



Väylävirasto  
Trafikledsverket

Opinnäytetyö  
8/2020

Jarno Kajanne

# RAUTATIEINFRASTRUKTUURIN KÄYTÖN RAJAKUSTANNUSTEN EKONOMETRISEN MALLINNUKSEN TARKASTELU





Jarno Kajanne

**Rautatieinfrastruktuurin käytön  
rajakustannusten ekonometrisen  
mallinnuksen tarkastelu**

Opinnäytetyö 8/2020

Väylävirasto  
Helsinki 2020

*Kannen kuva: Väyläviraston kuva-arkisto*

Verkojulkaisu pdf ([www.vayla.fi](http://www.vayla.fi))

ISSN 2490-1202

ISBN 978-952-317-829-8

Väylävirasto

PL 33

00521 HELSINKI

Puhelin 0295 34 3000

**Jarno Kajanne: Rautatieinfrastruktuurin käytön rajakustannusten ekonometrisen mallinnuksen tarkastelu.** Väylävirasto. Helsinki 2020. Opinnäytetyö 8/2020. 62 sivua ja 3 liitettä. ISSN 2490-1202. ISBN 978-952-317-829-8.

**Asiasanat:** rautatie, infrastruktuuri, ratamaksut, mallinnus

## Tiivistelmä

Valtion rautatieinfrastruktuurin käytön hinnoittelu perustuu ratamaksuun, joka koostuu Suomessa tällä hetkellä vain ratamaksun perusmaksusta. Perusmaksun yksikkökustannus määritetään ekonometrisesti mallinnetun kustannusfunktion rajakustannuksena. Rautatieinfrastruktuurin kustannusfunktio mallinnuksen ongelmat, joista yksi oli mallin heteroskedastisuus, toimivat tämän työn lähtökohtana. Työn päätavoitteena oli tarkastella erilaisten kustannusfunktioiden sopivuutta Suomen rautatieinfrastruktuurin käytöstä aiheutuvien rajakustannusten määrittelyssä.

Tutkimuksessa tehtiin aluksi kirjallisuuskatsaus aiemmin tehtyihin rautatieinfrastruktuurin rajakustannusmallinnuksiin, ratamaksun perusmaksun lainsäädännölliseen taustaan sekä kolmen Euroopan talousalueen jäsenmaan ratamaksukäytäntöihin. Työn empiirisessä osiossa tutkittiin logaritmi-, Box-Cox- sekä Box-Tidwell-muunneltujen kustannusfunktioiden sopivuutta Suomen rautatieinfrastruktuurin rajakustannusten ekonometriseen mallinnukseen. Logaritmimuunneltu kustannusfunktio valikoitui jatkojalostuksen kohteeksi.

Työn kirjallisuuskatsauksesta huomattiin, että ratamaksun perusmaksun määrittelyn lainsäädännölliset rajaukset painottuvat laskentaan sisällytettäviin kustannuksiin, eikä ratamaksun perusmaksun laskentamenetelmiin. Laskentamenetelmien sääntelyn vähyydestä riippumatta rautatieinfrastruktuurin rajakustannusmallinnusten menetelmät painottuvat Euroopassa vain muutamiiin eri kustannusfunktio-tyyppeihin ja ne eroavat pääosin teknisten muuttujien osalta. Työn merkittävimpänä tuloksena huomattiin, että ratamaksun perusmaksun ekonometrisen mallinnuksen voidaan toteuttaa logaritmimuunnellulla poikkileikkausaineistolla, jossa usean vuoden paneeliaineiston aikasarja muuttujista laskettiin yhteen.

Työn empiriaosiossa kehitetty logaritmimuunnellulle poikkileikkausaineistolle tehty rajakustannusmallinnus täytti kaikki lineaariselle regressiomallille asetetut oletukset sekä mallinnuksen tuottama rajakustannus vastaa suuruudeltaan Suomessa aiemmin määritettyjä ja ratamaksuhinnastossa olevia rajakustannuksia. Mallinnusperiaatetta hyödynnettiin vuoden 2022 ratamaksun perusmaksun ekonometrisessä laskennassa syksyllä 2020

---

**Jarno Kajanne: Granskning av den ekonometriska modelleringen av marginalkostnader i användningen av järnvägsinfrastrukturen.** Trafikledsverket. Helsingfors 2020. Lärodomsprov 8/2020. 62 sidor och 3 bilagor. ISSN 2490-1202, ISBN 978-952-317-829-8.

## Sammanfattning

Statens prissättning på användningen av järnvägsinfrastrukturen baserar sig på banavgiften, som för närvarande endast består av banavgiftens grunddel i Finland. Enhetskostnaden för grunddelen fastställs som marginalkostnaden för den ekonometriskt modellerade kostnadsfunktionen. Problemen i modelleringen av kostnadsfunktionen för järnvägsinfrastrukturen, vilka omfattade heteroskedasticiteten i modellen, var utgångspunkten för detta arbete. Det huvudsakliga syftet med arbetet var att granska olika kostnadsfunktioners lämplighet i definieringen av marginalkostnader som orsakas av användningen av Finlands järnvägsinfrastruktur.

I studien gjordes inledningsvis en litteraturoversikt av tidigare modelleringar av marginalkostnaden i järnvägsinfrastrukturen, den rättsliga bakgrunden till banavgiftens grunddel och banavgiftspraxis i tre medlemsstater i Europeiska ekonomiska samarbetsområdet. I det empiriska avsnittet i arbetet undersöktes logaritm-, Box-Cox- och Box-Tidwell-modifierade kostnadsfunktioners lämplighet för ekonometrisk modellering av marginalkostnaderna i Finlands järnvägsinfrastruktur. Den logaritmmodifierade kostnadsfunktionen valdes för vidarebearbetning.

I litteraturoversikten konstaterades det att betoningen i de rättsliga begränsningarna i fastställandet av banavgiftens grunddel ligger på de kostnader som ska inkluderas i beräkningen, och inte på beräkningsmetoderna för banavgiftens grunddel. Trots den ringa regleringen kring beräkningsmetoder ligger betoningen i metoderna för modelleringar av marginalkostnader i järnvägsinfrastrukturen i Europa på endast ett fåtal olika kostnadsfunktionsformer och de skiljer sig från varandra främst i fråga om tekniska variabler. Det viktigaste resultatet av arbetet var observationen om att den ekonometriska modelleringen av banavgiftens grunddel kan utföras med logaritmmodifierad tvärsnittsdata, i vilken en tidsserie med för flera års paneldata kring variablerna räknades ihop.

Den marginalkostnadsmodellering som gjorts för logaritmmodifierat tvärsnittsdata, som utvecklats i det empiriska avsnittet i arbetet, motsvarade alla antaganden som fastställdes för den linjära regressionsmodellen, och den marginalkostnad som togs fram i modelleringen överensstämmer med de marginalkostnader som tidigare fastställdes i Finland och som finns i banavgiftsprislistan. Modelleringsprincipen användes vid den ekonometriska beräkningen av banavgiftens grunddel för år 2022 på hösten 2020.

**Jarno Kajanne: Evaluating the econometric modeling of the marginal cost of railway infrastructure use.** Finnish Transport Infrastructure Agency. Helsinki 2020. Thesis 8/2020. 62 pages and 3 appendices. ISSN 2490-1202, ISBN 978-952-317-829-8.

## Abstract

The charging principles of the use of railway infrastructure are based on the basic infrastructure charge, which is the only charge currently implemented in Finland. The basic infrastructure charge is set using a cost function, which is estimated using econometric methodology. The issues of the econometric model, for example, heteroscedasticity, were the starting point of this thesis. The main objective of this thesis was to evaluate the suitability of different cost functions for estimating the marginal cost of the railway infrastructure in Finland.

The research started with a literature review of existing railway infrastructure marginal cost modeling and with a review of the legislative side of railway infrastructure charges. The econometric modeling practices of the marginal costs of railway infrastructure in three European Economic Area members were also presented in the literature review. In the empirical part of the study, logarithmic, Box-Cox, and Box-Tidwell transformed cost functions were applied to the Finnish railway infrastructure data. The logarithmic cost function was chosen for further study.

The literature review showed that the legislative side of the basic railway infrastructure charge focuses on the costs that can be included in the calculation giving very little attention to the methodology of the econometric modeling. Although there are almost no restrictions for the econometric methodology, the international railway infrastructure marginal cost research focuses on only a few different forms of cost functions and differ mainly in terms of the technical variables included in the modeling. The key result of the empirical research was that the modeling of the marginal cost of railway infrastructure use in Finland can be successfully modeled with cross-sectional data, in which the panel data time series have been added up.

All of the linear regression assumptions held true for the logarithmic cross-sectional data model. The marginal costs calculated by the model are similar to previous basic infrastructure charges in Finland. The methodology presented in this study was used in the econometric modeling of the Finnish basic infrastructure charge in 2022 timetable.

## Esipuhe

Tämän diplomityön on tehnyt Jarno Kajanne Lappeenrannan-Lahden teknillinen yliopiston LUT, School of Engineering Science, tuotantotalouden koulutusohjelmassa.

Työn tilasi Väylävirasto ja tarkastajina toimivat professori Pasi Luukka ja KTT Jyrki Savolainen. Väyläviraston yhteyshenkilöinä toimivat Joni Tefke, Harri Lahelma ja Hanna Askola.

Helsingissä, joulukuussa 2020

Väylävirasto  
Väylien käyttöpalvelut



## Sisällysluettelo

1	JOHDANTO .....	9
1.1	Aiheeseen liittyvää kirjallisuutta.....	9
1.2	Työn tavoitteet ja rajaukset .....	14
1.3	Työn rakenne .....	15
2	RATAMAKSUN MÄÄRITTELY JA SEN LAINSÄÄDÄNNÖLLINEN TAUSTA.....	16
2.1	Ratamaksun perusteet .....	16
2.2	Ratamaksun perusmaksun määrittämisen sääntely .....	17
2.3	Ratamaksun perusmaksun laskemistavat.....	18
2.3.1	Tekninen mallinnus .....	19
2.3.2	Ekonometrinen mallinnus .....	20
3	RAUTATIEINFRASTRUKTUURIN KUSTANNUSFUNKTION MALLINNUKSEN TEORIAA.....	22
3.1	Lineaarinen regressio.....	24
3.2	Lineaarisen regression oletukset.....	25
3.3	Regressiomallinnuksen selitysvoima .....	27
4	VALITTUJEN EUROOPAN MAIDEN RATAMAKSUKÄYTÄNTÖJÄ.....	28
4.1	Norjan ratamaksun perusmaksu.....	28
4.1.1	Norjan ratamaksun perusmaksuun sisällytettävät kustannukset .....	29
4.1.2	Norjan kustannusfunktion ekonometrinen mallinnus.....	31
4.2	Ranskan ratamaksun perusmaksu .....	32
4.2.1	Ranskan ratamaksun perusmaksuun sisällytettävät kustannukset .....	33
4.2.2	Ranskan kustannusfunktion ekonometrinen mallinnus .....	33
4.3	Suomen ratamaksun perusmaksu .....	36
4.3.1	Suomen ratamaksun perusmaksuun sisällytettävät kustannukset .....	36
4.3.2	Suomen kustannusfunktion ekonometrinen mallinnus.....	37
5	SUOMEN RAUTATIEINFRASTRUKTUURIN KÄYTÖN RAJAKUSTANNUSTEN MALLINNUS .....	39
5.1	Laskenta-aineisto.....	39
5.2	Mallinnuksen menetelmät .....	41
5.3	Mallinnuksen jatkojalostaminen .....	47
5.4	Mallinnusten tulosten yhteenveto .....	52
5.5	Rajakustannus ja kustannusjousto .....	53
6	JOHTOPÄÄTÖKSET .....	55
7	YHTEENVETO .....	58
	LÄHTEET .....	59
	LIITTEET	
Liite 1	Koko verkon laajuiset suorat kustannukset	
Liite 2	Muut kuin huomioon otettavat kustannukset	
Liite 3	Mallinnusten yhteenvedot	

## Kuvaluettelo

Kuva 1.	Rajakustannusten määrittäminen ekonometrisella menetelmällä.....	20
Kuva 2.	Ylläpitorakenne.....	29
Kuva 3.	Suomen rataverkko.....	40
Kuva 4.	Mallin 1 visualisointi .....	43
Kuva 5.	Mallin 2 visualisointi .....	44
Kuva 6.	Mallin 3 visualisointi .....	46
Kuva 7.	Mallin 1a visualisointi.....	48
Kuva 8.	Mallin 1b visualisointi.....	50
Kuva 9.	Mallin 1c visualisointi .....	51

## Taulukkoluetelo

Taulukko 1.	Rajakustannusmallinnukset ylläpitokustannuksilla .....	10
Taulukko 2.	Rajakustannusmallinnukset ylläpitokustannuksilla jatkuu.....	11
Taulukko 3.	Rajakustannusmallinnukset ylläpito- ja korvausinvestointikustannuksilla .....	12
Taulukko 4.	Kustannuksia aiheuttavien tekijöiden vaikutukset kustannustyypeihin.....	30
Taulukko 5.	Norjan regressiomallien parametrit .....	31
Taulukko 6.	Parhaiten suoriutuneiden mallien selitysasteet.....	32
Taulukko 7.	Ranskan ekonometristen mallien muuttujat .....	35
Taulukko 8.	Mallinnusten selitysasteet ja kustannusjoustot .....	35
Taulukko 9.	Perusmaksun laskentaan sisältyvät kustannukset.....	37
Taulukko 10.	Muuttujien tilastolliset tunnusluvut.....	41
Taulukko 11.	Rautatieinfrastruktuurin rajakustannusten mallinnukset .....	42
Taulukko 12.	Mallin 1 testien tulokset .....	44
Taulukko 13.	Mallin 2 testien tulokset .....	45
Taulukko 14.	Mallin 3 testien tulokset .....	46
Taulukko 15.	Poikkileikkausaineiston tilastolliset tunnusluvut .....	47
Taulukko 16.	Mallin 1a testien tulokset.....	49
Taulukko 17.	Mallin 1b testien tulokset .....	50
Taulukko 18.	Mallin 1c testien tulokset.....	52
Taulukko 19.	Mallinnusten oletusten yhteenveto.....	52
Taulukko 20.	Poikkileikkausaineistomallinnuksen kustannusjousto ja rajakustannukset.....	54

# 1 Johdanto

Valtion rataverkon kunnossapito ja kehittäminen sekä rautatieinfrastruktuurin käytön hinnoittelu ovat Suomessa Väyläviraston vastuulla. Rautatieinfrastruktuurin käytön hinnoittelu perustuu ratamaksuun. Ratamaksu koostuu perusmaksusta sekä perusmaksun alennuksista ja korotuksista, lisämaksusta ja investointimaksusta. Rautatieinfrastruktuurin hinnoittelu on säänneltyä Suomen lainsäädännön sekä Euroopan unionin direktiivien ja täytäntöönpanoasetusten myötä. Ratamaksun perusmaksun tulee perustua suoraan rautatieliikenteestä aiheutuviin kustannuksiin sekä sen määrittävät ovat säänneltyjä ja niiden on oltava dokumentoituja.

Suomen tämänhetkinen ratamaksu koostuu ainoastaan perusmaksusta. Perusmaksun perusosa on ekonometrisesti regressioanalyysillä mallinnetun radanpidon kustannusfunktion rajakustannus. Itsenäisinä muuttujina perusmaksun mallinnuksessa ovat rataosajaon mukaisen rataosan liikennemäärä bruttotonneina ja rataosan pituus. Riippuvana muuttujana toimii mallinnuksessa radanpidon kustannukset. Ratamaksun perusmaksu on rajakustannus, joka saadaan kustannusfunktioista osittaisderivoimalla se liikennemäärän suhteen. Nykyään käytössä oleva perusmaksun ekonometrinen mallinnus perustuu Tiina Idströmin vuonna 2002 julkaisemaan Pro Gradu -tutkielmaan 'Suomen ratamaksun uudistaminen – ekonometrinen analyysi rataverkon käytön rajakustannuksista'. Idströmin (2002) esittelemä mallinnus hyödyntää Cobb-Douglas kustannusfunktioita logaritmisessa muodossa.

Ratamaksun perusmaksun tarkastelu on juuri nyt ajankohtaista osana Väyläviraston rautateiden hinnoitteluprojektia. Perusmaksun vuosien 2019 ja 2020 määrittämistä ja tasoa on kyseenalaistettu VR:n ja hinnoittelua valvovan Rautateiden sääntelyelimen toimesta. Sääntelyelin on esittänyt ongelmakohtaksi ekonometrisen mallinnuksen heteroskedastisuuden, joka tulisi korjata mahdollisesti mallinnustapaa muuttamalla. Perusmaksun ja sen määrittelyn ongelmakohtia ovat myös kustannusseurannan ja raportoinnin puutteellisuus sekä joidenkin mallinnukseen sisällytettävien kustannusten erittely. (Karjanlahti & Kukkarinen 2019, 21)

## 1.1 Aiheeseen liittyvää kirjallisuutta

Rautatieliikenteen infrastruktuurille aiheuttamien rajakustannusten estimointi on ollut ekonometrisen tutkimuksen kohteena jo 1960-luvulta lähtien (Borts 1960). Tässä alaluvussa tarkastellaan aiheeseen liittyvässä kirjallisuudessa hyödynnettyjä metodeja ja tärkeimpiä tuloksia. Kirjallisuuskatsaus perustuu 2000-luvulla tehtyihin tutkimuksiin, joissa käsitellään rautatieinfrastruktuurin käytön rajakustannusten ekonometristä mallinnusta. Nämä artikkelit on koostettu hyödyntäen LUT-tiedekirjaston tietokantoja, Googlen hakupalveluja ja Euroopan talousalueen jäsenmaiden rautateiden verkkoselostusten liitteistä löytyviä menetelmämuistioita.

Lähes kaikki ratainfrastruktuurin käytön rajakustannusten tutkimukset on toteutettu maakohtaisesti kyseisen maan rautatieinfrastruktuurin dataa hyödyntäen. Ainoa poikkeus tähän on Wheatin ja Smithin (2008) tutkimus, jossa toteutettiin mallinnuksia kansainvälistä dataa hyödyntäen. Rajakustannusten ja kustannusjoukon määrittelyyn lähestymistapana toimii rautatieinfrastruktuurin

kustannusfunktion estimointi ekonometristen mallien avulla. Rajakustannukset ja kustannusjoustot johdetaan malleista liikenteen tai muiden itsenäisten muuttujien suhteen. Kustannusjousto saadaan osittaisderivoimalla kustannusfunktio kiinnostuksen kohteena olevan muuttujan suhteen ja tästä voidaan edelleen laskea rajakustannus kertomalla kustannusjousto keskimääräisellä kustannuksella. Wheatin ja muiden (2009) mukaan aiheeseen liittyvän kirjallisuuden tuottamista kustannusjoustoista löytyy suurta hajontaa. Tulosten erilaisuudet tutkimusten välillä ovat lähtöisin datan laadun ja sen luokittelun eroavaisuuksista sekä valitusta ekonometrisesta metodologiasta.

Yleisesti ekonometrisessa mallinnuksessa kustannusdatana käytetään rautatieinfrastruktuurin ylläpidosta aiheutuvia kustannuksia. Näitä kustannuksia ovat kaikki kustannukset, jotka ovat lähtöisin rautatieinfrastruktuurin käytettävyyden turvaamisesta. Kustannukset ovat yleisesti huomioitu yhteenlasketuina, eikä niitä ole mallinnuksessa jaoteltu kustannuserittäin. (Silavong et al. 2014, 1) Poikkeuksena Silavong ja muut (2014) toteuttivat erillisen ekonometrisen mallinnuksen ylläpitokustannuksille, vaihteiden ja tasoristeysten kustannuksille, merkinantojärjestelmän kustannuksille, ajojohtojärjestelmän kustannuksille ja liikenteenohjauksen kustannuksille. Tämä tutkimus on esitetty tarkemmin kappaleessa 4.2., jossa käydään Ranskan ratamaksun perusmaksun mallinnustapa läpi.

Johanssonin ja Nilssonin (2004) tekemä rautatieinfrastruktuurin käytön rajakustannusmallinnus on toiminut perustana sittemmin yleistyneille rautatieinfrastruktuurin käytön rajakustannustutkimuksille. Johansson ja Nilsson hyödynsivät alun perin vuonna 2002 julkaistussa paperissaan Christensenin ja muiden (1972) esittelemää Translog-kustannusfunktion käyttöä ylläpitokustannusfunktion estimoinnissa. Riippuvana muuttujana mallinnuksessa on infrastruktuurin kustannukset ja itsenäisinä muuttujina liikennöinti sekä infrastruktuurin ominaisuudet. Tämä mallinnustapa on sittemmin vakiintunut (Munduch et al. 2002; Daljord, 2003; Tervonen ja Idström 2004; Andersson 2006; Marti et al. 2006; Smith et al. 2008; Gaudry et al. 2009; Andersson 2009; Link 2009; Marti et al. 2009; Wheat ja Smith 2009; Andersson et al. 2012; Silavong et al. 2014; Andersson et al. 2016; Odolinski et al. 2018; Sternad et al. 2018). Taulukkoon 1 on koostettu rautatieliikenteen pelkkien ylläpitokustannusten rajakustannusten mallinnoon liittyvät tutkimukset. Tutkimuksissa kustannusdataan ei ole siis sisällytetty korvausinvestointikustannuksia.

*Taulukko 1. Rajakustannusmallinnukset ylläpitokustannuksilla*

<b>Tutkimus</b>	<b>Maa</b>	<b>Funktiomuoto</b>	<b>Kustannusjousto</b>
Johansson ja Nilsson (2004)	Ruotsi	Translog	0,169
Johansson ja Nilsson (2004)	Suomi	Translog	0,167
Munduch et al. (2002)	Itävalta	Log-lineaarinen Cobb-Douglas vuorovaikutustermillä	0,270
Gaudry ja Quinet (2003)	Ranska	Rajoittamaton yleistetty Box-Cox	0,37
Tervonen ja Idström (2004)	Suomi	Log-lineaarinen Cobb-Douglas	0,133 – 0,175

Tutkimus	Maa	Funktiomuoto	Kustannusjousto
Marti ja Neu-enschwander (2006)	Sveitsi	Log-lineaarinen Cobb-Douglas	0,200 (malli 1) 0,285 (malli 2)
Andersson (2006)	Ruotsi	Translog sisältäen satunnaisvaikutukset	0,204

Taulukossa 2 on jatkoa taulukon 1 ylläpitokustannusten rajakustannusten mallinnuksille.

*Taulukko 2. Rajakustannusmallinnukset ylläpitokustannuksilla jatkuu*

Tutkimus	Maa	Funktiomuoto	Kustannusjousto
Wheat ja Smith (2008)	Iso-Britannia	Logaritminen neliö- ja kuutiotermeillä	0,239 0,378
Wheat ja Smith (2008)	Kansainvälinen	Translog	0,33
Andersson (2009)	Ruotsi	Box-Cox	0,23 0,199 (henkilö) 0,058 (rahti)
Link (2009)	Itävalta	Box-Cox	0,40
Marti et al. (2009)	Sveitsi	Box-Cox	0,23 0,127 (henkilö) 0,058 (rahti)
Silavong et al. (2014)	Ranska	Box-Cox	0,57 (henkilö) 0,17 (rahti)
Sternad et al. (2018)	Slovenia	Log-lineaarinen Box-Cox	0,18

Taulukoista 1 ja 2 nähdään, että rautatieliikenteen infrastruktuurille aiheuttamien rajakustannuksen mallinnuksissa on ollut käytössä neljä eri funktiomallinusta. Ne ovat Translog-kustannusfunktio, log-lineaarinen Cobb-Douglas-kustannusfunktio, Box-Cox-kustannusfunktio ja logaritminen funktiomuoto. Tutkimusten ollessa kronologisessa järjestyksessä nähdään, että ajan saatossa on siirrytty suosimaan Box-Cox-kustannusfunktioita. Jokaisessa tutkimuksessa on jätetty tuotantopanosten hintataso mallinnuksen ulkopuolelle puutteellisen datan tai hintojen merkityksettömyyden vuoksi. Kustannusdatan ja ekonometristen mallien muuttujien, mallin määritysten ja lopullisen mallin valintamenettelyn eroavaisuuksien takia tutkimusten tulosten vertailu on haastavaa (Silavong et al. 2014, 2). Tästä seuraten Wheat ja muut (2009, 30) suosittavat vertailuarvoina käytettävän kustannusjoustoja. Kustannusjousto kuvaa prosenttiosuutta, jolla kustannukset kasvavat liikenteen lisääntyessä yhdellä prosentilla. Kustannusjouston etuna on se, että niiden arvot eivät ole riippuvaisia toiminnan suuruusluokasta, kun taas rajakustannukset ovat. Taulukoista 1 ja 2 huomataan, että tutkimusten kustannusjoustot asettuvat pääosin välille 0,1–0,5. Lisäksi huomioitavaa on, että henkilöliikenteen kustannusjousto on yli kaksinkertainen kaikissa tutkimuksissa, joissa data on jaoteltu henkilö- ja rahtiliikenteen välille.

Taulukko 3. Rajakustannusmallinnukset ylläpito- ja korvausinvestointi-kustannuksilla

Tutkimus	Maa	Funktiomuoto	Kustannusjousto
Tervonen ja Idström (2004)	Suomi	Log-lineaarinen Cobb-Douglas	0,267 – 0,291
Marti ja Neuenchwander (2006)	Sveitsi	Log-lineaarinen Cobb-Douglas	0,265
Andersson (2006)	Ruotsi	Translog sisältäen satunnaisvaikutukset	0,302
Tervonen ja Pekkarinen (2007)	Suomi	Log-lineaarinen Cobb-Douglas ja Translog, molemmat satunnaisvaikutuksilla	0,18 (Cobb-Douglas) 0,11 – 0,13 (Translog)
Wheat ja Smith (2008)	Iso-Britannia	Log-lineaarinen Cobb-Douglas	0,48

Taulukkoon 3 on koostettu artikkelit, joissa tarkastellaan korvausinvestointi- ja ylläpitokustannuksia yhdessä. Lähestymistapana kaikissa on kustannuserien yhteenlaskeminen. Translog-kustannusfunktiota on käytetty yhteiskustannusten mallinnukseen kahdessa artikkelissa (Andersson 2006; Tervonen ja Pekkarinen 2007). Kaikki muut artikkelit suosivat Cobb-Douglas-kustannusfunktiota. Taulukosta havaitaan, että kustannusjoustot asettuvat keskimäärin välille 0,1–0,5. Väli on sama kuin mallinuksissa, joissa oli mallinnettu ainoastaan radanpidon ylläpitokustannuksia.

Ongelmallinen kustannuserä on yleisesti kirjallisuudessa ollut korvausinvestointikustannukset (Silavong et al. 2014, 1). Korvausinvestoinnit ovat kausittaisia, eli niitä ei esiinny joka rataosalla joka vuosi. Korvausinvestointien huomioiminen jossain muodossa on kuitenkin tärkeää, sillä ratamaksun on tarkoitus perustua kaikkiin liikennöinnistä aiheutuviin suoriin kustannuksiin. Niiden luonteen vuoksi korvausinvestointikustannusten huomioimisesta löytyy kirjallisuudesta eroavaisuuksia: osa jättää kyseiset kustannukset kokonaan huomioimatta ja osa laskee ne mukaan ylläpitokustannusdataan. Uusimmissa tutkimuksissa on tutkittu myös korvausinvestointikustannusten erilaisia mallintamismenetelmiä (Wheat et al. 2009, 35-36).

