

Vesi- ja ympäristötekniikan maisterin koulutusohjelma

# Vesijohtojen käyttöiän arviointi saneeraustietojen pohjalta

Diplomityö

---

Elli-Noora Tiainen

Copyright ©2023 Elli-Noora Tiainen

---

**Author** Elli-Noora Tiainen

---

**Title of thesis** Service life assessment of a water supply network based on renovation data

---

**Programme** Water and Environmental Engineering

---

**Thesis supervisor** Professor Riku Vahala

---

**Thesis advisor(s)** TkT Tuija Laakso

---

**Collaborative partner** Ramboll Finland

---

**Date** 23.5.2023

**Number of pages** 55

**Language** Finnish

---

### **Abstract**

Finland's water supply network is aging and most of the network is about to be rehabilitated for the first time. When timed optimally, rehabilitation maintains the condition of the infrastructure and is resource efficient. Determining the actual need for rehabilitation for water pipelines is challenging because the network is hidden underground, and up-to-date data on its condition is often unavailable. Mathematical methods can be used to model the development of a specific network's rehabilitation needs and to estimate its actual service life.

This Master's thesis explored how the service life of a water supply network can be estimated with the help of lifetime analysis. This was the first time that service lifetimes of water supply networks were estimated using rehabilitation data from Finnish water utilities. Three water utilities participated in the study: Alva company Ltd, Kuopion Vesi Ltd and Hangon Vesi. The service life of the Alva and Kuopio water supply networks was modelled using the Weibull distribution. One model for the whole network as well as separate models for the most common material groups were created for both utilities. Furthermore, the Cox Proportional Hazards Model was used to evaluate the role of different independent variables in connection to service life.

Based on these models, the actual service life of a water supply network was estimated to be close to the current expected lifetime of 60 years, but more precise predictions would require better data quality. The biggest uncertainty was caused by inadequate rehabilitation data, which significantly affected the reliability of the estimates. Notwithstanding, lifetime analysis using rehabilitation data has the potential to support rehabilitation decisions and resource allocation, but the quality of available asset data directly affects the reliability of the estimates. Water utilities are encouraged to systematically collect asset data which can then be used in lifetime analysis.

---

**Keywords** water supply network, service life, rehabilitation, asset data

---

---

**Tekijä** Elli-Noora Tiainen

---

**Työn nimi** Vesijohtojen käyttöiän arviointi saneeraustietojen pohjalta

---

**Koulutusohjelma** Vesi- ja ympäristötekniikka

---

**Vastuupettaja/valvoja** Professori Riku Vahala

---

**Työn ohjaaja(t)** TkT Tuija Laakso

---

**Yhteistyötaho** Ramboll Finland

---

**Päivämäärä** 23.5.2023

**Sivumäärä** 55

**Kieli** suomi

---

### **Tiivistelmä**

Suomen vesijohdot ikääntyvät ja suurin osa vesijohtoverkostosta on tulossa ensi kertaa saneerattavaksi. Optimaalisesti ajoitettuna saneeraus ylläpitää infrastruktuurin kuntoa ja on resurssitehokasta. Todellisen saneeraustarpeen määrittäminen on kuitenkin hankalaa, koska vesijohdot sijaitsevat piilossa maan alla eikä ajantasaista kuntotietoa ole yleensä saatavilla. Matemaattisten menetelmien avulla on mahdollista mallintaa saneeraustarpeen kehittymistä sekä arvioida verkoston todellista käyttöikä.

Diplomityössä vesijohtoverkoston käyttöikä arvioitiin elinaika-analyysin avulla. Käyttöiän mallinnus tehtiin ensimmäistä kertaa suomalaisen vesijohtoverkoston saneeraustietojen pohjalta. Lopputyöhön osallistuneet vesilaitokset olivat Alva-yhtiöt Oy, Kuopion Vesi Oy ja Hangon Vesi. Alvan ja Kuopion vesijohtoverkostojen käyttöiän jakautumista mallinnettiin Weibull-jakauman avulla. Kummallekin laitokselle luotiin yksi koko verkoston käyttöikä kuvaava malli sekä omat mallit yleisimmille materiaalityypille. Lisäksi eri selittäjien vaikutusta käyttöiän ennusteseen arvioitiin Coxin suhteellisen vaaran mallin avulla.

Mallinnustulosten perusteella vesijohtojen todellisen keskimääräisen käyttöiän arvioitiin olevan lähellä oletettua 60 vuotta, mutta tarkempien ennusteiden tekeminen edellyttää laadukkaampaa dataa. Suurin epävarmuustekijä mallinnuksessa oli puutteelliset saneeraustiedot, jotka vaikuttivat merkittävästi ennusteiden luotettavuuteen. Saneeraustietoihin pohjautuvalla elinajanmallinnuksella on potentiaalia toimia saneerauspäätösten ja resurssien kohdentamisen tukena, mutta mallinnustulosten luotettavuus edellyttää riittävästi dataa verkostossa tehdyistä saneerauksista. Vesilaitoksia kannustetaan verkostotietojen systemaattiseen keräämiseen, jotta aineistoa voidaan hyödyntää käyttöiän mallinnuksessa.

---

**Avainsanat** vesijohtoverkosto, käyttöikä, saneeraustarve, verkostotieto

---

# Sisällys

|       |  |    |
|-------|--|----|
| 1     | Johdanto.....  | 1  |
| 2     | Suomen vesijohtojen saneeraustarve.....                | 2  |
| 2.1   | Kuntotutkimukset .....                                 | 2  |
| 2.2   | Saneeraustarpeen arvioita .....                        | 2  |
| 2.3   | Saneerausinvestointien arvioita.....                   | 3  |
| 2.4   | Verkostotietojen kerääminen.....                       | 3  |
| 2.5   | Saneeraustarpeen jakautuminen .....                    | 4  |
| 2.6   | Saneerausten rahoitus.....                             | 4  |
| 2.7   | Johtopäätökset .....                                   | 5  |
| 3     | Vesijohtojen kunnonmallinnus.....                      | 6  |
| 3.1   | Kunnonmallinnuksessa käytettävät menetelmät .....      | 6  |
| 3.2   | Katkaistu ja sensuroitu lähtöaineisto .....            | 7  |
| 3.3   | Esimerkkejä vesijohtojen käyttöiän mallinnuksesta..... | 9  |
| 3.3.1 | Suppeat lähtötiedot .....                              | 9  |
| 3.3.2 | Tärkeimmät selittäjät .....                            | 11 |
| 3.3.3 | Putkiosien pituus.....                                 | 12 |
| 3.3.4 | Validointi ja kalibrointi .....                        | 12 |
| 3.3.5 | Yhteenveto .....                                       | 13 |
| 4     | Aineisto ja menetelmät.....                            | 15 |
| 4.1   | Weibull .....  | 15 |
| 4.2   | Coxin suhteellisen vaaran malli .....                  | 17 |
| 4.3   | Käytetyt kirjastot.....                                | 17 |
| 4.4   | Hyvyyden arviointi .....                               | 18 |
| 4.5   | Lähtöaineistot .....                                   | 20 |
| 4.6   | Aineistojen tulkinta.....                              | 20 |
| 5     | Alva .....   | 22 |
| 5.1   | Lähtöaineiston yleiskatsaus .....                      | 22 |
| 5.1.1 | Saneeraustiedot .....                                  | 22 |
| 5.1.2 | Mallinnetut materiaalit .....                          | 24 |
| 5.2   | Mallinnustulokset .....                                | 26 |
| 5.2.1 | Weibull.....   | 26 |

|       |  |    |
|-------|--|----|
| 5.2.2 | Weibull-ennuste putkiryhmittäin .....  | 27 |
| 5.2.3 | Coxin suhteellisen vaaran malli .....  | 28 |
| 5.2.4 | Hyvyyden arviointi .....   | 30 |
| 6     | Kuopion Vesi .....   | 33 |
| 6.1   | Lähtöaineiston yleiskatsaus .....  | 33 |
| 6.1.1 | Saneeraustiedot .....  | 33 |
| 6.1.2 | Mallinnetut materiaalit .....  | 34 |
| 6.2   | Mallinnustulokset .....  | 36 |
| 6.2.1 | Weibull.....   | 36 |
| 6.2.2 | Weibull-mallit putkiryhmittäin.....  | 37 |
| 6.2.3 | Coxin suhteellisen vaaran malli .....  | 38 |
| 6.2.4 | Hyvyyden arviointi .....   | 40 |
| 7     | Hangon Vesi .....  | 43 |
| 7.1   | Lähtöaineiston yleiskatsaus .....  | 43 |
| 7.2   | Aineiston tarkastelun tulokset .....   | 46 |
| 8     | Epävarmuustekijät .....  | 47 |
| 9     | Johtopäätökset .....   | 49 |
| 9.1   | Verkostotiedot käyttöiän mallinnuksessa.....   | 49 |
| 9.2   | Käyttöiän mallinnuksen rooli Suomen vesijohtojen saneeraustarpeen arvioinnissa ..... | 50 |
| 10    | Yhteenveto .....   | 52 |
| 11    | Lähteet .....  | 53 |

# 1 Johdanto

Ikääntyvän infrastruktuurin on pitkään katsottu olevan kriittisin Suomen vesihuollon haasteista (Kuulas ym. 2018). Verkostorakentaminen oli aktiivisimmillaan 1970-luvulla, ja vesijohtoverkoston kokonaismäärä on ollut kasvussa tästä lähtien (Maa- ja metsätalousministeriö 2008). Nyt, 50 vuotta myöhemmin, Suomen vesihuoltolaitokset ovat uuden haasteen kynnyksellä. Suuri määrä verkostoja on saavuttamassa samanaikaisesti arvioidun käyttöikänsä lopun, ja verkostojen saneeraustarve on muuttumassa jatkuvaksi kaikilla vesihuoltolaitoksilla (Berninger ym. 2018). Useissa selvityksissä on todettu nykyisen saneeraustahdin olevan riittämätön (Kuulas ym. 2018). Saneeraustarpeen arviot perustuvat oletettuun 60 vuoden käyttöikään, mutta verkoston todellinen käyttöikä voi olla tätä enemmän tai vähemmän.

Verkkotietojärjestelmien kehittymisen ansiosta myös osa Suomen vesilaitoksesta on kerännyt verkostojensa kunnosta ja toiminnasta dataa jo vuosikymmeniä. Tämän systemaattisesti kerätyn datan pohjalta voidaan luoda matemaattisia malleja, joiden avulla voidaan arvioida verkoston kunnan kehittymistä tulevaisuudessa. Nyt, saneeraustarpeen huipun lähestyessä, on ensimmäistä kertaa mahdollisuus hyödyntää tätä kerättyä dataa ennakoivien mallien luomiseen, ja arvioida vesijohtojen todellista käyttöikää suomalaisen aineiston pohjalta. Arvioita voidaan hyödyntää tulevien saneerauspäätösten ja resurssien kohdentamisen tukena. Tällaisella mallinnuksella on siten potentiaalia olla toimintavarmaa ja kestäväää vesihuoltoa tukeva työkalu.

Diplomityössä mallinnetaan putkien käyttöikää olemassa olevien saneeraustietojen pohjalta, jotka on saatu kolmelta eri suomalaiselta vesilaitokselta. Työssä tuotetaan tietoa vesijohtojen käyttöiän jakautumisesta, mikä puolestaan tukee arvioita tarvittavista saneerausinvestoinnista ja esimerkiksi kuntotutkimuksiin varattavista resursseista. Työn tavoitteena on arvioida, kuinka hyvin olemassa olevat aineistot soveltuvat käyttöiän mallinnukseen ja mitä mahdollisuuksia vesijohtojen käyttöiän mallinnus tuo vesijohtoverkoston saneeraustarpeen arviointiin. Työssä arvioidaan luotujen mallien luotettavuutta ja puuttuvien historiatietojen sekä muusta kuin huonosta kunnosta tehtyjen saneerausten vaikutusta lopputulokseen. Lisäksi työssä pohditaan, miten vesihuoltolaitosten kannattaa kerätä verkoston kunto ja saneeraustietoja, jotta aineisto tukisi käyttöiän mallinnusta mahdollisimman hyvin.

Lopputyö rakentuu seuraavasti: 1) kirjallisuuskatsaus Suomen vesijohtojen saneeraustarpeesta ja vesijohtojen käyttöiän mallinnuksesta, 2) mallinnusmenetelmien esittely, 3) lähdemateriaalien sekä mallinnustulosten esittely, 4) epävarmuustekijöiden arviointi 5) pohdintaa aineistojen soveltuvuudesta käyttöiän mallinnukseen, 6) yhteenveto.

## **2 Suomen vesijohtojen saneeraustarve**

Saneeraus on osa vesihuoltolaitosten riskienhallintaa, sillä sen avulla voidaan ennaltaehkäistä verkoston huonosta kunnosta aiheutuvia putkirikkoja. Verkoston huono kunto saattaa johtaa vedenjakelun katkeamiseen, veden laadun heikkenemiseen tai saastumiseen. (Berninger ym. 2018.) Osia verkostoista on saneerattu huonolaatuisten materiaalien ja asennusvikojen takia tai katusaneerauksien yhteydessä, mutta valtaosa verkostoista on tulossa ensimmäistä kertaa saneerausikään (Määttä 2021). Saneeraustarpeen estimointiin voidaan hyödyntää kuntotutkimuksia, asiantuntija-arvioita tai matemaattisia malleja.

### **2.1 Kuntotutkimukset**

Vesijohtojen nykyisen kunnan ja saneeraustarpeen määrittäminen on vaikeaa, sillä vesijohtoverkostot sijaitsevat pääosin piilossa maan alla. Yksittäisten putkien kuvaaminen vaatii erikoisjärjestelyjä kuten putken huolellista tyhjennystä ja desinfiointia. Eri putkimateriaalit, halkaisijat ja verkostorakenteet edellyttävät erilaisia kuntotutkimusmenetelmiä. Vesijohtojen kunnosta voidaan kuitenkin tehdä päätelmiä eri menetelmin. Verkoston vuotovesitilannetta voidaan arvioida alueittain, mikä tukee yksityiskohtaisempien kuntotutkimusmenetelmien kohdistamista niitä vaativiin vesijohtoihin. Näitä tarkempia menetelmiä ovat mm. maatutka, jonka avulla voidaan selvittää putken vuotokohtia tarkemmin ja lämpökamera, jonka avulla voidaan saada sijainnin lisäksi tietoa myös mahdollisista vuodoista. Tutkimuksia kohdistetaan vain verkoston joihinkin alueisiin tai linjoihin, joten ne eivät suoraan kerro koko verkoston kuntoa. Kuntotutkimuksella kerättyjä kuntotietoja voidaan hyödyntää saneerauspäätösten tekoon lyhyellä aikavälillä tai ennustaa tulevaa tarvetta erilaisten matemaattisten mallien avulla. Verkoston jatkuvat kuntotutkimukset ovat ennustavista malleista huolimatta tarpeellisia, sillä kuntotutkimusten avulla malleja voidaan päivittää ja ennusteiden oikeellisuus varmistaa. (Berninger ym. 2018.)

### **2.2 Saneeraustarpeen arvioita**

Vesijohtoverkoston saneeraustarvetta kansallisella tasolla on arvioitu viimeksi vuonna 2020 hankkeessa “Vesihuollon investointitarpeet vuoteen 2040” (Kuulas ym. 2020). Tuolloinen arvio perustui oletettuun 60 vuoden käyttöikään. Todellisuudessa osa putkista kestää arvioitua käyttöikää pidempään ja osa lyhyempään. Suomen vesijohdoista noin neljännes on rakennettu ennen vuotta 1980 (Kuulas ym. 2020), joten 60 vuoden käyttöikäoletuksella nämä kaikki on saneerattava vuoteen 2040 mennessä. Nykyinen

saneeraustaso on riittämätön, jos suuren osan verkostoista odotetaan saavuttavan saneerausiän samaan aikaan.

ROTI 2017- raportissa riittäväksi saneeraustasoksi arvioitiin 3 % verkostopituudesta vuodessa seuraavaksi kymmeneksi vuodeksi (Suomen Rakennusinsinöörien Liitto RIL ry. 2017). Suomessa on kunnallista vedenjakuverkkoa 107 000 km (Berninger y. 2018), joten 3 % saneeraustahti tarkoittaisi noin 3200 km saneerattavaa verkostopituutta. Jos verkoston todellinen käyttöikä on enemmän kuin oletettu 60 vuotta, ei saneeraustaso välttämättä ole riittämätön.

### **2.3 Saneerausinvestointien arvioita**

Vesihuoltoverkostojen investointitarve rakentuu saneeraustarpeen arviosta ja verkostosaneerauksen kustannuksista. Vesihuollon investointitarpeen arvioinnin haasteiksi on todettu datan puute ja hajanaisuus, minkä vuoksi arviot investointitarpeista eroavat usein toisistaan jopa useita kertaluokkia (OECD 2018). ROTI 2021 -raportissa arvioidaan, että Suomen vesihuollon investoinnit on nostettava jatkossa 770 miljoonaan euroon vuodessa nykyisestä 400 miljoonasta. Vuosien 2020–2040 arvioitu investointitarve on jopa yli viisinkertainen kuin nykyinen vuosittainen investointi. (Suomen Rakennusinsinöörien Liitto RIL ry. 2021.) Koska verkoston kuntoa ja saneeraustarvetta ei tunneta tarkasti, sisältyy arvioihin epävarmuutta.

Saneerauskustannuksille löytyy likimääräisiä arvioita kirjallisuudesta. Saneeraus on halvempaa harvaan asutuilla alueilla kuin kaupunkiympäristössä, joten erikokoisten laitosten välillä voi olla suuriakin eroja saneerauskustannuksissa. Yhden kilometrin saneeraushinnaksi on arvioitu 80 000–337 000 e/km. On hyvä muistaa, että kustannusten virhearviot kertautuvat pidemmillä matkoilla, ja että eri aikaan tehdyt saneerausvelka-arviot eivät ole vertailukelpoisia, sillä verkoston uusiutuminen ja leviäminen muuttaa materiaali- ja ikäjakaumaa. Saneerausvelalla tarkoitetaan jo käyttöikänsä ylittänyttä mutta vielä saneeraamatonta verkoston osuutta. Kustannusarvioiden epämääräisyys hankaloittaa vesihuoltolaitosten valmistautumista tuleviin saneerauksiin. (Kuulas ym. 2020.)

### **2.4 Verkostotietojen kerääminen**

Suomessa laitosten suurin haaste vesijohtojen saneerausvelan määrittämiseen on puutteelliset verkostojen ikä- ja kuntotiedot (mm. Kuulas ym. 2020 & Berninger ym. 2018). Datalähtöisen omaisuudenhallinnan osuus erityisesti pienillä ja keskisuurilla laitoksilla on toistaiseksi matalaa. Suomessa on olemassa ympäristökeskuksen ylläpitämä vesihuoltolain (119/2001) 20 d §

mukainen vesihuollon tietojärjestelmä Veeti, johon kerätään tietoja kootusti. Järjestelmiin kerätyt tunnusluvut eivät kuitenkaan sisällä kohdennetumpaa vesijohtojen kuntotietoa, joten ne tukevat heikosti käyttöiän mallinnusta. Toistaiseksi kaikki vesilaitokset eivät ole tietoja keränneet. (Kuulas ym. 2020.) Jos verkoston kunnosta ei ole kerätty tietoa systemaattisesti, myös vesihuoltoverkostojen henkilöstön ikääntyminen ja eläköityminen luo omat haasteensa, kun tieto verkoston historiasta katoaa heidän mukanaan (Määttä 2021). Jotta lähtötietoa voitaisiin hyödyntää verkoston kunnan ja saneeraustarpeen arvioinnissa, tietojen keräämisen ja esittämisen täytyy olla nykyistä systemaattisempaa. Tähän kannustaa myös kansainvälinen standardointijärjestö ISO, jonka vedenjakelun standardi (ISO-24516-1, 2016) kehottaa vesilaitoksia seuraamaan keskeisiä tunnuslukuja ja hyödyntämään tätä tietoa omaisuudenhallinnassa. Suomeen on 12. tammikuuta 2023 tullut voimaan Valtioneuvoston asetus vesihuollon tietojärjestelmästä ja tiedottamisesta (Valtioneuvoston asetus vesihuollon tietojärjestelmästä ja tiedottamisesta 6/2023), joka edellyttää vesihuoltolaitosten vastaavan siitä, että verkoston perustiedot ovat tietojärjestelmässä ajan tasalla. Tietojärjestelmän säännöllinen päivittäminen mahdollistaa myös saneeraustarpeen paremman arvioinnin jatkossa.

