

Inflaation ennustaminen neuroverkkomallilla
taloudellisen taantuman aikana

Markus Tommi Korhonen

Helsingin yliopisto

Valtiotieteellinen tiedekunta

Taloustiede

Maisterintutkielma

Syyskuu 2019



HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

Tiedekunta/Osasto – Fakultet/Sektion – Faculty Valtiotieteellinen tiedekunta		Laitos – Institution – Department Taloustieteen laitos	
Tekijä □ – Författare – Author Markus Tommi Korhonen			
Työn nimi – Arbetets titel – Title Inflaation ennustaminen neuroverkkomallilla taloudellisen taantuman aikana			
Oppiaine – Läroämne – Subject Taloustiede			
Työn laji – Arbetets art – Level Maisterintutkielma		Aika – Datum – Month and year Syyskuu 2019	Sivumäärä – Sidoantal – Number of pages 64
Tiivistelmä – Referat – Abstract			
<p>Hintavakauden saavuttamisesta on tullut keskuspankkien tärkeimpiä tehtäviä kaikkialla maailmassa, ja useat keskuspankit pyrkivät tiettyyn, hyvin määriteltyyn inflaatiotavoitteeseen. Samoin Euroopan keskuspankki pyrkii rahapolitiikallaan pitämään inflaation kahden prosentin tuntumassa. Inflaatiotavoite kuitenkin vaatii sen, että inflaatiota voidaan ennustaa mahdollisimman tarkasti.</p> <p>Koneoppimismetodeihin kuuluvat neuroverkkomallit ovat osoittautuneet olemaan monilla aloilla hyviä ennustemalleja. Inflaation ennustamisessa neuroverkkomallien tulokset ovat kuitenkin olleet ristiriitaisia. Aiempi tutkimus inflaation ennustamisesta on myös keskittynyt lähinnä Yhdysvaltojen ja muiden yksittäisten maiden inflaatioon. Tutkimusta ei ole myöskään tehty inflaation ennustamisesta eri suhdannetilanteissa neuroverkkomallien avulla. Tässä tutkielmassa tutkittiinkin neuroverkkomallin kykyä ennustaa inflaatiota koko euroalueella vuosien 2008-2009 taantuman aikana.</p> <p>Tutkielman aineistona käytettiin euroalueen harmonisoidusta kuluttajahintaindeksistä muodostettua inflaatioaikasarjaa vuosilta 1997-2010. Tutkielmassa epälineaarinen neuroverkko rakennettiin aiemmasta kirjallisuudesta vakiintuneella metodilla, jossa mallin valinta suoritettiin käyttämällä erillistä aineistoa. Valitulla mallilla simuloitiin aitoa ennustetilannetta käyttämällä euroalueen taantuman aikaista testiaineistoa. Ennusteet tehtiin myös taantuman jälkeiselle noususuhdanteelle, jotta eri suhdannetilanteita voitiin vertailla. Lisäksi samat ennusteet tehtiin ekonometriassa vakiintuneella lineaarisella mallilla, johon neuroverkkomallia verrattiin käyttämällä aiemmasta kirjallisuudesta tuttuja arviointikriteerejä ja tilastollisia testejä.</p> <p>Tutkielmassa selvisi, että neuroverkkomalli tuottaa hyvin tarkkoja ennusteita inflaatiolle kaikilla tutkielmassa käytetyillä ennusteväleillä. Neuroverkkomallin ennusteet ovat myös parempia, jos käytettävä aineisto on kausitasoitettu. Neuroverkkomalli tekee pienempiä ennustevirheitä noususuhdanteen aikana kuin taantumassa, mutta erot eri suhdannetilanteissa eivät ole kovin suuria. Neuroverkkomallin ennusteet eivät kuitenkaan poikkea yksinkertaisen lineaarisen mallin tekemistä ennusteista tilastollisesti merkitsevästi kummassakaan suhdannetilanteessa. Näin ollen neuroverkkomallin ei voida päätellä toimivan eri tavalla taloudellisessa taantumassa kuin muissa suhdannetilanteissa.</p> <p>Tutkielman tulosten perusteella neuroverkkomallia ei voida suositella keskuspankkien inflaatioennustemalliksi, koska mallin valinta ja testaaminen vievät yksinkertaista lineaarista mallia enemmän aikaa, mutta ennustetulokset eivät ole lineaarista mallia parempia. Tulokset antavatkin todisteita siitä, että inflaatio on euroalueella lineaarinen prosessi, jolloin epälineaariset mallit eivät tuota ennusteisiin lisähyötyä. Neuroverkkomallit voivat kuitenkin antaa hyvän työkalun keskuspankkien toiminnan arvioimiseen, koska niiden tuottamat ennusteet ovat tarkkoja pitemmillekin aikaväleille.</p>			
Avainsanat – Nyckelord – Keywords Inflaatio, aikasarja, ennustaminen, koneoppiminen, taantuma, neuroverkot			

Sisältö

1 Johdanto	3
2 Teoria	6
2.1 Neuroverkkomallit	6
2.2 Neuroverkon estimointi	8
2.3 Inflaatio ja rahapolitiikka	10
2.4 Inflaation ennustaminen	13
2.5 Inflaation ennustaminen neuroverkkomallien avulla	15
3 Metodit ja tutkimusaineisto	21
3.1 Neuroverkkomalli	21
3.2 Vertailumalli	25
3.3 Tutkimusaineisto	27
3.4 Ennustamisstrategia	30
3.5 Arviointikriteerit	32
4 Analyysi	36
4.1 Neuroverkkomalli ja ennusteet	36
4.1.1 Mallin valinta	36
4.1.2 Ennusteet	39
4.2 ARMA-malli ja ennusteet	40
4.2.1 Mallin valinta	40
4.2.2 Ennusteet	45
4.3 Ennusteiden vertailu	46
5 Pohdinta	49
6 Johtopäätökset	53

1 Johdanto

Tässä maisterintutkielmassa tutkitaan inflaation ennustamista koneoppimistekniikoihin kuuluvalla neuroverkkomallilla taloudellisen taantuman aikana. Tutkielmassa rakennetaan eteenpäin syöttävä neuroverkkomalli, jolla pyritään ennustamaan euroalueen inflaatiota ns. suuren taantuman (*The Great Recession*) aikana eli vuosina 2008-2009 euroalueella sekä taantuman jälkeisenä aikana.

Taloustieteessä ja erityisesti ekonometrian tutkimuksessa taloudellisten aikasarjojen analysointi on ollut tärkeässä asemassa. Makrotaloustieteessä aikasarjoja on käytetty sekä ilmiöiden ja muuttujien kuvailuun että eri muuttujien riippuvuuksien tutkimiseen ja ennustamiseen. Tässä tutkielmassa keskitytään jälkimmäiseen eli inflaation ennustamiseen ekonometrian ja makrotaloustieteen kehikossa.

Kokonaisten kansantalouksien aggregaattimuuttujia tutkivassa makrotaloustieteessä inflaatio eli rahan arvon heikkeneminen on ollut yksi suurista mielenkiinnon kohteista. Inflaatio ja odotukset tulevasta inflaatiosta vaikuttavat voimakkaasti ihmisten kulutus- ja sijoituspäätöksiin, joten inflaation tulevien arvojen tarkka ennustaminen on kuulunut inflaatiotutkimuksen ydinkysymyksiin.

Inflaation ennustaminen vaikuttaa myös keskuspankkien harjoittamaan rahapolitiikkaan. Hintavakauden ylläpitämisestä on tullut keskuspankkien tärkein tehtävä. Hintavakauden saavuttamiseksi keskuspankit käyttävät useimmiten työkaluja, joissa inflaation tulevien arvojen ennustaminen on tärkeää. Lisäksi keskuspankin rahapolitiikan vaikutukset näkyvät viiveellä, joten inflaation ennustamisesta on tullut erittäin tärkeä osa keskuspankkien toimintaa ja keskuspankkien toiminnan arvioimista.

Inflaation ennustaminen ei ole kuitenkaan ollut helppoa. Taloustieteen teoriaan perustuvat mallit eivät ole kyenneet ennustamaan inflaatiota merkittävästi paremmin kuin yksinkertaiset inflaation aiempiin arvoihin perustuvat mallit. Vaikuttaisi siltä, että mallien ennustevoima riippuu paljon ympäröivästä taloudellisesta tilanteesta. Tarkempaa tutkimusta inflaation ennustamisesta esimerkiksi taloudellisten taantumien aikana ei kuitenkaan ole tehty.

Neuroverkkomallit ovat kasvattaneet kiinnostusta viimeisen kahden vuosikymme-

nen aikana monissa taloustieteen sovelluksissa, mutta inflaation ennustamiseen niitä on käytetty vielä suhteellisen vähän, ja tulokset ovat olleet melko ristiriitaisia. Neuroverkkomallien oletetaan toimivan hyvin tilanteissa, joissa dataa generoiva prosessi on monimutkainen ja epälineaarinen. Tästä syystä neuroverkkomallit voivatkin sopia hyvin taloudellisten taantumien aikaiseen ennustamiseen.

Tässä tutkielmassa pyritään paikkaamaan aiemman tutkimuksen aukkoja. Ensinnäkin neuroverkkomallien kykyä ennustaa euroalueen inflaatiota on tutkittu erittäin vähän, koska tutkimuksen pääpaino on ollut Yhdysvaltojen inflaatiossa. Euroopan keskuspankin rahapolitiikka on kuitenkin sidoksissa inflaation tuleviin arvoihin, joten todistettavasti hyvien ennustemallien rakentaminen parantaisi keskuspankin harjoittamaa rahapolitiikkaa.

Toiseksi inflaation ennustamista ei ole tutkittu ollenkaan taloussuhdanteiden eri vaiheissa. Jos eri ennustemallit toimivat paremmin eri suhdannetilanteissa, voisi keskuspankki käyttää tietoa suhdannetilanteesta hyväkseen inflaation ennustamisessa ja antaa enemmän painoarvoa ennusteille, jotka tuottavat tarkempia ennusteita sen hetkisessä suhdannetilanteessa.

Henkilökohtaisesti aihe on kirjoittajalle tärkeä, koska kirjoittaja työskentelee data-analyytikkona rahoitusyhtiössä, joten inflaation ennustamismallien tuntemus on ammatillinen etu. Lisäksi kirjoittaja on kiinnostunut taloustieteessäkin yleistyneistä koneoppimismenetelmistä ja niiden mahdollisuuksista rahapolitiikan sovelluksiin.

Tutkielmassa pyritään vastaamaan seuraaviin tutkimuskysymyksiin:

1. Kuinka hyvin neuroverkkomalli ennustaa euroalueen inflaatiota?
2. Eroaako neuroverkkomallin ennustetarkkuus taloussuhdanteiden eri vaiheissa?
3. Ennustaako neuroverkkomalli inflaatiota merkittävästi paremmin taantumien aikana kuin lineaarinen ARMA-malli?

Tämän maisterintutkielman rakenne on seuraava. Luvussa kaksi esitellään neuroverkkomallien ja niiden rakentamisen teoriaa sekä aiempaa tutkimusta inflaation ennustamisesta sekä neuroverkkomalleilla että muilla taloustieteen ja ekonometrian malleilla.

Luvussa kolme käsitellään tässä tutkielmassa käytettävä aineisto, tutkielmassa käytettävä neuroverkkomalli ja ARMA-vertailumalli. Luvussa neljä esitellään tämän tutkielman tulokset, ja luvussa viisi vertaillaan tuloksia aiempaan tutkimukseen ja pohditaan tulosten yleistettävyyttä. Luvun kuusi johtopäätökset päättävät tutkielman.

2 Teoria

2.1 Neuroverkkomallit

Epälineaarista malleista keinotekoiset neuroverkot (*artificial neural network*) ovat kasvattaneet suosiotaan viimeisen kahden vuosikymmenen aikana. Neuroverkkomallit ovat matemaattisia malleja, jotka on alun perin kehitetty imitoimaan aivojen toimintaa, mikä edelleen selittää yleisellä tasolla neuroverkkomallien toimintaa. Neuroverkot keräävät aineistosta tietoa, joka heikentää tai vahvistaa neuroverkon eri osien yhteyksiä, joiden avulla neuroverkko suorittaa sille annettuja tehtäviä (Haykin 1999). Nykyisin neuroverkot enää harvemmin pyrkivät aivojen mallintamiseen, mutta niille on löydetty sovelluskohteita sekä talouden että biologian tutkimuksesta.

Taloudellisessa tutkimuksessa neuroverkkoja on käytetty aiemmin rahoituksen soveluksiin kuten konkurssien ja rahoitusvaikeuksien ennustamiseen sekä vakuutuspetosten todennäköisyyden arviointiin. Myös aikasarjaennustamisessa neuroverkkomallien suosio on kasvanut, mutta inflaation ennustamiseen niitä on sovellettu vielä hyvin vähän (Tkáč ja Verner 2016). Tkáč ja Verner (2016) ovat tehneet koontia vuosien 1995 ja 2015 välillä julkaistuista neuroverkkotutkimuksista talouden alalla, mutta 412 tutkimusartikkelista vain kuusi koski inflaatiota.

Neuroverkot koostuvat toisiinsa kytköksissä olevista yksiköistä, joita kutsutaan neuroneiksi. Kytkökset neuronien välillä voivat olla melkein millaisia vain, mikä tekee neuroverkoista joustavia aineistosta löytyville erilaisille yhteyksille. Neuroverkkomalleja on siis monia, mutta tavallisin tapa muodostaa neuroverkkomalli on jakaa neuronit kolmelle tai useammalle tasolle: sisäänottotasolle (*input layer*), yhdelle tai useammalle kätketylle tasolle (*hidden layer*) ja tulostustasolle (*output layer*) (Haykin 1999). Tällaisia monitasoisia havaittamalleja käytetään erityisesti ennustetarkoituksessa (Zhang ym. 1998).

Aikasarjaennustamisessa sisäänottotaso koostuu aikasarjan aiemmista arvoista, joiden kuvaus ennusteeksi tapahtuu kätketyillä tasoilla. Tulostustaso antaa lopullisen ennusteen halutulle periodille. Aikasarjaennustamisessa eteenpäin syöttävät monitasoiset neuroverkot ovat olleet suosituimpia. Niissä informaatio liikkuu sisäänottotasolta kät-

ketyn tason kautta tulostustasolle, mutta ei missään vaiheessa toiseen suuntaan, vaikka sekin olisi mahdollista neuroverkkomallien moninaisuuden puitteissa (Zhang 2012).

Eteenpäin syöttävää kolmitasoista neuroverkkomallia voidaan kuvata myös matemaattisesti. Esitys pohjautuu Bishopin (1995) vastaavaan, mutta esitys on muokattu soveltumaan aikasarjaneuroverkolle, jossa on yksi kätkeyty taso. Eteenpäin syöttävän neuroverkon sisäänottotasolla aikasarjan P aiemmasta arvosta muodostetaan yhtä monta painotettua summaa kuin kätkeyty tason neuroneita on eli tässä tapauksessa J kappaletta:

$$a_j = \sum_{p=0}^P w_{jp} y_{t-p} + w_{0j} \quad (1)$$

Tässä yhtälössä a_j on niin sanottu aktivaatio, jonka alaindeksi merkitsee kätkeyty tason neuronian; w_{jp} on aikasarjan p periodia aiemman arvon painokerroin neuronissa j ; y_{t-p} on aikasarjan p periodia aiempi arvo ja w_{0j} on niin sanottu harhatermi neuronissa j , joka toimii samoin kuin vakio tavallisessa regressiossa. Neuroverkon aktivaatiot muistuttavat hyvin paljon aikasarja-analyysin AR-mallia.

Neuroverkon kätkeytyllä tasolla aktivaatioita muokataan kuitenkin niin kutsutulla aktivointifunktiolla h , joka tuottaa aktivaatioista uuden vektorin, joka on muotoa:

$$z_j = h(a_j) \quad (2)$$

Arvot z_j ovat neuroverkon kätkeyty tason tulosteita. Teoriassa mikä tahansa differentioituva funktio voi toimia neuroverkkomallin aktivointifunktiona, mutta käytännössä vain joitain rajoitettuja, monotonisesti kasvavia ja differentioituvia funktioita käytetään (Zhang ym. 1998). Zhang ym. (1998) suosittelevat aikasarjaennustamiseen hyperbolista tangenttifunktiota kätkeytylle tasolle. Malliyhtälössä (2) tämä tarkoittaa sitä, että

$$h(x) = \frac{\exp(x) - \exp(-x)}{\exp(x) + \exp(-x)} \quad (3)$$

Kätkeyty taso puolestaan muodostaa tästä tulosteesta uuden aktivaation samalla

tavalla kuin sisäänottotaso:

$$a_k = \sum_{j=1}^J w_{kj} z_j + w_{0k} \quad (4)$$

Tässä aktivaatiossa k merkitsee tulostustason neuronien määrää, joka on ennustemalleissa lähes aina yksi. Tämän jälkeen tulostustason aktivaatiofunktio f muodostaa uusia aktivaatioista tulosteen, joka on neuroverkkomallin ennuste.

$$y_{t+k} = f(a_k) \quad (5)$$

Zhang ym. (1998) suosittelevat tulostustason aktivaatiofunktioiksi identiteettifunktiota.

$$f(x) = x \quad (6)$$

Neuroverkot, jotka käyttävät identiteettifunktiota tulostustason aktivointifunktiona eivät kykene ennustamaan aikasarjoja, joissa on trendi, joten aineisto pitää muokata tästä syystä stationaariseksi (Zhang ym. 1998). Kun edelliset yhtälöt yhdistetään, saadaan neuroverkon funktionaalinen muoto yhdessä lausekkeessa:

$$y_{t+k} = f\left(\sum_{j=1}^J w_{kj} h\left(\sum_{p=0}^P w_{ji} y_{t-p} + w_{j0}\right) + w_{k0}\right) \quad (7)$$

2.2 Neuroverkon estimointi

Yhtälöstä (7) nähdään, että neuroverkkomalleissa ennuste riippuu aikasarjan aikaisemmista arvoista ja mallin painokertoimista. Koska aikasarjan aikaisemmat arvot on annettu, mallin estimointi tarkoittaa painokertoimien estimointia. Tämä tapahtuu käytännössä minimoimalla ennustevirhefunktio painokertoimien suhteen. Ennustevirhefunktiona toimii perinteisissä ekonometrisissa malleissakin käytetty jäännöseliösummafunktio:

$$E = \frac{1}{2} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (8)$$

jossa y_i on aikasarjan todellinen arvo ja \hat{y}_i on neuroverkon tuottama ennuste (Bishop 1995).

Neuroverkon estimointia kutsutaan valvotuksi estimoinniksi (*supervised learning*), koska vastemuuttujan arvo on estimoinnin aikana saatavilla ja sitä käytetään mallin parametrien estimoinnissa hyödyksi (Zhang ym. 1998). Neuroverkon estimointi on rajoittamaton epälineaarinen minimointiongelma, jossa jäännösneliösummafunktiota minimoidaan tietyllä estimointialgoritmilla. Estimointialgoritmia joudutaan käyttämään, koska minimointiongelmaa ei ole analyttistä ratkaisua, vaan joudutaan turvautumaan numeerisiin menetelmiin. Neuroverkkomalleissa voidaan käyttää monia eri algoritmeja, mutta suosituin on vastavirta-algoritmi (*backpropagation algorithm*), jonka tarkoitus on löytää globaali minimi.

Vastavirta-algoritmissa on kaksi vaihetta: eteenpäin ja taaksepäin virtaus. Ensin eteenpäin virtauksessa neuroverkon painokertoimet valitaan satunnaisesti ja esimerkiksi vektorit syötetään neuroverkon läpi käyttäen yhtälöä (7), josta saadaan neuroverkon tuottamat ennusteet. Tämän jälkeen taaksepäin virtauksessa tulostustason virhe lasketaan kaavalla

$$\delta_k = y_{t+1} - \hat{y}_{t+1} \quad (9)$$

Yhtälössä $\delta_k = \frac{\partial E_n}{\partial a_k}$ eli virhefunktion osittaisderivaatta tulostustason aktivoinnin suhteen. Seuraavaksi kätkeyn tason virhe arvioidaan kaavalla:

$$\delta_j = h'(a_j) \sum_k w_{kj} \delta_k \quad (10)$$

Yhtälössä $\delta_j = \frac{\partial E_n}{\partial w_{jp}}$ on virhefunktion osittaisderivaatta kätkeyn neuronin j viiveen p painokertoimen suhteen, ja $h'(a_j)$ on aktivointifunktion derivaatta kätkeyn tason aktivaation suhteen. Tämän jälkeen lasketaan derivaatta

$$\frac{\partial E_n}{\partial w_{ji}} = \frac{\partial E_n}{\partial a_j} \frac{\partial a_j}{\partial w_{jp}} = \delta_j z_i \quad (11)$$

Tämän jälkeen muodostetaan virhefunktion gradientti ja päivitetään painokertoimet gradienttimenetelmällä:

$$\mathbf{w}^{\tau+1} = \mathbf{w}^{\tau} - \eta \nabla E(\mathbf{w}^{\tau}) \quad (12)$$

Algoritmi toistaa tätä prosessia niin kauan, että validaatio-otoksen virhefunktion arvo on mahdollisimman pieni. Yhtälössä (12) η on oppimiskerroin, joka määrittelee kuinka paljon kertoimet muuttuvat gradientin osoittamaan suuntaan. Korkeampi oppimistaso nopeuttaa mallin estimointia, mutta liian korkeana se voi estää painokertoimien konvergoitumisen. Lisäksi yhtälössä $\nabla E(\mathbf{w}^T)$ on virhefunktion gradientti ja \mathbf{w} on painokertoimien matriisi (Bishop 1995).