Pelkkien korvausinvestointikustannusten rajakustannusten mallinnukseen on käytetty nurkkaratkaisuja (Andersson et al. 2012) ja elinaika-analyysiä (Andersson et al. 2016). Nurkkaratkaisumallilla tarkoitetaan jossain määrin sensuroitua regressiomallia. Nurkkaratkaisuilla viitataan tilanteeseen, missä esimerkiksi kustannus määräytyy jonkun ekonomisen päättäjän, kuten yksilön tai yrityksen, päätöksen perusteella. Se saa arvon nolla positiivisella todennäköisyydellä ja on jatkuva satunnainen muuttuja saaden vain positiivisia arvoja. Esimerkkeinä nurkkaratkaisusta toimivat yrityksen tutkimuksen ja kehityksen rahoituksen määrä sekä yksilön henkivakuutukseen valitsema summa. (Woolridge 2002, 518) Andersson ja muut (2012) toteuttivat nurkkaratkaisumallinnuksen kolmen eri regressiomallin avulla: Tobit-mallin, kaksiosaisen mallin ja Heckit-mallin avulla. Kaikki mallit muokattiin sopiviksi käytössä olleelle datalle, jonka riippuva muuttuja saa paljon todellisia nolla-arvoja (Andersson et al. 2012, 17) Elinaika-analyysi perustuu korvausinvestointikustannusten mallinuksessa korvausinvestointien välisen ajan todennäköisyysjakauman mallintamiseen rataosittain. Mallinnus toteutettiin Weibull-mallilla perustuen radan laadun huonontumisen

joustokertoimeen, joka kuvaa korvausinvestointitarpeiden välisen ajan keston prosentuaalista muutosta suhteessa prosentin muutokseen liikennöinnissä. (Andersson et al. 2016, 76)

Odolinski ja muut (2018a) toteuttivat kahden edellä mainitun korvausinvestointien huomioimiseen käytetyn menetelmän vertailun 18 vuoden datalla Ruotsin rautatieverkon infrastruktuurista ja sen käytön kustannuksista. Huomioitava lopputulos oli, että nurkkaratkaisujen avulla määritetty bruttotonnikohtainen rajakustannus on yli kaksinkertainen verrattuna elinaika-analyysin vastaavaan. Tämä on seurausta siitä, että nurkkaratkaisumallit ottavat huomioon investointien kustannusten systemaattiset vaihtelut, jotka selittyvät osittain liikenteellä, kun taas elinaika-analyysi ei. Lisäksi kyseiset menetelmät vaativat eri tasoilla eroteltua dataa. Elinaika-analyysin toteuttamiseen tarvitaan tarkkaa dataa korvausinvestointien elinkaarista, mutta kustannusdatan ei tarvitse olla samalla tarkkuudella. Nurkkaratkaisumalli taas vaatii kustannusdatan erottelun, mutta investointien elinkaaritietoja ei tarvita. Lopputulemana on se, että nurkkaratkaisumallit ovat vähemmän rajoittavia infrastruktuurin haltijan päätösten seurausten selittämisessä verrattuna elinaika-analyysimalleihin. Tämän seurauksena nurkkaratkaisumalleja suositellaan käytettäväksi korvausinvestointien mallintamisessa. Mikäli data ei ole eroteltavissa nurkkaratkaisumallien vaativalle tasolle ja elinaika-analyysit ovat ainoa mahdollinen mallinnustapa, infrastruktuurin haltijan tulee ottaa huomioon se, että mallinnustapa vähättelee liikenteen vaikutusta kustannuksiin. (Odolinski et al. 2018a, 17)

Viimeisimmässä korvausinvestointikustannuksiin liittyvässä tutkimuksessa Odolinski ja Wheat (2018b) mallintavat korvausinvestointikustannukset yhdessä radanpitokustannusten kanssa vektoriautoregressiivisen (VAR) mallin avulla. Vektoriautoregressiivinen malli on moniyhtälöaikasarjamalli, jossa itsestä muuttujaa selittää oman ja muiden itsenäisten muuttujien viiveet. VAR-metodologiaa sovelletaan yleisesti muun muassa valuuttakurssien ja talouspolitiikan analysointiin ja laajamittaiseen ennustamiseen. (Suomen Pankki 1989, 13) Mallinnuksessa käytetään logaritmista funktiomuotoa aiempien rautatieinfrastruktuurin rajakustannusten tutkimusten suositusten mukaisesti (Odolinski et al. 2018b, 23). Odolinskin ja Wheatin tutkimus on ensimmäinen, jossa hyödynnetään kokonaisvaltaista dynaamista mallinnusta. Olennainen huomio tuloksista on se, että kun huomioidaan kustannusten samanaikainen vaihtelu ja tulevien kustannusten vaihtelut liikenteen suhteen, kustannusjousto ja liikenteestä aiheutuva rajakustannus ovat huomattavasti korkeampia kuin perinteisillä staattisilla malleilla mallinnettuna. Tulokset ovat lupaavia ja tekijät esittävät dynaamisten mallien olevan tärkeitä tulevaisuuden tutkimusten kohteita. (Odolinski et al. 2018b, 28)

Yhteenvedon kirjallisuuden tarkastelusta voidaan todeta, että tutkimusten painopiste on vuosien varrella siirtynyt lyhyen aikavälin rajakustannusten ja kustannusjouston määrittämiseen, kun taas varhaiset tutkimukset painottuivat rautatieoikeellisuuden tehottomuuden tarkasteluun. Rautatieinfrastruktuurin käytön rajakustannusten mallinnusta hallitsee logaritmuunnellut Box-Cox-, Cobb-Douglas- ja Translog-kustannusfunktiot. Ajan saatossa tutkimuksissa ollaan siirrytty pääosin suosimaan Box-Cox-muunnosta, joka sisällyttää logaritmuunnokset erityistapauksissa. Viime aikoina tutkimuskohteeksi on keskitynyt korvausinvestointien kustannusten sisällyttäminen mallinnuksiin niiden ongelmallisen luonteen vuoksi. Korvausinvestointien ongelmallisten kustannusten mallinnukseen on esitetty ratkaisuksi nurkkaratkaisuja, elinaika-analyysiä ja viimeisimpänä vektoriautoregressiivistä mallinnusta. Tässä alaluvussa

esiteltyjen kustannusfunktioiden, mallinnusten ja muunnosparametrien matemaattiset taustat on avattu tarkemmin tämän työn luvussa 3.

## 1.2 Työn tavoitteet ja rajaukset

Työn tavoitteena on ensin tarkastella rautatieinfrastruktuurin käytön rajakustannusten määrittelyn lainsäädännöllinen tausta sekä siihen liittyvät asetukset ja rajoitukset. Lainsäädännöllisten rajoitusten esittelyn jälkeen työssä on tavoitteena tarkastella Euroopan talousalueen jäsenmaiden rautatieinfrastruktuurin käytön rajakustannusten ekonometrisen mallinnuksen menetelmiä. Työssä esitellään Norjan, Suomen ja Ranskan rautatieinfrastruktuurin käytön rajakustannusten ekonometrinen mallinnustapa.

Työn päätavoitteena tavoitteena on tarkastella erilaisten kustannusfunktioiden ja menetelmien sopivuutta rautatieinfrastruktuurin käytöstä aiheutuvien rajakustannusten määrittämisessä Suomen tapauksessa. Tarkastelu alkaa Suomessa jo käytössä olevan ja kirjallisuudessakin yleisesti hyödynnetyn logaritmi-muunnellun kustannusfunktion estimoimisella ja testaamisella. Tämän lisäksi muita työssä tarkasteltavia kustannusfunktio-tyyppejä ovat Box-Cox-muunneltu kustannusfunktio sekä Box-Tidwell-muunneltu kustannusfunktio, jotka ovat myös laajalti käytettyjä aiheeseen liittyvässä kirjallisuudessa. Työn mallinnukset toteutetaan R-ohjelmointikielellä.

Aiemmin Suomessa käytössä olleessa mallinnustavassa on ollut huomattavana ongelmana mallinnuksen heteroskedastisuus sekä jäännöstermien normaalijakautumattomuus. Tässä työssä mallinnusten tavoitteena on siis löytää toimiva mallinnustapa, joka hyödyntää pienimmän neliösumman estimointimenetelmää sekä täyttää kaikki menetelmälle asetetut oletukset. Eri mallinnukset arvostetaan seuraavin perustein:

- Mallinnuksen tilastollinen selitysaste
- Mallille asetettujen oletusten täytyminen
- Tulosten kohtuullisuus

Mallinnuksen tilastollisella selitysasteella viitataan  $R^2$ -lukuun. Luku kertoo, kuinka suuren osuuden itsenäiset muuttujat pystyvät selittämään riippuvan muuttujan vaihtelusta. Mallille asetetut oletukset määräytyvät lineaarisen regressiomallinnuksen standardioletusten mukaisesti. Nämä kyseiset oletukset ja  $R^2$ -luku esitellään tarkemmin kappaleessa 3.2. Kansainvälisiä tutkimustuloksia ja metodeja tarkasteltiin kirjallisuuskatsauksessa ja tulosten kohtuullisuudella viitataan siihen, että kustannusjousto ja selittävyysasteet ovat suurusluokaltaan kansainvälistä tasoa.

Työn tarkastelu keskitetään vain itse ekonometrisen mallinnuksen tarkasteluun eli muut ratamaksun määrittämiseen liittyvät kehittämistarpeet jätetään tarkastelun ulkopuolelle. Ulkopuolelle jääviä asioita on siis muun muassa ratamaksun perusmaksun laskentaan sisällytettävien kustannusten kerääminen. Ratamaksun laskentaan sisällytettävät kustannukset ovat kuitenkin esitelty, mutta niitä tai niiden suodatusprosessia ei ole tarkemmin esitelty.



Edellä mainittujen työn tavoitteiden pohjalta on kehitelty seuraavat kolme tutkimuskysymystä:

1. Mitkä ovat ratamaksun perusmaksun määrittämisen lainsäädännölliset rajoitukset?
2. Millä tavoilla Euroopan talousalueen jäsenmaissa on mallinnettu ekonometrisesti rautatieinfrastruktuurin käytön rajakustannuksia?
3. Mikä on sopiva mallinnustapa Suomen ratamaksun perusmaksulle, eli rautatieinfrastruktuurin liikenteestä aiheutuville rajakustannuksille?

Ensimmäisellä tutkimuskysymyksellä haetaan vastausta ratamaksun perusmaksun, eli tässä työssä liikenteestä aiheutuvan rajakustannuksen lainsäädännöllisiin rajoituksiin. Vastausta haetaan lainsäädännöstä ja asetuksista, jotka koskevat ratamaksun perusmaksun määrittelyä. Toisella tutkimuskysymyksellä etsitään vastausta Euroopan talousalueen jäsenmaiden rautatieinfrastruktuurin käytön rajakustannusten ekonometristen mallinnusten käytäntöihin. Kolmas tutkimuskysymys käsittelee Suomen rautatieinfrastruktuurin liikenteestä aiheutuvien rajakustannusten mallinnusta.

## 1.3 Työn rakenne

Johdannon jälkeisessä luvussa 2 käsitellään ratamaksun perusmaksun lainsäädännöllisiä rajoituksia. Luvussa käydään läpi asiaa koskevia lainsäädäntö sekä Euroopan komission asetuksia. Lisäksi luvussa käydään läpi eurooppalaisen rautatiealan itsenäisten sääntelyviranomaisten ryhmän, IRG-Railin ratamaksun ekonometriselle mallinnukselle suosittelemia toimintatapoja.

Kolmannessa luvussa siirrytään ekonometristen mallinnusten teoriaan. Luku käsittelee tässä työssä käytettävien ja esiteltävien kustannusfunktioiden muotoa sekä erinäisiä funktioille toteutettavia muunnoksia. Luvussa käydään läpi myös työssä käytettävän tilastollisen analyysimenetelmän, eli lineaarisen regression, teoriaa. Teorian lisäksi luvussa esitetään lineaariselle regressiolle asetetut standardioletukset, joiden täytyminen on yksi tämän työn mallinnusten päämääristä.

Neljäs luku käsittelee Euroopan talousalueen jäsenmaiden ratamaksun perusmaksun ekonometrisen mallinnuksen toteutustapoja. Luvussa käydään läpi kolmen maan, Norjan, Ranskan ja Suomen, menetelmät, joilla ratamaksun perusmaksu on mallinnettu. Läpikäynti sisältää suhteellisen yksityiskohtaisen esittelyn käytössä olevista ekonometrisista malleista sekä erittelyn mallinnuksiin sisällytettävistä kustannuksista.

Viidennessä luvussa siirrytään itse rautatieinfrastruktuurin liikenteestä aiheutuvien rajakustannusten mallinnukseen. Luvussa esitellään käytössä oleva kustannus- ja liikennöintiaineisto, joilla ekonometrinen rajakustannusmallinnus toteutetaan. Lisäksi luvussa käydään läpi mallinnuksen menetelmät ja tulokset. Lopulta luvussa jatkojalostetaan aluksi rakennettuja mallinnuksia ja käydään tulokset läpi.

Työn kuudennessa luvussa vastataan tutkimuskysymyksiin ja esitetään työn tärkeimmät johtopäätökset ja pohdintaa. Lopulta viimeisessä seitsemännessä luvussa esitetään työstä yhteenveto.

## 2 Ratamaksun määrittely ja sen lainsäädännöllinen tausta

Ratamaksulla tarkoitetaan valtion rataverkon käytöstä perittyä maksua. Rautatiepalvelujen käytöstä perittävien maksujen asettamisesta ja hinnoittelusta säädetään Euroopan parlamentin ja neuvoston direktiivissä 2012/34/EU yhteisestä eurooppalaisesta rautatiealueesta. Maksujen asettamisesta ja hinnoittelusta säädetään direktiivin 7 artiklassa ja edelleen hinnoittelun reunaehdoista artikloissa 29-37. Rataverkon haltija voi periä vähimmäiskäyttöpalveluista perittäviä maksuja ohjaavine porrastuksineen, kapasiteetin varausmaksuja rataverkon ja rautatiepalvelujen käytöstä, lisämaksuja erikseen määritellyiltä liikenteen segmenteiltä sekä oheis- ja lisäpalveluista perittäviä maksuja. (Tervonen 2016, 12)

Ratamaksun perusmaksun suuruutta ja perusteita säädellään Euroopan parlamentin ja neuvoston direktiivissä 2012/34/EU yhtenäisestä eurooppalaisesta rautatiealueesta, Euroopan komission täytäntöönpanoasetuksessa 2015/909 rautatieliikenteen harjoittamisesta suoraan aiheutuvien kustannusten yksityiskohtaisista laskentasäännöistä sekä Suomen raideliikennelaissa (1302/2018). (Karjanlahti & Kukkarinen 2019, 3)

### 2.1 Ratamaksun perusteet

Suomen raideliikennelain (1302/2018) 139 §:ssä säädetään ratamaksun koostuvan perusmaksusta ja muista ratamaksuun liittyvistä erikseen säädellyistä maksuista. Muita erikseen säädetyjä maksuja ovat perusmaksun alennukset ja korotukset, lisämaksu ja investointimaksu. Ratamaksua peritään rataverkon ylläpidosta aiheutuvien kustannusten rahoittamiseen ja rautatieliikenteen harjoittaja maksaa sen rataverkon haltijalle, jonka rataverkkoa tai palveluja on käytetty.

Ratamaksun perusmaksun tulee perustua suoraan rautatieliikenteen harjoittamisesta aiheutuviin kustannuksiin ja sen perusteella tulee tarjota vähimmäiskäyttömahdollisuuksiin sisältyvät palvelut, jotka ovat säädetty raideliikennelain 132 §:n 1 momentissa. (Raideliikennelaki 2018) Raideliikennelain määritelmän mukaan vähimmäiskäyttömahdollisuuksien palvelut ovat säädetty rautatiemarkkinadirektiivin (2012/34/EU) liitteen II kohdassa 1, jonka mukaan vähimmäiskäyttömahdollisuuksiin sisältyvät:

- a) rautateiden ratakapasiteettia koskevien hakemusten käsittely,
- b) oikeus käyttää myönnettyä kapasiteettia,
- c) rautatieinfrastruktuurin käyttö, mukaan lukien verkon risteysasemat ja vaihteet,
- d) junaliikenteen ohjaus, mukaan lukien merkinanto, liikenteen ohjaus, lähettäminen ja viestinanto sekä tietojen antaminen junien liikkumisesta,
- e) sähkönsiirtolaitteiden käyttö kuljetusvirran saamiseksi,
- f) kaikki muut tiedot, joita tarvitaan sellaisen liikenteen harjoittamiseen, jota varten kapasiteetti on myönnetty. (Rautatiemarkkinadirektiivi 2012/34/EU)

Edellä mainittujen vähimmäiskäyttömahdollisuuksien lisäksi ratamaksun perusmaksun vastikkeena tulee turvata myös pääsy seuraaviin palvelupaikkoihin:

- a) matkustaja-asetat, niihin kuuluvat rakennukset ja muut tilat, mukaan lukien matkatietopalvelut ja soveltuvat tilat lipunmyyntipalveluille
- b) tavaraliikenneterminaalit,
- c) järjestelyratapihat ja junanmuodostuslaitteet, mukaan lukien vaihtotyö-laitteet,
- d) varikkosivuraiteet,
- e) huoltotilat, lukuun ottamatta perusteellisen huoltopalvelun tiloja, jotka on tarkoitettu suurnopeusjunien tai muun tyyppisen erityisiä tiloja edellyttävän liikkuvan kaluston huoltoon,
- f) muut tekniset laitteet, mukaan lukien puhdistus- ja pesulaitteet,
- g) rautatietoimintaan liittyvät meri- ja sisävesisatamien varusteet,
- h) taukotilat,
- i) tankkauspalvelut ja polttoaineen toimittaminen näihin palveluihin, josta perittävät maksut on erotettava laskuissa polttoaineen tankkauksessa perittävistä maksuista. (Rautatiemarkkinadirektiivi 2012/34/EU)

Raideliikennelaissa säädetään myös rataverkon haltijan velvollisuudesta maksujen ja niiden perusteiden julkaisemisesta verkkoselostuksessaan. (Raideliikennelaki 142 §)

## 2.2 Ratamaksun perusmaksun määrittämisen sääntely

Alaluvussa 2.1 mainitut vähimmäiskäyttöpalvelut tulee rautatiemarkkinadirektiivin (2012/34/EU) 29-34 artiklojen mukaan hinnoitella rataverkon haltijalle suoraan liikenteen harjoittamisesta aiheutuvien kustannusten perusteella. Junaliikenteen harjoittamisesta suoraan aiheutuviin kustannuksiin sisältyvät kustannukset on määritelty tarkemmin Euroopan komission täytäntöönpanoasetuksessa 2015/909. Täytäntöönpanoasetuksella yhtenäistetään maksun laskentaperusteita Euroopan unionin jäsenvaltioissa.

Euroopan komission täytäntöönpanoasetuksen 3 artiklassa on esitetty kustannukset, jotka voidaan laskea koko verkon laajuisiksi suoriksi kustannuksiksi ja näiden kustannusten määritelmät löytyvät liitteestä 1. Täytäntöönpanoasetuksen artiklan 3 kohdassa 1 määritellään koko verkon laajuisiksi suoriksi kustannuksiksi "erotus toisaalta vähimmäiskäyttömahdollisuuksien palvelujen tarjoamisesta ja palvelupaikkoja yhdistävän infrastruktuurin käyttöoikeuksista aiheutuvien kustannusten sekä toisaalta 4 artiklassa tarkoitettujen muiden kuin huomioon otettavien kustannusten välillä". Täytäntöönpanoasetuksen artiklassa 4 (liite 2) mainitaan kustannukset, jotka tulee jättää koko verkon laajuisten suorien kustannusten tarkastelun ulkopuolelle. Artiklan 4 k) -kohdassa mainitaan, että sellaisen rautatieliikenteen kustannuksiin, jossa ei käytetä sähkönsyöttölaitteistoa, ei saa sisältyä sähkönsyöttölaitteiston kustannuksia. Tämän artiklan noudattaminen johtaa siihen, että sähkövetoiselle liikenteelle pitää määritellä erillinen lisähinta perusmaksun lisäksi, ja dieselvetoiselle lisähintaa ei voi kohdentaa.

Ratamaksun perusmaksun alennuksia ja korotuksia voidaan myöntää kannustintyyppisesti ja ohjaavasti muun muassa perustuen junakalustoon. Alennuksia ratamaksuun voidaan myöntää vähämeluiselle junakalustolle ja melua rajoittavan jarruteknologian jälkiasennusta varten. Ratamaksusta myönnettäviä alennuksia voidaan myöntää myös kannustamaan eurooppalaisen junien automaattisen kulunvalvonnan standardin mukaisen järjestelmän käyttöönottoa varten. Alennuksia voidaan ottaa käyttöön myös uuden rautatieliikenteen kehittämisen kannustamiseksi tai vastaavasti kannustimeksi käyttämään vähäliikenteisiä rataosia. Alennukset tulee myöntää kaikille rautatien harjoittajille tasapuolisesti ja syrjimättömästi ja ne saavat koskea vain tiettyjen rataverkon osilta perittäviä maksuja. (Raideliikennelaki 140 §)

Korotettua ratamaksua voidaan periä ajoittain tai kokonaisuudessaan ylikuorimitettujen rataosien käytöstä. Ratamaksun korotuksen perusteena on tässä tapauksessa kattaa ylikuormituksen poistamisesta aiheutuvat kustannukset. Myös uusilta rataosilta voidaan periä korotettua ratamaksua niiden rakentamisesta aiheutuneiden investointikustannusten kattamiseksi. Perusmaksun alennukset tai korotukset ovat lähtökohtaisesti suuruudeltaan korkeintaan 80 prosenttia perusmaksun tasosta. (Raideliikennelaki 140 §)

Rataverkon haltijalla on mahdollisuus periä myös lisämaksuja raideliikennöinnistä, mikäli markkinat niin sallivat. Lisämaksujen käyttöönotto vaatii rataverkon haltijalta markkinoiden segmentointia sekä lisähintojen vaikutusten arviointia. Maksujen pitää olla sillä tasolla, että myös sellaiset rautatieliikenteen harjoittajat, jotka pystyvät maksamaan vain ratamaksun perusmaksun sekä sen katetuoton, voivat käyttää rataverkkoa. Lisämaksujen käyttöönottoa varten täytyy siis rataverkon haltijan osoittaa markkinoiden maksukyky. Tämän lisäksi verkkoselostuksessa tulee osoittaa lisämaksujen vastikkeena kyseisessä markkina- luokassa tarjottavat lisä- ja erityispalvelut. (Raideliikennelaki 141 §)

## 2.3 Ratamaksun perusmaksun laskemistavat

Perusmaksun laskemistavat ovat myös säännelty Euroopan komission täytäntöönpanoasetuksessa 2015/909. Perusmaksun tulee olla koko verkon tai verkon osien keskimääräisten suorien yksikkökustannusten suuruinen. Keskimääräiset suorat yksikkökustannukset on täytäntöönpanoasetuksen 5 artiklan mukaan mahdollista laskea jakamalla koko verkon laajuiset suorat kustannukset tosiasiallisten tai arvioitujen junakilometrien tai bruttotonnikilometrien kokonaismäärällä. Edellä mainitulla menetelmällä viitataan vähennyslaskumenetelmään. Toinen tapa on, 6 artiklan mukaan, laskea suorat yksikkökustannukset vankkaan näyttöön perustuvan teknisen tai ekonometrisen kustannusmallinnuksen avulla. Jälkimmäinen vaihtoehto edellyttää, että on mahdollista osoittaa rautateiden sääntelyelimelle, että suoriin yksikkökustannuksiin sisältyvät vain sallitut kustannukset. (Karjanlahti & Kukkarinen 2019, 4-5)

Suorien rautatieliikenteestä aiheutuvien yksikkökustannusten mallinnukseen Euroopan komission täytäntöönpanoasetuksen 2015/909 puitteissa on julkaistu tulkintoja eurooppalaisen rautatiealan itsenäisten sääntelyviranomaisten ryhmän, IRG-Railin toimesta. IRG-Rail on vuonna 2011 perustettu rautateiden sääntelyelinten ryhmä, johon kuuluu nykyään 31 Euroopan maan sääntelyelimet (IRG-Rail 2020). Ryhmä tekee Euroopan komission säännöksiin ja direktiiveihin tulkintoja ja kannanottoja, joihin kansalliset rautatieliikenteen sääntelyelimet voivat

tukeutua päätöksissään. IRG-Rail julkaisi vuonna 2016 artikkelin rautatieliikenteen suorien kustannusten laskemista mukaillen Euroopan komission täytäntöönpanoasetusta 2015/909. Artikkelissa esitellään IRG-Railin käsitys täytäntöönpanoasetuksen artikkelissa 6 mainituista mallinnustavoista. Artikkelin ensimmäinen metodologia viittaa täytäntöönpanoasetuksen tekniseen, eli niin kutsuttuun insinöörimallinnukseen. (IRG-Rail 2016, 5)

### 2.3.1 Tekninen mallinnus

Teknisellä eli niin kutsulla insinöörimenetelmällä viitataan insinööritieteitä hyödyntävään rataverkon kulumisen tekniseen mallintamiseen. Insinöörimenetelmät nojaavat luonnontieteisiin, erityisesti fysiikkaan, etsien riippuvuussuhteita junaliikennöinnin ja rataverkon infrastruktuurin kulumisen kesken. Menetelmien hyödyntäminen vaatii teknistä ymmärrystä mallinnettavista ilmiöistä, sillä rautatieinfrastruktuurin kulumista mallinnetaan junan raiteisiin ja muihin komponentteihin kohdistamalla voimilla. Voimat vaihtelevat erilaisten junatyypin, raiteiden ja painolastien mukaan, joten insinöörimenetelmien hyödyntämiseen tarvitaan tarkkaa dataa. Tarvittavaa dataa on esimerkiksi liikennöivien junien määrä, junatyyppit, akselipainot ja radan ominaisuudet. Insinöörimenetelmän lopullisena tavoitteena on saada rakennettua tekninen mallinnus, josta saadaan arvio riippuvuudesta liikenteen ja radan kulumisen välille. (IRG-Rail 2016, 5)

IRG-Railin (2016, 6) mukaan insinöörimenetelmän avulla voidaan lopulta laskea yksikkökustannukset liikennesuoritetta kohden, kun menetelmän tuottamat riippuvuusarvot kohdistetaan kustannuksiin. Menetelmä on kuitenkin kapeakatseinen, sillä se olettaa kaikkien kustannusten riippuvan fyysisestä kulumisesta. Tätä varten IRG-Rail (2016, 6) esittää artikkelissa insinöörimallit yhdistettäväksi ei-ekonometriin top-down menetelmiin, joita käytetään kustannusten allokointiin. Top-down-menetelmillä viitataan artikkelissa kustannusten allokointiin ja suodatukseen. Kustannusten allokoinnin top-down menetelmien ensimmäinen vaihe on koota radanpidon kustannusdata rataverkon haltijan kirjanpitojärjestelmästä. Mikäli kaikki vähimmäiskäyttöpalveluista aiheutuvat kustannukset eivät ole suoraan saatavilla, tulee ne suodattaa koko radanpidon kustannusdatamassasta. Kustannusdatan suodatus toteutetaan poistamalla kustannukset, jotka eivät aiheudu vähimmäiskäyttöpalveluiden tarjoamisesta sekä ne kustannukset, jotka eivät ole suoraan riippuvaisia liikenteestä. Edelleen suodatus jatkuu poistamalla aineistosta täytäntöönpanoasetuksen 2015/909 artikkelissa 4 (liite 2) mainitut kielletyt erät, joita ei saa ottaa huomioon koko verkon laajuisia suoria kustannuksia laskettaessa.

