Pro gradu -tutkielma
Ilmakehätieteiden maisteriohjelma
Meteorologia

Kansalaisten Netatmo-kesäsadehavaintojen laatuanalyysi

Sarai Kiema

29.2.2024

Ohjaajat: Kaisa Ylinen, Jouni Räisänen ja Marja Bister

Tarkastajat: Jouni Räisänen ja Heikki Järvinen

Helsingin Yliopisto
Matemaattis-luonnontieteellinen tiedekunta

PL 64 (Gustaf Hällströmin katu 2a)
00014 Helsingin yliopisto

HELSINGIN YLIOPISTO — HELSINGFORS UNIVERSITET — UNIVERSITY OF HELSINKI

Tiedekunta — Fakultet — Faculty	Koulutusohjelma — Utbildningsprogram — Degree programme
Matemaattis-luonnontieteellinen tiedekunta	Ilmakehätieteiden maisteriohjelma Meteorologia

Tekijä — Författare — Author

Sarai Kiema

Työn nimi — Arbetets titel — Title

Kansalaisten Netatmo-kesäsadehavaintojen laatuanalyysi

Työn laji — Arbetets art — Level	Aika — Datum — Month and year	Sivumäärä — Sidantal — Number of pages
Pro gradu -tutkielma	29.2.2024	42

Tiivistelmä — Referat — Abstract

Ilmatieteen laitoksen ylläpitämä Suomen virallisten sademittarien havaintoverkko on harva ja sen laajentaminen vaatisi enemmän resursseja. Kansalaisten sadehavaintojen hyödyntäminen olisi yksi keino laajentaa havaintoverkkoa ja siten muun muassa parantaa sääennusteiden laatua. Tässä tutkielmassa selvitettiin kansalaisten Netatmo-kesäsadehavaintojen käyttökelpoisuutta Suomessa.

Vuosien 2019–2022 kesä-, heinä- ja elokuun Netatmo-sadehavaintojen laatua tutkittiin vertailemalla niitä Ilmatieteen laitoksen automaattisadeasemien havaintoihin. Vertailua tehtiin tilastollisten suureiden, keskiarvon, korrelaation ja absoluuttisen keskivirheen, avulla. Ennen varsinaista analyysia pyrittiin rajauksilla selkeyttämään aineistoa sekä poistamaan siitä selvästi virheellisiä Netatmo-sadeasemia ja -havaintoja, kuten yli 150 mm:n tunti- ja yli 200 mm:n vuorokausisademäärät.

Pääsääntöisesti Netatmo-sadehavainnot näyttävät tilastollisten suureiden valossa hyviltä, sillä esimerkiksi 75 % Netatmo- ja lähimpien Ilmatieteen laitoksen asemien välisistä vuorokausisateiden korrelaatioista oli vähintään 0.6. Netatmo-havaintojen välinen vaihtelu oli kuitenkin suurempi kuin Ilmatieteen laitoksen asemien havaintojen, mikä kertoo osan Netatmo-havainnoista olevan virheellisiä. Virheitä löytyi useita erilaisia. Yleisesti monien Netatmo-asemien havaittiin aliarvioivan sademäärää, koska keskimäärin Netatmo-asemat olivat mitanneet sateita vajaat 10 % vähemmän kuin niiden vertailuasemat. Lisäksi Netatmo-asemien havainnoissa oli huomattavasti enemmän pieniä 0.1 mm mittauksia kuin Ilmatieteen laitoksen vertailuasemilla eikä osa asemista ollut mitannut mitään 0.1 mm virhemittauksia lukuunottamatta. Jotkut Netatmo-asemat puolestaan mittasivat yksittäin tai jopa jatkuvasti virheellisiä hyvin suuria sademääriä. Osa asemista myös yliarvioi sademäärää, sillä asemien sateet korreloivat hyvin vertailuasemien sateiden kanssa ollen vain paljon suurempia. Toisaalta joidenkin Netatmo-asemien realistiset sadehavainnot oli mitattu eri aikoihin kuin vertailuasemien sateet, joten asemien koordinaatit voivat olla väärät. Välillä taas Netatmo-havaintojen laatu muuttui ajan myötä, sillä kyse on kansalaisten havainnoista. Asemat saattoivat ensin tuottaa hyviä havaintoja ja sitten huonoja tai päinvastoin.