Vastavirta-algoritmia pidetään yleisesti toimivana estimointialgoritmina, ja se onkin ollut suosituin valinta aikaisemmissa tutkimuksissa. Vastavirta-algoritmi voi kuitenkin olla melko hidas, ja siksi joissain tutkimuksissa on käytetty esimerkiksi toisen asteen metodeja, jotka konvergoituvat globaaliin minimiin nopeammin (Zhang ym. 1998). Samoin inflaatiota ennustavissa aikaisemmissa tutkimuksissa on käytetty muitakin algoritmeja. Esimerkiksi Nakamura (2005) käytti Levenberg-Marquadt-algoritmia Yhdysvaltojen inflaation neuroverkkoonustamisessa.

Myös itse vastavirta-algoritmiin on tehty muutoksia, jotka ovat nopeuttaneet mallin konvergoitumista. Esimerkiksi Riedmiller ja Braun (1993) ovat kehittäneet kimmoisan vastavirta -algoritmin (*resilient backpropagation*), joka käyttää vastavirta-algoritmin derivaattatermien arvojen sijaan vain niiden etumerkkejä. Riedmiller ja Braun (1993) osoittavat myös, että kimmoisa vastavirta -algoritmi toimii nopeammin ja tarkemmin kuin perinteinen vastavirta-algoritmi tai moni muu estimointialgoritmi.

2.3 Inflaatio ja rahapolitiikka

Hintavakaudella tarkoitetaan yleisesti hyödykkeiden hintojen vaihtelun vähäisyyttä tai jopa inflaation ja deflaation puuttumista kokonaan. Inflaatiolla tarkoitetaan yleistä hintojen nousua ja rahan arvon heikkenemistä taloudessa. Käytännössä inflaatio tarkoittaa sitä, että sadalla eurolla saa ostettua enemmän hyödykkeitä tänään kuin samalla rahalla huomenna. Deflaatio on vastaavasti hintojen yleistä laskua.

Inflaatiolle on monia mittareita. Tavallisin inflaation mittari on kuluttajahintaindeksi, joka mittaa tyypillisen kuluttajan kulutuskorin hinnan kuukausittain. Joissain mittareissa kulutuskorista poistetaan hinnoiltaan epävakaita tuotteita kuten ruoka ja polttoaine niin sanotun pohjainflaation mittaamiseksi. Euroalueella on käytössä harmo-

nisoitu kuluttajahintaindeksi, joka mittaa euroalueen maiden tyypillisen kulutuskorin hintaa (ECB 2011).

Hintavakaudella on monia positiivisia vaikutuksia talouteen. Fischer (1993) on osoittanut, että inflaatio vaikuttaa negatiivisesti talouskasvuun pienentämällä investointeja ja tuottavuuden kasvua. Erityisesti odottamaton inflaatio aiheuttaa epävarmuutta yhteiskunnassa, mikä voi johtaa resurssien tehottomaan allokaatioon ja varallisuuden satunnaiseen uudelleenjakautumiseen (Briault 1995).

Rahapolitiikalla voidaan kuitenkin vaikuttaa inflaatioon. Rahapolitiikalla tarkoitetaan keinoja, joilla pyritään vaikuttamaan talouteen rahan tarjonnan kautta. Teoriassa rahapolitiikan tavoitteena voisi siis olla niin työllisyyden kuin talouskasvunkin parantaminen. Friedmanin (1968) mukaan keskuspankkien tulisi kuitenkin tavoitella asioita, joihin ne pystyvät oikeasti vaikuttamaan. Maat, joissa rahan tarjonta on ollut suurta, ovat kärsineet myös korkeasta inflaatiosta, joten Friedman (1968) päätteli, että inflaatio on lähtökohtaisesti riippuvaista rahan tarjonnasta, ja rahan tarjonta ei vaikuta pitkällä aikavälillä muihin makromuuttujiin. Raha on niin sanotusti neutraalia pitkällä aikavälillä. Tämän takia myös Euroopan keskuspankki on ottanut hintavakauden ensisijaiseksi tavoitteeksi rahapolitiikassaan. Euroopan keskuspankki pystyy vaikuttamaan pitkällä aikavälillä inflaatioon, mutta ei esimerkiksi reaalipalkkoihin tai työllisyyteen (ECB 2011).

Keskuspankit ovat historian saatossa käyttäneet sekä rahamäärätavoitetta että myös vaihtokurssitavoitetta hintavakauden saavuttamiseksi. Viimeisten vuosikymmenten aikana monet keskuspankit ovat kuitenkin ottaneet käyttöönsä inflaatiotavoitteen eli julkisesti ilmoitetun inflaation arvon tai arvovälin, jolle inflaatio pyritään saamaan tietyllä aikavälillä rahapolitiikan avulla. Inflaatiotavoitteen yksi tarkoitus on ankkuroida yksityisen sektorin inflaatio-odotukset, jotta ne eivät vaikuttaisi hintavakauteen. Inflaatiotavoite vaatii kuitenkin avoimuutta ja vastuullisuutta keskuspankilta, jotta yksityinen sektori pystyy luottamaan inflaatiotavoitteen saavuttamiseen (Svensson 2010).

Euroopan keskuspankki ei ole ottanut virallisesti inflaatiotavoitetta rahapolitiikkansa työkaluksi, mutta se toteuttaa hyvin pitkälti samanlaista rahapolitiikkaa kuin inflaatiotavoitteeseen sitoutuneet keskuspankit. Euroopan keskuspankin tavoite on pitää har-

monisoidulla kuluttajahintaindeksillä mitattu inflaatio 2 %:n tuntumassa keskipitkällä aikavälillä. Lisäksi Euroopan keskuspankki korostaa läpinäkyvyyden, itsenäisyyden ja vastuun merkitystä omassa rahapolitiikassaan inflaatio-odotusten ankkuroimiseksi (ECB 2011). Euroopan keskuspankki ei suoranaisesti huolehdi esimerkiksi talouskasvusta tai työllisyydestä, vaikka monet inflaatiotavoitteeseen sitoutuneet keskuspankit niin tekevätkin (Svensson 2010), mutta se pyrkii luomaan täystyöllisyyden ja vakaan talouskasvun mahdollistavan ympäristön hintavakauden avulla. Lisäksi reaalityaloudesta huolehtiminen on koko eurojärjestelmän tehtävä (ECB 2011).

Svensson (1997) on osoittanut, että inflaatiotavoite on käytännössä kuitenkin aina inflaatioennustetavoite. Inflaatiotavoitteen käyttö on vaikeaa, koska inflaatiota on hankala ohjata ja seurata. Rahapolitiikan päätökset näkyvät viiveellä ja rahapolitiikka ei ole ainoa inflaatioon vaikuttava asia. Inflaatioennustetavoite on siinä mielessä optimaalinen, että se korreloi voimakkaasti tulevaisuuden inflaation kanssa, inflaatio-odotuksia on helpompi ohjata kuin nykyistä inflaatiota, inflaatio-odotukset ovat helpommin havaittavissa kuin inflaatio, ja inflaatioennusteet lisäävät keskuspankin läpinäkyvyyttä. Inflaation ennustaminen antaa tässä mielessä myös keinon arvioida keskuspankin sekä nykyistä että menneitä toimintaa, ja rahapolitiikan päätöksistä tulee ennustettavampia.

Myös Euroopan keskuspankki toteaa, että rahapolitiikan tulee perustua tulevaisuuden näkyviin rahapolitiikan vaikutusten viiveen vuoksi. Siksi Euroopan keskuspankin rahapolitiikan tavoitteet on asetettu keskipitkälle aikavälille, joka voi vaihdella tilanteen mukaan. Vaikka Euroopan keskuspankki ei keskity rahapolitiikassaan ainoastaan inflaatioennusteeseen, vaan ottaa huomioon sekä reaalityalouden tilanteen että rahapolitiikan näkymät kahden pilarin strategiassaan, inflaatioennusteilla on tärkeä rooli rahapolitiikassa. Euroopan keskuspankki julkaiseekin säännöllisesti inflaatioennusteen seuraavalle vuodelle (ECB 2011).

Inflaation ennustaminen on siis tärkeää sekä rahapolitiikan että keskuspankkien toiminnan arvioinnin kannalta. Jos inflaatioennuste poikkeaa merkittävästi esimerkiksi inflaatiotavoitteesta, keskuspankki voi kasvattaa tai pienentää rahan tarjontaa, mikä tulee vaikuttamaan viiveellä tulevaisuuden inflaatioon. Lyhyellä aikavälillä keskuspankki

pyrkii vaikuttamaan inflaatioennusteisiin, johon yksityinen sektori perustaa omat päätöksensä. Parempien inflaatioennusteiden tekeminen siis hyödyttää sekä keskuspankkeja että yksityisen sektorin toimijoita.

2.4 Inflaation ennustaminen

Taloustieteessä inflaation ennustemallit voidaan jakaa ryhmiin sen mukaan mitä muuttujia mallissa käytetään inflaation ennustamiseen. Ennusteet voivat perustua pelkästään inflaation aiempiin arvoihin, reaalityalouden aktiivisuutta kuvaaviin muuttujiin kuten työttömyysasteeseen tai bruttokansantuotteeseen, muiden tekemiin ennusteisiin inflaatiosta tai aivan muihin inflaatiota oletettavasti selittäviin muuttujiin (Stock & Watson 2009).

Inflaation aiempiin arvoihin perustuvat mallit ovat yksinkertaisia, joten niitä käytetään yleisesti muiden ennustemallien vertailukohtana. Uusi ennustemalli ei tuo lisäarvoa, jos yksinkertainen inflaation aiempiin arvoihin perustuva malli tuottaa paremman ennusteen. Hyvä esimerkki tällaisista malleista on ensimmäisen asteen autoregressiivinen AR(1)-malli

$$\pi_{t+1} = \phi_1 \pi_t + \epsilon_t \tag{13}$$

jossa tulevaisuuden inflaatiota ennustetaan pelkästään inflaation kuluvan periodin arvolla. Yhtälössä ϵ_t on valkoista kohinaa vakiokeskiarvolla 0 ja vakiovarianssilla, ja kerroin ϕ täyttää ehdon $|\phi| < 1$ (Brockwell & Davis 2002).

Aktiivisuusmuuttujiin perustuvat mallit ovat olleet inflaation ennustamisen keskiössä taloustieteessä, koska ne perustuvat taloustieteen teoriaan inflaatiosta. Phillipsin käyrään perustuvasta inflaatiomallista, jossa inflaatiota ennustetaan inflaation aiemmilla arvoilla ja työttömyydellä, on tullut käytetyin malli inflaation ennustamisessa (Stock & Watson 2009). Inflaation teoriaan perustuvien mallien suosiota voidaan perustella sillä, että ne selittävät inflaation yhteyttä muihin talouden muuttujiin. Siksi Phillipsin käyrään perustuvia ennustemalleja on käytetty laajasti myös keskuspankkien inflaatioennusteissa.

Talouden aktiivisuuteen perustuvien mallien ongelma on ollut se, että ne eivät ole ennustaneet tulevaisuuden inflaatiota systemaattisesti paremmin kuin yksinkertaiset pelkkään inflaatioon perustuvat mallit. Yhtenä kuuluisimpana esimerkkinä Atkeson ja Ohanian (2001) vertailivat kolmea yleisesti käytettyä Phillipsin käyrään perustuvaa mallia yksinkertaiseen ennustesääntöön, jossa inflaation odotettiin pysyvän muuttumattomana edellisestä vuodesta. Yksikään tutkituista ennustemalleista ei pärjännyt yksinkertaiselle ennusteelle, ja Atkesonin ja Ohanianin mukaan Phillipsin käyrään perustuvat mallit eivät ole kolikonheittoa parempia ennustamaan inflaatiota.

Stock ja Watson (2009) ovat tiivistäneet tähän ongelmaan liittyviä tutkimuksia 15 vuoden ajalta, ja he toteavat, että Phillipsin käyrään perustuvien mallien ennustevoima riippuu inflaation ja yleisesti muun talouden tilanteesta. Phillipsin käyrään perustuvat mallit toimivat parhaiten aikoina kun työttömyys on kaukana tasapainotasostaan, ja vastaavasti huonommin silloin kun työttömyys on lähellä tasapainotasoa. Stock ja Watson (2007) tiivistävät tämän ongelman siihen, että inflaation ennustamisesta on tullut helpompaa inflaation vaihtelun pienentymisen vuoksi, jolloin lineaariset yksiulotteiset mallit toimivat hyvin. Toisaalta taas ennustamisesta on tullut vaikeampaa, koska useamman selittävän muuttujan mallien on vaikeampi tuottaa lisähyötyä verrattuna yksinkertaisiin ennustesääntöihin.

Sen lisäksi, että taloustieteen teoriaan perustuvat ennustemallit eivät ole aina olleet yksinkertaisia ennustesääntöjä parempia, monet muutkin ennusteenluomistavat voivat tuottaa parempia tuloksia. Ang ym. (2007) ovat osoittaneet, että kyselyihin perustuvat ennusteet tuottavat tarkempia ennusteita tulevaisuuden inflaatiosta kuin mikään perinteisistä ennustemalleista mainittuna yllä. Jopa tavallisten ihmisten odotukset ovat osuneet paremmin oikeaan kuin monimutkaiset matemaattiset mallit. He perustelevat tätä ilmiötä sillä, että kyselyihin varastoituu paljon enemmän informaatiota kuin talouden havaittuun toimintaan. Samaan lopputulokseen tulevat myös Faust ja Wright (2013), jotka toteavat, että subjektiiviset inflaatioennusteet ovat selkeästi lähempänä inflaation todellista arvoa kuin taloustieteen malleihin perustuvat ennusteet.

Zhang (2012) toteaaakin, että ei ole olemassa yhtä parasta mallia, jota voidaan käyttää kaikkiin tilanteisiin, vaan paras mahdollinen ennustemalli riippuu monesta tekijästä.

Siksi inflaation ennustamisessa on ehdotettu myös epälineaaristen mallien käyttämistä. Jos inflaation dataa generoiva prosessi on lähtökohtaisesti epälineaarinen, taloustieteen lineaariset mallit eivät kykene mallintamaan sitä oikein. Stock ja Watson (2009) kuitenkin toteavat, että epälineaariset mallit eivät ole onnistuneet parantamaan ennustevoimaa merkittävästi. Phillipsin käyrään perustuvien ennustemallien analyysistä voidaan kuitenkin olettaa, että paras ennustemalli riippuu ennustetilanteesta. Tässä tutkielmassa keskitytäänkin hyvin rajattuun tilanteeseen, taloudelliseen taantumiaan, jolloin epälineaaristen mallien hyödyt voivat olla suuremmat kuin tavallisesti.

2.5 Inflaation ennustaminen neuroverkkomallien avulla

Yksi varhaisimmista neuroverkkojen sovellutuksista inflaation ennustamiseen on Aiken (1999), joka ennusti Yhdysvaltojen kuluttajahintaindeksillä mitattua inflaatiota yhden kvartaalin eteenpäin erilaisten hintaindeksien ja rahan tarjonnan avulla. Aikenein mukaan tällainen neuroverkkomalli pystyy ennustamaan inflaatiota melko tarkasti absoluuttisella virheellä mitattuna. Tutkimuksessa ei kuitenkaan otettu huomioon neuroverkkomallien mallinvalintakysymyksiä, vaan tulokset saatiin käyttämällä valitun tilastollisen ohjelman oletusarvoja. Samoin selittävät muuttujat on valittu visuaalisella arvioinnilla mahdollisista inflaatiota ennustavista muuttujista.

Vastaavasti Gupta ja Kashyap (2015) ennustivat Yhdysvaltojen lisäksi Iso-Britannian, Ranskan, Saksan, Italian, Japanin ja Kanadan inflaatiota vuoden päähän, mutta pyrkivät etsimään parhaan mahdollisen yhdistelmän sisääntulotason, kätketyn tason ja tulostason neuronien määrälle eri maissa. Gupta ja Kashyap (2015) käyttivät tutkimuksessaan kuukausittaista kuluttajahintaindeksiä vuosilta 1996-2013. Tulokset osoittivat, että neuroverkot pystyvät ennustamaan inflaatiota todella tarkasti kun neuroneiden määrä on valittu mahdollisimman optimaalisesti. Arviointikriteereinä käytettiin virheiden neliösummaa, neliösumman keskiarvon neliöjuurta sekä absoluuttisten virheiden keskiarvoa (ks. osio 3.5). Saatuja tuloksia ei kuitenkaan verrattu muihin inflaatioennustemalleihin, joten epälineaaristen neuroverkkomallien tuomaa lisähyötyä ei ole voitu arvioida.

Monissa tutkimuksissa onkin käytetty erilaisia vertailumalleja, joihin neuroverko-

jen ennustevoimaa on verrattu. Choudhary ja Haider (2012) vertailivat kahta erilaista neuroverkkomallia ja kahta näiden yhdistelmää yksinkertaiseen yhtälön (13) mukaiseen AR(1)-malliin. He käyttivät kuukausittaista kuluttajahintaindeksiä 28 OECD-maasta vuosilta 1991-2008, ja tekivät ennusteita kuukauden, kolmen kuukauden ja vuoden päähän. keskineliövirheen neliöjuurella sekä absoluuttisella keskiprosenttivrheellä mitattuna neuroverkkomalli voitti yksinkertaisen AR(1)-mallin 45 %:ssa tapauksista yksi ja kolme kuukautta eteenpäin. AR(1)-malli oli kuitenkin parempi ennustamaan inflaation arvon 23 %:ssa tapauksista, ja lopuissa maissa eroa ennusteiden välille ei voitu tehdä. Tutkimus antaa todisteita siitä, että neuroverkkojen hyödyt riippuvat käytetystä tutkimusaineistosta.

Toisaalta ennustevälin kasvattaminen parantaa neuroverkkomallien suoritusta, koska kun ennusteet tehtiin vuosi eteenpäin, neuroverkot olivat ehdottomasti parempi ennustemalli 71 %:ssa tutkituista maista molemmilla arviointikriteereillä (Choudhary ja Haider 2012). Rahapolitiikka tähtää yleisesti pitemmille aikaväleille, joten neuroverkkojen käytölle voi olla tässäkin mielessä käyttöä inflaation ennustamisessa.

Samoin Nakamura (2005) toteaa, että neliövirheiden keskiarvolla arvioituna neuroverkkomalli ennustaa Yhdysvaltojen inflaatiota paremmin kuin useat yksinkertaiset AR-mallit, joissa käytetään aiempia inflaation arvoja enintään kahdeksalta edelliseltä periodilta. Aiemmista tutkimuksista poiketen Nakamura (2005) käyttää inflaation mittarina Yhdysvaltojen kvartaalin välein mitattua BKT-deflaattoria vuosilta 1960-2003. Ennusteet tehtiin 1-4 kvartaalia eteenpäin. Nakamura (2005) toteaa, että neuroverkkomallin valinnalla ja mallin estimoinnilla on tärkeä rooli neuroverkon ennusteille. Tämä tukee Guptan ja Kashyapin (2015) havaintoa optimaalisen neuronimäärän valinnasta.

Nakamuran (2005) ja Choudharyn ja Haiderin (2012) tutkimuksissa ei ole kuitenkaan suoritettu tilastollisia testejä erojen luotettavuuden selvittämiseksi. Myös AR(p)-malli voi olla liian yksinkertainen käytettäväksi vertailumallina. Jos tutkittu ennustemalli tuottaa pienellä erolla parempia ennusteita kuin yksinkertaisin mahdollinen vertailumalli, tutkittua ennustemallia ei voida kuitenkaan pitää vielä parhaana mahdollisena mallina.