Seuraavassa vaiheessa suodatetuille kustannuksille määritetään kustannusajurit perustuen insinöörimallin antamiin riippuvuusarvioihin. Kustannuksen riippuvuus liikenteestä voidaan esittää prosenttina, joka ilmaisee riippuvuuden liikenteen suhteen. Esimerkiksi asiantuntija voi määrittää oman teknisen ymmärryksensä myötä sähkökunnossapidon erotusjaksojen riippuvuuden liikenteestä olevan 90% seurauksena suorasta fyysisestä kontaktista, kosketusvoimasta ja värähtelystä (Väylävirasto 2019, 17). Huomioitavaa on kuitenkin, että kaikki junaliikenteestä riippuvat kustannukset eivät ole suoraan johdettavissa rautateiden kulumisen tai muiden fyysisten ominaisuuksien perusteella. Esimerkkinä kyseisistä kustannuksista ovat henkilöstökulut, joita on tarvittu junien aikataulutukseen tai reititykseen. Näiden kustannusten sisällyttäminen mallinnukseen tapahtuu kuitenkin samalla tavalla perustuen asiantuntijan tietämykseen, vaatien kuitenkin vankat ja objektiiviset taustat sisällyttämisen perusteista.

### 2.3.2 Ekonometrinen mallinnus

IRG-Rail (2016, 7) esittelee toisena metodologiana näkemyksensä ekonometrisestä mallinnuksesta. Ekonometrinen rajakustannusten määrittelyprosessi on havainnollistettu kuvassa 1. Prosessi alkaa datasettien keräyksellä ja suodattamisella. Kustannusdata tulee kerätä vain soveltuvista eristä, eli suoraan rautatieliikenteen harjoittamisesta aiheutuvista eristä, eikä missään nimessä täytäntöönpanoasetuksen 2015/909 artiklassa 4 mainituista kielletyistä eristä. Kustannusdataa tulee kerätä vähintään vuoden ajalta ja sen laajuus tulee olla riittävä rautatieverkon tasolla, jotta mallinnus voidaan suorittaa. Pidemmältä ajanjaksolta saatavilla oleva data parantaa mallinnuksen tarkkuutta ja tuloksia. Liikennöintidata tulee kerätä samalta ajanjaksolta ja samalta laajuudelta kuin edellä mainitut kustannusdatat, jotta riippuvuussuhteita voidaan mallintaa. Yleisesti liikennöinnin yksikkönä on käytössä joko (brutto)tonnikilometrit tai junakilometrit, joita suositellaan kerättävän eri liikennöintityypeille erikseen. Liikennöintityypeillä viitataan henkilö- ja tavaraliikenteeseen. Mikäli mahdollista, voidaan kerätä dataa myös rataverkon infrastruktuurin eri osista. Täytäntöönpanoasetuksen 2015/909 artikla 5 mahdollistaa infrastruktuurin osia kuvaaviksi muuttujiksi radan muuttujat, erityisesti radan kaarresäteet. Tämä rautatieinfrastruktuurin ominaisuuksia kuvaava data tulee myös kerätä samalle ajanjaksolle ja laajuudelle kuin yllä mainitut datat. Seuraavassa vaiheessa nämä kolme datasettiä tulee yhdistää samaan datasettiin, jotta voidaan toteuttaa ekonometrinen mallinnus. (IRG-Rail 2016, 7)

Vaihe	Tavoite	Kuvaus
1 a	Kustannusdatan keräys	<ul style="list-style-type: none"> <li>Data operationaalisista, ylläpito- ja korvausinvestointikustannuksista</li> <li>Suosittelavaa jakaa em. kustannukset tyyppien mukaisesti</li> <li>Tulee olla kerättyä rautatieverkon laajuisesti ja vähintään vuoden ajalta</li> <li>Kiinteät ja kielletyt kustannukset voidaan poistaa jo tässä vaiheessa</li> </ul>
1 b	Liikennöintidatan keräys	<ul style="list-style-type: none"> <li>Tulee kerätä samalta ajalta ja samalle yksikölle kuin yllä mainitut kustannukset</li> <li>Yleensä yksikkönä tonnikilometrit tai junakilometrit, suositeltavaa kerätä eri liikennöintityypeille erikseen</li> </ul>
1 c	Muun datan keräys	<ul style="list-style-type: none"> <li>Dataa verkon infrastruktuurin osa-alueista (esim. ratojen lukumäärä ja tyyppi, komponenttien ikä etc.) samoille yksiköille kuin em. Datat</li> <li>Myös mahdollista kerätä esimerkiksi sääolosuhteista ja maanosista dataa</li> </ul>
2	Datan järjestely	<ul style="list-style-type: none"> <li>Kustannus-, liikennöinti- ja muu data tulee yhdistää samaan datasettiin, jotta ekonometrista mallinnusta voidaan hyödyntää</li> </ul>
3	Ekonometrisen estimointi	<ul style="list-style-type: none"> <li>Ekonometrisen mallien estimointi. Riippuva muuttuja on kustannustyyppi ja itsenäiset muuttujat ovat liikennemäärät ja kontrollimuuttujia voivat olla esimerkiksi verkon infrastruktuurin eri osa-alueet, ilmasto tai maanosat.</li> </ul>
4 a	Jouston johtaminen	<ul style="list-style-type: none"> <li>Ekonometrisen mallin tuloksista on mahdollista johtaa kustannusten jousto liikennöinnin suhteen.</li> <li>Jousto esittää prosenttiosuuden, jolla kustannukset riippuu liikennöinnin määrästä</li> </ul>
4 b	Rajakustannusten laskeminen	<ul style="list-style-type: none"> <li>Viimeisenä lasketaan yksikkökustannukset. Rajakustannus kuvaa kustannusta per liikenneyksikkö (esimerkiksi bruttotonnikilometri)</li> <li>Jos jousto on jo laskettu, rajakustannukset voidaan laskea jousto kertaa keskimääräiset kustannukset (as elasticities times avg costs)</li> </ul>

Kuva 1. *Rajakustannusten määrittäminen ekonometrisellä menetelmällä (IRG-Rail 2016, 10)*

---

Kolmannessa vaiheessa suoritetaan ekonometrinen estimointi. Ekonometrisen mallin riippuvana muuttujana tulee olla kustannukset ja itsenäisinä muuttujina edellä mainittuja liikennöintidatan muuttujia. Itsenäisiä muuttujia ovat siis esimerkiksi liikennemäärät ja bruttotonnit. Lisäksi itsenäisinä muuttujina voidaan käyttää rataverkon infrastruktuuria kuvaavia muuttujia, kuten esimerkiksi vaihteiden lukumääriä tai maantieteellisiä sijainteja. (IRG-Rail 2016, 8)

Lopulta ekonometrisesta mallista johdetaan kustannusten jousto liikennemäärien suhteen. Joustolla viitataan prosenttiosuuteen, jolla kustannukset riippuvat liikennemäärästä. Rajakustannukset on mahdollista laskea kustannusjouston avulla, kertomalla jousto keskimääräisellä kustannuksella. Toinen tapa on osittaisderivoida mallinnettu kustannusfunktio liikennettä kuvaavan muuttujan tai muuttujien suhteen ja täten laskea rajakustannus, eli kustannus, joka aiheutuu yhden yksikön muutoksesta bruttotonnikilometriä kohden. (IRG-Rail 2016, 9)

Tässä työssä toteutetaan kuvasta 1 vaiheet 2–4b. Kustannusdatan, liikennöintidatan ja muun datan keräys toteutetaan taustalla eikä sitä dokumentoida tarkemmin tässä työssä.

### 3 Rautatieinfrastruktuurin kustannusfunktion mallinnuksen teoriaa

Lähtökohtana radanpidon kustannusfunktion määrittämisessä on valita muuttajat, joille on löydettävissä yhteys kustannuksiin. Kustannuksiin vaikuttavia muuttujia voidaan teoriassa luetella useita, mutta rajoittavina tekijöinä kustannusfunktion määrittelyssä toimii muuttujien numeerisen määrittämisen haastavuus sekä saatavilla olevan aineiston laajuus.

Kustannusfunktio on matemaattinen rakenne, joka yhdistää kustannukset ( $C$ ) tuotantopanoksiin ( $Y$ ) ja tuotantopanosten hintoihin ( $W$ ). Kustannusfunktio voidaan esittää ilman yksilöivää matemaattista funktiota seuraavasti (kaava 1):

$$C = f(Y, W) \quad (1)$$

Rautatieinfrastruktuurin kustannusfunktioiden pohjana yleisimmin käytetty funktiomuoto on Christensenin ja muiden (1972) kehittämä Translog-tuotantoteknologiaa kuvaava kustannusfunktio. Translog kuvaa kustannusten ( $C$ ) suhdetta  $m$ -ulotteiseen tuotantopanosten vektoriin ( $Y$ ) ja  $n$ -ulotteiseen tuotantopanosten hintavektoriin ( $W$ ). Translog-kustannusfunktion avulla voidaan esittää  $U$ :n muotoinen kustannuskäyrä sekä siitä voidaan johtaa yksinkertaisesti kustannusjousto. Kyseinen kustannusfunktio on esitetty kaavassa 2. (Nash 2015, 11)

$$\begin{aligned} \ln C = & \alpha_0 + \sum_{i=1}^m \alpha_i \ln Y_i + \sum_{j=1}^n \beta_j \ln W_j + \\ & \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m \alpha_{ij} (\ln Y_i \ln Y_j) + \\ & \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \beta_{ij} (\ln W_i \ln W_j) + \\ & \sum_{i=1}^m \sum_{k=1}^n \delta_{ik} (\ln Y_i \ln W_k) \end{aligned} \quad (2)$$

Kaavassa 2 kokonaiskustannuksia itsenäisinä tekijöinä toimivat tuotantopanosten ja tuotantopanosten hintojen logaritmien lisäksi niiden risti- ja neliötermit. Translogia käyttäessä ei tarvitse etukäteen kiinnittää sitä, missä määrin panokset ovat keskenään korvautuvia, koska substituution aste saadaan mallinnuksen tuloksena (Tervonen & Pekkarinen 2007, 23). Substituution aste viittaa substituoituihin, joka karkeasti tarkoittaa kuinka helposti tuotantopanoksia voidaan vaihtaa toisiinsa (Lescohier 1935, 5). Mitä korkeampi substituoituihin, sitä pienempi vaikutus tuotantotekijöiden keskenään vaihtamisella on, kun tuotannon taso pysyy samana. Kaava 2 ilman ristituloa ja neliötermiä on Translog-funktion erityistapaus, yleisesti käytetty Cobb-Douglas-kustannusfunktio (Nash 2015, 21). Cobb-Douglas kuvaa tuotantokustannukset minimoivaa tuotantoteknologiaa kahden tuotantopanoksen ja yhden tuotoksen tapauksessa. Tuotantopanokset ovat korvattavissa keskenään, koska tuotantopanosten substituoituihin on Cobb-Douglas-kustannusfunktion tapauksessa vakio, yksi. (Tervonen & Pekkarinen 2007, 23)

Usein kustannusfunktioista jätetään tuotantopanosten hintavektori pois puutteellisen datan takia tai siksi, että voidaan olettaa tuotantopanosten hintojen olevan samalla tasolla eri kustannuserien välillä (Johansson & Nilsson 2004, 279; Andersson 2009, 5; Nash 2015, 21). Tätä linjaa mukailien Translog-teknologiaa hyödyntävä kustannusfunktio voidaan kirjoittaa muotoon (kaava 3)



$$\ln C = \alpha_0 + \sum_{i=1}^m \alpha_i \ln Y_i + \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m \alpha_{ij} (\ln Y_i \ln Y_j) \quad (3)$$

missä  $i$  ja  $j$  ovat rataosan indeksejä.  $Y_i$  on rataosan  $i$  tuotantopanosvektori ja  $C$  on kokonaiskustannukset.

Asettamalla Translog-kustannusfunktiolle rajoite  $\alpha_{ij} = 0$  voidaan testata, onko Cobb-Douglas-kustannusfunktio sopiva aineistolle. Mikäli huomataan, että neliötermien kertoimet eivät poikkea nolasta, on mallinnukseen sopiva muoto Cobb-Douglas-kustannusfunktio. Tämä supistunut muoto on seuraava (kaava 4):

$$\ln C = \alpha_0 + \sum_{i=1}^m \alpha_i \ln Y_i \quad (4)$$

Translog- ja Cobb-Douglas-kustannusfunktiot olettavat tehokkaimman datan muunnoksen olevan logaritminen. Lisäksi Translog- ja Cobb-Douglas-kustannusfunktioiden heikkouksia ovat muun muassa nolla-arvojen käsittely, joka ei ole mahdollista logaritmisessä muodossa. Translogin ongelmana on merkittävän multikollineaarisuuden esiintyminen, joka johtuu erilaisista termeistä, niiden neliöistä ja ristituloista. (Nash 2015, 11-12) Multikollineaarisuuden esiintymisen ja nolla-arvojen heikon käsittelyn vuoksi vaihtoehtoisiksi kustannusfunktion muodoiksi on esitetty Box-Cox-muunneltua kustannusfunktiota. Box-Cox-muunneltu kustannusfunktio perustuu riippuvan muuttujan muunnokseen kaavalla (kaava 5)

$$y^{(\lambda)} = \begin{cases} \frac{y^\lambda - 1}{\lambda}, & \text{kun } \lambda \neq 0 \\ \log y, & \text{kun } \lambda = 0 \end{cases} \quad (5)$$

Box-Cox-funktiomuoto sopii usein paremmin mallinnuksiin kuin Translog-kustannusfunktio, sillä se sisältää Translogin logaritmuunnoksen erityistapauksessa, kun muunnosparametri on nolla. (Nash 2015, 12) Muunnosparametrin estimointi toteutetaan etsimällä optimaalinen arvo, jolla saadaan muunnoksen seurauksena muuttujat seuraamaan normaalijakaumaa. Etsintä voidaan tehdä esimerkiksi Shapiro-Wilk testin tai suurimman uskottavuuden menetelmän avulla. (Asar et al. 2017, 7) Box-Tidwell-mallissa muunnetaan myös jokin itsenäinen tai kaikki itsenäiset muuttujat samalla kaavalla 5, kukin muuttuja kuitenkin omalla muunnosparametrilla. Box-Tidwell-muunneltu kustannusfunktio voidaan esittää seuraavasti (kaava 6):

$$C_i^{(\lambda)} = \alpha + \beta_i Y_i^{(\lambda)} \quad (6)$$

missä  $i$  on rataosan indeksi.  $Y_i$  on rataosan  $i$  tuotantopanosvektori ja  $C$  on rataosan  $i$  kustannukset. Yläindeksi kuvastaa muuttujan muunnosparametrin muusta Box-Cox-muunnosta.

### 3.1 Lineaarinen regressio

Regressioanalyysi on tilastotieteen menetelmä, jonka avulla voidaan analysoida jonkin tekijän tai muuttujan riippuvuutta toisista muuttujista tai tekijöistä, kun riippuvuus on tilastollisesti merkitsevää. Riippuvuutta kuvaava regressiomalli on tilastollinen malli, joka pyrkii selittämään jonkin riippuvan muuttujan tai tekijän havaittujen arvojen vaihtelun joidenkin itsenäisten muuttujien tai tekijöiden havaittujen arvojen vaihtelun avulla. (Mellin 2006, 267) Regressiomalleille löytyy useita käyttökohteita tilastotieteissä, muun muassa aikasarjojen analyysissä ja ennustamisessa sekä ekonometriassa (Mellin 2006, 277). Teoreettisesta näkökulmasta regressiomalli, joka ei edellytä mitään matemaattista funktiomuotoa, voidaan esittää seuraavasti (kaava 7):

$$Y = f(X_1, X_2, \dots, X_k, \varepsilon) \quad (7)$$

missä  $Y$  on riippuva tekijä tai muuttuja,  $X_1, X_2, \dots, X_k$  ovat itsenäisiä muuttujia tai tekijöitä ja  $\varepsilon$  ovat mallin virhetermejä eli jäännöksiä (Welc & Esquerdo, 1). Yleinen lineaarinen malli, jolla pyritään selittämään riippuvan muuttujan havaittujen arvojen vaihtelu useamman itsenäisen muuttujan havaintojen arvojen vaihtelun avulla, voidaan esittää seuraavasti (kaava 8):

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + \varepsilon_i \quad (8)$$

missä  $y_i$  on riippuvan muuttujan arvo havaintoyksikössä  $i$ .  $x_{ij}$  on  $j$ :nnen itsenäisen muuttujan arvo havaintoyksikössä  $i$ .  $\varepsilon_i$  on jäännöstermin arvo havaintoyksikössä  $i$ .  $\beta_0$  on vakio regressiokerroin ja  $\beta_j$  on itsenäisen muuttujan  $x_j$  regressiokerroin, missä  $j = 1, 2, \dots, k$ . Kaavassa 8 esitelty yleinen lineaarinen malli voidaan esittää kompaktisti matriisisein seuraavassa muodossa (kaava 9):

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (9)$$

missä  $\mathbf{y}$  on  $n \times 1$  vektori  $n$ :stä riippuvan muuttujan havainnosta.  $\mathbf{X}$  on  $n \times (k + 1)$  matriisi itsenäisten muuttujien havainnoista  $n$ :lle riippuvalle muuttujalle.  $\boldsymbol{\beta}$  on  $(k + 1) \times 1$  regressiokerroinien vektori.  $\boldsymbol{\varepsilon}$  on  $n \times 1$  virhetermien vektori.

Yleisen lineaarisen mallin parametrien estimointi toteutetaan tavallisesti pienimmän neliösumman menetelmällä (PNS). Pienimmän neliösumman menetelmä perustuu mallin virhetermien neliösumman minimoimiseen. Virhetermien neliösumma voidaan esittää seuraavasti lineaarisen mallin matriisiesityksen avulla (kaava 10):

$$\boldsymbol{\varepsilon}^2 = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \quad (10)$$

mistä voidaan ratkaista derivaamalla  $\boldsymbol{\beta}$ , joka minimoi virhetermien neliösumman. Merkitsemällä derivaatta nolaksi saadaan ratkaisuksi pienimmän neliösumman estimaattori. Pienimmän neliösumman estimaattori matriisimuodossa voidaan esittää seuraavasti (kaava 11):

$$\boldsymbol{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \quad (11)$$

## 3.2 Lineaarisen regression oletukset

Mellin (2006, 337-338) esittelee lineaarista mallia koskevat 6 standardioletusta, jotka takaavat, että tavanomaisia estimointi- ja testausmenetelmiä voidaan käyttää hyväksytysti mallin analysointiin. Nämä oletukset ovat seuraavat:

1. itsenäisten muuttujien arvot ovat kiinteitä, eli ei-satunnaisia vakioita
2. itsenäisten muuttujien välillä ei ole lineaarisia riippuvuuksia
3. kaikilla jäännös- eli virhetermeillä on sama odotusarvo
4. mallin homoskedastisuus, jonka mukaan kaikilla jäännös- eli virhetermeillä on sama varianssi
5. korreloimattomuus, jonka mukaan jäännös- eli virhetermit eivät korreloi keskenään
6. normalisuus, jonka mukaan mallin jäännös- eli virhetermit ovat normaalijakautuneita

Yllä mainittujen oletusten 1-5 täytyessä voidaan todeta Gauss-Markovin teoreeman mukaisesti pienimmän neliösumman estimaattorin olevan tehokkain harhaton lineaarinen estimaattori (BLUE, best linear unbiased estimator). Toisin sanoen teoreeman ehtojen täytyessä pienimmän neliösumman estimaattori on paras, eikä vaihtoehtoisia estimaatiomenetelmiä tarvitse harkita (Brooks 2014, 91). Kuudes oletus jäännöstermien normaliuudesta yhdessä oletusten neljä ja viisi kanssa mahdollistaa luottamusvälien laskemisen.

Kuudesta mainitusta standardioletuksesta oletuksia 4-6 voidaan ja tulee testata tilastollisesti. Oletuksessa 4 oletetaan mallin homoskedastisuus ja mikäli oletus ei päde, malli on heteroskedastinen. Heteroskedastisuudesta seuraa regressiokertoimien pienimmän neliösumman estimointimenetelmän tehottomuus (Mellin 2006, 338). PNS-menetelmä antaa silti harhattomat regressiokertoimien estimaatit, mutta keskivirheet ja niistä tehdyt johtopäätökset voivat olla harhaanjohtavia. (Brooks 2014, 184-185) Heteroskedastisuuden testaamiseen löytyy lukuisia testejä, joista yksi käytetyimmistä on Breusch-Paganin testi. Testi alkaa ottamalla estimoidun mallin virhetermien neliöt ylimääräiseen regressioon. Ylimääräisessä regressiossa itsenäisinä muuttujina toimivat alkuperäisen mallin itsenäiset muuttujat. Tämän ylimääräisen mallin  $R^2$ -arvon avulla voidaan tehdä F-testi, jonka merkitsevyys kertoo heteroskedastisuudesta. (Breusch & Pagan 1979, 1288-1290) Mikäli heteroskedastisuutta ilmenee mallissa, sitä voi korjata esimerkiksi tekemällä muuttujille logaritmuunnoksen tai vähentämällä muuttujien arvosta jonkun sopivan luvun. Näillä toimenpiteillä saadaan äärihavainnot skaalattua sopivammiksi. (Brooks 2014, 186)

Oletus 5, jonka mukaan mallissa ei esiinny jäännöstermien välistä korrelaatiota on myös testattava tilastollisesti. Tämä oletus viittaa mallin autokorrelaatioon, eli virhetermin ja edellisen tai edellisten havaintojen virhetermien suhteeseen. Yksi yleinen autokorrelaation tunnistavista testeistä on Durbin-Watsonin testi, joka testaa ensimmäisen asteen autokorrelaatiota. (Brooks 2014, 199) Durbin-Watson-testiä voidaan pitää mielekkäänä pääasiasaassa vain aikasarja-aineistolle, joten tässä työssä ei tarkemmin tarkastella testin tuloksia. Testin tulokset on kuitenkin esitetty lopussa liitteessä.

Jäännöstermien normaalijakautuneisuus on kuudes tilastollisesti testattava oletus. Kuudes oletus ei vaikuta pienimmän neliösumman estimaattorin tehokkuuteen eikä harhaisuuteen, mutta mikäli jäännöstermit poikkeavat merkittävästi normaalijakaumasta, mallinnuksen kerrointen luottamusvälit voivat olla

liian kapeita tai leveitä. Normaalijakautuneisuuden häiriöitä voi aiheuttaa muutama äärihavainto, sillä parametrien estimointi perustuu jäännöstermien minimointiin.

Yksi yleisimmistä normaalisuuden testaamiseen käytetyistä testeistä on Jarque-Beran testi. Testi perustuu jäännöstermien jakauman vinoudesta ja huipukkuudesta laskettuun arvoon. Vinous viittaa jakauman epäsymmetrisyyttä tai symmetrisyyttä kuvaavaan tunnuslukuun ja huipukkuus on jakauman huipukkuutta kuvaava tunnusluku. Normaalijakauma ei ole vino ja sen huipukkuuden tunnusluvun arvo on 3. (Brooks 2014, 209) Vinouden tunnusluku  $S$  ja huipukkuuden tunnusluku  $K$  voidaan laskea kaavoilla (kaavat 12 ja 13)

$$S = \frac{E[\varepsilon^3]}{(\sigma^2)^{3/2}} \quad (12)$$

$$K = \frac{E[\varepsilon^4]}{(\sigma^2)^2} \quad (13)$$

missä  $\varepsilon$  on virhetermi ja  $\sigma^2$  on mallinnuksen virheiden varianssi. Kaavoista 12 ja 13 voidaan edelleen laskea Jarque-Beran testisuure kaavalla (kaava 14)

$$J - B = \frac{N}{6} \left( S^2 + \frac{(K-3)^2}{4} \right) \quad (14)$$

missä  $S$  on vinous ja  $K$  on huipukkuus. sekä  $N$  on havaintojen lukumäärä. Testin arvot seuraavat  $\chi^2$ -jakaumaa kahdella vapausasteella. Testissä nollahypoteesina on jakauman normaalisuus ja tämä hylätään, mikäli virhetermit ovat merkittävästi vinoja tai huipukkaita (tai molempia). Tämä tarkoittaa 5 prosentin merkitsevyytasolla, että testin arvoja tulee verrata kriittiseen arvoon 5,99, eli nollahypoteesin toteutuessa J-B-testin arvon tulee olla pienempi kuin 5,99. (Brooks 2014, 210)

Edellä mainittujen lisäksi erittäin tärkeä matemaattinen oletus on mallin lineaarisuus (Nyblom 2015, 14; Ballance 2020, 4). Mikäli riippuvan ja itsenäisten muuttujien välinen riippuvuus ei ole lineaarinen, regressioanalyysin tulokset vähättelevät tai yliarvioivat todellista riippuvuussuhdetta. Tästä seuraa kohonnut riski tyyppin 1 ja 2 virheille mallinnuksessa. (Osborne & Waters 2002, 2) Tyyppin 1 virhe viittaa nollahypoteesin hylkäämiseen, vaikka se on totta, eli kyseessä on false-positive. Tyyppin 2 virhe viittaa tilanteeseen, jossa nollahypoteesia ei hylätä, vaikka se tulisi hylätä, eli kyseessä on false-negative. Lineaarista riippuvuutta voidaan tarkastella regressioanalyysin jäännöstermien visualisoinnin avulla.

Tässä työssä tarkastellaan myös, ettei mallinnuksen muuttujien välillä ole täydellistä multikollineaarisuutta. Multikollineaarisuus viittaa itsenäisten muuttujien väliseen korrelaatioon. Korkea muuttujien välinen korrelaatio aiheuttaa mallille korkean selitysasteen ilman, että riippuvat muuttujat ovat tilastollisesti merkittäviä. (Brooks 2014, 218) Multikollineaarisuutta voidaan testata varianssin inflaatiotekijällä (Variance Inflation Factor, VIF) (Mellin 2006, 402-403). Multikollineaarisuuden tulkinta varianssin inflaatiotekijällä on kolmivaiheinen. Ensimmäisessä vaiheessa alkuperäisen lineaarisen regression kaikille itsenäisille muuttujille ajetaan omat pienimmän neliösumman regressiot, missä riippuvana muuttujana toimii yksi alkuperäisistä itsenäisistä muuttujista ja itsenäisinä muuttujina toimivat loput itsenäiset muuttujat. Esimerkiksi kolmen itsenäisen muuttujan lineaarisen regression ensimmäisen itsenäisen muuttujan regressio on seuraava (kaava 15):

$$X_1 = \beta_0 + \beta_1 X_2 + \beta_2 X_3 + \varepsilon \quad (15)$$

missä  $\beta_0$  on vakiotermi ja  $\varepsilon$  on regression jäännös-, eli virhetermi. Varianssin inflaatiotekijän tulkinnan toisessa vaiheessa näiden regressioiden selitysasteen avulla voidaan laskea jokaisen muuttujan kohdalla varianssin inflaatiotekijä seuraavasti (kaava 16):

$$VIF_i = \frac{1}{1-R_i^2} \quad (16)$$

missä  $R_i^2$  on alkuperäisen regression  $i$ :nnen itsenäisen muuttujan kaavan 15 mukaisen ylimääräisen regression selitysaste. Kolmannessa, eli viimeisessä vaiheessa tulkitaan varianssin inflaatiotekijöiden arvoja. Yleisesti nyrkkisääntönä voidaan pitää sitä, että mikäli varianssin inflaatiotekijä saa arvon yli 10, multikollinearisuudesta saattaa olla haittaa (Mellin 2006, 403). Multikollinearisuuteen voidaan reagoida monella tapaa: se voidaan jättää huomioimatta, kollineaarinen muuttuja voidaan jättää ulos mallista tai suuresti korreloivat muuttujat voidaan muuttaa niiden suhteeksi (Brooks 2014, 219).

