

Fölin runkolinjojen täsmällisyyden analysointi hyödyntäen joukkoliikenteen avointa dataa

Maantiede

LuK-tutkielma

Johannes Salovaara

17.6.2026

Turku

Tutkinto-ohjelma, oppiaine: Luonnontieteiden kandidaatti, maantiede

Tekijä: Johannes Salovaara

Otsikko: Fölin runkolinjojen täsmällisyyden analysointi hyödyntäen joukkoliikenteen avointa dataa

Ohjaaja: Harri Tolvanen

Sivumäärä: 39 sivua + liitteet 3 sivua

Päivämäärä: 17.6.2026

Joukkoliikenteen kehitystyö on merkittävä väline kaupunkiseutujen liikkumisen edistämiseksi ja liikenteen päästöjen vähentämiseksi. Täsmällisyys kasvattaa joukkoliikenteen käyttöastetta ja matkustajamäärää. Turun seudun joukkoliikenteen (Föli) uuden runkolinjaston liikennöinti alkoi 1.7.2025. Uusien runkolinjojen täsmällisyyden analysointi on yksi näkökulma linjaston toimivuuden tarkasteluun ja avoimen datan hyödyntämiseen.

Tutkimuksessa tarkasteltiin Turun seudun joukkoliikenteen liikennöinnin täsmällisyyttä linja- ja sijaintikohtaisesti Fölin datarajapinnasta saatavilla olevan avoimen datan avulla. Tutkimuskohteina ovat Fölin uudet runkolinjat sekä niiden reitit. Tutkimuksen tavoitteena oli selvittää, miten täsmällistä liikenne on ja esiintyykö liikennöinnin täsmällisyydessä ongelmia tai linjojen välisiä eroja, sekä etsiä mahdollisia mahdollisille eroille ja ongelmille selityksiä.

Tutkielman aineistona on reaaliaikaisesta datasta kerätty ja analyysia varten jatkokäsitelty paikkatietoaineisto, joka sisältää pistemäistä tietoa Fölin linja-autojen sijainneista ja myöhästymisistä linjaliikenteessä. Kerätystä liikennöintidatasta laskettiin linjakohtaisia tunnuslukuja, joiden avulla linjojen täsmällisyyttä verrattiin toisiinsa. Sijaintikohtainen tarkastelu toteutettiin karttatarkasteluna paikkatietoa visualisoimalla.

Tutkimuksessa havaittiin sekä eroja että samankaltaisuuksia linjojen välillä. Osalla linjoista liikennöinti oli toisia linjoja täsmällisempää. Linjat, joilla myöhästymiset olivat suuria, kulkivat usein myös etuajassa. Myöhästymiset keskittyivät usealla linjalla tiettyyn ajosuuntaan. Kaupunkien keskusta-alueilla korostuivat myöhästymiset ja aikataulusta etuajassa kulkemista tapahtui enemmän maaseutumaisilla alueilla. Linjojen reiteillä, aikatauluilla sekä liikennöintiajoilla huomattiin olevan vaikutusta liikennöinnin täsmällisyyteen.

Avainsanat: joukkoliikenne, täsmällisyys, runkolinjasto, avoin data, Föli, Turku

Sisällysluettelo

1	Johdanto	5
2	Teoreettinen tausta ja viitekehys	7
2.1	Joukkoliikennejärjestelmät ja Föli	7
2.1.1	Joukkoliikennejärjestelmät yleisesti	7
2.1.2	Turun seudun joukkoliikenne (Föli)	7
2.2	Runkolinjat joukkoliikenteessä	8
2.2.1	Runkolinjat yleisesti	8
2.2.2	Turun seudun joukkoliikenteen runkolinjat	8
2.3	Täsmällisyys joukkoliikenteessä	10
2.4	Joukkoliikenteen avoin data	11
2.4.1	Joukkoliikenteen avoin data yleisesti	11
2.4.2	GTFS, GTFS-RT, SIRI ja NeTEx	12
2.4.3	Joukkoliikenteen dataan perustuva selvitys etenemisnopeuksista ja etuuksista	13
3	Aineistot ja menetelmät	15
3.1	Aineistot	15
3.1.1	Turun seudun joukkoliikenteen avoin data	15
3.1.2	Avoimesta datasta tuotettu paikkatietoaineisto	16
3.2	Menetelmät	17
3.2.1	Aineiston tuottaminen avoimesta datasta Python-ohjelmointikielen avulla	17
3.2.2	Datan analysointi – paikkatietomenetelmät	18
4	Tulokset	22
4.1	Fölin runkolinjojen täsmällisyyden arviointi yleisesti	22
4.2	Linjakohtaiset tulokset	24
4.3	Ongelmalliset sijainnit	29
5	Tulosten tarkastelu	30
5.1	Liikenteen täsmällisyyteen vaikuttavia tekijöitä	30
5.2	Vertailu aiempiin tutkimuksiin muissa kaupungeissa	32
5.3	Tulosten epävarmuudet ja ongelmat	33
5.4	Jatkotutkimusideoita	33

Lähteet	35
Liitteet	40
Liite 1. Python-ohjelmakoodi datan keräämiseksi ja tallentamiseksi	40

1 Johdanto

Toimiva joukkoliikennejärjestelmä on tärkeä osa nykyaikaista kaupunkia (Woltek ym. 2025: 278–280). Joukkoliikenteen käyttäjien kannalta toimiva liikennejärjestelmä mahdollistaa edullisen, tehokkaan ja ympäristöystävällisen liikkumisen sekä eri kulkumuotojen yhdistämisen.

Yhä kasvava yksityisautoilu aiheuttaa kuitenkin haasteita liikkumiselle ja elämiselle kaupungeissa: henkilöautoliikenteen määrä kaupungeissa kasvaa, tämän myötä ruuhkat, melu ja saasteet yleistyvät (Gössling 2020). Henkilöautot vievät kaupunkitilaa paljon joukkoliikennettä ja kevyttä liikennettä enemmän. Joukkoliikenteen kehittäminen on siis tärkeää, sillä tehokkaampi ja käyttöasteeltaan korkeampi joukkoliikenne vähentää kaupungeissa syntyviä päästöjä sekä liikenteen ruuhkia (Gössling 2020; Lyu ym. 2025).

Suomessa liikenneinfrastruktuuria on kehitetty 1950-luvulta alkaen pääasiassa autopainotteisesti ja henkilöautojen määrä lisääntyy Suomessa edelleen (Lehtonen ym. 2024). Pääkaupunkiseudulla joukkoliikenteen käytön osuus on korkea, mutta keskiuurissa kaupungeissa yli 65 % matkoista tehdään henkilöautolla. Turun kaupungin tavoitteena on olla hiilineutraali vuonna 2029 ja joukkoliikenteen kehittäminen on yksi tapa päästä lähemmäs tätä tavoitetta (Turun kaupunki 2022).

Turun seudulla on otettu ensiaskeleita joukkoliikennejärjestelmän kehittämisessä, kun uusi runkolinjasto otettiin käyttöön 1.7.2025 (Runkolinjasto | Föli 2026). Uudistuksen myötä liikenteen vuoroväli on monilla alueilla lyhentynyt ja vastaavasti liikenteen määrä lisääntynyt. Uuden järjestelmän liikennöinti ei ole kuitenkaan ollut ongelmattomaa: paikoin on kärsitty täysistä busseista tai myöhästymisistä, joillain linjoilla puolestaan linja-autojen kulkemisesta etuajassa. Ideaalitulanteessa liikenne ajetaan tietysti aikataulun mukaisesti, mutta liikenteen sujuvuuteen vaikuttavia tekijöitä on paljon.

Tästä syystä on mielekästä tarkastella Turun seudun joukkoliikennettä seuraaviin tutkimuskysymyksiin tukeutuen:

1) Miten täsmällisesti liikenne runkolinjoilla toimii?

2) Millä linjoilla tai missä sijainneissa täsmällisyyden ongelmia syntyy?

3) Mikä aiheuttaa täsmällisyyden vaihtelua?

Tässä tutkielmassa analysoin linja-autoliikenteen aikataulujen toteutumista ja liikenteen täsmällisyyttä Turun seudun joukkoliikenteen tuottamaa avointa reaaliaikaista dataa hyödyntäen. Tarkastelen liikennöintiä linja- ja paikkakohtaisesti.

2 Teoreettinen tausta ja viitekehys

2.1 Joukkoliikennejärjestelmät ja Föli

2.1.1 Joukkoliikennejärjestelmät yleisesti

Joukkoliikennejärjestelmät ovat monimutkaisia kokonaisuuksia, joihin vaikuttavat muun muassa lainsäädäntö, poliittiset päätökset, teknologian kehitys ja taloudelliset tekijät (Canitez 2019). Tällaisen järjestelmän päätoimijoita ovat hallinto, liikennöitsijät ja matkustajat. Joukkoliikennejärjestelmän toimintaa säätelee yleensä jokin joukkoliikenneviranomainen, jonka kanssa yhteistyössä liikennöitsijät tuottavat joukkoliikennepalveluita matkustajille. Joissain tapauksissa liikenne voi olla myös osittain tai kokonaan joukkoliikennettä hallinnoivan toimijan, kuten kunnan tai paikallisen hallinnon tuottamaa. Useimmiten joukkoliikennepalveluiden tuottajat valitaan kuitenkin avoimilla liikennöintisopimusten kilpailutuksilla, joihin osallistuneista yrityksistä valitaan kriteerien mukaisesti paras tuottamaan kyseisen sopimuksen määrittämää palvelua (Lodi ym. 2016).

Lisäksi toimiva joukkoliikennejärjestelmä edellyttää useiden erilaisten organisaatioiden yhteistyötä esimerkiksi kaupunkisuunnittelussa ja maankäytön suunnittelussa (Hrelja ym. 2020).

2.1.2 Turun seudun joukkoliikenne (Föli)

Föli on Turun seudulla vuodesta 2014 alkaen toiminut joukkoliikenneviranomainen, joka järjestää liikennettä seitsemän kunnan alueella (Turku, Kaarina, Raisio, Lieto, Naantali, Paimio ja Rusko) (Tietoa Fölistä 2026). Föli järjestää seudullisen joukkoliikenteen, suunnittelee aikataulut sekä kilpailuttaa liikennöintisopimuksille liikennöitsijät (Mutiganda ym. 2023; Tietoa Fölistä 2026).

Fölin liikenteessä ei ole vyöhykkeitä, sillä liikennöintialue on yhtenäinen kokonaisuus, jonka sisällä lipputuotteet ja niiden hinnat ovat kaikkialla samat (Mutiganda ym. 2023). Liikennöintialueen kunnat kehittävät liikennettä yhdessä, mutta jokaisella kunnalla on mahdollisuus päättää oman alueensa liikenteestä. Kunnat maksavat alueidensa sisällä kulkevasta liikenteestä itsenäisesti ja usean kunnan alueella kulkevat seutulinjat

rahoitetaan yhdessä perustuen muun muassa matkustajamääriin, etäisyyksiin sekä pysäkkien määriin siten, että jokainen kunta maksaa oman osuutensa.

2.2 Runkolinjat joukkoliikenteessä

2.2.1 Runkolinjat yleisesti

Joukkoliikenteen käyttäjän näkökulmasta selkeä joukkoliikennelinjasto perustuu runkoyhteyksiin ja niitä täydentävään liikenteeseen (Rosenberg ym. 2007: 24, 37).