Näin toteavat myös Álvarez-Díaz ja Gupta (2016), jotka tutkivat onko inflaation epä-

lineaarisuudella merkitystä Yhdysvaltojen kuukausittaisen inflaation ennustamisessa. He käyttivät AR(1)- ja satunnaiskulkumallien (*random walk*) lisäksi SARIMA-mallia, joka ottaa huomioon aineiston kausivaihtelun. He käyttivät aineistona kuukauden välein mitattua kuluttajahintaindeksiä tammikuusta 1980 joulukuuhun 2013. Tutkimuksessa huomattiin SARIMA-mallin tuottavan pienempiä ennustevirheitä absoluuttisella keskiprosenttivarheella mitattuna, mutta ero neuroverkkomalliin oli erittäin pieni sekä koulutusotoksen sisällä että sen ulkopuolella. Álvarez-Díaz ja Gupta (2016) huomasi- vatkin, että ero kahden ennusteen välillä ei ollut tilastollisesti merkitsevä, vaan sekä neuroverkko- että SARIMA-malli ennustivat inflaatiota yhtä hyvin.

Euroalueella Binner ym. (2005) vertailivat lineaaristen ARIMA- ja VAR-mallien en- nustekykystä epälineaarisiin neuroverkkoihin kolmen eri ennustevirhemittarin (keskine- liövirheen neliöjuuri, absoluuttisten virheiden keskiarvo ja absoluuttinen keskiprosent- tivirhe) avulla. Heidän aikasarjansa koostui kvartaalidatasta vuosilta 1981-2000. Jälleen erot ennusteiden välillä olivat pieniä, mutta Binner ym. (2005) myös päättelivät, että VAR-mallin informaatio sisältyy kokonaisuudessaan neuroverkkomallin informaatioon. Tutkimus antaa olettaa, että inflaatio on euroalueella epälineaarinen ilmiö, jolloin neu- roverkko osoittautuu lineaarisia malleja paremmaksi ennustemalliksi. Tämä antaa luot- toa euroalueen inflaation ennustamiseen neuroverkkojen avulla myös vuosien 2007-2009 taantumana aikana, jota Binnerin ym. (2005) tutkimus ei kata.

Neuroverkkomalleja on myös vertailtu laajempaan joukkoon erilaisia malleja. Mos- hiri ja Cameron (2000) kilpailuttivat neuroverkkomallia kolmen strukturaalisen mallin ja kolmen aikasarjamallin kanssa käyttäen Kanadan inflaatioaikasarjaa vuosilta 1970- 1994. Moshiri ja Cameron (2000) tutkivat sekä dynaamista että staattista ennustamista absoluuttisen keskivirheen ja keskineliövirheen neliöjuuren avulla yksi, kolme ja kaksi- toista kuukautta eteenpäin.

Dynaaminen ennustaminen tarkoittaa sitä, että mallin avulla ennustetaan yksi pe- riodi eteenpäin, mutta ennustetta käytetään tuottamaan seuraavien periodien ennusteita kunnes halutun periodin ennuste voidaan tuottaa. Dynaamisen ennustamisen vaih- toehtona olisi staattinen ennustaminen, jossa tämän hetken havaintoja käytettäisiin suoraan 12 periodin päähän menevään ennusteeseen (Moshiri ym. 1999).

Dynaamisessa ennustamisessa keskineliövirheen neliöjuurella mitattuna ARIMA-malli ennusti parhaiten yhden kuukauden eteenpäin, BVAR-malli kolme kuukautta eteenpäin ja neuroverkkomalli 12 kuukautta eteenpäin. Keskineliövirheellä mitattuna ARIMA-malli ennusti parhaiten yhden kuukauden eteenpäin, BVAR- ja neuroverkkomalli olivat yhtä hyviä kolme kuukautta eteenpäin, ja neuroverkko päihitti muut mallit 12 kuukautta eteenpäin (Moshiri & Cameron 2000).

Staattisessa ennustamisessa ARIMA-malli oli muita malleja parempi keskineliövirheen neliöjuurella mitattuna sekä kolme että kaksitoista kuukautta eteenpäin. Neuroverkkomallit onnistuivat muita paremmin tai yhtä hyvin vain yhdellä kuukaudella eteenpäin keskineliövirheellä mitattuna (Moshiri & Cameron 2000). Tutkimuksessa neuroverkkojen syötteenä käytettiin vertailumallien mukaisia muuttujia, mikä voi vinouttaa tuloksia lineaaristen mallien hyväksi. Jos tutkimuksessa olisi käytetty neuroverkkomalleille optimaalisia muuttujia tai inflaation viivearvoja, tulokset olisivat voineet olla toisenlaisia.

Samoin He ym. (2012) vertailivat neuroverkkomalleja AR-, ARIMA-, ARIMA-GARCH- ja eksponentiaalinen tasoitus (*exponential smoothing*) -malleihin Yhdysvaltojen inflaation ennustamisessa yksi ja kaksi kuukautta eteenpäin. He käyttivät aineistona Yhdysvaltojen kuukausittaista inflaatiota tammikuusta 1990 huhtikuuhun 2011. Tässä tapauksessa neuroverkkomalli hävisi ARIMA- ja ARIMA-GARCH-malleille kaikilla ennusteväleillä ja -mittareilla. Erot ennusteiden välillä eivät olleet suuria, ja ero neuroverkkomalliennusteen ja parhaimman ennustemallin välillä oli 0,0229 prosenttiyksikköä. He ym. (2012) eivät tutkineet eroja tai niiden tilastollista merkitsevyyttä, joten tutkimuksen perusteella ei voida olla varmoja ovatko erot ennusteissa todellisia.

Miten neuroverkkomallit pärjäävät sitten muille epälineaarille malleille? Binner ym. (2006) vertailivat neuroverkkomallia epälineaariseen Markov-vaihto autoregressiiviseen malliin (*Markov switching autoregressive model*, MS-AR) sekä AR-malliin eri ennusteväleillä samalla aineistolla kuin Nakamura (2005). He tekivät ennusteita samoin 1-4 kvartaalia eteenpäin. Binner ym. (2006) päättelivät, että MS-AR-malli on parempi inflaation ennustamisessa kuin neuroverkko, koska se voitti yksinkertaisen AR-mallin useammassa tapauksessa kuin neuroverkko. Tulokset ovat myös tilastollisesti merkit-

seviä suurimmassa osassa tapauksia. Neuroverkkomalli oli lyhyillä ennusteväleillä melkein yhtä hyvä kuin MS-AR-malli, mutta pitemmillä väleillä ero kasvoi. Pitemmillä aikaväleillä myös neuroverkkomallin ja AR-mallien erot kaventuivat, mikä on ristiriidassa esimerkiksi Choudhary ja Haiderin (2012) tutkimukseen. Toisaalta Binnerin ym. (2006) tutkimuksessa tarkasteltiin ainoastaan Yhdysvaltoja, joten Choudharyn ja Haiderin (2012) löytämät tulokset voivat näkyä muiden maiden vaikutuksesta.

Neuroverkkomalleja on myös vertailtu keskenään. Moshiri ym. (1999) vertailivat perinteistä eteenpäin syöttävää neuroverkkomallia VAR-mallin ja strukturaalisen mallin lisäksi kahteen erilaiseen neuroverkkomalliin: uusiutuvaan neuroverkkoon, jossa informaatio liikkuu neuroverkossa myös taaksepäin, ja neuroverkkoon, jossa käytettiin harvinaisempaa aktivointifunktiota. Moshiri ym. (1999) käyttivät kanadalaista dataa vuosilta 1970-1990, ja muodostivat ennusteita yhden, kolmen ja kahdentoista kuukauden päähän. He tutkivat sekä dynaamista että staattista ennustamista. Dynaamisessa ennustamisessa keskineliövirheen neliöjuurella mitattuna voittajia olivat eteenpäin syöttävä neuroverkko ja VAR-malli yksi periodi eteenpäin ja uusiutuva neuroverkko kolme ja kaksitoista periodia eteenpäin. Keskineliövirheellä mitattuna VAR-malli oli paras yksi periodi eteenpäin ja uusiutuva neuroverkko kolme ja kaksitoista kuukautta eteenpäin. Staattisessa ennustamisessa ei ollut yhtä selkeää voittajaa millään ennusteperiodilla tai -mittarilla. Mänttämisen arvoista on se, että staattisessa ennustamisessa uusiutuva neuroverkko tuotti keskimäärin kaksi kertaa suurempia ennustevirheitä kuin tavallinen eteenpäin syöttävä neuroverkko.

Kirjallisuuden perusteella on siis selvää, että selkeästi yhtä parasta mallia ei ole löydetty, vaan mallin, muuttujien ja arviointikriteerien valinta vaikuttaa siihen onko neuroverkkomallista hyötyä inflaation ennustamisessa. Toisaalta erot ovat olleet hyvin pieniä lineaaristen mallien ennusteisiin nähden, mikä kasvattaa epäilyksiä neuroverkkomallien tuomasta lisäarvosta. Aiemmissä tutkimuksissa ei ole otettu kuitenkaan huomioon talouden suhdannetilannetta, ja esimerkiksi inflaation ennustamista taloudellisten taantumien aikana ei ole tutkittu ollenkaan. Tutkimuksissa käytettävät aikasarjat sisältävät selkeästi taantumia, ja esimerkiksi Álvarez-Díaz ja Gupta (2016) huomioivat vuosien 2007-2009 taantumien olemassaolon aineistossaan, mutta eroihin ennusteissa eri

suhdannetilanteissa ei ole otettu kantaa.

Myös koko euroalueen inflaation ennustaminen on jäänyt tutkimuksessa vähemmälle huomiolle, vaikka Euroopan keskuspankki tekee rahapoliittiset päätökset koko euroalueen kokonaistoiminnan perusteella. Tutkimuksen pääpainopiste onkin ollut Yhdysvaltojen ja muiden yksittäisten suurten maiden inflaation ennustamisessa. Vain Binner ym. (2005) vertailivat neuroverkkoa lineaarisiin malleihin euroalueen inflaation ennustamisessa, mutta heidänkään tutkimuksensa ei kattanut niin sanotun suuren taantumman aikaa. Tämä tutkielma pyrkii vastaamaan näihin puutteisiin tutkimalla erityisesti inflaation ennustamista vuosien 2007-2009 taantumman aikana euroalueella.

3 Metodit ja tutkimusaineisto

3.1 Neuroverkkomalli

Tässä osiossa esitellään tarkemmin tutkielmaan valittuja tilastollisia malleja ja tutkielmassa käytettävää tutkimusaineistoa. Lisäksi esitellään strategiat, joilla tutkielman johdannossa esitettyihin tutkimuskysymyksiin vastataan.

Tässä tutkielmassa neuroverkkomalliksi on valittu luvussa 2.1 esitelty eteenpäin syöttävä neuroverkkomalli, jonka aktivointifunktiona käytetään kaavan (3) mukaista hyperbolista tangenttifunktiota. Eteenpäin syöttävä neuroverkko on valittu esimerkiksi uusiutuvan neuroverkon tai muun monimutkaisemman mallin sijaan, koska eteenpäin syöttävä neuroverkko on ollut suosituin aiemmassa kirjallisuudessa. Tällöin tutkimuksen tuloksia voidaan verrata myös aiempiin tuloksiin vastaavanlaisista tutkimuksista.

Neuroverkkomalleilla on monia hyviä puolia, jotka tekevät niistä potentiaalisia vaihtoehtoja perinteisille ennustemalleille. Neuroverkkomallit eivät esimerkiksi perustu mihinkään tiukkoihin oletuksiin. Neuroverkkomallit sopivat siksi hyvin tilanteisiin, joissa taustalla oleva dataa generoiva prosessi on tuntematon, mutta havaintoja on paljon (Zhang ym. 1998). Tavallisesti ekonometriset mallit vaativat oletuksia esimerkiksi virhetermien todennäköisyysjakaumista tai kertoimien arvoista, joten neuroverkkomallien käyttö on tässä mielessä helpompaa.

Neuroverkkomallit ovat myös niin sanottuja universaaleja funktionaalisia approksimaattoreita. Tämä tarkoittaa neuroverkkomallien kohdalla sitä, että neuroverkot voivat teoriassa approksimoida yhdellä kätjetyllä tasolla mitä tahansa jatkuvaa funktiota millä tahansa tarkkuudella, jos aktivointifunktio on jatkuva, rajoitettu, ja se ei ole vakiofunktio (Hornik 1991). Neuroverkot voivat siis löytää hyvinkin monimutkaisia yhteyksiä datasta, mikä palvelee tosielämän ilmiöiden ennustamista. Tämä pätee erityisesti epälineaariin yhteyksiin, joita lineaariset mallit eivät aina kykene mallintamaan riittäväällä tarkkuudella (Zhang 2012). Jos inflaatio on epälineaarinen ilmiö, neuroverkot voivat parantaa sen ennustettavuutta huomattavasti.

Neuroverkkomallien huonoja puolia on se, että ne voivat vaatia isoja aineistoja toimiakseen parhaalla mahdollisella tavalla. Toisekseen ne eivät välttämättä tuo mitään

lisähyötyä lineaaristen yhteyksien ennustamiseen, vaan yksinkertaisemmat mallit toimivat silloin paremmin. Kolmanneksi neuroverkkomallit eivät kerro mitään eri muuttujien välisestä yhteydestä, koska neuroverkkomallin kertoimia ei voi tulkita yhtä helposti kuin vaikka regressiomallin. Lisäksi neuroverkkomallien joustavuus voi muodostua ongelmaksi, jos malli oppii koulutusaineiston erilaiset yhteydet liian hyvin, jolloin mallin ennustevoima ja yleistettävyyys heikkenee. Tällöin mallia vaivaa ylisovittaminen (*overfitting*) (Zhang ym. 1998).

Yksi neuroverkkomallien huonoista puolista on myös se, että mallinvalintaan ei ole kehitetty mitään metodia, jolla objektiivisesti paras malli monien joukosta valitaan. Entistä monimutkaisempaa asiasta tekee se, että tutkijan täytyy kokeilla eri valintoja ja parhaan mallin löytämiseksi, mikä vie aikaa. Joitain nyrkkisääntöjä ja vinkkejä on kuitenkin kehitetty mallinvalinnan helpottamiseksi, ja tässä tutkielmassa käytettävässä neuroverkkomallissa pyritään seuraamaan näitä sääntöjä aiempiin inflaation ennustamista käsitteleviin tutkimuksiin peilaten. Epäselvissä tilanteissa pyritään valitsemaan yksinkertaisempi ratkaisu, koska tarkoituksena on tehdä mahdollisimman yksinkertainen malli, joka ennustaa inflaatiota mahdollisimman hyvin.

Yksi tärkeimmistä valinnoista neuroverkon suunnittelussa on kätkeytyjen tasojen, sisäänottotason neuroneiden, kätkeytyjen tasojen neuroneiden ja tulostustason neuroneiden lukumäärä. Nämä ovat myös neuroverkkojen vaikeimpia valintoja, koska juuri näiden parametrien valintaan ei ole olemassa mitään objektiivista ohjetta.

Tässä tutkielmassa käytetään yhtä kätkeytyä tasoa. Teoriassa yhdellä kätkeytyllä tasolla pitäisi pystyä approksimoimaan mitä tahansa jatkuvaa funktiota (Hornik 1991). Toisaalta Zhang ym. (1998) perustelevat kahden kätkeytyksen tason käyttämistä paremman ennustetarkkuuden saavuttamiseksi. Lisäksi yksi kätkeytyksen taso voi tarvita paljon kätkeytyksiä neuroneita toimiakseen kunnolla, mikä toisaalta kasvattaa ylisovittamisen riskiä. Myös useamman tason käyttäminen kasvattaa ylisovittamisen riskiä, jos koulutusotos on melko lyhyt kuten makromuuttujien aikasarjoissa usein on. Kaikki inflaation ennustamiseen keskittyneet aiemmat neuroverkkotutkimukset ovat käyttäneet vain yhtä kätkeytyä tasoa, joten niin tehdään myös tässä tutkimuksessa.

Taloustieteen teoriaan perustuvassa neuroverkossa sisäänottotason neuronien luku-

määrä olisi selittävien muuttujien lukumäärä, mutta aikasarjaennustamisessa sen määrittäminen on vaikeampaa. Sisäänottotason ja kätketyn tason neuronien määrän selvittäminen vaatii usein kokeilua eri yhdistelmillä. Kaastra ja Boyd (1996) neuvovat, että neuroverkon koulutusotoksen havaintojen lukumäärän tulisi olla vähintään kaksi kertaa ja mielellään neljä kertaa yhtä suuri kuin neuroverkon painokertoimien lukumäärä, mikä asettaa ylärajan sisäänottotason neuronien lukumäärälle. Kätketyn tason neuroneiden lukumäärälle on joitain satunnaisia sääntöjä. Esimerkkinä säännöistä kätketyllä tasolla voi olla $[0.5n, 3n]$ tai \sqrt{n} verran neuroneita, jos sisäänottotason neuroneiden lukumäärä on n (Kaastra ja Boyd, 1996).

Álvarez-Díaz ja Gupta (2016) etsivät parasta mallia enintään 20 neuronin joukosta sisäänottotasolla, Binner ym. (2005) tutkivat neuroverkkoja, joissa on 1-5 kätketyä neuronua, ja Gupta ja Kashyap (2015) etsivät sopivinta mallia joukosta, jossa oli enintään kymmenen neuronua sekä sisäänottotasolla että kätketyllä tasolla. Tässä tutkielmassa tullaan käyttämään vastaavaa strategiaa, jossa etsitään paras malli joukosta, jossa neuroneiden lukumäärät vaihtelevat sisäänottotasolla ja kätketyllä tasolla 1-20 neuronin välillä siten, että kätketyllä tasolla on enintään yhtä monta neuronua kuin sisäänottotasolla.

Tulostustason neuroneiden lukumäärä vastaa tehtyjen ennusteiden lukumäärää. Tässä tutkielmassa tehdään ennusteita ainoastaan yhden periodin välein enintään 12 kuukautta eteenpäin, joten tulostustason neuroneiden lukumäärä on yksi. Aikasarjaennustamisessa neuroverkkomallilla ei myöskään suositella käytettävän useampaa tulostustasoa, vaan jokaiselle ennustevalille tulisi rakentaa oma neuroverkkomalli (Lin & Grange, 1994, ks. Zhang 2012). Dynaamisessa ennustamisessa tätä ongelmaa ei ole, koska neuroverkko rakennetaan vain tuottamaan ennuste yksi kuukausi eteenpäin.

Tässä tutkielmassa neuroverkkomallin estimoimisessa käytetään kimmoista vastavirta-algoritmia. Perinteinen gradienttilaskua käyttävä vastavirta-algoritmi on ollut suosituin estimointialgoritmi neuroverkkomallien tutkimuksessa myös inflaatioennustamisen ulkopuolella (Tkac & Verner 2015), joten myös tässä tutkielmassa sitä pidetään algoritmivalinnan lähtökohtana. Perinteinen vastavirta-algoritmi voi kuitenkin olla erittäin hidas, mikä voi aiheuttaa ongelmia esimerkiksi mallin valinnassa, jossa joudutaan esti-

moimaan satoja malleja. Kimmoisa vastavirta -algoritmi on todennetusti nopeampi ja robustimpi kuin esimerkiksi perinteinen vastavirta (Riedmiller & Braun 1993), mutta siinä on myös hyvänä puolena se, että algoritmin parametreja kuten oppimiskerrointa ei tarvitse valita itse, mikä vähentää neuroverkon mallinvalintakysymyksiä tehden valinnasta hieman helpompaa (Riedmiller 1994).

Globaalin minimin saavuttaminen ei välttämättä ole kuitenkaan helppoa. Algoritmi voi juuttua lokaaliin minimiin, vaikka iterointien määrää kasvatettaisiin. Kaastra ja Boyd (1996) suosittelevatkin valitsemaan iterointien määräksi niin suuren, että uudet iteroinnit eivät enää tee suurta muutosta painokertoimiin. Lisäksi he suosittelevat valitsemaan 5-10 eri alkuarvausta painokertoimille, joista estimointi aloitetaan. Tällöin globaalin minimin löytyminen on varmempaa. Tässä tutkielmassa tehdään kymmenen eri alkuarvausta mallin painokertoimille, jotta neuroverkkomallin valinta olisi mahdollisimman luotettavaa.

Tavallisesti neuroverkkomallien käytössä aikasarja jaetaan kolmeen osaan: Koulutus-, testi- ja validaatio-otoksiin. Koulutusotos on suurin näistä kolmesta, ja sitä käytetään neuroverkkomallin estimoimiseen. Neuroverkkomalli kuitenkin estimoidaan koulutusdatalla siten, että se minimoi ennustevirheen validaatio-otoksessa ylisovittamisen riskin pienentämiseksi. Neuroverkkomallin estimoinnin jälkeen saatua neuroverkkomallia käytetään testiotoksen arvojen ennustamiseen, millä mitataan neuroverkkomallin todellista ennustevoimaa (Kaastra & Boyd 1996).

