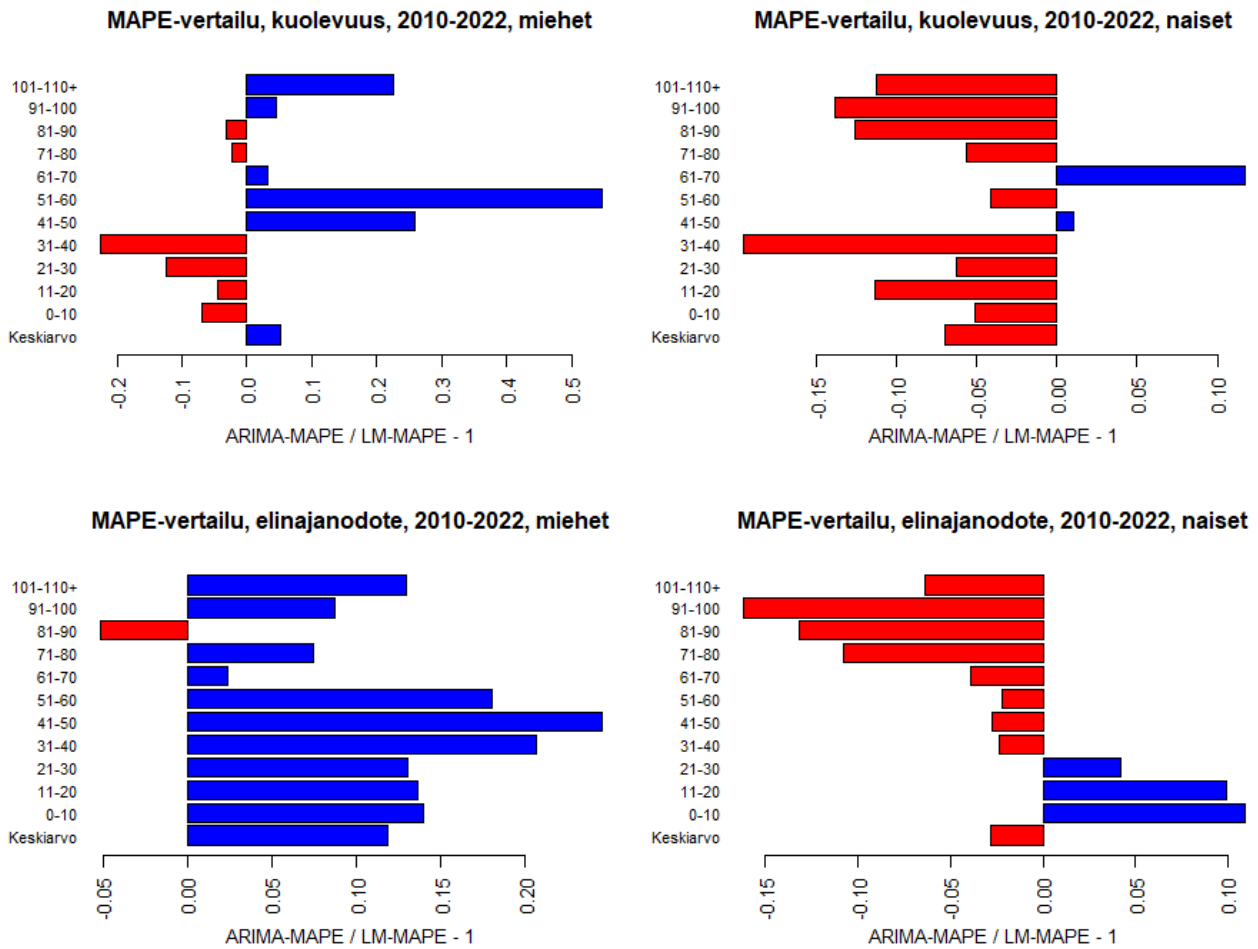
¹⁹Voidaan esimerkiksi kysyä, että miten ja missä määrin tässä tarkasteltujen tilastollisten mallien tulisi osata ennakoida koronapandemian kaltaisia tilanteita, sillä vastaavaa ei ole tapahtunut mallien sovitusaikaväleillä.

²⁰Lee ja Carter (1992) [8] (670, taulukko B2) arvioivat, että Lee-Carter -mallin aikafaktorin k_t ennusteen keskihajonta kattaa runsaan kolmanneksen log-kuolevuuden yhden vuoden ennusteen kokonaiskeskihajonnasta, kun taas yli kahdenkymmenen vuoden ennusteissa noin yhdeksänkymmentä prosessia keskihajonnasta syntyy aikafaktorin ennusteesta.



Kuva 5: ARIMA(0,1,1)- ja Lee-Miller -mallien ennustamat elinajanodotteet vuosivälin 2010—2022 toteumaan verrattuna, 0-vuotiaat sekä 80-vuotiaat miehet ja naiset. Mallit on sovitettu vuosivälille 1955—2009.

Elinajanodotteiden ennusteiden kuvaajia esitetään pidemmältä aikaväliltä liitteen F kuvassa 25 0-, 20- ja 40-vuotiaille. Tässä kuvassa sekä kuvassa 5 näkyy selvästi korona-aikana (vuosina 2021 ja 2022) tapahtuva elinajanodotteiden huomattava laskeminen. Esimerkiksi kuvan 5 perusteella vastasyntyneiden naisten sekä 80-vuotiaiden miesten ja naisten elinajanodotteet ovat korona-aikana saavuttaneet Lee-Miller -ennusteiden 95 prosentin luottamusvälien alarajat. Toisaalta korona-ajan toteumat jäävät selvästi ARIMA-ennusteiden 95 prosentin luottamusvälien sisään, mutta kvalitatiivisesti arvioiden vielä suhteellisen kauas luottamusvälien alarajoista. Koska koronapandemia on kuitenkin mallien sovitusaikaväliin nähden poikkeuksellinen tapahtuma, on pandemiavuosien elinajanodotteiden arvioiminen suhteessa ennustejakaumiin huomattavan haastavaa. Lisäksi erityisesti ARIMA-mallien luottamusvälien muodostamisessa käytetään karkeita oletuksia (ks. liite D), mikä osaltaan vaikuttaa myös luottamusvälien tarkkuuteen.



Kuva 6: ARIMA(0,1,1)- ja Lee-Miller -mallien MAPE-arvojen vertailu. Mallit on sovitettu vuosivälille 1955—2009.

ARIMA- ja Lee-Miller -ennusteiden (odotusarvojen) tarkkuutta toteutuneisiin kuolevuuksiin ja elinajanodotteisiin nähden voidaan arvioida MAPE-mitalla (*Mean Absolute Percentage Error*), joka lasketaan ikäkohtaisesti kaavalla

$$MAPE_x = \frac{1}{H} \sum_{h=1}^H \left| \frac{\hat{y}_{x,T+h|T} - y_{x,T+h}}{y_{x,T+h}} \right|, \quad (19)$$

jossa $y_{x,t}$ kuvaaja ikäkohtaista kuolevuutta tai elinajanodotetta. Mitä pienempi ennusteen MAPE-arvo on, sitä tarkempi ennuste on ollut. Kuvassa 6 esitetään edellä kuvattujen Lee-Miller ja ARIMA(0,1,1)-mallien MAPE-arvojen välisistä suhteista kymmenen ikävuoden välein lasketut keskiarvot vuosien 2010—2022 ajalta ($T = 2009$ ja $H = 13$). Tulosten perusteella Lee-Miller -mallin ennuste on ollut miehillä keskimäärin ARIMA-mallia tarkempi etenkin elinajanodotteissa. Sen sijaan naisilla ARIMA-malli on ollut keskimäärin tarkempi etenkin kuolevuuden ennusteissa.

Lopuksi vertaillaan AR-osan sisältävää ARIMA(1,1,1)-mallia Lee-Miller ja Lee-Carter -malleihin siten, että mallit sovitetaan vuosivälille 1960—2022. Liitteen F kuvasta 26 nähdään, että mallien ennusteet reagoivat eri tavoilla 90-vuotiaiden koronapandemian aikaiseen kuolevuuden nousuun. ARIMA-ennusteet palautuvat (epälineaarisesti) vuoden 2022 kohonneelta tasolta lähelle mallin sovitusaikavälin mukaista kuolevuuden kehityksen keskimääräistä uraa. Hieman vastaavasti Lee-Carter -mallin (lineaarinen) ennuste asettuu etenkin miesten tapauksessa vuoden 2022 havaittua kuolevuutta alhaisemmalle uralle, sillä mallin ennuste lähtee viimeisestä sovitetun kuolevuuden pisteestä (ks. liite C ja ns. *jump-off* -virhe). Lee-Miller -mallin (lineaarinen) ennuste pysyy puolestaan korona-ajan mukaisella kohonneella tasolla, sillä mallin ennuste lähtee viimeisestä havaitusta kuolevuuden pisteestä. Luonnollisesti vasta kuolevuuden tuleva kehitys tulee näyttämään, mikä ennusteista osuu lopulta parhaiten oikeaan.²¹

Taulukoissa 3 ja 4 vertaillaan ARIMA(1,1,1)-, Lee-Miller ja Lee-Carter -mallien ennustamia elinajanodotteita sekä ennusteiden 95 prosentin luottamusvälejä (ennusteiden kuvajia pidemmältä ajalta näkee myös liitteen F kuvasta 27 0-, 20- ja 40-vuotiaille). Taulukoiden perusteella ARIMA-malli ennustaa 2060-vuonna syntyneille miehille noin puolen vuoden verran ja naisille vajaan vuoden verran muita malleja pidempiä elinajanodotteita, kun taas Lee-Miller ja Lee-Carter -ennusteiden väliset erot jäävät tätä pienemmiksi. Ennusteiden 95 prosentin luottamusvälejä vertailemalla havaitaan, että Lee-Miller -ennusteilla on selvästi kapeimmat luottamusvälit. ARIMA-ennusteen luottamusvälit ovat puolestaan muita malleja leveämpiä ennusteen ensimmäisinä vuosina, mutta horisontin pidentyessä Lee-Carter -ennusteen luottamusvälit muuttuvat ARIMA-ennusteita leveämmiksi. Näiden vertailujen perusteella ARIMA-ennusteiden luottamusvälit vaikuttavatkin asettuvan suhteellisen järkevälle tasolle Lee-Miller ja Lee-Carter -mallien väliin.

Tässä työssä esiteltävät ARIMA-mallit ovat ikäkohtaisia, jolloin ikäkohtaisten kuolevuusai-
 kasarjojen välisiä riippuvuuksia ei (suoraan) mallinneta. Valittu lähestymistapa on käytännöllinen mutta puutteellinen, sillä ikäkohtaisten kuolevuuksien välillä havaitaan riippuvuutta. Esimerkiksi liitteen F kuvista 28 ja 29 nähdään, että ikäkohtaisten log-kuolevuuksien muutosten väliset korrelaatiot ovat olleet voimakkaita ja positiivisia juuri vanhimpien ikäluokkien välillä (n. 80-vuotiaille ja tätä vanhemmille), kun taas nuoremmille ikäluokille korrelaatiot ovat olleet heikompia. Tulevaisuudessa ARIMA-mallikehikkoa tulisikin kehittää niin, että riippuvuusmallinnus otetaan käyttöön vähintäänkin vanhempien ikäluokkien välille esimerkiksi moniulotteisia ARMA-malleja (ks. Brockwell ja Davis (2002) [3], luku 7) tai copula-riippuvuusrakenteita (ks. esimerkiksi Lin ym. (2015) [10]) soveltamalla. Huomioiden tässä työssä käytettävät karkeat oletukset ARIMA-mallien elinajanodotteiden ennusteiden luottamusväleihin liittyen (ks. liite D), riippuvuusrakenteiden tarkempi mallintaminen johtaisi todennäköisesti myös elinajanodotteiden ennusteiden luottamusvälien kapenemiseen (toki lopputulos riippuu paljolti myös sovellettavasta riippuvuusmallista).²²

²¹Ennusteiden tulevan (erityisesti lyhyen aikavälin) tarkkuuden näkökulmasta ratkaiseva tekijä on, että jääkö pandemian aikainen piikki kuolevuudessa hetkelliseksi vai pysyvämmäksi ilmiöksi. Tätä työtä kirjoitettaessa ei ollut vielä tarkkaa tietoa siitä, miten kuolevuus tulee kehittymään koronapandemian jälkeen. Kuitenkin esimerkiksi Leen ja Millerin (2001) [9] tutkimuksessa on havaittu, että ns. "normaaliaikoina" ennusteen sisältämä *jump-off* -virhe on heikentänyt ennusteen tarkkuutta.

²²Lin ym. (2015) [10] mallinsivat kuolevuuden ikäkohtaisten AR-mallien välistä riippuvuutta copularakenteilla. Mielenkiintoisesti he havaitsivat, että yksinkertaisin eli staattinen Gaussinen copula sopi heidän tarkastelemaan aineistoon parhaiten. Lisäksi he havaitsivat, että yksinkertaisenkin riippuvuusrakenteen lisääminen AR-malliin pienensi pitkäikäisyysriskin VaR-mittaa.

Sovitus: 1960 - 2022

Ikä	Vuosi	ARIMA(1,1,1)			Lee-Miller			Lee-Carter		
		2.5%	Ennuste	97.5%	2.5%	Ennuste	97.5%	2.5%	Ennuste	97.5%
0	2023	77.61	79.33	80.86	78.46	78.87	79.27	78.05	78.93	79.79
0	2040	79.09	82.05	84.59	79.77	81.57	83.21	77.73	81.74	85.07
0	2060	81.29	84.84	87.74	81.75	84.32	86.53	78.67	84.59	88.92
40	2023	39.31	40.54	41.74	39.74	40.11	40.47	39.42	40.21	40.97
40	2040	40.56	42.96	45.18	40.91	42.52	44.02	39.14	42.73	45.79
40	2060	42.52	45.50	48.15	42.69	45.04	47.08	39.97	45.35	49.37
80	2023	7.72	8.32	8.94	7.83	7.96	8.08	7.98	8.25	8.52
80	2040	8.20	9.24	10.34	8.24	8.84	9.42	7.89	9.17	10.41
80	2060	8.92	10.30	11.73	8.90	9.85	10.76	8.17	10.22	12.11



Taulukko 3: ARIMA(1,1,1)-, Lee-Miller ja Lee-Carter -mallien ennustamat elinajanodotteet, miehet.

Sovitus: 1960 - 2022

Ikä	Vuosi	ARIMA(1,1,1)			Lee-Miller			Lee-Carter		
		2.5%	Ennuste	97.5%	2.5%	Ennuste	97.5%	2.5%	Ennuste	97.5%
0	2023	83.02	84.62	85.95	83.52	83.97	84.41	83.08	83.96	84.81
0	2040	84.54	87.04	89.21	84.39	86.36	88.12	82.33	86.36	89.64
0	2060	86.62	89.52	92.02	85.95	88.79	91.13	82.77	88.82	93.02
40	2023	44.10	45.25	46.33	44.29	44.70	45.10	43.86	44.67	45.45
40	2040	45.38	47.50	49.47	45.08	46.91	48.57	43.18	46.89	49.99
40	2060	47.24	49.84	52.20	46.53	49.20	51.42	43.59	49.20	53.22
80	2023	9.48	10.03	10.59	9.51	9.71	9.92	9.29	9.70	10.10
80	2040	10.10	11.25	12.43	9.90	10.86	11.77	8.95	10.87	12.63
80	2060	11.05	12.58	14.16	10.66	12.13	13.46	9.15	12.17	14.70



Taulukko 4: ARIMA(1,1,1)-, Lee-Miller ja Lee-Carter -mallien ennustamat elinajanodotteet, naiset.