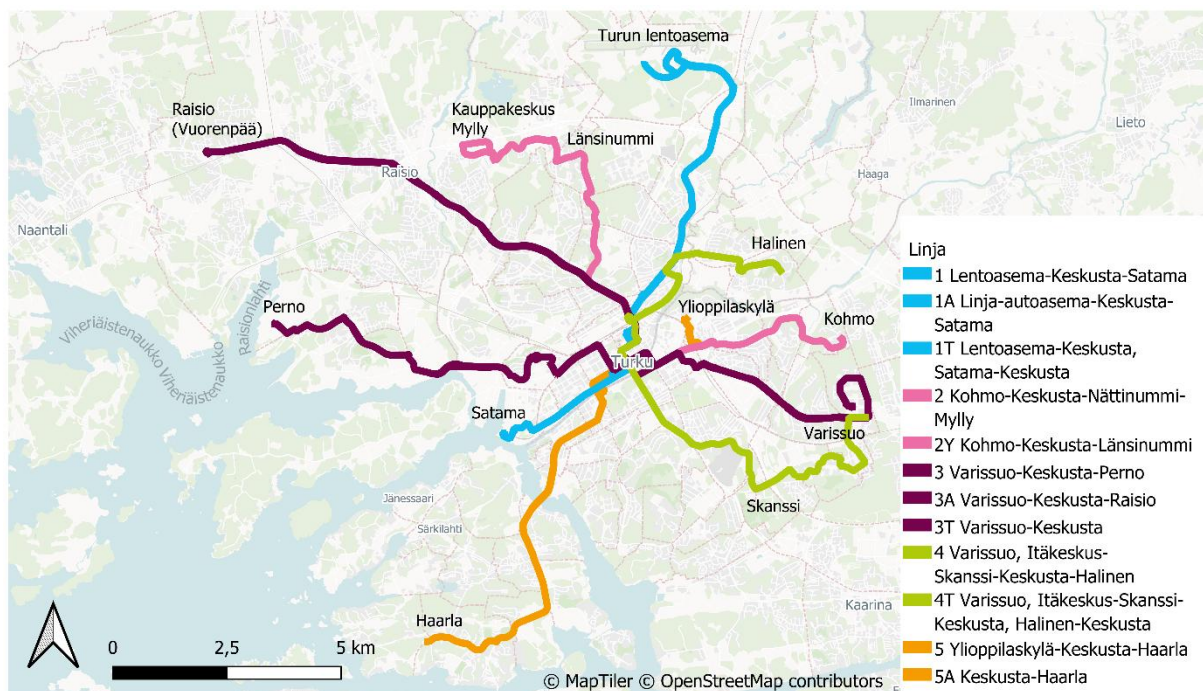
Runkolinjat voivat olla tiheästi kulkevaa raide- tai bussiliikennettä, jotka muodostavat liikenteen pääverkon ja linjastoa täydentävät muut harvemmin kulkevat linjat. Keskeistä runkolinjaston toimivuudessa on sujuvat vaihtoyhteydet runkoliikenteen sisäisesti ja muun liikenteen kanssa. Kasvavilla kaupunkialueilla on myös tarvetta poikittaisille runkoyhteyksille. Näiden tarve lisääntyy, sillä liikenne ei suuntaudu enää pelkästään keskusta-alueelle ja sieltä pois.

Runkoliikenteen pitäisi myös erottua muusta liikenteestä tiheämmin ja täsmällisemmin liikennöitynä (Rosenberg ym. 2007: 24, 37). Runkolinjojen pitäisi olla matka-ajan suhteen kilpailukykyisiä autoliikenteen kanssa, siksi runkoliikennettä kehittäessä on tarve lisätä joukkoliikenteen etuuksia muuhun liikenteeseen nähden. Tämä lisää myös liikenteen kannattavuutta ja kustannustehokkuutta.

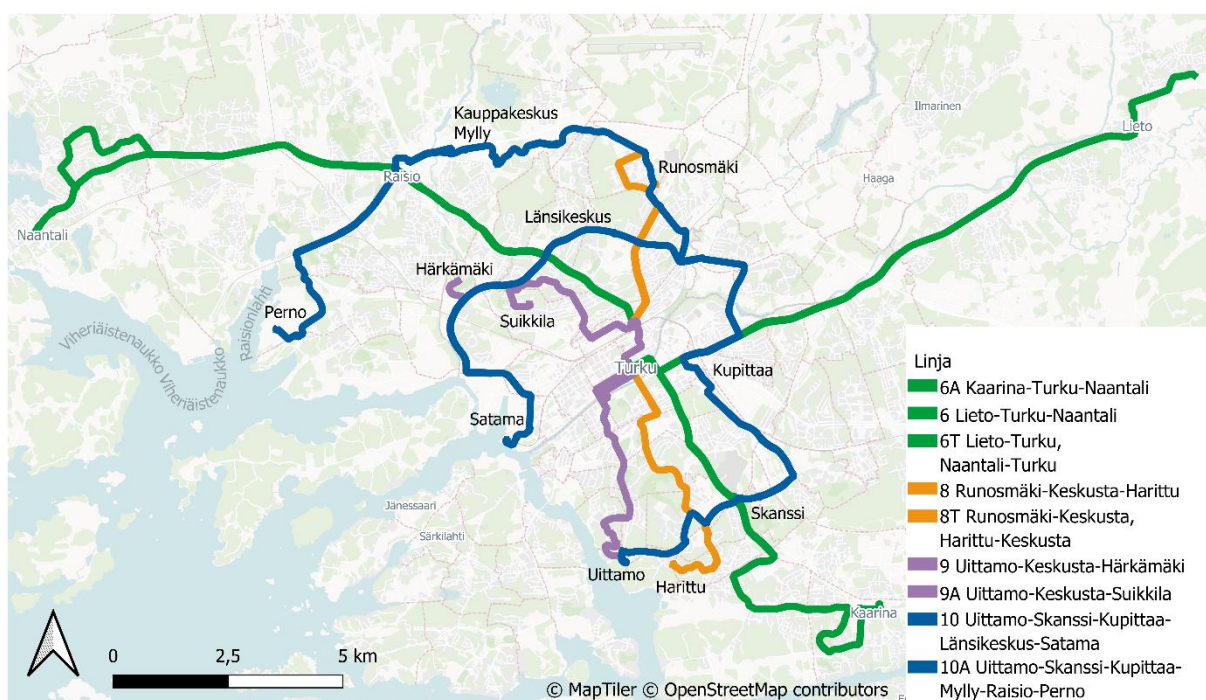
2.2.2 Turun seudun joukkoliikenteen runkolinjat

Fölin runkolinjat ovat linjat 1–10 kirjainvariantteineen, pois lukien linja 7, jota ei ainakaan tällä hetkellä ole olemassa (Runkolinjasto | Föli 2026). Linjat 1–9 ovat keskustan läpi kulkevia ja lähinnä pitkittäissuuntaisia reittejä (kuva 1; kuva 2). Linjat 10 ja 10A ovat uuden linjaston suurin uudistus, sillä ne tuovat Fölin liikennöintialueelle muita linjoja yhdistävät kehämäiset runkoyhteydet, jotka mahdollistavat etenkin poikittaissuuntaisen ja keskustan kautta kulkemattoman liikenteen.

Osa runkolinjoista on haaroitettu toisesta päästä kahtia siten, että joka toisen vuoron määränpää on eri paikassa (Turun seudun joukkoliikenteen aikataulut 2025). Näin suurimalle osuudelle reitistä on saatu tiheä vuoroväli, mutta myös pienemmän kysynnän alueelle saadaan järjestettyä liikennettä. Näin on toteutettu linjaparien 3 ja



Kuva 1: Fölin runkolinjat 1–5 kirjainvariantteineen.



Kuva 2: Fölin runkolinjat 6–10 kirjainvariantteineen.

3A, 6 ja 6A, 9 ja 9A sekä 10 ja 10A liikenne. Linja 3A tosin on vain ruuhka-ajan liikenteessä Raisioon ja Varissuolle lisäkapasiteettia tuova linja. Linjan 3 vuoroväli on jo yksinään samalla tasolla muiden linjaparien kanssa.

Edellä kuvatun mukaisesti linjat 6 ja 6A muodostavat tiheään kulkevan yhteyden Naantali-Raisio-Turku osuudelle, mutta Liedon ja Kaarinan suuntaan eriävillä haaroilla vuoroväli on pidempi (Turun seudun joukkoliikenteen aikataulut 2025). Samaan tapaan

linjat 9 ja 9A muodostavat tiheään liikennöidyn yhteyden Turun keskustan ja Uittamon välillä. Härkämäen ja Suikkilan päissä reitit haarautuvat erilleen ja täällä vuoroväli on myös linjojen yhteistä osuutta harvempi. Myös linjojen 10 ja 10A yhteisellä reittiosuudella on tiheä vuoroväli, mutta Raisio-Perno- ja Länsikeskus-Satama-reittihaaroilla liikennöidään harvemmalla vuorovälillä.

Runkolinjoilla 1, 2, 4 ja 5 ei ole haarautuvia reittivariantteja ja linjojen liikennöinti on kaikkina liikennöintiäikoina tiheää (Turun seudun joukkoliikenteen aikataulut 2025). Ruuhka-aikaan liikennöidään lisäksi linjoja 1A ja 5A. Nämä tuovat lisäkapasiteettia väleille satama-keskusta-linja-autoasema ja keskusta-Haarla.

Lisäksi linjoilla 1, 3, 6 ja 8 on Turun keskustaan päättyvät reittivariantit, jotka voi erottaa muista linjan vuoroista linjatunnuksen perään lisäystä T-kirjaimesta. Linja 3T ajetaan ainoastaan Varissuolta Kauppatorille, muiden linjojen T-versiot ajetaan kummastakin suunnasta Turun keskustaan. Tämän tutkielman tarkastelujaksolla liikennöitiin kuitenkin vain linjaa 6T.

Runkolinjaston suunnittelussa on otettu huomioon mahdollinen tuleva nivelbusseilla tehtävä liikennöinti ja Föli on jo neuvotellut liikennöitsijöiden kanssa nivelbussien koekäytöstä (Turun kaupunkiseudun joukkoliikennelautakunta 2026a–b).

Tulevaisuuden varalta runkolinjoille on varattu myös linjanumerot 11–19 (Runkolinjasto | Föli 2026). Kaupungin kasvaessa tiheitä runkoyhteyksiä tullaan todennäköisesti tarvitsemaan lisää. Lisäksi linjalle 5 on suunnitteilla muutoksia: linjaa on tarkoitus jatkaa Ylioppilaskylän päässä myös Ylioppilaskylän itäpuolelle, sekä muuttaa joka toisen vuoron reittiä Hirvensalossa palvelemaan paremmin saaren keskiosaa (Turun kaupunkiseudun joukkoliikennelautakunta 2026c). Myös linjalle 10A on suunnitteilla reittimuutos: Skanssin kauppakeskuksen jälkeen linja käännettäisiin Uittamon suunnan sijaan linjan 4 tapaan Lauste-Varissuo-suuntaan siten, että linja päättyisi lopulta linjojen 3 ja 3A käyttämälle Varissuon päätepysäkille Katteluksenkadulle.

2.3 Täsmällisyys joukkoliikenteessä

Täsmällisyys on keskeinen joukkoliikenteen palvelun laatua määrittelevä tekijä (Wotek ym. 2025: 279). Sillä tarkoitetaan aikataulussa suunniteltujen lähtö- ja saapumisaikojen

toteutumisen astetta. Täsmällisyys on suoraviivainen tapa kuvailla ja mitata liikennöinnin laatua. Wotek ym. (2025) mukaan täsmällisyys voidaan jakaa subjektiiviseen ja objektiiviseen täsmällisyyteen. Subjektiivisella täsmällisyydellä tarkoitetaan joukkoliikenteen käyttäjän havainnoimaa täsmällisyyttä, joka syntyy aikataulun lähtö- ja saapumisaikojen vertailusta oikeisiin saavutettuihin aikoihin. Objektiivinen täsmällisyys tarkoittaa mitattua eroa aikataulutettujen ja ajoneuvojen oikeasti saavuttamien aikojen välillä.

Täsmällinen joukkoliikenne on aikataulun mukaisesti ajettavaa liikennettä, jossa liikenneväline on pysäkillä joukkoliikenteen käyttäjälle tarjotun tiedon mukaisesti (Rosenberg ym. 2007). Täsmällisen liikenteen tuottaminen vaatii usean liikenteen osa-alueen yhteistyötä: vaihtoyhteyksien eri linjojen ja liikennemuotojen välillä tulee onnistua ja vaihtoaikojen on oltava lyhyitä. Liikenteen täytyy olla sujuvaa ja matkustajainformaation paikkansapitävää. Täsmällisyys vaikuttaa myös joukkoliikenteen kysyntään eli sen käyttäjämääriin (Chakrabarti 2015).

Raportissaan joukkoliikenteen palvelutasosta Rosenberg ym. (2007) toteavat, että joukkoliikennettä ei saisi ajaa edellä aikataulusta (Rosenberg ym. 2007). Tämä on tärkeää, jotta vaihtoyhteydet toimivat. Fölin reittiopas antaa oletuksena linjojen väliseksi vaihtoajaksi 5 minuuttia (Fölin reittiopas 2026). Analyysissani tarkastelen aineistoa siitä näkökulmasta, että liikenne on kohtuullisen täsmällistä, jos se kulkee aikataulun mukaisesti tai alle viisi minuuttia aikataulusta jäljessä. Tällöin Fölin reittioppaan ehdottamien vaihtojen pitäisi onnistua, olettaen, että ne tapahtuvat samalla tai viereisellä pysäkillä.

2.4 Joukkoliikenteen avoin data

2.4.1 Joukkoliikenteen avoin data yleisesti

Joukkoliikenteen avoin data on ilmaiseksi verkossa jaettavaa dataa, joka sisältää tietoa esimerkiksi joukkoliikenteen suunnitelluista aikatauluista ja joukkoliikennevälineiden liikkeistä (Lock ym. 2020; Jiang ym. 2025). Tarkemmin tietoa on esimerkiksi paikkatietona pysäkeistä sekä ajettavien vuorojen lähtöpaikoista ja määränpäistä, sekä

näiden aikatauluista. Dataa voidaan linkittää eri lähteistä ja tiedostoista toisiinsa yhteensopivien ID-tunnusten avulla.