### 3.3 Regressiomallinnuksen selitysvoima

Regressiomallinnuksen selitysvoimaa kuvaavia testejä on useita. Näistä yleisimmin käytettyjä ovat  $R^2$ -luku ja F-testi.  $R^2$ -luvulla tarkoitetaan mallin selitystasetta. Luku kertoo, kuinka suuren osuuden riippuvan muuttujan vaihtelusta itsenäiset muuttujat pystyvät selittämään.  $R^2$ -luku saadaan laskemalla mallin tuottamien ennustearvojen sekä riippuvan muuttujan arvojen korrelaation neliö.  $R^2$ -luvun ongelma on se, että selitysosuus nousee aina, kun regressioanalyysiin lisätään itsenäisiä muuttujia, vaikka lisätyt muuttujat eivät todellisuudessa nostaisikaan selityskykyä. Tätä ongelmaa varten mallien keskinäistä vertailua varten on korjattu  $R^2$ -luku (adjusted  $R^2$ ), joka huomioi mallinnuksen sisältämien itsenäisten muuttujien lukumäärän. Se on aina arvoltaan yhtä suuri tai pienempi kuin  $R^2$ -luku. Korjattu  $R^2$ -luku lasketaan kaavan 17 mukaisesti (Mellin 2006, 378-379).

$$R^2 = 1 - \frac{n-1}{n-p} * \frac{SSE_p}{SST} \quad (17)$$

missä  $SSE$  on jäännösneliösumma mallista ja  $SST$  on muuttujan vaihtelua kuvaava kokonaisneliösumma.  $n$  on mallinnuksen havaintojen lukumäärä ja  $p$  on mallinnuksen itsenäisten muuttujien lukumäärä.

F-testi kertoo, pystytäänkö regressioanalyysin muuttujilla selittämään riippuvan muuttujan vaihtelua. Se on tilastollinen testi, jolle saadaan merkitsevyystaso. F-testi lasketaan  $R^2$ -luvun avulla seuraavasti (kaava 18):

$$F = (n - 2) \frac{R^2}{1-R^2} \quad (18)$$

Korjattua  $R^2$ -lukua ja F-testiä hyödynnetään tässä työssä mallinvalintakriteereinä. Korjatun  $R^2$ -luvun selitysasteen mukaan vertailtavista malleista paras on se, joka maksimoi selitysasteen arvon.

## 4 Valittujen Euroopan maiden ratamaksukäytäntöjä

Euroopan unionin jäsenmaiden ratamaksukäytäntöjä sääntelee aiemmin esitelty Euroopan parlamentin ja neuvoston direktiivi 21.11.2012/34/EU yhtenäisestä eurooppalaisesta rautatiealueesta sekä Euroopan komission täytäntöönpanoasetus (EU) 12.6.2015/909 rautatieliikenteen harjoittamisesta suoraan aiheutuvien kustannusten laskentasäännöistä. Suurimmassa osassa Euroopan talousalueen maita rautatieinfrastruktuurin hinnoittelu perustuu rajakustannuksiin (IRG-Rail 2018, 10). Rajakustannusten mallinnustapa vaihtelee jäsenmaissa täytäntöönpanoasetuksen mahdollistamien tapojen välillä. Ekonometrinen rajakustannusmallinnus on käytössä Suomessa, Ranskassa, Hollannissa, Norjassa ja Ruotsissa. (IRG-Rail 2018, 28)

Yksityiskohtaisempaan tarkasteluun tähän lukuun valikoituvat Norjan, Ranskan ja Suomen ratamaksun perusmaksun mallinnusmenetelmät. Näiden kolmen maan rautatieinfrastruktuurin haltijat tarjoavat yksityiskohtaiset menetelmämuistiot ratamaksun perusmaksun ekonometrisesta mallinnuksesta. Suurimmassa osassa Euroopan talousalueen jäsenmaita yksityiskohtaisia menetelmämuistioita ei ole saatavilla ollenkaan tai niitä ei ole käännetty ulkomaisille kielille. Tässä luvussa tarkastellaan Norjan, Ranskan ja Suomen ratamaksun perusmaksun mallintamiseen sisällyttämiä kustannuksia, parametreja sekä rajakustannuksen mallintamistapoja.

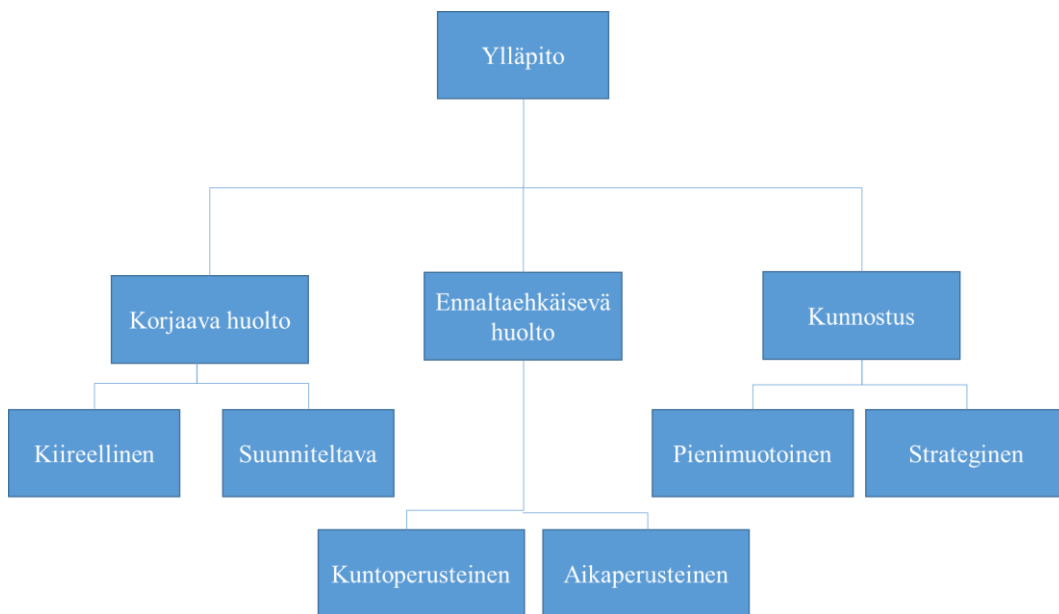
### 4.1 Norjan ratamaksun perusmaksu

Norjan rautatieinfrastruktuuri on valtionyhtiö Bane NORin omistuksessa. Bane NOR vastaa Norjan kansallisen rautatieverkon suunnittelusta, kehittämisestä, hallinnosta, toiminnasta ja ylläpidosta. Lisäksi yhtiön vastuulla on liikenteen ja rautateiden kiinteistöjen hallinta ja kehittäminen. (Bane NOR 2020) Ratamaksun vuoden 2019 eri luokat, siihen sisältyvät kustannukset ja maksun taso avataan rautateiden aikataulukauden 2021 verkkoselostuksessa. Ratamaksun perusmaksun määrittelytapa on ekonometrinen mallinnus ja koko määrittelyprosessi on kuvattu tarkemmin vuonna 2017 julkaistussa infrastruktuurin maksujen käyttöönottosuunnitelmassa. (Bane NOR 2019; Bane NOR 2017)

Norjan ratamaksun perusmaksu on jaoteltu liikennetyypeittäin henkilö- ja rahtiliikenteelle erillisiin maksuihin. Edelleen henkilö- ja rahtiliikenteeseen jaotellusta on jatkettu kolmen eri alueen jaoteluun. Aluejaotellut ovat Oslon paikallisalue, Ofotenin linja ja jäljelle jäävät raidealueet. (Bane NOR 2019, 8) Ofotenin linja on vain malmikuljetuksiin tarkoitettu linja ja se eroaa niin huomattavasti Norjan muusta rautatieverkosta liikenteen ja ominaisuuksien osalta, että sen kustannukset on eriytetty ja niistä on tehty erillinen ekonometrinen mallinnus. Ofotenin linja aiheuttaisi huomattavaa vääristymää koko verkon ekonometrisen mallinnukseen, koska siellä kulkee keskimäärin kaksi kolmasosaa kaikesta Norjan rautatieliikenteen kuormituksesta. Ofotenin linjalla on lisäksi 523 metrin korkeusero kahden terminaalien välillä, sekä se on ainoa linja, joka on suunniteltu kestämään 30 tonnin akselipainon. (Bane NOR 2017, 22)

#### 4.1.1 Norjan ratamaksun perusmaksuun sisällytettävät kustannukset

Ratamaksun perusmaksun laskelmiin sisällytettävät kustannukset on määritetty aiemmin esitellyn Euroopan komission täytäntöönpanoasetuksen 2015/909 säännösten mukaisesti. Huomioitavia laskelmista pois jääviä kustannuksia ovat liikenteenohjauksen kustannukset sekä pääoma- ja kiinteät kustannukset. Bane NOR:n kustannusrakenne ei kuitenkaan sovi sinällään laskelmaan soveltuvien kustannusten erittelyyn, joten käyttöönottosuunnitelmassa on esitetty yksityiskohtainen kuvaus, miltä osin kustannukset sisällytetään laskelmiin ja mikä jätetään ulos. (Bane NOR 2017, 16) Kuvassa 2 on esitetty Bane NOR:n ylläpitorakenne ylläpitotyypeittäin.



Kuva 2. Ylläpitorakenne (Bane NOR 2017, 18)

Korjaava huolto jaetaan rakenteessa kiireelliseen ja suunniteltavaan korjaavaan huoltoon. Kiireellinen huolto viittaa huoltotilanteisiin, jotka tulee toteuttaa välittömästi vikojen tai vaaratilanteiden ilmentyessä ja joiden tarvetta on vaikea ennustaa. Suunniteltava huolto taas viittaa vikojen ja ongelmien, jotka eivät aiheuta välittömiä ongelmia, huoltamiseen ja korjaamiseen. Ennaltaehkäisevän huollon tarkoituksena on ylläpitää tietoisuutta rautatieverkon tilasta ja kehityksestä ja se on jaettavissa kunto- ja aikaperusteiseen ennaltaehkäisevään huoltoon. Ennaltaehkäisevää huoltoa toteutetaan jaksottaisesti tai tarvittaessa kuntoperusteisesti ja esimerkkejä siitä ovat tarkastukset, mittaukset, siistiminen ja voitelu. Kunnostuksella viitataan korvausinvestointeihin, joilla korvataan korjaavassa ja ennaltaehkäisevässä huollossa olevia laitteistoja. Pienimuotoisia kunnostuksia toteutetaan tietyn rautatieverkon kuntotason ylläpitämiseksi, esimerkiksi melun vaimentamiseksi ja turvallisuuden ylläpitämiseksi. Strategiset kunnostukset viittaavat pidemmän aikavälin investointeihin, joilla saatetaan kohteet uutta vastaavaan tilaan hyödyntäen uusinta tarjolla olevaa teknologiaa. Esimerkkejä strategisista kunnostuksista ovat rautatieasemien remontoinnit, siltojen korjaukset ja rautateiden uusimiset. (Bane NOR 2017, 18)

Mainittujen ylläpitotyyppien kustannuserät eivät kokonaisuudessaan sovellu suoraan junaliikenteen aiheuttamien yksikkökustannusten laskennan pohjaksi, sillä kustannusten laajuus ei ole kokonaisuudessaan selitettävissä infrastruktuurin käytöllä. Kustannuksia aiheuttavat tekijät ovat lajiteltu neljään eri luokkaan:

1. käytöstä aiheutuvan kuluman ja rikkoutumisen aiheuttamat kustannukset
2. luonnollisen hajoamisen aiheuttamat kustannukset
3. infrastruktuurin palvelutason korottamisesta aiheutuvat kustannukset
4. muiden kuin yllä mainittujen syiden aiheuttamat kustannukset

Käytöstä aiheutuvan kuluman ja rikkoutumisen aiheuttamat kustannukset ovat suoraan riippuvaisia liikennöintimääristä, joten ne luovat perustan junaliikenteestä suoraan riippuvaisten yksikkökustannusten laskemiselle. Luonnollisen hajoamisen aiheuttamat kustannukset riippuvat osittain tai täysin ikääntymisestä tai luonnollisista ilmiöistä, eivätkä siten ole kokonaisuudessaan soveltuvia laskelmiin. Infrastruktuurin palvelutason korottamisesta aiheutuvat kustannukset taas ovat seurausta päätöksistä parantaa palvelutasoa, joten ne eivät sovellu myöskään laskelmiin. Muista syistä aiheutuvat kustannukset voidaan määritellä pääomakustannuksiksi, jotka ovat kiinteitä lyhyellä aikavälillä, eivätkä siten sovi yksikkökustannusten laskelmiin. Taulukossa 4 on esitetty, kuinka suurelta osin edellä mainitut tekijät vaikuttavat ylläpitotyyppien kustannusten määrään. (Bane NOR 2017, 19)

*Taulukko 4. Kustannuksia aiheuttavien tekijöiden vaikutukset kustannustyypeihin (Bane NOR 2017, 18)*

	Infrastruktuurin palvelutason kohottaminen	Luonnollinen hajoaminen	Käytöstä aiheutuva kuluma ja rikkoutuminen
Kiireellinen huolto	0%	50%	100%
Suunniteltava huolto	0%	50%	100%
Aikaperusteinen huolto	0%	100%	50%
Kuntoperusteinen huolto	0%	100%	50%
Pienimuotoinen kunnostus	50%	100%	50%
Strateginen kunnostus	100%	50%	50%

Taulukko havainnollistaa, miten eri kustannustyyppien kokonaiskustannukset riippuvat useista eri tekijöistä. Lopulta kaikki kustannukset ovat jossain määrin johdettavissa liikennöinnin määrästä, sillä ilman liikennöintiä kaikki toiminta lopetettaisiin, eikä siitä aiheutuisi kustannuksia. Norjan yksikkökustannusten kustannusdatasta jätetään lopulta ulos korvausinvestointikustannukset (pienimuotoinen ja strateginen kunnostus), sillä ne ovat riippuvaisia kertyneestä liikennehistoriasta, eikä tarkasteltavan vuoden liikennemäärästä. (Bane NOR 2017, 20) Korvausinvestointien sisällyttämättä jättäminen johtaa hieman pienempiin rajakustannuksiin, mutta se on yleisesti käytetty tapa infrastruktuurien hinnoittelussa (Bane NOR 2017; Link & Maibach 1999)



#### 4.1.2 Norjan kustannusfunktion ekonometrinen mallinnus

Kustannusten mallinnus voidaan toteuttaa tuotantomuuttujien ja teknisten muuttujien avulla rakennetulla kustannusfunktiolla. Taulukossa 5 seuraavalla sivulla on kuvattu Bane NOR:n kuuden kustannusfunktiota kuvaavan regressiomallin muuttujat. Kustannusfunktioita rakennetaan useita, jotta pystytään tarkastelemaan erilaisten liikenteen kuormitusten vaikutuksia. Datana kustannusfunktioimallinnuksissa on käytetty vuosien 2014, 2015 ja 2016 arvoja. Kustannusfunktion itsenäiseen muuttujaan, kustannuksiin, on sisällytetty aiemmin aluvussa 3.2.1. kuvattujen korjaavan huollon ja ennaltaehkäisevän huollon aiheuttamat kustannukset. Itsenäiset muuttujat ovat jaoteltu tuotantomuuttujiin ja teknisiin muuttujiin. Tuotantomuuttujat ovat muuttujia, joiden suhteen rajakustannukset johdetaan. Tekniset muuttujat kuvaavat ratainfrastruktuurin rataosien teknisiä ominaisuuksia. Mallit 2, 4 ja 5 perustuvat absoluuttisiin arvoihin: puhtaaseen lineaariseen funktioon, toisen asteen polynomiseen funktioon ja kolmannen asteen polynomiseen funktioon vastaavasti. Kaikki mallit ovat estimoitu myös riippuvan ja itsenäisten muuttujien luonnollisten logaritmien kanssa. (Bane NOR 2017, 24-25)

Taulukko 5. Norjan regressiomallien parametrit (Bane NOR, 24-25)

Malli	Tuotantomuuttujat	Tekniset muuttujat
Malli 1	<ul style="list-style-type: none"> <li>- Junien lukumäärä</li> <li>- Junan keskimääräinen paino</li> </ul>	<ul style="list-style-type: none"> <li>- Kilometr määrä</li> <li>- Vaihteiden lukumäärä per kilometri</li> <li>- Tunnelikilometrien osuus</li> <li>- Alle 500 metrin taipuisuuksien lukumäärä per kilometri</li> <li>- Painotettu nopeus</li> <li>- Dummy-muuttuja: sähköistetty</li> <li>- Hintaindeksi</li> </ul>
Malli 1b	- Mallin 1 muuttujat sisältäen neliö- ja ristitermit	Samat kuin mallissa 1
Malli 2	- Bruttotonnit	Samat kuin mallissa 1
Malli 3	<ul style="list-style-type: none"> <li>- Henkilöliikenteen bruttotonnit</li> <li>- Rahtiliikenteen bruttotonnit</li> </ul>	Samat kuin mallissa 1
Malli 3b	- Mallin 3 muuttujat sisältäen neliö- ja ristitermit	Samat kuin mallissa 1
Malli 4	<ul style="list-style-type: none"> <li>- Bruttotonnit (<math>x</math>)</li> <li>- Bruttotonnit (<math>x^2</math>)</li> </ul>	Samat kuin mallissa 1
Malli 5	<ul style="list-style-type: none"> <li>- Bruttotonnit (<math>x</math>)</li> <li>- Bruttotonnit (<math>x^2</math>)</li> <li>- Bruttotonnit (<math>x^3</math>)</li> </ul>	Samat kuin mallissa 1

Kuvatut mallit on lajiteltu kolmeen eri kategoriaan, polynomisiin malleihin (2, 4 ja 5) perustuen absoluuttisiin arvoihin, polynomisiin malleihin perustuen arvojen luonnollisiin logaritmeihin ja kahden tuotantomuuttujan Translog-malleihin (1, 1b, 3 ja 3b). Yhteensä malleja on rakennettu 10 kappaletta, kun sisällytetään myös mallien 2, 4 ja 5 luonnolliset logaritimuunnokset. (Bane NOR 2017, 26) Ensimmäisen kategorian malleista parhaiten suoriutui kolmannen asteen polynominen malli (malli 5), toisen kategorian malleista toisen asteen polynominen malli perustuen luonnollisiin logaritmeihin (malli 2) ja kolmannen kategorian malleista parhaiten suoriutui kahden tuotantomuuttujan Translog-malli (3b). Parhaiten suoriutuneiden mallien regressioanalyysien selittävyysasteet on kuvattu taulukossa 6. (Bane NOR 2017, 26-30)

Taulukko 6. Parhaiten suoriutuneiden mallien selitysasteet

	Malli 5	Malli 2	Malli 3b
$R^2$ -luku	0,396	0,379	0,404

Regressioanalyyseistä parhaan selittävyysasteen,  $R^2$ -luvun, sai malli 3b. Kahden tuotantomuuttujan Translog-malli selittää siis 40,4% kustannusten kokonaisvaihtelusta. Tuotantomuuttujina tässä mallissa ovat henkilö- ja rahtiliikenteen bruttotonnit, kyseisten bruttotonnien neliöt sekä interaktiomuuttuja, joka on henkilöliikenne kerrottuna rahtiliikenteellä. Teknisinä muuttujina ovat raidekilometrit, alle 500 metrin taipuisuuksien määrä per kilometri, tunnelikilometrien osuus, vaihteiden lukumäärä, painotettu nopeus, hintaindeksi ja dummy-muuttuja sähköistyksestä. Tämä mallinnus 3b päättyi myös ratamaksun perusmaksun mallinnustavaksi. Mallinnuksella määritettiin ratamaksun perusmaksu aikataulukausille 2018-2021. (Bane NOR 2017, 47)

## 4.2 Ranskan ratamaksun perusmaksu

Ranskan rautatieoperaattorina toimii SNCF (Société Nationale des Chemins de Fer Français), joka harjoittaa henkilö- ja tavaraliikennettä valtiollisella rataverkolla. SNCF:n organisaatio koostuu seitsemästä eri toimialasta. Toimialat ovat Réseau, Gares & Connexions, Fret SNCF, Voyageurs, Geodis ja Keolis. Toimialoista Réseau vastaa rataverkon infrastruktuurin ylläpidosta, kehityksestä ja liikenteenohjauksesta. Réseaurin vastuualueella on myös ratainfrastruktuurien maksujen määrittely. (SNCF 2020) Vuoden 2021 verkkoselostuksen liitteestä löytyy perusteet ratainfrastruktuurin maksujen eri luokille ja määrittelytavoille (SNCF Réseau 2019a).

Ranskan ratamaksun perusmaksu on jaoteltu kolmeen osaan. Osat ovat dieselvetoisen liikenteen maksu, sähkövetoisen liikenteen maksu ja rojalti sähkövetoisuuden tarjonnasta ja jakelusta. Ensimmäisen edellä mainitun maksun yksikkönä toimii bruttotonnikilometrit. Kaksi jälkimmäistä osaa kohdistetaan vain sähkövetoiselle liikenteelle. (SNCF Réseau 2019b, 117-118)

#### 4.2.1 Ranskan ratamaksun perusmaksuun sisällytettävät kustannukset

SNCF Réseau (2019a, 3) verkkoselostuksen maksut-liitteessä viitataan Euroopan komission rautatiemarkkinadirektiiviin 2012/34/EU kustannuksia määrittäessä. Ratamaksun perusmaksuun sisällytetään vain ja ainoastaan suoraan liikenteestä aiheutuneet kustannukset. Kustannukset on jaoteltu niin, että sähkövetoisen liikenteen aiheuttamia kustannuksia ei peritä dieselvetoisen liikenteen perusmaksussa. Dieselvetoisen liikenteen maksuun sisällytettäviin kustannuksiin lukeutuvat radan, kytkinlaitteiden, sekä merkinanto- ja ajojohtojärjestelmän kustannukset, operatiivisen liikenteenohjauksen kustannukset sekä kiskojen, vaihteiden ja risteysten korvausinvestointien kustannukset. Korvausinvestointien kustannukset on simuloitu uusimisaikataulujen perusteella, sillä yksittäisen vuoden korvausinvestointikustannusten sisällyttäminen ei onnistu niiden jaksollisen luonteen vuoksi. (SNCF Réseau 2019a, 7-8)

Sähkövetoisen liikenteen kustannuksiin sisällytetään kaikki sähkölaitteistoista aiheutuvat kunnossapito- ja korvausinvestointikustannukset. Sähkövetovoiman ja jakelun rojaltiin sisällytetään kaikki kustannukset, jotka aiheutuvat sähköenergian toimittamisesta. Rojaltin tarkoituksena on kompensoida häviöt sähköjärjestelmissä. (SNCF Réseau 2019b, 118)

#### 4.2.2 Ranskan kustannusfunktion ekonometrinen mallinnus

Ranskan ratamaksun perusmaksun määrittelytapaa on tarkasteltu vuonna 2014 julkaistussa tieteellisessä artikkelissa, missä se on kuvattu yksityiskohtaisesti (Silavong et al. 2014). Kustannusfunktio oletetaan muotoon  $Y = f(Q, P, X)$ , missä  $Y$  kuvaa lyhyen aikavälin kustannuksia,  $Q$  tuotantovektoria,  $P$  panosten hintavektoria ja  $X$  kiinteiden panosten vektoria, eli infrastruktuurin ominaisuuksia. Lyhyen aikavälin kustannusfunktio voidaan saattaa muotoon  $Y = f(Q, X)$ , sillä materiaalin, työvoiman ja energian hintojen oletetaan olevan lyhyen aikavälin tarkastelussa yhtenäisiä koko rautatieverkon tasolla. (Silavong et al. 2014, 2).

Tarkennetuksi funktiomuodoksi on valittu Box-Cox funktiomuoto, jotta välteetään tuotantotekniikan ja joustojen vakioomuotoisuuden rajoitukset. Artikkelin mallissa käytetyssä funktiossa käytetään kahta erillistä Box-Cox parametriä, theetaa ( $\theta$ ) riippuvalle muuttujalle ja lambdaa ( $\lambda$ ) itsenäisille muuttujille. Funktion lopullinen muoto on kuvattu kaavassa 19. (Silavong et al. 2014, 3)

$$\forall i, Y_i^{(\theta)} = \alpha + \sum_{k=1}^{p_1} \beta_k X_i^k^{(\lambda)} + \sum_{k=p_1+1}^p \gamma_k Z_i^k + \varepsilon_i \quad (19)$$

missä  $i$  on rataosan numero,  $Y_i$  on rataosan  $i$  kustannukset. Kaavassa itsenäisistä muuttujista  $p_1$  muuttujat ( $X$ ) Box-Cox muunnetaan ja  $p - p_1$  dummy-muuttujat ( $Z$ ) jätetään muuntamatta. Mallin virheiden  $\varepsilon_i$  oletetaan olevan itsenäisiä ja jakautuneet tasaisesti noudattaen normaalijakaumaa. Mallin parametrit ovat estimoitu käyttäen suurimman uskottavuuden estimointimenetelmää. Kustannusjoustoksi liikenteen suhteen esitetään kaava 20. (Silavong et al. 2014, 3)

$$\varepsilon_i = \frac{\partial \ln Y_i}{\partial \ln Q_i} = \beta_Q Q_i^\lambda Y_i^{-\theta} \quad (20)$$

missä  $Q_i$  on tuotantovektori, joka koostuu liikennöintiä kuvaavista muuttujista. Tuotantovektori  $Q_i$  koostuu siis kaavassa 20 muodossa  $X_i^k$  esitellyn vektorin liikennöintiä kuvaavista muuttujista, joita ovat esimerkiksi bruttotonnikilometrit.