4 Yhteenveto

Tämän SHV-työn tavoitteena on tarkastella ARIMA-mallien soveltumista suomalaisten miesten ja naisten kuolevuuden ja elinajanodotteiden ennustamiseen. Lisäksi tässä työssä vertaillaan ARIMA-mallien tuottamia ennusteita Lee-Carter ja Lee-Miller -mallien tuottamiin ennusteisiin. Muutamissa viimeaikaisissa akateemisissa tutkimuksissa on havaittu, että yksiulotteisilla (nollan faktorin) aikasarjamalleilla toteutetut kuolevuuden ennusteet pärjäävät toteumatesteissä paremmin kuin Lee-Carter -malleilla toteutetut kuolevuuden ennusteet. Tämän havainnon jatkotarkastelu toimiikin pääasiallisena motivaationa tämän tutkimuksen toteuttamiselle.

Aikasarja-analyysien perusteella suomalaisten miesten ja naisten ikäkohtaisia log-kuolevuuden aikasarjoja voidaan mallintaa ARIMA(p,1,q)-malleilla. Informaatiokriteerien perusteella parhaiten aikasarjoihin sopivat yksinkertaiset ARIMA-spesifikaatiot, kuten ARIMA(0,1,1) tai ARIMA(1,1,1). Eri ARIMA-spesifikaatioiden tuottamia kuolevuuden ja elinajanodotteiden ennusteita vertailemalla havaittiin, että ennusteiden odotusarvojen väliset erot ovat pääsääntöisesti pieniä, mutta ennusteiden luottamusväleissä voi olla hieman enemmän eroja. Valittua ARIMA-spesifikaatiota enemmän kuolevuuden ennusteisiin näyttääkin vaikuttavan mallien sovitusaikavälillä vallitseva kuolevuuden trendi.

ARIMA(0,1,1)-mallia verrattiin Lee-Miller -malliin siten, että mallit sovitettiin vuosien 1955—2009 aineistoon, jolloin molempien mallien ennusteita voitiin verrata vuosien 2010—2022 vertailuajavälillä toteumaan. ARIMA-mallilla toteutettujen ennusteiden (odotusarvojen) havaittiin olleen vertailuajavälillä keskimäärin naisille tarkempia ja Lee-Miller -ennusteiden miehille tarkempia. Yleisesti ARIMA-ennusteiden (95 prosentin) luottamusvälien havaittiin olevan erityisesti ennustehorisontin alkupäässä Lee-Miller -ennusteiden luottamusvälejä leveämpiä. Tämän arvioitiin johtuvan mallien ennustevirheiden erilaisista oletuksista. Kvalitatiivisesti arvioiden ARIMA-mallien leveämmät luottamusvälit kattoivat esimerkiksi korona-ajan poikkeukselliset havainnot Lee-Miller -mallia useammin, joskin samalla ARIMA-ennusteiden luottamusvälit vaikuttivat paikoin jo liiankin leveiltä.

ARIMA(1,1,1)-mallia verrattiin Lee-Miller ja Lee-Carter -malleihin siten, että mallit sovitettiin vuosien 1960—2022 aineistoon. Tällöin ARIMA-mallit ennustivat, että esimerkiksi 90-vuotiaiden kuolevuudet palautuisivat muutamassa vuodessa korona-ajan kohonneelta tasolta takaisin sovitusaikavälillä mukaiselle keskimääräiselle uralle. Sen sijaan vastaavat Lee-Miller -ennusteet pysyivät korona-ajan kohonneella tasolla ja Lee-Carter -ennusteet sijoituivat karkeasti edeltävien ennusteiden väliin. Lopulta ARIMA-mallit ennustivat esimerkiksi vuonna 2060 syntyville vajaan vuoden verran pidempiä elinajanodotteita kuin Lee-Carter tai Lee-Miller -mallit. Leveytensä puolesta ARIMA-ennusteiden luottamusvälit asettuivat Lee-Carter ja Lee-Miller -mallien väliin.

Tämän työn tulosten perusteella yksinkertaiset ARIMA-mallit vaikuttavat sopivan hyvin suomalaisten miesten ja naisten kuolevuuksien ja elinajanodotteiden ennustamiseen, mutta mallikehikossa on yhä myös kysymysmerkkejä ja kehitettävää. ARIMA-mallien suorituskyvyn johdonmukainen arvioiminen vaatisi mallien testaamista eri ajankohtina sekä pidemmällä aikavälillä. Lisäksi ennusteiden luottamusvälien arvioiminen vaatisi tarkempaa tilastollista testaamista. Teknisesti mallikehikkoa voitaisiin kehittää esimerkiksi lisäämällä ikäkohtaisten ARIMA-mallien välille tarkempaa riippuvuusrakenteiden mallintamista.

Lähteet

- [1] Bell, W. (1997). Comparing and assessing time series methods for forecasting age-specific fertility and mortality rates. *Journal of Official Statistics*, Vol. 13 (3), 279–303.
- [2] Booth, H. – Tickle, L. (2008) Mortality Modelling and Forecasting: a Review of Methods. *Annals of Actuarial Science*, Vol. 3, 3–43.
- [3] Brockwell, P. J. – Davis, R. A. (2002) *Introduction to Time Series and Forecasting*. 2. p. Springer, New York.
- [4] Carter, L. – Lee, R. (1992) Modeling and forecasting US sex differentials in mortality. *International Journal of Forecasting*, Vol. 8, 393–411.
- [5] Giacometti R. – Bertocchi, M. – Rachev, S. T. – Fabozzi, F. J. (2012) A comparison of the Lee-Carter model and AR–ARCH model for forecasting mortality rates. *Insurance: Mathematics and Economics*, Vol. 50, 85–93.
- [6] Hyndman, R. J. – Athanasopoulos, G. (2021) *Forecasting: principles and practice*. 3. p. OTexts, Melbourne, Australia. <OTexts.com/fpp3>, haettu 15. maaliskuuta 2024.
- [7] Kwiatkowski, D. – Phillips, P. C. B. – Schmidt, P. – Shin, Y. (1992) Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, Vol. 54 (1–3), 159–178.
- [8] Lee, R. – Carter, L. (1992) Modelling and forecasting U.S. mortality. *Journal of American Statistical Association*, Vol. 87, 659–671.
- [9] Lee, R. – Miller, T. (2001) Evaluating the performance of the Lee-Carter method for forecasting mortality. *Demography*, Vol. 38: 537–549.
- [10] Lin, T. – Wang, C.-W. – Tsai, C. C.-L. (2015) Age-specific copula-AR-GARCH mortality models. *Insurance: Mathematics and Economics*, Vol. 61, 110–124.
- [11] Sirén, T. (2013) *Henkivakuutusyhtiöiden kuolevuustutkimus K2012*. SHV-työ, Suomen Aktuaariyhdistys. <https://www.actuary.fi/julkaisut/shv-tyot/henkivakuutusyhtioiden-kuolevuustutkimus-k2012/SHV_115_Siren_Tarja.pdf>, haettu 15. maaliskuuta 2024.
- [12] Syuhada, K. – Hakim, A. (2021) Stochastic modeling of mortality rates and Mortality-at-Risk forecast by taking conditional heteroscedasticity effect into account. *Heliyon*, Vol. 7.
- [13] Wilmoth J.R. – Andreev K. – Jdanov D. – Gleij D.A. – Riffe, T. ja avustajat Boe C. – Bubenheim M. – Philipov D. – Shkolnikov V. – Vachon P. – Winant C. – Barbier M. (2021) *Methods Protocol for the Human Mortality Database*, 6. versio.

<<https://www.mortality.org/File/GetDocument/Public/Docs/MethodsProtocolV6.pdf>>, haettu 15. maaliskuuta 2024.

A Kuolevuusaineisto ja elinajanodotteiden laskeminen

Kuolevuusennusteiden pohjalla käytetty Suomen periodikuolevuusaineisto on kerätty *Human Mortality Database* -tietokannasta (lyhyesti HMD). Täten aineiston muuttujat on määritelty HMD:n metodiprotokollan mukaisesti (ks. Wilmoth ym. (2021) [13], luku 6).

- $P(x, t)$ on populaatio eli vuonna t (tammikuun 1. päivänä) elossa olevat x -ikäiset.
- $D_L(x, t)$ on vuonna t niiden x -ikäisten kuolleiden lukumäärä, jotka saavuttivat iän x vuonna t .
- $D_U(x, t)$ on vuonna t niiden x -ikäisten kuolleiden lukumäärä, jotka saavuttivat iän x vuonna $t - 1$.
- $D(x, t) = D_L(x, t) + D_U(x, t)$ on x -ikäisten kuolleiden lukumäärä vuonna t .
- $E(x, t) = \frac{1}{2}[P(x, t) + P(x, t+1)] + \frac{1}{6}[D_L(x, t) - D_U(x, t)]$ on x -ikäisten eletyt vuodet vuonna t eli altistuma (engl. *exposure*). Tätä kaavaa käytetään, mikäli tietoa syntymääjankohdista ei ole HMD:n aineistossa saatavilla kuukausitasolla. Kuukausitasoisen datan tapauksessa altistuma lasketaan HMD:n metodiprotokollan kaavojen 50—56 mukaisesti.
- $m_{x,t} = \frac{D(x,t)}{E(x,t)}$ on x -ikäisten keski-kuolevuus vuonna t .

Tässä työssä kuolevuusaineistona käytetään HMD:n tasoitettua kuolevuusaineistoa (ns. *Life tables*), jossa erityisesti vanhojen ikäluokkien kuolevuuksia on tasoitettu datan rajallisuuden takia. Näitä tasointumenetelmiä kuvataan tarkemmin HMD:n metodiprotokollan [13] luvussa 7.

Elinajanodotteet määritellään periodiaineiston pohjalta (ns. periodielinajanodote). Elinajanodotteiden laskenta perustuu oletukseen, että kuolinhetket vuoden aikana ovat tasajakautuneita. Tällöin yhden vuoden selviämistodennäköisyyden aktuaariestimaattori x -ikäiselle vuonna t saadaan keski-kuolevuudesta $m_{x,t}$ kaavalla

$$p_{x,t} = \max\left(0, \frac{2 - m_{x,t}}{2 + m_{x,t}}\right). \quad (20)$$

Elinajanodote x -ikäiselle vuonna t saadaan puolestaan kaavalla

$$e_{x,t} = \frac{1}{2} + \sum_{k=1}^{\infty} {}_k p_{x,t}, \quad (21)$$

jossa

$${}_k p_{x,t} = \prod_{n=0}^{k-1} p_{x+n,t}. \quad (22)$$

B ARIMA-mallien ennusteiden luottamusvälit

Ennusteen luottamusvälien kuvaaminen aloitetaan esittämällä ARMA-prosessin ennustevirhe, jonka arvioimisessa olennaisessa osassa on ARMA-prosessin muuntaminen päättymättömäksi summaksi eli MA(∞)-prosessiksi. Mikäli $y_t \sim ARMA(p, q)$ täyttää stationaarisuuden ehdot, sille pätee, että

$$y_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}, \quad (23)$$

jossa $\psi_0 = 1$ (ks. Brockwell ja Davis (2002) [3] luku 3.1 ARMA-mallin stationaarisuuden ehtoihin ja MA(∞)-esitykseen liittyen). Tällöin ennusteelle $\hat{y}_{T+h|T}$ pätee

$$\hat{y}_{T+h|T} = \sum_{j=h}^{\infty} \psi_j \epsilon_{t+h-j}, \quad (24)$$

jolloin puolestaan ennustevirheelle $y_{T+h} - \hat{y}_{T+h|T}$ pätee

$$y_{T+h} - \hat{y}_{T+h|T} = \sum_{j=0}^{\infty} \psi_j \epsilon_{t+h-j} - \sum_{j=h}^{\infty} \psi_j \epsilon_{t+h-j} = \sum_{j=0}^{h-1} \psi_j \epsilon_{t+h-j}. \quad (25)$$

Ennustevirheen odotusarvo on nolla, jolloin sen varianssi saadaan yksinkertaisesti muotoon

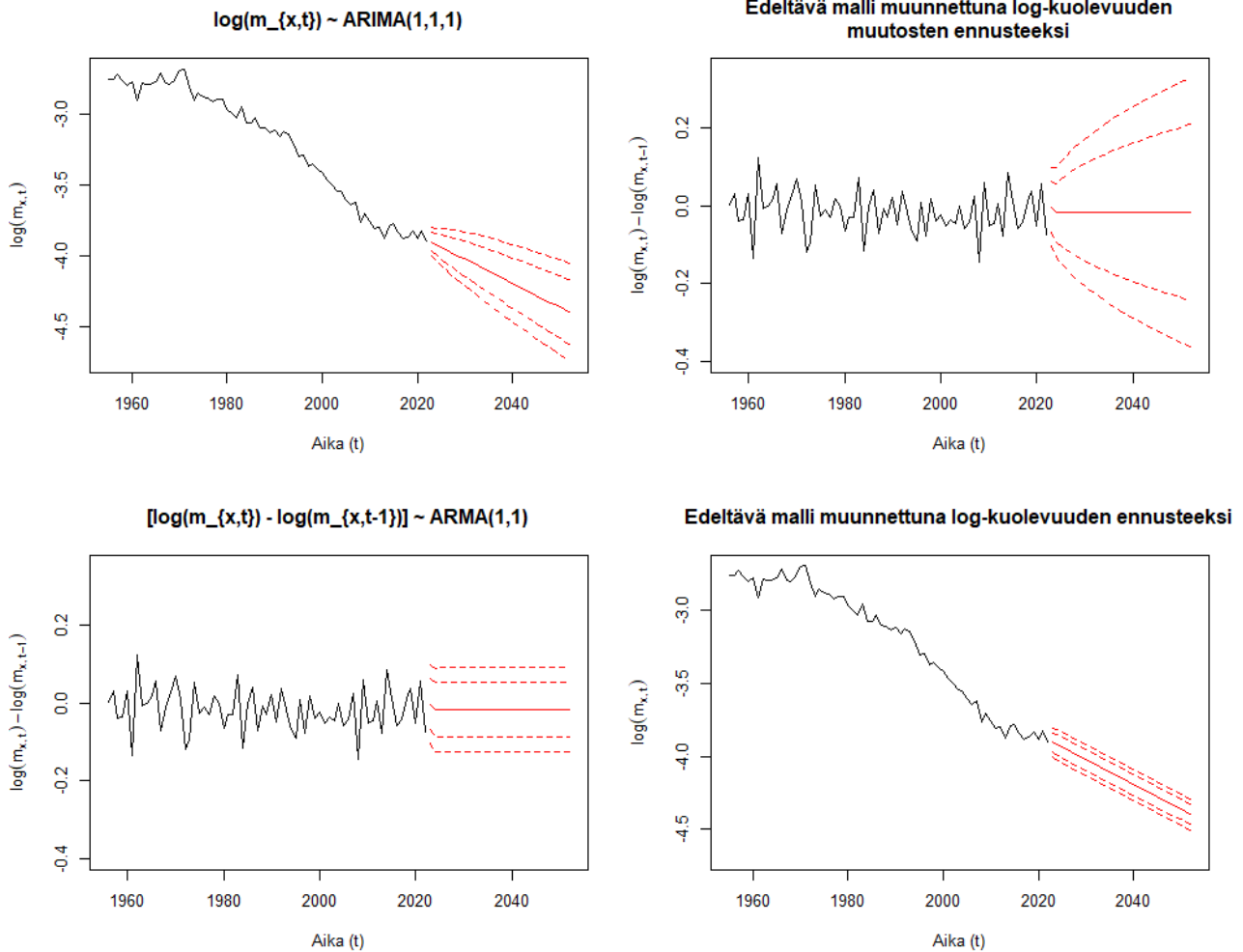
$$\sigma_h^2 = Var[y_{T+h} - \hat{y}_{T+h|T}] = \sigma^2 \sum_{j=0}^{h-1} \psi_j^2. \quad (26)$$

Tällöin normaalijakaumaoletuksen alla, ts. $\epsilon_t \sim nid(0, \sigma^2)$, ennusteen luottamusväli saadaan kaavalla $\hat{y}_{T+h|T} \pm N_{1-\alpha/2} \sigma_h$.²³ Ennustevirheestä lasketun luottamusvälin osalta on hyvä huomata, ettei tässä huomioida mallin parametreihin liittyvää epävarmuutta.