## **2.5 Saneeraustarpeen jakautuminen**

Saneeraustarve ei jakaudu tasaisesti Suomen kunnissa. Optimaalinen saneerausajankohta voi vaihdella eri laitosten välillä, jos vastuualueen verkostot on rakennettu pääosin eri aikakaudella (Berninger ym. 2018). Saneeraus on usein ajankohtaista myöhemmin pienillä laitoksilla, joilla ei ole nykyisen arvon mukaan tarpeellisia resursseja verkoston ylläpitoon. Itse asiassa 20 suurinta laitosta omistaa 80 % vesihuollosta, mutta 23 suurinta laitosta omistaa vain 20 % vedenjakeluverkostosta (22 000 km). (Berninger ym. 2018.) Pienemmät laitokset omistavat siten suhteessa suuremman osan vesijohtoverkostoista, vaikka myydyn veden ja asiakkaiden määrä onkin vähäisempi. Tuoreimmassa rakennetun ympäristön tila eli ROTI 2023- raportissa todetaankin, että vesilaitosten yhdistymisen tarve on kasvanut Suomen eri puolilla riittävien resurssien varmistamiseksi (Suomen Rakennusinsinöörien Liitto RIL ry. 2023).

## **2.6 Saneerausten rahoitus**

Vesihuoltolain (119/2001) mukaan käyttömaksujen tulee kattaa pitkän aikavälin huoltoinvestoinnit eli saneerausten kustannukset, mutta tällä hetkellä moni vesihuoltolaitos ei täytä kyseistä kustannusvaatimusta (Silfverberg 2017). Suomessa vedenjakelun kustannukset ovat käyttäjille toistaiseksi suhteellisen halpoja. Veden käyttömaksujen mediaanikustannus on vain 0,11 e/m<sup>3</sup>, kun muualla Euroopassa mediaanikustannukset vaihtelevat 0,09 ja

0,51 välillä (European Commission 2013). Tarvittavien saneerauskustannuksen kattamiseksi vesihuollon vesimaksujen tulisi Vesilaitosyhdistyksen mukaan olla jopa 150 % nykyisestä (Silferberg 2017).

## **2.7 Johtopäätökset**

Optimaalisesti ajoitettuna saneeraus ylläpitää vesihuoltolaitoksen toimintavarmuutta ja on suunniteltu resurssiviisaasti ja taloudellisesti (Kuulas ym. 2020). Sopiva saneeraustaso tukee myös vesihuoltopalveluiden kohtuuhintaisuutta jatkossa, sillä vuotuiset saneerauskustannukset jakaantuvat tasaisemmin, kun saneeraustarpeen huippu alenee. Jos optimaalisessa saneeraustasossa ei pysytä, on riskinä vesihuollon toimintavarmuuden heikkeneminen ja saneerausvelan kasaantuminen, mikä tarkoittaa, että kerralla saneerattavaa voi olla kohta huomattava paljon (Kuulas ym. 2018). Myöskään saneeraaminen yli tarpeen ei ole resurssien hallinnan kannalta kestävää vaan optimaalisessa tilanteessa saneeraus perustuu todelliseen tarpeeseen (Berninger ym. 2018). Suurin haaste saneeraustarpeen arvioinnissa on puutteelliset verkostojen ikä ja kuntotiedot (mm. Kuulas ym. 2020 & Berninger ym. 2018).

### 3 Vesijohtojen kunnonmallinnus

Vesihuollossa on tarve datalähtöisen omaisuudenhallinnan lisäämiselle, jotta tulevaan saneeraustarpeen kasvuun voidaan valmistautua mahdollisimman hyvin. Kerättyä verkostotietoa voidaan hyödyntää esimerkiksi kunnonmallinnuksessa. Vesijohtojen kunnonmallinnuksen tavoitteena on toimia saneerauspäätösten, ennakoivien kunnossapitotoimien ja tarkempien kuntotutkimusten kohdentamisen tukena (Magi ym. 2022). Mallinnuksella on myös potentiaalia toimia tukena arvioidessa, miten saneerauskustannukset kannattaa jakaa, jos vesijohtoverkosta joudutaan saneeraamaan kunnan kannalta ennenaikaisesti esim. tierakentamisen yhteydessä.

#### 3.1 Kunnonmallinnuksessa käytettävät menetelmät

Kunnon arviointiin käytettävät ennustavat mallit voidaan jakaa fysikaalisiin, tilastollisiin ja datalähtöisiin malleihin (Giraldo-González & Rodríguez 2020). Hyvin epävarmoihin ja huonosti ymmärrettyihin tapauksiin voidaan myös käyttää muita datalähtöisiä menetelmiä mm. sumeaa logiikkaa ja heuristiikkaa, jotka perustuvat pitkälti asiantuntija-arvioihin (St. Clair & Sinha 2012). Fysikaaliset mallit vaativat usein näytepalojen ottoa ja laboratoriotutkimuksia, ja mallit sopivat kuvaamaan vain tietyn alueen verkostoa tai tiettyä putkea (St. Clair & Sinha 2012). Koska maan alla sijaitsevista putkista on vaikea saada fysikaalisiin malleihin riittävästi kuntotietoa, vesijohtojen kunnonmallinnukseen käytetään usein tilastollisia tai datalähtöisiä menetelmiä (Winkler ym. 2018). Datalähtöisillä malleilla tarkoitetaan mm. tekoälyyn ja koneoppimiseen perustuvia menetelmiä. Koneoppimisen avulla on potentiaalia luoda malleja, jotka tuottavat tarkkoja tuloksia selittäjien monimutkaisten riippuvuuksien avulla, mutta toistaiseksi tilastolliset mallit ovat tuottaneet tarkempia tuloksia. (Giraldo-González & Rodríguez 2020.) Tämä johtuu todennäköisesti mallien kouluttamisen vaatiman datan suuresta määrästä eikä kunnonmallinnukseen ole toistaiseksi saatavilla riittävästi aineistoa.

Tilastollisia malleja pidetään siten varteenotettavimpana vaihtoehtona vesihuoltoverkostojen kunnonmallinnuksessa (St. Clair & Sinha 2012; Vladeanu & Koo 2015). Mallit on kehitetty ennustamaan verkoston kunnon kehittymistä tulevaisuudessa useaa eri parametria hyödyntäen (Vladeanu & Koo 2015). Tilastollinen malli luodaan tyypillisesti putkiryhmittäin. Putkiryhmät voidaan muodostaa esimerkiksi materiaalin, halkaisijan tai asennusajankohdan mukaan. (St. Clair & Sinha 2012.) Mallinnettava tekijä voi olla putkirikko, putken tekninen käyttöikä tai kuntoluokitus (Ana & Bauwens 2010). Useimmiten mallinnettava kohde on putkirikkojen määrä, jota pidetään indikaattorina putken todellisen fysikaalisen kunnan

huononemisesta (Vladeanu & Koo 2015). Mallintamalla saadaan tällöin tietoa häiriön todennäköisyydestä valitussa putkiryhmässä. Osassa malleista riippuva muuttuja, kuten putken käyttöikä, voidaan yhdistää putken ominaisuus- ja kuntotietoihin (St. Clair & Sinha 2012), jolloin myös niiden vaikutusta putken käyttöikäen voidaan arvioida. Mallintamalla ei saada täydellistä ennustetta, mutta sen avulla voidaan arvioida huonon kunnan todennäköisyyttä (Berninger ym. 2018). Riittävät historiatiedot ovat tilastollisten mallien kynnysehto ja mallien epävarmuustekijät ovat sitä suurempia mitä vähemmän historiatietoja on saatavilla. Tämä pätee erityisesti uudemmille verkostoille tai sellaisille, joiden tietoa on vasta alettu keräämään. (St. Clair ja Sinha 2012.) Mallin tuottamasta ennusteesta saadaan luotettavampi, kun sitä päivitetään uuden kuntotiedon pohjalta (Berninger ym. 2018).

Elinaika-analyysi (survival analysis, time-to-event analysis) on tilastotieteen osa-alue, jossa analysoidaan tarkasteltavan tapahtuman, esim. vesijohtojen tapauksessa käyttöiän päättymisen tai putkirikon, todennäköisyyttä eri ajan hetkillä. Elinaika-analyysi kehitettiin alun perin ennustamaan väestönkasvua, mutta 1980-luvulla sen käyttöä alettiin soveltaa myös infrastruktuurin käyttöiän mallinnukseen. Mallit soveltuvat hyvin vesijohtojen käyttöiän mallintamiseen, koska ne mahdollistavat oikealta sensuroidun lähtödatan käyttämisen. (Baur & Herz 2002.) Vesijohtojen käyttöikä mallintava elinaikamalli kertoo, kuinka suuri osa samanikäisistä putkista on tietyllä ajanhetkellä edelleen hyväkuntoisia ja käytössä, ja kuinka suuri osa huonossa kunnossa tai saneerattu (Berninger ym. 2018). Euroopassa yleisintä on elinajanmallinutus putkiryhmittäin Herzin mallilla, joka on laadittu erityisesti vesihuoltoverkostojen käyttöiän mallinnukseen. Myös Weibull-jakaumaa käytetään usein käyttöiän jakautumisen mallintamiseen. (Vladeanu ja Koo 2015).

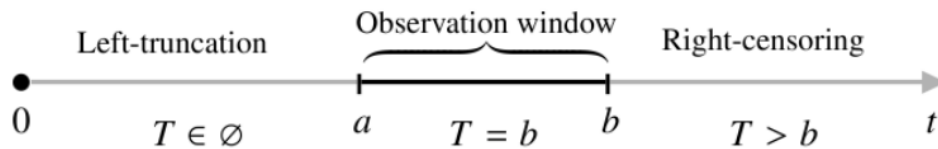
### **3.2 Katkaistu ja sensuroitu lähtöaineisto**

Käyttöiän mallinnuksen luotettavuus riippuu lähtöaineiston määrästä ja laadusta. Kappaleessa 2 pohdittiin systemaattisen verkostotietojen tallentamisen merkitystä, mutta toinen mallinnustuloksiin vaikuttava tekijä on mallinnettavan tapahtuman ajankohta suhteessa aineiston keräämisen ajankohtaan.

Vesilaitokset ovat kansainvälisestikin keränneet vesijohtoverkostojensa kuntotietoja vasta noin vuodesta 1985 lähtien, kun verkkotietojärjestelmät kehittyivät (Le Gat, Kropp & Poulton 2013). Koska vesijohtojen käyttöikä voi vaihdella suuresti, on todennäköistä, että tiedot ennen vuotta 1985 tehdyistä saneerauksista puuttuvat. Tällä tavoin epätäydellistä lähtöaineistoa kutsutaan käyttöiän mallinnuksessa vasemmalta katkaistuksi (left-truncacion) (Le Gat ym. 2013). Puuttuvat saneeraustiedot muuttavat elinaikamallin käyrää ja luovat todellisuutta positiivisemmän ennusteen (Le Gat ym. 2013). Tällöin

puhutaan ns. selviytymisharhasta (survival bias) (Introduction to survival analysis 2022).

Toisaalta osaa aineiston vesijohdoista ei ole vielä saneerattu tarkasteluhetkellä, mikä tarkoittaa, että mallinnettava tapahtuma eli lopputyön tapauksessa putken saneerausajankohta ei vielä ole tapahtunut. Tällöin puhutaan oikealta sensuroidusta aineistosta. (Le Gat ym. 2013). Jos oikealta sensuroituneet putket jätettäisiin huomioimatta kunnonmallinnuksessa, saataisiin lopputulokseksi todellisuutta lyhyempi käyttöikä (Introduction to survival analysis 2022). Kuva 1 havainnollistaa, miten katkaisu ja sensurointi riippuvat saneerausajankohdan löytymisestä aineistossa: jos putki on saneerattu ennen ajankohtaa  $a$ , on kyseessä vasemmalta katkaiseminen, jos taas saneeraus tapahtuu vasta tarkasteluajanjakson jälkeen eli ajankohdan  $b$  oikealla puolella, on kyseessä oikealta sensurointi.



Kuva 1. Havainnekuva sensuroinnin vaikutuksesta saneerausikään  $T$ . (Le Gat ja Eisbeis 2013)

- $T$  on vesijohdon saneerausajankohta.
- $T < a$ , vasemmalta katkaistu
- $T > b$ , oikealta sensuroitu
- $a < T < b$ , saneerausajankohta löytyy aineistosta
- $[a, b]$ , aineiston aikaikkuna

Sensuroidun tai katkaistun lähtöaineiston takia saneerausajankohdat on arvioitava koko vesijohtoverkostolle vain tiedossa olevien tapahtumien avulla, mikä vaikuttaa mallin luotettavuuteen (Le Gat ym. 2013). Mallin tuloksia on siten syytä tarkastella kriittisesti, mikäli verkostoa on saneerattu merkittävästi ennen 1980-lukua tai verkosto on pääosin poikkeuksellisen nuori. Verkostojen rakentaminen oli Suomessa vilkkaimmillaan 1970-luvulla, joten on todennäköistä, että valtaosalla Suomen laitoksista saneeraustarpeen huippu on vasta edessä, eikä saneerauksia ole siten ainakaan kaikilla laitoksilla tehty merkittävästi ennen kuin kunto- ja saneeraustietoja on alettu keräämään.

### 3.3 Esimerkkejä vesijohtojen käyttöiän mallinnuksesta

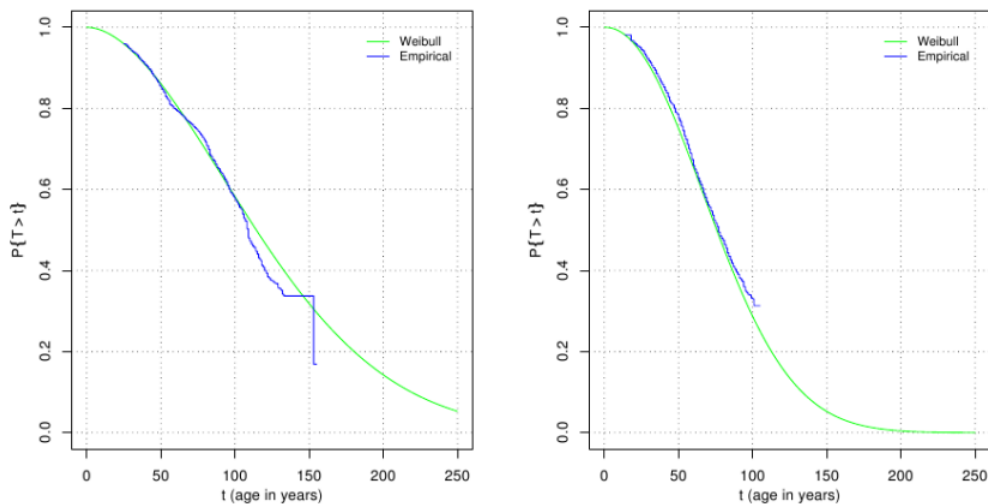
Kyseisessä kirjallisuuskatsauksen osiossa on esitelty lyhyesti erilaisia kansainvälisiä tutkimuksia vesijohtojen elinajanmallinnuksesta. Useimmissa tutkimuksista hyödynnettiin samanlaista mallinnusmenetelmää (Weibull) kuin lopputyössä. Vaikka käsitellyissä tutkimuksissa mallinnettiin useimmiten putkirikkoja, ovat niiden johtopäätökset hyödyllisiä myös lopputyön käyttöiän mallinnuksen kannalta. Vastaavaa saneeraustietoihin perustuva elinajanmallinnusta on vielä tehty hyvin vähän, ja suurin osa viemäriverkostolle eikä vesijohtoverkostolle. Kappaleessa esitellään tutkimuksia, joissa on samanlaisia reunaehtoja kuin diplomityössä.

#### 3.3.1 Suppeat lähtötiedot

Le Gat ja Eisenbeis (2000) esittivät verkoston kunnossapitotietojen hyödyntämistä elinajan mallinnuksessa putkimateriaaleittain. Mallinnusten toimivuutta vertailtiin kahden vesilaitoksen, Charente-Maritimen ja Lausannen, välillä. Toisella laitoksista oli tiedossa kuntohistoriatietoja pidemmältä ja toisella lyhyemmältä (vain 5–10 vuotta) ajalta. Lähtötietoina oli mm. putken tyypillisiä ominaisuustietoja kuten pituus, ikä, materiaali, paine, halkaisija, ja asennustapa, ympäristön tietoja kuten maaperä ja liikenteen määrä sekä kuntotietoja kuten aikaisempien putkirikkojen määrä. Putken ominaisuustietoja käytettiin mallin selittäjinä. Tapaustutkimuksissa elinajanmallinnukseen käytettiin Weibullin suhteellisen vaaran mallia (Weibull Proportional Hazard Model). Charente-Maritimen mallinnus osoitti, että Weibull-malli tuottaa luotettavia tuloksia myös suppeammilla lähtötiedoilla, kun aikaisemmissa tutkimuksissa mallia oli käytetty vain tapauksissa, joissa putkirikotietoja oli tallennettu pidemmän aikaa. Lisäksi työssä osoitettiin, että elinajanmallinnus onnistuu luotettavasti pelkästään putken ominaisuustiedoilla, mutta ympäristötietojen avulla ennustetta voidaan tämentää.

Restrepo ym. (2009) tutkimuksessa mallinnettiin raudasta valmistettujen vesijohtojen kuntoa. Mallinnusta varten kerättiin paljon dataa maaperästä ja verkostosta, ja selittäjiksi valittiin maaperän aggressiivisuus, putken nykyinen kunto ja putken ikä tarkasteluhetkellä. Mallinnukseen hyödynnettiin tilastollista analyysimenetelmää: suhteellisen osituksen otantamenetelmä (Proportionate Stratified Sampling Method). Verkoston nykyinen kunto oli arvioitu silmämääräisesti ulko- ja sisäpuolelta ja koepaloja vietiin laboratorioon jatkotutkimuksia varten. Verkoston ennustettu kunto esitettiin lopulta paikkatietomuodossa. Käytetyn menetelmän todettiin olevan hyödyllinen työkalu kunnossapito-ohjelmien kehittämisessä. Myös tässä tutkimuksessa datan puutteen, erityisesti putkirikkojen syistä ja kunnosta, koettiin olevan suurimpia haasteita, ja todettiin tilastollisten mallien onnistumisen pitkälti riippuvan lähtöaineiston datan laadusta.

Le Gat ym. (2013) hyödynsivät Weibull- elinaikamallia ja Turnbullin parametrifunktiota ranskalaisten teräsputkien käyttöiän mallintamiseen kahtena eri ajanjaksona (Kuva 2). Lähtöaineistona hyödynnettiin julkisesti saatavilla olevaa, katkaistua ja oikealta sensuroitua dataa. Mallien tuloksia vertailtiin ja todettiin parametrifunktion mallin eduksi sen kyky mukautua tarkkaan lähtöaineistoon, ja puutteeksi ettei se kykene luomaan ennusteita. Tästä syystä parametrifunktion mallia suositeltiin parametrifunktion Weibull-mallin validoinnin tueksi. Kuten kuvasta 2 näkyy Turnbullin parametrifunktio (sinisellä) näyttää Weibull-mallin (vihreällä) suuntaa ja tukee näin myös ennusteen oikeellisuutta, jos aineiston historiatietoa on riittävästi.