Avointa dataa on saatavissa monissa eri muodoissa riippuen joukkoliikenteen järjestäjästä, kaupungista ja valtiosta. Staattista eli ei-reaaliaikaista dataa on saatavilla hyvin kattavasti ympäri Suomen (Fintraffic 2026). Reaaliaikaisen datan saatavuus vaihtelee paljon alueittain. Joukkoliikenteen reaaliaikaista dataa on Suomessa saatavilla ainakin Turun seudulta, pääkaupunkiseudulta, Tampereen seudulta, Oulusta, Jyväskylästä, Joensuusta, Lahdesta ja Kuopiosta (Documentation | Waltti GTFS Realtime 2026; Avoin data | Vilkku 2026; TSJL - API 2026; Avoin data | HSL | HSL.fi 2026). Muualta pohjoismaista dataa on saatavilla vielä kattavammin: Norjasta reaaliaikaista dataa on kattavasti lähes koko maasta ja Ruotsistakin usean eri läänin alueelta (Entur AS 2025; NeTEx Regional | Trafiklab 2025).

2.4.2 GTFS, GTFS-RT, SIRI ja NeTEx

Yleisimmät käytössä olevat joukkoliikenteen avoimen datan muodot ovat GTFS ja GTFS-RT, nämä ovat alun perin Googlen ja TriMetin yhteiseen kehitystyöhön perustuvia datatyyppejä erityisesti joukkoliikenteen tarpeisiin (Park ym. 2020: 371–372; Jiang ym. 2025; About - General Transit Feed Specification 2026). GTFS (General Transit Feed Specification, aluksi Google Transit Feed Specification) on avoin standardi, joka määrittelee tavan jakaa joukkoliikenteen avointa dataa sekä mahdollistaa eri joukkoliikennetoimijoiden datan yhteensopivuuden tätä hyödyntävien sovellusten kanssa. Dataa on kahta eri tyyppiä: GTFS-data on staattista, harvemmin muuttuvaa dataa, joka sisältää tiedot esimerkiksi jokaisen viikonpäivän ajettavista linjoista ja niiden aikatauluista, sekä pysäkkien paikkatietoja. GTFS-RT on reaaliaikaista dataa, joka koostuu yleensä ajoneuvojen sijainti- ja liikkumistiedoista, pysäkkien vuorotiedoista sekä häiriötiedotteista (Lock ym. 2020). GTFS-data on suunniteltu erityisesti informaation välittämiseen joukkoliikenteen matkustajille (Overview - General Transit Feed Specification 2026). Nykyisin GTFS-standardin kehitystyöstä vastaa MobilityData, joka on vuonna 2019 perustettu voittoa tavoittelematon yhteisö (History | MobilityData 2026).

Vastaava eurooppalainen datastandardimalli on Public Transport Reference Data Model (lyhennettynä Transmodel). Tämä malli mahdollistaa erilaisten joukkoliikenteen tietojärjestelmien välisen yhteensopivuuden ja tiedonvälityksen (About – Transmodel 2026). GTFS:ää vastaavaa tietoa sisältävät mallin osastandardit SIRI ja NeTEx. Näiden lisäksi on standardeja esimerkiksi reittioppaiden reittisuunnitteluun ja joukkoliikenteen raakadatan välitykseen (Standards for implementation – Transmodel 2026).

SIRI on standardi, jota käytetään liikennöinnin suunnitellun tai reaaliaikaisen tilanteen tiedonvälitykseen eri tietojärjestelmien välillä (SIRI – Transmodel 2026). SIRI-dataa on saatavilla alueesta ja toimijasta riippuen muun muassa ajoneuvojen reaaliaikaisesta etenemisestä liikenteessä ja reaaliaikaisista pysäkkien aikataulutiedoista. Sitä voidaan hyödyntää esimerkiksi vaihtoyhteyksien toteuttamisessa ja liikennöintiin liittyvien tiedotteiden välityksessä.

NeTEx (Network Timetable Exchange) on tekninen määritelmä joukkoliikenteen aikataulujen ja niihin liittyvien tietojen välittämiseen etenkin joukkoliikenteen käyttäjille (NeTEx – Transmodel 2026). Lisäksi NeTExin avulla voidaan välittää tietoa esimerkiksi pysäkkien sijainneista, joukkoliikenteen liikennöitsijöistä ja joukkoliikenteen hinnoittelusta.

2.4.3 Joukkoliikenteen dataan perustuva selvitys etenemisnopeuksista ja etuuksista

Varsinais-Suomen liitto teetti vuonna 2021 Sitowise Oy:llä selvityksen Turun seudun joukkoliikenteen vaihtopaikoista, liityntä- ja solmupysäkeistä ja joukkoliikenteen liikennevaloetuksista (Hyvönen ym. 2021: 1–6). Selvitys toteutettiin yhteistyössä Varsinais-Suomen ELY-keskuksen, Fölin ja useiden Varsinais-Suomen kuntien kanssa. Selvityksen ensimmäisessä osa-alueessa keskityttiin joukkoliikenteen solmu- ja vaihtopaikkoihin ja niiden kehittämiseen (Hyvönen ym. 2021: 71–74). Toinen osa-alue keskittyi liikennevalojen joukkoliikenne-etuuksien sekä joukkoliikenteen sujuvuuden kehittämiseen. Analyysissä yhdistettiin Fölin liikennöintidataa ja suunniteltua aikatauludataa. Selvityksessä etsittiin syitä linjan 7 (vastaa nykyistä linjaa 6A) hitaalle etenemiselle sekä tutkittiin mahdollisten joukkoliikenne-etuuksien käyttöönottoa. Lisäksi tarkasteltiin linjan 18 (nykyinen linja 8) etenemistä ennen liikennevaloetuuksien

käyttöönottoa ja sen jälkeen. Selvityksen ohella tuotiin myös esille ideoita joukkoliikennedatan kehittämiseksi.

Selvitystyön tuloksena esitettiin uusia kehitysmahdollisuuksia ja toimenpide-ehdotuksia joukkoliikenteen liityntäpysäköintiin, vaihtoyhteyksiin ja reittimuutoksiin. Föli-liikenteen etenemisen huomattiin olevan tiiviissä kaupunkiympäristössä hitaampaa kuin maaseudulla (Hyvönen ym. 2021: 79). Liikennevaloilla ja pysäkeillä huomattiin olevan suurta vaikutusta matkanopeuteen. Selvitystyön ohella kehitettiin sovellus vuorojen keskimääräisen etenemisen tarkastelemiseksi eri vuorokaudenaikoina arkin ja viikonloppuisin. Linjan 7 tarkastelussa huomattiin aikataulun olevan liian tiukka ja luotiin ehdotuksia linjan etenemisen sujuvoittamiseksi (Hyvönen ym. 2021: 81). Linjan 18 liikennevaloetuuksien tarkastelussa tutkittiin linja-auton viipymisaikaa risteysalueilla ennen etuuksien käyttöönottoa ja sen jälkeen (Hyvönen ym. 2021: 82–83). Tulokset olivat vaihtelevia, paikoin valoetuudet nopeuttivat etenemistä ja paikoin hidastivat sitä. Selittävänä tekijänä tässä oli esimerkiksi ajoneuvon tulosuunta risteykseen tai se, mihin kohtaan liikennevalojen vaihtumissykliä ajoneuvo osuu. Joukkoliikennedatan kehitysehdotuksina esitettiin muun muassa ajoneuvodatan paikannustarkkuuden lisäämistä päivitysten aikaväliä pienentämällä ja raakapaikannusdatan tallentamista pidemmältä aikaväliltä tulevia analyyseja varten (Hyvönen ym. 2021: 91).

3 Aineistot ja menetelmät

3.1 Aineistot

3.1.1 Turun seudun joukkoliikenteen avoin data

Turun seudun joukkoliikenne tarjoaa avointa dataa julkisesti osoitteessa data.foli.fi, ja se on vapaasti käytettävissä Creative Commons Nimeä 4.0 -lisenssin mukaisesti. Datarajapinnasta on saatavilla viittä erityyppistä dataa: GTFS-data on staattista harvoin muuttuvaa dataa joukkoliikenteen reiteistä, aikatauluista ja pysäkeistä. SIRI-tyyppistä dataa on saatavilla ajoneuvojen reaaliaikaisesta liikennöinnistä sekä pysäkeille saapuvien vuorojen saapumisajoista pysäkkikohtaisesti. ALERTS-data sisältää tiedotteita, viestintää ja vuoroperuutuksia. GEOJSON-data on paikkatietoa muun muassa Fölin palvelupisteistä sekä liikennöintialueen rajoista. Tutkimukseni kannalta keskeiset datalähteet ovat Fölin ajoneuvoseurannan data (SIRI-VM, eli Vehicle Monitoring) ja GTFS-data.

Fölin SIRI-VM datasta on saatavilla hyvin monenlaista tietoa liittyen jokaiseen liikenteessä olevaan ajoneuvoon, sen sijaintiin ja sen ajamaan vuoroon. Tutkimukseni kannalta keskeisimmät datakentät ovat *vehicleref* (ajoneuvon yksilöllinen tunniste datassa), *latitude* ja *longitude* (ajoneuvon reaaliaikainen sijainti koordinaatteina), *delaysecs* (tieto ajoneuvon kulkemisesta etuajassa tai myöhässä sekunteina), *lineref* (linjanumero), *originname* (lähtöpaikan nimi) sekä *destinationname* (määränpään nimi). Ajoneuvoseurannan tiedot päivittyvät parhaimmillaan viiden sekunnin välein.

Havaitsin datassa jonkin verran epäluotettavuutta, hetkellisesti dataa ei ole ollut saatavilla ollenkaan, ja toisinaan ajoneuvojen sijainnit sekä myöhästymistieto eivät täsmänneet ollenkaan kyseisellä hetkellä ajettaviin vuoroihin. Datassa esiintyy myös ajoneuvoja, joilla ei jostain syystä ole sijaintitietoa, tällöin koordinaatit sisältävät datakentät puuttuvat kokonaan. Toisinaan ajoneuvojen koordinaatteihin perustuva sijainti ja tieto seuraavasta pysäkestä olivat myös ristiriidassa keskenään: ilmoitettu seuraava pysäkki saattoi sijaita aivan eri paikassa kuin ajoneuvon koordinaatteihin perustuva sijainti. Tällöin myös myöhästymistieto näytti epäluotettavia arvoja. Tällaisissa tilanteissa pystyin suodattamaan kerätystä datasta ajoneuvoja pois

ajoneuvotunnisteen avulla. Pääasiassa data on kuitenkin luotettavaa, enkä usko näillä pienillä ongelmilla olevan vaikutusta analyysini paikkansapitävyyteen, kunhan osaan ottaa ne huomioon.

Ajoneuvoseurannan datan lisäksi hyödynnän analyysivaiheessa GTFS-datan shapes.txt-tiedostossa olevaa dataa. Tämä tiedosto sisältää kaikki Fölin linjojen reitit koordinaattipisteinä linjojen alusta loppuun.

3.1.2 Avoimesta datasta tuotettu paikkatietoaineisto

Avoimesta datasta tuottamani aineisto sisältää pistemäistä paikkatietoa Fölin liikennöintialueelta kuuden päivän ajalta. Aineisto on kerätty 6.3.-12.3.2026 välisenä aikana (perjantai–torstai). Päivältä 11.3.2026 ei ole teknisten ongelmien vuoksi dataa. Data on kerätty Fölin avoimen datan ajoneuvoseurannasta 15 minuutin välein ja tallennettu jokaiselta havainnointipäivältä erikseen shapefile-tiedostoon. Seuraavan vuorokauden puolelle jatkuvat vuorot ovat siis tallennettuna osittain kahden eri vuorokauden tiedostoon, mutta tällä ei ole tutkimukseni kannalta merkitystä. Yhdistin myöhemmin kaikkien tarkastelupäivien tiedot yhteen shapefile-tiedostoon analyysia varten.

Tuottamani tiedostot sisältävät seuraavat tiedot, jotka on käytännössä tallennettu suoraan ajoneuvoseurannasta ja siksi myös datakentät on nimetty pitkälti alkuperäisen datan tavoin: *vehicleref* eli ajoneuvon yksilöllinen tunniste, *lineref* eli ajettavan vuoron linjatunnus, *destination* eli linjan määränpää, *delay* eli ajoneuvon myöhästymistieto desimaalilukuna (muutettu minuuteiksi), *next_stop* eli seuraavan pysäkin nimi ja *timestamp* eli tieto kyseisen datapisteen tallennuspäivämäärästä ja kellonajasta. Nämä tiedot on tallennettu koordinaattien lisäksi jokaisen datapisteen kohdalle.

Tiedostossa, johon kaikkien päivien data on yhdistetty, on tiedossa myös alkuperäisen tason/tiedoston nimi. Lisäksi laskin jokaiselle datapisteelle pyöristetyn myöhästymisarvon *delay_min*, joka kertoo myöhästymisen kokonaislukujen tarkkuudella minuutteina. Tämä arvo helpotti datan tarkastelua, mutta en kuitenkaan hyödyntänyt sitä varsinaisessa analyysissä.