Käytännössä laskentakaava ennusteen luottamusvälille on helppo esittää vielä ensimmäisen periodin ennusteelle ($h = 1$), sillä luottamusväli on muotoa $\hat{y}_{T+1|T} \pm \hat{\sigma} N_{1-\alpha/2}$, jossa $\hat{\sigma}$ on havaitusta residuaaliprosessista estimoitu keskihajonta. Toisesta ennusteperiodista alkaen luottamusvälin tarkan laskentakaavan esittäminen vaatii summan $\sum_{j=0}^{h-1} \psi_j^2$ avaamista, jonka kaava puolestaan riippuu ARMA-mallin spesifikaatioista (ts. p :n ja q :n arvoista). Käytännössä luottamusvälien käyttäytymisestä voidaan todeta, että ARMA(p, q)-mallin ennusteen luottamusvälit konvergoituvat saman suuruiseksi ennustehorisontin h kasvaessa tarpeeksi suureksi (ks. kuva 7).

ARIMA($p, 1, q$)-mallin ennusteen luottamusväli on ensimmäisen periodin osalta sama kuin ARMA-mallille, ts. $\hat{y}_{T+1|T} \pm \hat{\sigma} N_{1-\alpha/2}$. Sen sijaan ennusteperiodin ollessa tätä suurempi ($h \geq 2$) luottamusväleissä täytyy huomioida prosessin epästationaarisuus, jolloin myös luottamusvälien analyttinen tarkastelu vaikeutuu. Käytännössä ARIMA($p, 1, q$)-prosessin epästationaarisuudesta johtuen ennusteiden luottamusvälit levenevät ennusteperiodin h kasvaessa. Tämä on myös suurin syy sille, että log-kuolevuuden ennustaminen toteutetaan nimenomaan prosessille $\log(m_{x,t}) \sim ARIMA(p, 1, q)$ eikä muutosprosessille $\Delta \log(m_{x,t}) \sim ARMA(p, q)$. Kuvassa 7 esitetään ARIMA(1,1,1)- ja ARMA(1,1)-ennusteiden luottamusvälien eroja käyttäen 70-vuotiaiden miesten log-kuolevuuden aikasarjaa esimerkkinä. ARIMA-mallin ennusteen luottamusvälien tarkemman kuvaamisen osalta lukijaa kehoitetaan tutustumaan Brockwellin ja Davisin (2002) [3] kirjan lukuun 6.4.

²³Kuolevuuden ennusteen luottamusväli on siis yksinkertaisesti muotoa $\hat{m}_{x,T+h|T} \exp(\pm N_{1-\alpha/2} \sigma_h)$.



Kuva 7: ARIMA(1,1,1)- (ylemmät kuvat) ja ARMA(1,1)-malleilla (alemmat kuvat) toteutetut log-kuolevuuden ennusteet sekä 80 prosentin ja 95 prosentin luottamusvälit. Ylempistä kuvista nähdään, että ARIMA-ennusteen ($\log(m_{x,t}) \sim ARIMA(1,1,1)$) luottamusväli levenee ennustehorisontin h kasvaessa. Alemmista kuvista nähdään, että ARMA-ennusteen ($\Delta \log(m_{x,t}) \sim ARMA(1,1)$) luottamusväli asettuu nopeasti vakiouralle. Mallit on sovitettu 70-vuotiaiden miesten log-kuolevuuden ja log-kuolevuuden muutosten aikasarjoihin vuosiväliltä 1955—2022. ARIMA-mallin ennusteen pohjalta log-kuolevuuden muutosten luottamusväli lasketaan kaavalla $\Delta \log(\hat{m}_{x,t}^q) = \log(\hat{m}_{x,t}^q) - \log(\hat{m}_{x,t-1}^q)$, jossa $\hat{m}_{x,t}^q$ vastaa luottamusvälin mukaista kuolevuuden ennusteen kvantiilia, ts. $q = \alpha/2, q = 1 - \alpha/2$ ($\alpha = 20\%$ tai 5%). Vastaavasti ARMA-mallin ennusteen pohjalta log-kuolevuuden luottamusväli saadaan kaavalla $\log(\hat{m}_{x,t}^q) = \log(\hat{m}_{x,t-1}^q) + \Delta \log(\hat{m}_{x,t}^q)$.

C Lee-Carter ja Lee-Miller -mallit

Lee-Carter -mallissa (1992) [8] kuolevuutta mallinnetaan ikä- ja aikakohtaisilla faktoreilla yhtälöllä

$$\log(m_{x,t}) = a_x + b_x k_t + \epsilon_{x,t}, \quad (27)$$

jossa a_x on kuolevuuden ikäkohtainen keskimääräinen taso, b_x on kuolevuuden muutoksen ikäkohtainen intensiteetti (ikäfaktori), k_t on kuolevuuden suhteellinen taso vuonna t (aikafaktori) ja $\epsilon_{x,t} \sim nid(0, \sigma_{\epsilon;x}^2)$. Mallissa faktorit b_x ja k_t estimoidaan niin sanotulla singulaariarvohajotelmalla (engl. *Singular Value Decomposition*). Lee-Carter -mallin yli-identifioituvuuden takia parametreja rajoitetaan ehdoilla $\sum_t k_t = 0$ ja $\sum_x b_x = 1$. Lisäksi alkuperäisessä Lee-Carter -mallissa k_t asetetaan siten, että mallilla saadut kuolleiden lukumäärät vastaavat havaintoaineiston mukaisia kuolleiden lukumääriä.²⁴

Lee-Carter -mallilla ennustaminen tapahtuu faktorin k_t avulla siten, että faktorin oletetaan noudattavan tässäkin työssä esiteltävää satunnaiskulkuprosessia (ts. ARIMA(0,1,0)), jolloin

$$k_t = d + k_{t-1} + e_t, \quad (28)$$

jossa $e_t \sim nid(0, \sigma_e^2)$. Termin d estimaatti saadaan yksinkertaisesti yhtälöstä $d = (k_T - k_1)/(T - 1)$ ja ennusteen $\hat{k}_{T+h|T}$ arvo yhtälöstä $\hat{k}_{T+h|T} = k_T + dh$. Tällöin Lee-Carter -mallin kuolevuusennuste saadaan yhtälöstä

$$\hat{m}_{x,T+h|T} = \exp[a_x + b_x(k_T + dh)] = \exp[a_x + b_x k_T] \exp[b_x dh] = \hat{m}_{x,T} \exp[b_x dh]. \quad (29)$$

Toisin kuin Lee-Carter -mallissa, Lee-Miller -mallin (2001) [9] ennusteessa sovitettujen kuolevuuden $\hat{m}_{x,T}$ tilalle vaihdetaan periodilla T havaittu todellinen kuolevuus eli $m_{x,T}$.²⁵

Edeltävän ennusteyhtälön perusteella log-kuolevuuden muutoksen ennusteeksi saadaan

$$\log(\hat{m}_{x,T+h|T}) - \log(\hat{m}_{x,T}) = b_x dh = b_x \frac{h(k_T - k_1)}{T - 1}. \quad (30)$$

Ennusteen liikkua yhden periodin verran eteenpäin, log-kuolevuus muuttuu aina vakiokertoimen $b_x d$ verran. Kuolevuuden ennusteen kehitys perustuu täten kuolevuuden suhteellisen tason k_t sovituspäivästä $(1, \dots, T)$ havaitun keskimuutoksen d ja ikäkohtaisen intensiteetin b_x väliin tulon. Kuvasta 8 nähdään, että Lee-Miller -mallin ennustama log-kuolevuuden muutos asettuu heti ennustehorisontin alussa vakiouralle.

Lee ja Carter (1992, 699–670) [8] esittivät, että Lee-Carter -mallin log-kuolevuuden ikäkohtaisen ennusteen epävarmuus eli ennustevirheen varianssi $\sigma_{E;x,T+h}^2$ saadaan yhtälöstä

$$\sigma_{E;x,T+h}^2 = \sigma_{\epsilon;x,T+h}^2 + \sigma_{a;x}^2 + b_x^2 \sigma_{k;T+h}^2 + \sigma_{b;x}^2 (k_{T+h}^2 + \sigma_{k;T+h}^2), \quad (31)$$

jossa varianssit $\sigma_{a;x}^2$ ja $\sigma_{b;x}^2$ viittaavat parametrin a_x ja ikäfaktorin b_x estimointiin liittyvään epävarmuuteen. Varianssi $\sigma_{k;T+h}^2$ viittaa puolestaan aikafaktorin k_t ennusteen epävarmuuteen,

²⁴Tässä tutkimuksessa tätä ehtoa ei kuitenkaan sovelleta, jolloin Lee-Carter -malli toteutetaan kuten henkivakuutuksen K2004-referenssikuoletutkimuksessa (ks. Sirén (2013) [11], luku 4.2).

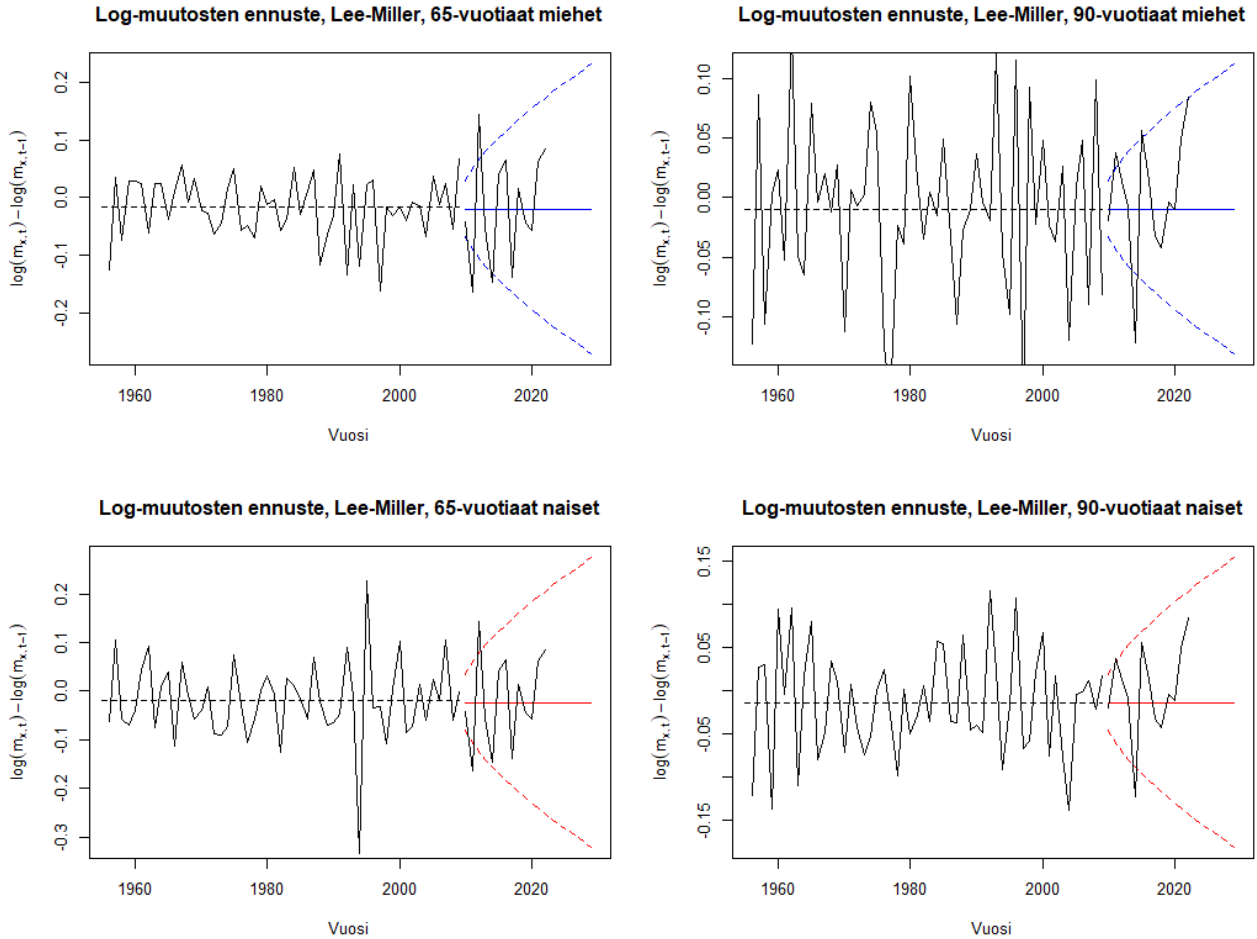
²⁵On hyvä huomata, ettei ennusteyhtälöön tehtävä korjaus ole ainoa ero Lee-Carter ja Lee-Miller -mallien välillä, vaan eroa syntyy myös parametrin k_t sovittamisessa. Käytännössä Lee-Miller -mallissa parametrit k_t sovitetaan siten, että mallin mukaiset periodikohtaiset vastasyntyneiden elinajanodotteet vastaavat lähtöaineistosta havaittuja vastasyntyneiden elinajanodotteita.

joka itsessään on summa satunnaiskulkuprosessin virhetermin e_t varianssista sekä trenditermin d estimoimiseen liittyvästä neliöidystä virheestä, eli $\sigma_{k;T+h}^2 = h\sigma_e^2 + h^2\sigma_d^2$.

Lee ja Carter (1992, 699–670) [8] arvioivat, että pidemmissä eli yli kahdenkymmenen vuoden ennusteissa aikafaktorin ennusteen epävarmuus tuottaa noin yhdeksänkymmentä prosenttia tai enemmän koko ennusteen epävarmuudesta. Sen sijaan yhden vuoden ennusteissa aikafaktoriin k_t liittyvän epävarmuuden arvioitiin kattavan runsaan kolmanneksen koko ennusteen epävarmuudesta. Koska Lee ja Carter tarkastelivat tutkimuksessaan pääsääntöisesti pitkiä kuolevuuden ennusteita, huomioivat he ennusteissaan vain aikafaktorin ennusteeseen liittyvän epävarmuuden. Tällä oletuksella Lee-Carter -mallin ennusteen luottamusvälit saadaan muotoon

$$\hat{m}_{x,T+h|T} \exp \left[\pm N_{1-\alpha/2} b_x \sqrt{(h\sigma_e^2 + h^2\sigma_d^2)} \right]. \quad (32)$$

Tämän työn yhteydessä Lee-Carter ja Lee-Miller -mallien luottamusvälit tuotetaan vastaavasti käyttäen vain aikafaktorin k_t ennusteeseen liittyvää epävarmuutta. Vertailun vuoksi luottamusvälejä arvioidaan kuitenkin myös siten, että ikäkohtaisen kertoimen b_x tilalla käytetään kerrointa $b_{\max} = \max(b_{20}, b_{21}, \dots, b_{110+})$ eli ikävälillä 20–110+ suurinta kerrointa (ks. Sirén (2013, 31) [11]).