Kaavasta 20 voidaan edelleen johtaa rajakustannus kertomalla jousto rataosan keskimääräisellä kustannuksella. Lopputuloksena on rataosan  $i$  rajakustannus, joka on esitetty kaavassa 21. (Silavong et al. 2014, 3)

$$MC_i = \epsilon_i \frac{Y_i}{Q_i} = \beta_Q Q_i^{\lambda-1} Y_i^{1-\theta} \quad (21)$$

Ekonometrisessa mallinnuksessa käytetään kolmea erityyppistä dataa: kustannusdataa, liikennöintidataa ja teknisiä ominaisuuksia kuvaavaa dataa. Malleja rakennetaan viisi eri kustannuskategoriaita varten. Kustannuskategoriat ovat radan ylläpito, vaihteiden ja risteysten ylläpito, merkinantojärjestelmän ylläpito, ajojohtojärjestelmän ylläpito sekä liikenteenohjaus. Riippuva muuttuja on malleissa ylläpitokustannus jaettuna rataosan pituus, eli euroa per ratakilometri. Ainoa eriävä mallinnus on vaihteiden ja risteysten ylläpidon malli, jonka kustannusfunktiossa riippuva muuttuja on kustannus per laite. (Silavong et al. 2014, 4)

Liikennedatana, eli itsenäisinä muuttujina, käytetään liikennemääriä ja junien lukumäärää. Junien liikennöinti esitetään bruttotonnikilometreinä tai junakilometreinä ja se on jaoteltu liikennöintityypin (pikajunat, lähijunat, tavarajunat yms.) ja segmentin (henkilö- ja tavaraliikenne) mukaan. Infrastruktuurin teknisiä ominaisuuksia kuvaavana datana käytetään rataosan pituutta, kiskojen lukumäärää, vaihteiden pituuksia, vaihteiden ja risteysten lukumääriä, kiskon asennustyyppiä, kiskojen painoa, signaalilaatikon tyyppiä, sähköistyksen tyyppiä, ratapölkkyjen ikää ja raiteiden ikää. Lisäksi ainoa suorituskykyyn liittyvä muuttuja on rataosan maksiminopeus. Jokaiseen viiteen malliin on valittu kustannuserään vaikuttavat teknisiä ominaisuuksia kuvaavat muuttujat, eli kaikki edellä mainitut muuttujat eivät ole jokaisessa kustannusfunktiossa mukana. (Silavong et al. 2014, 4) Mukana mallinuksissa on myös maantieteellisiä ja ajan heterogeenisyyttä kuvaavia muuttujia. Viiden eri mallin muuttujat ovat koostettuna taulukoon 7 seuraavalle sivulle.

Taulukko 7. Ranskan ekonometristen mallien muuttujat (Silavong et al. 2014, 4)

Malli	Riippuva muuttuja	Itsenäiset muuttujat	Dummy-muuttujat
Radan ylläpito (Malli 1)	Kustannus per ratakilometri	Rahtiliikennetonnit, henkilöliikennetonnit, rataosan ikä, rata- pölkkyjen ikä, maksiminopeus	Suurnopeuslinja, ilmasto- tyypit
Vaihteiden ja risteysten ylläpito (Malli 2)	Kustannus per vaihde/risteys	Rahtiliikennetonnit, henkilöliikennetonnit, vaihteen/risteyksen ikä, vaihteen/risteyksen pituus	vuosi 2007, vuosi 2008
Merkinantojärjestelmän ylläpito (Malli 3)	Kustannus per kilometri	Signaalilaatikon ikä, liikennevirta junien lukumääränä, signaalilaatikoiden lukumäärä, vaihteiden ja risteysten lukumäärä, signaalilaatikon tyyppin prosenttiosuus	Signaalilaatikon tyypit, vuosi 2007, vuosi 2008
Ajojohtojärjestelmän ylläpito (Malli 4)	Kustannus per kilometri	Liikennevirta junien lukumääränä, nopeusrajoitus	1.5kV, 25kV, vuosi 2007, vuosi 2008
Liikenteenohjaus (Malli 5)	Kustannus per keskus	Liikennevirta junien lukumääränä, vaihteiden, risteysten ja signaalilaatikoiden lukumäärä	Signaalilaatikoiden tyypit

Edellä esiteltyjen mallien selitysasteet ja kustannusjoustop on koostettu taulukoon 8. Mallien 1 ja 2 kohdalla on esitetty henkilöliikenteen ja rahtiliikenteen kustannusjoustop. Kustannusjoustop vaihtelevat välillä 0,08-0,65, joka on suurempi vaihteluväli kuin kirjallisuudessa yleisesti huomattu 0,1-0,5 (Silavong et al. 2014, 6). Tämä vaihtelu ei kuitenkaan ole yllättävää, sillä eri osa-alueet riippuvat eri suhteessa liikenteestä. Vaihteluvälin pienentämiseksi Silavong ja muut (2014, 6) esittävät globaalin kustannusfunktion estimoinnin, eli kaikkien mallinnuksen muuttujien yhdistämisen yhteen kustannusfunktion. Taulukossa selittävyysasteet ovat kansainvälisessä linjassa, sillä ne asettuvat välille 0,32-0,75.

Taulukko 8. Mallinnusten selitysasteet ja kustannusjoustop (Silavong et al. 2014, 6)

Mittari	Malli 1	Malli 2	Malli 3	Malli 4	Malli 5
$R^2$	0,39	0,38	0,52	0,32	0,75
Kustannusjousto	0,57 / 0,17	0,57 / 0,17	0,48	0,08	0,65

Silavongin ja muiden (2014) tekemää yllä esiteltyä mallinnusta hyödynnettiin ratamaksujen määrittelyssä vuoteen 2016 asti, jolloin asiasta teetettiin toinen jatkok tutkimus. Tämän tutkimuksen menetelmämuistio on saatavilla ainoastaan ranskan kielellä, joten sitä ei ole mahdollista esitellä tässä työssä tarkemmin (Frontier Economics 2017). Työn suosituksena kustannusfunktion Box-Cox-muodosta luovuttiin ja ratamaksun perusmaksun ekonometrisessa mallinnuksessa käytetään nykyään Translog-tuotantoteknologiaa hyödyntävää kustannusfunktiota. Translog-tuotantoteknologiaa hyödyntävän kustannusfunktion suurimmaksi eduksi luettiin sen parametrien helpompi tulkinta verrattuna Box-Cox-muunneltuun kustannusfunktioon. (SNCF Réseau 2019a, 8)

## 4.3 Suomen ratamaksun perusmaksu

Suomessa radanpitoviranomaisena ja hallinnassaan olevan rataverkon radanpittäjänä toimii Väylävirasto. Väylävirasto vastaa rautatiejärjestelmän toimivuudesta, liikennöinnin turvallisuudesta ja alueiden tasapainoisesta ja kestävästä kehityksestä. Viraston vastuulla on myös rautateiden hinnoittelu. (Väylävirasto 2020a, 12)

Suomessa ratamaksu koostuu ainoastaan perusmaksusta. Perusmaksu on jaoteltu rautatiemarkkinadirektiivin mukaisesti erillisiksi maksuiksi sähkövetoiselle ja dieselvetoiselle liikenteelle. Ekonometrista mallinnustapaa hyödynnetään tällä hetkellä vain kaikelta liikenteeltä perittävän perusmaksun perusosan määrittelyssä. Sähkövetoisien liikenteen lisämaksun laskennassa hyödynnetään vähennyyslaskumenetelmää. (Väylävirasto 2020a, 75-76)

### 4.3.1 Suomen ratamaksun perusmaksuun sisällytettävät kustannukset

Ratamaksun perusmaksun ekonometrisen mallinnuksen kustannusaineisto koostuu kunnossapitokustannuksista ja korvausinvestointikustannuksista. Kustannusdatasta on kuitenkin suodatettu pois sähkövetoisien liikenteen kunnossapitokustannukset, sillä ekonometrinen mallinnus toteutetaan vain dieselvetoiselle liikenteelle. Taulukossa 9 on eritelty tarkemmin kustannusdatan sisältämät kustannuserät ja siitä pois suodatettavat sähkökunnossapidon aiheuttamat kustannukset. (Väylävirasto 2019, 10-11)

Taulukko 9. Perusmaksun laskentaan sisältyvät kustannukset (Väylävirasto 2019)

Sisältyy aineistoon	Ei sisälly aineistoon (sähkökunnossapito)	Korvausinvestointikustannukset, jotka sisältyvät aineistoon
<ul style="list-style-type: none"> <li>- päällysrakenteen kunnossapito</li> <li>- vaihteiden ja vaihteenlämmityksen kunnossapito</li> <li>- radan varusteiden ja laitteiden kunnossapito</li> <li>- siltojen kunnossapito</li> <li>- alus- ja pohjarakenteen sekä rautatiealueiden kunnossapito</li> <li>- raideliikenteen ohjaus- ja turvalaittejärjestelmien kunnossapito</li> <li>- kunnossapidon materiaalit: mm. kiskot, pölkkyt ja sepeli</li> </ul>	<ul style="list-style-type: none"> <li>- siltojen sähkötekni- ninen kunnossapito</li> <li>- sähkörata-järjestelmien kunnossapito</li> <li>- radan vahvavirtalaitteiden kunnossapito</li> <li>- radan muiden erikoisjärjestelmien kunnossapito</li> <li>- erikseen tilattavat sähkökunnossapidon työt: mm. ajo- ja kannatinlankojen ja ripustimien vaihdot</li> <li>- materiaalit: mm. ajo- ja kannatinlangat</li> <li>- ajo- ja kannatinlankojen vaihdot</li> </ul>	<ul style="list-style-type: none"> <li>- päällysrakenteen uusiminen ja puhdistus</li> <li>- kiskojen, pölkkyjen, ajolankojen ja kannatinlankojen uusiminen</li> <li>- vaihteiden uusiminen ja korjaukset</li> <li>- kiskojen ja vaihteiden hionnat</li> <li>- materiaalit: mm. kiskot, pölkkyt, sepeli ja muut rakenteelliset materiaalit</li> </ul>

Vuonna 2018 tehdyn laskennan kustannusaineisto on koostettu Väyläviraston kustannustenlaskentajärjestelmästä vuosilta 2013-2016. Kustannukset on kohdennettu rataosille rataosienumeroiden perusteella. Rataosia on kustannusaineistossa lopulta yhteensä 120.

#### 4.3.2 Suomen kustannusfunktion ekonometrinen mallinnus

Kustannusfunktion mallintamisessa itsenäisinä muuttujina toimivat liikennesuoritteet ja raidekilometrit. Liikennesuoritetiedot koostuvat tavaraliikenteestä, tavaraliikenteen päivystyksestä sekä kaukoliikenteen ja lähiliikenteen henkilöliikenteestä. Liikennesuoritemuuttuja koostuu edellä mainituista tekijöistä koostetusta bruttotonnimuuttujasta. Muuttuja sisältää vuotuiset junakaluston ja lastin yhteispainon bruttotonnimäärät rataosittain jaoteltuna. Raidekilometri- muuttuja on koostettu rataosien linjapituuksista ja moniraiteisuuksista. Raidekilometrit on laskettu kertomalla rataosan linjapituus raiteiden lukumäärällä.

Suomen mallinuksissa käytössä oleva kustannusfunktio on muodoltaan Cobb-Douglas-kustannusfunktio. Mallinuksessa on siis käytössä logaritimuunnos kaikille muuttujille. Mallinnettava kustannusfunktio on lineaarisessa muodossa esitetty kaavassa 22.

$$\ln C_i = \alpha + \beta_{rd\_km} \ln rd\_km_i + \beta_{brt} \ln brt_i + \varepsilon_i \quad (22)$$

missä  $C_i$  on rataosan  $i$  kustannukset.  $\alpha$ ,  $\beta_{rd\_km}$  ja  $\beta_{brt}$  ovat regressioanalyysissä estimoitavat parametrit.  $rd\_km_i$  on rataosan  $i$  raidekilometrien määrä, joka on rataosan pituus kerrottuna raiteiden lukumäärällä.  $brt_i$  on rataosan liikennesuoritteiden määrä bruttotonneina.  $\varepsilon_i$  on estimoidun mallin residuaali, eli virhetermi.

Kaavasta 22 on johdettu rajakustannus, eli suoraan liikennesuoritteesta aiheutuva kustannus osittaisderivoimalla kyseinen kustannusfunktio liikennesuoritteen suhteen. Tämä johdettu osittaisderivaatta ilmaisee, kuinka radanpidon kustannukset muuttuvat yhden junaliikennesuoritteen seurauksena. Rajakustannusfunktio on esitetty kaavassa 23. (Väylävirasto 2019, 15)

$$MC_i = \frac{\partial \ln C_i}{\partial \ln brt_i} = \beta_{brt} \frac{e^{\alpha + \beta_{rd\_km} \ln rd\_km_i + \beta_{brt} \ln brt_i + \frac{\sigma^2}{2}}}{brt_i rt\_km_i} \quad (23)$$

missä  $\sigma^2$  on mallinnuksen virhetermin varianssin estimaatti. Tästä seuraa rajakustannuksen estimaatit erikseen jokaiselle aineiston sisältämälle rataosalle. Rataosien rajakustannukset vaihtelevat merkittävästi seurauksena huomattavista eroavaisuuksista raidepituuksissa, liikennesuoritteissa ja kustannuksista. Tästä seuraten lopullinen rajakustannus, eli ratamaksun perusmaksu, on laskettu painottamalla rataosakohtaisia rajakustannuksia rataosien liikennesuoritteiden määrällä (kaava 24).

$$MC = \frac{\sum_{i=1}^n brtkm_i MC_i}{\sum_{i=1}^n brtkm_i} \quad (24)$$

missä  $brtkm_i$  on bruttotonnikilometrit rataosittain. Bruttotonnikilometrit saadaan rataosittain laskettua kertomalla rataosan liikennesuoritteet rataosan raidekilometreillä.

Mallin estimointi on toteutettu yhdistetyllä pienimmän neliösumman menetelmällä. Mallinnuksen  $R^2$ -luku, eli selitysaste on 0,564. Ratamaksun perusmaksun tasoksi saatiin 0,1259 senttiä bruttotonnikilometriä kohden. Mallinnus on kuitenkin ongelmallinen sen sisältämän heteroskedastisuuden vuoksi, joten perusmaksulle ei voida määrittää luottamusvälejä, eikä mallin parametrien nollasta poikkeavuutta voida testata. (Väylävirasto 2019, 19)



## 5 Suomen rautatieinfrastruktuurin käytön rajakustannusten mallinnus

Ratainfrastruktuurin käytön rajakustannusten mallintamiseksi laaditaan aineisto, joka koostuu Väyläviraston rataverkkoa, rataverkon liikennöintiä ja radanpitoa kuvaavista tiedoista. Aineiston käsittely ja mallinnuksen menetelmät toteutetaan mukaillen luvussa 2 esiteltyjä ratamaksun määrittelyä lainsäädännöllisiä reunaehtoja. Tässä luvussa esitellään käytössä oleva laskenta-aineisto sekä menetelmät, joilla rautatieinfrastruktuurin rajakustannusten mallinnus toteutetaan. Lisäksi tässä luvussa toteutetaan rautatieinfrastruktuurin käytön rajakustannusten mallinnukset, niiden jatkojalostaminen ja vertailu.

Huomattavaa on, että tässä työssä käytössä oleva ja tässä esiteltävä aineisto ei vastaa täydellisesti minkään vuoden virallisen ratamaksun laskennassa käytettyä aineistoa. Huomattavin eroavaisuus edelliseen luvussa 4.3.2 esiteltyyn Suomen kustannusfunktion ekonometriseen mallinnukseen on rataosajako. Aiemmin rataverkko oli jaettu laskentaa varten 120 rataosaan, kun taas tässä työssä käytössä oleva aineisto sisältää 108 rataosaa. Toinen eroavaisuus on paneeliaineiston aikasarja, joka on tässä työssä vuosilta 2013–2018, kun se aiemmassa mallinnuksessa oli vuosilta 2013–2016.

### 5.1 Laskenta-aineisto

Mallinnuksessa käytettävä data koostuu rataosille kohdennetuista radanpidon kustannuksista (sisältäen kunnossapidon ja korvausinvestointien kustannukset), rataverkon ominaisuuksista rataosittain eriteltyinä sekä vuosittaisista liikennesuoritteista rataosittain eriteltyinä. Käytössä olevan paneeliaineiston tiedot on kerätty vuosilta 2013–2018 ja se sisältää yhteensä 648 havaintoa, 6 vuoden havainnot 108 rataosalta.

Mallinnuksessa itsenäiset muuttujat kuvaavat liikennöintiä ja rataverkon ominaisuuksia. Rataverkon ominaisuuksia kuvaavat tiedot koostuvat rataverkon rataosajakoista, rataosien linjapituuksista, rataosien raidepituuksista, rataosien raiteiden lukumääristä sekä rataosien sähköistyksestä. Näistä tiedoista on koostettu 108 rataosalle raidekilometrimuuttuja, jota hyödynnetään laskennassa. Raidekilometrit ovat laskettu kertomalla rataosan pituus raiteiden lukumäärällä ja näin saadaan rataosan kokonaisraidekilometrit. Suomen rataverkko koostuu 5926 ratakilometristä, joista suljettuja ja ei-liikennöityjä ratoja on 275 ratakilometriä (Väylävirasto 2020b). Liikennöidyn rataverkon pituus on siis 5651 ratakilometriä, josta laskennassa on mukana 5639,36 ratakilometriä. Mallinnuksessa hyödynnettävä rataosajako kattaa 99,7% Suomen koko liikennöidyn rataverkon laajuudesta. Suomen rataverkko vuonna 2018 on kuvattu kokonaisuudessaan kuvassa 3 (Liikennevirasto 2018). Kuvassa punertavat rataosat ovat sähköistetyt ja harmaat rataosat ovat liikenteeltä suljettuja rautateitä.



Kuva 3. Suomen rataverkko (Liikennevirasto 2018)

Rataverkon ominaisuuksista pyrittiin keräämään lisätietoa vaihteiden, tasoisteusten ja siltojen lukumäärästä. Nämä tiedot koostettiin sijoittamalla ne rataverkolle kilometripaalujen mukaisesti ja siitä eteenpäin kohdistamalla ne rataosille. Tämä sijoittelumenetelmä ei kuitenkaan tuottanut täydellisiä lukumääriä johtuen sijaintitietojen epätäydellisyyksistä. Tuloksetkaan eivät olleet näillä muuttujilla parempia, joten tästä syystä nämä kyseiset lisämuuttujat päätettiin jättää mallinnoista pois.

Liikennöintiä kuvaamaan kerätyt liikennesuoritetiedot on koostettu rataosittain junakaluston ja lastin yhteispainosta. Tiedot ovat saatavilla suoritetyyppiä oteluna seuraavasti:

- tavaraliikenne,
- tavaraliikenne – päivystys (vaihtotyöliikenne),
- henkilöliikenne – kaukoliikenne ja
- henkilöliikenne – lähiliikenne.

Mallinuksissa liikennesuoritteita ei kuitenkaan eritellä edellä mainitun jaotellun mukaisesti, vaan ne on summattu yhteen. Tästä yhteenlaskusta muodostuu bruttotonnimuuttuja, joka kuvastaa kokonaisliikennöintiä. Liikennöinti on saatavilla kullekin vuodelle 108 rataosan jaon mukaisesti.

Mallinnuksen riippuva muuttuja, eli kustannustiedot on koostettu korvausinvestointi- ja kunnossapitokustannuksista. Tämän työn liitteessä 2 esitellyt ratamaksun laskennan kustannusaineiston kielletyt erät on poistettu materiaalista. Myös sähkökunnossapidon kustannukset on jätetty näiden laskelmien materiaalien ulkopuolelle. Lopulta aineistoon on sisällytetty seuraavat kunnossapitokustannuserät korvausinvestointikustannusten lisäksi:

- päällysrakenteen kunnossapito,
- vaihteiden ja vaihteenlämmityksen kunnossapito,
- radan varusteiden ja laitteiden kunnossapito,
- siltojen kunnossapito,
- alus- ja pohjarakenteen sekä rautatiealueiden kunnossapito,
- raideliikenteen ohjaus- ja turvalaitejärjestelmien kunnossapito sekä
- kunnossapidon materiaalit.

Edellä mainittua erittelyä ei ole hyödynnetty mallinuksissa, vaan nämä kustannuserät on summattu yhteen. Tästä seurauksena mallinuksissa on eurot yhteensä-muuttuja, joka sisältää jokaisen rataosan vuosittaiset kokonaiskustannukset. Taulukossa 10 on kuvattu edellä mainittujen muuttujien tilastolliset tunnusluvut.

*Taulukko 10. Muuttujien tilastolliset tunnusluvut*

<b>Muuttuja</b>	<b>Keskiarvo</b>	<b>Keskihajonta</b>	<b>Minimi</b>	<b>Maksimi</b>
Kustannukset	1 572 000	2 868 000	37 000	29 875 000
Bruttotonnit	7 423 000	10 511 000	0	84 730 000
Raidekilometrit	58,4	40,0	5	183

Tilastollisista tunnusluvuista nähdään, että liikennöintiä kuvaava bruttotonnimuuttuja sisältää nolla-arvoja. Nämä rivit aiheuttanevat ongelmia mahdollisten parametrien muunnosten kanssa. Rivien poistaminen ja lukumäärä dokumentoidaan mallinnusten yhteydessä.

## 5.2 Mallinnuksen menetelmät

Tämän työn taustalla toteutettiin huomattava määrä erilaisia parametreja ja estimointimenetelmiä sisältämiä mallinuksia, joiden tulokset eivät olleet kuitenkaan niin merkittäviä, että niitä tulisi tässä työssä esitellä. Tähän työhön valikoituneet mallinuksissa käytettävät muunnokset ovat aiemmin esitellyt logaritmuunnos, Box-Cox-muunnos sekä Box-Tidwell-muunnos sekä riippuvalle että itsenäisille muuttujille. Nämä valikoituivat menetelmiksi Euroopan talousalueen jäsenmaiden ratamaksukäytäntöjen, mallinnustapojen yksinkertaisuuden sekä aiheeseen liittyvien tutkimusten perusteella. Taulukossa 11 seuraavalla sivulla on esitelty alkuvaiheessa rakennettavat mallinnukset.

Taulukko 11. Rautatieinfrastruktuurin rajakustannusten mallinnukset

Malli	Itsenäiset muuttujat	Muunnos
Malli 1	Bruttotonnit, raidekilometrit	Logaritmimuunnos
Malli 2	Bruttotonnit, raidekilometrit	Box-Cox muunnos
Malli 3	Bruttotonnit, raidekilometrit	Box-Tidwell muunnos

Kaikissa mallinuksissa on muuttujina vain liikennöinnin bruttotonnit ja raidekilometrit. Vaihteiden, tasoristeysten ja siltojen lukumääriä kokeiltiin työn taustalla lisämuuttujina, mutta ne päätettiin jättää mallinnusten ulkopuolelle. Syynä muuttujien ulosjättämiselle on se, että lisämuuttujille ei löytynyt mallinuksissa tilastollista merkitsevyyttä, eikä muuttujien rataosittaiset lukumäärät ole täydellisiä. Lisäksi mallinuksessa arvostetaan yksinkertaisuutta, mikäli yksinkertaisella mallilla saavutetaan monimutkaisemman mallin kanssa samankaltaiset tulokset.

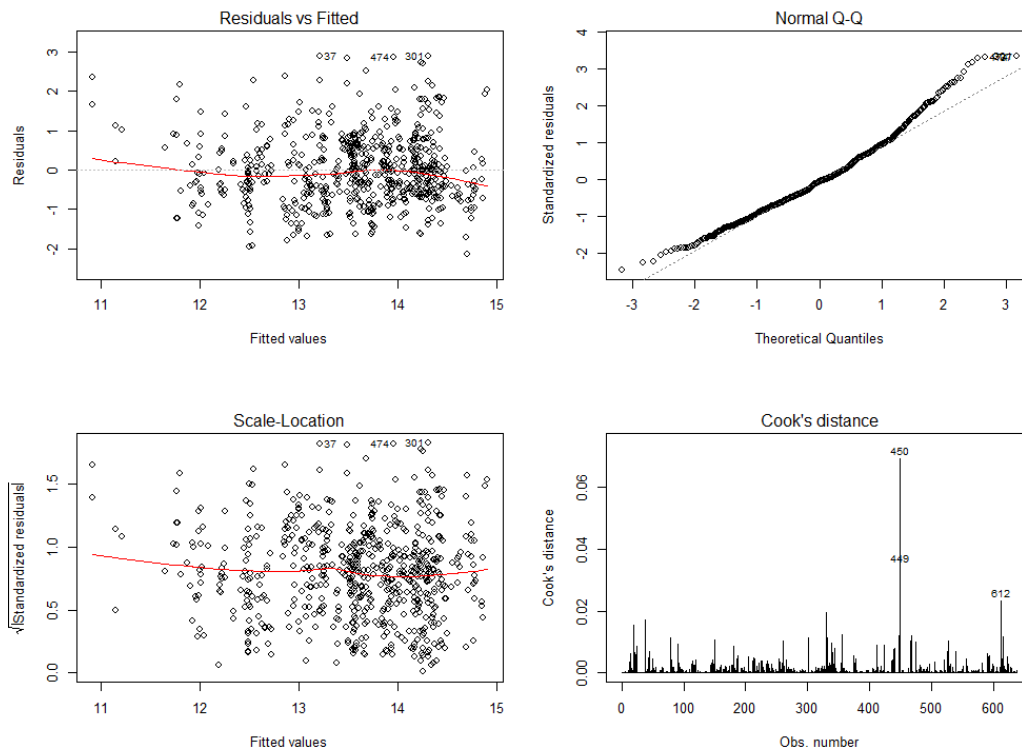
Malleja rakentuu lopulta tässä vaiheessa kolme kappaletta eri muuttujien parametrimuunnoksilla. Ensimmäinen malli on toteutettu muuttujien logaritmimuunnoksilla, toinen malli muuttujien Box-Cox-muunnoksella ja kolmas malli muuttujien Box-Tidwell-muunnoksilla. Mallinnusten havaintojen lukumäärä on 635, sillä nolllaliikenteiset rataosat, joita on yhteensä 10, on poistettu aineistosta muuttujien muunnosten vuoksi.

Ensimmäisessä mallissa on oletettu rautatieinfrastruktuurin kustannusten noudataavan Väyläviraston (2019) esimerkin mukaisesti seuraavaa logaritmimuunneltua Cobb-Douglas-kustannusfunktioimuotoa (kaava 25):

$$\ln C_i = \alpha_0 + \beta_{brt} x_i + \beta_{rd\_km} l_i + \varepsilon_i \quad (25)$$

missä  $C_i$  on rataosan  $i$  kustannukset. Kustannusfunktion itsenäisten muuttujien pienet kirjaimet viittaavat muuttujien logaritmimuotoon, jolloin  $x_i = \ln brt_i$  ja  $l_i = \ln rd\_km_i$ .  $brt_i$  on rataosan liikennesuoritteiden määrä bruttotonneina ja  $rd\_km_i$  on rataosan  $i$  raidekilometriä, joka on rataosan pituus kerrottuna raiteiden lukumäärällä. Kustannusfunktion kertoimet  $\beta_{brt}$  ja  $\beta_{rd\_km}$  ovat regressioanalyysissä estimoitavat parametrit. Mallinuksen vakioterminä on  $\alpha_0$  ja  $\varepsilon_i$  on estimoidun mallin residuaali, eli jäännöstermi.

Logaritmimuunnellun mallin yhteenveto löytyy liitteestä 3. Mallin selitysaste, eli korjattu  $R^2$ -luku on 0,43. Mallin selitysaste on siis suuruudeltaan samalla tasolla muiden maiden rautatieinfrastruktuurin rajakustannusten mallinuksien kanssa. F-testin arvo on mallinuksessa 238,1 ja p-arvo on lähes 0. Mallin läpäisee siis F-testin ja itsenäiset muuttujat selittävät riippuvan muuttujan vaihtelua hyvin. Ensimmäisen mallin jäännöstermien ja poikkeavien havaintojen kuviot on esitetty kuvassa 4.