Kuva 2. Weibull-elinaikamalli ja Turnbullin parametrifunktion malli esitettynä ranskalaisille teräsputkille vuosille 1995–2007 (vasemmalla) ja 1985–2002 (oikealla) (Le Gat ym. 2013)

Valis ym. (2017) mallinsivat puolalaisen vedenjakeluverkoston putkirikkojen määrää käyttäen 15 vuoden ajalta kuukausittain kerättyjä putkirikkotietoja. Lähtötietojen mainittiin olevan hyvin rajallisia ja katkaistuja. Tutkimuksen tavoitteena oli arvioida putkirikkojen määrän kehitystä verkostossa ajan funktiona. Tähän käytettiin kahta menetelmää: Markovin ketjua ja Kalmanin rekursiota. Markovin ketjun käyttämisen mallinnuksessa teki haasteelliseksi se, ettei lähtöaineistosta selvinnyt putkirikkojen syytä. Kalmanin rekursion mainittiin kuitenkin sopivan hyvin vastaaviin tapauksiin, joissa lähtöaineistoa on rajallisesti.

### 3.3.2 Tärkeimmät selittäjät

Vanrenterghem-Raven (2007) arvioi miten Weibullin suhteellisen vaaran malli ennustaa putkirikkojen määrää vesijohdossa. Mallin opetukseen hyödynnettiin Long Islandin vedenjakeluverkoston olemassa olevaa dataa. Analyysin tavoitteena oli tunnistaa vedenjakeluverkoston kunnon heikkenemiseen vaikuttavia tekijöitä ja arvioida eri selittäjien tilastollista merkitystä. Analyysin perusteella huomattiin, että joillakin ympäristötekijöillä, kuten putken etäisyydellä valtatiehen, metrotunneleihin tai sijaintiin vesistöissä, oli havaittavissa yhteys putkirikkojen todennäköisyyteen, mutta tilastollinen vaikutus oli suhteellisen pieni verrattuna aiempaan putkirikkohistoriaan ja putkimateriaaliin, joka hyvin usein valitaankin selittäjäksi. Sijainnin risteys- tai ruuhka-alueella ei havaittu olevan yhteydessä putkirikkoihin, mikä todennäköisesti johtuu putkiosien lyhyemmästä pituudesta taajama-alueella. Lisäksi putken iän huomattiin olevan yksinään huono selittäjä, sillä asennusten laatu saattaa vaihdella aikakausittain, joten jotkut vanhemmat putket voivat kestää keskimäärin pidempään.

Wang ym. (2010) hyödynsivät bayesilaista päättelyä eri selittäjien suhteellisen vaikutuksen arviointiin. Tulosten perusteella putken ikä ja halkaisija olivat elinajanmallinnuksessa ominaisuustiedoista oleellisimpia, kun taas tien kaistojen määrä, kaivantosyvyys ja sähkövaraus hylättiin, niiden vähäisemmän painoarvon takia.

Pelletier ym. (2003) käyttivät kolmen eri vesilaitoksen dataa elinajanmallinukseen aiemmin toteutuneiden putkirikkojen perusteella. Weibull-jakaumaa hyödynnettiin ennustamaan ensimmäisten rikkojen määrää ja eksponentiaalijakaumaa tätä myöhempien. Elinaikamallin lisäksi tapahtuman havainnollistamiseen käytettiin tiheys- ja kertymäfunktioita. Lähtötiedoiksi valittiin halkaisija, pituus, materiaali, asennusvuosi, maaperä ja maankäyttö putken yläpuolella. Tulosten perusteella pystyttiin tunnistamaan kunnonkehitykseen vaikuttavia trendejä. Ennen vuotta 1960 asennetut putket osoittautuivat kestävämmiksi kuin vuoden 1960 jälkeen asennetut, minkä perusteella voidaan päätellä käytettyjen materiaalien ja/tai asennusmenetelmien muuttuneen. Toisaalta on mahdollista, että tutkimuksessa käytetty aineisto on vasemmalta katkaistu ja tulokseen vaikuttaa selvitysmisharha (survival bias).

Kleiner ja Rajani (2008) arvioivat eri muuttujien vaikutusta putkirikkojen määrään epähomogeenisen Poissonin mallin avulla. Mallin luomiseen käytettiin kanadalaisen vesilaitoksen dataa. Parametrit jaettiin kolmeen luokkaan: putkiriippuvaisiin (esim. halkaisija), aikariippuvaisiin (mm. maan kosteus ja jäätyminen) sekä näiden kombinaatioon (aikaisemmat putkirikot). Tulokset osoittivat, että mallit kannattaa muodostaa putki- tai

putkiryhmäkohtaisesti, jotta eri selittäjien painoarvoa voidaan analysoida. Tosin mallinnus yksittäisessä vesijohdossa tapahtuneiden putkirikkojen perusteella ei ole ainakaan toistaiseksi realistista datan vähäisen määrän vuoksi. Jako putkiryhmiin voidaan tehdä esimerkiksi materiaalin tai halkaisijan perusteella, jolloin mm. putkirikkojen määrän kehittymistä verkostossa voidaan arvioida paremmin (Kleiner ja Rajani 2008).

### 3.3.3 Putkiosien pituus

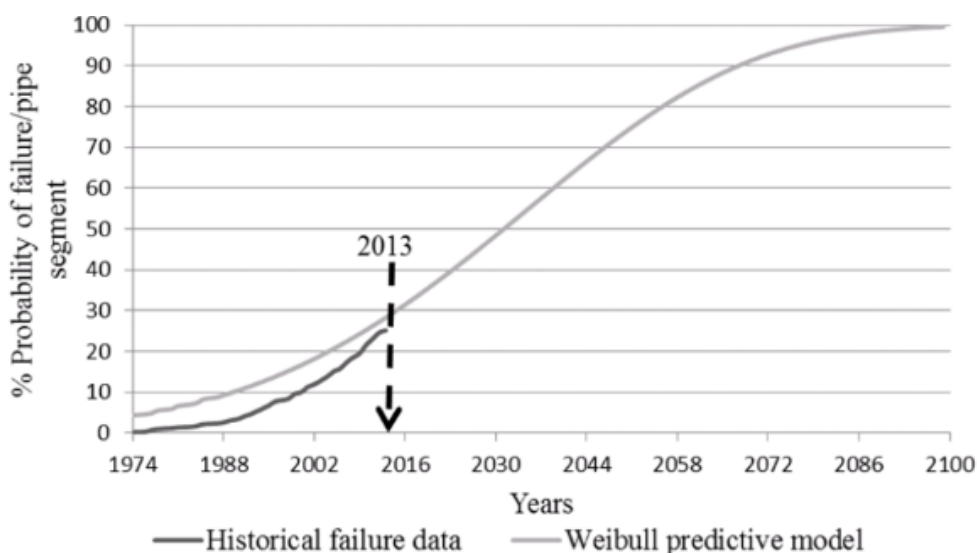
Poulton, Le Gat ja Bemond (2007) arvioivat vesijohdon putkiosuuden pituuden vaikutusta elinajanmallin putkirikkojen ennustustarkkuuteen. Putkirikkoja ennustettiin segmenteittäin LEYP-methodin (Linear Extension of the Yule Process) avulla, segmentin iän kehityksestä luotiin Weibull-malli ja eri muuttujien vaikutusta havainnollistettiin Coxin suhteellisen riskin mallissa. Tapaustutkimus tehtiin ranskalaiselle vesilaitokselle, jonka keräämästä aineistosta tarvittiin vähintään putken halkaisija ja pituus, asennusvuosi sekä putkirikon tyyppi ja ajankohta. Tutkijoiden harmiksi vesilaitoksen keräämässä datassa oli tulkinnanvaraisuuksia, joten osa aineistosta oli jätettävä kokonaan pois tai putken kuntotiedot oli arvioitava läheisten putkien tietojen perusteella eli ns. ketjutettava, minkä mainittiin heikentäneen mallin ennustustarkkuutta. Tulokset osoittivat mallin soveltuvan yhtä hyvin niin pitkille kuin lyhyille putkiosuuksille.

Wang ym. (2009) arvioivat putken pituuden ja muiden muuttujien vaikutusta putkirikkojen vuosittaisen määrään. Eri materiaalityypistä luotiin viisi regressiomallia. Putken muuttujien vaikutusta putkirikkojen määrään arvioidtiin herkkyysanalyysin avulla. Tutkimuksessa havaittiin, että vuosittainen putkirikkojen määrä on suurempi lyhyemmille putkille kuin pidemmille, ja määrä kasvaa putken ikääntyessä. Putkirikkojen todennäköisyys ei siis riipu lineaarisesti putkiosan pituudesta, joten eri pituisten putkien mallintaminen yhdessä ei ole välttämättä ongelma. (Wang ym. 2009.)

### 3.3.4 Validointi ja kalibrointi

Vladeanun ja Koon tutkimuksessa (2015) vertailtiin kahta tilastollista menetelmää vesijohtojen kunnonmallinnuksessa: Weibull-jakauman ja logistisen regression käyttöä. Työssä mallinnettiin putkirikkojen todennäköisyyttä vesijohdoissa. Lähtöaineisto koostui 100 vuotta vanhojen valurautaputkien putkirikkotiedoista kahdelta eri asennusaikakaudelta ja se saatiin amerikkalaiselta vesilaitokselta. Weibull-jakauman pohjalta luotua ennustetta vertailtiin todellisiin historiatietoihin. Kuvasta 3 näkyy mallintamalla saatu käyrä ja verkoston todelliset historiatiedot vuoteen 2013 saakka. Tulokset validoitiin vertailemalla aineiston viimeisen kuuden vuoden todellisia putkirikkoja mallien ennusteeseen. Tutkimuksen validointivaihe osoitti, että Weibull-

jakauma tuotti luotettavampia tuloksia kuin logistinen regressio, tosin tulokseen vaikutti mahdollisesti tutkimuksessa mallinnetun aineiston vähäinen määrä: 278 putkirikkoa 1105 putkiosaa kohden.



Kuva 3. 1900–1910 asennettujen valurautaputkien putkirikkojen todennäköisyys Weibull-jakaumalla kuvattuna (Vladeanu & Koo 2015)

Bruaset ym. (2017) tutkivat, voiko vesijohtoverkoston elinaikamallin epävarmuutta vähentää kalibroimalla sitä empiirisellä kuntotiedolla. Kalibroinnin osoitettiin tuottavan tarkempia tuloksia. Elinaikamallin puutteiksi mainittiin, ettei se ota huomioon ulkoisten muuttujien, kuten muuttuvan palvelutasotavoitteen, vaikutusta tulevaisuuden kunnostustarpeisiin. Tämän aspektin yhdistäminen mallinnukseen uskottiin johtavan tarkempiin ennusteisiin. Mainittujen puutteiden aiheuttavan epävarmuuden arvioitiin kuitenkin pienenevän, kun mallia kalibroidaan päivitetyllä tiedolla. Mallin päivittämistä uusilla tiedoilla suositeltiin vähintään 10 vuoden välein.

### 3.3.5 Yhteenveto

Useassa tutkimuksista todettiin elinajanmallinnuksen luotettavuuden riippuvan pitkälti lähtöaineiston laadusta ja määrästä (mm. Restrepo ym. 2009). Tämän tiedettiin jo etukäteen olevan haaste myös lopputyössä tehtävän mallinnuksen kannalta. Lopputyössä mallinnukseen hyödynnetäänkin Weibull-jakaumaa, jonka esitellyissä tutkimuksissa huomattiin tuottavan luotettavia tuloksia sekä lyhyemmän kuntohistorian (mm. Le Gat & Eisenbeis, 2000) että pienemmän otannan aineistolla (mm. Vladeanu & Koo 2015). Vesijohtoverkoston elinajanmallinnuksessa putken iän todettiin olevan yksinään huono selittäjä (Vanrenterghem-Raven ym. 2007; Pelletier ym. 2003), sillä erot eri aikakausien asennusten laadussa vaikuttavat saneeraustarpeeseen.

Jotta eri selittäjien merkitystä voidaan arvioida, kannattaa mallinnus tehdä myös putkiryhmittäin oleellisimman selittäjän perusteella (Kleiner & Rajani, 2008). Esitettyjen tutkimusten perusteella mallinnus onnistuu luotettavasti pelkästään putken ominaisuustiedoilla, mutta tarkempien ympäristö- ja sijaintitietojen avulla on mahdollista täsmentää ennustetta (mm. Le Gat & Eisenbeis, 2000). Mallien puuteeksi mainittiin, etteivät ne ota huomioon ulkoisten tekijöiden vaikutusta, joten on suositeltavaa päivittää mallin antama ennustetta päivitetystä tiedolla (Bruaset ym. 2017).

## 4 Aineisto ja menetelmät

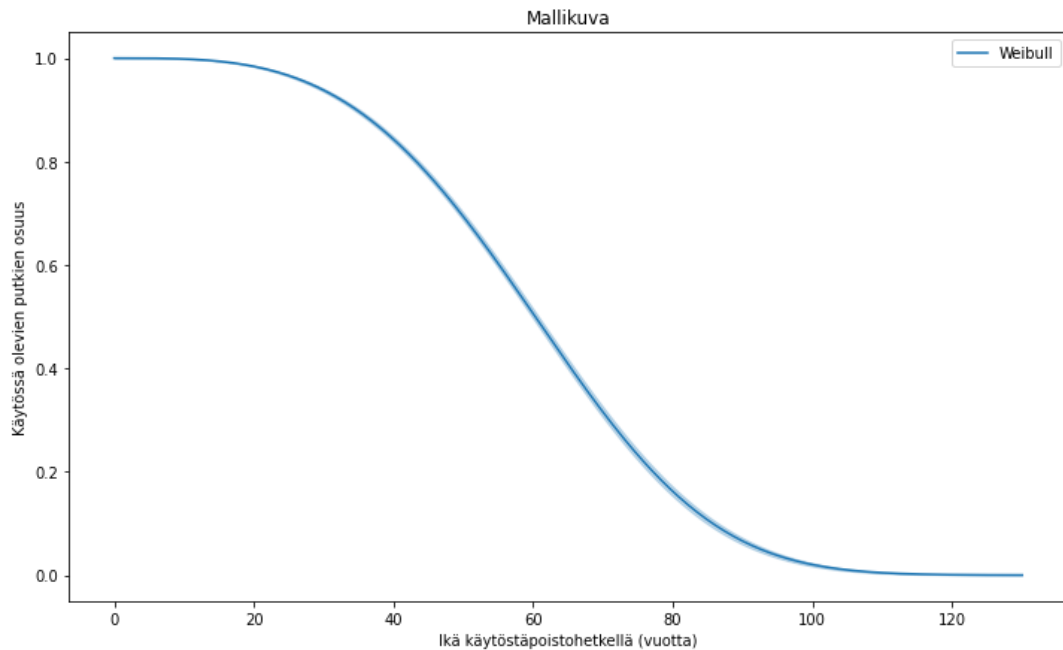
Diplomityössä arvioidaan vesijohtojen käyttöiän jakautumista kahden suomalaisen vesilaitoksen keräämän aineiston avulla. Lopputyöhön osallistuvat laitokset ovat Alva-yhtiöt Oy, Kuopion Vesi Oy ja Hangon Vesi.

Mallintamiseen käytetään Weibull-jakaumaa ja Coxin suhteellisen vaaran mallia. Mallinnus tehdään Python-ohjelmointikielellä, ja saatujen mallien luotettavuutta arvioidaan matemaattisin menetelmin. Mallinnuksen ja lähötödatan käsittelyn tueksi haastatellaan myös vesilaitosten edustajia. Haastatteluiden avulla pyritään selvittämään mallinnustuloksiin mahdollisesti vaikuttavia aineistojen ominaisuuksia mm. miten data on kerätty, mitä dataa puuttuu ja miksi ja onko verkostomateriaaleissa ollut eri vuosikymmeninä merkittäviä eroja.

### 4.1 Weibull

Weibullin parametrinen jakauma ennustaa tilastollisen tapahtuman todennäköisyyttä ajan funktiona. Jakaumaa hyödynnetään laajalti vesijohtoverkostojen käyttöiän mallinnuksessa. (Le Gat & Eisbeis 2000). Weibull-malli valikoitui lopputyöhön sen parametrien tarjoaman joustavuuden ansiosta sekä koska sitä on suhteellisen helppo käsitellä laskennallisesti useimmilla ohjelmointikielillä.

Lopputyössä tehtävässä käyttöiän mallinnuksessa mallinnettava tapahtuma on putken saneerausikä ja mallinnus toteutetaan putkiryhmittäin, kuten vesijohtojen kunnonmallinnuksessa on usein tapana (St. Clair & Sinha 2012). Tällöin Weibull-elinaikamalli kuvaa, miten tietyn putkiryhmän käyttöikä jakautuu ajan funktiona eli mikä on käytössä olevien putkien osuus kaikista kerran asennetuista. Esimerkiksi kuvan 4 mallin perusteella alle 20-vuotiaista putkista lähes kaikki ovat vielä käytössä, mutta 60 ikävuoteen mennessä käytössä olevien putkien määrä on laskenut lähelle nollaa. Mallin kalibrointiin käytetään olemassa olevia vesijohtoverkoston tietoja ja se validoidaan vertailemalla sen tuloksia empiirisesti tietyn aikavälin todellisiin arvoihin.



Kuva 4. Esimerkki elinaikamallista

Weibull-jakauma määräytyy kolmen muuttujan avulla: aika tapahtumahetkellä  $t$ , muotoparametri  $k$  ja asteikkoparametri  $\lambda$ . Jakauman yhtälö (1) on esitetty alla. Muotoparametri mahdollistaa myös vinot jakaumat (Vladeanu ja Koo 2015). Käyttöiän mallinnuksessa  $t$  kuvaa putken ikää,  $k$  muutoksen suuntaa eli sitä lisääntykö vai väheneekö käyttöikänsä päähän tulevien putkien määrä ajan myötä (failure rate increasing) ja  $\lambda$  kuvaa hajontaa. Mallin parametrien arvot määräytyvät, kun aineisto sovitetaan malliin. Koska jokaisella alueella on erityiset piirteensä, jotka vaikuttavat vesijohtoverkoston saneeraustarpeeseen, on epätodennäköistä, että yhtä universaalista kaikille sopivaa mallia olisi olemassa (Rostum 2000, Le Gat ja Eisebeis 2020). Tämän takia parametrien arvot tulee selvittää jokaiselle laitokselle erikseen.

Jakauman yhtälö (Vladeanu ja Koo 2015):

$$f(t) = \frac{k}{\lambda} \left(\frac{t-c}{\lambda}\right)^{k-1} e^{-\left(\frac{t-c}{\lambda}\right)^k} \quad (1)$$

- $f(t)$  on käyttöiän todennäköisyyden tiheysfunktio
- $t$  käyttöiän päättymisen ajankohta mallinnuksen alusta
- $k$  määrää käyrän muodon (symmetrinen tai vino)
- $c$  kuvaa putken asennusajankohtaa (putken ikä =  $t-c$ )

## 4.2 Coxin suhteellisen vaaran malli

Lopputyössä pääpaino on käyttöiän mallinnuksessa Weibull- elinaikamallin avulla, mutta lähdemateriaalia käytettiin myös toisen parametrin elinaikamallin opetukseen. Koska lopputyössä käytetty Weibull-malli ei kykene arvioimaan yksittäisten selittäjien vaikutusta vesijohtojen käyttöikäen (muuten kuin mallintamalla ne erikseen putkiryhmittäin) valittiin työhön Coxin suhteellisen vaaran malli (Cox Proportional Hazards Model). Cox-malli arvioi jokaisen putken käyttöiän jakautumista erikseen ja sen avulla selittäjien painoarvoa on mahdollista vertailla keskenään. Malli kertoo saneeraustarpeen todennäköisyyden sopivalla kertoimella selittäjästä riippuen (Survival regression 2022). Malli osaa tulkita sekä binäärisiä että jatkuvia selittäjiä, joten numeeriset selittäjät eivät vaatineet enempää käsittelyä, vaikka selittäjien arvot olisivat keskimäärin hyvin eri suuruisia.

## 4.3 Käytetyt kirjastot

Lopputyössä mallinnus tehdään käyttämällä python-ohjelmointikielen Lifelines- kirjastoa. Aineistot luettiin ohjelmaan csv-muodossa, johon käytettiin pythonin csv-moduulia. Aineistojen matemaattiseen käsittelyyn käytettiin lisäksi numpy, matplotlib ja pandas- kirjastoja.

Lifelines- kirjastosta löytyy valmis WeibullFitter-funktio. Funktio tarvitsee tiedon putken saneerausikästä tai tämänhetkisestä iästä (T) sekä totuusarvon (0 tai 1) putken saneeraustilasta (E). Kyseiset tiedot arvioitiin aineiston perusteella ja lisättiin samaan datakehykseen.