Kuva 8: Log-kuolevuuden muutosten ennusteet ja 95 prosentin luottamusvälit Lee-Miller -mallilla (K2012), 65- ja 90-vuotiaat miehet ja naiset. Kuvista nähdään, että Lee-Miller -mallin ennusteesta laskettu log-kuolevuuden muutoksen ennuste asettuu välittömästi vakiouralle. Ennusteiden vakiourat ovat esimerkkitapauksissa lähes samoja kuin soviteväliltä havaitut keskimääräiset log-kuolevuuden muutokset (kuvissa musta katkoviiva). Log-kuolevuuden muutosten luottamusvälit lasketaan kaavalla $\Delta \log(\hat{m}_{x,t}^q) = \log(\hat{m}_{x,t}^q) - \log(\hat{m}_{x,t-1}^q)$, jossa $\hat{m}_{x,t}^q$ vastaa luottamusvälin mukaista kuolevuuden ennusteen kvantiilia ts. $q = \alpha/2, q = 1 - \alpha/2$ ($\alpha = 5\%$). Kuviin on piirretty myös log-kuolevuuden muutosten toteumat vuosilta 2010–2022.

D Elinajanodotteiden ennusteiden luottamusvälit

Tässä työssä elinajanodotteiden ennusteiden luottamusvälit lasketaan siten, että kaikkien ikäkohtaisten kuolevuuksien oletetaan realisoituvan samanaikaisesti ennustejakaumien kvantiilien q mukaisille tasoille. Toisin sanoen, kaikille mallinnetuille ikäkohtaisille kuolevuuksille pätee

$$\hat{m}_{x,T+h}^q = \hat{m}_{x,T+h} \exp(N_q \sigma_{x,h}), \quad (33)$$

jossa N_q on standardinormaalijakauman q -kvantiili, $\hat{m}_{x,T+h}$ on ikäkohtaisen kuolevuuden ennustettu odotusarvo vuonna $T+h$, $\sigma_{x,h}$ on ikäkohtaisen ennustevirheen keskihajonta vuonna $T+h$. Tällöin elinajanodotteiden ennustejakaumien q -kvantiilit $\hat{e}_{x,T+h}^q$ x -ikäisille ennusteen vuonna $T+h$ saadaan liitteen A mukaisilla kaavoilla

$$\hat{p}_{x,T+h}^q = \max\left(0, \frac{2 - \hat{m}_{x,T+h}^q}{2 + \hat{m}_{x,T+h}^q}\right), \quad (34)$$

$$\hat{e}_{x,T+h}^q = \frac{1}{2} + \sum_{k=1}^{\infty} {}_k p_{x,T+h}^q, \quad (35)$$

ja

$${}_k p_{x,T+h}^q = \prod_{n=0}^{k-1} \hat{p}_{x+n,T+h}^q. \quad (36)$$

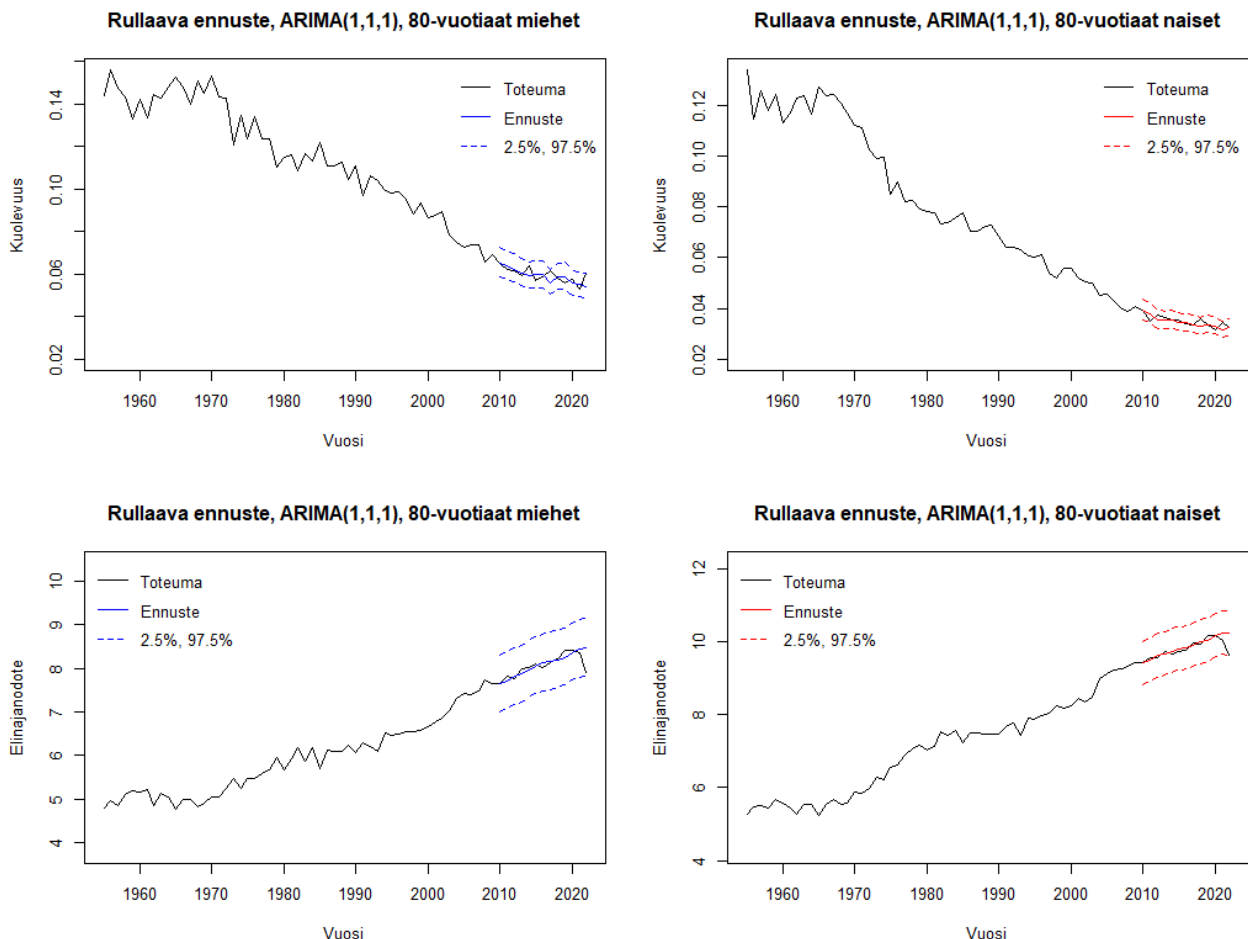
Esitetyn lähestymistavan hyvä puoli on se, että elinajanodotteiden luottamusvälit saadaan muodostettua samoilla yksinkertaisilla ja vertailukelpoisilla kaavoilla kaikille tässä työssä tarkasteltaville malleille. Lisäksi tekniikka vastaa Lee-Carter ja Lee-Miller -mallien oletusta, jonka mukaan ennusteiden epävarmuus syntyy kaikille ikävuosille yhteisen aikafaktorin ennusteen kautta (ks. liite C).²⁶ Toisaalta ARIMA-mallikehikon näkökulmasta esitettyä menettelytapaa voidaan pitää karkeana, sillä tämän työn ikäkohtaisessa ARIMA-mallinnuksessa ei lähtökohtaisesti tehdä oletuksia mallien välisistä riippuvuuksista. Täten ARIMA-mallikehikossa elinajanodotteiden ennustejakaumia voitaisiin todennäköisesti parantaa, ja sovellettavasta riippuvuusra-kenteesta riippuen todennäköisesti myös kaventaa, mikäli ikäkohtaisten ARIMA-mallien välille lisättäisiin tarkempaa riippuvuusmallinnusta.

²⁶Elinajanodotteiden ennusteiden luottamusvälejä perustellaan tarkemmin Leen ja Carterin (1992) [8] tutkimuksen liitteissä A ja B.

E Rullaava ARIMA-ennuste

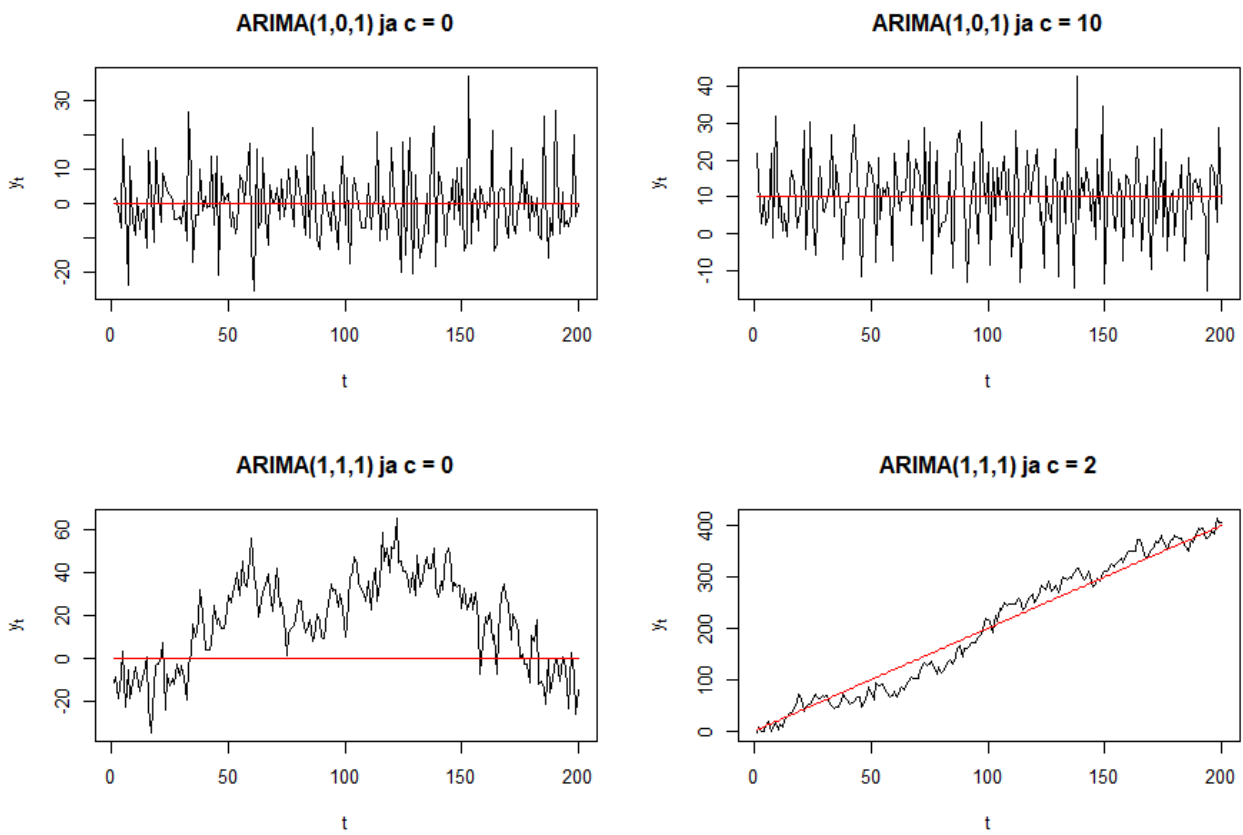
Luvussa 3 esitetyt kuolevuuden ja elinajanodotteiden ennusteet on toteutettu usean vuoden ennustehorisontille ($h \geq 2$). Sen sijaan tässä liitteessä esitetään kuolevuuden ja elinajanodotteen ennusteet toteutettuna siten, että yhden vuoden ennuste tehdään vuosittain uudelleen lähtien viimeisestä havaitusta pisteestä ($h = 1$). Käytännössä näin toteutettuja rullaavia ennusteita voitaisiin käyttää esimerkiksi vastuuvelan kuolevuus- ja pitkäikäisyysriskiä arvioitaessa (ts. yhden vuoden *value-at-risk*). Tällöin ennustejakauman volatiliteetin mallintamisessa voitaisiin puolestaan hyödyntää esimerkiksi GARCH-malleja.

Kuvassa 9 on esitetty 80-vuotiaiden miesten ja naisten kuolevuuden ja elinajanodotteiden rullaavat ennusteet ja ennusteiden 95 prosentin mukaiset luottamusvälit ARIMA(1,1,1)-mallilla. Mallit on sovitettu vuosivälille 1955—2009 ja ennusteita verrataan vuosien 2010—2022 toteutuneita havaintoja vastaan. Kuvaajista nähdään, että rullaavasta luonteestaan huolimatta ennusteet ovat suhteellisen vakaita. ARIMA-mallien rullaavat ennusteet eivät ole pystyneet ennakoimaan korona-aikana havaittua merkittävää kuolevuuden kasvua ja elinajanodotteiden laskua, joskin havaitut arvot ovat pysyneet kutakuinkin ennusteen luottamusvälien sisällä.

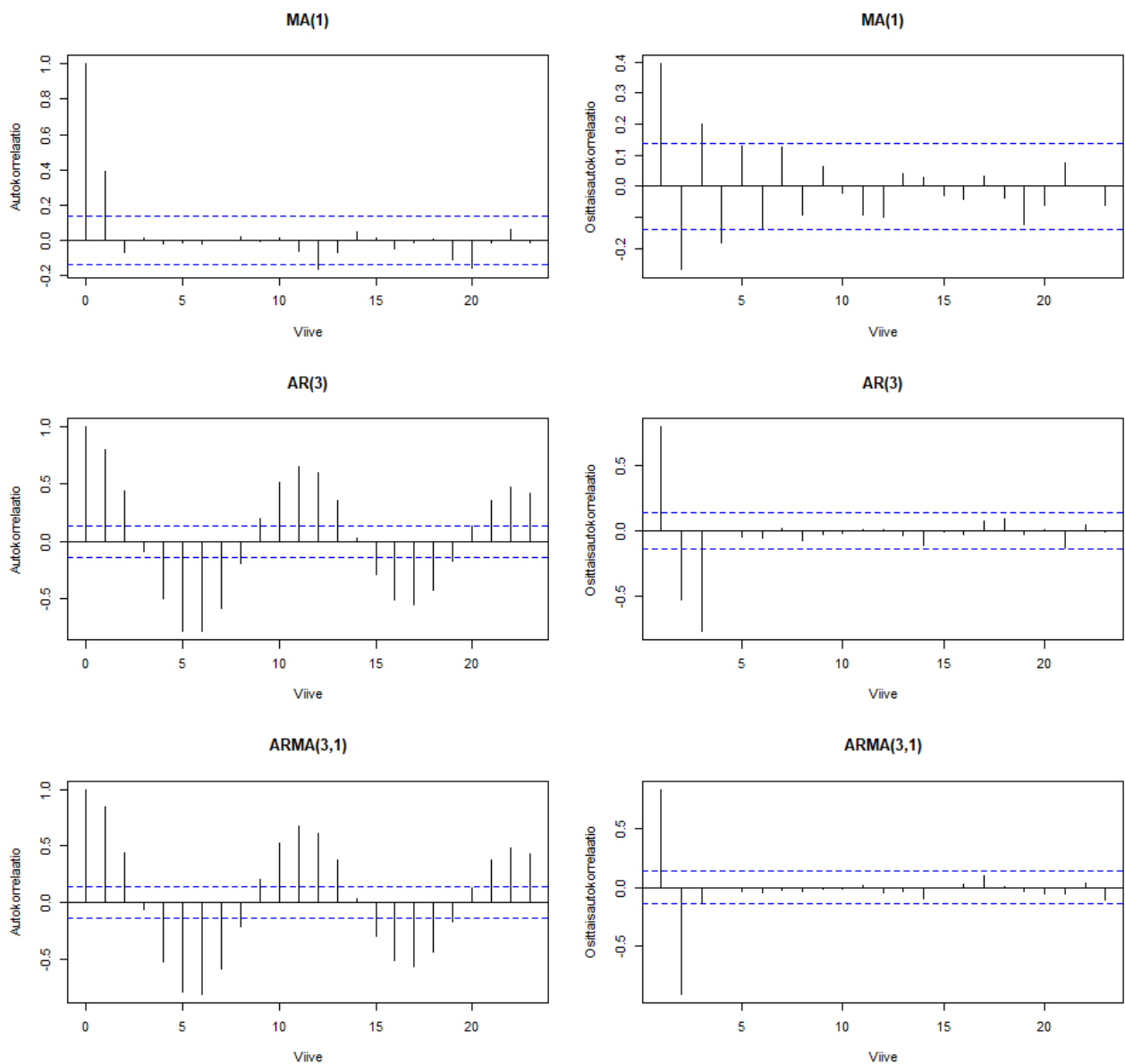


Kuva 9: Rullaava ARIMA(1,1,1)-ennuste 80-vuotiaiden miesten ja naisten kuolevuudelle ja elinajanodottele.