Kerättyjä datapisteitä on yhteensä 32 584, mutta tähän sisältyy myös Fölin datasta tulevaa virheellistä tai ei-paikkansapitävää tietoa, joka täytyy myöhemmässä analyysivaiheessa suodattaa pois.

3.2 Menetelmät

3.2.1 Aineiston tuottaminen avoimesta datasta Python-ohjelmointikielen avulla

Kirjoitin Python-ohjelmointikielellä ohjelman (liite 1), joka keräsi dataa automaattisesti Fölin avoimen datan rajapinnasta 15 minuutin välein ja tallensi kerätyt tiedot jokaisen tarkasteluvuorokauden lopussa Shapefile-tiedostoiksi.

Ohjelma lähettää Fölin ajoneuvoseurannan rajapinnalle datapyynnön ja saa vastauksena JSON-muotoista dataa, joka sisältää tietoa kaikista kyseisellä hetkellä Fölin liikenteessä olevista ajoneuvoista. Tämä data täytyy parsia JSON-muotoisesta datasta sellaiseksi, että sen käsittely Python-kielellä onnistuu. Käsiteltävä data sisältää ajoneuvon ja sen ajaman vuoron tietoa avain-arvo-pareina.

Tämän jälkeen ohjelma suodattaa datasta vain haluamieni ajoneuvojen tietoja. Ohjelma käy läpi jokaisen ajoneuvon ja valitsee jatkokäsittelyyn ajoneuvodataa seuraavien piirteiden mukaisesti:

Ensimmäisenä ohjelma tarkistaa, että ajoneuvolla on datassa sijaintitieto. Jostakin syystä joiltain autoilta sijaintitieto puuttuu kokonaan. Analyysissani tällaiset ajoneuvot ovat hyödyttömiä, sillä niiden data ei ole sijainnin puuttumisen takia luotettavaa, eikä niiden avulla voida muodostaa pistemuotoista paikkatietodataa, kun paikkatieto puuttuu. Ajoneuvot, joilla sijaintitietoa ei ole, ohitetaan siis datankeräysvaiheessa kokonaan.

Tämän jälkeen ohjelma tarkistaa, että ajoneuvo on ajamassa jotakin runkolinjaa. Jos olisin valinnut tarkasteluun kaikki Fölin liikennöimät linjat, kokonaisuutta olisi ollut todella hankala hallita ja datapisteiden määrä olisi kasvanut todella suureksi.

Käytännössä tämä tapahtuu tarkastamalla, että ajoneuvon kohdalla datakentän *lineref* arvo vastaa jonkin runkolinjan linjatunnusta.

Seuraavaksi ohjelma valitsee datasta haluamani parametrit, ja tallentaa ne ohjelman välimuistiin eräänlaisena ajoneuvo-objektina. Käytännössä siis *selected_data*-nimiseen tietorakenteeseen tallennetaan tarkasteluhetken runkolinjoja ajavat ajoneuvot.

Vuorokauden lopussa ohjelma tallentaa kaikki vuorokauden aikana kerätyt tiedot pistemuotoisena paikkatietona shapefile-tiedostoon. Jokainen tallennettu ajoneuvo-objekti käydään läpi ja sen tiedot lisätään tallennettavaan tiedostoon datapisteeksi. Kun kaikki vuorokauden aikana tallennetut tiedot on käyty läpi, tiedosto tallennetaan. Tämän jälkeen tietorakenne, johon ajoneuvo-objektit oli tallennettu, tyhjennetään ja tallennusprosessi alkaa alusta.

Ohjelmakoodi on osin kirjoitettu mukaillen Helsingin Yliopiston GeoPython - AutoGIS -kurssin ohjeita Creative Commons Attribution-ShareAlike 4.0 International -lisenssin mukaisesti (Tenkanen 2016) ja The Math Labin (2025) datankäsittelyohjeisiin perustuen. Käytin koodin kehitystyössä virheiden ja niiden aiheuttamien ongelmien ymmärtämisen tukena OpenAI:n ChatGPT-tekoälyä.

3.2.2 Datan analysointi – paikkatietomenetelmät

Datan keräämisen jälkeen olin tuottanut kuusi shapefile-tiedostoa, yhden kultakin havainnointipäivältä. Havaintopisteitä näissä tiedostoissa oli yhteensä 32 584. Yhdistin nämä tiedostot yhdeksi aineistoksi käyttäen QGIS:n vektoritasojen yhdistystyökalua (merge). Kirjoittamani koodi tallensi tiedot käyttäen EPSG:4326-koordinaattijärjestelmää. Tämä ei kuitenkaan ollut järkevä valinta pisteiden visualisoinnin kannalta, joten tallensin tiedot uutena vektoritasona projisoiden ne TM35FIN-koordinaattijärjestelmään (EPSG:3067).

Kaikki paikkatietoaineistoon kerätyt pisteet eivät olleet analyysiin sopivia keräysvaiheessa tehdyistä suodatuksista huolimatta, sillä pisteitä tallentui myös linjojen reittien ulkopuolelle oleviin sijainteihin. Toisinaan reaaliaikadatassa olevat autot saattoivat ilmoittaa olevansa runkolinjalla, vaikka niiden koordinaatteihin perustuva sijainti oli muualla. Pisteiden rajaaminen ennen varsinaista analysointia oli siis tarpeen.

Tällaista rajausta varten tarvitsin uuden karttatason linjojen reiteistä. Tämän luominen onnistui QGIS:n GTFS GO -lisäosalla oletusasetuksia käyttäen. Tämä lisäosa tuottaa

GTFS-datan pakatusta zip-tiedostosta GeoJSON-muotoiset stops- ja routes-nimiset tiedostot. Nämä tiedostot sisältävät siis vektorimuotoiset pysäkki- ja reittitiedot. Tässä analyysissä en hyödyntänyt pysäkkitietoja. Tallensin reittitiedot sisältävän karttatason shapefile-muotoiseksi tiedostoksi siten, että jokainen reitti oli erillisenä vektoriobjektina tiedostossa. Tämän jälkeen valitsin uudelta karttatasolta runkolinjojen reitit ja muodostin puskurityökalua (buffer) käyttäen 25 metrin puskurin valittujen kohteiden ympärille. GPS-paikannuksessa voi olla epätarkkuutta, jolloin reitillä olevan auton sijainti voi olla hieman sivussa reitistä, siksi pisteiden valintaan käytettävän puskurin on oltava tarpeeksi suurella etäisyydellä reitistä. Liian suurella puskurilla mukaan olisi voinut tulla myös pisteitä, joita nimenomaan oli tarkoitus suodattaa pois.

Seuraavaksi suodatin linjakohtaisesti virheellisen sijainnin omaavat pisteet pois käsiteltävästä aineistosta puskurien avulla. Suodatusprosessi on kuvattu alla käyttäen esimerkkinä runkolinjaa 8. Tämä prosessi toistettiin jokaiselle linjalle.

Valitsin puskuritasosta linjan reittiä mukailevan puskuriobjektin. Tämän jälkeen valitsin ajoneuvodatan tasosta kaikki pisteet, jotka leikkaavat tämän kyseisen reitin.

Oletusasetuksiin täytyi lisätä valinta ”käytetään vain valittuja kohteita”. Näin sain valittua kaikki linjan 8 reitin kanssa kohdakkain osuvat pisteet. Halusin kuitenkin suodattaa pois pisteitä, jotka ilmoittavat ajoneuvon linjaksi linjan 8, mutta joiden sijainti ei ole kyseisen linjan reitillä. Tätä varten valintatyökaluista täytyi valita ”Tee käänteinen kohteiden valinta”. Sitten avasin ajoneuvodatasta kerätyn pistemuotoisen aineiston attribuuttitaulun, ja valitsin työkalun ”select by expression”. Asetin valintalausekkeeksi ”*lineref*” = ’8’. Sitten valitsin ”Suodata nykyistä valintaa”. Näin valituksi tulivat linjan 8 reitin ulkopuolella olevista pisteistä ne, joissa linjatunnukseksi oli ilmoitettu linja 8. Linjalta 8 tällaisia pisteitä löytyi esimerkiksi linja-autoasemalta, missä autot olivat pysäköintialueella tauolla, vaikka ilmoittivat datassa olevansa linjalla. Lopuksi poistin nämä pisteet ajoneuvodatan tasolta.

Seuraavaksi rajasin kerätyistä aineistosta pisteitä pois *delay*-kentän arvoja hyödyntämällä. Valitsin rajauksen alarajaksi 15 minuutin etuajan (arvo -15,00) ja ylärajaksi 40 minuutin myöhästymisen (arvo 40,00). Tein tämän siksi, että tulkitsin näitä suurempia etuaikaa tai myöhästymistä omaavien pisteiden olevan käytännössä

virheellisiä ja kuvaavan lähinnä epärealistista liikennöintitilannetta. Tein rajauksen ”select by expression”-työkalulla käyttäen hakulauseketta `"delay" < '-15' OR "delay" > '40'`. Lopuksi poistin valitut rivit tiedostosta.

Seuraavaksi poistin päätepysäkkien lähistöiltä datapisteitä. Autot jäivät usein päätepysäkeille tai näitä edeltäville pysäkeille odottamaan lähtöä seuraavalle vuorolle, jolloin myöhästymisarvo kasvaa pysäkille saapumisen jälkeen. Lisäksi autot kirjautuvat usein vuorolle myös etuajassa, jolloin etuaikaa esittäviä datapisteitä syntyy jo ennen lähtöä. Näiden pisteiden poistaminen oli käytännössä käsityönä ja karttatarkasteluna tehtävää työtä: tarkastelin linja kerrallaan pisteitä päätepysäkkien ympäristössä, ja poistin niitä sellaisilta alueilta, joihin virheellisiä pisteitä kasautui paljon. Hyvä esimerkki oli Uittamon päätepysäkki, joka on neljän runkolinjan päätepysäkki: autoja on pysäkillä käytännössä aina odottamassa seuraavaa lähtöä, aiheuttaen dataan joko myöhästymisiä tai etuajassa olevia datapisteitä.

Linjojen 2 ja 4 kanssa tämä toimenpide vaati hieman enemmän työtä, sillä näiden linjojen päätepysäkkien sijainti on muiden runkolinjojen reittien varrella. Käytännössä näissä piti suodattaa kartalta tehtävää pisteiden valintaa vielä käyttäen linjatunnusta ja ”select by expression”-työkalua.

Näiden rajausten jälkeen aineistossa oli edelleen virheellisiä datapisteitä. Nämä olivat sellaisia, joissa auton sijainti oli ilmoittamansa reitin varrella, mutta sijaintitieto ja tieto seuraavasta pysäkistä olivat ristiriidassa keskenään. Tällaisissa tapauksissa ajoneuvojen ilmoittamat myöhästymistiedot olivat myös selvästi aina virheellisiä. Esimerkiksi auto tunnuksella 220942 ilmoitti seuraavaksi pysäkikseen lähes kaikissa datapisteissä ainoastaan linjan ensimmäisen pysäkin, vaikka auton sijaintitieto kertoi sen olevan aivan muualla. Arvioin tällaisten autojen kohdalla tiedot niin epäluotettaviksi, että autojen pisteet oli syytä poistaa analyysissä käytettävästä aineistosta kokonaan.

Kävin aineistoa läpi tarkastelemalla sen attribuuttitaulua ja pisteitä kartalla. Järjestin datan taulussa suurimmasta myöhästymisarvosta pienimpään ja kävin suuria myöhästymisarvoja läpi. Tarkastin yksi kerrallaan kaikki datapisteet, joissa ajoneuvojen ilmoitettu myöhästymistieto oli yli kymmenen minuuttia. Otin talteen

ajoneuvotunnuksia, jotta pystyin tarkastelemaan dataa myöhemmin ajoneuvokohtaisesti. Vertasin autojen ilmoittamaa tietoa seuraavasta pysäkistä niiden sijaintiin. Autoja, jotka ilmoittivat toistuvasti keskenään ristiriitaista tietoa seuraavasta pysäkistä ja sijainnista, löytyi tässä tarkastelussa noin kymmenen kappaletta.

Seuraavaksi tarkastelin virheellisiä tietoja aiheuttaneiden ajoneuvojen kaikkia tallentuneita datapisteitä. Tätä tarkastelua helpotti QGIS:ssä seuraavan pysäkin nimen asettamista karttatason nimiöiksi. Näin jokaisen pisteen kohdalle kartalle tuli näkyville pysäkkitiedon nimi, ja pystyin vertaamaan sijaintia tähän pysäkkiin. Havaitsin neljän auton kohdalla pisteiden olevan suurimmaksi osaksi tai kokonaan virheellisiä. Poistin näiden ajoneuvojen pisteet datasta kokonaan.