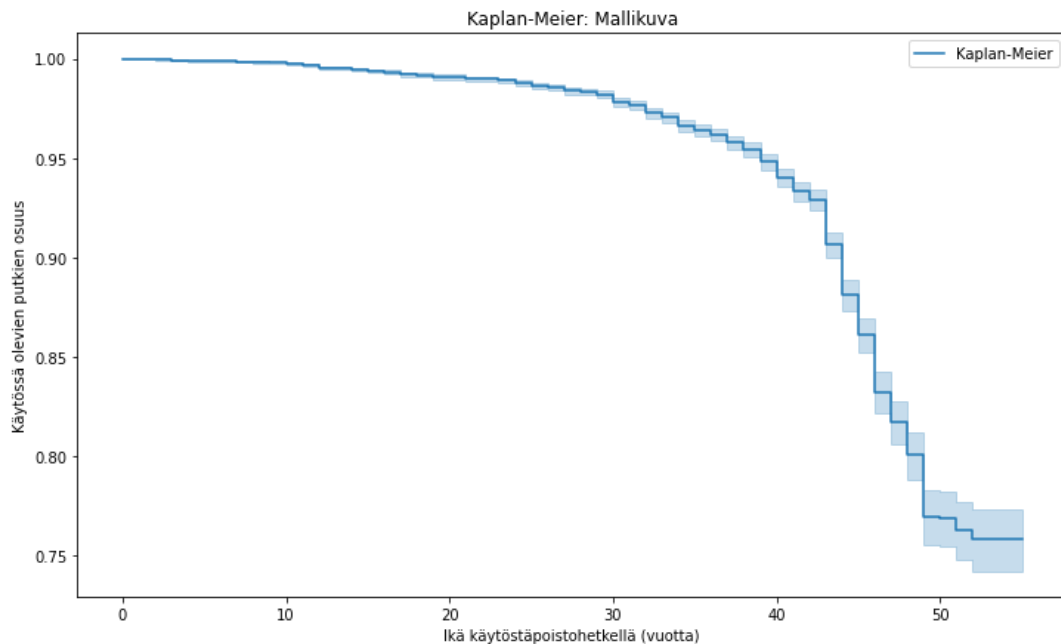
Käyttöikää mallinnettiin myös CoxPHFitter-funktioilla. T- ja E-muuttujien lisäksi malleihin valittiin selittäjiksi asennusvuosi, alkuperäinen asennusvuosi, käytöstäpoistovuosi, materiaali, pituus ja halkaisija. Joidenkin selittäjien alkuperäiset nimet poikkesivat hieman vesilaitosten aineistoissa. Mallit kykenevät tulkitsemaan vain numeerisia selittäjiä, joten mallinnusta varten valitut materiaalit oli muutettava kategorisiksi muuttujiksi. Koska valittuja materiaalikategorioita ei ollut montaa, voitiin tähän käyttää one-hot-encoding-menetelmää, joka tekee yhdestä kategorisesta muuttujasta vain nollia ja ykkösiä sisältäviä attribuutteja (datagy, 2022). Muut mallinnukseen sisällytetyt selittäjät olivat valmiiksi numeerisia.

Mallinnus tehtiin erikseen opetus- ja testiaineistolle. Opetettujen mallin tuloksia luettiin usean eri Lifelines-kirjastosta löytyvän funktion avulla.

## 4.4 Hyvyyden arviointi

Mallille suoritetaan tyypillisesti hyvyyden arviointi määrittämään, miten hyvin malli noudattaa havaintojen todellista jakaumaa. Hyvä malli kuvaa aineistoa hyvin ja sen pohjalta voidaan tehdä luotettavia ennusteita. (lifelines 2022)

Weibull-elinaikamallin hyvyyttä arvioitiin empiirisesti vertailemalla mallia Kaplan-Meier kertymäfunktioon. Oikein sovitettuna Weibull-mallin alku mukailee kertymäfunktioita likimääräisesti. Kaplan-Meier on parametrin elinaikamalli, joka kuvaa aineiston tilaa sellaisena kuin se on, ja on siksi hyvä tuki ennustavien elinaikamallien validoimiselle. Vesijohtojen käyttöiän mallinnuksessa Kaplan-Meierin kertymäfunktio kuvaa käytössä olevien putkien määrää verkostossa tiettyinä ajanhetkenä, mutta malli ei kykene ennustamaan tulevaisuuteen kuten parametriset elinaikamallit. Kun viimeinen aineistossa saneeratuksi merkitty putki saavuttaa käyttöikänsä lopun, putoaa kertymäfunktion käyrä alas, kuten kuvasta 5 on nähtävissä (mallikuva).



Kuva 5. Kaplan-Meier: Mallikuva

Coxin suhteellisen vaaran mallin hyvyyttä arvioitiin C-indeksin avulla. Ennen mallinnusta data jaettiin satunnaisesti testi- ja opetusaineistoon. Mallin opetukseen käytettiin 80 % koko aineistosta ja validointiin 20 % aineistosta. C-indeksin avulla voidaan varmistua siitä, että opetusaineistolla opetettu malli ei ole ylisovitettu vaan yleistyy hyvin testiaineistoon. C-indeksi asettaa mallin ennustamat arvot järjestykseen ja arvioi järjestyksen

oikeellisuutta luvulla 0 ja 1 välillä. Mitä korkeampi tulos sitä parempi malli on indeksin perusteella. (lifelines 2022.) Weibull-jakauman arviointiin se ei sovellu, sillä malli tuottaa koko verkostolle yhden elinajanennusteen. C-indeksin käyttö edellyttää ennustettuja arvoja, jotka voidaan laittaa järjestykseen.

Useamman mallinnusmenetelmän käyttö mahdollistaa mallien hyvyyden keskinäisen vertailun. Weibull-mallin ja Coxin suhteellisen vaaran mallin AIC ja log-todennäköisyyden (log-likelihood) arvot laskettiin ja niitä vertailtiin toisiinsa. AIC:n tapauksessa pienempi arvo kuvaa parempaa mallia, log-todennäköisyyden tapauksessa suurempi. Menetelmillä lasketut luvut eivät itsessään kuvaa mallin hyvyyttä, mutta sopivat juuri useamman mallin keskinäiseen vertailuun. Malleja vertailtaessa on huomioida, että Weibull-mallin opettamiseen käytettiin koko käytössä olevaa aineistoa ja Coxin tapauksessa aineisto jaettiin opetus- ja testausaineistoon.

## 4.5 Lähtöaineistot

Tarkasteluita varten lopputyöhön mukaan lähteneiltä vesilaitoksilta pyydettiin verkoston kaikkia putkia koskevat tiedot, sekä nykyään käytössä olevien että jo saneerattujen tai käytöstä poistettujen. Kattavan aineiston avulla on mahdollista arvioida koko verkoston elinkaarta. Mallinnusta varten tarvitaan vähintään tieto saneerattujen putkien asennus- ja saneerausvuosista. Aineistot pyydettiin lähtökohtaisesti taulukko- tai paikkatietomuodossa.

Vesilaitoksilta toivottiin seuraavia tietoja:

- putken perustiedot
  - tunnusluku
  - materiaali
  - halkaisija
  - rakennusvuosi
- muut putkeen kytketyt tiedot, jotka voisivat toimia käyttöikää selittävinä tekijöinä, esim.
  - asennussyvyys
  - sijainti vesistössä, tien alla tms. (sijaintitieto sellaiseen ei ollut välttämätön)
  - aiempi kunnossapitohistoria
  - mahdolliset muut kuntotiedot, esimerkiksi putkessa olleiden putkirikkojen määrä
  - tieto putkirikon syystä, jos tällainen on tiedossa
- putken saneeraustieto
  - jos on saneerattu, niin milloin
  - tieto saneerauksen syystä, jos tällainen on tiedossa
- tieto putken hydraulisesta toiminnasta (mallinnuksen kautta, esimerkiksi keskimääräisessä käyttötilanteessa), esimerkiksi
  - virtausnopeus
  - paine

## 4.6 Aineistojen tulkinta

Vesijohtojen käyttöiän elinajanmallinnuksessa mallinnettava tapahtuma on saneerausikä. Putken käyttöikä tulkitaan siis loppuvan joko saneeraukseen tai käytöstä poistoon. Lähtöaineistoon oli merkitty putken saneeraustilasta kertova ominaistieto, esim. ”saneerattu sujuttamalla”, ”maasta kaivettu”, ”uudisrakennettu”, mutta koska sekä Alvan että Kuopion tapauksessa saneeraustiedoissa oli puutteita erityisesti saneerausvuosien osalta, eikä malli voi huomioida putkiosaa saneeratuksi ilman tietoa saneerausvuodesta, oli saneerauksen tulkintaa yksinkertaistettava. Malli tulkitsee putkiosan saneeratuksi, jos sille löytyy saneerausvuosi ja käytöstä poistetuksi, jos sille

löytyy käytöstäpoistovuosi riippumatta putken merkitystä tilasta. Putki voi siis olla merkitty saneeratuksi, mutta malli tulkitsee putken edelleen käytössä olevaksi, jos siltä puuttuu asennusvuosi.

Alvan ja Kuopion Veden aineistoissa sujutetuille putkille oli merkitty erikseen sisäputken uusi materiaali (usein PEL tai PEH) saneeratun vanhan materiaalin lisäksi, joten myös nämä putket haluttiin huomioida mallinnuksessa. Esimerkiksi valurautaputken sisään sujutettu PEH-putki tulkitaan mallissa uudisrakennetuksi. Ottamalla myös sujutetut putket huomioon mallinnuksessa maksimoitiin lähtötietojen antama hyöty. On kuitenkin mahdollista, että sujuttamalla asennettujen putkien tuleva käyttöikä poikkeaa muiden saman materiaalien putkien käyttöiästä. Sujuttamalla asennettuun PEH-putkeen on saattanut asennusvaiheessa tulla naarmuja tai muuta vauriota, joka lyhentää sen käyttöikää verrattuna auki-kaivuulla asennettuun PEH-putkeen. Lopputyössä käytössä oleva aineisto ei ole tarpeeksi kattava näiden mallintamiseen omana kategorianaan, mutta aineiston kartuttua mallinnus voidaan tulevaisuudessa tehdä näille erikseen.

Sekä Alvan että Kuopion lähtöaineistoihin tehtiin manuaalisesti pieniä korjauksia materiaali- ja/tai asennusvuositietoihin yhdessä käytyjen keskustelun perusteella. Alvan aineistossa oli esimerkiksi jonkin verran 1960-luvulla rakennettua putkea merkitty PEH/PEL materiaaliksi, vaikka materiaalia ei ole vielä ollut käytössä. Kyseessä oli todennäköisesti merkintävirhe, joten kyseisten putkien materiaali korjattiin harmaaksi valuraudaksi (VAL), joka on verkostossa yleisin tuohon aikaan käytössä ollut materiaali. Alkuperäisestä aineistosta on lisäksi jätetty pois muutamia selkeästi virheellisiä arvoja, jotka todennäköisesti johtuvat inhimillisistä virheistä. Esimerkiksi vesijohdot, joille laskettu tämänhetkinen käyttöikä oli selvästi epärealistinen, negatiivinen luku tai tuhansia vuosia, jätettiin mallinnuksesta pois. Aineistoissa oli mukana myös tontti- ja kiinteistöjohtoja, mutta mallinnuksen kohteeksi valittiin vain vedenjakelujohdot.

Pituuden huomioimiseen elinajanmallinnuksessa ei ole yksinkertaista tapaa, joten putket huomioidaan mallissa yksittäisinä putkiosina. Yksi vaihtoehto olisi ryhmitellä putket muutamaaan kategoriaan putkipituuden mukaan. Tämän lopputyön tilanteessa havaittiin, ettei aineistoa ollut tarpeeksi useamman putkiryhmän mallintamiseen, joten työssä priorisoitiin vesijohtojen mallintamista materiaalityypittain.

## 5 Alva

Alva-yhtiöt Oy vastaa Jyväskylän alueen energian ja veden tuotannosta ja jakelusta. Yhtiön omistaa Jyväskylän kaupunki ja vuoteen 2019 se tunnettiin nimellä Jyväskylän Energia. (Alva). Alvan omistuksessa on lähes 1200 km vesijohtoverkostoa (Alva-yhtiöt Oy 2022).

### 5.1 Lähtöaineiston yleiskatsaus

Alvan aineistossa on tiedot noin 100 000 putkiosalle. Aineistosta löytyy vesijohtojen tyypillisimpien ominaisuustietojen lisäksi tietoa putken lähiympäristöstä ja mahdollisesta läheisestä katualueesta. Lopputyön kannalta tärkeimmät omaisuustiedot ovat putken laji, raportointitila, asennusvuosi, saneerausvuosi, käytöstäpoistovuosi, pituus, halkaisija ja materiaali. Verkoston virtaus- ja painetietoja oli lisäksi kerätty noin puolelta vesijohdoista, mutta näitä ei voitu huomioida mallinnuksessa tietojen vähyyden takia.

#### 5.1.1 Saneeraustiedot

Lukuja verkoston yleiskatsauksesta on esitetty taulukossa 1. Aineistoon on merkitty putken tämänhetkinen käyttötila, joka voi olla esimerkiksi ”uudisrakennettu”, ”saneerattu” tai ”poistettu käytöstä”. Saneeraukselle ja käytöstä poistolle on useita eri kategorioita. Kuten taulukosta on nähtävissä, saneerautuksi merkittyjä on aineistossa yli 21 000, joista vain noin 2800:lle on kirjattu sekä saneeraus- että asennusvuosi. Saneerausikä on siis laskettavissa vain 13 % saneeratuista putkista. Suurin osa vesijohdoista, joille löytyi sekä saneeraus- että asennusvuosi, oli sujuttamalla saneerattuja. Käytöstä poistetuista sekä asennus- että käytöstäpoistovuosi löytyi 88 %:lle putkiosista. Yhteensä käyttöikä mallinnuksessa huomioitiin siten vain 35 % putkiosista, jotka olivat jo tulleet käyttöikänsä päähän, mikä on mallinnuksen kannalta merkittävä epävarmuustekijä.

Taulukko 1: Alvan aineiston yleiskatsaus

|   | Putkiosuuksien lkm. | %   |
|---|---------------------|-----|
| Yhteensä  | 107876              |     |
| Merkitty saneeratuiksi                            | 21091               | 100 |
| Saneerausikä laskettavissa                        | 2769                | 13  |
| Merkitty käytöstä poistetuksi                     | 8431                | 100 |
| Käytöstäpoistoikä laskettavissa                   | 7445                | 88  |
| Kaikki saneeratut ja käytöstä poistetut           | 29522               | 100 |
| Saneerausikä- tai käytöstäpoistoikä laskettavissa | 10214               | 35  |

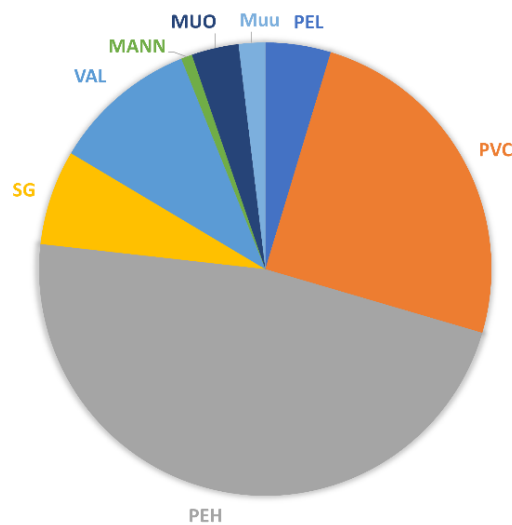
Saneeratut ja käytöstä poistetut on esitetty taulukossa 1 erillisinä kategorioina, sillä näiden käyttöiät luetaan malliin hieman eri tavalla, mutta käytännössä näillä kategorioilla ei ole mallinnustulosten kannalta merkitystä. Lisäksi sama putkiosa on voitu merkitä sekä saneeratuksi että käytöstä poistetuksi, mikä on huomioitu mallinnuksessa siten, että putkiosa luetaan malliin kaksi kertaa. Yhden putken käyttöikä lasketaan ensin asennuksesta saneeraukseen ja toisen saneeraushetkestä käytöstä poistoon. Tarkat kilometrimääräiset saneerauspiteudet saneerauskategorioittain on kuvattu taulukossa 2.

Taulukko 2: Alvan saneerausmäärät (Alva-yhtiöt Oy 2022)

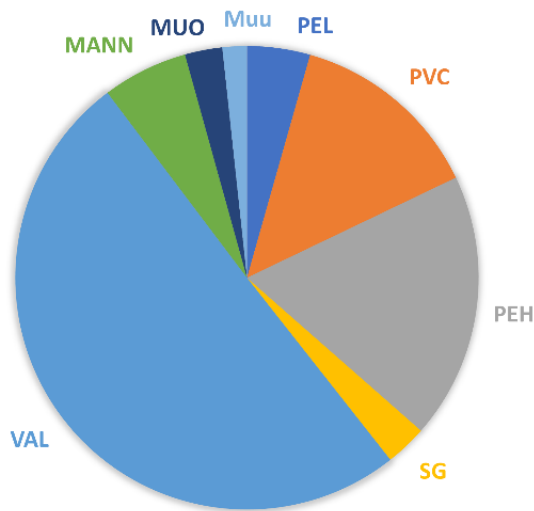
| Raportointitila                     | Määrät, km (tilanne 7.6.2022) |
|-------------------------------------|-------------------------------|
| Maasta kaivettu                     | 0,05                          |
| Saneerattu uusimalla eri sijaintiin | 53,41                         |
| Käytöstä poistettu, saneerattu      | 54,38                         |
| Saneerattu sujuttamalla             | 33,99                         |
| Käytöstä poistettu                  | 35,43                         |
| Maasta kaivettu, saneerattu         | 0,23                          |
| Saneerattu vaakaporaamalla          | 1,39                          |
| Saneerattu pinnoittamalla           | 2,15                          |
| Kaikki                              | 181                           |

### 5.1.2 Mallinnetut materiaalit

Lopputyössä arvioidaan vesijohtojen käyttöikää myös putkiryhmittäin materiaalin mukaan. Aineistossa oli eri materiaalien kategorioita 12 kpl, mutta näistä mallinnukseen valittiin 7 saneerausmäärältään suurinta ryhmää. Nämä materiaalit ovat harmaa valurauta (VAL), pallografiittivalurauta (SG), Mannesmann-teräsputket (MANN), muovilaadut PVC, PEH ja PEL sekä yleinen muovikategoria (MUO), joka voi sisältää mitä tahansa käytetyistä muoveista. MUO-materiaaliryhmän sisällä olevien muovien jakaumasta ei ole tarkempaa tietoa. Materiaaleihin viitataan myöhemmin pääosin lyhennyksin selkeyden vuoksi. Kuvassa 6 on esitetty koko vesijohtoverkoston materiaalijakauma ja kuvassa 7 saneerattujen ja käytöstä poistettujen vesijohtojen materiaalijakauma. Valittujen materiaalien osuus on kummassakin tapauksessa 98 %. Mallinnukseen valitut materiaalit kattavat siis huomattavan osan verkostosta.



Kuva 6. Alva: Koko vesijohtoverkoston materiaalijakauma



Kuva 7. Alva: saneerattujen ja käytöstä poistettujen vesijohtojen materiaali-jakauma

Elinajanmallinnus edellyttää riittävästi dataa, joten pienempien putkiryhmien mallinnukseen sisältyy väistämättä epävarmuutta. Kuvista 6 ja 7 on siten syytä huomioida yksittäisen materiaalin määrä. Yli puolet saneeratuista putkista on harmaata valurautaa, vaikka niiden osuus kaikista asennetuista on vain noin 10 %. Tämä johtunee siitä, että verkoston vanhimmat putket ovat tyypillisesti valurautaa ja se oli suosituin rakennusmateriaali vielä 60–70 –luvulla, jolloin verkostorakentaminen oli ahkerimmillaan. Suuri osa vanhimmista putkista on siten jo saneerattu.