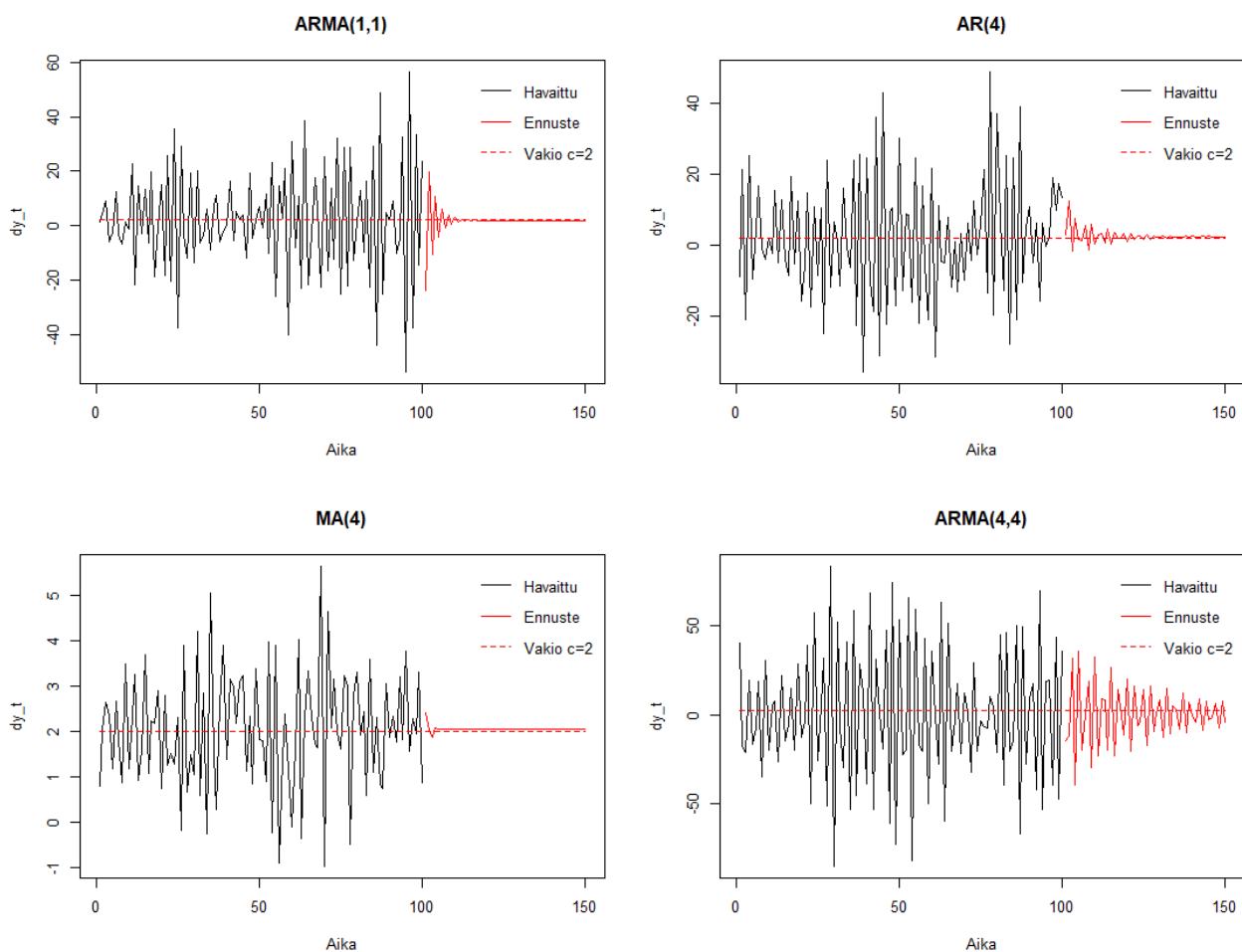
F Kuvakokoelma



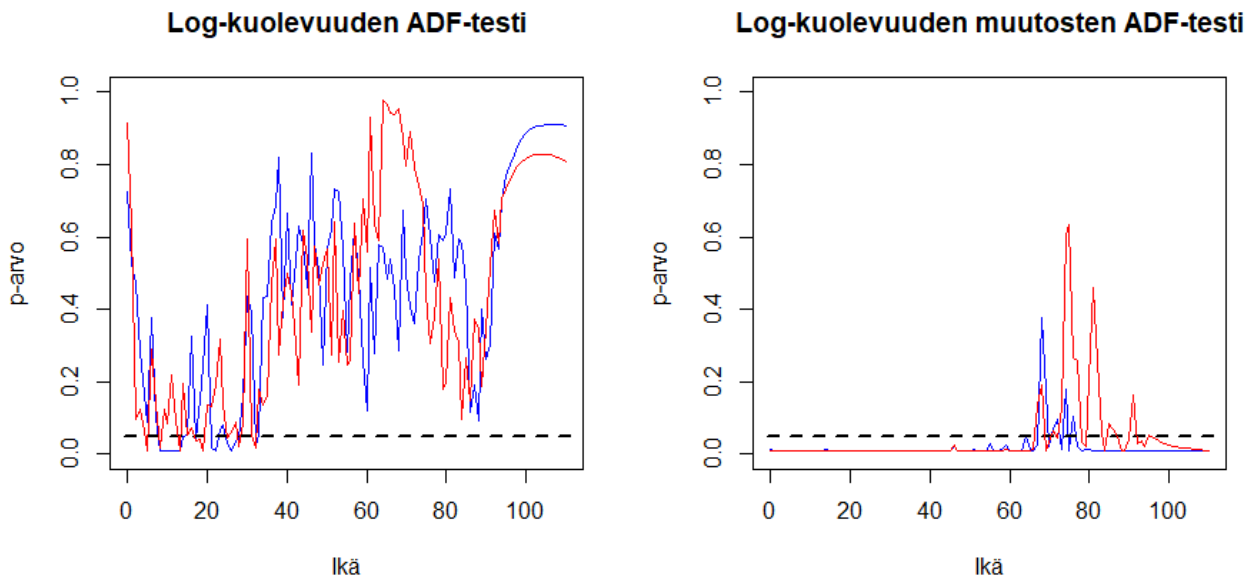
Kuva 10: Simuloituja esimerkkejä ARIMA(1,d,1)-prosesseista.



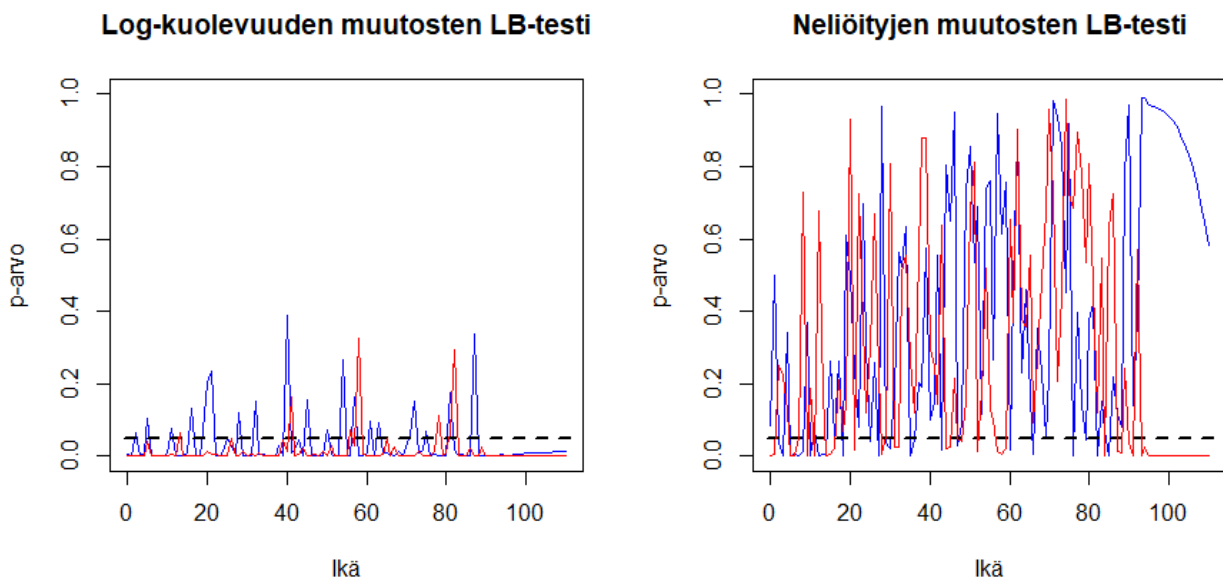
Kuva 11: Esimerkkejä simuloitujen MA(1)-, AR(3)- ja ARMA(3,1)-prosessien estimoiduista otosautokorrelaatio- ja osittaisautokorrelaatiofunktioista luottamusväleineen. MA(1)-prosessin otosautokorrelaatiofunktiossa on selvä katkos viipeen 1 kohdalla ja osittaisautokorrelaatiofunktio on siniaaltoinen. Sen sijaan AR(3)-prosessin otosautokorrelaatiofunktio on siniaaltoinen ja osittaisautokorrelaatiofunktiossa on selvä katkos viipeen 3 kohdalla. ARMA(3,1)-prosessin tapauksessa autokorrelaatiofunktioista ei voida selvästi päätellä sopivaa $p:n$ ja $q:n$ arvoa (jos jotain, niin funktiot jopa viittaisivat AR(2)-prosessiin tässä tapauksessa).



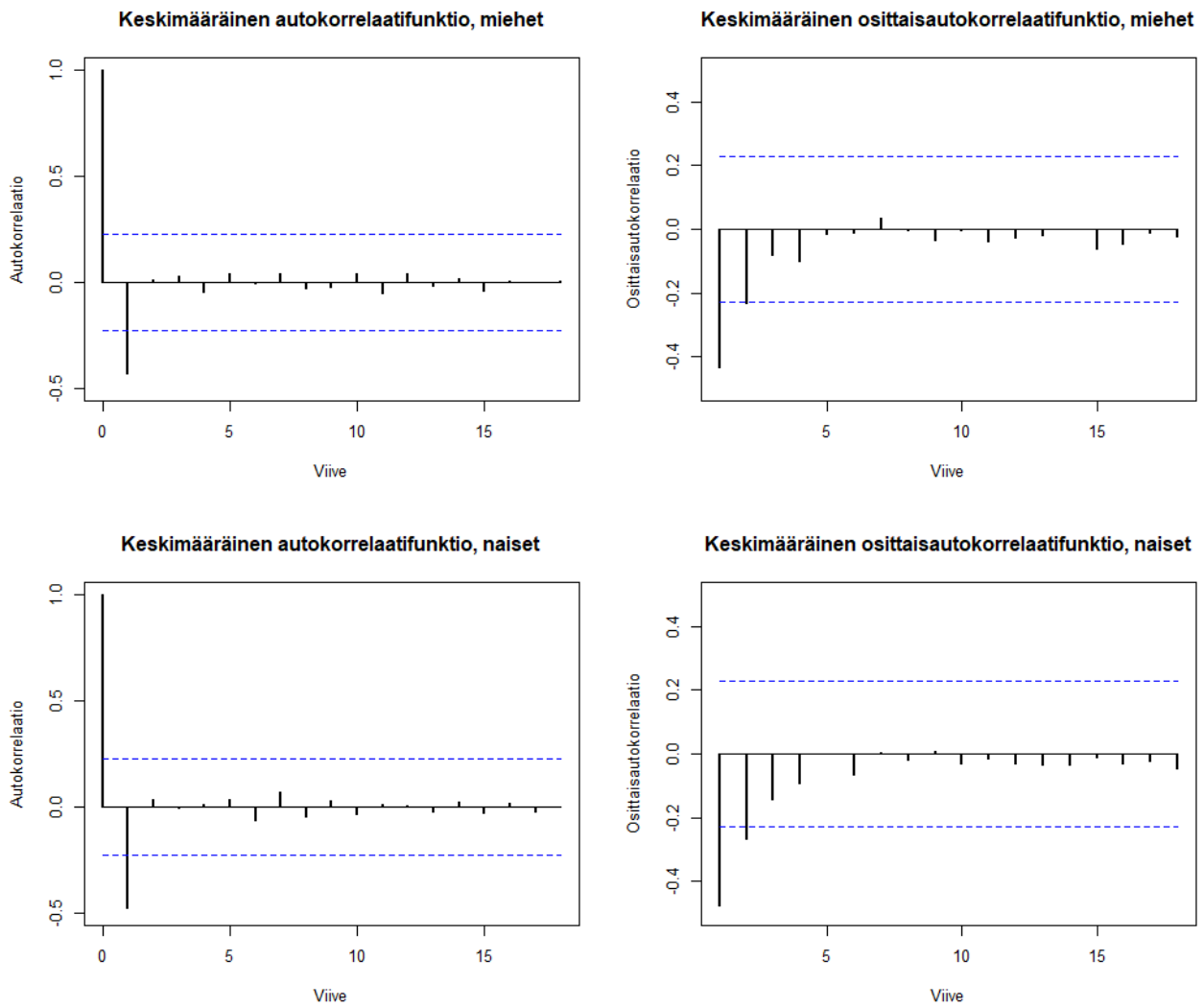
Kuva 12: ARMA-mallin spesifikaation vaikutus ennusteeseen. Kuvaajista nähdään, että stationaarisen prosessin tapauksessa ennuste (punainen) palautuu sitä nopeammin vakiouralle c (punainen katkoviiva), mitä vähemmän mallissa on parametreja eli mitä pienempiä q :n ja erityisesti p :n arvot ovat. Sekä MA- että AR-prosessien parametrit on asetettu arvoiksi -0.7 , 0.4 , 0.5 ja -0.1 (tässä järjestyksessä). Teoreettisen vakion c arvo on puolestaan 2 .