Jako koulutus- ja validaatio-otoksiin noudattelee usein tiettyjä nyrkkisääntöjä, jotka ovat kuitenkin aina mielivaltaisia. Useimmiten koulutusotos sisältää 70-90 % aineistosta ja validaatio-otos näin ollen 10-30 %. Tässä tutkielmassa koulutusotokselle annetaan 80 % estimoimiseen ja validointiin tarkoitettua aineistosta, ja validaatio-otokselle annetaan 20 % koko koulutus- ja validointiaineistosta, koska aikasarja ei ole kovin pitkä, ja neuroverkkomalli tarvitsee mahdollisimman paljon havaintoja koulutusotokseen toimiakseen optimaalisesti.

3.2 Vertailumalli

Tässä tutkielmassa käytetään vertailumallina autoregressiivistä liikkuvan keskiarvon (*autoregressive moving average*, ARMA) mallia

$$X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \dots + \theta_q Z_{t-q} \quad (14)$$

jossa $\{X_t\}$ on stationaarinen aikasarja (inflaatio tässä tutkielmassa), $\{Z_t\}$ on valkoista kohinaa keskiarvolla 0 ja vakiovarianssilla, ja polynomeilla $(1 - \phi_1 z - \dots - \phi_p z^p)$ ja $(1 + \theta_1 z + \dots + \theta_q z^q)$ ei ole yhteisiä tekijöitä (Brockwell & Davis 2002).

ARMA-malli on lineaarinen malli, ja se on erittäin hyödyllinen vertailumalliksi neuroverkkomalleille, koska se ei ole liian yksinkertainen (Álvarez-Díaz & Gupta 2016). Toisaalta ARMA-malli ei ole myöskään liian monimutkainen. Jos neuroverkkomalli pystyy ennustamaan inflaatiota paremmin kuin ARMA-malli, voidaan sanoa, että neuroverkkomallit ovat hyödyllisiä euroalueen inflaation ennustamisessa.

Koska ARMA-malli ja neuroverkkomalli eroavat selkeästi toisistaan, molempien mallien valinnat suoritetaan aiemmassa kirjallisuudessa näille malleille vakiintuneilla tavoilla, jotta tutkimuksen tulokset eivät vinoutuisi suosimaan jompaa kumpaa mallia. ARMA-mallit ovat sen verran yleisiä aikasarjaennustamisessa, että niille on muodostunut vakiintuneita tapoja mallinvalintaan.

Ehkä vakiintunein tapa rakentaa ARMA-malli on Box-Jenkins-metodologia. Box-Jenkins-metodologiassa ensimmäisenä askeleena on aikasarjan tutkiminen mahdollisten kausivaihteluiden ja muiden epästationaarisuutta aiheuttavien tekijöiden löytämiseksi, ja näiden tekijöiden poistaminen aikasarjasta. Ensimmäinen askel on erittäin tärkeä, koska ARMA-malli on tarkoitettu stationaarisille aikasarjoille (Box ym. 2015).

Seuraavaksi tarkoitus on löytää mallin viiveiden lukumäärä. Tarkoituksena on löytää sopivat viiveet aikasarjan aikaisemmille arvoille sekä mallin aikaisemmille virhetermeille. Viiveiden valintaan voidaan käyttää monia tapoja, mutta yleisimmät keinot ovat katsoa aikasarjan autokorrelaatio- ja osittaisautokorrelaatiofunktioita sekä vertailla eri mallien informaatiokriteerejä.

Autokorrelaatio- ja osittaisautokorrelaatioiden teoreettisia ominaisuuksia voidaan

käyttää viiveparametrien valinnassa. AR(p)-mallilla osittaisautokorrelaatiofunktio katkeaa viiveen p jälkeen, mutta autokorrelaatio pienenee pikkuhiljaa. Samalla tavalla MA(q)-prosessilla autokorrelaatio tipahtaa nolnaan viiveen q jälkeen, mutta osittaisautokorrelaatio häviää vähitellen. Jos aikasarjan tuottanut prosessi sisältää sekä autoregressiivisen että liikkuvan keskiarvon komponentit, autokorrelaatio- ja osittaisautokorrelaatiofunktiot hiipuvat vähitellen nolnaan (Box ym. 2015).

Yleisimmät ARMA-mallien valinnassa käytetyt informaatiokriteerit ovat Akaiken (AIC) ja Bayesin (BIC) informaatiokriteerit.

$$\text{AIC} = \ln(\hat{\sigma}^2) + r \frac{2}{n} \quad (15)$$

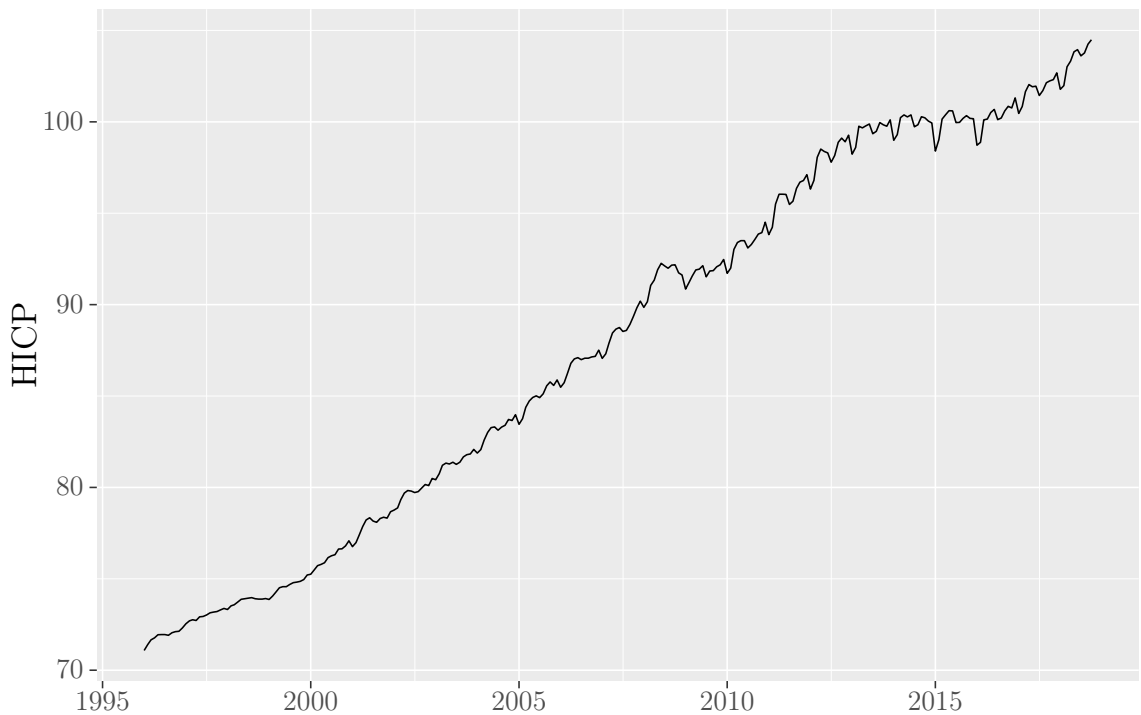
$$\text{BIC} = \ln(\hat{\sigma}^2) + r \frac{\ln(n)}{n} \quad (16)$$

Informaatiokriteeri siis koostuu mallin estimoidusta varianssista σ^2 sekä niin kutsutusta rangaistustermistä, joka koostuu mallin havaintojen määrästä n ja mallin viiveiden lukumäärästä r . Informaatiokriteeri on sitä pienempi mitä pienempi mallin varianssi on. Toisaalta uusien viiveiden lisääminen pienentää aina mallin varianssia, joten viiveiden lisääminen kasvattaa rangaistustermiä ja näin ollen myös koko informaatiokriteerin arvoa. Akaiken ja Bayesin informaatiokriteerien laskukaavoista nähdään, että Bayesin informaatiokriteeri rankaisee viiveiden lisäämisestä malliin enemmän kuin Akaiken informaatiokriteeri (Box ym. 2015).

Mallin viiveiden valinnan jälkeen valittu ARIMA-malli estimoidaan suurimman uskottavuuden menetelmällä, ja valitulle mallille tehdään diagnostiset tarkastelut. Mallin oletuksena on, että virhetermit noudattavat valkoinen kohina -prosessia. Näin ollen ARIMA-mallin diagnostisissa tarkasteluissa pitää varmistaa, että valitun mallin residuaalien välillä ei ole enää autokorrelaatiota tai muitakaan riippuvuuksia. Jos valitun mallin oletukset eivät päde, joudutaan mallin viiveet valitsemaan uudestaan jollain toisella tavalla ja tarkistamaan, että mallin oletukset pätevät tässä uudessa mallissa (Box ym. 2015). Kun ARIMA-malli on valittu ja se on todettu dataan sopivaksi, voidaan valitun mallin avulla tehdä ennusteita halutuille ajanjaksoille samoin kuin neuroverkkomallilla.

3.3 Tutkimusaineisto

Kuluttajahintaindeksi on ehdottomasti suosituin inflaation mittari, ja sitä on käytetty kaikissa vastaavissa tutkimuksissa. Kaastra ja Boyd (1996) suosittelevat talouden makromuuttujien ennustamisessa neuroverkkomallilla käytettäväksi kuukausittaista tai kvartaalista dataa. Tässä tutkielmassa aineistona käytetään euroalueen kuukausittaista harmonisoitua kuluttajahintaindeksiä, joka on julkisesti saatavilla Eurostatin aineistotietokannasta aikavälille tammikuu 1996 - lokakuu 2018 (Eurostat 2018). Euroopan keskuspankin päätökset perustuvat harmonisoituun kuluttajahintaindeksiin, joten se toimii luonnollisesti tämän tutkielman inflaation mittarina.



Kuva 1: Euroalueen harmonisoitu kuluttajahintaindeksi

Tutkimuksessa pyritään ennustamaan inflaatiota hyödyntämällä ainoastaan sen aikaisempia arvoja. Joissain tutkimuksissa on käytetty myös muita muuttujia inflaation ennustamiseen, mutta useiden muuttujien käyttö voi aiheuttaa myös ongelmia. Ensimmäinen ongelma on muuttujien valinta. Taloustieteen teoria voi antaa suosituksia mahdollisista ennustavista muuttujista, mutta Stock ja Watson (1999) ovat huomanneet,

että inflaatiota voi ennustaa melkein 200:lla eri muuttujalla, joten sopivien ennustavien muuttujien valinta näiden joukosta on hankalaa. Monimuuttujamenetelmät eivät ole kuitenkaan kyenneet ennustamaan inflaatiota erityisen hyvin verrattuna yksinkertaisiin inflaation aiempia arvoja käyttäviin malleihin (Faust & Wright 2013).

Toinen ongelma on neuroverkkomallin vertaaminen muihin malleihin. Jos eri malleja halutaan vertailla, pitäisi niissä käyttää samaa informaatiota, jotta vertailu olisi luotettavaa. Silloin muuttujat pitäisi valita jonkin tietyn mallin mukaan optimaalisesti, mikä aiheuttaa tuloksissa harhaa sen mallin hyväksi, jonka mukaan muuttujat on valittu (Binner ym. 2005).

Lisäksi ennustettaessa inflaatiota muiden muuttujien avulla, inflaation ja ennustavien muuttujien suhteen tulisi pysyä samana yli ajan (He ym. 2012). Tämä ei ole kuitenkaan uskottava oletus, joten inflaation aiemmissä arvoissa pysyminen on turvallisempi vaihtoehto.

Kuvassa 1 on euroalueen harmonisoitu kuluttajahintaindeksi tammikuusta 1996 lokakuuhun 2018 vuoden 2015 hinnoissa. Kuviosta nähdään selvästi, että kuluttajahintaindeksiaikasarja ei ole stationaarinen, vaan aikasarjalla on selkeä kasvava trendi ja siinä on kausivaihtelua.

Kuluttajahintaindeksiaikasarjan epästationaarisuuden voi todentaa myös tilastollisilla testeillä. Dickey-Fuller-testisuure (Dickey & Fuller 1979), joka testaa nollahypoteesia siitä, että aikasarjalla on yksikköjuuri, ei ole tilastollisesti merkitsevä 10 viiveellä.

Trendin ja yksikköjuuren vuoksi harmonisoidun kuluttajahintaindeksin käyttö suoraan ei ole mahdollista, koska ARMA-malli vaatii aikasarjan olevan stationaarinen. Samoin Zhang (2012) ja Remus ja O'Connor (2001) suosittelevat lineaaristen trendien siivoamista aineistosta ennen neuroverkkomallilla ennustamista, vaikka neuroverkkomallien pitäisi pystyä teoriassa mallintamaan myös lineaarisia trendejä ja muita epästationaarisuuksia aiheuttavia tekijöitä hyvin.

Taulukko 1: Dickey-Fuller-testi kuluttajahintaindeksiaikasarjalle

Viiveet	Testisuureen arvo	p-arvo
10	-0.6395	0.8186

Aikasarja-aineisto täytyy siis esikäsitellä käytettävien mallien mukaan sopivaksi. Lisäksi tässä tutkielmassa on tarkoitus ennustaa inflaatiota, ei kuluttajahintaindeksiä. Tavallinen tapa tuottaa inflaatioaikasarja kuluttajahintaindeksistä on ottaa kuluttajahintaindeksin arvoista luonnollinen logaritmi ja ensimmäinen differenssi, jotta saadaan kuukausittaista inflaatiota kuvaava muuttuja:

$$\pi_t = \ln(HICP_t) - \ln(HICP_{t-1}) \quad (17)$$

Ensimmäinen differenssi poistaa aineistosta lineaarisen trendin ja yksikköjuuren, jotka aiheuttaisivat aikasarjan epästationaarisuuden (Kaastra & Boyd 1996).

Inflaatioaikasarjassa on kuitenkin kausivaihtelua, joka voi tuottaa ongelmia ainakin ARMA-mallin käytössä. Teoriassa kausivaihtelun ei tulisi vaikuttaa neuroverkkomallin ennustekykyyneen, mutta Zhang ja Qi (2005) ovat todistaneet, että aikasarjan kausivaihtelun siivoaminen aineistosta parantaa ennustetuloksia. Siksi tässäkin tutkielmassa inflaatioaikasarjasta poistetaan myös kausivaihtelukomponentti.

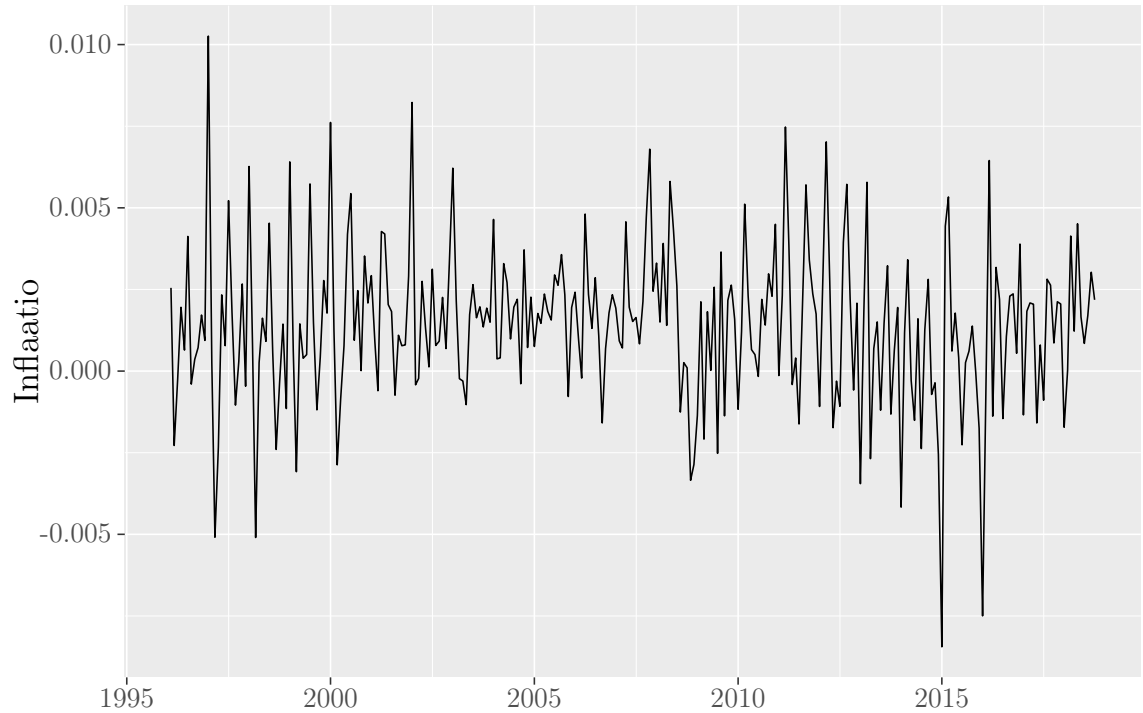
Koska kausivaihtelun ei teoriassa tulisi olla neuroverkkomallille ongelma, halutaan tässä tutkielmassa tutkia myös mallien ennusteita kausitasoittamattomalle inflaatioaikasarjalle. Molemmat mallit estimoidaan siis sekä kausitasoitettulle että kausitasoittamattomalle aineistolle, ja sama analyysi tuotetaan molemmissa tapauksissa tutkimusten tulosten robustisuuden parantamiseksi. Kausivaihtelun sisältävän aikasarjan analyysi on tämän tutkielman liitteenä.

Kuvassa 2 piirretty kausivaihtelusta siivottu inflaatioaikasarja vaikuttaa myös stationaariselta, joten tälle aikasarjalle voidaan suorittaa samanlaiset testit kuin alkuperäiselle kuluttajahintaindeksin aikasarjalle. Dickey-Fuller-testin p-arvo on pienempi kuin 0.01, joten voidaan päätellä, että inflaatioaikasarja ei ole yksikköjuurta.

Taulukko 2: Dickey-Fuller-testi inflaatioaikasarjalle

Viiveet	Testisuureen arvo	p-arvo
10	-5.7984	<0.01

Koska harmonisoidun kuluttajahintaindeksin arvot eivät muutu periodista toiseen kovin paljon, differenssi saattaa aineiston välille $[-1, 1]$. Monet neuroverkkomalleissa



Kuva 2: Euroalueen kuukausittainen inflaatio

käytettävät aktivointifunktiot vaativat sisääntulevan aineiston olevan tietyllä välillä, joten perinteiset muunnokset ovat tarpeellisia myös neuroverkkojen käytön kannalta.

Taulukko 3: Euroalueen harmonisoitu kuluttajahintaindeksi ja inflaatio

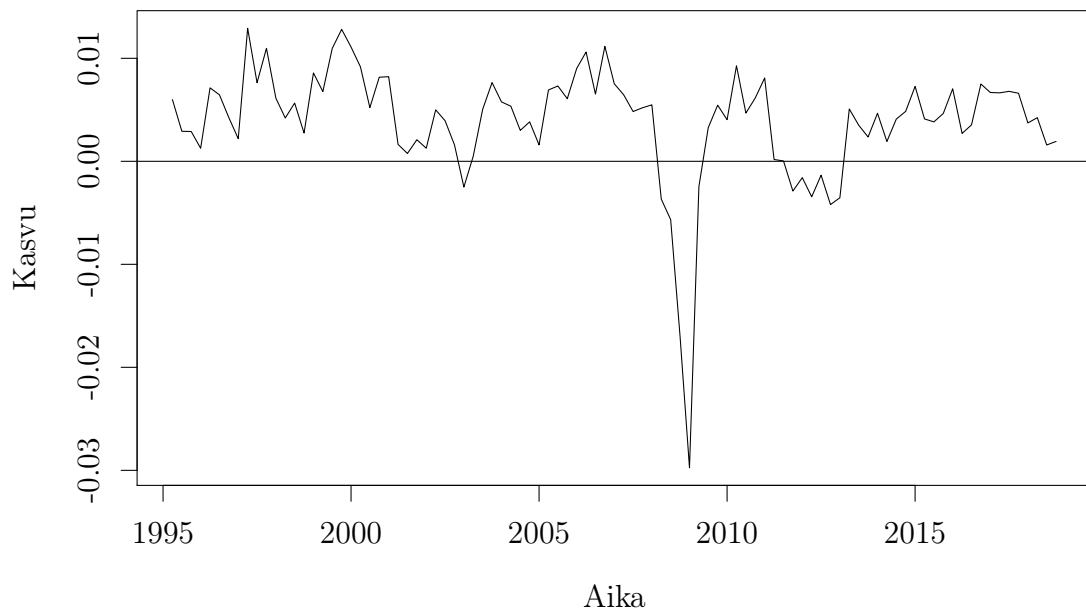
Aikasarja	Keskiarvo	Varianssi	Keskihajonta	Minimi	Maksimi
HICP	88.04	107.39	10.36	71.08	104.49
Inflaatio	0.00139	0.0000059	0.0024	-0.0084	0.010

3.4 Ennustamisstrategia

Tässä tutkielmassa halutaan simuloida taloudellisen taantuman aikaista ennustamista. Tässä tutkielmassa taantuma määritellään aikana, jolloin kahden peräkkäisen vuosineljänneksen bruttokansantuotteen kasvu on ollut negatiivista. Kuvassa 3 on euroalueen bruttokansantuotteen kasvu vuosina 1995-2018.