Tehtyäni linja- ja ajoneuvokohtaiset tarkastelut ja virheiden poistot, sain tuloksena aineiston, jonka paikkansapitävyys on vähintään riittävällä tasolla. Tätä aineistoa pystyin analysoimaan tarkemmin: kopioin kaikkien jäljellä olevien datapisteiden tiedot Exceliin ja laskin aineiston perusteella linjakohtaisesti etuajassa kulkevien, ajallaan kulkevien ja myöhässä kulkevien havaintopisteiden määrät, keskiarvot ja osuudet. Lisäksi tarkastelin etuajassa ja myöhässä olevia pisteitä ajosuunnan mukaan.

4 Tulokset

4.1 Fölin runkolinjojen täsmällisyyden arviointi yleisesti

Fölin runkolinjojen liikennöinti vaikuttaa olevan pääosin ainakin kohtuullisen täsmällistä, datapisteistä yli viiden minuutin myöhästymistiedon omaavia oli alle kolme prosenttia (taulukko 1). Aikataulussa ja enintään viisi minuuttia myöhässä olevia pisteitä oli noin 56 prosenttia. Aikataulua edellä ilmoittavaa liikennöintiä oli kuitenkin yllättävän suuri osa, noin 41 prosenttia pisteistä.

Taulukko 1: Yli viiden minuutin myöhästymisten, 0–5 minuutin välillä olevien myöhästymisten ja etuajassa olevien pisteiden osuudet linjoittain. Taulukko on järjestetty yli viiden minuutin myöhästymisten osuuksien mukaan pienimmästä suurimpaan.

<i>Linja</i>	<i>yli 5 minuuttia myöhässä olevien pisteiden osuus</i>	<i>0–5 minuuttia myöhässä olevien pisteiden osuus</i>	<i>Etujassa olevien pisteiden osuus (alle 0 min)</i>
2Y	0,0 %	45,8 %	54,2 %
6T	0,0 %	66,7 %	33,3 %
5	1,4 %	57,4 %	41,3 %
10	1,5 %	46,6 %	51,9 %
9	1,6 %	56,8 %	41,6 %
8	1,6 %	59,8 %	38,6 %
9A	1,8 %	57,6 %	40,6 %
5A	2,1 %	45,0 %	52,9 %
1	2,5 %	72,5 %	25,0 %
3	2,7 %	47,8 %	49,6 %
10A	2,9 %	44,9 %	52,2 %
6	3,0 %	59,7 %	37,2 %
4	3,2 %	59,8 %	37,0 %
2	3,6 %	59,7 %	36,6 %
6A	4,6 %	62,3 %	33,0 %
3A	6,1 %	48,5 %	45,3 %
1A	6,5 %	81,3 %	12,2 %
<i>kaikki</i>	2,7 %	56,2 %	41,1 %

Huomionarvoista on, että etuaikaa ilmoittavien pisteiden etuaikatiedon keskiarvo on 1,56 minuuttia, eli keskimäärin etuajassa kulkeminen ei ole kovin suurta (taulukko 2). Vastaavasti kaikkien myöhästymisten (myöhästymistieto yli 0,00 minuuttia) keskiarvo

oli 1,76 minuuttia. Kaikkien pisteiden keskiarvoinen myöhästyminen oli 0,26 minuuttia eli hieman päälle 15 sekuntia.

Taulukko 2: Runkolinjojen myöhässä olevien, kaikkien ja etuajassa olevien datapisteiden myöhästymistiedon keskiarvot minuutteina linjoittain. Järjestetty myöhässä olevien pisteiden keskiarvojen mukaan.

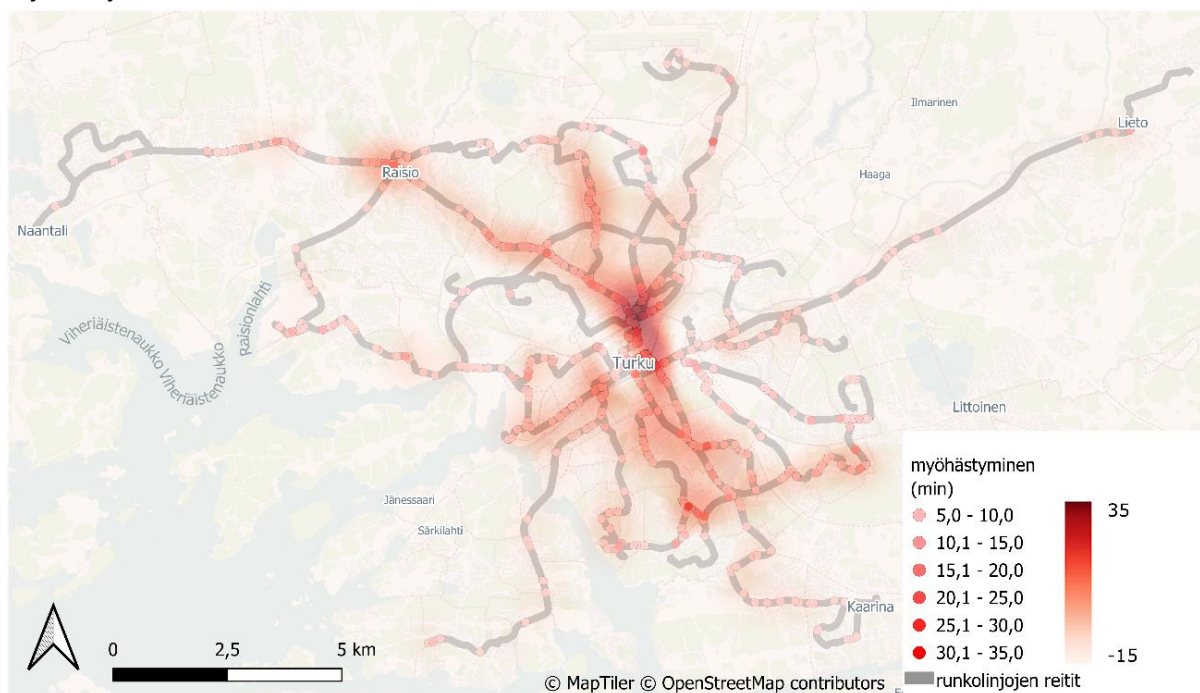
<i>Linja</i>	<i>Myöhässä (yli 0 min) olevien keskiarvo (min)</i>	<i>Kaikkien pisteiden keskiarvo (min)</i>	<i>Etujassa (alle 0 min) olevien keskiarvo (min)</i>
6T	0,30	0,03	-0,20
2Y	0,61	-0,40	-0,97
10	1,32	-0,08	-1,23
5	1,33	0,16	-1,08
9	1,45	0,15	-1,17
8	1,59	0,11	-1,90
9A	1,65	0,28	-1,17
1	1,66	0,80	-1,04
10A	1,66	0,11	-1,16
5A	1,74	-0,01	-1,49
6	1,85	0,30	-2,05
4	1,86	0,60	-1,26
3	1,87	-0,28	-2,25
2	1,93	0,69	-1,20
6A	2,09	0,61	-2,13
1A	2,16	1,07	-0,62
3A	2,76	0,29	-2,27
<i>kaikki</i>	1,76	0,26	-1,56

Sekä myöhästymispisteitä että etuajassa kulkevia pisteitä oli kaikkien reittien varrella pääosin tasaisesti (kuva 3). Suurin osa merkittävän myöhästymisen pisteistä sijoittuu Turun keskusta-alueen tuntumaan. Myöhästymisten ja etuajassa kulkemisten osuudet kuitenkin vaihtelevat jonkin verran eri linjojen välillä. Myös linjan ajosuunnalla vaikuttaa olevan jonkin verran vaikutusta täsmällisyyteen.

Suhteellisesti eniten myöhästymistä vaikuttaa olevan linjoilla 1A, 3A, 6A, 2, 4 ja 6 (taulukko 1). Näiden linjojen myöhästymisen keskiarvot ovat myös suurimmat.

Vastaavasti kaikista vähiten myöhästymisiä on linjoilla 2Y ja 6T, näiden linjojen kohdalle ei ole tallentunut yhtään datapistettä, joissa myöhästyminen olisi yli viisi minuuttia.

Kuva 3: Viisi minuuttia tai enemmän myöhässä olevat pisteet kartalla ja lämpökartta-visualisointi myöhästymisistä.



4.2 Linjakohtaiset tulokset

Linjalla 1 täsmällisyys on kaikista runkolinjoista toiseksi parhaalla tasolla, sillä 0–5 minuutin myöhästymisen omaavia pisteitä on kaikista linjan pisteistä yli 72 prosenttia. Etuajassa olevia pisteitä on 25 prosenttia ja yli viiden minuutin myöhästymisiä on vain 2,5 prosentilla pisteistä. Linjan keskiarvoinen myöhästymisen on noin 40 sekuntia ja keskiarvoinen etuaika noin minuutin. Ajosuunnalla ei ole juurikaan vaikutusta linjan täsmällisyyteen.

Linja 1A on sataman ja linja-autoaseman välillä kulkeva ruuhkalinja. Linjan suurien (yli viiden minuutin) myöhästymisten osuus on kaikista runkolinjoista suurin. Kuitenkin yli 81 prosenttia linjalta kerätyistä pisteistä ilmoittaa myöhästymistiedoksi 0–5 minuuttia. Etuajassa olevia pisteitä on hieman päälle 12 prosenttia. Keskiarvoinen myöhästymisen on kuitenkin melko maltillisella tasolla, hieman päälle kaksi minuuttia. Tälläkään linjalla ajosuunnalla ei ole merkitystä täsmällisyyden kannalta.

Linja 2 on myös kohtuullisen täsmällinen linja, vaikka sen merkittävien myöhästymisten osuus onkin yli kaikkien linjojen keskiarvon. Kohmon suuntaan myöhässä olevia pisteitä on määrällisesti enemmän kuin Myllyn suuntaan (taulukko 3). Vastaavasti Myllyn suuntaan on enemmän etuajassa olevia pisteitä kuin Kohmon suuntaan. Kumpaankin

Taulukko 3: Linjojen etuajassa ja myöhässä olevien datapisteiden määrät jaoteltuna ajosuunnan mukaan. Taulukossa ei ole mukana pisteitä, joiden myöhästymisarvo on nolla minuuttia.

<i>Linja</i>	<i>Suunta</i>	<i>Enemmän pisteitä</i>	<i>Pisteitä myöhässä (yli 0 min)</i>	<i>Pisteitä etuajassa (alle 0 min)</i>
1	Lentoasema	myöhässä	527	211
1	Satama	myöhässä	500	192
1A	Linja-autoasema	myöhässä	31	6
1A	Satama	myöhässä	34	9
2	Kohmo	myöhässä	727	379
2	Mylly	myöhässä	591	445
2Y	Kohmo	etuajassa	1	8
2Y	Länsinummi	etuajassa	4	5
3	Perno	myöhässä	722	642
3	Varissuo	etuajassa	542	773
3A	Raisio	myöhässä	196	157
3A	Varissuo	etuajassa	161	182
4	Halinen	myöhässä	575	370
4	Varissuo	myöhässä	591	385
5	Haarla	myöhässä	406	389
5	Ylioppilaskylä	myöhässä	466	393
5A	Haarla	etuajassa	46	60
5A	Keskusta	etuajassa	43	68
6	Lieto	myöhässä	530	271
6	Naantali	myöhässä	472	380
6A	Kaarina	myöhässä	591	306
6A	Naantali	myöhässä	585	313
6T	Turku	-	1	1
8	Harittu	myöhässä	711	520
8	Runosmäki	myöhässä	678	490
9	Härkämäki	etuajassa	157	192
9	Uittamo	myöhässä	204	152
9A	Suikkila	myöhässä	176	160
9A	Uittamo	myöhässä	174	150
10	Satama	etuajassa	402	535
10	Uittamo	etuajassa	415	469
10A	Perno	etuajassa	457	556
10A	Uittamo	etuajassa	479	575

suuntaan enemmistö pisteistä on myöhässä. Myöhästymiset ovat kuitenkin keskimäärin alle kaksi minuuttia. 0–5 minuutin myöhästymisen pisteitä on lähes 60 prosenttia kaikista linjan pisteistä.

Linja 3 on keräämäni aineiston perusteella hyvin keskimääräinen linja Fölin runkolinjastossa. Myöhästymiset ovat jakautuneet melko tasaisesti koko reitille ja merkittävien myöhästymisten keskiarvo on sama kuin kaikkien linjojen vastaava keskiarvo, eli 2,7 prosenttia. Etuajassa olevien pisteiden osuus on kuitenkin suurempi kuin 0–5 minuuttia myöhässä olevien pisteiden osuus. Myöhästymisiä on enemmän Pernon suuntaan kuin Varissuon suuntaan ja etuajassa olevia pisteitä on enemmän Varissuon suuntaan kuin Pernon suuntaan. Suurin osa Pernon ajosuunnan pisteistä on myöhässä, kun taas Varissuon suuntaan pisteitä on enemmän etuajassa. Linjan 3 etuajassa kulkevien pisteiden myöhästymisarvon keskiarvo on kaikista linjoista toiseksi pienin. Eli linjalla on etuaikaa keskimääräisesti toiseksi eniten.