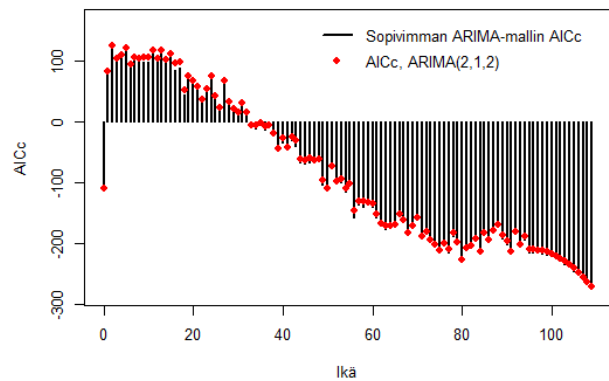
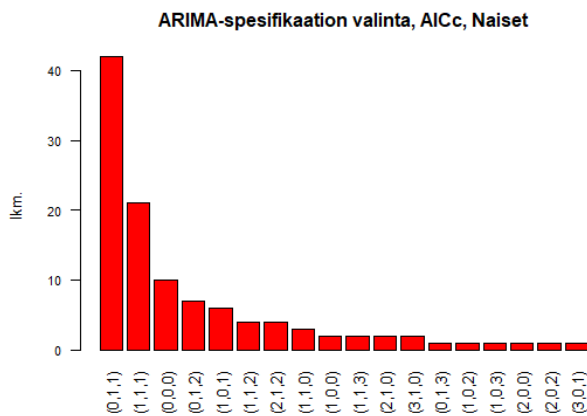
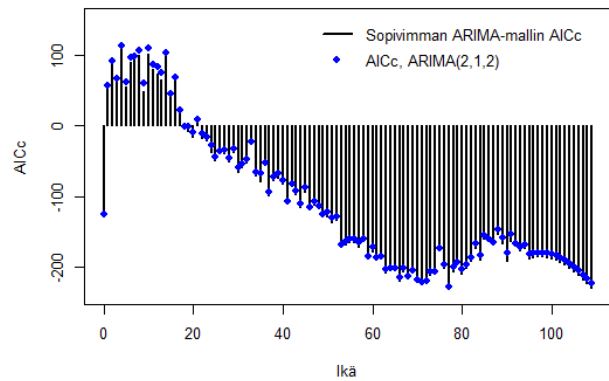
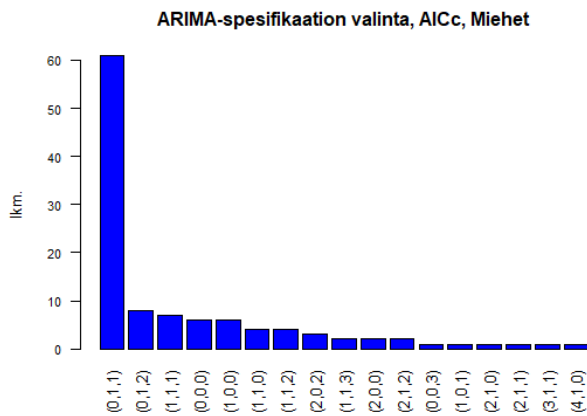
Kuva 13: Log-kuolevuuden sekä log-kuolevuuden muutosten yksikköjuuren ADF-testien ikäkohtaiset p-arvot miehille (sininen viiva) ja naisille (punainen viiva). Musta katkoviiva kuvaa viiden prosentin luottamustasoa.



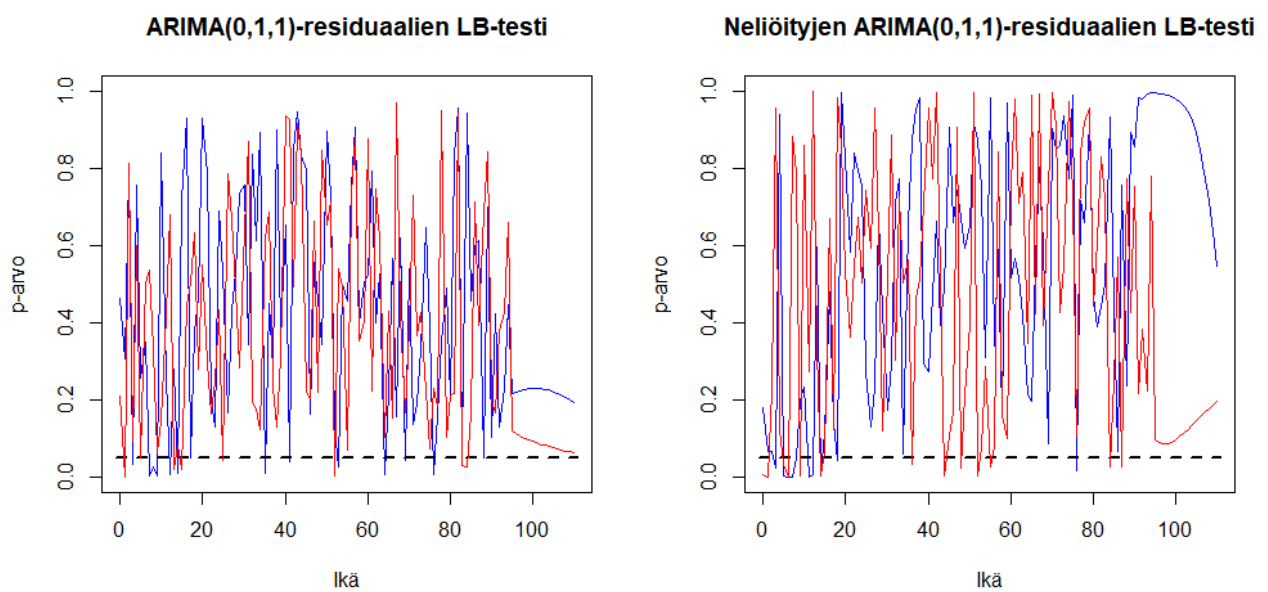
Kuva 14: Log-kuolevuuden muutosten lineaarisen ja epälineaarisen autokorreloituneisuuden LB-testien ikäkohtaiset p-arvot miehille (sininen viiva) ja naisille (punainen viiva). Musta katkoviiva kuvaa viiden prosentin luottamustasoa.



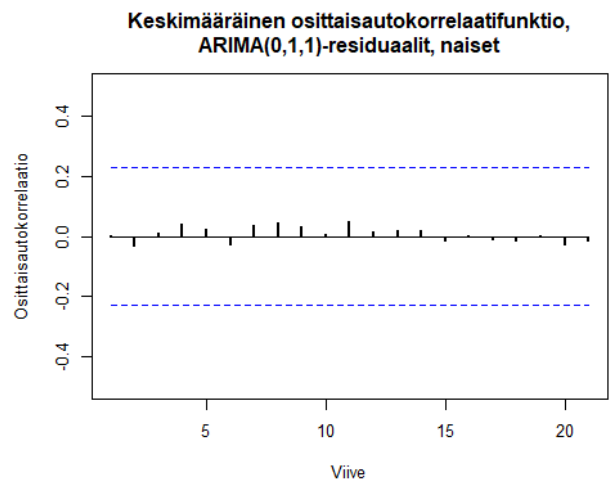
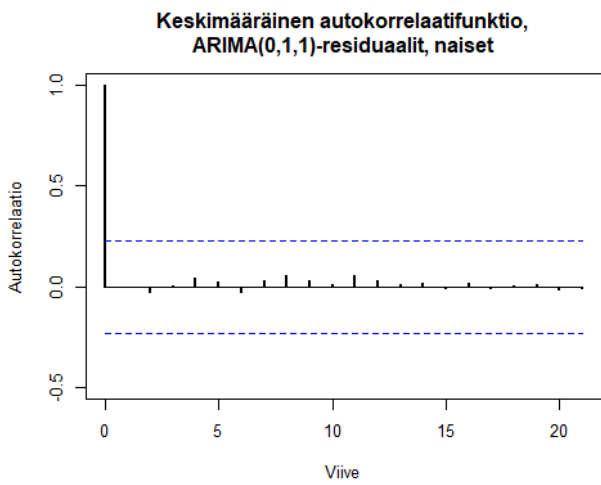
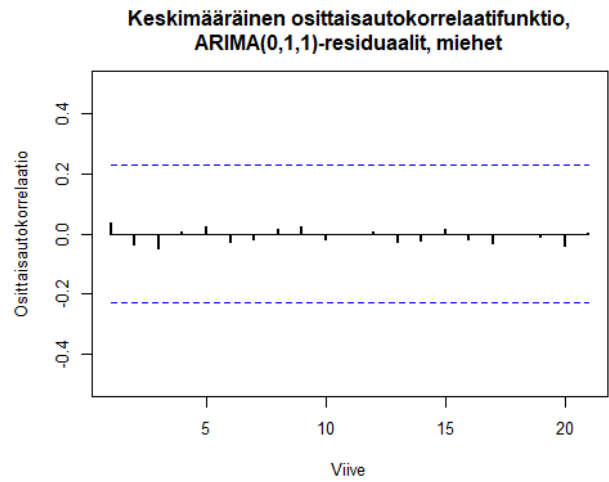
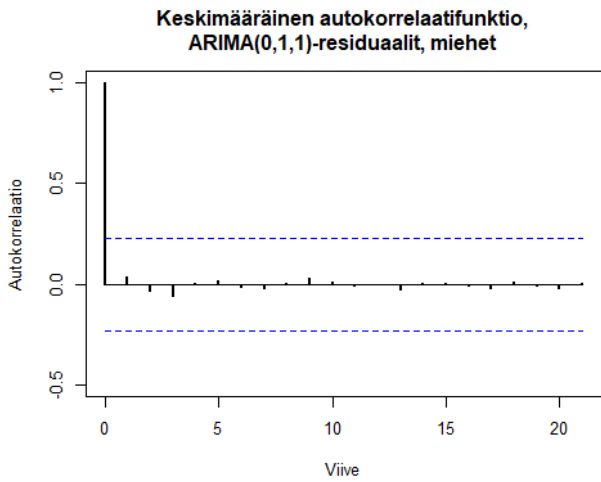
Kuva 15: Log-kuolevuuden muutosten keskimääräiset otosautokorrelaatio- ja osittaisautokorrelaatiofunktio. Keskimääräiset autokorrelaatiofunktio on laskettu sukupuolittain yli kaikkien ikäkohtaisten aikasarjojen autokorrelaatiofunktioiden. Siniset katkoviivat kuvaavat yksittäisten autokorrelaatioiden 95 prosentin luottamusvälejä.



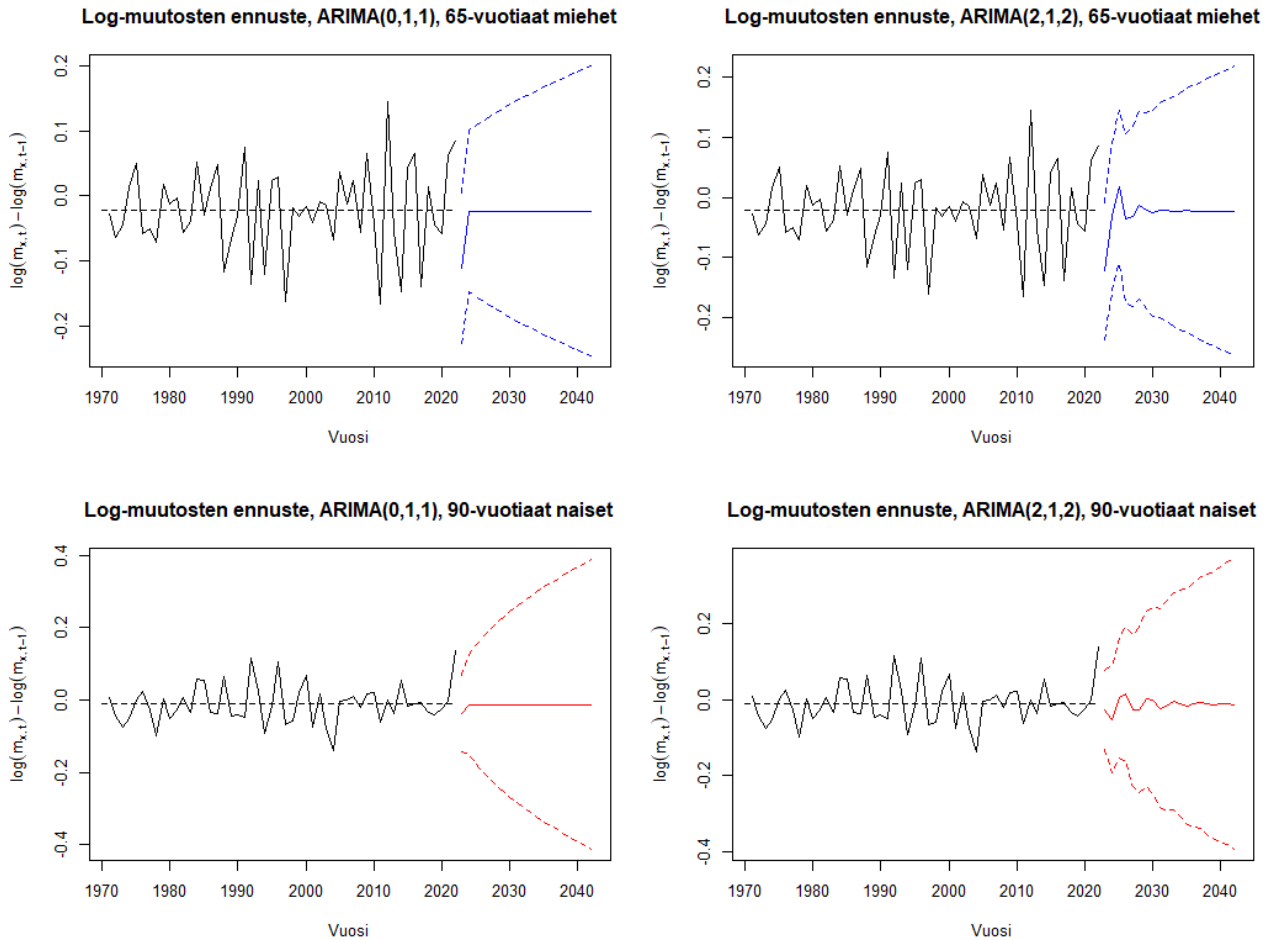
Kuva 16: Sopivimmat ARIMA-mallit auto.arima -funktion ja AICc-informaatiokriteerin perusteella. Vasemmanpuoleisissa kuvissa näytetään, kuinka usealle kuolevuusajaksarjalle kyseinen spesifikaatio oli sopivin. Oikeanpuoleisissa kuvissa mustat viivat kuvaavat sopivimman ARIMA-mallin AICc-arvoa ja värilliset pallot valitun vertailumallin eli ARIMA(2,1,2)-spesifikaation AICc-arvoa. Ylhäällä sinisellä kuvataan tulokset miehille ja alhaalla punaisella tulokset naisille.



Kuva 17: ARIMA(0,1,1)-mallin residuaalien lineaarisen ja epälineaarisen autokorreloituneisuuden LB-testien ikäkohtaiset p-arvot miehille (sininen viiva) ja naisille (punainen viiva). Musta katkoviiva kuvaa viiden prosentin luottamustasoa.

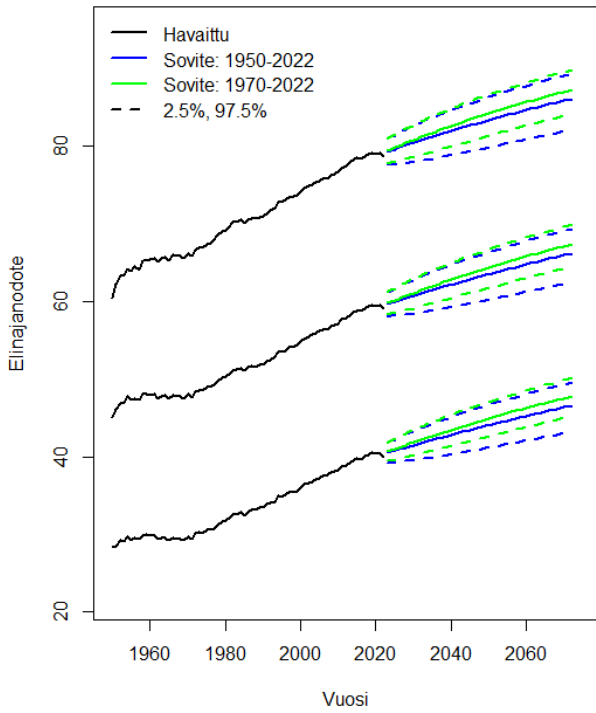


Kuva 18: ARIMA(0,1,1)-mallin residuaalien keskimääräiset otosautokorrelaatio- ja osittaisautokorrelaatiofunktio. Keskimääräiset autokorrelaatiofunktio on laskettu sukupuolittain yli kaikkien ikäkohtaisten residuaaliaikasarjojen autokorrelaatiofunktioiden. Siniset katkoviivat kuvaavat yksittäisten autokorrelaatioiden 95 prosenttin luottamusvälejä.