Euroalueella on ollut taantumia vuoden 2008 toisesta neljänneksestä vuoden 2009

Euroalueen bruttokansantuotteen kasvu



Kuva 3: Euroalueen bruttokansantuotteen kasvu

kolmanteen neljännekseen asti sekä vuoden 2011 neljänneltä neljännekseltä vuoden 2013 toiselle neljännekselle asti. Koska vuosien 2008-2009 taantuma on kuvionkin perusteella syvempi, käytetään tässä tutkielmassa tätä niin sanotun suuren taantuman aikaa simuloimaan inflaation ennustamista taantuman aikana.

Tässä tutkielmassa halutaan vertailla neuroverkkomallin tuottamia ennusteita eri suhdannetilanteissa, joten ennusteet tehdään valitulla mallilla myös noususuhdanteen aikana. Suuren taantuman jälkeinen nousukausi (vuoden 2009 kolmannesta neljänneksestä vuoden 2010 kolmannen neljänneksen loppuun) toimii hyvin vertailukohtana, koska ajanjaksot ovat lähellä toisiaan. Näin ollen samat valitut mallit pystyvät ennustamaan teoriassa arvoja myös nousukaudella hyvin eivätkä esimerkiksi rahapolitiikan muutokset ole ehtineet vaikuttaa inflaatioon merkittävästi.

Tässä tutkielmassa halutaan simuloida oikeaa tilannetta, jossa halutaan ennustaa inflaation arvoa kuukauden, kolmen kuukauden ja vuoden päähän, joten tutkielmassa käytetään rekursiivista ennustestrategiaa. Ennusteiden tekemiseen käytetään siis vain

informaatiota, joka olisi todellisuudessaakin ennustetilanteessa käytettävissä. Uusien ennusteiden tekemistä varten malliin otetaan uusia havaintoja, ja malli estimoidaan uudelleen. Tätä jatketaan niin kauan kunnes havaintoja riittää aineistossa. Mallien valinnat suoritetaan ennen taantumaa muodostuneella aineistolla. Esimerkiksi kun ennustetaan periodin 2008:12 arvoa vuoden päähän, käytetään aikasarjan havaintoja 1996:2-2007:12 valitun mallin parametrien estimointiin.

Tässä tutkielmassa käytetään dynaamista ennustamista staattisen ennustamisen sijaan. Pidemmille ajanjaksoille tehtävissä ennusteissa käytetään siis aiempien periodien ennusteita viiveinä. Dynaaminen ennustaminen on ehdottomasti vastaavassa kirjallisuudessa suositumpi metodi (mm. Binner ym. 2006, Gupta & Kashyap 2015, He ym. 2012), joten sitä käytetään myös tässä tutkielmassa.

Jotta ennustestrategia tulisi konkreettisemmaksi, esitetään ennustamisstrategia aikasarjan havaintoja käyttämällä. Aikasarjan havainnot 1996:2-2008:3 käytetään neuroverkkomallin valintaa koulutus- ja validaatio-otosten avulla. Tämän jälkeen otosikkunaa kasvatetaan yhden havainnon verran eteenpäin, eli havaintoja 1996:2-2008:4 käytetään valitun mallin mukaisen neuroverkon kouluttamiseen ja estimoitua mallia käytetään ennustamaan arvoja 2008:5, 2008:7 ja 2009:4. Näin käydään koko aikasarja läpi taantumien loppuun asti ja tutkitaan kuinka suuria ennustevirheitä neuroverkkomalli tekee. Samalla mallilla tehdään ennusteet myös taantumien jälkeiselle ajalle, ja tuloksia vertaillaan sekä vertailumallin vastaaviin tuloksiin että toisiinsa.

3.5 Arviointikriteerit

Ennusteiden arvioinnissa voidaan käyttää monenlaisia mittareita, jotka eroavat toisistaan ja jotka voivat vaikuttaa myös parhaimman mallin valintaan. Tämän takia monissa tutkimuksissa on käytetty useampaa eri arviointikriteeriä, jolla eri malleja on vertailtu.

Tavallisimmat arviointikriteerit ennusteille ovat absoluuttinen keskivirhe (*mean absolute error*, MAE), keskineliövirhe (*mean squared error*, MSE), keskineliövirheen neliöjuuri (*root of mean squared error*, RMSE), absoluuttinen keskiprosenttivirhe (*mean*

absolute percentage error, MAPE) ja neliövirheiden summa (*sum of errors*, SE).

$$MAE = \frac{\sum |e_t|}{N} \quad (18)$$

$$MSE = \frac{\sum (e_t)^2}{N} \quad (19)$$

$$RMSE = \sqrt{\frac{\sum (e_t)^2}{N}} \quad (20)$$

$$MAPE = 100 \frac{1}{N} \sum \left| \frac{e_t}{y_t} \right| \quad (21)$$

$$SE = \sum (e_t)^2 \quad (22)$$

Kriteereissä e_t on mallin ennustevirhe periodilla t , N on otoksen suuruus ja y_t on inflaation toteutunut arvo periodilla t .

Álvarez-Díaz ja Gupta (2016) käyttivät absoluuttista keskiprosenttivrhetä, Nakamura (2005) käytti keskineliövirhetä ja Binner ym. (2005) käyttivät aiemmin mainittujen lisäksi keskineliövirheen neliöjuurta neuroverkkotutkimuksissaan. Gupta ja Kashyap (2012) käyttivät tutkimuksessaan absoluuttista keskivirhetä, keskineliövirheen neliöjuurta sekä neliövirheiden summaa.

Keskineliövirhe ja sen neliöjuuri ovat suosittuja arviointikriteerejä ennusteille niiden teoreettisten ominaisuuksien vuoksi. Ne sopivat hyvin vertailemaan malleja, jotka on estimoitu käyttämällä samaa dataa. Keskineliövirheen neliöjuuri on parempi vaihtoehto kuin pelkkä keskineliövirhe, jos ennustevirheen arvio halutaan samassa mittayksikössä kuin itse aineisto (Hyndman & Koehler 2006).

Poikkeavat havainnot vaikuttavat kuitenkin keskineliövirheeseen voimakkaasti ja keskineliövirhe ei välttämättä ole luotettava, jos ennustemalli tekee yksittäisiä voimakkaasti positiivisia tai negatiivisia virheitä. Tämän takia Armstrong (2001) suosittelee käytettäväksi mieluummin absoluuttista keskiprosenttivrhetä, joka on myös aiemmassa vastaavassa kirjallisuudessa ollut suosittu.

Absoluuttinen keskiprosenttivirhe edustaa prosentuaalisiin virheisiin perustuvia arviointikriteerejä. Sen huonoja puolia ovat sen tuottamat suuret virhearvot, jos aikasarjan arvot ovat lähellä nollaa. Tällöin esimerkiksi äärettömän suuret arvot ovat mahdollisia. Sillä on myös tapana rangaista positiivisista virheistä kovemmin kuin negatiivisista, joten se ei ole erityisen käyttökelpoinen inflaatiomallien arvioinnissa (Hyndman & Koehler 2006). Se kuitenkin antaa kuvaa virheiden suhteellisesta suuruudesta verrattuna inflaation omiin arvoihin. Pienetkin erot todellisiin arvoihin voivat olla suuria, jos itse aikasarja koostuu vain pienistä arvoista.

Tässä tutkielmassa käytetään ensisijaisesti keskineliövirheen neliöjuurta ennustemallien vertailuun. Sekä neuroverkkomalli että ARMA-malli estimoidaan samalla aineistolla, joten keskineliövirheeseen perustuva kriteeri on perusteltu. Samassa mittayksikössä olevat virheet helpottavat niiden tulkintaa. Aiemmasta kirjallisuudesta voidaan päätellä, että molemmat ennustemallit eivät tee erityisen isoja ennustevirheitä, joten poikkeavan suurista virheistä ei todennäköisesti tarvitse huolestua. Inflaatioaikasarjassa on myös nollahavaintoja, joten absoluuttinen keskiprosenttivirhe tulisi olemaan ääretön tietyillä aikaväleillä, mikä estää eri mallien vertailun kokonaan.

Näiden arviointikriteerien käytöllä on kuitenkin rajoituksensa eikä useammankaan mittarin käyttö kerro vielä koko totuutta ennusteista ja niiden paremmuusjärjestyksestä, jos erot ovat pieniä. Esimerkiksi Binner ym. (2005) muodostivat regression, jossa todellista inflaatiota selitetään neuroverkkomallin ennusteella ja vertailumallin ennusteella. Regression kertoimien t-testien tilastollinen merkitsevyys kertoo kuinka paljon ennusteet sisältävät samaa informaatiota, ja kuinka tärkeitä mallit ovat inflaation ennustamisessa. Samoin Álvarez-Díaz ja Gupta (2016) käyttivät Diebold-Mariano-testiä, joka testaa nollahypoteesia siitä, että kahden mallin ennusteiden tarkkuus on sama.

Myös tässä tutkielmassa käytetään muokattua Diebold-Mariano-testiä mallien tuottamien ennustevirheiden arviointiin. Dieboldin ja Marianon (1995) alkuperäisessä testissä tutkitaan kahden ennustemallin ennustevirheiden keskiarvoja

$$\bar{d} = \frac{1}{n} \sum_{t=1}^n d_t \quad (23)$$

jossa d_t on keskineliövirheellä mitattujen ennustevirheiden erotus. Diebold ja Mariano

ovat osoittaneet, että keskiarvon asymptoottinen varianssi on $n^{-1}(\gamma_0 + 2 \sum_{k=1}^{h-1} \gamma_k)$, jossa γ_k on ennustevirheiden erotusten k . autokovarianssi. Diebold-Mariano-testin testisuure $h-1$ periodia eteenpäin tehtäville ennusteille on

$$S_1 = [n^{-1}(\hat{\gamma}_0 + 2 \sum_{k=1}^{h-1} \hat{\gamma}_k)]^{-0.5} \bar{d} \quad (24)$$

Diebold ja Mariano (1995) huomasivat myös, että nollahypoteesin ollessa voimassa, testisuure noudattaa asymptoottisesti standardoitua normaalijakaumaa. Harvey ym. (1997) ovat kuitenkin muokanneet alkuperäistä testisuureta sopimaan myös pidemmille ennusteväleille. Heidän testisuurensa on

$$S_1^* = \left[\frac{n + 1 - 2h + n^{-1}h(h - 1)}{n} \right] S_t \quad (25)$$

jossa S_t on alkuperäinen Dieboldin ja Marianon testisuure. Tämä testisuure sopii myös harhaisille ennusteille ja se mahdollistaa muidenkin kuin keskineliövirheen käytön tapiofunktiossa.

Muunnettua Diebold-Mariano-testiä voidaan käyttää neuroverkon tuottamien ennusteiden vertailuun taantumien ja noususuhdanteen aikana. Erojen neuroverkko- ja ARMA-mallin välillä tulisi olla samanlaiset taantumassa ja ei-taantumassa, jos mallit toimivat yhtä hyvin sekä taantumassa että noususuhdanteessa. Jos inflaation tuottama dataa generoiva prosessi on kuitenkin taantumassa epälineaarinen, neuroverkkomallin pitäisi teoriassa olla tilastollisesti merkitsevästi parempi kuin ARMA-malli tällä ajanjaksolla.

4 Analyysi

4.1 Neuroverkkomalli ja ennusteet

4.1.1 Mallin valinta

Tässä osiossa rakennetaan neuroverkkomalli inflaation aiemmista arvoista osiossa 3.1 kuvatulla tavalla käyttäen luvussa 3.3 kuvattua aineistoa. Inflaatioennusteet tehdään aikavälille 2008:4-2009:6, joten neuroverkko rakennetaan aikavälin 1996:3-2008:3 aineistoa käyttäen. Tämä aineisto jaetaan edelleen koulutus- ja validaatio-otoksiin siten, että koulutusotos sisältää 80 % ja validaatio-otos 20 % neuroverkon rakentamiseen tarkoitusta aineistosta. Tutkielman inflaatioaineistossa se tarkoittaa sitä, että koulutusotos sisältää aineiston 116 ensimmäistä havaintoa, ja mallin validaatio suoritetaan 29 havainnolla.

Mallinvalintavaiheessa estimoidaan joukko neuroverkkoja, joissa on 1-20 viivettä sisäänottotasolla ja 1-20 kätkeytyä neuronია kätkeytyllä tasolla siten, että kätkeytyjen neuroneiden määrä on maksimissaan yhtä suuri kuin sisäänottotason neuroneiden. Koska inflaatioaikasarjasta otettu koulutusotos ei ole kovin pitkä, malliin ei voida sisällyttää kovin monta viivettä ja kätkeytyä neuronია. 20 on oletettavasti riittävän suuri luku molemmille, jotta neuroverkkomalli pystyy approksimoimaan inflaatioaikasarjan taustalla olevaa funktiota.

Neuroverkkojen estimointi suoritetaan tekemällä kymmenen satunnaista alkuarvausta mallin parametreille ja minimoimalla mallin residuaalien neliösumma käyttämällä kimmoisa vastavirta -algoritmia. Algoritmiksi valittiin kimmoisa vastavirta -algoritmi, koska se on kehitetty suosituimmasta estimointialgoritmista, mutta se on todettu nopeammaksi ja tarkemmaksi kuin perinteinen vastavirta-algoritmi.

Neuroverkkomallin estimointi tapahtuu R-ohjelmalla nnfor-paketin avulla (Kourentzes 2019), joka on kirjoitettu aikasarjojen neuroverkkotutkimusta varten. Käytetty paketti perustuu R:n toiseen neuroverkkopakettiin neuralnetiin (Fritsch ym. 2019), jonka avulla neuroverkko voidaan rakentaa helposti annetuilla aktivointifunktioilla ja estimointialgoritmeilla.

Neuroverkkomallin valinta tapahtuu koulutusotoksella, mutta neuroverkkojoukosta valitaan malli, jonka keskineliövirhe on pienin validaatio-otoksessa. Keskineliövirheen käyttäminen mallinvalinnassa on luonnollinen ratkaisu, koska lopulta valitun neuroverkon tuottamia ennusteita arvioidaan keskineliövirheen avulla, ja esimerkiksi ARMA-mallin valinnassa hyödylliset informaatiokriteerit eivät ole järkeviä kriteerejä neuroverkkomallien valinnassa (Qi & Zhang 2001).

Taulukkoon 4 on kerätty eri viiveillä parhaiden mallien keskineliövirheet. Taulukosta nähdään, että neuroverkko, jossa on 17 viivettä ja 2 kätkeytyä neuronია, tuottaa validaatio-otoksessa pienimmän keskineliövirheen. Koulutusotoksen keskineliövirhe tälle mallille on 8.92×10^{-7} . Odotetusti validaatio-otoksen keskineliövirhe on huomattavasti suurempi, noin 0.018. Tulos osoittaa, että neuroverkko sopii erittäin hyvin koulutusotokseen, mutta se voi kärsiä ylisovittamisesta.

Taulukosta 4 nähdään myös se, että suurimmassa osassa viiveitä kätkeytyjen neuroneiden määrä on melko pieni. 20 tapauksesta validaatio-otoksen perusteella kuudessa-toista pitäisi valita alle kymmenen kätkeytyä neuronია, ja enimmillään kätkeytyjen neuroneiden määrä olisi 16. Voidaan siis päätellä, että yli 20 kätkeytyä neuronია ei tuottaisi parempia neuroverkkomalleja.

Taulukko 4: Validaatio-otoksen perusteella valitut mallit eri viiveillä

Viiveiden lkm	Kätkettyjen neuronien lkm	MSE validaatio-otoksessa
1	1	0.04078577
2	2	0.03867664
3	3	0.03633576
4	2	0.03923058
5	3	0.03627556
6	3	0.04110668
7	6	0.04908847
8	8	0.03309962
9	5	0.03937659
10	6	0.03714591
11	2	0.03324406
12	2	0.02444552
13	1	0.02567352
14	1	0.02876397
15	15	0.02687908
16	3	0.02281761
17	2	0.01801458
18	13	0.02323320
19	16	0.03377529
20	11	0.03644782

Estimoidun mallin painokertoimia ei esitellä tässä, koska ne eivät sisällä tutkimuksen kannalta erityistä informaatiota eikä niiden avulla ole mahdollista tehdä tulkintoja neuroverkon kytköksistä. Toisin kuin ARMA-mallin valinassa, neuroverkkomallille ei ole kehitetty diagnostisia testejä, joilla mallin oikeellisuutta voitaisiin testata. Validaatio-otoksen virheet ovat siis ainoa keino valitun mallin arvioimiseen. Validointiotoksen keskineliövirheet ovat sen verran pieniä, että malli selkeästi sopii ennennäkemättömään aineistoon melko hyvin. Mallia voidaan siis käyttää euroalueen inflaation ennusteiden

tekemiseen.

4.1.2 Ennusteet

Valittua neuroverkkoa käytetään tässä osiossa taantuman aikana ja sen jälkeen tehtäviin ennusteisiin. Taulukossa 5 ovat tiedot neuroverkkomallin tekemistä ennustevirheistä taantuman ajalle yhden, kolmen ja kahdentoista kuukauden päähän dynaamisesti ennustettuna. Taulukkoon on laskettu myös ennustevirheiden keskineliövirhe sekä keskineliövirheen neliöjuuri, joita käytetään mallien ennusteiden vertailuun.

Taulukosta nähdään, että neuroverkkomallin tuottamat ennustevirheet ovat keskimäärin lähellä nollaa. Absoluuttisten ennustevirheiden keskiarvo kuitenkin pienenee huomattavasti 12 kuukautta eteenpäin tehtävissä ennusteissa, joten voidaan päätellä, että neuroverkkomalli ennustaa inflaatiota paremmin pidemmille aikaväleille taantuman aikana. Ennustevirheiden varianssi ei muutu merkittävästi eri ennusteväleillä taantuman aikana.

Taulukko 5: Neuroverkkomallin ennustevirheet taantuman aikana

	1 kuukausi	3 kuukautta	12 kuukautta
Keskiarvo	-0.001179886	-0.002144159	0.0004461112
Varianssi	1.2672e-05	1.324646e-05	1.248617e-05
Keskihajonta	0.003559775	0.003639569	0.003533577
MSE	1.321933e-05	1.696078e-05	1.185277e-05
RMSE	0.003635839	0.004118347	0.003442785

Taulukossa 6 on vastaavasti tietoa valitun neuroverkkomallin ennusteista ajanjaksolle 2009:7-2010:9. Ennustevirheet ovat taulukon perusteella hyvin samansuuntaisia kuin taantuman aikaisissa ennusteissa. Taulukoiden perusteella neuroverkkomalli tekee kuitenkin pienempiä virheitä noususuhdanteen kuin taantuman aikana keskineliövirheen neliöjuurella mitattuna kaikilla ennusteväleillä. Pienet erot ennustevirheissä antavat olettaa, että talouden suhdannetilanteella ei ole merkitystä neuroverkkomallin ennusteisiin.

Kuvassa 4 on vielä havainnollistettu neuroverkkomallin tekemiä ennusteita verrattuna toteutuneisiin inflaation arvoihin taantuman aikana. Kuvasta nähdään, että erot eivät ole kovin suuria, ja neuroverkkomallin ennusteet taantuman aikana seuraavat hyvin inflaation nousuja ja laskuja.

Tutkielman liitteessä on tehty vastaava analyysi kausitasoittamattomalle inflaatioaikasarjalle. Kausitasoittamattomalle aikasarjalle valittu neuroverkkomalli ei eroa merkittävästi tämän osion ennustemallista, joten myös mallien tekemät ennusteet euroalueen inflaatiolle ovat hyvin samansuuntaisia. Neuroverkko tekee kuitenkin parempia ennusteita taantuman jälkeisenä aikana, jos aikasarja on kausitasoitettu ennen ennusteiden tekemistä.