Kuva 4. Mallin 1 visualisointi

Kuvan 4 vasemman yläreunan kuvio, jäännöstermit vastaan mallinnetut arvot, kuvastaa mallinnuksen lineaarisuutta. Kuvioista nähdään, että mallinnus on lineaarinen, sillä punainen viiva on kutakuinkin suora, eikä jäännöstermeissä esiinny selkeitä ryppäitä. Oikean yläreunan kvanttilikuvio (Q-Q-plot) kuvaa normaalijakauman nousevana suorana ja otoksen jäännöstermien arvot pisteinä. Kvanttilikuvioista havaitaan S-muotoinen käyrä jäännöstermien pisteissä, joten mallinnuksen virheissä on liian vähän tai liikaa suuria virheitä molempiin suuntiin. Kvanttilikuvion perusteella voidaan todeta, että jäännöstermien normaalijakautuneisuudessa on puutteita mallinnuksessa, eikä täten malli täytä pienimmän neliösumman estimointimenetelmän oletuksia täydellisesti. Tätä päätelmää tukee myös Jarque-Bera-testin tulos 48,9, jonka tulisi olla normaalijakautuneessa tilanteessa alle arvon 5,99.

Vasemman alareunan kuvio on esittää jäännöstermien jakaumaa mallinnettujen arvojen välillä. Kuvioista nähdään, että jäännöstermien pisteet jakautuvat poikiviivan molemmiin puolin suhteellisen tasaisesti ilman selkeitä ryppäitä. Kuvion jakaumasta voidaan päätellä, että mallinnuksen varianssi on suhteellisen tasainen mallinnettujen arvojen välillä, eli mallinnus on homoskedastinen. Tätä päätelmää tukee myös Breusch-Paganin testin p-arvo 0,581, joka viittaa mallin heteroskedastisuus-hypoteesin hylkäämiseen. Kuvan 4 oikean alareunan kuvio kertoo poikkeavien havaintojen esiintyvyyden Cookin etäisyysmitan avulla. Kuvion mukaan mallissa on kolme havaintoa, jotka erottuvat selkeästi muista havainnoista. Autokorrelaatiotestien tulosten analysointi ei ole näissä mallinnuksissa merkityksellistä, sillä kyseessä ei ole aikasarja-aineisto. Mallinnuksen oletusten toteutumisen testit on koostettu taulukkoon 12 ja ne löytyvät myös liitteestä 3. Varianssin inflaatiotekijä kertoo myös, että mallissa ei esiinny muuttujien välistä multikollineaarisuutta. Varianssin inflaatiotekijöiden arvot löytyvät liitteestä 3.

Taulukko 12. Mallin 1 testien tulokset

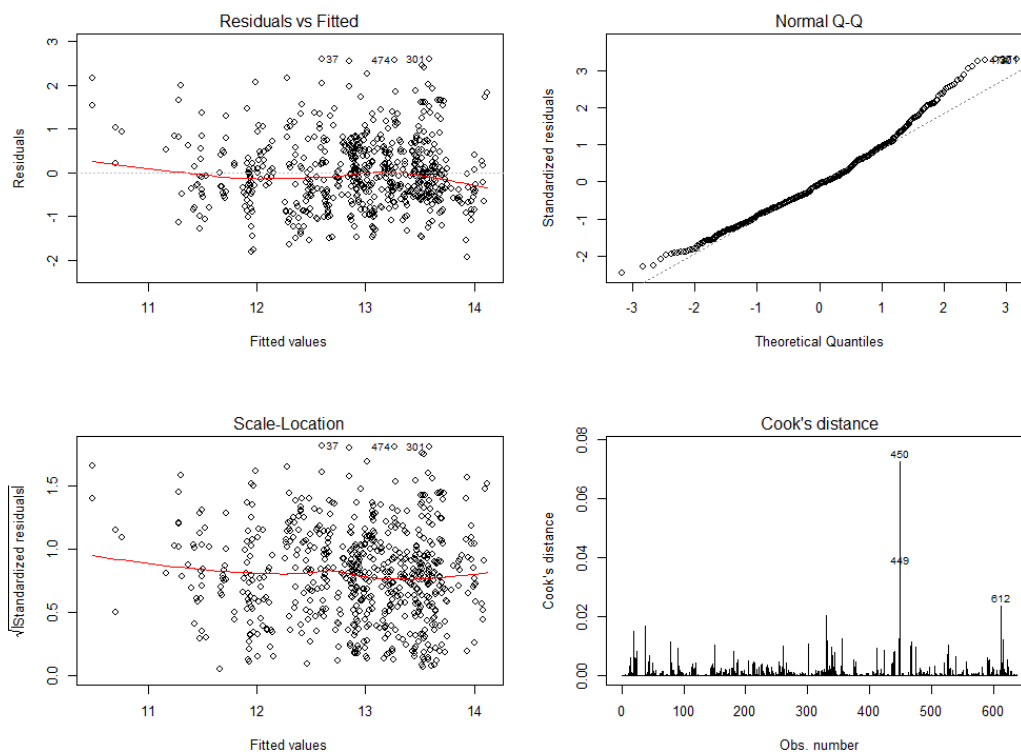
Testi	Testisuure	p-arvo
Jarque-Bera-testi	48,9	$2,5 \times 10^{-11}$
Breusch-Pagan-testi	1,09	0,6
F-testi	238,1	$2,2 \times 10^{-16}$

Toinen mallinnus on Box-Cox-muunneltu malli. Mallin kustannusfunktio oletetaan Silavongin ja muiden (2014) tutkimuksen mukaisesti muotoon (kaava 26):

$$C_i^{(\lambda)} = \alpha_0 + \beta_{brt} X_i^{(\lambda)} + \beta_{rd\_km} L_i^{(\lambda)} + \varepsilon_i \quad (26)$$

missä  $C_i$  on rataosan  $i$  kustannukset,  $X_i$  on rataosan  $i$  liikennöinti bruttotonneina ja  $L_i$  on rataosan  $i$  raidekilometripituus. Muuttujien yläindeksi viittaa muuttujien Box-Cox-muunnokseen. Muunnosparametri estimoidaan Shapiro-Wilk-testin avulla. Kustannusfunktion kertoimet  $\beta_{brt}$  ja  $\beta_{rd\_km}$  ovat regressioanalyysissä esitöitavat parametrit. Mallinnuksen vakioitermi on  $\alpha_0$  ja  $\varepsilon_i$  on estimoidun mallin residuaali, eli virhetermi.

Mallissa on sama muunnosparametri  $\lambda$  sekä itsenäisille että riippuvalle muuttujalle. Muunnosparametrin arvo on  $-0,0073$  ja se on estimoitu Shapiro-Wilk-testin avulla. Mallin selittävyysaste, eli korjattu  $R^2$ -luku on sama kuin ensimmäisellä logaritimuunnellulla mallilla, eli  $0,43$ . F-testin arvo on toisessa mallinnuksessa  $239,5$  ja p-arvo on lähes  $0$ . Mallinnus läpäisee siis F-testin ja itsenäiset muuttujat selittävät riippuvan muuttujan vaihtelua hyvin. Toisen mallin jäännöstermien ja poikkeavien havaintojen kuviot on esitetty kuvassa 5.



Kuva 5. Mallin 2 visualisointi

Kuvasta 5 huomataan, että Box-Cox-mallissa on kutakuinkin samankaltaiset jäännöstermikuviot kuin ensimmäisellä mallilla. Tämä selittyy sillä, että muunnosparametri  $\lambda$  on niin lähellä nollaa, että muunnos on lähes vastaava logaritmuunnoksen kanssa. Mallin 2 ongelma on siis sama kuin logaritmuunnollun, eli jäännöstermien normaalijakautuneisuusoletus ei täyty. Kvantiilikuviossa esiintyy S-muotoinen jakauma jäännöstermien välillä. Jarque-Bera-testisuureen arvo on Box-Cox-mallilla 44,62, joka vahvistaa jäännöstermien normaalijakautumattomuuden. Poikkeavien havaintojen kuviosta erottuu samat havainnot kuin ensimmäisellä mallilla. Mallin lineaarisuus ja homoskedastisuus ovat hyvällä tasolla. Homoskedastisuudesta kertoo sekä kuvio jäännöstermien jakaumasta mallinnettujen arvojen välillä, että Breusch-Paganin testin p-arvo 0,4892, joka hylkää mallin heteroskedastisuushypoteesin. Mallin 2 oletusten testien tulokset on koostettu taulukkoon 13 ja ne löytyvät myös liitteestä 3. Multikollineaarisuus ei myöskään ole mallissa ongelma, jonka kertoo muuttujien varianssin inflaatiotekijät. Mallinnuksen muuttujien varianssin inflaatiotekijät löytyvät liitteestä 3.

Taulukko 13. Mallin 2 testien tulokset

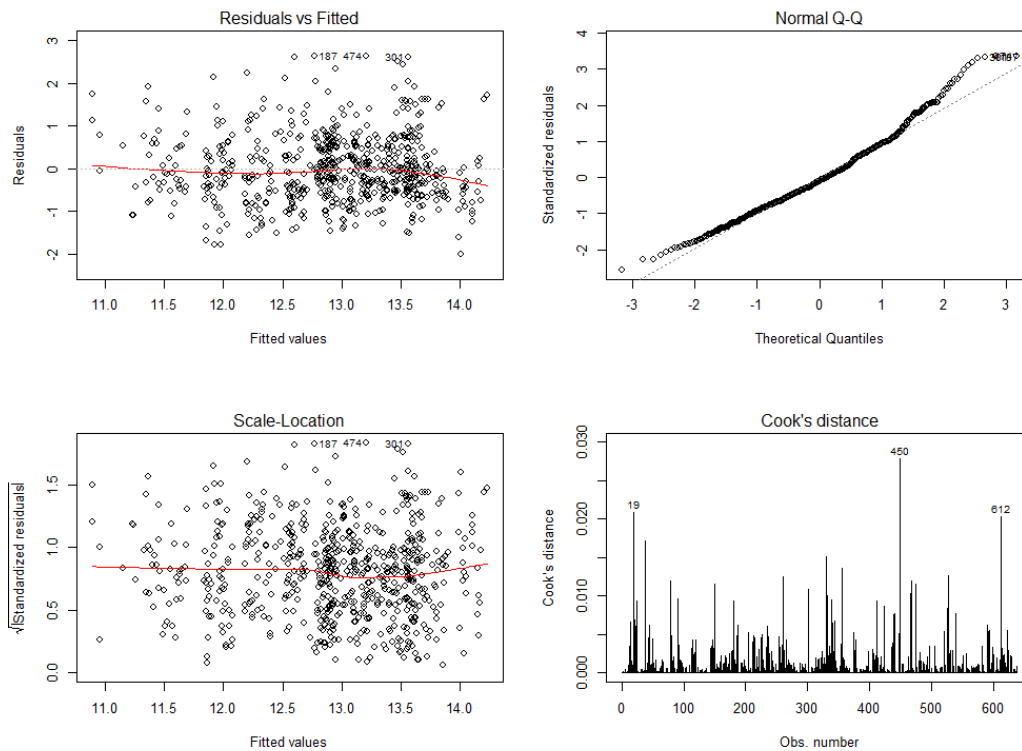
Testi	Testisuure	p-arvo
Jarque-Bera-testi	44,6	$2,0 \times 10^{-10}$
Breusch-Pagan-testi	1,4	0,5
F-testi	238,1	$2,2 \times 10^{-16}$

Kolmannessa mallinnuksessa muuttujille toteutettiin Box-Tidwell-muunnos. Box-Tidwell-muunneltu kustannusfunktio esitetään muodossa (kaava 27):

$$C_i^{(\lambda)} = \alpha_0 + \beta_{brt} X_i^{(\theta_{brt})} + \beta_{rd\_km} L_i^{(\theta_{rd\_km})} + \varepsilon_i \quad (27)$$

missä on samat muuttujat kuin Box-Cox-muunnellussa mallissa. Ainoana eroavaisuutena aiempaan kustannusfunktioon on se, että jokaiselle muuttujalle esitetään oma muunnosparametri. Kustannusfunktion yläindeksit kuvaavat Box-Tidwell-muunnoksia. Muunnosparametrit on etsitty funktion suurimman uskottavuuden menetelmällä.

Itsenäisen muuttujan muunnosparametrin arvo on sama kuin edellisessä mallinnuksessa, eli -0,0073. Bruttotonniin muunnosparametri on 0,24 ja raidekilometrien muunnosparametri on 0,024. Kolmannen mallin selittävyysaste, eli korjattu  $R^2$ -luku on 0,43. F-testin arvo on mallinnuksessa 244,8 ja p-arvo on lähes 0. Mallinnus läpäisee siis F-testin ja itsenäiset muuttujat selittävät riippuvan muuttujan vaihtelua hyvin. Mallin jäännöstermien ja poikkeavien havaintojen kuviot on esitetty kuvassa 6.



Kuva 6. Mallin 3 visualisointi

Mallinnuksen jäännöstermeissä jatkuu samankaltaisuus kahden edellisen mallin kanssa. Lineaarisuus- ja homoskedastisuusoletukset täyttyvät kuvioiden perusteella. Vasemman yläreunan jäännöstermien ja mallinnettujen arvojen kuviossa punainen viiva on suhteellisen suora, eikä vasemman alareunan kuviossa ilmene selkeitä ryppäitä. Homoskedastisuuspäätelmää tukee myös Breusch-Pagan-testisuureen p-arvo 0,8, joka johtaa heteroskedastisuushypoteesin hylkäämiseen. Kuitenkin tässäkin mallissa on ongelmana jäännöstermien normaalijakautuneisuus, joka ei oikean yläreunan kuvion perusteella täyty. Kuviossa esiintyy jälleen S-muotoa jäännöstermien arvojen kuviossa. Myös Jarque-Beran normaalijakautuneisuustestin testisuureen arvo 45,9 tukee kuviosta tehtyä johtopäätelmää. Eroavaisuutta edellisiin mallinnoiksiin löytyy poikkeavien havaintojen kuviossa, jossa Cookin etäisyysmitat ovat keskimäärin matalammalla tasolla sekä sieltä erottuu eri poikkeavat havainnot. Mallin 3 oletusten testien tulokset on koostettu taulukkoon 14 ja ne löytyvät myös liitteestä 3. Multikollineaarisuudesta kertovat varianssin inflaatiotekijät löytyvät mallinnuksen osalta myös liitteestä 3. Niiden perusteella myöskään tämän mallin muuttujien välillä ei ilmene multikollineaarisuutta.

Taulukko 14. Mallin 3 testien tulokset

Testi	Testisuure	p-arvo
Jarque-Bera-testi	45,9	$1,1 \times 10^{-10}$
Breusch-Pagan-testi	0,6	0,8
F-testi	244,1	$2,2 \times 10^{-16}$



Tässä luvussa toteutetuista mallinnoista yksikään ei täyttänyt kaikkia malleille asetettuja oletuksia. Kaikissa mallinnoissa ongelmana oli jäännöstermien normaalijakautuneisuusoletuksen hylkääminen. Seuraavassa luvussa pyritään kehittämään mallinnoita jatkojalostamalla niistä yksinkertaisempaa, eli logaritmista mallinnusta.

### 5.3 Mallinnuksen jatkojalostaminen

Kaikissa aiemmin esitellyissä malleissa ongelmana on jäännöstermien normaalijakautuneisuus, joka johtaa siihen, että pienimmän neliösumman menetelmän antamat estimaatit eivät mahdollista luottamusvälien laskemista. Jäännöstermien normaalijakautuneisuusoletuksen ongelma on todennäköisesti korvausinvestointien vuosittainen suuri vaihtelu, mikä aiheuttaa kustannuksiin pistemäisiä nousuja. Jäännöstermien normaalijakautuneisuutta yritetään ratkaista seuraavilla keinoilla:

1. Malli 1a: paneeliaineiston muuntaminen poikkileikkausaineistoksi yhdistämällä aineiston aikasarjan havainnot rataosittain
2. Malli 1b: poikkeavien äärihavaintojen poistaminen
3. Malli 1c: lisäämällä korvausinvestointien suuruutta kuvaava tekemuuttuja (dummy-muuttuja)

Näistä ratkaisutavoista rakentuu kolme uutta mallinnusta, joita tarkastellaan tarkemmin tässä luvussa. Kaikkien aiemmin rakennettujen mallien samankaltaisten tulosten vuoksi valitaan jatkojalostettavaksi rautatieinfrastruktuurin rakastannusten laskemisessa jo aiemmin käytössä ollut logaritminen mallinnus, jossa muuttujat muunnetaan logaritmuunnoksella. Malleista kaksi ensimmäistä, mallit 1a ja 1b, noudattavat samaa funktiomuotoa kuin aiemmin esitelty malli 1. Mallissa 1c funktiomuoto on seuraava (kaava 28):

$$\ln C_i = \alpha_0 + \beta_{brt} x_i + \beta_{rd\_km} l_i + \beta_{korv\_inv} D_i + \varepsilon_i \quad (28)$$

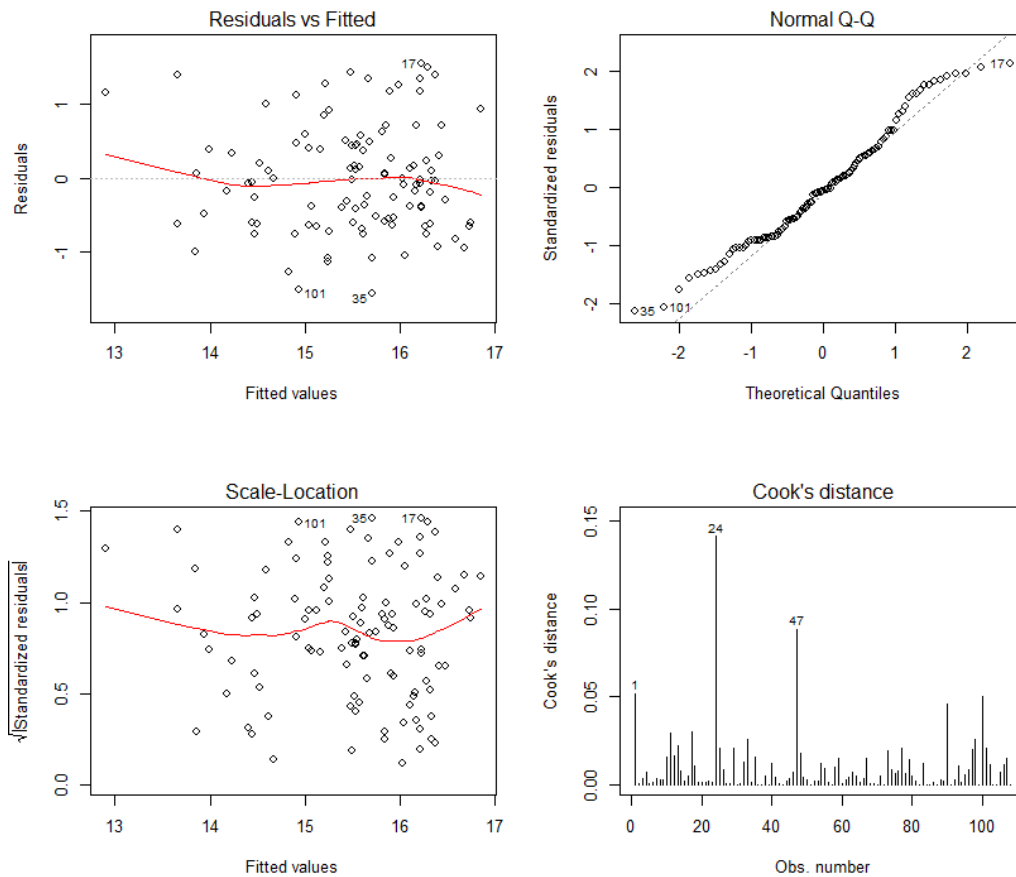
missä on korvausinvestointeja kuvaava muuttuja  $D_i$ , joka saa arvon 0, jos korvausinvestointien kokonaissumma on rataosalla yli 400 000 euroa. Muuten funktiomuoto on vastaava, kuin kaavassa 26 esitelty logaritmuunneltu kustannusfunktio.

Ensimmäisen ratkaisun mallissa paneeliaineisto muunnetaan poikkileikkausaineistoksi yhdistämällä havainnot rataosittain. Tästä seuraa, että mallissa havaintojen lukumäärä on rataosajaon mukainen lukumäärä, eli 108. Aikasarja yhdistetään summaamalla kustannukset ja bruttotonnit yhteen kuuden vuoden ajanjaksolta. Raidekilometrimuuttuja on sama kuin kaikissa aiemmissa malleissa. Tämä menetelmä tasoittaa korvausinvestointien vuosittaisen vaihtelun ongelmia mallinnoissa. Rakentuvan poikkileikkausaineiston tilastolliset tunnusluvut on kuvattu taulukossa 15.

Taulukko 15. Poikkileikkausaineiston tilastolliset tunnusluvut

Muuttuja	Keskiarvo	Keskihajonta	Minimi	Maksimi
Kustannukset	9 434 000	11 294 000	377 000	53 310 000
Bruttotonnit	44 540 000	62 973 000	20 000	465 780 000

Poikkileikkausaineistoksi yhdistetyn datan mallinnus tuotti korjatuksi selittävyysasteeksi 0,53. Tämä on korkeampi kuin aiempien mallinnusten selittävyysasteet, joten huomataan, että poikkileikkausaineiston mallinnus toimii selittävyyden osalta paremmin kuin paneeliaineiston mallinnus. F-testin arvo on mallinnuksessa 61,8 ja p-arvo on lähes 0. Mallinnus läpäisee siis F-testin ja itsenäiset muuttujat selittävät riippuvan muuttujan vaihtelua hyvin. Mallinnuksen jäännöstermien ja poikkeavien havaintojen kuviot on esitetty kuvassa 7.



Kuva 7. Mallin 1a visualisointi

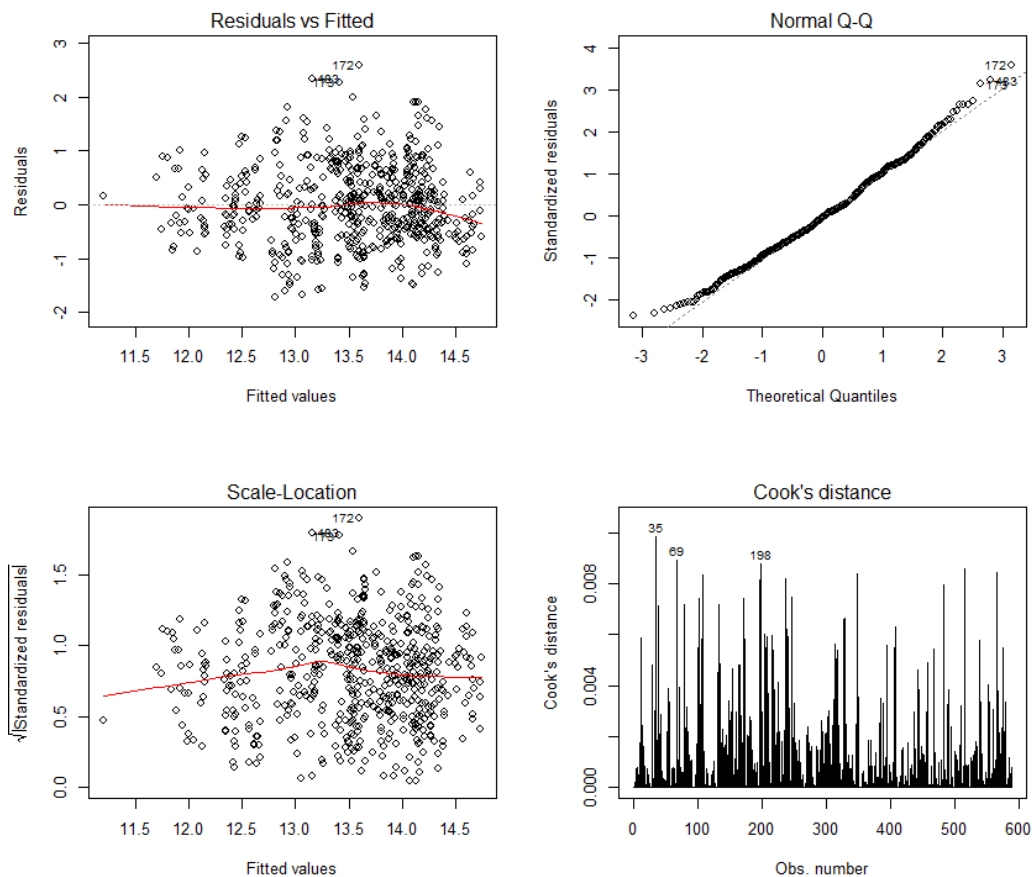
Kuvassa esitetyt kuviot näyttävät, että mallinnuksen oletukset täyttyvät lineaarisuuden, heteroskedastisuuden ja jäännöstermien normaalijakauman osalta. Vasemman yläreunan kuvio kertoo lineaarisuudesta, sillä kuvion punainen viiva ei poikkea merkittävästi suorasta. Oikean yläreunan kvantiilikuvio viittaa parempiin tuloksiin jäännöstermien normaalijakautuneisuuden osalta kuin alkuperäiset mallit, eikä jäännöstermien arvoissa näy merkittävää S-muotoisuutta. Tätä tukee myös jäännöstermien Jarque-Bera-testisuuren arvo 2,9, joka vahvistaa normaalijakautuneisuusoletuksen täyttymisen. Vasemman alareunan kuviossa ei myöskään esiinny selkeitä rypäitä sekä Breusch-Paganin testisuuren p-arvo on mallinnuksessa 0,71, eli myös homoskedastisuusoletus täyttyy. Cookin etäisyysmittojen kuvaaja näyttää kolme huomattavan suurta poikkeavuutta havainnoissa, jotka selittyvät korkealla keskihajonnalla. Huomattavin poikkeama on rataosissa 24 ja 47. Mallin 1a oletusten testien tulokset on koostettu taulukkoon 16 ja ne löytyvät myös liitteestä 3. Mallinnuksen varianssin inflaatiotekijöiden mukaisesti mallissa ei myöskään esiinny multikollineaarisuutta muuttujien välillä. Varianssin inflaatiotekijät löytyvät liitteestä 3.

Taulukko 16. Mallin 1a testien tulokset

Testi	Testisuure	p-arvo
Jarque-Bera-testi	2,9	0,2
Breusch-Pagan-testi	0,7	0,7
F-testi	61,8	$2,2 \times 10^{-16}$

Seuraavassa mallinnuksessa poistetaan poikkeavimmat havainnot alkuperäisestä logaritimuunnellusta mallinnuksesta (malli 1). Poikkeavia havaintoja voidaan etsiä Cookin etäisyysmitan avulla, joka kertoo mitkä havainnot kannattaa ottaa erikoistarkasteluun (Mellin 398-399). Poikkeavien havaintojen poisto toteutetaan laskemalla koko aineiston Cookin etäisyysmitat ja tarkastelemalla aineistosta havainnot, jotka ylittävät kolminkertaisen etäisyysmittojen keskiarvon. Havainnot, jotka ylittävät tämän etäisyysmittojen kolminkertaisen keskiarvon ovat äärihavaintoja, joiden arvot ovat huomattavan matalia tai suuria. Esimerkkinä menetelmällä poistettaviksi ehdottamista havainnoista on vuosien 2016 ja 2017 Murtomäki-Otanmäki rataosan havainnot, jotka sisältävät vain 10 000 bruttotonnin liikennöinnin, joka vastaa vain muutamaa junaa. Kustannuksia nämä havainnot sisältävät noin 39 000 ja 54 000 euroa. Toisen ääripään poistettavaksi ehdotetun havainnon esimerkkinä toimii vuoden 2014 Lahti-Kouvola rataosan havainto, jossa liikennöintiä on 16 miljoonan bruttotonnin edestä ja kustannuksia rataosalla on vain noin 283 000 euroa. Menetelmällä poistetut havainnot on siis tarkastettu ja ne sisältävät joko huomattavan korkeat tai matalat liikennöintimäärät ja/tai kustannukset tai niiden suhde on huomattavan erilainen muusta aineistosta.