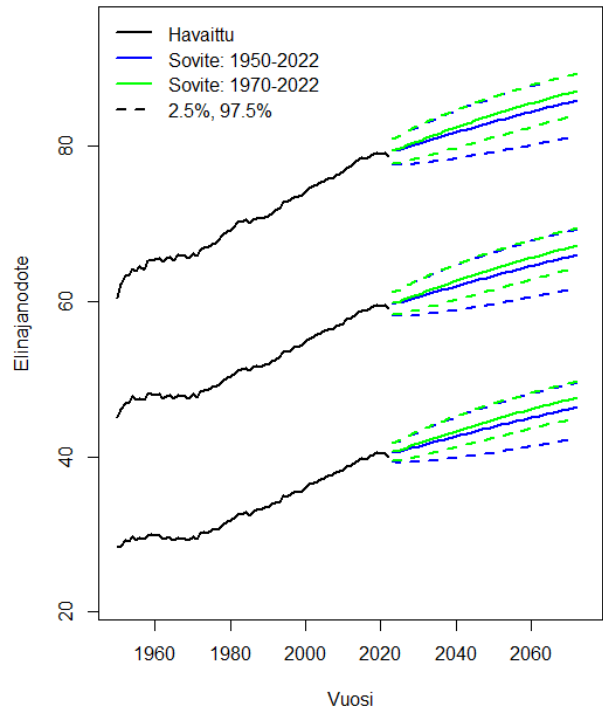


Kuva 19: Log-kuolevuuden muutosten ennusteet ja 95 prosentin luottamusvälit ARIMA(0,1,1)-mallilla ja ARIMA(2,1,2)-mallilla, 65-vuotiaat miehet ja 90-vuotiaat naiset. Kuvista nähdään, että ARIMA(0,1,1)-ennuste asettuu nopeasti vakiouralle, kun taas ARIMA(2,1,2)-ennusteessa on miesten tapauksessa havaittavissa pientä heiluntaa vielä 2030-luvulla ja naisten tapauksessa jopa 2040-luvulla. Ennusteiden vakiourat vastaavat soviteväleiltä havaittuja keskimääräisiä log-kuolevuuden muutoksia (kuivissa musta katkoviiva). Log-kuolevuuden muutosten luottamusvälit lasketaan kaavalla $\Delta \log(\hat{m}_{x,t}^q) = \log(\hat{m}_{x,t}^q) - \log(\hat{m}_{x,t-1}^q)$, jossa $\hat{m}_{x,t}^q$ vastaa luottamusvälin mukaista kuolevuuden ennusteen kvantiilia ts. $q = \alpha/2, q = 1 - \alpha/2$ ($\alpha = 5\%$).

Miesten elinajanodotteet ja ennusteet, ARIMA(0,1,1)
0-, 20-, 40-vuotiaat

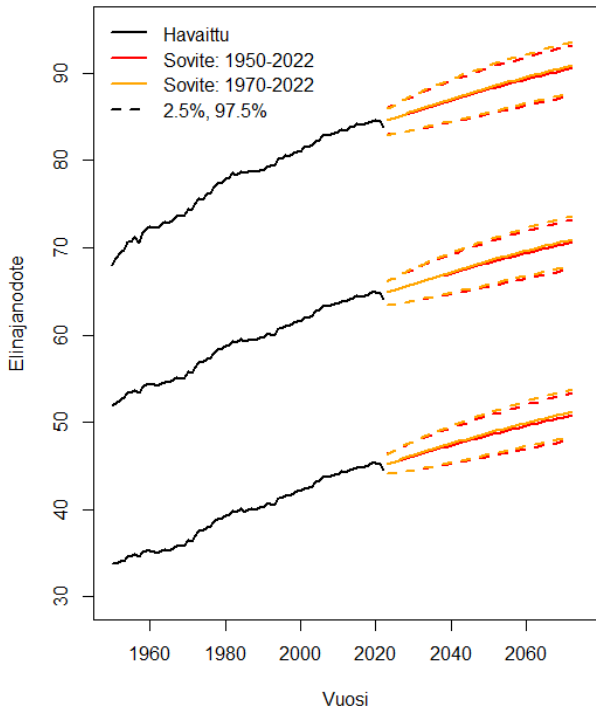


Miesten elinajanodotteet ja ennusteet, ARIMA(2,1,2)
0-, 20-, 40-vuotiaat

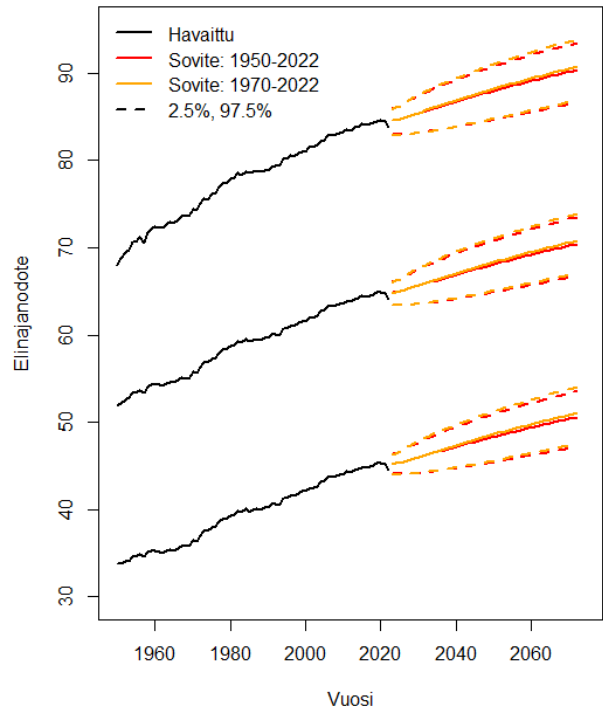


Kuva 20: Elinajanodotteen ennusteet miehille eri ARIMA-spesifikaatiolla.

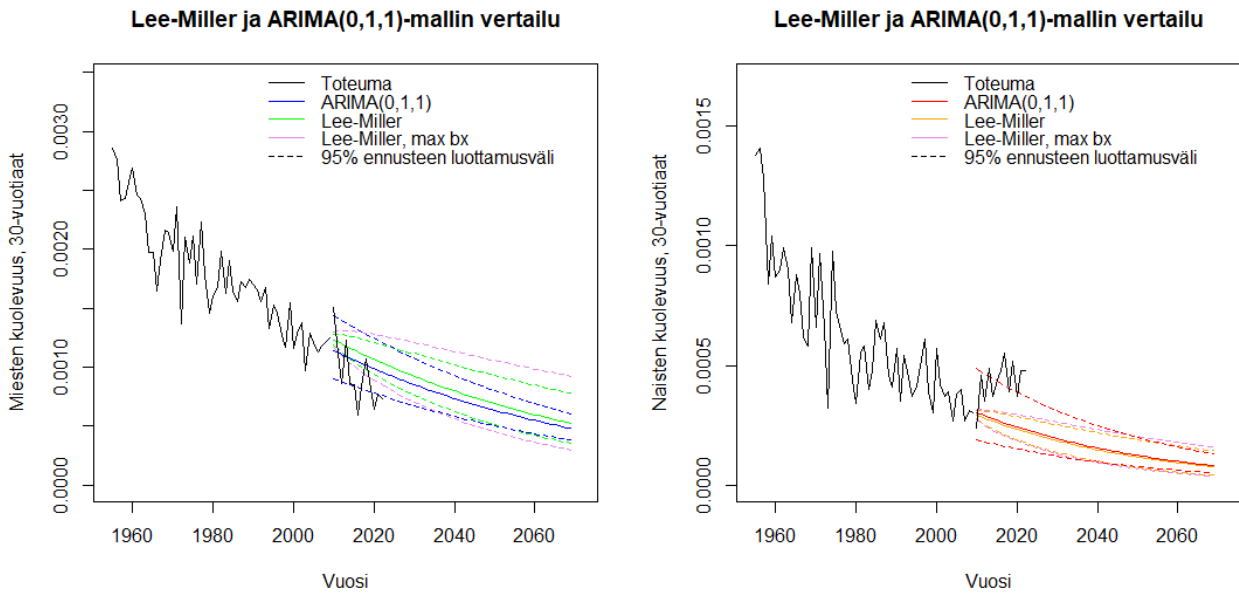
Naisten elinajanodotteet ja ennusteet, ARIMA(0,1,1)
0-, 20-, 40-vuotiaat



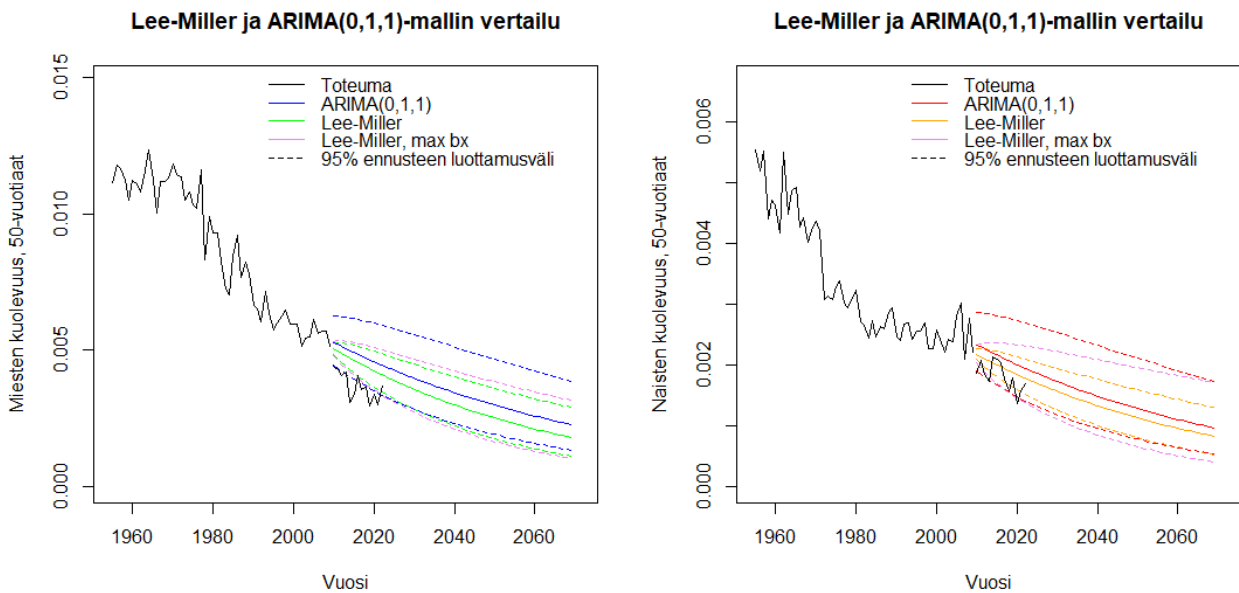
Naisten elinajanodotteet ja ennusteet, ARIMA(2,1,2)
0-, 20-, 40-vuotiaat



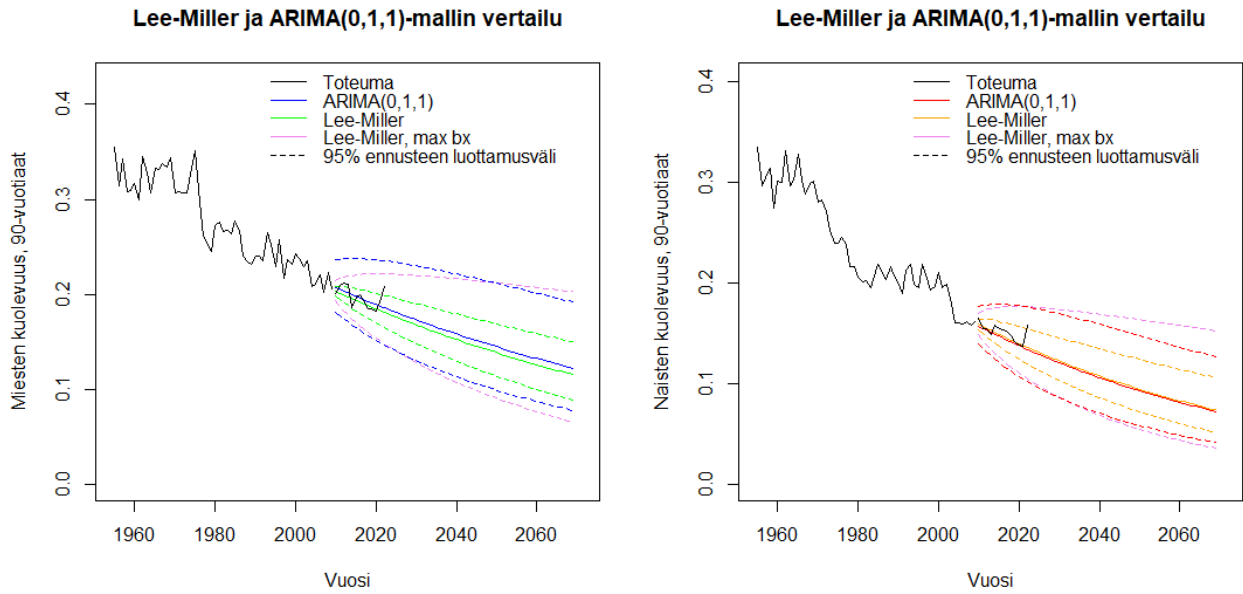
Kuva 21: Elinajanodotteen ennusteet naisille eri ARIMA-spesifikaatiolla.



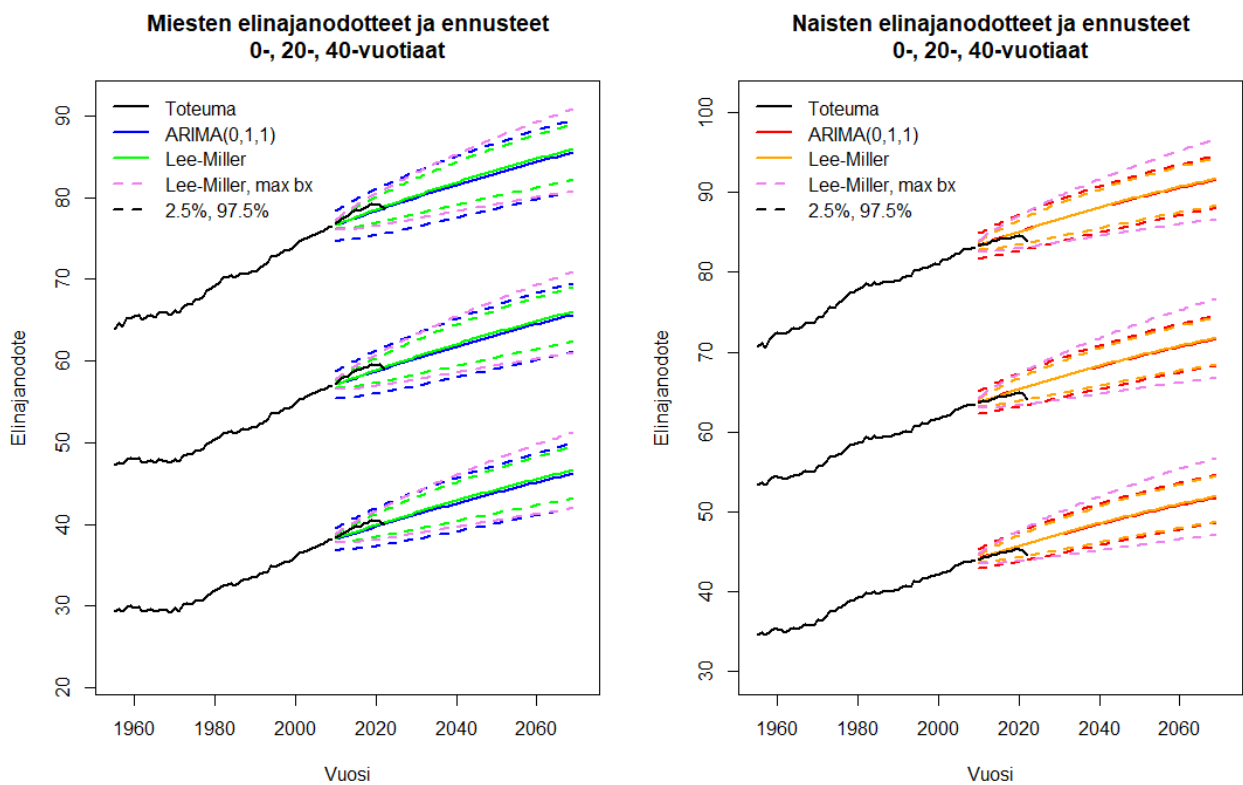
Kuva 22: ARIMA(0,1,1)- ja Lee-Miller -mallien ennustamat kuolevuudet vuosivälin 2010—2022 toteumaan verrattuna, 30-vuotiaat naiset ja miehet. Mallit on sovitettu vuosivälille 1955—2009.