Taulukko 6: Neuroverkkomallin ennustevirheet noususuhdanteen aikana

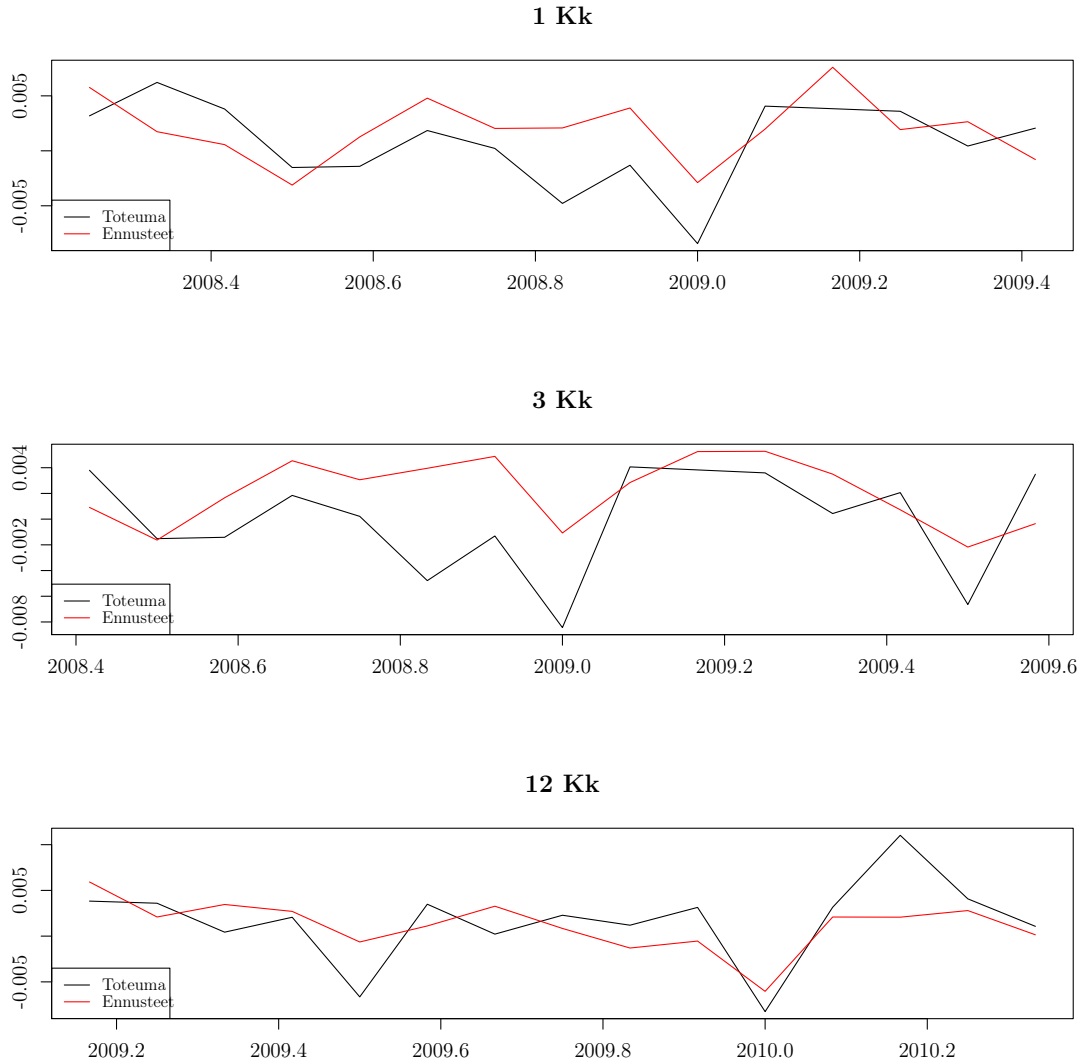
	1 kuukausi	3 kuukautta	12 kuukautta
Keskiarvo	-0.0001910604	0.0006573873	-4.292852e-05
Varianssi	9.624883e-06	4.783652e-06	5.924994e-06
Keskihajonta	0.0031024	0.002187156	0.002434131
MSE	9.019728e-06	4.8969e-06	5.531837e-06
RMSE	0.003003286	0.002212894	0.002351986

4.2 ARMA-malli ja ennusteet

4.2.1 Mallin valinta

Tässä osiossa muodostetaan ennusteet euroalueen inflaatiolle sekä taantuman että kasvukauden aikana käyttämällä ARMA-mallia, joka esiteltiin osiossa 3.2 käyttämällä aineistoa osiosta 3.3. ARMA-mallin rakentamisessa käytetään Box-Jenkins-metodia, josta on muodostunut erittäin yleinen tapa ARMA-mallien valinnalle.

Box-Jenkins-metodin ensimmäisenä tavoitteena on mallin identifioiminen, joka tarkoittaa mallin stationaarisuuden varmistamista ja mallin viiveparametrien p ja q valintaa (Box ym. 2015). Osiossa 3.3 muodostettiin stationaarinen aikasarja euroalueen kuukausittaiselle inflaatiolle ottamalla harmonisoidun kuluttajahintaindeksisarjasta lo-



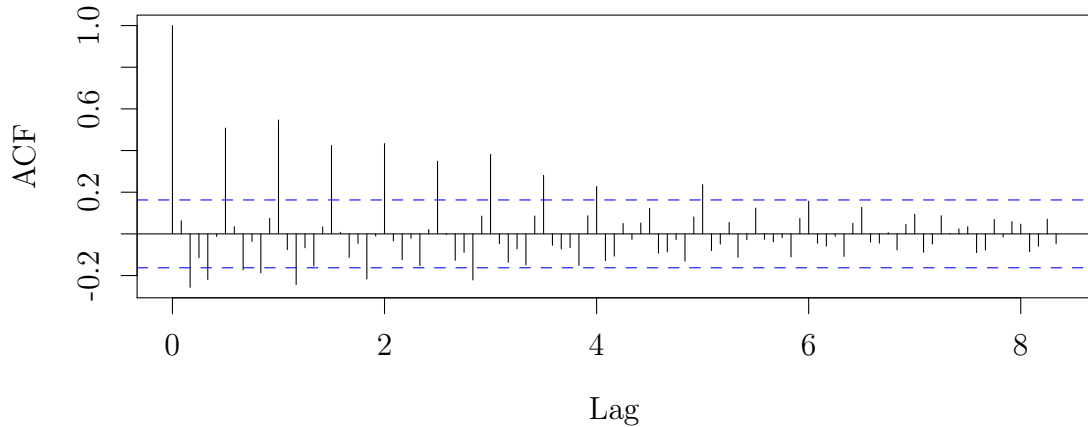
Kuva 4: Neuroverkkomallin ennusteet ja todelliset arvot taantuman aikana

garitmi ja ensimmäinen differenssi sekä poistamalla aikasarjasta kausivaihtelu. Tämä muunnos varmisti, että inflaatioaikasarja todella on stationaarinen kuten osiossa 3.3 nähtiin.

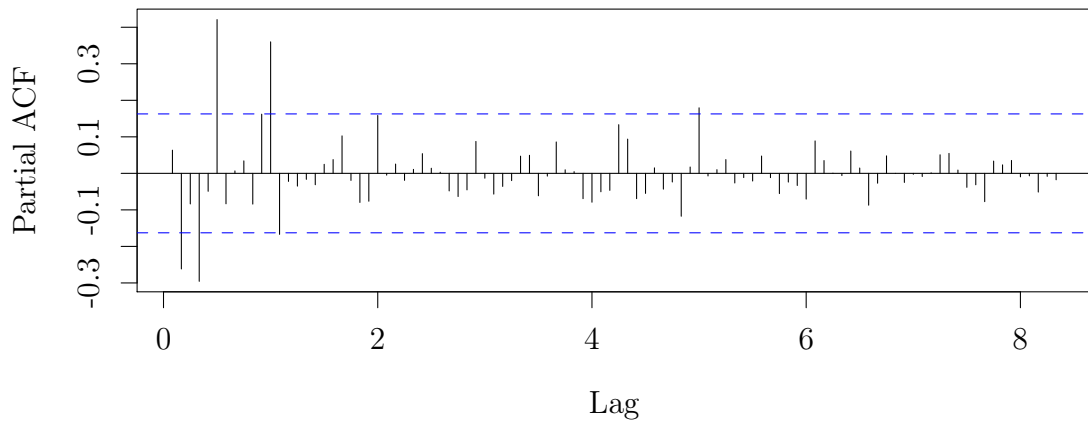
Mallin viiveparametrien valinnassa voidaan käyttää inflaatioaikasarjan autokorrelaatio- ja osittaisautokorrelaatiofunktioita. Kuvassa 4 on inflaatioaikasarjan autokorrelaatio- ja osittaisautokorrelaatiofunktioita. Aikasarjan autokorrelaatiofunktio näyttäisi laskevan vähitellen nollaan, joten vain liikkuvan keskiarvon komponentin omaava malli ei ole realistinen valinta. Inflaatioaikasarjan osittaisautokorrelaatio ei anna yhtä selkeää ar-

vausta viiveille. Useat osittaisautokorrelaation viiveet ovat tilastollisesti merkitseviä, mutta funktio ei tipahda nollaan yllättäen tai häviää vähitellen.

Autokorrelaatio



Osittaisautokorrelaatio



Kuva 5: Inflaatioaikasarjan autokorrelaatio- ja osittaisautokorrelaatio

Koska autokorrelaatio- ja osittaisautokorrelaatiofunktiot eivät antaneet selkeää vastausta viiveparametrien valinnalle, voidaan käyttää informaatiokriteerejä mahdollisen mallin valinnassa. Tässä mallinvalintametodissa estimoidaan joukko ARMA-malleja, joille lasketaan informaatiokriteerin arvo. Mallijoukosta valitaan malli, jonka informaatiokriteeri on pienin. Tässä tutkielmassa käytetään perinteisiä Akaiken ja Bayesin informaatiokriteerejä.

Valinta tehdään ARMA-malleista, joissa on enintään 20 viivettä sekä enintään viisi aiempaa virhetermiä, koska inflaatioaikasarjan autokorrelaatio- ja osittaisautokorrelaatio-funktiot antoivat olettaa, että valittu malli ei tarvitse kovin montaa viivettä liikkuvan keskiarvon komponenttiin.

Molemmilla Akaiken ja Bayesin informaatiokriteereillä malli ARMA(13,0) osoittautui parhaaksi. Koska molemmat informaatiokriteerit olivat mallista yksimielisiä, valitaan se jatkotarkastelua varten. ARMA(13,0)-malli estimoidaan käyttämällä suurimman uskottavuuden menetelmää. Taulukossa 7 on tietoa mallin ennustevirheistä estimointiotoksessa.

Taulukko 7: ARMA(13,0)-mallin ennustevirheet estimointiotoksessa

Mittari	Arvo
ME	1.21127e-05
RMSE	0.001547987
MAE	0.001191922

Box-Jenkins-metodissa mallin valinnan viimeisenä tasona on valitun mallin diagnostinen tarkastelu. Mallin diagnostisilla tarkasteluilla tarkoitetaan keinoja, joilla varmistetaan siitä, että valittu malli täyttää mallin ehdot ja sopii hyvin aineistoon (Box ym. 2015).

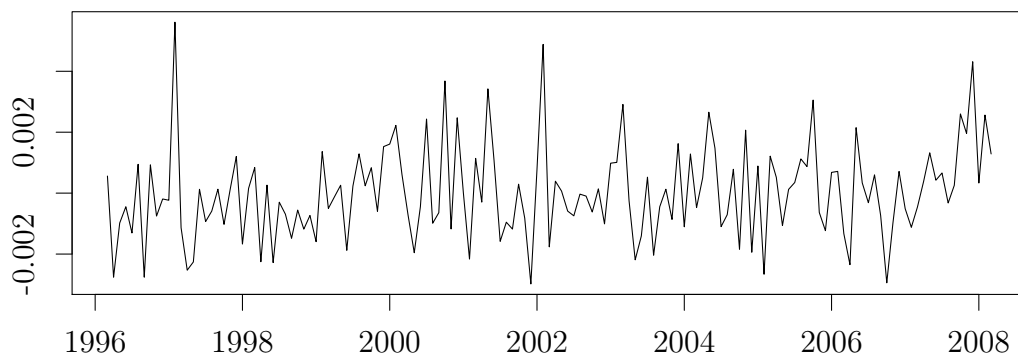
Valitun mallin diagnostiset tarkastelut perustuvat lähinnä mallin residuaalien tutkimiseen. ARMA-mallin oletuksena on, että mallin virhetermit noudattavat valkoinen kohina -prosessia.

Kuvassa 5 on ARMA(13,0)-mallin residuaalien aikasarja sekä autokorrelaatiofunktio. Residuaalien aikasarja näyttää päällisin puolin stationaariselta nollan ympärillä. Residuaalien autokorrelaatiofunktio osoittaa, että residuaalien välillä ei selkeästi ole autokorrelaatiota. Sama voidaan todentaa myös Ljung-Box-testillä, joka testaa nollahypoteesia onko aineisto itsenäisesti jakautunut. Taulukossa 8 ovat residuaaliaikasarjan Ljung-Box-testin tulokset, joista nähdään, että testiarvo ei ole tilastollisesti merkitsevä perinteisillä luottamustasoilla, mikä antaa todisteita residuaalien itsenäisestä jakaumasta.

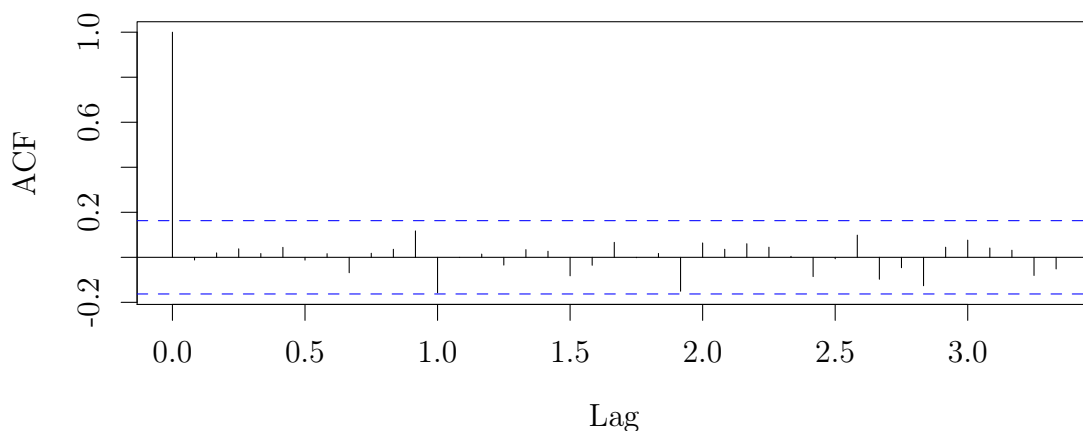
Taulukko 8: Ljung-Box-testi ARMA(13, 0)-mallin residuaaleille

Testisuureen arvo	Vapausasteet	p-arvo
10.489	7	0.1625

ARMA(13,0) residuaalit



Residuaaliautokorrelaatio



Kuva 6: ARMA(13, 0)-mallin residuaalien aikasarja ja autokorrelaatio

Näistä diagnostisista tarkasteluista voidaan päätellä, että ARMA-mallin oletukset pitävät paikkansa ja ARMA(13, 0)-mallia voidaan käyttää inflaatioennusteiden tekemiseen.

4.2.2 Ennusteet

Inflaatioennusteet taantuman aikana tehdään ARMA(13,0)-mallilla luvussa 4.2 kuvulla tavalla ajanjaksolle 2008:4-2009:6. Taulukossa 9 on perustiedot ennustevirheistä kuukauden, kolmen kuukauden ja kahdentoista kuukauden päähän menevistä ennusteista taantuman aikana. Taulukosta nähdään, että ARMA-malli tekee keskimäärin hyvin pieniä ennustevirheitä taantuman aikana. Samoin ennusteiden varianssi on erittäin pieni. Yhden ja kolme kuukautta eteenpäin menevissä ennusteissa ennusteet ovat keskimäärin liian suuria ja vuoden päähän menevissä ennusteissa liian pieniä.

Taulukko 9: ARMA(13,0)-mallin ennustevirheet taantuman aikana

	1 kuukausi	3 kuukautta	12 kuukautta
Keskiarvo	-0.001361995	-0.00232965	0.000151599
Varianssi	1.281733e-05	1.323115e-05	1.265778e-05
Keskihajonta	0.00358013	0.003637465	0.003557777
MSE	1.381788e-05	1.777635e-05	1.183691e-05
RMSE	0.00371724	0.004216201	0.00344048

Samoin ennusteet tehtiin niin sanotun suuren taantuman jälkeiselle ajalle, jolloin euroalueen bruttokansantuote kasvoi. Taulukossa 10 ovat tiedot ei-taantuman aikana tehdyistä ennusteista. ARMA-malli ennusti taantuman jälkeisellä ajalla huomattavasti paremmin keskiarvon perusteella yksi ja kolme kuukautta eteenpäin. Vuodella eteenpäin menevissä ennusteissa ARMA-malli oli edelleen parempi noususuhdanteessa, mutta ero on paljon pienempi. Myös ennustevirheiden varianssi oli pienempi taantuman jälkeen kolme ja kaksitoista kuukautta eteenpäin tehdyissä ennusteissa.

On kuitenkin huomioitava, että keskineliövirheen neliöjuurella mitattuna ARMA-malli toimi paremmin taantuman jälkeen kuin taantuman aikana. Esimerkiksi yhden kuukauden päähän menevien ennusteiden keskineliövirheen neliöjuuri oli noin 0.0037 taantuman aikana ja 0.0034 taantuman jälkeen. Ero kasvaa kolmen kuukauden päähän tehtävissä ennusteissa. Eroja ei voida kuitenkaan vertailla täysin toisiinsa, koska ennusteet on tehty eri ajanjaksolle ja eri aineistolla.

Taulukko 10: ARMA(13,0)-mallin ennustevirheet noususuhdanteen aikana

	1 kuukausi	3 kuukautta	12 kuukautta
Keskiarvo	-9.722039e-05	-4.354527e-05	0.0002550767
Varianssi	1.295072e-05	8.569063e-06	5.695986e-06
Keskihajonta	0.003598711	0.002927296	0.002386627
MSE	1.209679e-05	7.999689e-06	5.381318e-06
RMSE	0.003478044	0.002828372	0.002319767

4.3 Ennusteiden vertailu

Tässä osiossa vertaillaan neuroverkkomallin ja ARMA-mallin ennusteita osion 3.5 kriteereillä. Taulukkoon 11 on kerätty NN(17, 2)- ja ARMA(13, 0)-mallien keskineliövirheen neliöjuurella mitattujen ennustevirheiden erotukset. Positiivinen arvo tarkoittaa sitä, että ARMA-mallin keskineliövirheen neliöjuuri on pienempi. Vastaavasti negatiivinen luku merkitsee neuroverkon olevan parempi samalla mittarilla.

Taulukko 11: Neuroverkko- ja ARMA-mallien RMSE:n erotus

	1 kuukausi	3 kuukautta	12 kuukautta
Taantuma	-8.140091e-05	-9.785345e-05	2.304639e-06
Ei-taantuma	-0.0004747582	-0.0006154781	3.221895e-05

Taulukosta nähdään, että näiden mallien ennustevirheissä ei ole erityisen suuria eroja. Tämä antaa todisteita siitä, että molemmat mallit sopivat yhtä hyvin inflaation ennustamiseen euroalueella. Taulukosta myös nähdään, että neuroverkko päihittää ARMA-mallin taantuman aikana yksi ja kolme kuukautta eteenpäin tehtävissä ennusteissa, mutta ARMA-malli on parempi vuoden päähän tehtävissä ennusteissa.

Taantuman jälkeisellä ajanjaksolla tilanne on lähes vastaava. ARMA-malli ennustaa inflaatiota paremmin kuin neuroverkko 12 kuukautta eteenpäin, mutta yksi ja kolme kuukautta eteenpäin tehtävissä ennusteissa neuroverkko on voittaja. Erot keskineliövirheen neliöjuurissa ovat kuitenkin isompia noususuhdanteen aikana tehdyissä ennusteissa, mikä tarkoittaa ehkä sitä, että jompikumpi malleista toimii toista paremmin

noususuhdanteessa.

Koska erot valitussa arviointikriteerissä ovat erittäin pieniä, erojen todenperäisyyttä täytyy tutkia pidemmälle. Tähän käytetään luvussa 3.5 esiteltyä muokattua Diebold-Mariano-testiä, jolla vertaillaan ARMA-mallin ja neuroverkon ennustevirheitä sekä taantumaa että talouskasvun aikana kaikille kolmelle ennustevälille. Testissä käytetään tapiofunktiona virheiden neliösummaa, jota käytettiin myös neuroverkkoeestimoinnin tapiofunktiona.

Taulukossa 12 ovat testin tulokset taantumaa ajalle. Millään ennustevälillä testi ei ole tilastollisesti merkitsevä tavanomaisilla luottamustasoilla, joten voidaan päätellä, että neuroverkkomalli ja ARMA-malli eivät tuota erilaisia ennusteita taantumaa aikana. Erot keskineliövirheen neliöjuurissa voidaan siis päätellä johtuvan vain satunnaisuudesta.

Taulukko 12: Diebold-Mariano-testi taantumaa aikana

Ennusteväli	Testisuureen arvo	p-arvo
1 kk	-0.21571	0.8323
3 kk	-0.3378	0.7405
12 kk	0.0073353	0.9943

Taulukossa 13 ovat vastaavat Diebold-Mariano-testin tulokset talouskasvun ajalle. Kuten taantumankin aikana, kaikilla ennusteväleillä testin p-arvot ovat perinteisten rajojen yläpuolella, joten myöskään noususuhdanteen aikana ei voida sanoa olevan tilastollisesti merkitsevää eroa tutkittujen mallien ennusteissa.