Vastaavasti suuria myöhästymisiä vaikuttaa olevan linjalla 3A enemmän Raision suuntaan kuin Varissuon suuntaan. Raision suuntaan myös suurin osa pisteistä on myöhässä, Varissuon suuntaan taas etuajassa. Tällä linjalla etuajassa olevien pisteiden keskiarvo on kaikista pienin eli linja on keskimäärin kaikista linjoista eniten etuajassa. Toisaalta linjan keskiarvoinen myöhästyminen on kaikista linjoista suurin, lähes kolme minuuttia, ja linjan merkittävien myöhästymisten osuus on kaikista linjoista toiseksi suurin, hieman päälle kuusi prosenttia.

Linjan 4 myöhästymiset ovat jakautuneet reitin varrelle tasaisesti, mutta ne ovat jonkin verran suurempia kuin enemmistöllä linjoista. Linjan myöhästymisten keskiarvo on kaikkien linjojen kuudenneksi suurin ja linjan merkittävien myöhästymisten osuus on suurempi kuin kaikkien linjojen vastaava osuus. Ajosuunnalla ei ole juurikaan vaikutusta linjan täsmällisyyteen.

Linja 5 on täsmällisyydeltään runkolinjojen parhaimmistoa. Yli viiden minuutin myöhästymisten osuus on linjalla vain 1,4 %, joka on ainoastaan yö- ja ilta-aikoina liikennöiviä linjoja lukuun ottamatta paras tulos. Keskimääräiset myöhästymiset ja etuajassa kulkemiset ovat vain hieman yli minuutin. Kaikkien linjan 5 pisteiden

myöhästymisarvojen keskiarvo on alle kymmenen sekuntia. Ajosuunnalla ei ole merkitystä linjan täsmällisyyteen.

Linjan 5A kaikkien pisteiden keskiarvo on hyvin lähellä nollaa, mutta hieman negatiivisen puolella. Myöhästymisten keskiarvo on alle kaksi minuuttia ja hyvin lähellä kaikkien linjojen keskiarvoa. Linjalla on kumpaankin ajosuuntaan enemmän etuajassa kuin myöhässä olevia pisteitä. Myös linjan etuajassa olevien pisteiden osuus on kaikkien linjojen vastaavaa keskiarvoa suurempi.

Linjojen 6 ja 6A välillä on yllättävän paljon eroa. Kaarinan suuntaan haarautuvalla 6A:lla myöhästymisiä on selvästi enemmän kuin linjan 6 Lietoon kulkevalla reitillä. Myös myöhästymisten keskiarvo on 6A:lla suurempi, kuin linjalla 6. Kummallakin linjalla kaikkiin ajosuuntiin olevista pisteistä enemmistö on myöhässä. Linjalla 6A sekä myöhässä että etuajassa olevia pisteitä on kuitenkin huomattavasti enemmän kuin linjalla 6. Kerätyn datan perusteella linjan 6T liikennöinti on erittäin täsmällistä, mutta havaintomääräkin on vain kolme pistettä. Etuajassa ja myöhässä olevia pisteitä on kumpaakin yksi ja keskimääräinen myöhästymisarvo linjalla on alle 2 sekuntia.

Linjan 8 täsmällisyys on hyvällä tasolla, keskimääräinen myöhästymisarvo linjalla on noin seitsemän sekuntia. Lähes 60 prosenttia pisteistä on 0–5 minuutin myöhästymisarvon aikaikkunan sisällä. Merkittäviä myöhästymisiä oli vain 1,6 prosenttia. Etuajassa kulkemisen osuus on kaikkien linjojen keskiarvoa pienempi. Myöhässä olevien pisteiden keskiarvo on hieman yli puolitoista minuuttia. Ajosuunnalla ei ole merkitystä linjan täsmällisyyteen.

Linjoilla 9 ja 9A myöhästymiset ovat suurempia Turun keskusta-alueella ja sen eteläpuolella, kuin keskustan pohjoispuolella. Kummallakin linjalla täsmällisyys on hyvällä tasolla, 0–5 minuutin myöhästymisiä on kummallakin linjalla yli 55 % linjan pisteistä. Myöhästymisiä on linjalla 9A hieman linjaa 9 enemmän, mutta yli viiden minuutin myöhästymisiä on tallentunut linjalla 9A vain yhden pisteen verran linjaa 9 enemmän (taulukko 4). Linjan 9 kaikkien pisteiden myöhästymisarvo on keskimäärin noin 10 sekuntia ja linjalla 9A keskimäärin hieman alle 30 sekuntia. Linjalla 9 Härkämäen suuntaan enemmistö pisteistä on etuajassa ja Uittamon suuntaan myöhässä. Linjalla 9A kummankin ajosuunnan pisteistä enemmistö on myöhässä.

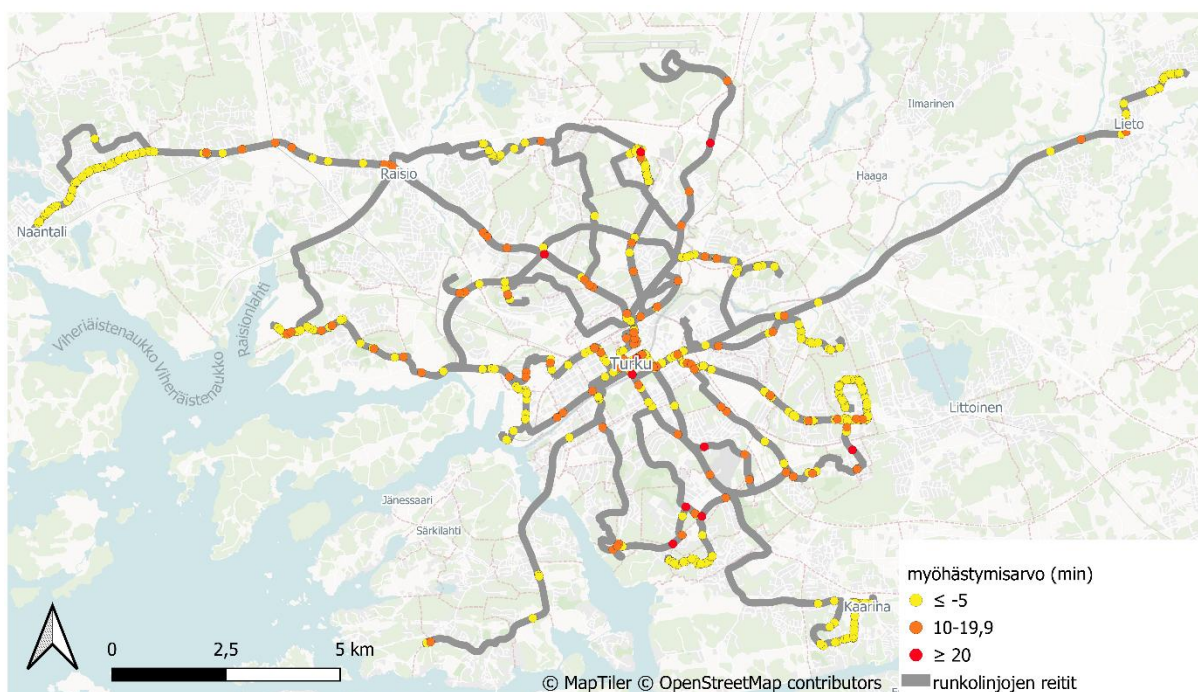
Taulukko 4: Kaikkien datapisteiden, yli 5 minuutin myöhästymistiedolla, 0–5 minuutin myöhästymistiedolla, ja alle 0 minuutin myöhästymistiedolla olevien datapisteiden lukumäärät linjoittain.

<i>Linja</i>	<i>Kaikkiaan pisteitä</i>	<i>pisteet yli 5min</i>	<i>pisteet 0–5 min</i>	<i>pisteet alle 0 min</i>
1	1612	40	1169	403
1A	123	8	100	15
2	2249	82	1343	824
2Y	24	0	11	13
3	2854	76	1363	1415
3A	748	46	363	339
4	2043	66	1222	755
5	1895	26	1087	782
5A	242	5	109	128
6	1748	53	1044	651
6A	1874	87	1168	619
6T	3	0	2	1
8	2617	42	1565	1010
9	827	13	470	344
9A	764	14	440	310
10	1933	29	900	1004
10A	2166	62	973	1131
yhteensä	23 722	649	13 329	9744

Linjalla 10A näyttää olevan suurempia myöhästymisiä kuin linjalla 10. Linjalla 10A myöhästymisiä on myös linjaa 10 enemmän. Keskiarvoinen myöhästymisen on linjalla 10 noin yksi minuutti ja 20 sekuntia, linjalla 10A vastaava arvo on noin 20 sekuntia suurempi. Linjalla 10 merkittävien myöhästymisten osuus on 1,5 prosenttia ja linjalla 10A 2,9 prosenttia. Etuajassa kulkevien vuorojen osuudet ovat kuitenkin lähes yhtä suuret kummallakin linjalla. Kummankin linjan kaikkiin ajosuuntiin enemmistö pisteistä on etuajassa. Ajosuunnalla ei ole vaikutusta täsmällisyyteen. Myöhästymispisteitä on kummallakin linjalla erityisesti Kupittaalla.

4.3 Ongelmalliset sijainnit

Analyysissa ei noussut esiin erityisiä ongelmapaikkoja runkolinjaston reiteillä, myöhästymisten jakautuminen on reiteillä pääosin tasaista. Raision ja Turun keskusta-alueilla myöhästymispisteitä on kuitenkin paljon (kuva 4). Maaseutumaisemilla ja vähemmän tiiviisti rakennetuilla alueilla, kuten esimerkiksi Hämeentiellä Turun ja Liedon välillä, myöhästymispisteitä on vähemmän. Linjojen 6 ja 6A päissä, eli Naantalissa, Liedossa ja Kaarinassa etuajassa olevia pisteitä on paljon. Myös linjan 8 kummassakin päässä, eli Runosmäessä ja Haritussa, etuajassa olevia pisteitä on useita.



Kuva 4: Runkolinjojen viisi minuuttia tai enemmän etuajassa olevat ja kymmenen minuuttia tai enemmän myöhässä olevat pisteet.

5 Tulosten tarkastelu

5.1 Liikenteen täsmällisyyteen vaikuttavia tekijöitä

Linja-autoliikenteen täsmällisyyteen voivat vaikuttaa monet tekijät, kuten esimerkiksi muu liikenneympäristö, linjojen liikennöintiajat, linjojen reitit, suunnitellut aikataulut ja matkustajamäärät (Gershenson & Pineda 2009; Yetiskul & Senbil 2012; Park ym. 2020: 367–370). Turun seudulla suurimmat tekijät ovat luultavasti muu liikenne sekä liikennöintiajat. Lisäksi linjojen pysähtymistiheys esimerkiksi liikennevaloihin ja pysäkeille vaikuttavat niiden etenemiseen.

Ruuhka-aikaan liikennöivät linjat 1A, 3A ja 5A ovat erityisen alttiita muun liikenteen aiheuttamalle häiriölle. Lisäksi ruuhka-aikoina matkustajamäärät ovat suuremmat, jolloin pysäkeille pysähtytään useammin sekä matkustajien autoon ja autosta nousuun kuluu enemmän aikaa. Toisaalta näiden linjojen tarkoitus on lisätä jo tiheän liikenteen kapasiteettia, jolloin esimerkiksi myöhästymisen merkitys ei ole niin suuri, mutta tietysti epätäsmällisyydestä on silti haittaa. Erityisesti linjalla 1A muulla liikenteellä on luultavasti suuri vaikutus, sillä Turun satama-alue on laivojen ollessa satamassa ruuhkainen (WSP Finland Oy 2020: 20). Tämä korostuu entisestään tutkielman tarkastelujaksolla sataman lähellä olevan työmaan vaikutuksesta (Collin 2025).

Tuloksia tarkastellessa kävi ilmi, että usein linjoilla, joilla myöhästymiset ovat yleisiä tai suuria, myös etuajassa kulkevaa liikennettä on paljon. Fölin aikatauluissa ajoajat ovat useimmiten kaikkina kellonaikoina samat (Turun seudun joukkoliikenteen aikataulut 2025). Tällaisissa tapauksissa linja-autot voivat olla vähemmän ruuhkaisina aikoina helpommin etuajassa ja vastaavasti ruuhkaisina aikoina ne jäävät helpommin aikataulusta jälkeen. Tilannetta voisi helpottaa muuttamalla linjojen ajoaikoja kellonajan mukaan.

Yö- ja ilta-aikaan sekä aikaisin aamulla liikennöivät linjat 2Y ja 6T olivat data-aineiston perusteella usein etuajassa. Tätä varmasti selittää myös se, että näihin vuorokaudenaikoihin muu liikenne on vähäistä ja matkustajamäärät kohtuullisen pienet. Lisäksi vuoroja on määrällisesti muita linjoja vähemmän, jolloin myöhästymisen

riski on pienempi. Toisaalta, jos vuoroväli on esimerkiksi yölinjalla harvempi, tällöin täsmällisyys on tärkeämpää kuin päivisin tiheään liikennöidyllä linjalla.