Saneerattujen putkien osuus kaikista samaa materiaalia olevista putkista on vielä koostettu taulukkoon 3. Taulukosta näkyy, miten esimerkiksi Mannesmann-teräsputkista, joita käytettiin materiaalina erityisesti 1900-luvun alkupuoliskolla, on jo saneerattu suurin osa. Toisaalta PEH- ja SG-putkista on saneerattu vasta 4 %, mikä johtunee siitä, että niitä on asennettu enemmän vasta viime vuosikymmeninä. Uudempien putkimateriaalien kohdalla käyttöään mallinnukseen epävarmuutta aiheuttaa saneerausten suhteellisesti pieni määrä.

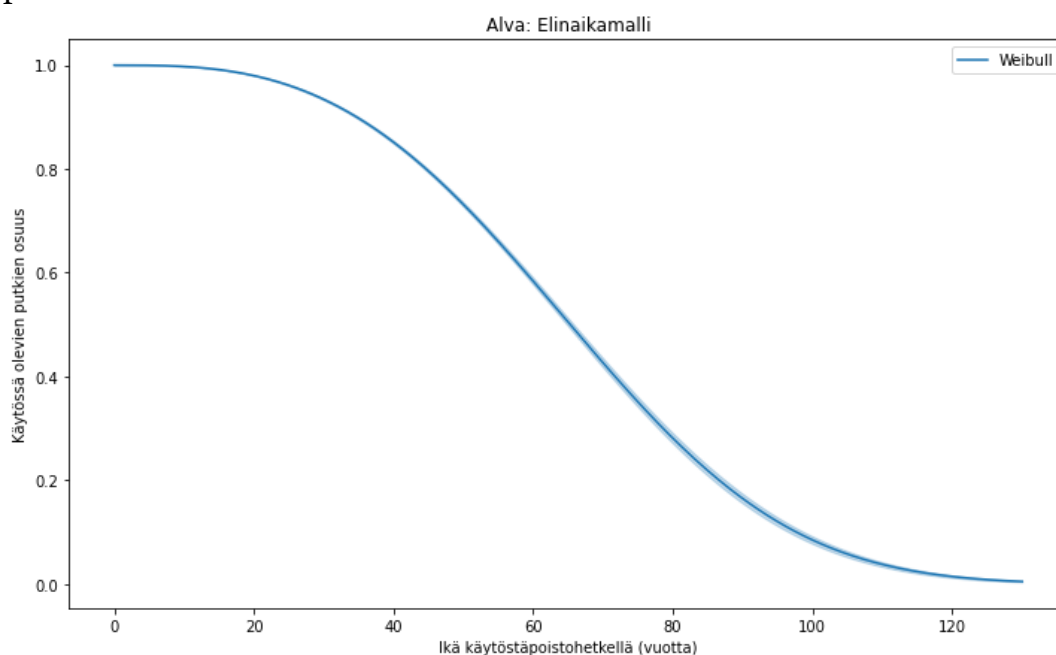
Taulukko 3. Alva: Saneerattujen osuus kaikista samaa materiaalia olevista vesijohdoista

| Materiaali | %  |
|------------|----|
| PEL        | 10 |
| PVC        | 6  |
| PEH        | 4  |
| SG         | 4  |
| VAL        | 51 |
| MANN       | 79 |
| MUO        | 8  |
| Kaikki     | 11 |

## 5.2 Mallinnustulokset

### 5.2.1 Weibull

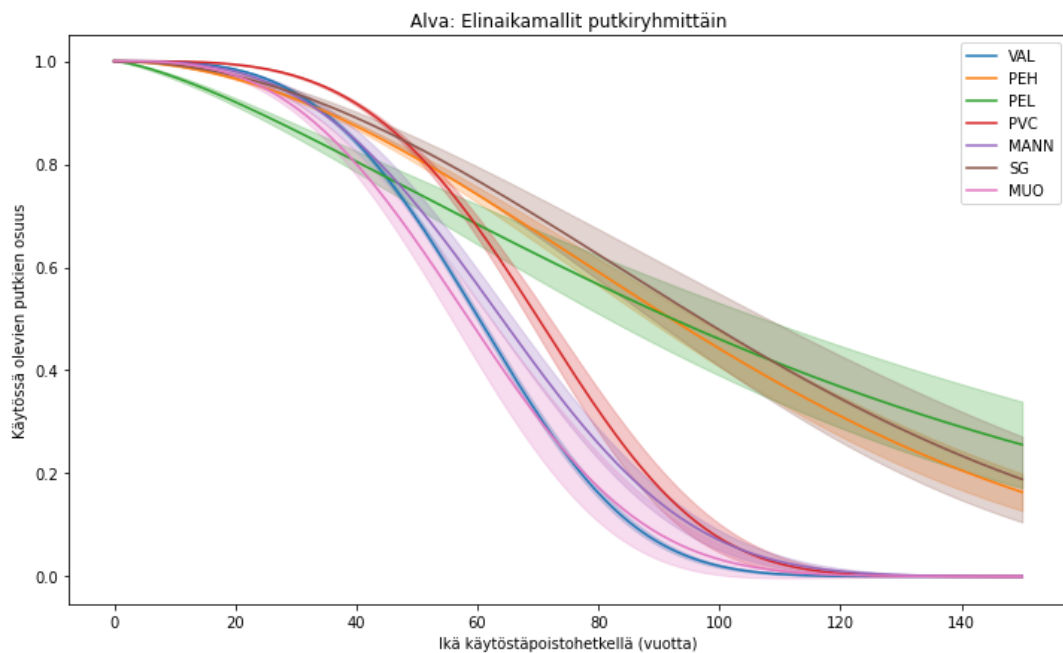
Kuvassa 6 on esitetty Alvan aineistoon sovitettu Weibull-käyrä, joka kuvaa verkoston saneerausvolyymin kehittymistä ajan funktiona ja ennustaa tätä tulevaisuuteen. Lukemalla käyrää nähdään, että 20 vuoden jälkeen lähes kaikki putket ovat vielä käytössä, mutta tämän jälkeen käytössä olevien vesijohtojen määrä laskee, kunnes viimeiset putket tulevat saneerattaviksi noin 100 vuoden tienoilla. Mallin perusteella verkoston eloonjäämisajan mediaani (mean survival time) on 62 vuotta. Malli siis ennustaa, että tämän ikäisenä 50 % verkostosta on tullut käyttöikänsä päähän eli on saneeraustarpeessa.



Kuva 6. Alva: Weibull-elinaikamalli

## 5.2.2 Weibull-ennuste putkiryhmittäin

Käyttöiänmallinnus tehtiin myös putkiryhmittäin. Kuvassa 7 näytetään materiaalikohtaiset Weibull-käyrät. Kuvaajassa on myös käyrien luottamusvälit, jotka on merkitty himmeämmällä sävyllä. Eri materiaalien käyrät visualisoivat hyvin, miten tapahtumien eli saneerausten määrä ja materiaalityypin mukaan kuuluvien putkien määrä vaikuttavat käyttöiänmallinnuksen luotettavuuteen.



Kuva 7. Alva: Weibull materiaalityypittain

Harmaan valuraudan (VAL) ennuste on luottamusvälien kapeuden perusteella luotettavin. Harmaan valuraudan eloonjäämisajan mediaani on mallin perusteella noin 60 vuotta (ks. Taulukko 4). VAL-putkia oli sekä asennettu että saneerattu eniten muihin materiaaleihin verrattuna. Koska suurin osa materiaalityypin putkista on verkoston vanhimmasta päästä, on niitä myös tullut saneerattavaksi suuri määrä, noin puolet kaikista asennetuista valurautaputkista. Toisaalta on mahdollista, että kaikkia aiempia saneerauksia ei ole tiedossa.

PVC, MANN ja MUO- materiaalityypien käyrät mukailevat harmaan valuraudan käyrää, mutta niillä on suuremmat luottamusvälit. Mainittujen materiaalityypien eloonjäämisajan mediaanit ovat 60–70- vuotta (ks. Taulukko 4). Syy suurempiin luottamusväleihin on, että mm. MUO- ja MANN-putkia on aineistossa suhteellisen alhainen määrä. Lisäksi PVC- ja MUO-putkilla saneerattujen osuus on vielä suhteellisen alhainen. On myös hyvä

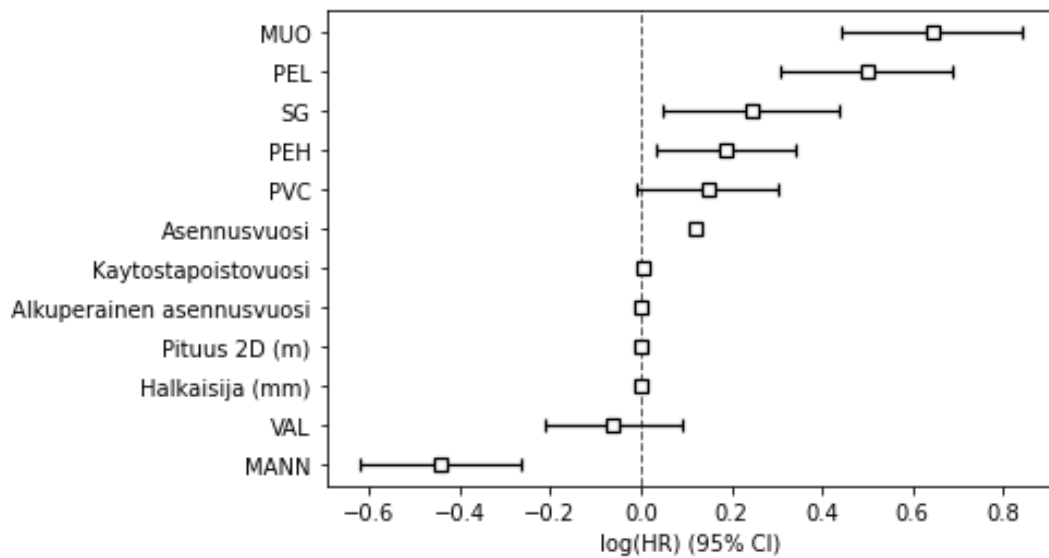
muistaa, että MUO-materiaaliryhmän putket voivat sisältää mitä tahansa käytettyä muovia, joten sen vertailukelpoisuus muihin materiaaleihin on kyseenalainen. Suurimmat elinaikamallien epävarmuusvälit olivat materiaaliryhmien SG, PEH ja PEL putkilla. Kaikkia kolmea materiaalia on asennettu pääosin vasta viime vuosikymmeninä, joten putkia on toistaiseksi saneerattu vain vähän, mikä hankaloittaa käyttöiän mallintamista luotettavasti. Näistä jokaisen materiaaliryhmän eloonjäämisajan mediaanin ennusteet ovat yli 90 vuotta (ks. taulukko 4).

Taulukko 4. Eri putkiryhmien eloonjäämisajan mediaanit

| Materiaali | Eloojäämisajan mediaani |
|------------|-------------------------|
| VAL        | 61                      |
| PEH        | 92                      |
| PEL        | 92                      |
| PVC        | 70                      |
| MANN       | 64                      |
| SG         | 97                      |
| MUO        | 60                      |

### 5.2.3 Coxin suhteellisen vaaran malli

Coxin suhteellisen vaaran mallin tavoitteena on analysoida eri selittäjien painoarvoa. Alvan lähtöaineistosta valittiin mallinnukseen seuraavat selittäjät: asennusvuosi, alkuperäinen asennusvuosi, käytöstäpoistovuosi, materiaali, pituus ja halkaisija. Selittäjille saadut kertoimet ja luottamusvälit on esitetty kuvassa 8. Materiaalit on esitetty omina luokkinaan, koska malliin voi syöttää vain numeerisia arvoja. Kuten kuvasta on nähtävissä, materiaaleille lasketut kertoimet poikkeavat eniten nolasta eli niillä on eniten vaikutusta käyttöiän ennusteeseen, mutta kertoimen suuruus vaihtelee materiaaleittain. Lisäksi asennusvuosi vaikuttaa käyttöiän ennusteeseen pidentävästi. On kuitenkin hyvä huomata, että kyseinen selittäjä sisältää sekä asennus- että saneerausvuositietoja vesilaitoksen merkintätavan takia. Sen sijaan halkaisijalla, pituudella, käytöstäpoistovuodella tai alkuperäisellä asennusvuodella ei ollut merkitystä käyttöiän ennusteen kannalta. Mallin tulokset tukevat lopputyössä tehtyä päätöstä tarkastella vesijohtojen käyttöiän jakautumista materiaaliryhmittäin. Koska halkaisijalla tai pituudella ei ole tulosten mukaan vaikutusta käyttöiän ennusteeseen, ei erillinen mallinnus niiden putkiryhmillä ole tarpeen.



Kuva 8. Alva: Selittäjien kertoimet luottamusväleineen

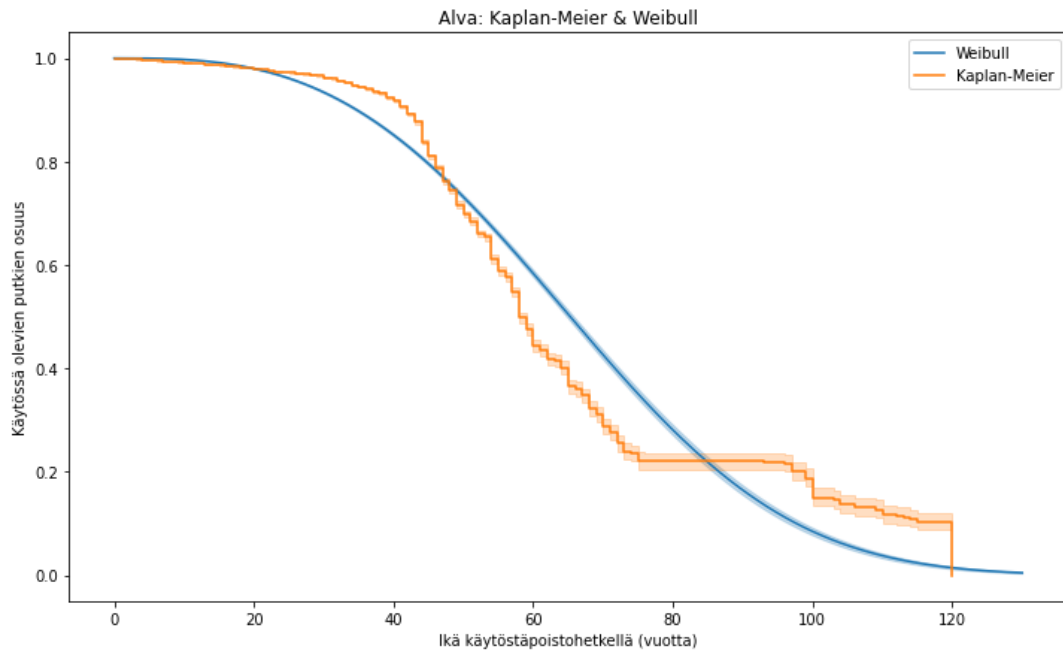
Taulukossa 5 esitetään mallinnetut selittäjät ja niiden lasketut kertoimet (coef), ekponentiaaliset kertoimet ( $\exp(\text{coef})$ ), luottamusvälien ala- (coef lower 95 %) ja ylärajat (coef upper 95 %) sekä kertoimien p-arvot. Taulukko tarjoaa mahdollisuuden arvioida eri kertoimien luotettavuutta tai tilastollista merkitsevyyttä, kuten matemaattisissa termeissä ilmaistaan. Tulokset ovat merkitseviä vain niille selittäjille, joiden p-arvo on suurempi kuin 0,05. Taulukosta nähdään, että harmaan valuraudan (VAL) ja PVC-muovin materiaalityyppien sekä putken pituuden tulokset eivät olleet tilastollisesti merkitseviä. Harmaan valuraudan saama tulos voi olla yllättävä, koska verkostossa on paljon kyseistä materiaalia ja noin puolet niistä on saneerattu. VAL-putkiryhmä on kuitenkin Alvan aineistossa se materiaalityyppi, jonka saneerattujen putkien tiedot puuttuvat eniten, mikä aiheuttaa ristiriitaisuutta kertoimien arvioinnissa. On kuitenkin tärkeää muistaa, että selittäjien merkitsevyys riippuu käytetystä aineistosta ja että toisella aineistolla tulokset voivat olla erilaisia.

Taulukko 5. Alva: Coxin suhteellisen vaaran mallin tuloste

|                           | coef  | exp(coef) | coef lower 95 % | coef upper 95 % | p     |
|---------------------------|-------|-----------|-----------------|-----------------|-------|
| Asennusvuosi              | 0.12  | 1.13      | 0.12            | 0.12            | <0.05 |
| Alkuperäinen asennusvuosi | 0     | 1         | 0               | 0               | <0.05 |
| Kaytostapoistovuosi       | 0     | 1         | 0               | 0               | <0.05 |
| Pituus 2D (m)             | 0     | 1         | 0               | 0               | 0.09  |
| Halkaisija (mm)           | 0     | 1         | 0               | 0               | <0.05 |
| VAL                       | -0.1  | 0.9       | -0.25           | 0.05            | 0.18  |
| PEH                       | 0.13  | 1.13      | -0.03           | 0.28            | <0.05 |
| PEL                       | 0.46  | 1.59      | 0.27            | 0.65            | <0.05 |
| PVC                       | 0.09  | 1.1       | -0.06           | 0.25            | 0.24  |
| MANN                      | -0.43 | 0.65      | -0.6            | -0.25           | <0.05 |
| SG                        | 0.19  | 1.21      | 0               | 0.39            | <0.05 |
| MUO                       | 0.58  | 1.79      | 0.38            | 0.78            | <0.05 |

#### 5.2.4 Hyvyyden arviointi

Weibull-jakauman hyvyyttä arvioitiin empiirisesti Kaplan-Meier-mallin avulla. Kaplan-Meier kuvaa mallinnukseen käytettyä aineistoa sellaisena kuin se on ja Weibull pyrki sovittamaan tähän aineistoon likimääräisen käyrän. Kummatkin käyrät on esitetty kuvassa 9. Kaplan-Meier käyrästä nähdään, että ensimmäiset putket ovat tulleet saneerattaviksi noin 20 vuoden tienoilla, noin 45 vuoden jälkeen käytössä olevien vesijohtojen määrä on lähtenyt jyrkempään laskuun ja viimeiset putket on saneerattu noin 120 vuoden ikäisinä. Noin vuosien 75–95 osalta käyrä pysyy täysin vakiona, mikä tarkoittaa, että Alvan lähtöaineistossa ei ole saneerausmerkintöjä mainitun ikäisille vesijohdoille. Alvan aineistoa tarkastelemalla selvisi, että Alvan aineistosta puuttuu saneeraustietoja suurelle määrälle vesijohtoja, jotka tunnetaan saneeratuiksi. Nämä puuttuvat tiedot näkyvät Kaplan-Meier käyrällä tasaisena alueena.



Kuva 9. Alva: Kaplan-Meier kertymäfunktio ja Weibull-elinaikamalli

Käyriä vertailemalla voidaan nähdä, miten puuttuvan saneeraustiedot vaikuttavat käyttöiän ennusteeseen. Jos käytössä olevien putkien määrä verkostossa olisi kehittynyt vuoden 75 jälkeen yhtä jyrkästi kuin vuotta 75 ennen olisi käytössä olevien putkien määrä esimerkiksi 100 vuoden tienoilla paljon vähäisempi. Puuttuvien saneeraustietojen takia Weibull-jakauma kuvaa käyttöiän jakautumista selkeästi loivemmin kuin mitä ensimmäisen 75 vuoden kehitys antaa olettaa. Puuttuvat saneeraustiedot johtavat siis tavallista optimistisempaan käyttöiän ennusteeseen. Weibull-jakauman puutteeksi voidaanankin mainita se, ettei se kykene ottamaan huomioon mahdollista puuttuvaa dataa.

Coxin suhteellisen vaaran mallia käytettiin eri selittäjien painoarvon määrittämiseen. Ennen mallinnusta aineisto jaettiin testi- ja opetusaineistoon, jotta mallin hyvyttä voitiin arvioida C-indeksin avulla. Sekä testi- että opetusaineiston C-indeksi pyörityi arvoon 0,96, mikä tarkoittaa, että malli yleistyi hyvin testiaineistolle eikä ollut ylisovitettu.