Taulukko 13: Diebold-Mariano-testi talouskasvun aikana

Ennusteväli	Testisuureen arvo	p-arvo
1 kk	-0.87925	0.3941
3 kk	-0.89364	0.3866
12 kk	0.047507	0.9628

Vaikka mallien tuottamissa ennusteissa oli eroa sekä taantumaa että noususuhdanteen aikana, erot olivat pieniä eivätkä ne olleet tilastollisesti merkitseviä. Voidaan siis

sanoa, että neuroverkko- ja ARMA-malli ennustavat euroalueen inflaatiota yhtä hyvin sekä taantumassa että noususuhdanteessa kun inflaatioaikasarja on kausitasoitettu.

Kausitasoittamattomalla aikasarjalla tilanne on hieman erilainen (ks. liite). ARMA-mallin tuottamat ennusteet olivat tilastollisesti merkitsevästi parempia 12 kuukautta eteenpäin taantumän aikana. Taantumän jälkeen ARMA-mallin ennusteet olivat Diebold-Mariano-testin mukaan parempia sekä kolme että kaksitoista kuukautta eteenpäin. Voidaan päätellä, että vaikka aikasarjan kausitasoittaminen parantaa neuroverkko-mallin ennusteita siten, että ne ovat yhtä hyviä ARMA-mallin tekemien ennusteiden kanssa.

5 Pohdinta

Tässä osiossa vastataan tutkielman johdannossa esitettyihin tutkimuskysymyksiin edellisen luvun tulosten valossa. Lisäksi näitä tuloksia verrataan aikaisempaan kirjallisuuteen. Erityisesti vertailua tehdään ainoaan aiempaan koko euroaluetta koskevaan Binnerin ym. (2005) tutkimukseen. Lisäksi pohditaan tulosten luotettavuutta ja yleistettävyyttä sekä mahdollisia suosituksia rahapolitiikkaan.

Kuinka hyvin neuroverkkomalli ennustaa euroalueen inflaatiota? Edellisen luvun tulosten perusteella hyvin. Neuroverkkomallin ennustevirheet olivat erittäin pieniä sekä taantuman aikana että sen jälkeen tehdyissä ennusteissa. Tässä tutkielmassa saadut tulokset ovat linjassa Binnerin ym. (2005) tutkimuksen kanssa. Erona Binnerin ym. (2005) tutkimukseen oli se, että tässä tutkielmassa käytettiin vain inflaation aiempia arvoja tulevan inflaation ennustamiseen. Voidaan kuitenkin päätellä, että myös tällainen yksiulotteinen neuroverkkomalli kykenee ennustamaan euroalueen inflaatiota todella hyvin. Vastaavia tuloksia on saatu myös muita maita koskevissa tutkimuksissa.

Ennustaako neuroverkkomalli inflaatiota merkittävästi paremmin taantuman aikana kuin lineaarinen ARMA-malli? Tähän vastaus on ei. Neuroverkkomallin tuottamat ennusteet olivat parempia keskineliövirheen neliöjuurella mitattuna yksi ja kolme kuukautta eteenpäin verrattuna ARMA-malliin, mutta mallien tuottamat ennustevirheet eivät olleet tilastollisesti merkitsevästi erilaisia. Myöskään vuoden päähän menevissä ennusteissa ei ollut tilastollisesti merkitsevää eroa. Kausitasoittamattomalla aikasarjalla erot olivat samanlaisia, mutta vertailumalli oli tilastollisesti merkitsevästi parempi pidemmillä ennusteväleillä. Neuroverkkomalli toimii siis paremmin, jos tutkittavasta aikasarjasta on poistettu kausivaihtelu.

Tässä tutkielman tulokset ovat ristiriidassa Binnerin ym. (2005) tutkimukseen, jossa parhaimmat neuroverkkomallit olivat parempia kuin ARIMA-mallit. Erot tuloksissa voivat johtua eroista valituissa malleissa. Binner ym. (2005) käyttivät malleja, joissa on myös muita muuttujia mukana. Tässä tutkielmassa on käytetty myös pidempää aikasarjaa kuin aiemmassa kirjallisuudessa, joten teoreettisesti ajateltuna tämän tutkielman neuroverkon pitäisi toimia paremmin kuin lyhyemmällä aikasarjalla koulutetun mallin.

Toisaalta Binner ym. (2005) eivät tutkineet tulostensa tilastollista merkitsevyyttä samalla tavalla kuin tässä tutkielmassa. Erot heidän tutkimuksessaan neuroverkkomallien ja vertailumallien välillä olivat pieniä, joten mallien paremmuusjärjestystä on vaikea todentaa ilman tilastollisia testejä. Tässä tutkielmassa käytettiin ennusteiden vertailuun yleisesti käytettyä Diebold-Mariano-testiä, joka osoitti mallien keskineliövirheellä mitattujen erojen olevan otoksesta johtuvaa satunnaisuutta eikä neuroverkkomallin paremmuutta.

Entä onko neuroverkkomallin ennusteissa eroja taantumien ja noususuhdanteiden aikana? Keskineliövirheen neliöjuurella mitattuna neuroverkkomalli teki pienempiä virheitä taantumien jälkeisenä aikana kaikilla ennusteväleillä. Tämä antaa todisteita siitä, että neuroverkko voisi toimia paremmin muissa suhdannetilanteissa kuin taantumien aikana. Ennusteet eivät kuitenkaan ole suoraan verrattavissa toisiinsa.

Toinen mahdollisuus vastata kysymykseen löytyy Diebold-Mariano testeistä. Taantumien ja noususuhdanteiden välillä ei ollut eroja testien lopputuloksessa. Ei voida siis suoraan päätellä, että neuroverkko toimisi eri tavalla taantumien aikana kuin talouskasvussa. Kausitasoittamattomalla aineistolla kolme kuukautta eteenpäin tehtävissä ennusteissa vertailumalli oli tilastollisesti merkitsevästi parempi noususuhdanteessa, mutta samanlaista eroa ennusteissa ei ollut taantumien aikana. Tämä antaa olettaa, että neuroverkkomalli toimisi keskipitkillä aikaväleillä paremmin taantumien aikana kuin noususuhdanteissa yksinkertaiseen ARMA-malliin verrattuna.

Mutta johtuvatko erot keskineliövirheen neliöjuuressa kuitenkin suhdannetilanteesta? Tätä on erittäin hankala sanoa tämän tutkielman valossa. Tutkimuskysymystä on täysin mahdotonta tutkia niin, että voitaisiin tehdä kausaalipäätelyä, koska samaa ajanjaksoa ei voida havaita sekä taantumassa että noususuhdanteessa. Tässä tutkielmassa on pyritty minimoimaan ajan vaikutuksia ottamalla vertailukohta heti taantumien jälkeen, jolloin esimerkiksi rahapolitiikan muutokset eivät ole ehtineet vaikuttaa merkitsevästi inflaatioon.

Tässä tutkielmassa keskityttiin myös vain euroalueen inflaatioon kokonaisuutena. Yksittäisten euromaiden inflaatioissa voi olla eroja taantumien ja noususuhdanteiden välillä neuroverkkomallilla ennustettaessa, mutta kysymys ei ole siinä mielessä merkitykselli-

nen, koska euroalueen rahapolitiikkaa harjoitetaan koko euroalueen inflaation pohjalta. Tuloksia ei voida myöskään yleistää euroalueen ulkopuolelle.

Tutkielma antaa kuitenkin todisteita aiemmille tutkimuksille siitä, että neuroverkot sopivat inflaation ennustamiseen hyvin. Tämä tutkielma antaa myös todisteita siitä, että neuroverkko toimisi yhtä hyvin sekä taantumassa että noususuhdanteessa. Tutkielma antaa kuitenkin myös tukea aiemmille tutkimuksille kuten Álvarez-Díaz ja Gupta (2016) ja He ym. (2012), että neuroverkkomalli ei kuitenkaan tuota parempia ennusteita kuin ARIMA-malli. Toisaalta päinvastoin kuin Moshirin ja Cameronin (2000) tutkimuksessa, tässä tutkielmassa neuroverkko toimi merkittävästi huonommin pidemmällä aikavälillä kuin ARMA-malli. Tutkielma antaa myös todisteita aikasarjan kausitasoittamisen hyödyistä neuroverkkomallin ennusteille kuten Zhangin ja Qin (2005) tutkimuksessa.

Vaikuttaisi siltä, että tämän tutkielman perusteella Stockin ja Watsonin (2007) yleistys inflaatiosta pitää paikkansa myös euroalueella. Ennustemallit tuottavat erittäin pieniä ennustevirheitä, mutta monimutkaisempien mallien (kuten neuroverkkojen) käytöstä ei tule merkittävää lisähyötyä. Tämä vaikuttaisi pätevän niin taantumassa kuin noususuhdanteessakin.

Tulos voi myös johtua siitä, että inflaatio euroalueella on lineaarinen ilmiö. Jos inflaation dataa generoiva prosessi olisi epälineaarinen, neuroverkkomallin pitäisi teoriassa tuottaa parempia ennusteita kuin ARMA-mallin. Lineaarisen dataa generoivan prosessin tapauksessa neuroverkkomallin ennusteiden pitäisi olla huonompia tai yhtä hyviä kuin ARMA-mallin, mikä on totta tämän tutkielman valossa.

Toisena selityksenä tutkielman tuloksille voi olla se, että neuroverkkomallit tarvitsevat paljon dataa, jotta ne toimisivat optimaalisesti. Euroalueen inflaatioaikasarja on lyhyt verrattuna esimerkiksi Yhdysvaltojen vastaavaan, joten tulokset voisivat olla erilaisia tulevaisuudessa, jos vastaavanlaisia taantumia euroalueella ilmaantuu.

Tutkielman tulosten perusteella yksinkertaisten eteenpäin syöttävien neuroverkkojen käyttöä esimerkiksi keskuspankkitoiminnassa ei voida suositella. Rahapolitiikkaan vaikuttavat inflaatioennusteet, joten optimaalinen rahapolitiikka vaatisi, että inflaatiota pystytään ennustamaan mahdollisimman tarkasti. Tämän tutkielman perusteella neuroverkkomallit eivät ennusta inflaatiota paremmin kuin ARMA-mallit, joten niitä

ei voida suositella käytettävän lineaaristen mallien sijasta.

Lisäksi neuroverkkomallien rakentamiseen menee paljon aikaa, mutta tulokset eivät ole parempia kuin hyvin yksinkertaisten mallien, jotka voidaan estimoida perinteisillä menetelmillä hyvinkin nopeasti. Lisäksi ennustamiseen käytettyä neuroverkkoa ei voida arvioida ennen kuin se on tehnyt ennusteita, koska mallin sopivuuteen ei ole testejä kuten ARMA-mallilla.

Keskuspankkien toiminnan arvioinnissa neuroverkkomalli olisi siinä mielessä hyödyllinen, että sen rakentaminen ei vaadi estimoitujen parametrien tulkintaa tai huomattavasti muuta osaamista tilastotieteestä tai ekonometriasta. Vaikka neuroverkkomalli ei olisikaan tarkoin mahdollinen ennustemalli euroalueen inflaatiolle, antaa se työkalun keskuspankin harjoittaman rahapolitiikan arvioinnille.

6 Johtopäätökset

Tässä tutkielmassa tutkittiin miten eteenpäin syöttävä neuroverkkomalli ennustaa euroalueen inflaatiota taloudellisen taantuman aikana. Tarkoituksena oli tutkia miten tarkkoja ennusteita neuroverkkomalli tekee taantuman aikana, ovatko ennusteet parempia kuin lineaarisen vertailumallin, ja eroaako ennustetarkkuus jotenkin taantuman ja noususuhdanteen välillä. Oletuksena neuroverkkomallin tulisi tehdä parempia ennusteita taantuman aikana kuin lineaarisen vertailumallin, jos inflaation dataa generoiva prosessi on epälineaarinen taloudellisessa taantumassa.

Tutkielmassa rakennettiin eteenpäin syöttävä neuroverkko käyttämällä aineistona euroalueen harmonisoidusta kuluttajahintaindeksistä muodostettua inflaatioaikasarjaa ennen niin sanottua euroalueen suurta taantumaa (2008-2009). Mallinvalinnassa hyödynnettiin koneoppimiskirjallisuudessa vakiintunutta validaatio-otostaktiikkaa, jossa mallin parametrit valitaan käyttämällä erillistä otosta parhaan ennusteen antavan mallin löytämiseksi. Lisäksi samalla aineistolla valittiin lineaarinen ARMA-malli, jota käytettiin vertailukohtana neuroverkolle inflaation ennustamisessa.

Tutkielmassa selvisi, että neuroverkkomalli tekee hyvin tarkkoja ennusteita sekä taantumassa että noususuhdanteessa aiemmassa kirjallisuudessa käytetyillä arviointikriteereillä. Se ei kuitenkaan kyennyt tuottamaan tilastollisesti merkitsevästi parempia ennusteita kuin lineaarinen vertailumalli taantumassa tai sen jälkeen. Lyhyellä ennustevälillä neuroverkkomalli oli yhtä hyvä kuin yksinkertainen lineaarinen malli, ja vuoden päähän ennusteet olivat merkitsevästi huonompia. Kolme kuukautta eteenpäin tehtävissä ennusteissa neuroverkkomalli toimi paremmin taantuman aikana kuin nousuhdanteessa kun aikasarjasta ei poistettu kausivaihtelua, mikä näkyi myös ennustemalleja vertailevassa tilastollisessa testissä.

Tutkielman tulosten perusteella voidaan olettaa, että euroalueen dataa generoiva prosessi on lineaarinen sekä taantumassa että nousuhdanteessa. Tuloksista voidaan myös päätellä, että neuroverkkomallien käyttö inflaation ennustamisessa ei ole mielekästä keskuspankin näkökulmasta, koska neuroverkkomallien rakennus vie enemmän resursseja kuin perinteisten lineaaristen mallien, mutta tulokset ovat huonompia ra-

hapolitiikan kannalta merkittävillä aikaväleillä. Keskuspankkien toimintaa arvioidessa neuroverkkomallit voivat tuoda jotain hyötyä, vaikka ne eivät olisikaan tarkimpia mahdollisia malleja.

Tämä tutkielma luo kuitenkin pohjaa lisätutkimukselle aiheesta. Tutkielmassa käytettiin aiemmassa kirjallisuudessa suosituinta neuroverkkomallityyppiä, joten tulevaisuudessa tutkimukselle voisi kehittää jatkoa käyttämällä esimerkiksi uusiutuvia neuroverkkoja. Vastaavanlaisia tutkimuksia voitaisiin toteuttaa käyttämällä myös muita uusia koneoppimistekniikoita kuten tukivektorikoneita, jotka ovat yleistyneet ekonometriassa.

Lisäksi tulevaisuudessa voitaisiin tutkia euroalueen maita yksittäin tai käyttää esimerkiksi euroalueen kaikkia taantuma-aikoja tutkimuksessa pelkän suuren taantuman sijaan. Lisäksi voitaisiin käyttää muiden maiden, esimerkiksi Yhdysvaltojen, inflaatio-dataa, jota voi olla paljon euroaluetta enemmän.

Lähteet

- [1] Aiken M. (1999). Using a neural network to forecast inflation. *Industrial Management & Data Systems* 99:7, 296-301.
- [2] Álvarez-Díaz M. ja R. Gupta (2016). Forecasting US consumer price index: does nonlinearity matter? *Applied Economics* 48:46, 4462-4475.
- [3] Ang, A., G. Bekaert ja M. Wei (2007) Do macro variables, asset markets, or surveys forecast inflation better? *Journal of Monetary Economics* 54:4, 1163-1212.
- [4] Armstrong J. (2001) Evaluating Forecasting Methods. Teoksessa: *Principles of Forecasting*, toim. J. Armstrong. Springer, Boston, MA.
- [5] Atkeson, A. ja L.E. Ohanian (2001), Are Phillips Curves Useful for Forecasting Inflation? *Federal Reserve Bank of Minneapolis Quarterly Review* 25:1, 2-11.
- [6] Binner J., R. Bissoondeal, T. Elger, A. Gazely ja A. Mullineux (2005). A comparison of linear forecasting models and neural networks: an application to Euro inflation and Euro Divisia. *Applied Economics* 37:6, 665-680.
- [7] Binner J., C. Elger, B. Nilsson ja J. Tepper (2006). Predictable non-linearities in U.S. inflation. *Economics Letters* 93, 323-328.
- [8] Bishop, C. (1995). *Pattern recognition and machine learning*. Springer, New York.
- [9] Box, G., G. Jenkins, G. Reinsel ja G. Ljung (2015), *Time Series Analysis: Forecasting and Control*, John Wiley & Sons. New Jersey.
- [10] Briault, C. (1995) The costs of inflation. *Quarterly bulletin/Bank of England* 35, 33-45.
- [11] Brockwell, P. ja R. Davis (2002) *Introduction to Time series and forecasting*. Springer, New York.
- [12] Choudhary, M. ja A. Haider (2012). Neural network models for inflation forecasting: an appraisal. *Applied Economics* 44:20, 2631-2635.

- [13] Dickey D. ja W. Fuller (1979). Distribution of the Estimators for Autoregressive Time Series with a Unit Root, *Journal of the American Statistical Association*, 74:366a, 427-431.
- [14] Diebold, F. ja R. Mariano (1995). Comparing predictive accuracy. *Journal of Business & economic statistics*, 13:3, 253-263.
- [15] European central bank (2011) *The monetary policy of ECB*. Frankfurt am Main.
- [16] Eurostat (2018) Harmonised index of consumer prices (HICP). <https://ec.europa.eu/eurostat/web/hicp/data/database>. Haettu 20.11.2018.
- [17] Faust, J. ja J. Wright (2013) Forecasting inflation. *Handbook of Economic Forecasting*, 2:A, 2-56.
- [18] Fischer, S. (1993). The role of macroeconomic factors in growth. *Journal of Monetary Economics* 32:3, 485-512.
- [19] Friedman, M. (1968) The role of monetary policy. Teoksessa *Essential Readings in Economics* 215-231.
- [20] Fritsch S., F. Guenther, M. Wright, M. Suling ja S. Mueller (2019). neuralnet: Training of Neural Networks. R package version 1.44.2.
- [21] Gupta, S. ja S. Kashyap (2015). Forecasting inflation in G-7 countries: an application of artificial neural network. *Foresight* 17:1, 63-73.
- [22] Harvey, D., S. Leybourne ja P. Newbold (1997). Testing the equality of prediction mean squared errors. *International Journal of forecasting*, 13:2, 281-291.
- [23] Haykin S. (1999). *Neural Networks, A Comprehensive Foundation*. Pearson, Delhi.
- [24] He, Q., H. Shen ja Z. Tong (2012). Investigation of Inflation Forecasting. *Applied Mathematics & Information Sciences* 6:3, 649-665.
- [25] Hornik, K. (1991) Approximation Capabilities of Multilayer Feedforward Networks. *Neural Networks* 4, 251-257.

- [26] Hyndman, R. ja A. Koehler (2006). Another look at measures of forecast accuracy. *International journal of forecasting* 22:4, 679-688.
- [27] Kaastra I. ja M. Boyd (1996). Designing a neural network for forecasting financial and economic time series. *Neurocomputing* 10, 215-236.
- [28] Kourentzes N. (2019). nnfor: Time Series Forecasting with Neural Networks. R package version 0.9.6.
- [29] Lin J.-L. ja C. Granger (1994) Forecasting from non-linear models in practice. *Journal of Forecasting* 13:1, 1-9.
- [30] Moshiri, S. ja N. Cameron (2000). Neural Network Versus Econometric Models in Forecasting Inflation. *Journal of Forecasting* 19, 201-217.
- [31] Moshiri, S., N. Cameron ja D. Scuse (1999). Static, Dynamic and Hybrid Neural Networks in Forecasting Inflation. *Computational Economics* 14, 219-235.
- [32] Nakamura, E. (2005). Inflation forecasting using a neural network. *Economics Letters* 86, 373-378.
- [33] Qi M. ja G. Zhang (2001). An investigation of model selection criteria for neural network time series forecasting. *European Journal of Operational Research* 132, 666-680.
- [34] Remus W. ja M. O'Connor (2001) Neural networks for time-series forecasting. Teoksessa *Principles of Forecasting*. toim. J. Armstrong. Springer, Boston.
- [35] Riedmiller M. (1994) Rprop - Description and Implementation Details. *Technical Report* University of Karlsruhe.
- [36] Riedmiller M. ja H. Braun (1993) A Direct Adaptive Method for Faster Backpropagation Learning: The RPROP Algorithm. *Proceedings of the IEEE international conference on neural networks*. Vol. 1993.
- [37] Stock, J.H. ja M.W. Watson (1999) Forecasting inflation. *Journal of Monetary Economics* 44. 293-335.