Linjojen 1, 5 ja 8 sujuvaan liikennöintiin vaikuttavat todennäköisesti erityisesti linjojen reitit. Näiden linjojen reitit ovat hyvin suoraviivaisia, eikä niiden varrella ole juurikaan Turun keskusta-alueen lisäksi paikkoja, joissa eteneminen olisi erityisen hidasta. Turun keskusta-alueen myöhästymisten tarkastelussa on syytä ottaa huomioon, että myöhästymispisteitä on keskustassa paljon erityisesti siksi, että suurin osa runkolinjoista kulkee keskusta-alueen läpi (kuva 3).

Linjalla 6A reitti ja matkustajamäärät sekä muun joukkoliikenteen reitit vaikuttavat luultavasti suurempaan myöhästymisten osuuteen. Myöhästyminen saattaa aiheutua erityisesti Kaarinan keskusta-alueella sekä Kaarinan ja Piispanristin välillä.

Nopeusrajoitus on täällä verrattain alhainen ja alueelta ei ole muita joukkoliikenneyhteyksiä Turun suuntaan. Kaarinaan kyllä kulkee muitakin linjoja, mutta ne jäävät linjan 6A reitistä osittain kauemmas. Linjan 6A ongelmia (entinen linja 7) on pohdittu jo vuonna 2021 tehdyssä selvityksessä (Hyvönen ym. 2021: 81).

Linjan 10 ja 10A täsmällisyyden erot selittyvät luultavasti myös linjojen reiteillä. Fölin aikatauluissa linjan 10 ajoajaksi on ilmoitettu noin 75 minuuttia ja linjan 10A ajoajaksi noin 80 minuuttia (Turun seudun joukkoliikenteen aikataulut 2025). Linja 10A saattaisi vaatia ainakin joinakin kellonaikoina enemmän ajoaikaa. Myöhästymisiä saattavat aiheuttaa erityisesti Raision keskustan ja kauppakeskus Myllyn alueet. Näillä alueilla on ruuhka-aikaan paljon liikennettä, ja joukkoliikenne joutuu kulkemaan kaiken muun liikenteen seassa.

Maaseutumaisilla alueilla ja keskusta-alueiden ulkopuolella myöhästymispisteitä on hieman vähemmän. Hyvönen ym. (2021: 79) toivat saman ilmiön esille selvityksessään. Tämä johtuu luultavasti siitä, että näillä alueilla ruuhka-ajan merkitys on pienempi ja nopeusrajoitukset ovat korkeampia. Myöskään pysäkeillä ei välttämättä pysähdytä yhtä usein, ja autoon tai autosta nousevia matkustajia on vähemmän. Tämä saattaa osin selittää linjojen 6 ja 6A eroja: linja 6A kulkee Kaarinan keskusta-alueella pidemmän matkan ja alhaisemmat nopeusrajoitukset ulottuvat Kaarinassa kauas keskustan

ulkopuolelle. Liedossa keskusta-alueella ajoa on vähemmän ja nopeusrajoitukset pääosin korkeampia.

Liikenteen luotettavuus ja täsmällisyys sekä matkanopeus ovat matkustajan kannalta tärkeitä piirteitä joukkoliikenteelle (Cats ym. 2015; Abenoza ym. 2017). Ne vaikuttavat merkittävästi matkustajatytyväisyyteen ja siten myös joukkoliikenteen käyttöasteeseen. Diab ym. (2015: 298) toteavat tutkimuksessaan, että matkustajat kokevat pysäkeillä tapahtuvan odotusajan hukatuksi ajaksi. Siksi liikenteen täsmällisyyden kehittäminen on matkanopeuden kehittämistä tärkeämpää.

5.2 Vertailu aiempiin tutkimuksiin muissa kaupungeissa

Jiang ym. (2025) käyttivät tutkimuksessaan GTFS-RT dataa Gainesvillestä, Floridasta liikennevirtojen tarkastelussa. He onnistuivat hyödyntämään joukkoliikenteen reaaliaikaista dataa tieliikenteen etenemisen tarkastelussa. Joukkoliikennedatan avulla saavutetut tulokset heijastivat muilla menetelmillä kerätyn validointidatan tuloksia. Joukkoliikennedatalla ei kuitenkaan ole välttämättä saatavilla kaikilta alueilta tai kaikkina kellonaikoina. Tällä voi olla vaikutusta siihen, onko datan hyödyntäminen mielekästä vai ei. Tällainen liikennevirtojen etenemisen tutkimus voisi olla mahdollista joukkoliikenteen reaaliaikaisen datan avulla myös Turussa, sillä linja-autoliikenne kulkee hyvin pitkälti muun autoliikenteen seassa. Yksittäisiä bussikaistoja on ja linja-autoilla on paikoin käytössä liikennevaloetuuudet, mutta näiden vaikutukset joukkoliikenteen eduksi jäävät melko pieniksi.

Park ym. (2019) tutkivat linja-autoliikenteen myöhästymistä reaaliaikaisen GTFS-datan avulla Columbusissa, Ohiossa. Heidän tutkimustuloksissaan myöhästyminen keskittyi tietyille pysäkeille ja tiettyihin sijainteihin. Tämä eroaa saamistani tuloksista Turun seudulla, sillä oman aineistoni myöhästymiset olivat jakautuneet melko tasaisesti linjojen reiteille. Toisaalta heidän keräämänsä datan aikaresoluutio oli huomattavasti korkeampi ja datan keräys toteutettiin pidemmällä aikavälillä. Park ym. (2019) keräsivät aineistoa 20 viikonpäivän ja neljän viikonlopun ajalta. Keräämäni aineisto oli vain noin viikon ajalta, tällä aineistomäärällä pysäkkikohtainen analyysi ei ole mahdollista, eivätkä tulokset yhtä tarkkoja.

5.3 Tulosten epävarmuudet ja ongelmat

Paikkatietoaineiston kerääminen reaaliaikadatarajapinnasta ei ollut aivan ongelmatonta. Ongelmia aiheutui sekä dataa tallennettaessa että myöhemmin käsittelyvaiheessa. Tallennusvaiheessa ongelmat syntyivät lähinnä datan rakenteen epäsäännöllisyyksistä, esimerkiksi joidenkin ajoneuvojen kohdalta saattoi ajoittain puuttua jokin datakenttä. En osannut ottaa tätä aluksi ohjelmakoodini toteutuksessa huomioon, jolloin koko ohjelma kaatui tähän. Lisäksi keräämästäni datasta jäi joiltain päiviltä puuttumaan dataa aamuilta, sillä jokin virhe ohjelmani ja Fölin rajapinnan välillä aiheutti tallennuksen pysähtymisen. Epäilen tämän johtuvan esimerkiksi Fölin palvelimen uudelleenkäynnistymisestä, mutta tämä ei kuitenkaan tapahtunut säännöllisesti eikä esimerkiksi aina samaan kellonaikaan, joten en saanut oikein syytä tapahtuneelle selville.

Analyysivaiheessa ongelmia tuottivat datan toimintatavasta johtuvat virheelliset tiedot. Kun linja-auto saapui pääte pysäkille, sen ajama vuoro jäi usein dataan aktiiviseksi. Tämä aiheutti kyseisen ajoneuvon kohdalla myöhästymisarvon kasvamisen, vaikka tosiasiaassa ajoneuvo olisi saapunut pääte pysäkille aikataulussa. Joillain linjoilla ilmeni myös tapa jäädä pääte pysäkkiä edeltävälle pysäkille odottamaan paluulähtöä, tällöinkin ilmiö oli sama: myöhästymisarvot kasvoivat vaikkei oikeaa myöhästymistä tapahtunutkaan. Tästä syystä jouduin suodattamaan pääte pysäkeille ja paikoin niitä edeltäville pysäkeille tallentuneet pisteet analyysivaiheessa pois.

Lisäksi tietyiltä ajoneuvoilta tuleva data oli joko täysin tai suurimmaksi osaksi virheellistä, kun koordinaattien sijainti ja tieto seuraavasta pysäkistä olivat ristiriidassa keskenään. Tällöin myös myöhästymisarvo oli virheellinen. Nämä autot täytyi suodattaa kokonaan pois.

5.4 Jatkotutkimusideoita

Tässä tutkielmassa täsmällisyyden tarkastelu jää melko yleiselle tasolle, eikä kerätty data ole niin tarkkaa kuin se voisi olla. Liikenne vaikuttaa olevan täsmällisyyden kannalta hyvällä tasolla, mutta esimerkiksi tiettyyn linjaan keskittyvä tarkempi lisätarkastelu voisi olla hyödyllistä. Fölin ajoneuvoseurannan data päivittyy noin viiden

sekunnin välein, tässä tutkimuksessa dataa otettiin talteen 15 minuutin välein. Jos dataa kerättäisiin talteen esimerkiksi minuutin välein, voisi esimerkiksi aikataulusta jälkeen jääviä vuoroja analysoida paljon tarkemmin. Lisäksi dataa voisi kerätä pidemmältä ajanjaksolta, jotta datasta voisi tarkastella toistuvia ilmiöitä, ja esimerkiksi harvoin liikennöidyiltä linjoilta kertyisi enemmän tietoa.

Fölin matkustajamääristä on saatavilla matkakorteilla tehtyihin kyytiin nousuihin perustuvaa dataa (Turun seudun joukkoliikenteen käyttäjämääriä 2025). Dataa ei kuitenkaan ole vielä tarjolla uuden runkolinjaston liikennöinnin ajalta. Mikäli tällaista dataa julkaistaan myös tulevaisuudessa, sen avulla voisi selvittää matkustajamäärien vaikutusta liikennöinnin täsmällisyyteen.

Runkolinjat eivät yksinään kata koko liikennöintialueen tarpeita. Runkolinjat ovat joukkoliikennejärjestelmän perusta, mutta niitä täydentäviä linjoja on paljon. Myös näiden linjojen on oltava täsmällisiä, jotta vaihtoyhteydet runko- ja täydentävien yhteyksien välillä toteutuvat. Tästä syystä myös muiden kuin runkolinjojen täsmällisyyden tutkimus olisi Turun seudulla tärkeää.

Lähteet

- Abenoza, R. F., Cats, O. & Susilo, Y. O. (2017) Travel satisfaction with public transport: Determinants, user classes, regional disparities and their evolution. *Transport Research. Part A* 95 64–84. <http://dx.doi.org/10.1016/j.tra.2016.11.011>
- About - General Transit Feed Specification (2026) <https://gtfs.org/about/> 31.3.2026.
- About – Transmodel (2026) <https://transmodel-cen.eu/index.php/about/> 1.4.2026.
- Avoim data | HSL | HSL.fi (2026) <https://www.hsl.fi/hsl/avoim-data> 1.4.2026.
- Avoim data | Vilkku (2026) <https://vilkku.kuopio.fi/ohjeita-tietoja/avoim-data> 1.4.2026.
- Berggren, U., D’Agostino, C., Svensson, H. & Brundell-Freij, K. (2022) Intrapersonal variability in public transport path choice due to changes in service reliability. *Transportation* 49(1) 1517–1547. <https://doi.org/10.1007/s11116-021-10218-z>
- Canitez, F. (2019) Urban public transport systems from new institutional economics perspective: a literature review. *Transport Reviews* 39(4) 511–530. <https://doi.org/10.1080/01441647.2018.1552631>
- Cats, O., Abenoza, R. F., Liu, C. & Susilo, Y. O. (2015) Evolution of satisfaction with public transport and its determinants in Sweden: Identifying priority areas. *Transportation Research Record* 2538(1) 86–95. <https://doi.org/10.3141/2538-10>
- Chakrabarti, S. (2015) The demand for reliable transit service: New evidence using stop level data from the Los Angeles Metro bus system. *Journal of Transport Geography* 48(1) 154–164. <https://doi.org/10.1016/j.jtrangeo.2015.09.006>
- Collin, P. (2025) Turun linnan ympärillä alkaa lähes vuoden kestävä tietyö – sataman katujen uudistus katkaisee Linnankadun. Yle Uutiset. 8.9.2025. <https://yle.fi/a/74-20180972> 6.4.2026.
- Diab, E. I., Badami, M. G. & El-Geneidy, A. M. (2015) Bus transit service reliability and improvement strategies: Integrating the perspectives of passengers and transit agencies in North America. *Transport Reviews* 35(3) 292–328. <https://doi.org/10.1080/01441647.2015.1005034>
- Documentation | Waltti GTFS Realtime (2026) <https://opendata.waltti.fi/docs> 1.4.2026.
- Entur AS (2025) Data Portal Real-Time. https://data.entur.no/public/datasets/realtime_siri_et 1.4.2026.