Kuva 23: ARIMA(0,1,1)- ja Lee-Miller -mallien ennustamat kuolevuudet vuosivälin 2010—2022 toteumaan verrattuna, 50-vuotiaat naiset ja miehet. Mallit on sovitettu vuosivälille 1955—2009.

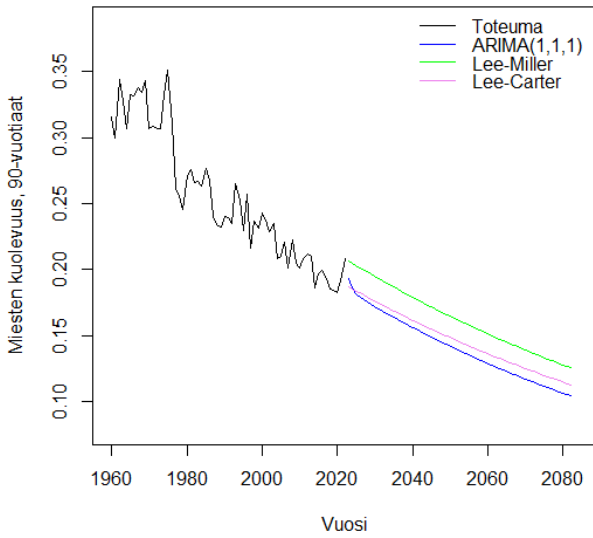


Kuva 24: ARIMA(0,1,1)- ja Lee-Miller -mallien ennustamat kuolevuudet vuosivälin 2010—2022 toteumaan verrattuna, 90-vuotiaat naiset ja miehet. Mallit on sovitettu vuosivälille 1955—2009.

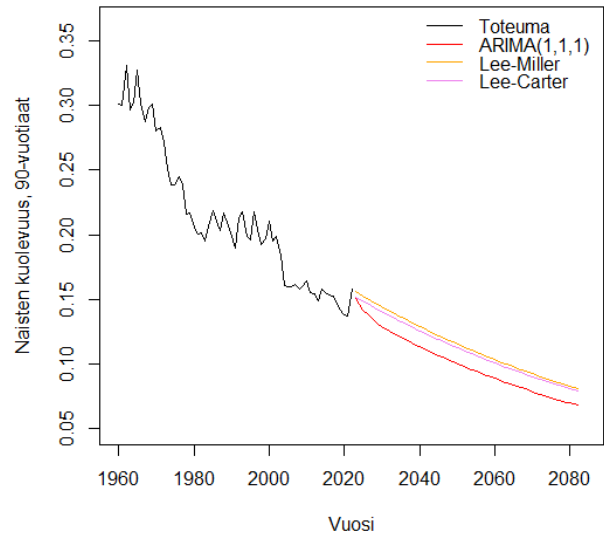


Kuva 25: Elinajanodotteen ennusteiden vertailu, ARIMA(0,1,1) ja Lee-Miller. Mallit on sovitettu vuosivälille 1955—2009.

Lee-Miller, Lee-Carter ja ARIMA(1,1,1)-mallin vertailu

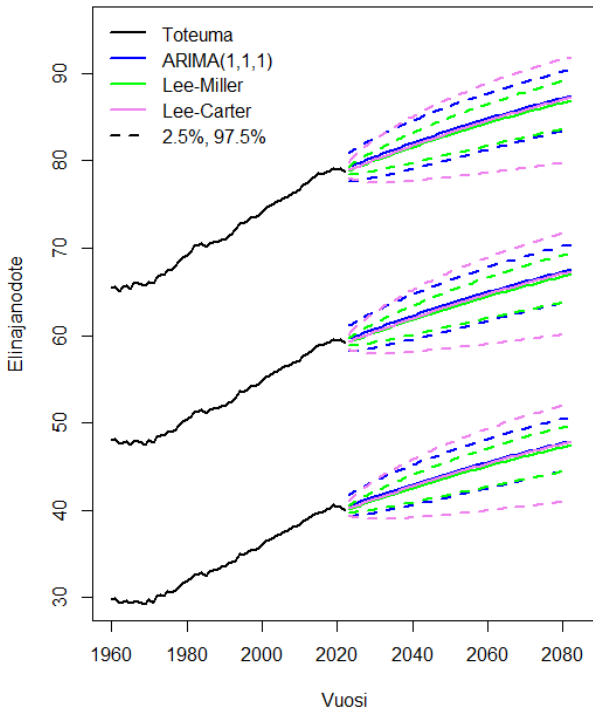


Lee-Miller, Lee-Carter ja ARIMA(1,1,1)-mallin vertailu

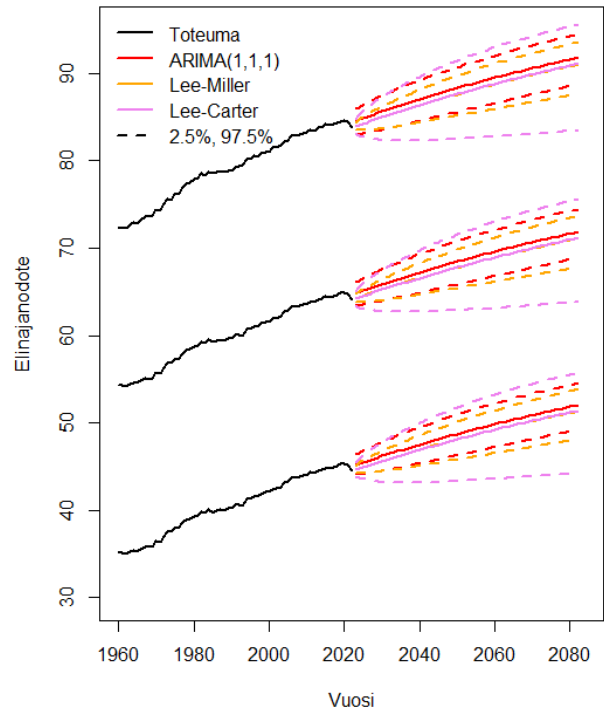


Kuva 26: ARIMA(1,1,1)-, Lee-Carter ja Lee-Miller -mallien kuolevuus, 90-vuotiaat miehet ja naiset. Mallit on sovitettu vuosivälille 1960—2022.

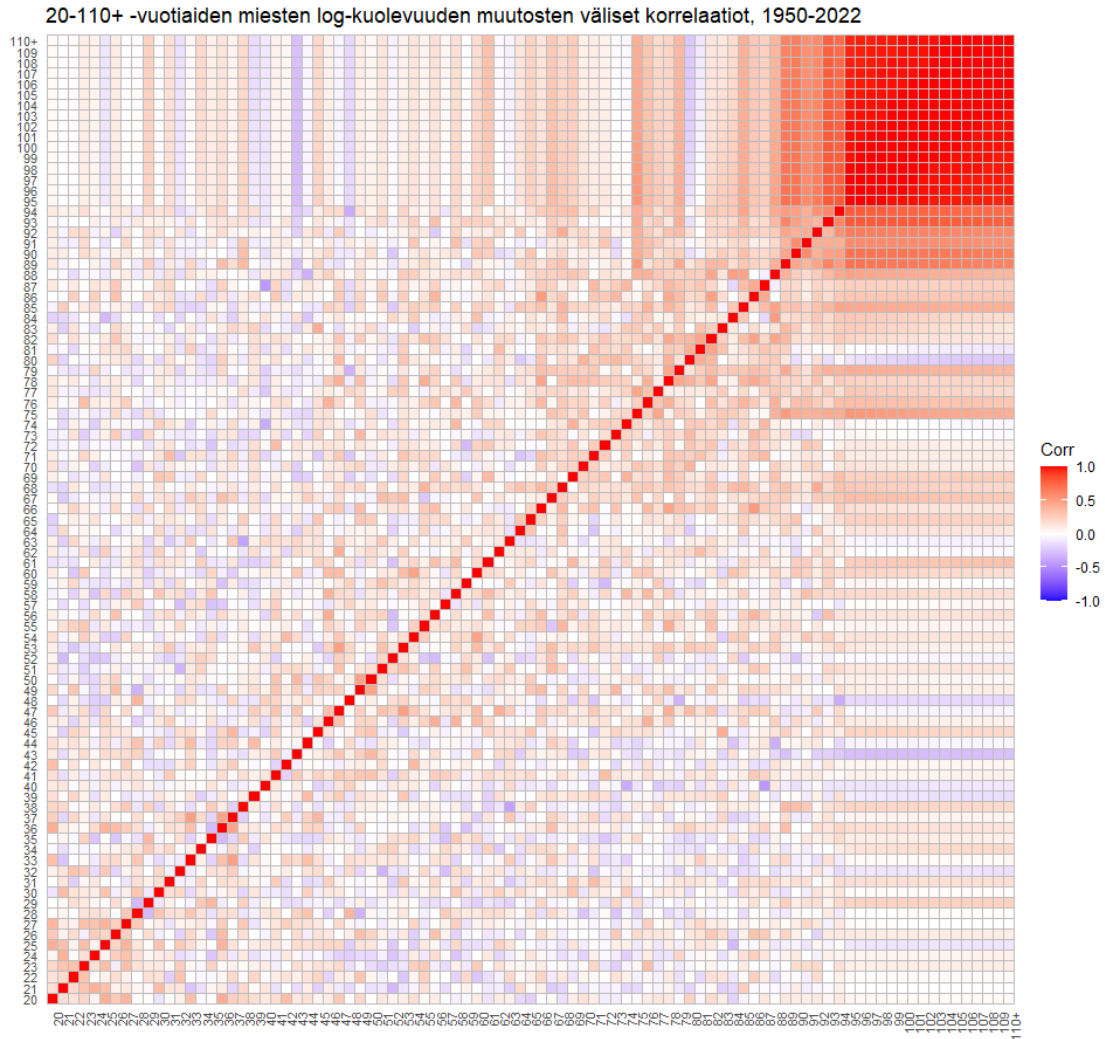
Miesten elinajanodotteet ja ennusteet
0-, 20- ja 40-vuotiaat



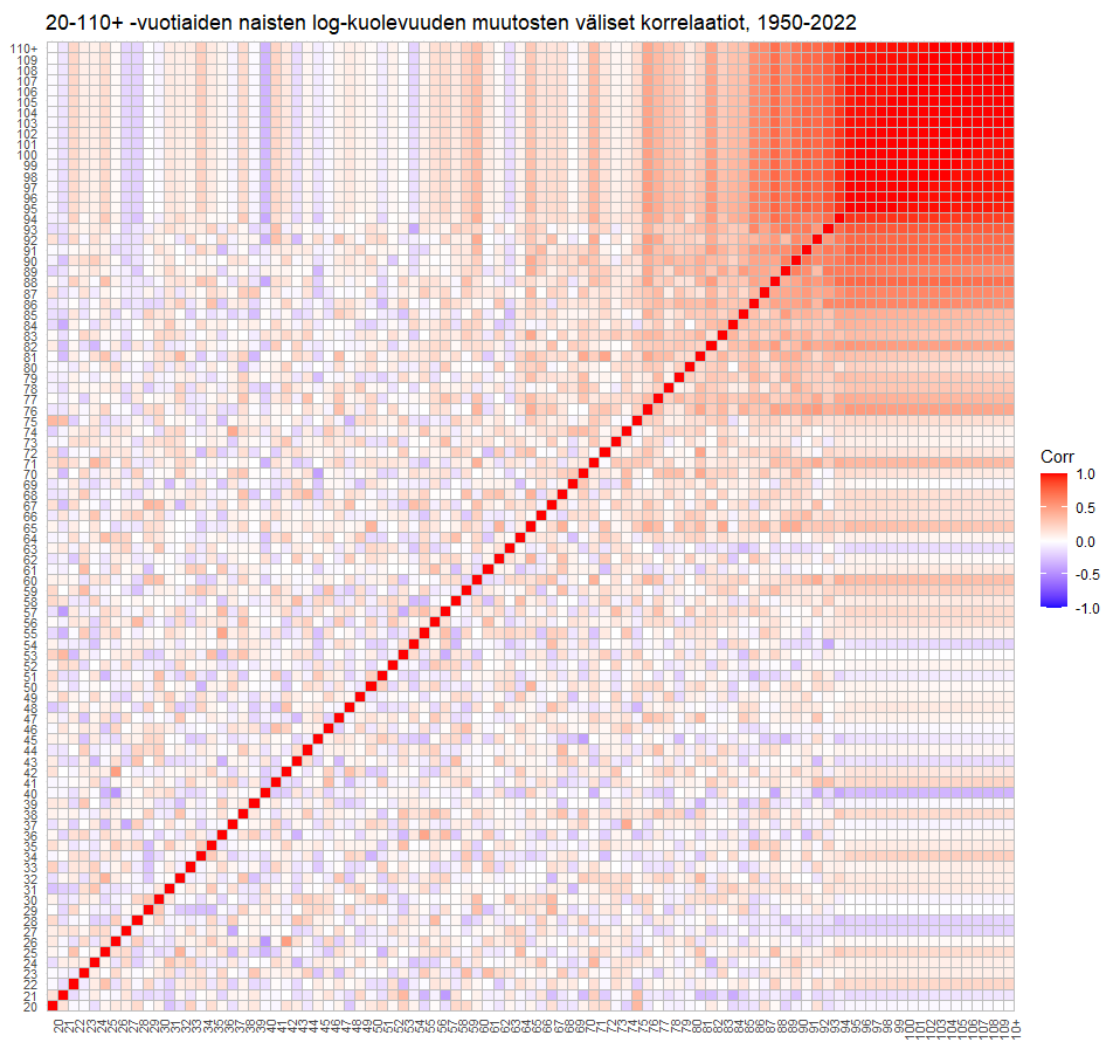
Naisten elinajanodotteet ja ennusteet
0-, 20- ja 40-vuotiaat



Kuva 27: ARIMA(1,1,1)-, Lee-Carter ja Lee-Miller -mallien elinajanodotteet miehille ja naisille. Mallit on sovitettu vuosivälille 1960—2022.



Kuva 28: Log-kuolevuuden muutosten korrelaatiomatriisi 20—110+ -vuotiaille miehille estimoituina vuosiväliltä 1950—2022. Kuvasta nähdään, että korrelaatio on selvästi voimakkainta vanhempien ikäluokkien välillä (erityisesti n. 90—110+ -vuotiaissa), kun taas nuoremmassa ikäluokissa korrelaatio on heikompaa.



Kuva 29: Log-kuolevuuden muutosten korrelaatiomatriisi 20—110+ -vuotiaille naisille estimoitu vuosiväliltä 1950—2022. Kuvasta nähdään, että korrelaatio on selvästi voimakkainta vanhempien ikäluokkien välillä (erityisesti n. 90—110+ -vuotiaissa), kun taas nuoremmissa ikäluokissa korrelaatio on heikompaa.