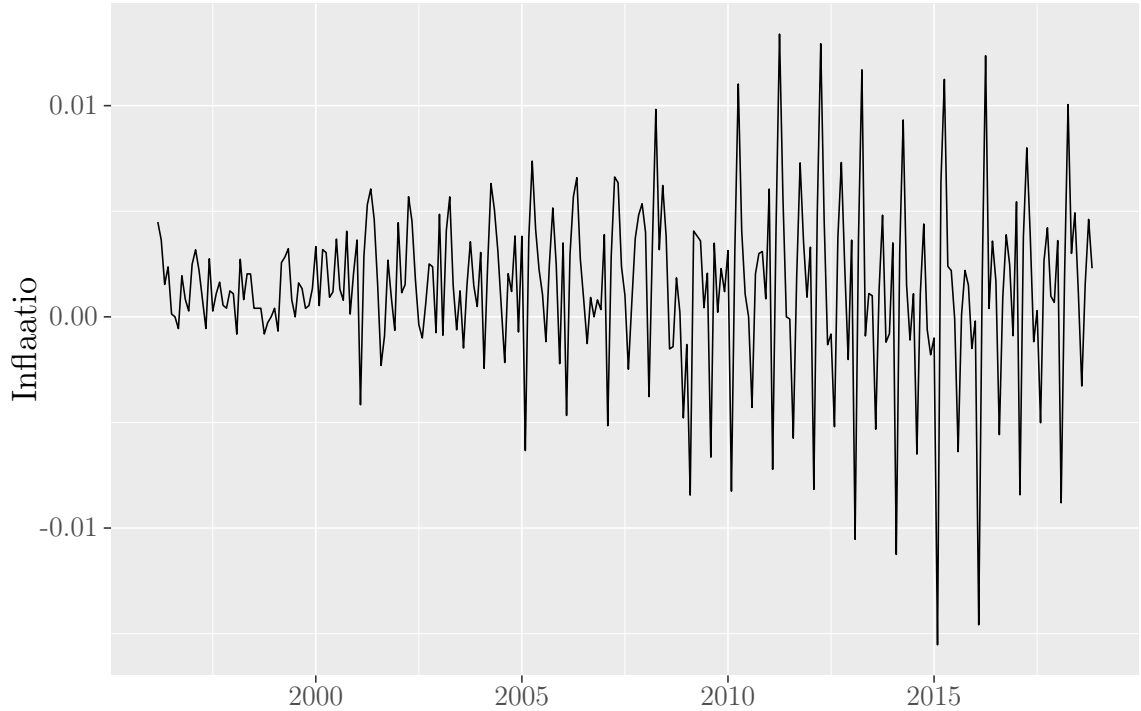
- [38] Stock, J.H. ja M.W. Watson (2007) Why has U.S. Inflation Become Harder to Forecast? *Journal of Money, Credit and Banking* 39:1, 3-33.
- [39] Stock, J.H. ja M.W. Watson (2009) Phillips Curve Inflation Forecasts. Teoksessa: *Understanding Inflation and the Implications for Monetary Policy*, toim. J. Fuhrer, Y. Kodrzycki, J. Little ja G. Olivei. MIT Press, Cambridge.
- [40] Svensson, L. (1997) Inflation forecast targeting: Implementing and monitoring inflation targets. *European Economic Review* 41:6, 1111-1146.
- [41] Svensson, L. (2010) Inflation Targeting. *Handbook of Monetary Economics* 3, 1237-1302.
- [42] Tkáč, M. ja R. Verner (2016). Artificial neural networks in business: Two decades of research. *Applied Soft Computing* 38, 788-804.
- [43] Zhang G., E. Patuwo ja M. Hu (1998). Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting* 14, 35-62.
- [44] Zhang, G. ja M. Qi (2005) Neural network forecasting for seasonal and trend time series. *European Journal of Operational Research* 160, 501-514.
- [45] Zhang, G. (2012) Neural Networks for Time-Series Forecasting. Teoksessa *Handbook of Natural Computing* toim. Rozenberg G., T. Bäck, J. Kok. Springer, Berlin, Heidelberg.

Liite: Kausitasoittamattoman inflaatioaikasarjan analyysi

Tässä osiossa tutkitaan neuroverkkomallin ennustekykyä taantuman aikana kausitasoittamattomalla aikasarjalla. ARMA-mallin oletuksena on aikasarjan stationaarisuus, joka ei pidä paikkaansa, jos aikasarjassa on kausivaihtelua (Brockwell & Davis 2002). Sitä vastoin neuroverkkomalli ei vaadi teoriassa kausivaihtelun poistamista aineistosta, koska neuroverkon joustavuuden pitäisi ottaa myös tutkittavan ilmiön kausivaihtelu huomioon (Gorr 1994, ks. Zhang 2012). Zhang ja Qi (2005) ovat kuitenkin tulleet tulokseen, että kausivaihtelun poistaminen aineistosta parantaa myös neuroverkkomallin ennustekykyä, koska neuroverkot eivät kykene käytännössä ottamaan kausivaihtelua huomioon tehokkaasti. Koska kausivaihtelu voi olla monimutkaista, eikä sitä välttämättä voida hävittää aineistosta kokonaan, on neuroverkkomallin ennustekykyä taantuman aikana hyvä tutkia myös kausitasoittamattomalla aineistolla, ja tutkia eroavatko tulokset näissä kahdessa tapauksessa tulosten robustisuuden parantamiseksi.

Tässä osiossa toistetaan luvun 4 analyysi käyttämällä kausitasoittamatonta euroalueen inflaatioaikasarjaa (kuvio 7), joka on tuotettu ottamalla ensimmäinen differenssi logaritmoidusta harmonisoidusta kuluttajahintaindeksistä. Mallien valinta suoritetaan ARMA-mallille osiossa 3.2 kuvailtua Box-Jenkins-metodia käyttäen, ja neuroverkon parametrit valitaan käyttämällä samaa koulutus- ja validointiotosta kuin osiossa 3.4. Ennusteet tehdään euroalueen taantuma-ajalle 2008:4-2009:6 sekä taantuman jälkeiselle kaudelle 2009:7-2010:9. Neuroverkkomallin ja ARMA-mallin ennusteita verrataan toisiinsa sekä taantuman aikana että taantuman jälkeen osiossa 3.5 kuvatuilla mittareilla: Käyttämällä lähinnä keskineliövirheen neliöjuurta sekä muokattua Diebold-Mariano-testiä.

Validaatio-otoksen perusteella paras neuroverkkomalli kausitasoittamattomalle inflaatioaikasarjalle on verkko, jossa on 18 viivettä ja kahdeksan kätkeytyä neuronია kätkeytyllä tasolla. Valittu malli ei viiveiden osalta erityisesti eroa kausitasoitettulle sarjalle valitusta mallista. Kätkeytyjen neuroneiden määrä on vain selkeästi suurempi, mikä voi johtua neuroverkon yrityksestä ottaa aikasarjassa oleva kausivaihtelu huomioon para-



Kuva 7: Euroalueen kuukausittainen kausitasoittamaton inflaatio

metrejä lisäämällä. Valitun neuroverkon keskineliövirhe validaatio-otoksessa oli 0.03534 ja koko koulutusotoksessa 1.98×10^{-9} .

Taulukoissa 14 ja 15 on koottu perustiedot neuroverkkomallin ennusteista taantuman ja sen jälkeiselle ajalle kausitasoittamattomalle inflaatiolle. Neuroverkkomallin keskineliövirheen neliöjuuri on hyvin samansuuntainen kuin kausitasoitettussa aikasarjassa taantuman aikana, mutta keskineliövirheen neliöjuuri on suurempi kausitasoittamattomassa sarjassa yksi ja 12 kuukautta eteenpäin, ja pienempi kolme kuukautta eteenpäin menevissä ennusteissa. Taantuman jälkeisenä aikana neuroverkkomalli tekee keskimäärin hieman pienempiä virheitä yksi ja 12 kuukautta eteenpäin menevissä ennusteissa. Noususuhdanteessa kausitasoittamattomassa aikasarjassa neuroverkko tekee kuitenkin selkeästi isompia virheitä keskimäärin kuin kausitasoitettussa, mikä antaa todisteita kausivaihtelun poistamisen puolesta ennen ennustamista mallilla.

Taulukko 14: Neuroverkkomallin ennustevirheet kausitasoittamattomalle aikasarjalle taantuman aikana

	1 kuukausi	3 kuukautta	12 kuukautta
Keskiarvo	3.284945e-05	-0.0004751515	0.004992893
Varianssi	1.614071e-05	1.015904e-05	7.091786e-05
Keskihajonta	0.00401755	0.003187324	0.008421274
MSE	1.506574e-05	9.707536e-06	9.111898e-05
RMSE	0.003881461	0.003115692	0.009545626

Taulukko 15: Neuroverkkomallin ennustevirheet kausitasoittamattomalle aikasarjalle noususuhdanteessa

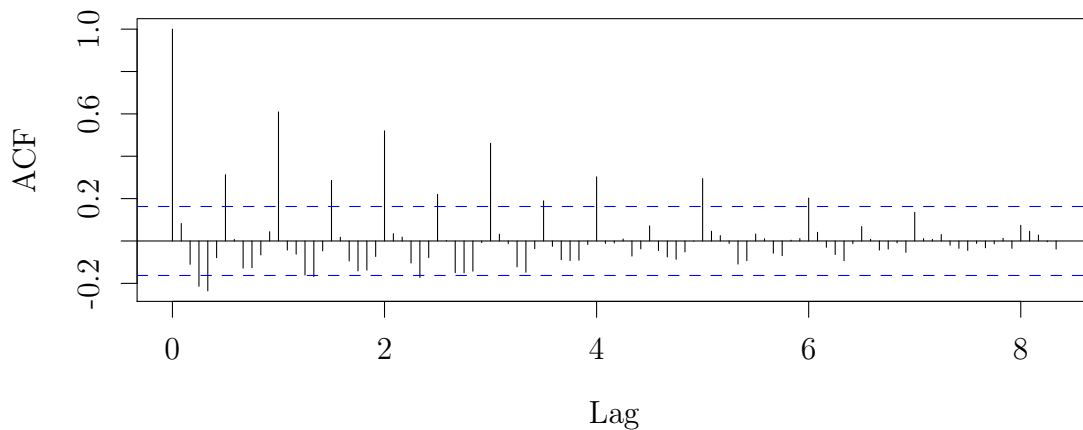
	1 kuukausi	3 kuukautta	12 kuukautta
Keskiarvo	0.0005378289	0.003192269	0.005783504
Varianssi	1.275754e-05	1.466872e-05	3.237841e-05
Keskihajonta	0.00357177	0.003829976	0.005690203
MSE	1.21963e-05	2.388139e-05	6.366877e-05
RMSE	0.00349232	0.004886859	0.007979271

Kausitasoittamattomalle inflaatioaikasarjalle valitaan ARMA-malli kuten osiossa 4.2.1. Kuviossa 8 ovat kausitasoittamattoman aikasarjan autokorrelaatio- ja osittaisautokorrelaatiofunktiot. Kuten kausitasoitettussa aikasarjassakin, kuviot eivät anna selkeää vastausta mallin parametrien valintaan, joten lopullinen valinta tehdään käyttämällä Akaiken ja Bayesin informaatiokriteerejä. Sekä Akaiken että Bayesin informaatiokriteerit osoittavat, että ARMA(12, 1)-malli olisi paras ARMA-malli inflaatioaikasarjalle. Malli eroaa kausitasoitettulle aikasarjalle valitusta mallista siten, että tässä mallissa on nyt myös mukana MA-prosessi.

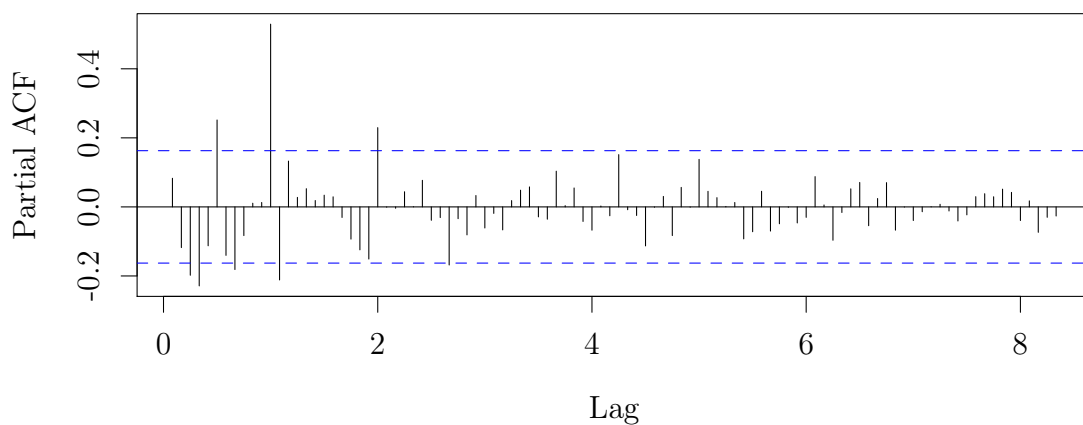
Kuviossa 9 ovat ARMA(12,1)-mallin residuaalit sekä residuaalien autokorrelaatiot. Residuaalit eivät osoita merkittävää autokorrelaatiota, minkä vahvistaa myös Ljung-Box-testi, jonka p-arvo on 0.126. Mallin voidaan olettaa sopivan hyvin kausitasoittamattomaan inflaatiodataan.

Taulukoissa 16 ja 17 ovat ARMA(12,1)-mallin tekemien ennusteiden yhteenveto

Autokorrelaatio



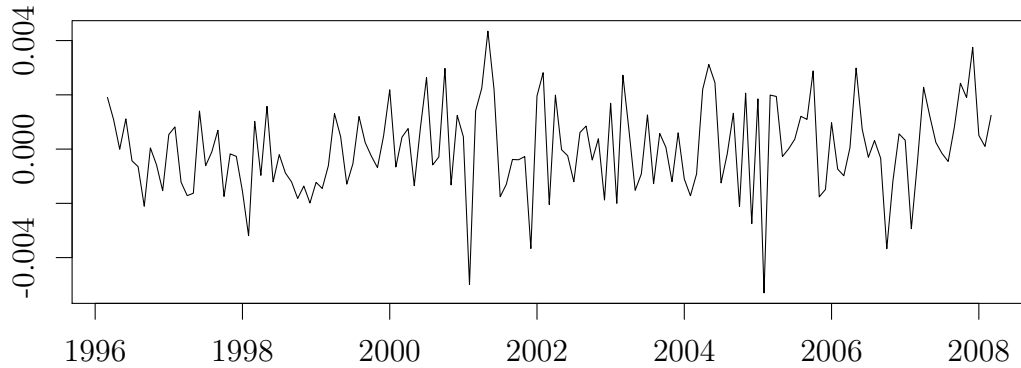
Osittaisautokorrelaatio



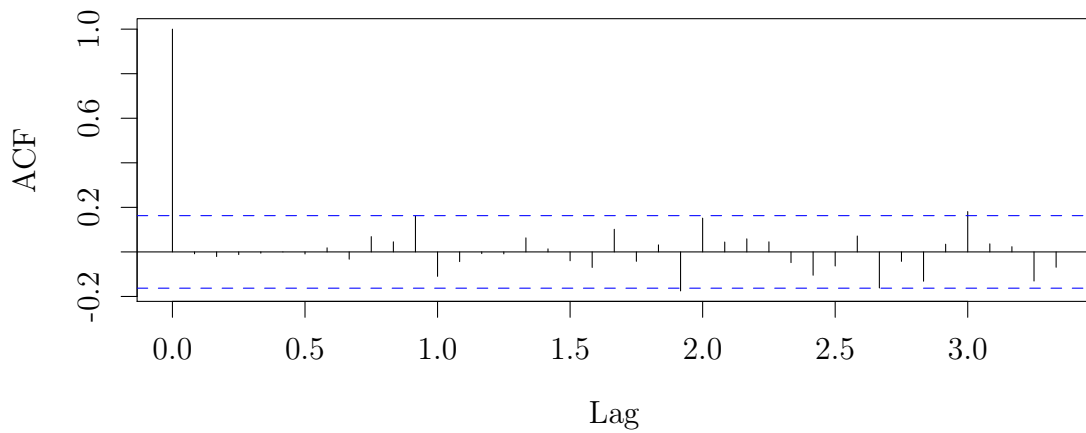
Kuva 8: Kausitasoitamattoman inflaatioaikasarjan autokorrelaatio- ja osittaisautokorrelaatio

taantumaa ja noususuhdanteen ajalle. Taantumaa aikana ARMA-malli tekee keskineliövirheen neliöjuurella mitattuna keskimäärin neuroverkkomallia suurempia virheitä yhden ja kolme kuukautta eteenpäin. Taantumaa jälkeisellä ajalla ARMA-malli on kuitenkin neuroverkkoa parempi ennustamaan inflaatiota kaikilla ennusteväleillä. Taantumaa aikana ARMA-mallin ennustevirheet ovat kuitenkin pienempiä kausitasoitettuissa aineistossa. Noususuhdanteessa erot keskineliövirheiden neliöjuurissa ovat mitättömän pieniä.

ARMA(12,1) residuaalit



Residuaaliautokorrelaatio



Kuva 9: ARMA(12, 1)-mallin residuaalien aikasarja ja autokorrelaatio

Taulukko 16: ARMA(12,1)-mallin ennustevirheet kausitasoittamattomalle aikasarjalle taantuman aikana

	1 kuukausi	3 kuukautta	12 kuukautta
Keskiarvo	-0.001625598	-0.002515147	0.0003286221
Varianssi	1.359301e-05	1.332062e-05	1.312615e-05
Keskihajonta	0.00368687	0.003649743	0.003623003
MSE	1.532938e-05	1.875854e-05	1.235906e-05
RMSE	0.003915275	0.004331113	0.003515546

Taulukko 17: ARMA(12,1)-mallin ennustevirheet kausitasoittamattomalle aikasarjalle noususuhdanteessa

	1 kuukausi	3 kuukautta	12 kuukautta
Keskiarvo	0.0001845285	0.0004089164	0.0002567571
Varianssi	1.280428e-05	8.601491e-06	5.634825e-06
Keskihajonta	0.003578306	0.00293283	0.002373779
MSE	1.198471e-05	8.195271e-06	5.325094e-06
RMSE	0.003461894	0.002862738	0.002307617

Ennustetulokset antavat olettaa, että neuroverkkomalli ennustaa inflaatiota paremmin taantuman aikana ainakin yksi ja kolme kuukautta eteenpäin. Vastaavasti voidaan olettaa, että ARMA-malli toimii paremmin muissa suhdannetilanteissa. Kuten luvussa 4.3., ennusteiden eroja testataan muunnellulla Diebold-Mariano-testillä, jolla tutkitaan eroavatko mallien tekemät ennusteet toisistaan tilastollisesti merkitsevästi. Taulukoissa 18 ja 19 ovat testin tulokset kuukauden, kolmen kuukauden ja vuoden päähän meneville ennusteille sekä taantumassa että taantuman jälkeisenä aikana. Yhden kuukauden päähän menevissä ennusteissa testisuure ei ole tilastollisesti merkitsevä, joten päätellään, että ARMA-malli ja neuroverkkomalli eivät tuota erilaisia ennusteita kun ennusteväli on yksi kuukausi, eikä suhdannetilanteella ole vaikutusta.

Kolmen kuukauden päähän menevissä ennusteissa Diebold-Mariano-testisuure on tilastollisesti merkitsevä perinteisillä luottamustasoilla vain taantuman jälkeisen ajan ennusteille. Neuroverkko- ja ARMA-mallit tuottavat siis samanlaisia ennusteita kolme kuukautta eteenpäin taantuman aikana, mutta noususuhdanteessa ARMA-malli on tilastollisesti merkitsevästi parempi.

Vuoden päähän menevissä ennusteissa testisuure on tilastollisesti merkitsevä sekä taantuman että noususuhdanteen aikana. ARMA-malli tuottaa siis parempia inflaatioennusteita vuoden päähän kuin neuroverkkomalli molemmissa suhdannetilanteissa.

Taulukko 18: Diebold-Mariano-testi taantuman aikana

Ennusteväli	Testisuureen arvo	p-arvo
1 kk	-0.049002	0.9616
3 kk	-1.5193	0.1509
12 kk	2.2254	0.043

Taulukko 19: Diebold-Mariano-testi noususuhdanteessa

Ennusteväli	Testisuureen arvo	p-arvo
1 kk	0.043519	0.9659
3 kk	2.2171	0.04368
12 kk	2.4001	0.03086