Weibull- ja Cox-mallien keskinäistä hyvyttä vertailtiin AIC- ja log-todennäköisyys arvojen avulla. Lasketut arvot on esitetty taulukossa 6. Kummankin perusteella Weibull-malli pärjäsikin vertailussa paremmin. Aineiston sovittaminen onnistui siis paremmin Weibull- jakauman kuin Cox-mallin avulla.

Taulukko 6. Alva: AIC ja log-todennäköisyys

|                | Weibull | Cox    |
|----------------|---------|--------|
| AIC            | 114067  | 116807 |
| log-likelihood | -57031  | -58391 |

## 6 Kuopion Vesi

Kuopion Veden toiminta-alueella on yli 100 000 vedenkäyttäjää ja se kattaa Kuopion keskustan lisäksi lukuisia taajama-alueita sekä Kuopiossa että Siilinjärvellä. Tämän lisäksi Kuopion Vesi toimittaa vettä useille alueiden vesiosuuskunnille. (Kuopion Vesi 2023). Vesihuoltoverkostoja on yhtiöllä yhteensä 1186 km, josta 871 km on vesijohtoverkosta. Vedenkulutus on kaiken kaikkiaan noin 17 000 kuutiota vuorokaudessa. (Pisaroina 2023).

### 6.1 Lähtöaineiston yleiskatsaus

Kuopion Veden aineistossa on tiedot noin 40 000 putkiosalle. Lähtötietoihin on merkitty asennus- ja saneerausvuosien lisäksi muutamia muita ominaisuuksitietoja kuten putken halkaisija ja materiaali sekä putken käyttötila, joka kertoo mahdollisesta saneeraustavasta.

#### 6.1.1 Saneeraustiedot

Tilastoja Kuopion Veden lähtöaineistosta on esitetty taulukossa 7. Saneeratuiksi merkityistä putkista käyttöikä on laskettavissa vain 20 %:lle putkista. Käytöstä poistetuiksi merkityistä käyttöikä on laskettavissa 67 %:lle. Kun yhdistettiin saneeratuiksi ja käytöstä poistetuiksi merkittyjen tiedot, saatiin mallinnettavia tapahtumia yhteensä 1107, mikä on noin 60 % kaikista verkoston käyttöikänsä päähän tulleista vesijohdoista ja vain 3 % kaikista mallinnukseen sisällytetyistä vesijohdoista. Vuosilukujen puuttuminen ja kirjatujen saneerausten suhteellisen pieni määrä aineistossa ovat merkittäviä epävarmuustekijöitä luotujen mallien luotettavuuden kannalta.

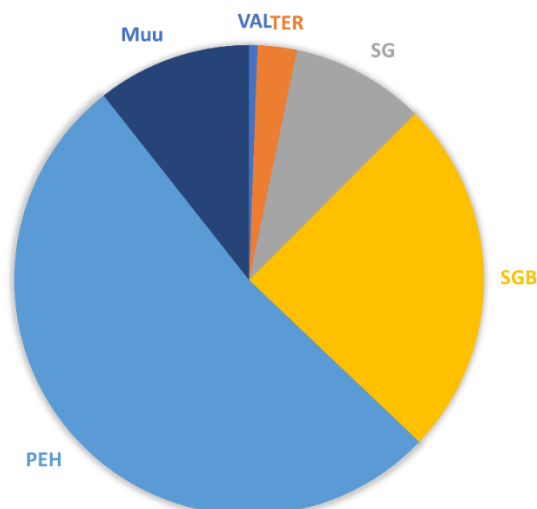
Taulukko 7. Kuopion Veden lähtöaineiston yleiskatsaus

|   | Putkiosien lkm. | %   |
|---|-----------------|-----|
| Yhteensä                                | 37719           |     |
| Merkitty saneeratuksi                   | 760             | 100 |
| Saneerausikä laskettavissa              | 147             | 20  |
| Merkitty käytöstä poistetuksi           | 1107            | 100 |
| Käytöstäpoistoikä laskettavissa         | 1002            | 67  |
| Kaikki saneeratut ja käytöstä poistetut | 1867            | 100 |
| Tapahtumat mallinnuksessa mukana        | 1107            | 60  |

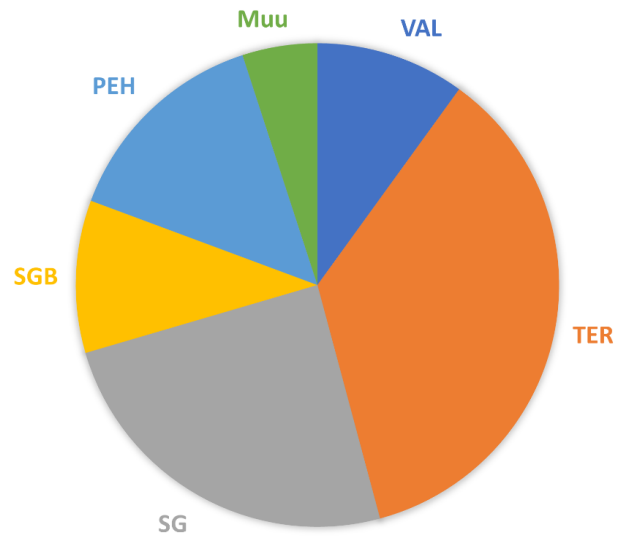
Kuten Alvankin tapauksessa, Kuopion aineistossa saneeraustapahtuman vuosi on tallennettu rakennusvuoden paikalle ja tiedot alkuperäisesti rakennusvuodesta on siirretty uuteen sarakkeeseen, mikä luo aineiston käsittelyyn monitulkintaisuutta. Jos putki on merkitty saneeratuksi, mutta sen tiedoista löytyy vain yksi vuosiluku, ei voida olla varmoja onko kyseessä asennus- vai saneerausvuosi.

### 6.1.2 Mallinnetut materiaalit

Lopputyössä mallinnus tehdään myös materiaaliryhmittäin. Verkostossa on 13 erilaista materiaalimerkintää, mutta näistä 5 saneerausmäärältään suurinta valittiin mallinnettavaksi. Valitut materiaalit ovat valurauta (VAL), teräs (TER), PEH-muovi, pallografiittivalurauta (SG) sekä betonoitu pallografiittivalurauta (SGB). Materiaaleihin viitataan myöhemmin pääosin lyhennyksin selkeyden vuoksi. Valittujen materiaalien osuus kaikista asennetuista vesijohdoista on esitetty kuvassa 12 ja osuus kaikista saneeratuista vesijohdoista kuvassa 13. Mallinnettavien materiaalien vesijohdot kattavat 89 % kaikista verkoston asennetuista vesijohdoista ja 95 % kaikista saneeratuista vesijohdoista. Materiaalikohtaisten mallien kohdalla aineiston määrän lisäksi kannattaa myös huomioida saneerattujen putkien osuus kaikista materiaalin asennetuista vesijohdoista, mikä on esitetty taulukossa 8. Vasta viime vuosina putkissa käytettyjen materiaalien kohdalla mallinnukseen kohdistuu enemmän epävarmuutta, koska materiaaliryhmän vesijohtoja on saneerattu vasta vähän.



Kuva 10. Kuopio: Koko vesijohtoverkoston materiaali-jakauma



Kuva 11. Kuopio: saneerattujen ja käytöstä poistettujen vesijohtojen materiaalityyppien jakauma

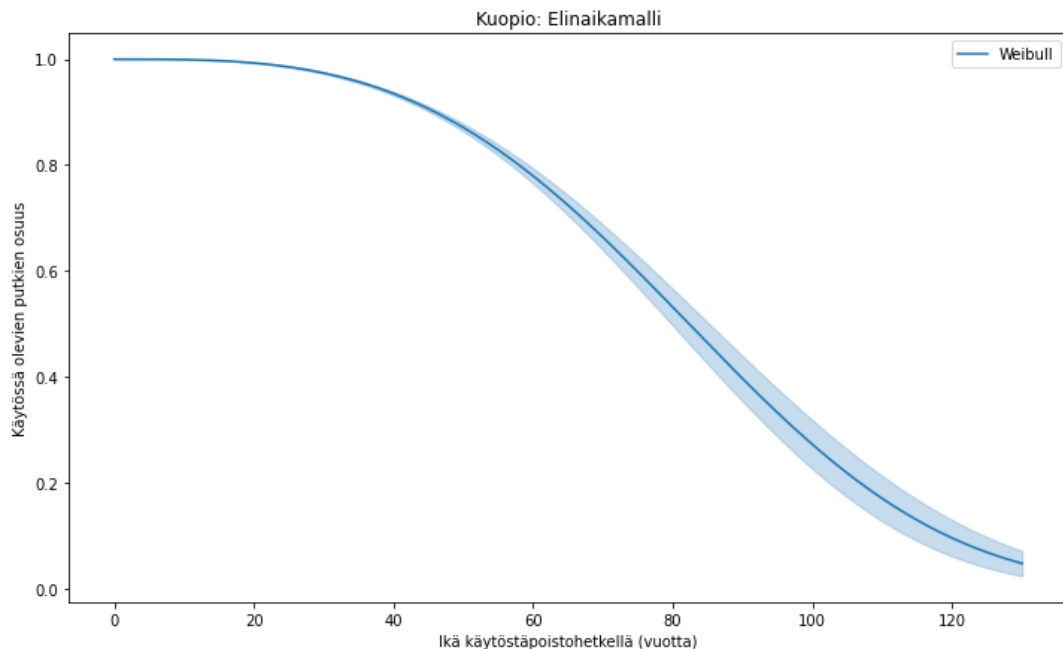
Taulukko 8. Kuopio: Saneerattujen osuus kaikista samaa materiaalia olevia vesijohdoista

| Materiaali | %    |
|------------|------|
| VAL        | 79   |
| TER        | 61   |
| SG         | 12   |
| SGB        | 2    |
| PEH        | 1    |
| Kaikki     | 0.05 |

## 6.2 Mallinnustulokset

### 6.2.1 Weibull

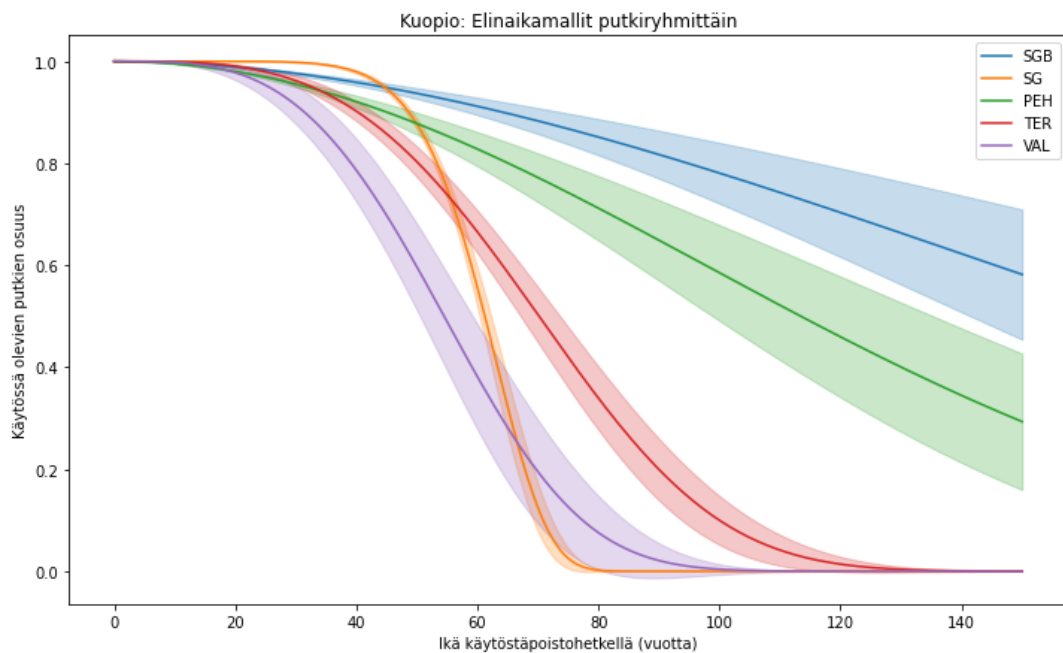
Kuvassa 12 näkyy aineistoon sovitettu Weibull-käyrä, joka kuvaa verkoston saneerausvolyymien kehittymistä ajan funktiona. Lukemalla käyrää nähdään, että 20 vuoden jälkeen lähes kaikki putket ovat vielä käytössä, jonka jälkeen käytössä olevien vesijohtojen määrä laskee, kunnes viimeiset tulevat saneerattaviksi noin 120 vuoden tienoilla. Mallin perusteella verkoston eloonjäämisajan mediaani (median survival time) on 82 vuotta, mikä on noin 20 vuotta enemmän kuin Alvalle saatu ennuste. Tämän ikäisenä arviolta 50 % verkostosta on tullut käyttöikänsä päähän eli on saneeraustarpeessa.



Kuva 12. Kuopio: Weibull-elinaikamalli

## 6.2.2 Weibull-mallit putkiryhmittäin

Vesijohtojen käyttöikää mallinnettiin myös putkiryhmittäin. Mallinnettaviksi materiaaleiksi valittiin eniten saneeratut materiaalityypit verkostossa. Materiaalikohtaiset käyrät luottamusväleineen on esitetty kuvassa 13. Vertailemalla eri materiaalityyppien käyriä voidaan nähdä, miten mm. mallinnettavien tapahtumien määrä sekä materiaalityypin vesijohtojen kokonaismäärä vaikuttavat ennusteen luotettavuuteen.



Kuva 13. Kuopio: Weibull putkiryhmittäin

Mallinnetuista käyristä tavallisella pallografiittivaluraudalla (SG) on pienimmät luottamusvälit. Mallin perusteella materiaalin eloonjäämisajan mediaani on 61 vuotta (ks. Taulukko 9). SG-putkista on saneerattu noin 12 %.

Teräs-(TER) ja valurautaputkista (VAL) on saneerattu huomattavasti suurempi osuus (60–80 %) kuin SG-putkista, mutta käyrien luottamusvälit ovat tästä huolimatta suuremmat. Suuremmat luottamusvälit johtunevat materiaalien pienemmästä kokonaismäärästä. Teräksen eloonjäämisajan mediaani on mallin perusteella 70 vuotta ja valuraudan 55 vuotta (ks. Taulukko 9).

PEH- ja SGB-materiaaleja on koko asennetussa verkostossa eniten (ks. Kuva 10), jopa yli 75 % kaikista asennetuista vesijohdoista, mutta materiaaleja on asennettu enemmän vasta viime vuosikymmeninä ja niiden vesijohdoista on tähän mennessä saneerattu vain muutama prosentti. Näiden materiaalityyppien aineistot olivat siten vahvasti oikealta sensuroituja, mikä luo

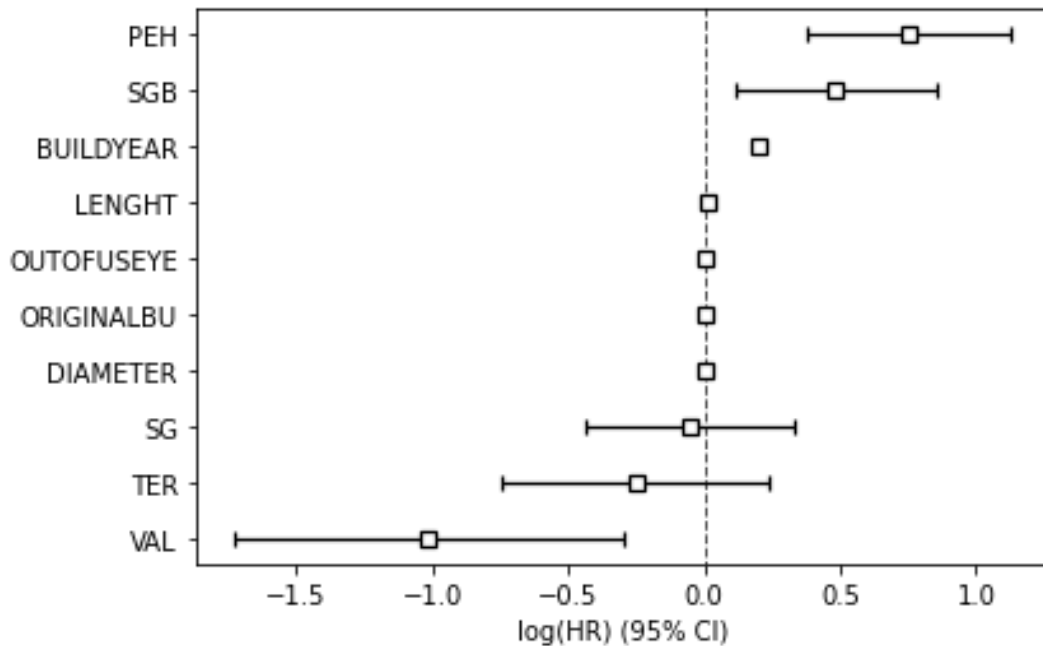
tyypillisesti todellisuutta optimistisempia ennusteita elinajanmallinnuksessa. On siten ymmärrettävää, että kyseisten materiaaliryhmien käyrien epävarmuusvälit ovat käyristä suurimmat. PEH:in eloonjäämisajan mediaaniksi saatiin 115 vuotta ja SGB:n huimat 170 vuotta (ks. Taulukko 9).

Taulukko 9. Kuopio: Eri putkiryhmiä eloonjäämisaikojen mediaanit

| Materiaali | Eloonjäämisajan mediaani |
|------------|--------------------------|
| SGB        | 170                      |
| SG         | 61                       |
| PEH        | 113                      |
| TER        | 70                       |
| VAL        | 55                       |

### 6.2.3 Coxin suhteellisen vaaran malli

Eri selittäjien painoarvoa arvioitiin Coxin suhteellisen vaaran mallin avulla. Kuopion Veden lähtöaineiston selittäjistä valittiin asennusvuosi (BUILDDYEAR), alkuperäinen asennusvuosi (ORIGINALBU), käytöstäpoistovuosi (OUTOFUSEYEAR), pituus (LENGHT), halkaisija (DIAMETER) ja materiaali, joka on huomioitu erillisinä selittäjinä jokaiselle mallinnetulle materiaalille. Cox-mallilla saadut kertoimet ja niiden vaihteluvälit eri selittäjille on esitetty kuvassa 14. Materiaaliryhmistä valuraudan (VAL), PEH-muovin, ja betonoidun pallografiittivaluraudan (SGB) kertoimet ja niiden vaihteluvälit poikkeavat nolasta, mikä tarkoittaa, että kyseisellä ominaisuustiedolla huomattiin olevan käyttöikä pidentävä (positiivinen kerroin) tai lyhentävä (negatiivinen kerroin) vaikutus. SG:lle ja TER:lle arvioitujen kertoimien vaihteluvälit jatkuvat nolasta saakka, joten niille ei mallintamalla löydetty käyttöikä pidentävää tai lyhentävää yhteyttä. Halkaisijalla, pituudella, käytöstäpoistovuodella tai alkuperäisellä asennusvuodella ei myöskään ollut vaikutusta käyttöikä ennusteeseen. Näiden ominaisuustietojen ja saneerausien välillä ei siis löytynyt matemaattista riippuvuutta näiden selittäjien osalta.