- Fintraffic (2026) Valtakunnallisen liikenteen koontipalvelun tietoaaineistot – GTFS-aineiston tietomäärät. <https://mobility.mobility-database.fintraffic.fi/fi> 4.4.2026.
- Fölin reittiopas (2026) Reittiehdotuksien asetukset. <https://reittiopas.foli.fi/> 4.4.2026.
- Gershenson, C. & Pineda, L. A. (2009) Why does public transport not arrive on time? The pervasiveness of equal headway instability. *PloS one* 4(10) e7292.
<https://doi.org/10.1371/journal.pone.0007292>
- Gössling, S. (2020) Why cities need to take road space from cars - and how this could be done. *Journal of Urban Design* 25(4) 443–448.
<https://doi.org/10.1080/13574809.2020.1727318>
- History | MobilityData (2026) <https://mobilitydata.org/history/> 31.3.2026.
- Hrelja, R., Khan, J., & Pettersson F. (2020) How to create functioning collaboration in theory and in practice – practical experiences of collaboration when planning public transport systems. *International Journal of Sustainable Transportation* 14(1) 1–13. <https://doi.org/10.1080/15568318.2018.1517842>
- Hyvönen, A., Lintusaari, M., Sisto, T., Haapakorva, P. & Rokka, H. (2021) *Turun seudun joukkoliikenteen vaihtopaikat, liityntä- ja solmupysäkit, sekä liikennevaloetuuudet -selvitys*. Sitowise Oy, Espoo. https://varsinais-suomi.fi/wp-content/uploads/2025/10/TuSePy_Turun_seudun_joukkoliikenteen_vaihtopaikat_2021_taukamuistio.pdf
- Jiang, S., Sun, Y., Wong, W., Xu, Y. & Zhao, X. (2025) Real-time urban traffic monitoring using transit buses as probes. *Transportation Research Record* 2679(2) 219–236.
<https://doi.org/10.1177/03611981241260708>
- Lehtonen, E., Helminen, V., Mesimäki, J., Tuominen, A. & Wallius E. (2024) Liikenteen uusien teknologioiden ja palveluiden vaikutusmekanismi. Liikennejärjestelmä sekä alue- ja yhdyskuntarakenne murroksessa. *Valtioneuvoston selvitys- ja tutkimustoiminnan julkaisusarja*. 2024:36. Valtioneuvoston kanslia, Helsinki.
<https://julkaisut.valtioneuvosto.fi/server/api/core/bitstreams/3d77569c-c8ae-4e6b-a0c7-79b5c0b5f73f/content>
- Lock, O., Bednarz, T. & Pettit, C. (2020) The visual analytics of big, open public transport data – a framework and pipeline for monitoring system performance in Greater Sydney. *Big Earth Data* 5(1) 134–159.
<https://doi.org/10.1080/20964471.2020.1758537>

- Lodi, A., Malaguti, E., Stier-Moses, Nicolás E. & Bonino, T. (2016) Design and control of public-service contracts and an application to public transport systems. *Management Science* 62(4) 1165–1187.
<http://dx.doi.org/10.1287/mnsc.2015.2174>
- Lyu, S., Huang, Y. & Sun, T. (2025) Urban sprawl, public transportation efficiency and carbon emissions. *Journal of cleaner production* 489(1) 144652.
<https://doi.org/10.1016/j.jclepro.2025.144652>
- Mutiganda, J. C., Skoog, M. & Igudia, E. (2023) Understanding the process of improving accessibility and affordability of inter-municipal bus transport system. *European Transport Research Review* 15(1) 24–33. <https://doi.org/10.1186/s12544-023-00598-x>
- NeTEx – Transmodel (2026) <https://transmodel-cen.eu/index.php/netex/> 1.4.2026.
- NeTEx Regional | Trafiklab (2025) 22.10.2025 <https://www.trafiklab.se/api/netex-datasets/netex-regional/> 1.4.2026.
- Overview - General Transit Feed Specification (2026) <https://gtfs.org/documentation/overview/> 31.3.2026.
- Park, Y., Mount, J. Liu, L., Xiao, N. & Miller H. J. (2020) Assessing public transit performance using real-time data: spatiotemporal patterns of bus operation delays in Columbus, Ohio, USA. *International Journal of Geographical Information Science* 34(2) 367–392.
<https://doi.org/10.1080/13658816.2019.1608997>
- Rosenberg, M, Rajamäki, R., Järvi, T. & VTT (2007) Suurten kaupunkiseutujen joukkoliikenteen kilpailukykyinen palvelutaso. Liikenne- ja viestintäministeriön julkaisuja 55/2007. Liikenne- ja viestintäministeriö, Helsinki.
<https://urn.fi/URN:NBN:fi-fe2016110428167>
- Runkolinjasto | Föli (2026) <https://www.foli.fi/fi/runkolinjasto> 4.4.2026.
- SIRI – Transmodel (2026) <https://transmodel-cen.eu/index.php/siri/> 12.3.2026.
- Standards for implementation – Transmodel 2026 <https://transmodel-cen.eu/index.php/standards-for-implementation/> 1.4.2026.
- Tenkanen H. (2016) GeoPython - AutoGIS. Helsingin Yliopisto. <https://automating-gis-processes.github.io/2016/index.html> 12.3.2026.

- The Math Lab (2025) API Requests & JSON Parsing in Python: A Guide in Data Collection. Medium 2.4.2025 <https://medium.com/@themathlab/api-requests-json-parsing-in-python-a-guide-in-data-collection-31e985981ea3> 12.3.2026.
- Tietoa Fölistä (2026) <https://www.foli.fi/fi/etsitk%C3%B6n%C3%A4it%C3%A4/tietoa-f%C3%B6list%C3%A4> 23.3.2026.
- TSJL - API (2026) <https://data.foli.fi/> 1.4.2026.
- Turun kaupunki (2026). Ilmastosuunnitelma 2029. Turun kaupungin kestävä ilmasto- ja energiatoimintasuunnitelma 2029. Kaupunginvaltuusto 16.5.2022 § 103. ISSN 2343-0710 <https://www.turku.fi/sites/default/files/document/turun-ilmastosuunnitelma-2029.pdf> 7.4.2026.
- Turun kaupunkiseudun joukkoliikennelautakunta (2026a) Pöytäkirja 01.04.2026 54 § Nivelkalustokokeilu Turun Kaupunkiliikenne Oy:n kanssa. 1.4.2026. <https://paatokset.turku.fi/turun-kaupunkiseudun-joukkoliikennelautakunta/poytakirja/5378511/pykala/54/> 2.4.2026.
- Turun kaupunkiseudun joukkoliikennelautakunta (2026b) Pöytäkirja 11.02.2026 16 § Kokonaismatkustajakapasiteetin kasvattaminen nivelkalustokokeilulla linjalla 3. 11.2.2026. <https://paatokset.turku.fi/turun-kaupunkiseudun-joukkoliikennelautakunta/poytakirja/5349983/pykala/16/> 2.4.2026.
- Turun kaupunkiseudun joukkoliikennelautakunta (2026c) Pöytäkirja 11.03.2026 37 § Turun seudun joukkoliikenteen peruskauden 2026 - 2027 aikataulut. <https://paatokset.turku.fi/turun-kaupunkiseudun-joukkoliikennelautakunta/poytakirja/5366442/pykala/37/> 4.4.2026.
- Turun seudun joukkoliikenteen aikataulut (2025) Aikataulu 1.7.2025 – 31.5.2026. Turun seudun joukkoliikenne, Föli.
- Turun seudun joukkoliikenteen käyttäjämääriä (2025) Suomi.fi-avointa. <https://avointa.suomi.fi/data/fi/dataset/turun-seudun-joukkoliikenteen-kayttajamaaria> 7.4.2026.
- Wotek, M., Hebel, K., Birr, K. & Nozari, H. (2025) The importance of punctuality in particular segments of public transport users: a case study from Gdynia (Poland). *Scientific Journal of Silesian University of Technology. Series Transport*. 127(1) 277–289. <https://doi.org/10.20858/sjsutst.2025.127.17>

WSP Finland Oy (2020) *Turun satama-alueen liikenteen ja liikkumisen yleissuunnitelma*

2020. WSP Finland Oy. <https://www.hupmobile-project.eu/sites/hupmobile/files/tulevif.pdf> 6.4.2026.

Yetiskul, E. & Senbil, M (2012) Public bus transit travel-time variability in Ankara (Turkey). *Transport policy* 23(1) 50–59.

<http://dx.doi.org/10.1016/j.tranpol.2012.05.008>

Liitteet

Liite 1. Python-ohjelmakoodi datan keräämiseksi ja tallentamiseksi

Ohjelmassa käytetyt kirjastot: requests, geopandas, shapely, fiona, schedule

```
import requests
import time
from datetime import datetime, date
import geopandas as gpd
from shapely.geometry import Point
from fiona.crs import from_epsg
import schedule

url = "https://data.foli.fi/siri/vm"

selected_data = []
timestamp = ""

# Funktio, joka hakee datan Fölin palvelimelta, käy datan läpi ajoneuvo
# kerrallaan ja tallentaa haluttujen
# ajoneuvojen tiedot välimuistiin, ja varmuuskopiona jokaisen datan hakukerran
# tiedot myös tekstitiedostona.
def fetchData():
    vehicles = {}

    for _ in range(5):
        response = requests.get(url)
        if response.status_code == 200:
            data = response.json()
            vehicles = data["result"]["vehicles"]
            responseTime = data["result"]["responsetimestamp"]
            timestamp = datetime.fromtimestamp(responseTime).strftime("%Y-%m-%dT%H_%M")
            break
        else:
            print(f"Virhe: {response.status_code}, {response.text}")
            print("Yritetään uudelleen...")
            time.sleep(5)

    for veh in vehicles:
        v = vehicles[veh]
        vehicleref = v["vehicleref"]
        longitude = v.get("longitude", 0)
        latitude = v.get("latitude", 0)

        if longitude != 0 and latitude != 0:
```

```

    vehicle = {}
    vehicle["vehicleref"] = vehicleref
    vehicle["longitude"] = longitude
    vehicle["latitude"] = latitude
    vehicle["lineref"] = v.get("lineref", "unknown line")
    vehicle["origin"] = v.get("originname", "unknown origin")
    vehicle["destination"] = v.get("next_destinationdisplay", "unknown
destination")
    vehicle["origintime"] = v.get("originaimeddeparturetime", 0)
    vehicle["desttime"] = v.get("destinationaimedarrivaltime", 0)
    vehicle["delaysecs"] = v.get("delaysecs", 0)
    vehicle["timestamp"] = timestamp
    vehicle["next_stoppointname"] = v.get("next_stoppointname", "")
    selected_data.append(vehicle)

    file_path = f'{timestamp}.txt'
    with open(file_path, "a", encoding="utf-8", newline="\n") as tiedosto:
        for vehicle in selected_data:
            datastring = f'{vehicle["vehicleref"]}:
lon:{vehicle["longitude"]}, lat:{vehicle["latitude"]},
line:{vehicle["lineref"]}, dest:{vehicle["destination"]},
stop:{vehicle["next_stoppointname"]}, delay: {vehicle["delaysecs"]},
timestamp: {vehicle["timestamp"]}\n'
            tiedosto.write(datastring)

# Funktio, joka tallentaa vuorokauden aikana kerätyt tiedot yhteen shapefile-
tiedostoon.
def dataToShapefile():
    global selected_data
    print(selected_data)
    newdata = gpd.GeoDataFrame()
    newdata['geometry'] = None
    index = 0

    for vehicle in selected_data:
        latitude = vehicle['latitude']
        longitude = vehicle['longitude']
        vehicleref = vehicle['vehicleref']
        lineref = vehicle['lineref']
        destination = vehicle['destination']
        delay = vehicle['delaysecs']/60
        timest = vehicle['timestamp']
        next = vehicle['next_stoppointname']
        origintime = vehicle['origintime']
        destinationtime = vehicle['desttime']

        point = Point(longitude, latitude)
        newdata.loc[index, 'geometry'] = point

```

```

newdata.loc[index, 'vehicleref'] = vehicleref
newdata.loc[index, 'lineref'] = lineref
newdata.loc[index, 'destination'] = destination
newdata.loc[index, 'delay'] = delay
newdata.loc[index, 'next_stop'] = next
newdata.loc[index, 'origin_time'] = origintime
newdata.loc[index, 'destinationtime'] = destinationtime
newdata.loc[index, 'timestamp'] = timest

index += 1

newdata.crs = from_epsg(4326)
timestamp_date = date.today()

output_path = f'./shapefiles/{timestamp_date}.shp'
newdata.to_file(output_path)

selected_data = []

# Kerätään dataa 15 minuutin välein. Vuorokauden lopussa (klo 23.59)
tallennetaan tiedot shapefile-tiedostoon.
schedule.every(15).minutes.do(fetchData)
schedule.every().day.at("23:59").do(dataToShapefile)

# Käynnistetään aikataulutetut toiminnot, kun tämä python-tiedosto
suoritetaan.
while True:
    schedule.run_pending()
    time.sleep(1)

```