Havaintoja poistetaan aineistosta siis yhteensä 50 ja jäljelle jäävien havaintojen lukumäärä on 558 ja näillä havainnoilla toteutetaan uusi mallinnus (malli 1b). Uuden mallinnuksen korjattu  $R^2$ -luku on 0,47, eli hiukan korkeampi, kuin alkuperäisellä logaritmisella mallilla. F-testin arvo on mallinnuksessa 263,6 ja p-arvo on lähes 0. Mallinnus läpäisee F-testin ja itsenäiset muuttujat selittävät riippuvan muuttujan vaihtelua hyvin. Kuvassa 8 on esitetty mallinnuksen jäännöstermien ja poikkeavien havaintojen kuviot.



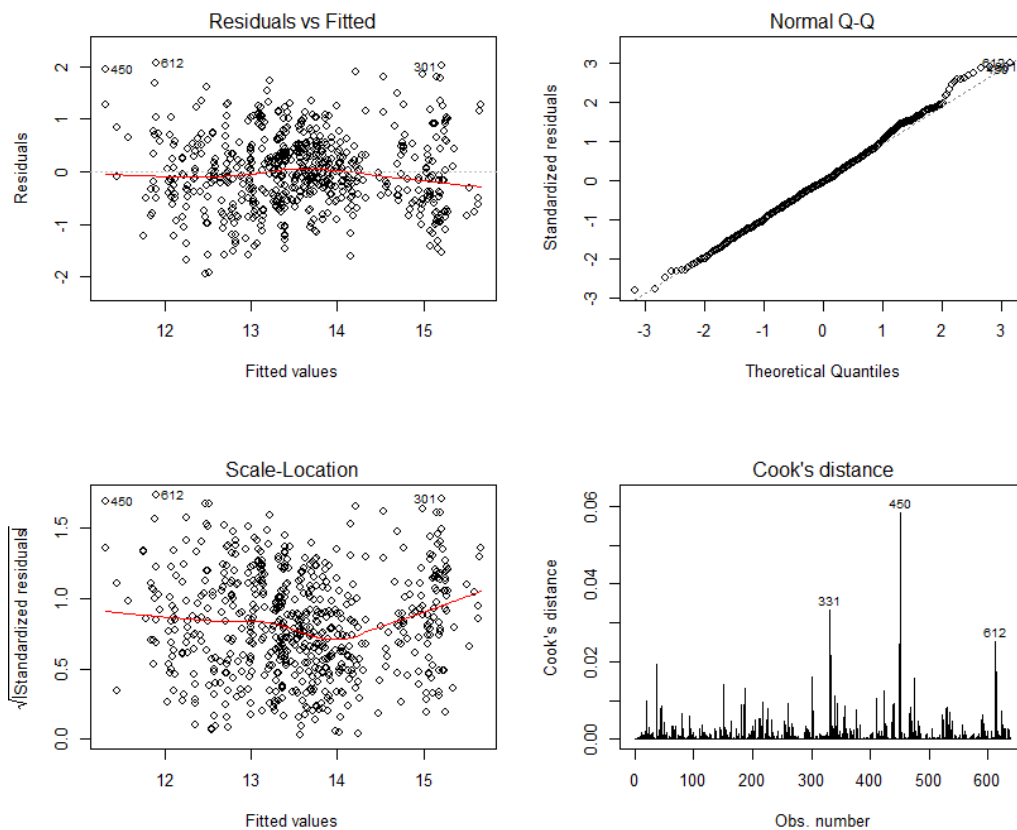
Kuva 8. Mallin 1b visualisointi

Kuvan oikean yläkulman kvantiilikuviosta nähdään, että jäännöstermien  $S$ -muoto on suoristunut verrattuna alkuperäiseen logaritmiseen malliin, joten normaalijakautuneisuusoletus voi täytyä. Jarque-Bera-testisuuren arvo on mallinnuksen virhetermeillä kuitenkin 12,105, joten normaalijakautuneisuusoletus ei testin perusteella täyty. Vasemman yläreunan kuvio tukee lineaarisuusoletusta suoran punaisen viivan myötä, sekä vasemman alareunan kuvio näyttää mallinnuksen homoskedastisuuden kuvion pisteiden suhteellisen tasaisen jakautuneisuuden myötä. Heteroskedastisuushypoteesin hylkäämistä tukee myös Breusch-Pagan-testin  $p$ -arvo 0,5. Oikean alareunan poikkeavien havaintojen kuvaajassa ei ole suurta huomioitavaa, kun suurimmat poikkeamat on jo aineistosta poistettu. Mallin 1b oletusten testien tulokset on koostettu taulukkoon 17 ja ne löytyvät myös liitteestä 3. Mallissa ei myöskään esiinny muuttujien välistä multikollinearisuutta. Mallinnuksen varianssin inflaatiotekijät löytyvät liitteestä 3.

Taulukko 17. Mallin 1b testien tulokset

Testi	Testisuure	$p$ -arvo
Jarque-Bera-testi	12,1	0,002
Breusch-Pagan-testi	1,3	0,5
F-testi	263,6	$2,2 \times 10^{-16}$

Kolmas vaihtoehtoinen mallin jatkojalostus on korvausinvestointien tasoa kuvaavan dummy- eli tekemuuttujan lisääminen. Dummy-muuttujan avulla voidaan sisällyttää analyysiin korvausinvestointien vaikutus kustannuksiin. Tässä mallinnuksessa toteutettu tekemuuttuja saa arvon 1, kun rataosan korvausinvestointien määrä on yli 400 000 euroa, muissa tilanteissa arvo on 0. Toteutuneen mallinnuksen korjattu  $R^2$ -luku, eli selittävyysaste on 0,64. F-testin testi-suureen arvo on mallinnuksessa 373,3 ja p-arvo on lähes 0. Mallinnuksen itsenäiset muuttujat selittävät riippuvan muuttujan vaihtelua hyvin ja mallinnus läpäisee F-testin. Kuvassa 9 on esitetty dummy-muuttujan sisältävän mallinnuksen jäännöstermien ja poikkeavien havaintojen kuvat.



Kuva 9. Mallin 1c visualisointi

Kuvan ensimmäinen kuvio vasemmassa yläreunassa kertoo mallin muuttujien suhteen olevan lineaarinen, eli lineaarisuusoletus täyttyy suoran punaisen viivan myötä. Myös ongelmana ollut jäännöstermien normaalijakautuneisuus näyttää kuvion perusteella täyttyvän, eikä S-muotoa näy kvantiilikuviossa. Tätä päätelmää tukee myös Jarque-Bera-testisuureen arvo 1,0, joka kertoo jäännöstermien normaalijakautuneisuudesta. Vasemman alareunan kuvio viittaa kuitenkin mallin heteroskedastisuuteen, eli homoskedastisuusoletus ei kuvion perusteella täyty tässä mallinnuksessa. Kuviossa on havaittavissa selkeä erillinen rypäs oikeassa reunassa. Breusch-Pagan-testin p-arvo  $3,2 \times 10^{-7}$  tukee tätä päätelmää, eli malli on testin ja kuvion mukaisesti heteroskedastinen. Mallinnuksen ongelmaksi havaitaan siis heteroskedastisuus, eikä malli täten täytä sille asetettuja oletuksia. Mallin 1b oletusten testien tulokset on koostettu taulukoon 18 ja ne löytyvät myös liitteestä 3. Tässäkään mallinnuksessa ei kuitenkaan esiinny multikollineaarisuutta muuttujien välillä. Mallinnuksen varianssin inflaatiotekijät löytyvät liitteestä 3.

Taulukko 18. Mallin 1c testien tulokset

Testi	Testisuure	p-arvo
Jarque-Bera-testi	2,6	0,276
Breusch-Pagan-testi	33,0	$3,2 \times 10^{-7}$
F-testi	373,3	$2,2 \times 10^{-16}$

## 5.4 Mallinnusten tulosten yhteenveto

Tässä tutkimuksessa on esitelty Suomen rautatieinfrastruktuurin mallinnukseen yhteensä kuusi mallinnusta, joista kolme viimeisintä perustuvat ensimmäisten mallinnusten jatkojalostamiseen. Mallinnuksia tehtiin logaritmi-, Box-Cox- ja Box-Tidwell-muunnoksilla. Näistä ensimmäinen, logaritmuunnos valikoitui tulosten yksinkertaisuutensa ja tulosten samankaltaisuuden vuoksi jatkojalostettavaksi. Jatkojalostuksen keinoja olivat käytössä olevan paneelidatiston yhdistäminen poikkileikkausaineistoksi summaamalla vuosittaiset kustannus- ja liikennöintidatat rataosittain, poikkeavimpien havaintojen poistaminen sekä korvausinvestointien tasoa kuvaavan tekemuuttujan (dummy-muuttuja) lisääminen mallinnukseen. Näiden edellä esitettyjen mallien soveltuvuutta kuvaavat tiedot on koostettu taulukkoon 19. Taulukossa on kerrottu oletusten kohdalla, täyttyykö oletus vai ei.

Taulukko 19. Mallinnusten oletusten yhteenveto

Malli	korjattu R <sup>2</sup>	Heteroskedastisuus	Jäännöstermien normaalijakautuneisuus	Multikollineaarisuus
Malli 1	0,43	✓	✗	✓
Malli 2	0,43	✓	✗	✓
Malli 3	0,43	✓	✗	✓
Malli 1a	0,53	✓	✓	✓
Malli 1b	0,47	✓	✗	✓
Malli 1c	0,64	✗	✓	✓

Taulukosta nähdään, että alkuperäisellä laskenta-aineistolla, josta oli poistettu vain nollaliikenteiset rataosat, kaikkien kolmen mallinnuksen korjattu selittävyysaste on sama, 0,43. Tämä on linjassa kansainvälisten ratainfrastruktuurin käytön rajakustannusten sekä Suomessa aiemmin toteutettujen mallinnusten kanssa. Kaikissa näissä kolmessa mallinnuksessa oli ongelmana jäännöstermien normaalijakautuneisuus, joka ei täytynyt missään mallinnuksessa. Mallinuksista missään ei kuitenkaan esiintynyt alkuperäistä ongelmaa, eli heteroskedastisuutta. Myöskään multikollineaarisuus ei ollut minkään mallin ongelma-kohta.

Kolmen ensimmäisen mallinnuksen jäännöstermien normaalijakautuneisuuden ongelmien vuoksi malli 1 valittiin jatkojalostuksen kohteeksi. Jatkojalostetuissa malleissa selittävyysaste parani, eli ne pystyivät selittämään kustannusten vaihtelua paremmin. Mallinnusten selittävyysasteet ovat kansainvälisesti yleisellä tasolla, eli tulokset ovat näiltä osin kohtuullisia. Mallissa 1c on selkeästi parempi selitysaste kuin alkuperäisellä logaritmuunnellulla mallilla. Poikkileikkausaineistolla mallinnus sekä korvausinvestointimuuttujan sisällyttäminen mallin-

nukseen paransivat ongelmana olleen jäännöstermien normaalijakautuneisuuden. Poikkeavien havaintojen poiston jälkeisessä mallinnuksessa jäännöstermien normaalijakautuneisuus parani hieman, mutta normaalijakautuneisuusoletus ei kuitenkaan täyty Jarque-Beran normaalijakautuneisuustestin perusteella. Korvausinvestointimuuttujan sisällyttäminen aiheutti malliin 1c heteroskedastisuutta, eli mallissa 1c ei täyty homoskedastisuusoletus. Missään jatkajalostetuista malleista ei esiintynyt myöskään multikollineaarisuutta.

Malleista lopulta siis vain yksi, poikkileikkausaineistolla toteutettu mallinnus, täytti kaikki mallille asetetut oletukset ja soveltuu näin rautatieinfrastruktuurin käytön rajakustannusten mallinnukseen. Malli sisältää 108 havaintoa, joista jokainen on koostettu rataosittain kustannukset ja bruttotonnit yhteenlaskemalla kuuden vuoden ajanjaksolta.

## 5.5 Rajakustannus ja kustannusjousto

Aiemmassa luvussa 5.3 rakennettiin yksi toimiva ekonometrinen mallinnus rautatieinfrastruktuurin liikennöinnistä aiheutuvien rajakustannusten laskemista varten. Tässä alaluvussa koostetaan esitellyn logaritmuunnellun poikkileikkausaineistomallinnuksen tuottamat rajakustannukset sekä koko rataverkon painotetun rajakustannuksen luottamusvälit.

Oletukset täyttävä mallinnus, poikkileikkausaineistomallinnus on logaritmuunneltu kustannusfunktio. Kustannusfunktioista voidaan ensin johtaa kustannusjousto osittaisderivoimalla funktio liikennesuorituksen suhteen. Liikennesuoritetta kuvaavana muuttujana toimii funktiossa bruttotonnit. Rataosakohtainen kustannusjousto mallinnuksen 1a kustannusfunktioista (kaavasta 25) voidaan siis esittää seuraavasti (kaava 29):

$$\epsilon_i = \frac{\partial \ln C_i}{\partial x_i} = \beta_{brt} \quad (29)$$

Kaavasta 29 voidaan jatkaa rajakustannusten laskemiseen kertomalla kustannusjousto rataosan keskimääräisellä kustannuksella. Tässä vaiheessa laskelmaan tulee lisätä mukaan rataosan pituus, sillä rajakustannusten laskutusyksikkönä toimii bruttotonnikilometrit. Bruttotonnikilometrit havainnoittain saadaan laskettua kertomalla liikennöinnin bruttotonnit rataosan pituudella. Tähän laskelmaan ei sisällytetä rataosan raiteiden lukumäärää, eli rataosan pituus eroaa raidekilometreistä. Edelleen havainnoittain rajakustannuksen estimaatti voidaan siis esittää seuraavalla tavalla (kaava 30):

$$MC_i = \beta_{brt} \frac{\hat{C}_i}{brtkm_i} \quad (30)$$

missä  $brtkm_i$  on havainnon  $i$  bruttotonnikilometrit.  $\hat{C}_i$  oletetaan olevan log-normaalijakautunut parametrein  $\ln \hat{C}_i \sim [\ln C_i - \epsilon_i, \sigma^2]$ , missä  $\sigma^2$  on virhetermin varianssi. Tästä seuraten  $\hat{C}_i$  estimaatti voidaan määrittää seuraavasti (kaava 31):

$$\hat{C}_i = \exp(\ln C_i - \epsilon_i + 0.5(\sigma^2)) \quad (31)$$

joka on käytännössä log-normaalijakauman odotusarvo. Kaavojen 30 ja 31 avulla saadaan siis laskettua yksittäisen havainnon rajakustannus. Koko rataverkon laajuinen rajakustannus on kuitenkin tarpeellinen ja se saadaan laskettua painottamalla keskimääräiset rajakustannukset bruttotonnikilometriä mukaisesti. Keskimääräisten painotettujen rajakustannusten laskentatapa on seuraava (kaava 32):

$$\overline{MC}_i = \beta_{brt} \frac{\sum_i \hat{C}_i}{\sum_i brtkm_i} \quad (32)$$

Taulukossa 20 on esitetty mallinnuksen 1a tuottama rajakustannus ja kustannusjousto. Rajakustannukset on esitetty taulukossa muodossa senttiä per bruttotonnikilometri kertomalla kaavasta saatu lopputulos sadalla. Rajakustannusten osalta bruttotonnikilometriä mukaisen painotuksen huomataan olevan erittäin tärkeää, sillä havainnoittain rajakustannusten vaihteluväli on huomattavan laaja. Tämä tarkoittaa, että eri rataosilla liikennöinti aiheuttaa jopa tuhatkertaiset kustannuserot.

*Taulukko 20. Poikkileikkausaineistomallinnuksen kustannusjousto ja rajakustannukset*

Malli	Kustannusjousto	Keskiarvo	Painotettu keskiarvo	Minimi	Maksimi
Malli 1a	0,221	1,17	0,115	0,0426	47,4

Kirjallisuuskatsauksessa huomattiin kansainvälisten raitinfraktuurien ekonometristen rajakustannusmallinnusten kustannusjoustojen asettuvan yleisesti välille 0,1–0,5. Taulukosta huomataan, että myös tässä tutkimuksessa tehty oletukset täyttävän mallinnuksen kustannusjousto asettuu myös tuolle välille, joten tulokset ovat myös näiltä osin kohtuullisia ja linjassa kansainvälisten tutkimusten kanssa.

Painotetun rajakustannuksen 95 prosentin luottamusvälin alaraja on 0,0170 senttiä bruttotonnikilometriltä ja yläraja on 0,674 senttiä bruttotonnikilometriltä, mikä viittaa siihen, että "oikeat" rajakustannukset sijoittuvat 95 prosentin todennäköisyydellä tälle välille. Väli on huomattavan laaja ja taulukossa esitetty painotettu keskiarvo, jota ratamaksun perusmaksuna käytetään, asettuu hyvin alas tässä luottamusvälissä johtuen mallin logaritmisuudesta. Tutkimuksen tuloksena saatu rajakustannusten painotettu keskiarvo vastaa suuruudeltaan Suomessa saatua ja ratamaksuhinnastossa olevia rajakustannuksia.



## 6 Johtopäätökset

Tämän tutkimuksen tarkoituksena oli tarkastella ratamaksun perusteita koskevat rajoitukset, tutkia muiden maiden käytäntöjä sekä lopulta tarkastella Suomen raitainfrastruktuurin käytön rajakustannuksen mallinnusta. Näihin kohtiin pyrittiin hakemaan ratkaisuja kolmen tutkimuskysymyksen avulla. Ensimmäinen tutkimuskysymys oli:

1. Mitkä ovat ratamaksun perusmaksun määrittämisen lainsäädännölliset rajoitukset?

Tutkimuksen toisen luvun perusteella voidaan todeta ratamaksun perusmaksun määrittämisen lainsäädännöllisten reunaehtojen löytyvän Euroopan parlamentin ja neuvoston direktiivistä 2012/34/EU yhtenäisestä eurooppalaisesta rautatiealueesta, Euroopan komission täytäntöönpanoasetuksesta 2015/909 rautatieliikenteen harjoittamisesta suoraan aiheutuvien kustannusten laskentasäännöistä sekä Suomen raideliikennelaista. Nämä reunaehdot koskevat ratamaksun perusmaksuun sisällytettäviä kustannuksia, ratamaksun perusmaksun suuruutta sekä laskentaperusteita. Ratamaksun perusmaksun tulee perustua suoraan liikenteen harjoittamisesta aiheutuviin kustannuksiin, jota vastaan rataverkon haltijan tulee tarjota liikennöitsijälle vähimmäiskäyttömahdollisuuksiin sisältyvät palvelut. Junaliikenteen harjoittamisesta suoraan aiheutuviin kustannuksiin sisältyvät kustannukset on eritelty tarkemmin Euroopan komission täytäntöönpanoasetuksessa 2015/909.

Ratamaksun perusmaksun määrittämiseen on mainittu karkeat peruslinjat lainsäädännössä ja lainsäädäntö jättää paljon tulkinnanvaraa. Ratamaksun perusmaksun määrittäjän vastuulla on tarkka kustannusten erittely ja suodattaminen ratamaksua varten, sillä lainsäädännössä määrätään, että ratamaksun perusmaksun tulee perustua vain suoraan liikenteen harjoittamisesta aiheutuviin kustannuksiin. Lisäksi näistä kustannuksista tulee poistaa erikseen täytäntöönpanoasetuksessa mainitut kielletyt erät, joita ei saa sisällyttää ratamaksun perusmaksun laskentaan. Kielletyt erät voivat olla monitulkintaisia, joten tämä lisää haastavuutta kustannusaineiston keräämisessä.

Lainsäädäntö painottuu suurelta osin laskentaan sisällytettäviin kustannuksiin, eikä niinkään itse ratamaksun perusmaksun laskentaan. Ratamaksun perusmaksun laskentatapoja mainitaan täytäntöönpanoasetuksessa 2015/909 kolme. Nämä laskentatavat ovat niin kutsutut vähennyslasku-, insinööri- ja ekonometrinen menetelmä. Näistä menetelmistä ekonometrinen ja tekninen kustannusmallinnus on mainittu erikseen yhdessä artikkelissa yhdellä lauseella ikään kuin ensisijaisen vähennyslaskumenetelmän sivussa ilman ohjeistusta itse kustannusmallinnusten menetelmiin. Tämä vähäinen sääntely kyseisten kahden laskentatavan osalta mahdollistaa hyvinkin erilaisten kustannusmallinnusten rakentamisen, sillä minkäänlaisia funktiomuotoja tai estimointimenetelmiä ei ole määritelty. Tässä työssä käytiin läpi IRG-Railin tulkinta tekniseen ja ekonometriseen kustannusmallinnukseen, mikä on vain yksi mahdollinen näkökulma. Toisaalta vähäinen sääntely antaa suhteellisen vapaat kädet ratamaksun perusmaksun teknisen tai ekonometrisen kustannusfunktion määrittelijälle, mutta taas toisaalta vaatii paljon menetelmien termien ja soveltamisen ymmärtämistä.

Toisena tutkimuskysymyksenä oli:

2. Millä tavoilla Euroopan talousalueen jäsenmaissa on mallinnettu ekonometrisesti rautatieinfrastruktuurin käytön rajakustannuksia?

Ekonometrinen kustannusmallinnus on käytössä viidessä Euroopan talousalueen jäsenmaassa. Tähän tutkimukseen valikoitui tarkempaan käsittelyyn kolmen Euroopan talousalueen jäsenmaan rautatieinfrastruktuurin käytön rajakustannusten ekonometrinen mallinnus. Tutkimuksen kohteena olivat Norjan, Ranskan ja Suomen rajakustannusten mallinnus. Maat valikoituivat niiden saatavilla olevien tarkkojen menetelmämuistioiden vuoksi.

Norjan rautatieinfrastruktuurin käytön rajakustannusmallinnus perustuu tarkasti eriteltyihin kustannuksiin sekä useisiin mallinnuksiin. Kustannusten erottelun taustalla on lainsäädännön asettamat reunaehdot. Norjassa tutkimuksen kohteena oli kuusi eri mallinnusta, jotka oli lajiteltu kolmeen eri kategoriaan: polynomisiin malleihin absoluuttisilla arvoilla, polynomisiin malleihin luonnollisilla logaritmeilla sekä kahden tuotantomuuttujan Translog-malleihin. Parhaiten suoriutuneet mallinnukset olivat kolmannen asteen polynominen malli, toisen asteen polynominen malli perustuen luonnollisiin logaritmeihin sekä kahden tuotantomuuttujan Translog-malli. Lopullisessa, käyttöön valikoituneessa rajakustannusten mallinnuksessa käytössä on Translog-malli, joka sisältää tuotantomuuttujina henkilö- ja rahtiliikenteen bruttotonnit, kyseisten bruttotonnien neliöt sekä interaktiomuuttuja, joka on henkilöliikenne kerrottuna rahtiliikenteellä. Teknisinä muuttujina mallissa ovat raidekilometrit, alle 500 metrin taipuisuuksien määrä per kilometri, tunnelikilometrien osuus, vaihteiden lukumäärä, painotettu nopeus, hintaindeksi ja dummy-muuttuja sähköistyksestä. Tämä mallinnus on edelleen käytössä ratamaksun perusmaksun laskennassa.

Ranskassa rautatieinfrastruktuurin käytön rajakustannusten mallinnus perustuu myös rautatiemarkkinadirektiivin 2012/34/EU määrittelemiin kustannuksiin ja ensimmäinen kattava tutkimus on toteutettu vuonna 2014. Vuonna 2014 toteutetun tieteellisen tutkimuksen myötä kustannusfunktion ekonometrisen mallinnuksen funktiomuodoksi valikoitui Box-Cox-muunneltu kustannusfunktio. Tutkimuksessa toteutettiin jokaiselle kustannuskategorialle oma kustannusfunktio, joita rakentui yhteensä viisi. Kustannuskategoriat olivat radan ylläpito, vaihteiden ja risteysten ylläpito, merkinantojärjestelmän ylläpito, ajojohtojärjestelmän ylläpito sekä liikenteenohjaus. Näissä kaikissa mallinnuksissa oli omat itsenäiset ja dummy-muuttujat, jotka olivat liitännäisiä kustannuskategoriaan. Myöhemmin, vuonna 2017 kustannusfunktion muodoksi valikoitui uuden tutkimuksen myötä Translog-kustannusfunktio, jota käytetään nykyään Ranskassa ratamaksun perusmaksun laskennassa.

Suomessa käytössä olevassa ratamaksun perusmaksun ekonometrisessä mallinnuksessa on myös tarkasti eriteltyt kustannukset, jotka perustuvat säännöksiin. Ratamaksun perusmaksun mallinnuksiin sisällytetään vain siihen sallitut kustannukset. Käytössä oleva mallinnustapa on logaritmimuunneltu kustannusfunktio, jossa on kaksi itsenäistä muuttujaa. Itsenäiset muuttujat ovat mallinnuksessa liikennöintiä kuvaava bruttotonnikilometrimuuttuja sekä radan ominaisuuksia kuvaava raidekilometrimuuttuja.

Tässä tutkimuksessa esitellyissä Euroopan talousalueen jäsenmaissa on siis koekeltu polynomisia mallinnuksia, Box-Cox-muunneltuja kustannusfunktioita, Translog-tuotantoteknologiaa hyödyntäviä kustannusfunktioita sekä logaritmimuunneltuja kustannusfunktioita. Nykyään Norjassa ja Ranskassa on käytössä

Translog-tuotantoteknologiaa hyödyntävät kustannusfunktiot ja Suomessa on käytössä logaritmimuunneltu kustannusfunktio. Kaikissa maissa riippuvana muuttujana toimii kustannukset ja itsenäisinä muuttujina toimivat liikennesuoritteita sekä ratainfrastruktuurin ominaisuuksia kuvaavia muuttujia. Rautatieinfrastruktuuria kuvaavat muuttujat vaihtelevat eri maiden välillä, mikä on seurausta tapauskohtaisesti saatavilla olevan datan laadusta ja laajuudesta.

Euroopan talousalueen jäsenmaissa rautatieinfrastruktuurin käytön rajakustannukset eroavat toisistaan pääosin teknisten muuttujien lukumäärän osalta. Mahdollisia ekonometrisia menetelmiä ei ole juurikaan rajoitettu lainsäädännössä, mutta silti menetelmät ovat vakiintuneet logaritmimuunneltuihin, Box-Cox-muunneltuihin sekä Translog-tuotantoteknologiaa hyödyntäviin kustannusfunktioihin.

Kolmantena tutkimuskysymyksenä oli:

3. Mikä on sopiva mallinnustapa Suomen ratamaksun perusmaksulle, eli rautatieinfrastruktuurin liikenteestä aiheutuville rajakustannuksille?