Kuva 14. Kuopio: Selittäjien kertoimien luottamusvälit

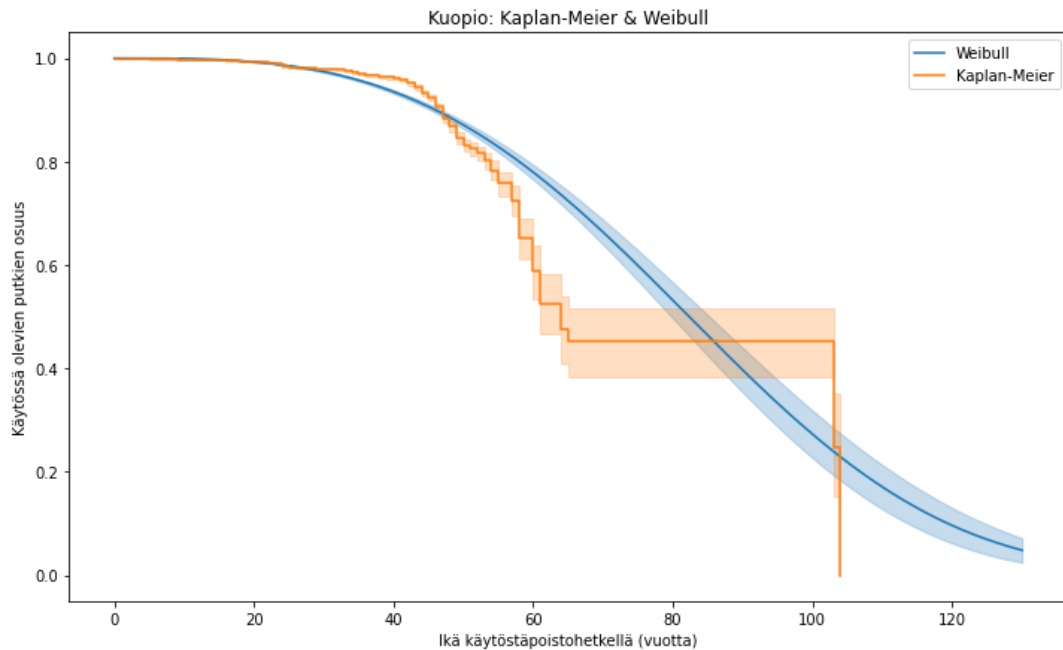
Taulukossa 10 on esitetty mallinnetut selittäjät ja niiden lasketut kertoimet (coef), eksponentiaaliset kertoimet ( $\exp(\text{coef})$ ), luottamusvälien ala- (coef lower 95 %) ja ylärajat (coef upper 95 %) sekä kertoimien p-arvot. Taulukon avulla voidaan tarkastella eri laskettujen kertoimien luotettavuutta tai tilastollista merkitsevyyttä, kuten matemaattisin termein sanotaan. Tulos on merkitsevä vain niille selittäjille, joille p-arvo on suurempi kuin 0,05. Taulukosta nähdään, että materiaaliyryhmistä teräksen (TER) ja pallografiittivaluraudan (SG) sekä halkaisijan ja pituuden kertoimet eivät ole tilastollisesti merkitseviä. Vaikka Kuopion aineistolla matemaattista riippuvuutta ei löytenyt, on mahdollista, että selittäjien kertoimet saavat tilastollisesti merkitsevät tulokset toisen vesilaitoksen aineistolla. Luotettavien tulosten saamiseksi onkin tärkeää tehdä vastaavaa mallinnusta myös kattavammalla aineistolla.

Taulukko 10. Kuopio: Coxin suhteellisen vaaran malli

|            | coef  | exp(coef) | coef lower 95 % | coef upper 95 % | p     |
|------------|-------|-----------|-----------------|-----------------|-------|
| BUILDEAR   | 0.19  | 1.21      | 0.18            | 0.2             | <0.05 |
| ORIGINALBU | 0     | 1         | 0               | 0               | <0.05 |
| OUTOFUSEYE | 0     | 1         | 0               | 0               | <0.05 |
| LENGHT     | 0     | 1         | -0.01           | 0.01            | 0.74  |
| DIAMETER   | 0     | 1         | 0               | 0               | 0.94  |
| SGB        | 0.55  | 1.73      | 0.2             | 0.9             | <0.05 |
| SG         | -0.04 | 0.96      | -0.39           | 0.31            | 0.83  |
| PEH        | 0.79  | 2.2       | 0.44            | 1.14            | <0.05 |
| TER        | -0.09 | 0.92      | -0.51           | 0.33            | 0.68  |
| VAL        | -0.45 | 0.63      | -1.07           | 0.16            | <0.05 |

#### 6.2.4 Hyvyyden arviointi

Kuopion Weibull-mallia vertailtiin empiirisesti Kaplan-Meier-kertymäfunktion, joka kuvaa käytössä olevien putkien todellista määrää tietyllä ajanhetkellä. Niiden kuvaajat on esitetty kuvassa 15. Kaplan-Meier-käyrästä nähdään, että Kuopion vesijohtoverkostossa ensimmäiset putket ovat tulleet saneerausikään noin 30 vuoden ikäisinä ja käytössä olevien putkien määrä laskee tasaisesti noin vuoteen 65 asti. 65–100-vuotiaita vesijohtoja ei ole käyrän perusteella juurikaan saneerattu, joten käyrä pysyy näiden vuosien osalta vakiona. Syy tähän on puutteelliset saneeraustiedot. Kuopion tapauksessa saneerausikä oli arvioitavissa vain noin 60 % vesijohdoista, jotka tunnettiin saneeratuiksi. Tämän lisäksi saneeraustiedot puuttuvat niiden vanhimpien putkien osalta kokonaan, jotka on saneerattu ennen kuin saneeraustietoja on alettu keräämään. Aineistossa on siis paljon vesijohtoja, joiden malli tulkitsee olevan käytössä, vaikka todellisuudessa ne ovat saneerattuja. Luonnollisesti tämä tekee aineistoon sovitetun Weibull-mallin ennusteesta optimistisemmän. Jos käytössä olevien putkien määrä kehittyisi vuoden 65 jälkeen samoin kuin ennen vuotta 65 olisi Weibull-mallin käyrä selkeästi jyrkempi ja sen antama käyttöiän ennuste siten lyhyempi.



Kuva 15. Kuopio: Kaplan-Meier kertymäfunktio ja Weibull-elinaikamalli

Puuttuvat saneeraustiedot selittävät myös miksi Alvan ja Kuopion käyttöiän mediaanit poikkeavat toisistaan jopa parikymmentä vuotta. Alvalla saneeratut vesijohtoja mukana mallinnuksessa oli 10 % kaikista vesijohdoista, kun Kuopiolla vastaava luku oli vain 3 %, mikä viittaa Kuopiolta puuttuvan kokonaan osa saneeraustiedoista, koska verkostot ovat keskimäärin hyvin samanikäisiä. Vertailemalla Alvan ja Kuopion Kaplan-Meier käyriä nähdään, että itse asiassa käyrät ovat hyvin lähellä toisiaan ensimmäiset 60 vuotta. Saneerausten määrä verkostossa on siis vuoteen 60 asti kehittynyt hyvin samalla tavalla. Tämän jälkeen Kuopion käyrä pysyi vakiona yli 100 vuoden kohdalle, mistä voidaan päätellä, että Kuopion mallissa on hyvin paljon vanhoja vesijohtoja käytössä, joista suuri osa on todennäköisesti todellisuudessa saneerattuja. Myös Alvan mallissa Kaplan-Meier pysyi vakiona noin vuodesta 75 vuoteen 95, mikä viittaa puuttuvaan aineistoon. Alvan osalta puuttuvaa aineistoa oli kuitenkin vähemmän, minkä takia Kuopion ennusteen voidaan nähdä olevan vielä optimistisempi.

Coxin suhteellisen vaaran mallin hyvyttä arvioitiin C-indeksin avulla. Sekä testi- että opetusaineistolle laskettu C-indeksi sai arvon 0,99. Opetusaineistolla sovitettu malli yleistyi siis hyvin myös testiaineistoon eikä ylisovittamista tapahtunut.

Kuopion Weibull- ja Cox-mallien hyvyttä vertailtiin AIC- ja log-todennäköisyys arvojen avulla, jotka on esitetty taulukossa 11. Laskettujen arvojen perusteella Cox-malli tuotti luotettavampia tuloksia. Cox-mallin avulla sovitettu malli vastasi siis paremmin alkuperäistä aineistoa kuin Weibull-malli.

Taulukko 11. Kuopio: AIC ja log-todennäköisyys

|                | Weibull | Cox   |
|----------------|---------|-------|
| AIC            | 15247   | 8670  |
| log-likelihood | -7622   | -4325 |

## 7 Hangon Vesi

Hangon Vesi on kunnallinen vesilaitos, jolla on vesijohtoverkosta yli 200 km. Hangon hallinnoimaan vesijohtoverkoston kuuluu sekä Hangon keskusta että Lappohja sekä varayhteys Tammisaareen. (Hangon Vesi, 2022). Hangon keräämät lähtöaineistot eivät riitä lopputyössä tehtävään käyttöiän mallinnukseen, joten Hangon osalta lopputyöhön on sisällytetty muuta aineiston pohdintaa.

### 7.1 Lähtöaineiston yleiskatsaus

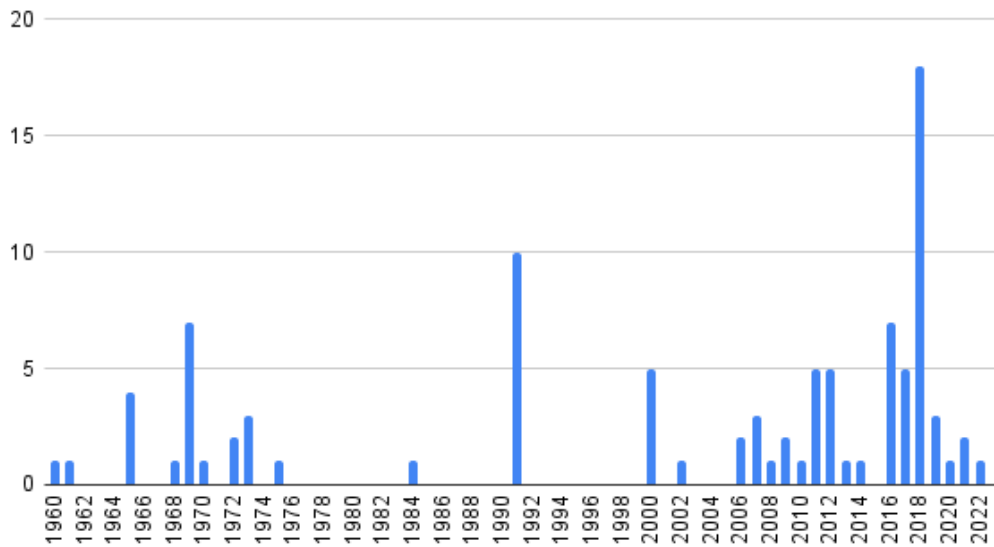
Tilastoja Hangon lähtöaineistosta on esitetty taulukossa 12. Aineistosta löytyy tiedot yli 14 000 vesijohdolle, mutta vain noin puolelle on merkitty rakennusvuosi. Tämän lisäksi saneerausvuodet on tallennettu virheellisesti alkuperäisen rakennusvuoden päälle, jolloin tieto alkuperäisestä asennusajankohdasta on kadonnut. Hangon tapauksessa puutteellinen saneeraustietojen tallennus on este käyttöiän mallinnukselle.

Taulukko 12. Hangon lähtöaineiston yleiskatsaus

|  | Putkiosat lkm | %   |
|--|---------------|-----|
| Yhteensä                                 | 14071         | 100 |
| Asennusvuodelliset                       | 7457          | 53  |
| *Merkitty rakennetuiksi 2000-2022        | 5442          | 39  |
| *Hangon Veden tai kaupungin omistuksessa | 4372          | 31  |

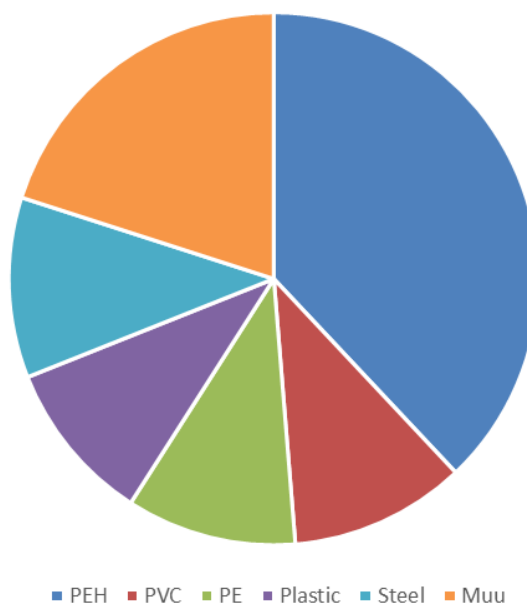
Haastattelujen perusteella selvisi, että Hangon Vesi on jo saneerannut suuren osan vesijohtoverkostostaan, ja että suurin osa vuodesta 2000 alkaen rakennetuiksi merkityistä on todellisuudessa saneerattuja vesijohtoja. Rakennusvuodet niiden vesijohtojen osalta, joilta tieto löytyy, on esitetty kuvassa 16. Aineistossa 73 % rakennusvuodellisista putkista on merkitty rakennetuiksi vuosina 2000–2022. Hangon Vesi on siis jo saneerannut suuren osan vesijohtoverkostostaan, mutta aineiston perusteella ei ole mahdollista luotettavasti arvioida, mitkä verkoston vesijohdoista on saneerattu tai minkä ikäisiä putket ovat silloin olleet.

### Hangon Vesi: vesijohtojen merkityt asennusvuodet



Kuva 16. Vesijohtojen asennusvuodet

Suurelle osalle Hangon Veden vesijohdoista on merkitty tieto materiaalista. Eri materiaalien osuus verkostosta on esitetty kuvassa 17. Tunnetut materiaalit ovat PEH, PVC, PE, muovi (plastic) ja teräs (steel). Vesijohtojen materiaalit ovat pääosin vasta viime vuosikymmeninä käytettyjä materiaaleja, koska vanhempien vesijohtojen sekä materiaali- ja asennusvuositiedot on poistettu aineistosta.



Kuva 17. Hanko: Vesijohtoverkoston materiaali-jakauma

Koska Hangon Veden aineistosta puuttuu tieto saneerausvuodesta, ei sitä voitu hyödyntää vesijohtojen käyttöiän mallinnuksessa. Hangon Veden aineistoa yritettiin hyödyntää lopputyössä myös muilla menetelmillä. Saneerattuja ja uudisrakennettuja vesijohtoverkoston alueita pyrittiin erottamaan paikkatiedon ja haastatteluissa saadun tiedon perusteella. Lisäksi vuototietojen yhdistämistä yritettiin putken tunnusluvun avulla. Kumpikaan menetelmä ei kuitenkaan mahdollistanut puuttuvien saneeraustietojen päättelyä.

## 7.2 Aineiston tarkastelun tulokset

Vesijohtojen käyttöiän mallinnus edellyttää, että vesijohtojen ikä saneeraus-hetkellä on laskettavissa aineiston perusteella. Vähimmäisvaatimus on siis putkiosien asennus- ja saneerausvuodet kattavalle osalle verkostoa. Hangon Veden lähtöaineistot eivät siten sovellu lopputyössä tehtävään käyttöiän mallinnukseen. Varsinaisten mallinnustulosten uupumisesta huolimatta Hangon Vesi toimii oivallisena esimerkkinä siitä, miten moni muukin suomalainen vesilaitos tällä hetkellä kerää verkostonsa kuntotietoja. Pohdintaa verkostotietojen keräämisen sopivista käytänteistä on esitetty lopputyön johtopäätöksissä.

Vesijohtojen käyttöiän mallinnuksella on potentiaalia toimia niidenkin vesilaitosten tukena, joilla itsellään ei ole vielä aineistoa kattavasti. Mallien avulla saatuja ennusteita voidaan tapauskohtaisesti hyödyntää myös toisen vesilaitoksen aineiston analysointiin. Analysoinnin luotettavuutta tukee, jos mallinnus on tehty kattavalla aineistolla ja vertailu voidaan tehdä samankaltaisten putkiryhmiä avulla. Koska Alvan ja Kuopion Veden mallinnetuissa materiaaleissa on osin samoja materiaaleja kuin Hangon Vedellä, voi mallintamalla saatuja ennusteita mahdollisesti hyödyntää myös Hangon vesijohtoverkoston käyttöiän arviointiin. Taulukko 13 esittää Alvan ja Kuopion Veden saamia eloonjäämisajan mediaaneja Hangon vesijohtoverkoston materiaaleille. Hangon Veden verkoston materiaalit ovat vesijohdoissa suhteellisen uusia, joten mallinnettujen materiaalien ennusteisiin liittyy luonnollisesti jonkun verran epävarmuutta. Esimerkiksi PEH-muovin käyrissä oli sekä Alvan että Kuopion Veden tapauksessa suhteellisen suuret luottamusvälit. On myös hyvä muistaa, että sekä Alvan että Hangon Veden tapauksessa muovimateriaaliryhmän tarkemmista materiaaleista tai niiden jakaumasta ei ole tietoa, joten samankaltaisesta nimestä huolimatta ennuste ei ole suoraan yleistettävissä toisen laitoksen materiaaliryhmään.

Taulukko 13. Eloojäämisajojen mediaanit Hangon Veden materiaaleille Alvan ja Kuopion Veden mallien perusteella

| Materiaali | Alva | Kuopion Vesi |
|------------|------|--------------|
| PEH        | 92   | 113          |
| PVC        | 70   | -            |
| Muovi      | 60   | -            |
| Teräs      | 64   | 70           |

## 8 Epävarmuustekijät

Lopputyössä merkittävin epävarmuustekijä on puutteelliset saneeraustiedot. Tyypillinen haaste aineiston käsittelyssä oli, että saneeratuiksi merkityiltä putkilta puuttui tieto saneerausvuodesta. Alvalla saneeratuiksi tiedetyistä putkista saneerausikä oli laskettavissa vain noin 35 %:lle ja Kuopion Vedellä puolestaan 60 %:lle putkista. Lopputyössä mallinnuksen epävarmuutta lisäsi saneeraustiedoiltaan puutteellisten putkiosien suuri määrä, mikä näkyi myös Kaplan-Meier-kertymäfunktioiden muodossa. Puuttuvien saneerausvuosien arviointi jälkikäteen ei ole mahdollista ilman verkostohistorian syvää tuntemusta, joten lopputyössä mallinnus jouduttiin tekemään puutteellisella aineistolla. Aineisto oli siis luettava malliin joko jättämällä kyseiset virheellisesti merkityt putkiosat pois tai huomioimana nämä mallissa edelleen käytössä olevina putkiosina, huolimatta siitä, olivatko ne todellisuudessa saneerattuja. Näiden vaihtoehtojen välillä ei ollut mallinnustulosten kannalta merkittävää eroa, mutta tuloksia tarkastellessa on syytä huomioida saneeraustietojen puutteellisuuden vaikutus. Käytännössä puuttuva saneerausvuosi on sama asia kuin puuttuva data, joka tekee käyttöään ennusteesta optimistisemmän.

Saneerausvuosien puuttuminen saneeratuiksi tunnistetuilta vesijohdoilta on eri asia kuin katkaistu data. Katkaistussa aineistossa osa saneeraustiedoista puuttuu, koska saneeraus on tapahtunut ennen kuin tietoja tapahtumasta on alettu keräämään. Tämä on todennäköistä erityisesti vanhimpien vesijohtojen kohdalla. Verkoston vanhimpien materiaalien käyriin kohdistuu tästä syystä epävarmuutta. Erityisesti Kuopion aineistossa oli aineistojen vertailun perusteella paljon katkaistua dataa. Katkaisun vaikutus käyttöäänmallinnuksessa näkyy tyypillisesti todellisuutta korkeampana saneerausikäen ennusteena. Tästä syystä Kuopion koko verkoston käyttöään ennuste oli Alvan ennustetta jopa parikymmentä vuotta optimistisempi.

Kun osaa vesijohdoista ei ole tarkasteluhetkellä vielä saneerattu, on aineisto oikealta sensuroitu. Weibull-malli soveltuu hyvin oikealta sensuroidun aineiston mallintamiseen, mutta mitä enemmän mallinnettavassa aineistossa on sensuroitu oikealta, sitä vähemmän aineistossa on mallinnettavia tapahtumia. Mallinnettavien tapahtumien vähäinen määrä aiheuttaa mallinnustulokseen epävarmuutta. Tämä koskee erityisesti pääosin nuoria vesijohtoverkostoja sekä niiden materiaalien malleja, joita on käytetty rakennusmateriaalina enemmän vasta viime vuosikymmeninä.