Kolmanteen tutkimuskysymykseen haettiin vastaus työn empiirisessä osiossa. Työn taustalla mallinnuksia teetettiin useita kymmeniä muun muassa erinäisillä parametreilla ja estimointimenetelmillä ilman mitään merkittäviä tuloksia. Tähän työhön esiteltäväksi valikoitui kolme mallinnusta, joita vertailtiin pienimmän neliösumman menetelmän oletusten täyttymisen sekä selitysasteen perusteella. Nämä kolme mallinnusta perustuivat muuttujien logaritmi-, Box-Cox- sekä Box-Tidwell-muunnoksiin. Edelleen näistä kolmesta mallinnuksesta valittiin samankaltaisten tulosten myötä yksinkertaisin mallinnus jatkojalostettavaksi. Jatkojalostettavaksi valikoitu jo aiemmin Suomessa rautatieinfrastruktuurin liikenteestä aiheutuvien rajakustannusten mallinnuksessa käytetty logaritmimuunneltu kustannusfunktio. Logaritmimuunnellulle kustannusfunktiosta tehtiin kolme uutta mallinnusta: ensimmäisessä yhdistettiin käytössä oleva paneeliaineisto poikkileikkausaineistoksi yhteen laskemalla vuosittaiset havainnot rataosittain, toisesta poistettiin poikkeavimmat äärihavainnot ja kolmanteen lisättiin korvausinvestointien tasoa kuvaava tekomuuttuja. Lopulta sopivimmaksi rautatieinfrastruktuurin liikenteestä aiheutuville rajakustannuksille havaittiin poikkileikkausaineistolla toteutettu logaritmimuunneltu mallinnus.

Tämä lineaariselle regressioanalyysille asetetut oletukset täyttävä mallinnus tehtiin siis aineistolle, jossa oli yhdistetty kuuden vuoden aikasarjan havainnot laskemalla ne yhteen. Tämä toimenpide muunsi käsillä olleen paneeliaineiston poikkileikkausaineistoksi, jonka yksi aikayksikkö on kuusi vuotta. Edelleen logaritmimuunneltuna tällä aineistolla teetetty regressioanalyysi tuotti kustannusfunktion, joka täytti kaikki mallinnukselle asetetut oletukset ja erityisesti korjasi Suomen ratamaksun perusmaksulaskelmissa aiempina vuosina ongelmana olleen heteroskedastisuuden. Poikkileikkausaineistoksi muunneltu aineisto tasoi korvausinvestointien aiheuttamaa vuosittaista kustannusten vaihtelua, joka oli omiaan tuottamaan ongelmia mallinnuksille.

Rautatieinfrastruktuurin käytön rajakustannusten ekonometrista mallinnusta voidaan pyrkiä kehittämään edelleen tarkastelemalla liikennesuoritteiden ajallista viivettä suhteessa kustannuksiin. Tämänlaisen tarkastelun avulla piste- ja viivemallien korvausinvestointikustannusten mallinnus onnistuisi mahdollisesti paremmin. Ajallisten viiveiden tarkastelu ei ollut tässä työssä mahdollista suhteellisen lyhyen aikasarjan vuoksi.

## 7 Yhteenveto

Tutkimuksen tavoitteena oli tarkastella Suomessa nykyään käytössä olevaa rautatieinfrastruktuurin ekonometrista mallinnusta. Tarkastelun tarve syntyi ekonometrisen mallinnuksen ongelmakohdista, joista suurin oli mallin heteroskedastisuus.

Tutkimus aloitettiin tarkastelemalla lainsäädännöllisiä reunaehtoja, jotka ratamaksun perusmaksun määrittelylle on asetettu. Työssä esiteltiin ratamaksun perusmaksun sääntelyä koskevat asetukset ja lait, joista merkittävimpiä ovat Euroopan parlamentin ja neuvoston direktiivi 2012/34/EU, Euroopan komission täytäntöönpanoasetus 2015/909 sekä Suomen rautatieinfrastruktuurilaki. Lainsäädännöllisten viitekehysten määrittelyn jälkeen työssä tutkittiin rajakustannusfunktioiden ja lineaarisen regression teoriaa. Teorian läpikäymisen jälkeen työssä tuotiin esille Norjan, Ranskan ja Suomen rautatieinfrastruktuurin käytön rajakustannusten ekonometrisen mallinnuksen menetelmiä.

Tutkimuksen empiirisessä osiossa tarkasteltiin kirjallisuuden perusteella valittuja erilaisia mallinnustapoja Suomen rautatieinfrastruktuurin käytön rajakustannusten mallinnuksessa. Mallinnuksia esiteltiin työssä aluksi kolme, joista valittiin tulosten samankaltaisuuden vuoksi jatkojalostettavaksi jo Suomessa käytössä ollut logaritimuunneltu malli. Jatkojalostuksessa rakentui kolme mallinnusta, joista yksi täytti kaikki mallille asetetut oletukset. Sopiva mallinnustapa perustui paneeliaineiston yhdistämiseen poikkileikkausaineistoksi yhteen laskemalla aikasarjan liikennöinti- ja kustannustiedot rataosittain.

Työssä kehitetty ratkaisu täyttää lainsäädännölliset reunaehdot, tulokset ovat kustannusjouston sekä mallin selittävyysasteen osalta kansainvälisellä tasolla ja mallinnuksen antama rajakustannus vastaa suuruudeltaan Suomessa aiemmin määritettyjä ja ratamaksuhinnastossa olevia rajakustannuksia. Työlle asetetut tavoitteet täyttyivät sekä työn tulokset tuottivat uuden mallinnustavan ratamaksun perusmaksun laskentaan. Mallinnusperiaatetta hyödynnettiin vuoden 2022 ratamaksun perusmaksun ekonometrisessa laskennassa syksyllä 2020. Ratamaksun perusmaksun ekonometrista mallinnusta voidaan kehittää edelleen tarkastelemalla liikennesuoritteiden ajallista viivettä suhteessa kustannuksiin, mikä ei tässä työssä ollut mahdollista suhteellisen lyhyen aikasarjan vuoksi.

## Lähteet

Andersson, M. 2006. Marginal cost pricing of railway infrastructure operation, maintenance and renewal in Sweden: From policy to practice through existing data. *Transportation Research Record. Journal of the Transportation Research Board*, 1943, Washington, DC, s. 1–11.

Andersson, M. 2009. Marginal costs of railway infrastructure wear and tear for freight and passenger trains in Sweden. Annex to Deliverable D8. CATRIN (Cost Allocation of Transport Infrastructure). Research project funded by Sixth EU Framework Programme, Stockholm: VTI.

Andersson, M., Smith, A., Wikberg, Å. & Wheat, P. 2012. Estimating the marginal cost of railway track renewals using corner solution models. *Transportation Research Part A*. 46(6), s. 954-964.

Andersson, M., Björklund, G. & Haraldsson, M. 2016. Marginal railway track renewal costs: A survival data approach. *Transportation Research Part A*. 87, s. 68-77.

Asar Ö., Ilk O. & Dag O. 2017. Estimating Box-Cox power transformation parameter via goodness-of-fit tests. *Communications in Statistics - Simulation and Computation*. 46(1), s. 91-10.

Bane NOR. 2017. Infrastructure charging implementation plan. Saatavilla: [https://www.banenor.no/globalassets/kundeportal/dokumenter/infrastrukturpriser/implementeringsplan\\_infrastrukturavgifter\\_20170714\\_engelsk.pdf](https://www.banenor.no/globalassets/kundeportal/dokumenter/infrastrukturpriser/implementeringsplan_infrastrukturavgifter_20170714_engelsk.pdf)

Bane NOR. 2019. Network Statement 2021. Saatavilla: <http://networkstatement.jbv.no/doku.php?id=2021>

Bane NOR. 2020. About us. Saatavilla: <https://www.banenor.no/en/startpage1/About-Us/>

Borts, G. (1960). The Estimation of Rail Cost Functions. *Econometrica*, 28(1), s. 108-131.

Breusch, T. & Pagan, A. 1979. A SIMPLE TEST FOR HETEROSCEDASTICITY AND RANDOM COEFFICIENT VARIATION. *Econometrica* (ennen 1986), 47 vol. 5, s. 1287.

Brooks, C. 2014. *Introductory Econometrics for Finance*. 3. painos. Cambridge: Cambridge University Press.

Christensen, L. R., Jorgenson, D. W. & Lau, L. J. 1973. Transcendental Logarithmic Production Frontiers. *The Review of Economics and Statistics*, 55(1), s. 28-45.

Daljord, Ö.B. 2003. Marginalkostnader i Jernbanenettet. Raportti 2/2003. Ragnar Frisch Centre for Economic Research, Oslo, Norja.

Euroopan parlamentin ja neuvoston direktiivi 21.11.2012/34/EU yhtenäisestä eurooppalaisesta rautatiealueesta. [Euroopan unionin virallinen lehti](#).

Frontier Economics. 2017. [ESTIMATION DES COÛTS MARGINAUX D'ENTRETIEN DU RESEAU FERRE NATIONAL](#).

Gaudry, M. & Quinet, E. 2003. Rail Track Wear-and-tear Costs by Traffic Class in France. Publication AJD-66. Universite de Montreal.

Gaudry, M. & Quinet, E. 2009. Track wear-and-tear cost by traffic class: Functional form, zero output levels and marginal cost pricing recovery on the French rail network. Federal Reserve Bank of St Louis, St. Louis.

Idström T. 2002. Suomen ratamaksun uudistaminen – ekonometrinen analyysi rataverkon käytön rajakustannuksista (pro gradu -tutkielma, Jyväskylän yliopisto).

Independent Regulator's Group – Rail (IRG-Rail). 2016. [An introduction to the calculation of direct costs in respect of implementing regulation 2015/909](#).

Independent Regulator's Group – Rail (IRG-Rail). 2020. [General Information](#).

Johansson, P. & Nilsson, J. 2004. An economic analysis of track maintenance costs. Transport Policy, 11(3), s. 277-286.

Karjanlahti, J. & Kukkarinen, A. 2019. [VR-Yhtymä Oy:n oikaisuvaatimus koskien Väyläviraston verkkoselostuksissaan julkaisemia sähkövetoisen liikenteen ratamaksuja vuosille 2019 ja 2020](#). Helsinki, Rautatiealan sääntelyelin.

Komission täytäntöönpanoasetus ([EU 12.6.2015/909 rautatieliikenteen harjoittamisesta suoraan aiheutuvien kustannusten laskentasäännöistä](#)). Euroopan unionin virallinen lehti.

Lescohier, D. D. 1935. The Theory of Wages. Political Science Quarterly, 50 Vol 2, s. 272-277.

Liikennevirasto. 2018. Kartta: [valtion rataverkko 1.1.2018](#) (pdf).

Link H. & Maibach M. 1999. [FINAL REPORT OF THE EXPERT ADVISORS TO THE HIGH LEVEL GROUP ON INFRASTRUCTURE CHARGING](#) (WORKING GROUP 1).

Link, H. 2009. Marginal costs of rail maintenance and renewals in Austria. Annex 3 to Rail Cost Allocation for Europe, Deliverable 8. CATRIN (Cost Allocation of Transport Infrastructure). Research project funded by Sixth EU Framework Programme, Tukholma: VTI.

Marti, M. & Neuenschwander, R. 2006. Case study 1.2E: track maintenance costs in Switzerland. Marginal Cost Case Studies for Road and Rail Transport, Deliverable D3. GRACE (Generalisation of Research on Accounts and Cost Estimation). Research project funded by Sixth EU Framework Programme.

Marti, M., Neuenschwander, R. & Walker, P. 2009. Track maintenance and renewal costs in Switzerland. Annex 1B to Rail Cost Allocation for Europe, Deliverable 8. CATRIN (Cost Allocation of Transport Infrastructure), Research project funded by Sixth EU Framework Programme, Stockholm: VTI.

Mellin, I. 2006. [Lineaarinen regressioanalyysi](#). TKK.

Munduch, G., Pfister, A., Sögner, L., & Stiassny, A. 2002. Estimating Marginal Costs for the Austrian Railway System. Vienna University of Economics, Department of Economics, Department of Economics Working Papers, 78.

Odolinski, K. & Wheat, P. 2018b. Dynamics in rail infrastructure provision: Maintenance and renewal costs in Sweden. *Economics of Transportation*, 14, s. 21-30.

Odolinski, K., Haraldsson, M., Yarmukhamedov, S. & Nilsson, J. 2018a. The marginal cost of track reinvestments in the Swedish railway network: Using data to compare methods. *Working papers in Transport Economics 2018:20*, CTS - Centre for Transport Studies Stockholm (KTH and VTI).

Osborne, J., & Waters, E. 2002. Four assumptions of multiple regression that researchers should always test. *Practical Assessment, Research & Evaluation*, 8(2).

[Raideliikennelaki 28.12.2018/1302](#). Finlex.

Silavong, C., Guiraud, L. & Brunel, J. 2014. [Estimating the marginal cost of operation and maintenance for French railway network](#). ITEA Conference, 2014.

Sternad M., Jagric T. & Rosi B. 2018. Railway usage charges based on marginal maintenance costs. *Proceedings of the Institution of Civil Engineers – Transport*, 171 (1), s. 3-10.

Suomen Pankki. 1989. Vektoriautoregressiiviset mallit: Menetelmä ja sovelluksia Suomen aineistolla. Helsinki.

SNCF. 2020. [Profile and key figures](#).

SNCF Réseau. 2019a. [CHARGING PRINCIPLES OF MINIMUM SERVICES](#).

SNCF Réseau. 2019b. [NETWORK STATEMENT OF THE NATIONAL RAIL NETWORK 2021 Timetable](#).

Tervonen J. & Idström T. 2004. Radan kulumisen rajakustannukset vuosina 1996-2002. Ratahallintokeskus. Helsinki. Julkaisusarjan julkaisu A 5/2004.

Tervonen J. & Pekkarinen S. 2007. [Radan kulumisen rajakustannukset 1997-2005](#). Ratahallintokeskus. Helsinki. Julkaisusarjan julkaisu A 2/2007.

Tervonen, J. 2016. Ratamaksun perusmaksun kehittäminen. Helsinki, Liikennevirasto.

Väylävirasto. 2019. [Rautatiealan sääntelyelimen edellyttämä ratamaksun perusmaksun uudelleenmäärittäminen Väyläviraston vuoden 2021 verkkoselostukseen, menetelmämuistio 29.10.2019](#).

Väylävirasto. 2020a. [Rautateiden verkkoselostus 2021](#). Väylävirasto. Helsinki. Väyläviraston julkaisuja 46/2019.

Väylävirasto. 2020b. [Rataverkko](#).

Welc, J. & Esquerdo, P. J. R. 2018. Applied Regression Analysis for Business: Tools, Traps and Applications. Cham, Sveitsi: Springer International Publishing AG. 286 s. ISBN 978-3-319-71155-3

Wheat, P. and Smith, A. 2008. Assessing the marginal infrastructure maintenance wear and tear costs for Britain's railway network. *Journal of Transport Economics and Policy*, 42 (2), s. 189–224.

Wheat, P., Smith, A., & Nash, C. (2009). Rail cost allocation for Europe. CATRIN (Cost Allocation of Transport Infrastructure cost) Deliverable 8. Funded by the European Commission 6th Framework Programme.

Wooldridge, J. M. 2002. *Econometric analysis of cross section and panel data*. Cambridge (MA): MIT Press.



# Koko verkon laajuiset suorat kustannukset

## 3 artikla

### Koko verkon laajuiset suorat kustannukset

1. Koko verkon laajuiset suorat kustannukset ovat erotus toisaalta vähimmäiskäyttömahdollisuuksien palvelujen tarjoamisesta ja palvelupaikkoja yhdistävän infrastruktuurin käyttöoikeuksista aiheutuvien kustannusten sekä toisaalta 4 artiklassa tarkoitettujen muiden kuin huomioon otettavien kustannusten välillä.
2. Jäsenvaltio voi päättää, että rataverkon haltijan on määritettävä 1 kohdassa tarkoitetut koko verkon laajuiset suorat kustannukset tehokkaan palvelutarjonnan kustannusten periaatteen mukaisesti.
3. Koko verkon laajuisten suorien kustannusten laskennassa huomioon otettavien omaisuuden arvojen on perustuttava alkuperäisiin arvoihin tai, jos tällaisia arvoja ei ole saatavilla tai jos käyvät arvot ovat alemmat, käypiin arvoihin. Omaisuuden alkuperäisten arvojen on perustuttava hintaan, jonka rataverkon haltija on maksanut ja dokumentoinut kyseisen omaisuuden hankintahetkellä. Velkahelpotuksen tapauksessa, jossa muu yksikkö ottaa vastatakseen rataverkon haltijan velat kokonaan tai osittain, rataverkon haltijan on käytettävä vastaava osa velkahelpotuksesta omaisuutensa arvon ja vastaavien koko verkon laajuisten suorien kustannusten alentamiseen. Poiketen siitä, mitä ensimmäisessä virkkeessä säädetään, rataverkon haltija voi soveltaa arvoja, mukaan lukien arvioidut arvot, käyvät arvot tai jälleenhankinta-arvot, jos se voi osoittaa sääntelyelimelle, että arvot on määritetty ja perusteltu läpinäkyvästi, luotettavasti ja objektiivisesti.
4. Rajoittamatta 4 artiklan soveltamista ja jos rataverkon haltija voi, muun muassa parhaiden kansainvälisten käytäntöjen perusteella, määrittää ja osoittaa läpinäkyvästi, luotettavasti ja objektiivisesti, että kustannukset aiheutuvat suoraan rautatieliikenteen harjoittamisesta, rataverkon haltija voi ottaa koko verkon laajuisten suorien kustannusten laskennassa huomioon erityisesti seuraavat kustannukset:
  - a. sellaisen henkilöstön kustannukset, jota tarvitaan tietyn rataosuuden pitämiseen saatavilla, jos jokin toiminnanharjoittaja pyytää saada harjoittaa erityistä rautatieliikennettä kyseisen rataosuuden tavanomaisten liikennöinti-aikojen ulkopuolella;
  - b. rautatiepalvelujen vuoksi kulumiselle altistuvien vaihteiden, mukaan lukien yksinkertaiset vaihteet, risteysvaihteet ja raideris-teykset, kustannusosuus;
  - c. ajolangan tai virtakiskon taikka molempien ja kannattavien ratajohtolaitteiden uusimisen tai kunnossapidon kustannusosuus, joka aiheutuu suoraan rautatieliikenteen harjoittamisesta;
  - d. sellaisen henkilöstön kustannukset, jota tarvitaan junareittien jakamiseen ja aikataulun laatimiseen, jos ne ovat aiheutuneet suoraan rautatieliikenteen harjoittamisesta.

- 
5. Tämän artiklan nojalla laskentaan sovellettavien kustannusten on perustuttava rataverkon haltijan suorittamiin tai arvioimiin maksuihin. Tämän artiklan nojalla laskettavat kustannukset on määritettävä tai arvioitava johdonmukaisesti samalta viitekaudelta saatujen tietojen perusteella.

# Muut kuin huomioon otettavat kustannukset

## 4 artikla

### **Muut kuin huomioon otettavat kustannukset**

1. Koko verkon laajuisia suoria kustannuksia laskettaessa rataverkon haltija ei saa ottaa huomioon seuraavia kustannuksia:
  - a. tietyn rataosuuden tarjontaan liittyvät kiinteät kustannukset, joista rataverkon haltijan on vastattava myös silloin, kun rautatieliikennettä ei ole;
  - b. kustannukset, jotka eivät liity rataverkon haltijan suorittamiin maksuihin. Kustannukset tai kustannuspaikat, jotka eivät liity suoraan vähimmäiskäyttömahdollisuuksien tai palvelupaikkoja yhdistävän infrastruktuurin käyttöoikeuksien tarjontaan;
  - c. kustannukset, jotka aiheutuvat maan ja muun kiinteän omaisuuden hankinnasta, myynnistä, purkamisesta, puhdistuksesta, uudelleenistutuksista tai vuokraamisesta;
  - d. koko verkkoa koskevat yleiskustannukset, mukaan lukien palkka- ja eläkekulut;
  - e. rahoituskustannukset;
  - f. kustannukset, jotka liittyvät tekniikan edistykseen tai vanhentumiseen;
  - g. aineettoman omaisuuden kustannukset;
  - h. radanvarren anturien, radanvarren viestintälaitteiden ja merkinantolaitteiston kustannukset, jotka eivät aiheudu suoraan rautatieliikenteen harjoittamisesta;
  - i. informaatiolaitteiden, muiden kuin radanvarren viestintälaitteiden tai televiestintälaitteiston kustannukset;
  - j. yksittäiseen ylivoimaiseen esteeseen, onnettomuuteen ja palveluhäiriöön liittyvät kustannukset, sanotun kuitenkin rajoittamatta direktiivin 2012/34/EU 35 artiklan soveltamista;
  - k. sähkönsyöttölaitteiston kustannukset, jotka eivät aiheudu suoraan rautatieliikenteen harjoittamisesta. Sellaisen rautatieliikenteen suoriin kustannuksiin, jossa ei käytetä sähkönsyöttölaitteistoa, ei saa sisältyä sähkönsyöttölaitteiston käyttökustannuksia;
  - l. kustannukset, jotka liittyvät direktiivin 2012/34/EU liitteessä II olevan 1 kohdan f alakohdassa mainittujen tietojen toimittamiseen, paitsi jos ne aiheutuvat rautatieliikenteen harjoittamisesta;
  - m. hallintokustannukset, jotka aiheutuvat direktiivin 2012/34/EU 31 artiklan 5 kohdassa ja 32 artiklan 4 kohdassa tarkoitetuista eriytettyjen maksujen järjestelmistä;
  - n. poistot, jotka eivät määräydy rautatieliikenteen harjoittamisesta aiheutuvan, rataverkon tosiasiallisen kulumisen perusteella;
  - o. rakennusteknisen infrastruktuurin kunnossapitoon ja uusimiseen liittyvä kustannusosuus, joka ei aiheudu suoraan rautatieliikenteen harjoittamisesta.

## Mallinnusten yhteenvedot

### Malli 1

#### MODEL INFO:

*Observations:* 638

*Dependent Variable:* ln\_eur

*Type:* OLS linear regression

#### MODEL FIT:

$F(2,635) = 238.10$ ,  $p = 0.00$

$R^2 = 0.43$

*Adj. R<sup>2</sup> = 0.43*

#### *Standard errors: OLS*

	Est.	S.E.	t val.	p
(Intercept)	7.66	0.34	22.44	0.00
ln_brt	0.21	0.02	9.85	0.00
ln_rdkm	0.73	0.04	18.38	0.00

#### studentized Breusch-Pagan test

data: malli\_log

BP = 1.0852, df = 2, p-value = 0.5812

#### Durbin-watson test

data: malli\_log

DW = 0.90871, p-value < 2.2e-16

alternative hypothesis: true autocorrelation is greater than 0

#### Jarque Bera Test

data: virhetermit\_log

X-squared = 48.85, df = 2, p-value = 2.468e-11

#### Variance Inflation Factor

ln\_brt ln\_rdkm

1.010326 1.010326

**Malli 2**MODEL INFO:*Observations:* 638*Dependent Variable:* eur\_bc*Type:* OLS linear regressionMODEL FIT: $F(2,635) = 239.50, p = 0.00$  $R^2 = 0.43$ *Adj. R<sup>2</sup> = 0.43**Standard errors: OLS*

	Est.	S.E.	t val.	p
(Intercept)	7.39	0.32	22.95	0.00
bcbrrt	0.21	0.02	9.85	0.00
bcrd_km	0.68	0.04	18.44	0.00

## studentized Breusch-Pagan test

data: malli\_bc2

BP = 1.43, df = 2, p-value = 0.4892

## Durbin-watson test

data: malli\_bc2

DW = 0.90796, p-value &lt; 2.2e-16

alternative hypothesis: true autocorrelation is greater than 0

## Jarque Bera Test

data: virhetermit\_bc2

X-squared = 44.62, df = 2, p-value = 2.046e-10

## Variance Inflation Factor

bcbrrt bcrd\_km

1.010502 1.010502

**Malli 3**MODEL INFO:*Observations:* 638*Dependent Variable:* eur\_bc*Type:* OLS linear regressionMODEL FIT: $F(2,635) = 244.83, p = 0.00$  $R^2 = 0.44$ *Adj. R<sup>2</sup> = 0.43**Standard errors: OLS*

	Est.	S.E.	t val.	p
(Intercept)	9.57	0.16	61.74	0.00
btbrt	0.01	0.00	10.21	0.00
btrd_km	0.62	0.03	18.75	0.00

## studentized Breusch-Pagan test

data: malli\_bt

BP = 0.5673, df = 2, p-value = 0.753

## Durbin-watson test

data: malli\_bt

DW = 0.90226, p-value &lt; 2.2e-16

alternative hypothesis: true autocorrelation is greater than 0

## Jarque Bera Test

data: virhetermit\_bt

X-squared = 45.881, df = 2, p-value = 1.089e-10

## Variance Inflation Factor

btbrt btrd\_km

1.006442 1.006442

**Malli 1a**MODEL INFO:*Observations:* 108*Dependent Variable:* ln\_eur*Type:* OLS linear regressionMODEL FIT: $F(2,105) = 61.82, p = 0.00$  $R^2 = 0.54$ *Adj. R<sup>2</sup> = 0.53**Standard errors: OLS*

	Est.	S.E.	t val.	p
(Intercept)	9.09	0.72	12.65	0.00
ln_brt	0.22	0.04	5.34	0.00
ln_rdkm	0.73	0.08	8.91	0.00

## studentized Breusch-Pagan test

data: malli\_log

BP = 0.68905, df = 2, p-value = 0.7086

## Durbin-watson test

data: malli\_log

DW = 1.4855, p-value = 0.003097

alternative hypothesis: true autocorrelation is greater than 0

## Jarque Bera Test

data: virhetermit\_log

X-squared = 2.9341, df = 2, p-value = 0.2306

## Variance Inflation Factor

ln\_brt ln\_rdkm

1.020309 1.020309

**Malli 1b**MODEL INFO:*Observations:* 588*Dependent Variable:* ln\_eur*Type:* OLS linear regressionMODEL FIT: $F(2,585) = 263.56, p = 0.00$  $R^2 = 0.47$ *Adj. R<sup>2</sup> = 0.47**Standard errors: OLS*

	Est.	S.E.	t val.	p
(Intercept)	8.02	0.33	24.51	0.00
ln_brt	0.18	0.02	8.96	0.00
ln_rdkm	0.74	0.04	20.80	0.00

## studentized Breusch-Pagan test

data: malli\_log

BP = 1.3241, df = 2, p-value = 0.5158

## Durbin-Watson test

data: malli\_log

DW = 0.99319, p-value &lt; 2.2e-16

alternative hypothesis: true autocorrelation is greater than 0

## Jarque Bera Test

data: virhetermit\_log

X-squared = 12.105, df = 2, p-value = 0.00235

## Variance Inflation Factor

ln\_brt ln\_rdkm

1.001343 1.001343



**Malli 1c**MODEL INFO:*Observations:* 638*Dependent variable:* ln\_eur*Type:* OLS linear regressionMODEL FIT: $F(3,634) = 373.27, p = 0.00$  $R^2 = 0.64$ *Adj. R<sup>2</sup> = 0.64**Standard errors: OLS*

	Est.	S.E.	t val.	p
(Intercept)	9.05	0.28	32.18	0.00
ln_brt	0.12	0.02	6.80	0.00
ln_rdkm	0.65	0.03	20.10	0.00
korv_dummy	1.34	0.07	19.19	0.00

studentized Breusch-Pagan test

data: malli\_log

BP = 32.978, df = 3, p-value = 3.255e-07

Durbin-watson test

data: malli\_log

DW = 1.0308, p-value &lt; 2.2e-16

alternative hypothesis: true autocorrelation is greater than 0

Jarque Bera Test

data: virhetermit\_log

X-squared = 2.5749, df = 2, p-value = 0.276

Variance Inflation Factor

ln_brt	ln_rdkm	korv_dummy
1.087706	1.031533	1.106778



Väylävirasto  
Trafikledsverket

ISSN 2490-1202  
ISBN 978-952-317-829-8  
[www.vayla.fi](http://www.vayla.fi)