Koska lopputyön tavoitteena on tehdä käyttöään mallinnusta olemassa olevien saneeraustietojen perusteella, on syytä muista, että saneerauksia tehdään myös muuten kuin huonon kunnan takia. Vesijohdot pyritään usein saneeraamaan kadun rakennuksen tai saneerauksen yhteydessä, kun pääsy

maan alle on helpoin, jotta hyväkuntoiseen tiehen ei tarvitse kajota saman tien uudelleen. Osa verkostosta on myös saatettu saneerata huonokuntoisten materiaalien tai asennusvikojen takia aikaisemmin. Jos vesijohtoja saneerataan kunnan puolesta ennenaikaisesti, vaikuttaa saneeraustietojen käyttö mallinnuksessa myös käyttöään ennusteen luotettavuuteen. Vaikka ennuste ei kerrokaan teknistä käyttöikä, kuvaa se kuitenkin toteutunutta saneerausvolyymia ja ennustaa sitä tulevaisuuteen. Jos saneerauksen syy olisi merkitty lähtöaineistoon, voitaisiin vesijohtoverkoston tekninen käyttöikä ja saneeraustarve mallintaa erikseen olettaen, että myös tulevaisuudessa saneerataan muista kuin huonosta kunnosta johtuvista syistä.

Arvioidessa aineistosta johtuvia epävarmuustekijöitä, on syytä myös mainita inhimillisistä virheistä johtuvat merkinnät. Aineistoon on saatettu tallentaa tietoja väärään sarakkeeseen tai muuten tehdä suurempia muutoksia erehdyksessä. Ennen mallinnusta aineistosta on yritetty tunnistaa ja poistaa selkeästi virheelliset arvot, eli merkinnät, jotka eivät vastaa selittäjän kategorialta tai eivät ole suuruudeltaan realistisia, mutta menetelmä ei takaa kaikkien virheiden löytymistä. Lisäksi on todennäköistä, että vanhojen putkien osalta asennustietoja on saatettu täydentää verkkotietojärjestelmään vasta myöhemmin asiantuntijan likimääräisen arvion perusteella. Myös merkintätavat eri työntekijöiden välillä saattavat vaihdella.

## 9 Johtopäätökset

### 9.1 Verkostotiedot käyttöiän mallinnuksessa

Lopputyössä tehdyssä mallinnuksessa lähtöaineiston käsittelyyn ja lukemiseen ohjelmaan meni paljon aikaa, koska lopputyöhön osallistuneet laitokset olivat keränneet verkostonsa kuntotietoja hieman eri käytäntein. Verkostotietojen hyödyntäminen elinajanmallinnuksessa helpottuu, jos tiedot kerätään yhdenmukaisesti. Lopputyössä käsitellyissä aineistoissa yksi epäjohtonmukaisuus ilmeni mm. saneeraustietojen tallentamistavassa. Kuopion Vesi ja Alva olivat siirtäneet tiedon alkuperäisestä asennusvuodesta putken uudeksi ominaisuustiedoksi ja tallentaneet tämän tilalle saneerausvuoden. Saneerausikä oli useimmiten laskettavissa aineistosta, mutta aineiston lukemiseen sisältyi monitulkintaisuutta erityisesti niiden saneeratuiksi merkittyjen putkien kohdalla, joilla saneerausikä tiedot olivat puutteelliset.

Kummastakin aineistosta puuttui paljon saneerausvuositietoja putkilta, jotka oli merkitty saneeratuiksi, mikä oli mallinnuksen kannalta suuri epävarmuustekijä. Näiden putkien osalta jäi lisäksi epäselväksi, puuttuuko putkiosalta asennusvuosi vai saneerausvuosi, tai onko saneerausmerkintä mahdollisesti itsessään virheellinen. Alvan ja Kuopion Veden tilanteessa päätettiin materiaalitetietojen perusteella, että puutteellisesti merkittyjen saneerattujen putkiosien tiedoista puuttuu saneerausvuosi eikä asennusvuosi. Parempi käytäntö olisi tallentaa saneerausvuosi omaan sarakkeeseensa, jotta tulkintavirheiltä vältytään ja aineiston lukeminen helpottuu. Muutos edellyttää kuitenkin verkkotietojärjestelmää, jossa tallennus vastaavalla tavalla on mahdollista.

Tärkein käytäntö verkostotiedon keräämisessä on sen toteutuminen. Verkostotietojen systemaattinen ja riittävä kerääminen on tärkeää, koska laadukkaalla aineistolla voidaan saavuttaa laadukkaita mallinnustuloksia. Vesijohtojen käyttöiän mallinnus edellyttää vähintään tietoa vesijohtojen asennus- ja saneerausvuosista. Lopputyöhön osallistuneen Hangon Veden aineisto ei soveltunut käyttöiän mallinnukseen, koska tieto saneerausvuodesta oli tallennettu asennusvuoden tilalle eikä tietoa alkuperäisestä asennusvuodesta ollut saatavilla, minkä takia vesijohtojen käyttöikä ei ollut laskettavissa.

Puutteelliset saneeraus- ja kuntotiedot puolestaan huonontavat mallien luotettavuutta. Jos vesilaitoksella on saneerattu paljon ennen verkostotietojen keräämistä, puuttuu aineistosta todennäköisesti myös paljon dataa. Jos verkosto on pääosin suhteellisen nuori, aiheuttaa saneerausten vähäinen määrä todennäköisesti malliin epävarmuutta.

Lopputyössä elinajanmallinnus tehtiin olemassa olevien saneeraustietojen perusteella, mutta mallin ennustetta voisi täsmentää kuntotietojen avulla. Vesilaitoksia tulisi siten kannustaa myös keräämään verkostonsa kuntotietoja. Käytännössä tietoja voi kerätä mistä vain putken kuntoon liittyvästä muuttujasta. Havaintoja voidaan saada vuodonetsinnän tai koepalan avulla tai silmämääräisesti saneerauksen yhteydessä. Erityisesti saneeraushetken kunnan dokumentoiminen tukee aineiston tulkintaa, sillä olisi hyödyllistä pystyä arviomaan, mistä syystä vesijohto on saneerattu. Merkintöjen kannattaa kuitenkin olla kategorisoitavissa, jotta niiden hyödyntäminen käyttöiän mallinnuksessa onnistuu. Esimerkiksi vapaamuotoiset sanalliset kuvaukset eivät ole käytännössä hyödynnettävissä.

## **9.2 Käyttöiän mallinnuksen rooli Suomen vesijohtojen saneeraustarpeen arvioinnissa**

Suomessa vesijohtojen käyttöiäksi arvioidaan usein 60 vuotta, vaikka todellinen käyttöikä jakautuu usealle vuosikymmenelle. Kyseisen diplomityön tavoitteena oli arvioida vesijohtojen todellista käyttöikää matemaattisten menetelmien ja saneeraustietojen avulla. Saatujen mallien perusteella käyttöiän eloonjäämisajan mediaani oli suurempi kuin 60 vuotta kaikille materiaali-ryhmille sekä koko mallinnetulle verkostolle. Puutteelliset saneeraustiedot olivat kuitenkin mallinnustulosten kannalta suuri epävarmuustekijä. Puutteelliset tiedot näkyvät aineistossa käytännössä puuttuvana datana, jonka takia mallin perusteella tehdyt ennusteet olivat tavallista positiivisempia. Niiden mallinnettujen materiaali-ryhmien osalta, joilla oli eniten luotettavaa saneeraustietoa, kuten Alvan harmaa valurata (VAL), ennusteet olivat lähellä arvioitua 60 vuotta. Näin ollen voidaan todeta vesijohtojen todellisen keskimääräisen käyttöiän olevan mallinnustulosten perusteella lähellä oletettua 60 vuotta, mutta tarkempaa käyttöiän arviota ei voi niiden perusteella tehdä luotettavasti juuri puuttuvan aineiston takia.

Monella suomalaisella vesilaitoksella on jo kerättynä vesihuoltoverkostojen ominaisuus- ja saneeraustietoja verkkotietojärjestelmään, mutta näitä tietoja ei vielä ole hyödynnetty vesijohtojen käyttöiän mallinnuksessa. Muillakin vesilaitoksilla on siis mahdollisuus arvioida resurssien tarvetta aiempaa luotettavammin matemaattisten menetelmien avulla. Ennustavien mallien avulla voidaan arvioida verkoston käyttöikää ja sitä kautta saneeraustarvetta. Kuitenkin, kuten lopputyössä huomattiin, luotettavien mallien tekeminen edellyttää kattavasti tietoa tehdyistä saneerauksista. Vesilaitosten vastuulla on varmistaa, että verkostotietoa kerätään systemaattisesti, jotta aineistoa voidaan hyödyntää vastaavassa mallinnuksessa. Verkostotiedon kerääminen on jatkuva prosessi, sillä tehtyjä malleja kannattaa myös kalibroida päivitetyn saneeraus- ja kuntotiedon avulla oikeiden tulosten varmistamiseksi. Jos verkoston saneeraustiedoissa on puutteita, voidaan tietoja täydentää

luotettavampien tuloksien saamiseksi. Esimerkiksi verkoston tarkasteleminen karttakuvien avulla ja saneeraus- ja kuntotietojen täydentäminen tunnistettujen alueiden osalta voi olla hyödyllistä. Useissa vesilaitoksissa on henkilökuntaa, joka tuntee verkoston hyvin ja pystyy arvioimaan verkoston asennus- ja saneerausajankohtia. Matemaattisten menetelmien käyttö todennäköisesti yleistyy tulevina vuosikymmeninä, joten verkostotietojen tallentaminen ja resurssointi datan laadun hallintaan kannattaa aloittaa viimeistään nyt.

## 10 Yhteenveto

Suomen vesihuoltoverkostojen saneeraustarpeen huippu lähestyy. Saneerauksiin valmistautuminen edellyttää uusien menetelmien määrätietoista hyödyntämistä saneeraustarpeen arvioinnissa. Diplomityössä vesijohtojen käyttöikää mallinnettiin kolmen suomalaisen vesilaitoksen saneeraustietojen pohjalta. Mallinnustulokset kuvaavat toteutunutta saneerausvolyyymia ja ennustavat sitä tulevaisuuteen. Aikaisemmissa tutkimuksissa suomalaisten vesijohtoverkostojen saneeraustarvetta on mallinnettu lähinnä putkirikkotietojen perusteella. Aikaisemmat vesijohtoverkoston käyttöiän arviot puolestaan perustuvat pitkälti asiantuntija-arvioihin.

Käyttöiän mallinnukseen käytettiin Weibull-jakaumaa ja selittäjien painoarvoa arvioitiin Coxin suhteellisen vaaran mallin avulla. Lisäksi lähtöaineiston tulkitsemisen tueksi haastateltiin vesilaitosten edustajia. Mallinnus tehtiin lopputyöhön osallistuneista vesilaitoksista vain kahdelle, sillä Hangon Veden lähtöaineisto ei täyttänyt mallinnustavan vaatimuksia. Alvan ja Kuopion Veden tapauksessa mallinnettiin ensin koko verkoston käyttöiän jakautumista yhteisesti. Alvan verkoston keskimääräinen ennuste oli noin 60- ja Kuopion noin 80 vuotta. Ero keskimääräisessä ennusteessa johtuu puuttuvan aineiston vaikutuksesta, mikä tyypillisesti tekee ennusteesta optimistisemmän. Verkoston eniten saneeratut materiaalit mallinnettiin erikseen. Eri materiaalien ennusteet ja ennusteiden luotettavuus vaihtelivat merkittävästi. Kaikkien materiaalityyppien keskimääräinen käyttöiän ennuste oli kuitenkin yli 60 vuotta. Mallinnustulosten perusteella vesijohtoverkoston keskimääräinen käyttöikä on lähellä aikaisemmin arvioitua 60 vuotta, mutta tarkempia luotettavia ennusteita ei pystytty lopputyössä tekemään.

Lähtöaineiston laadusta johtuvan merkittävän epävarmuuden takia lopputyössä tunnistettiin tarve samankaltaiselle mallinnukselle kattavamman aineiston avulla, jotta tarkempia tuloksia voidaan saavuttaa. Osalla Suomen vesilaitoksista on jo olemassa käyttöiän mallinnukseen sopivaa aineistoa, mutta monilla verkostotietoja kerätään edelleen puutteellisesti. Lopputyössä korostettiin verkostotietojen systemaattisen keräämisen tärkeyttä ja suositeltiin vakiintuneiden käytäntöjen noudattamista tiedon keräämisessä. Työssä tehty mallinnus antoi uuden näkökulman siitä, miten kerättyä verkostotietoa voidaan hyödyntää saneeraustarpeen arvioinnissa. Käyttöiän mallinnuksella on suuri potentiaali toimia arvokkaana työkaluna saneerauspäätösten tekemisessä, kun mallinnukseen käytetään laadukasta lähtöaineistoa.

## 11 Lähteet

Alva-yhtiöt Oy. (2022). Putket\_pitoaika-arvio\_29\_6\_2022\_lähetetty.xlsx [Alvan lähdemateriaali].

Baur, R. & Herz, R. (2002). Selective inspection planning with ageing forecast for sewer types. *Water science and technology*. 46(6-7). s.389-396.

Berninger, K., Laakso, T., Paatela, H., Virta, S., Rautiainen, J., Virtanen, R., Tynkkynen, O., Piila, N., Dubovik, M., Vahala, R. (2018). Tulevaisuuden kestävä vesihuolto–ennakointi, ohjaus ja järjestäminen. Valtioneuvoston selvitys- ja tutkimustoiminnan julkaisusarja 56/2018. 139 s. ISBN 978-952-287-607-2. Saatavissa: <http://urn.fi/URN:ISBN:978-952-287-607-2>

Bruaset, S. Sægrov S. & Ugarelli R. (2018). Performance-based modelling of long-term deterioration to support rehabilitation and investment decisions in drinking water distribution systems. *Urban Water Journal*. 15:1. 46–52 s. DOI: 10.1080/1573062X.2017.1395894

datagy. (2022). One-Hot Encoding in Sckiti-Learn with OneHotEncoder. Verkkosivu. Viitattu 4.2.2023. Saatavissa: <https://datagy.io/sklearn-one-hot-encode/>

European Commission. (2013). Resource and Economic Efficiency of Water Distribution Networks in the EU. Final Report (ENV.D.1/SER/2010/0029). European Union, Brysseli, Belgia.

Giraldo-González, M. & Rodríguez, J. (2020). Comparison of statistical and machine learning models for pipe failure modeling in water distribution networks. *Water*, 12(4). 1153. DOI:10.3390/w12041153.

Introduction to survival analysis. (2022). Lifelines. [Viitattu 9.10.2022]. Saatavissa: <https://lifelines.readthedocs.io/en/latest/Survival%20Analysis%20intro.html#introduction-to-survival-analysis>

Kleiner, Y. and Rajani, B. (2008). Prioritising individual water mains for renewal. In: *Proceedings of the ASCE/EWRI World Environmental and Water Resources*. 1–10 s. Honolulu, Havaiji. DOI:10.1061/40976(316)498.

Kuulas, A., Renko, T., Kuivamäki R. (2020). Vesihuollon investointitarpeet vuoteen 2040. Vesilaitosyhdistyksen monistesarja nro 63. 88 s. ISBN 978-952-6697-59-8.

Le Gat, Y. & Eisenbeis, P. (2000). Using maintenance records to forecast failures in water networks. *Urban Water*. 2. 173-181 s. DOI:10.1016/S1462-0758(00)00057-1.

Le Gat, Y., Kropp, I. Poulton, M. (2013). Is the service life of water distribution pipelines linked to their failure rate?. *Water Science & Technology. Water Supply*. 12. 386. DOI:10.2166/ws.2013.089

Maa- ja metsätalousministeriö. (2008). Vesihuoltoverkostojen nykytila ja saneeraustarve–YVES-tutkimuksen päivitys 2008. [Viitattu 21.1.2023]. Saatavissa: <https://docplayer.fi/5275414-Maa-ja-metsatalousministerio-vesihuoltoverkostojen-nykytila-ja-saneeraus-tarve.html>

Magi, V. Laakso, T. Metsävuori, J & Mattila, K. (2022). Vesijohtojen kunnan tutkiminen. Suomen Vesilaitosyhdistys ry (VVY). 73 s. Vesilaitosyhdistyksen monistesarja nro 70. ISBN 978-952-6697-77-2.

Määttä, A. (2021). Verkostotiedon merkitys vesihuoltolaitoksen omaisuudenhallinnan kehittämisen näkökulmasta. Diplomityö. Aalto-yliopisto. Espoo. 48 s.

Pelletier, G., Mailhot, A. & Villeneuve, J. (2003). Modeling water pipe breaks-three case studies. *Journal of Water Resources Planning and Management*, 129(2). 115–123 s.

Poulton, M., Le Gat, Y. and Bemond, B. (2007). The impact of pipe segment length on break predictions in water distribution systems. In: *Proceedings of the LESAM-2nd Leading Edge Conference on Strategic Asset Management*. Lisbon, Portugal

Restrepo, A. & Delgado, J. & Echeverria, F. (2009). Evaluation of Current Condition and Lifespan of Drinking Water Pipelines. *Journal of Failure Analysis and Prevention*. 9. DOI:541–548. 10.1007/s11668-009-9291-5.

Silfverberg, P. (2017). Vesihuollon suuntaviivat vuoteen 2020. Suomen Vesilaitosyhdistys ry (VVY). 49 s. Vesilaitosyhdistyksen monistesarja nro 44. ISBN 978-952-6697-31-4. ISSN 2242-7279.

St. Clair, A. & Sinha S. (2012). State-of-the-technology review on water pipe condition, deterioration and failure rate prediction models!. *Urban Water Journal*. 9:2. 85–112 s. DOI: 10.1080/1573062X.2011.644566

Suomen Rakennusinsinöörien Liitto RIL ry. (2017) Rakennetun omaisuuden tila 2017. [Viitattu 4.2.2023]. Saatavissa:

[https://www.ril.fi/media/2017/2017-vaikuttaminen/roti-2017/taustat/roti-2017\\_painettu-raportti.pdf](https://www.ril.fi/media/2017/2017-vaikuttaminen/roti-2017/taustat/roti-2017_painettu-raportti.pdf)

Suomen Rakennusinsinöörien Liitto RIL ry. (2021) Rakennetun omaisuuden tila 2021. [Viitattu 4.2.2023]. Saatavissa: [https://www.ril.fi/media/2021/vaikuttaminen/roti2021\\_low.pdf](https://www.ril.fi/media/2021/vaikuttaminen/roti2021_low.pdf)

Suomen Rakennusinsinöörien Liitto RIL ry. (2023) Rakennetun omaisuuden tila 2023. [Viitattu 4.2.2023]. Saatavissa: <https://www.ril.fi/media/2023/roti-2023/roti-2023.pdf>

Survival regression. (2022). Lifelines. [Viitattu 9.02.2023]. Saatavissa: <https://lifelines.readthedocs.io/en/latest/Survival%20Analysis%20intro.html#introduction-to-survival-analysis>

Valis, D. Hasilová, K. Forbelská, M. & Pietrucha-Urbanik, K. (2017). Modeling water distribution network failures and deterioration. 2017 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM). 924–928 s. DOI: 10.1109/IEEM.2017.8290027.

Valtioneuvoston asetus vesihuollon tietojärjestelmästä ja tiedottamisesta 6/2023. <https://www.finlex.fi/fi/laki/smur/2023/20230006>

Vanrenterghem-Raven, A. (2007). Risk factors of structural degradation of an urban water distribution system. *Journal of Infrastructure Systems*. 13(1). 55–64.

Vesihuoltolaki 119/2001

Vladeanu, G. & Koo, D. (2015). A Comparison Study of Water Pipe Failure Prediction Models Using Weibull Distribution and Binary Logistic Regression. DOI:1590-1601. 10.1061/9780784479360.146.

Wang, C. Niu, Z. Jia, H. & Zhang, H. (2010). An assessment model of water pipe condition using bayesian inference. *Journal of Zhejiang University-Science A (Applied Physics & Engineering)*. 11(7). 495–504 s.

Wang, Y., Zayed, T. and Moselhi, O. (2009). Prediction models for annual break rates of water mains. *Journal of Performance of Constructed Facilities*. 23(1). 47–54 s. DOI:10.1061/(ASCE)0887-3828(2009)23:1(47).