Kaikkiaan Netatmo-kesäsadehavainnot vaikuttavat käyttökelpoisilta, koska suurin osa havainnoista on hyviä. Netatmo-asemat myös saavat suuria sateita kiinni hyvin. Lisäksi huonoja havaintoja korrelaation perusteella tuottavat Netatmo-asemat ovat hajallaan eri puolilla Suomea ja hyviä asemia on kaikkialla enemmän. Koska virheellisiä Netatmo-asemia ja -havaintoja on silti varsin paljon, Netatmo-sadehavainnot tarvitsevat kattavaa laadun varmistusta ennen havaintojen hyödyntämistä. Laadun varmistusta voisi tehdä tämän tutkimuksen tavoin vertailemalla havaintoja tilastollisesti Ilmatieteen laitoksen asemien havaintoihin. Lisäksi Netatmo-havaintoja voisi verrata keskenään.

Avainsanat — Nyckelord — Keywords

Netatmo, Ilmatieteen laitos, sadeasema, sadehavainto, vertailu, tilastolliset suureet

Säilytyspaikka — Förvaringsställe — Where deposited

HELDA – Helsingin yliopiston digitaalinen arkisto

Muita tietoja — Övriga uppgifter — Additional information

Sisällys

1 Johdanto **1**

2 Kansalaisten Netatmo-sadeasemat **3**

3 Ilmatieteen laitoksen automaattisadeasemat **6**

4 Aineisto **8**

4.1 Rajaukset 10

5 Menetelmät **13**

6 Tulokset **15**

6.1 Tilastolliset suureet 15

6.1.1 Keskiarvot ja niiden erot 17

6.1.2 Korrelaatio 19

6.1.3 Absoluuttinen keskivirhe 23

6.2 Netatmo-asemien mittausten erilaisia virheitä 25

6.3 Sadehavaintojen suuruusjakauma 30

6.4 Suuret sadehavainnot 32

6.5 Tapaustutkimus 30.6.2020 35

7 Johtopäätökset **38**

1. Johdanto

Sadehavainnot antavat arvokasta tietoa moneen tarkoitukseen, kuten sään ja tulvien ennustamiseen. Lisäksi sadehavainnoista on hyötyä muun muassa maataloudelle, monille teollisuuden aloille ja yksityisille ihmisille. Suomen sadehavaintoverkosta ja siten virallisista sadehavainnoista vastaa Ilmatieteen laitos. Ilmatieteen laitoksella on noin 200 sadeasemaa, joista yli puolet on automaattisia ja loput manuaalisia. Koska sademäärä vaihtelee paljon paikan ja ajan suhteen, Ilmatieteen laitoksen sademittariverkko on liian harva mitatakseen sateita kattavasti. Säätutkat täydentävät sadehavaintoverkkoa. Niidenkään avulla tehdyt mittaukset eivät ole virheettömiä eivätkä ne siten voi korvata perinteisiä sademittauksia. Kattavamman sademittariverkoston rakentaminen ja ylläpitäminen vaatisi puolestaan enemmän resursseja, sillä automaattiset sademittarit ovat kalliita ja manuaalisten sadeasemien määrä taas on vuosien saatossa vähentynyt, kun havainnontekoon ei ole ollut enää jatkajia. Monilla ihmisillä on kuitenkin erilaisia sademittareita omilla pihoillaan. Yksi tällaisista sademittareista on kaupallinen Netatmo-sademittari. Tämän tutkielman tarkoituksena on selvittää, kuinka käyttökelpoisia kansalaisten Netatmo-sadehavainnot ovat kesällä Suomen alueella.

Netatmo-sademittareiden käyttökelpoisuutta tutkitaan vertailemalla kansalaisten Netatmo-sademittauksia Ilmatieteen laitoksen virallisiin automaattisiin sademittauksiin tilastollisten menetelmien avulla. Tarkoituksena on selvittää, olisiko mahdollista käyttää Netatmo-havaintoja virallisten havaintojen lisänä laajentamassa sadehavaintoverkkoa. Jos Netatmo-sadehavainnot osoittautuvat käyttökelpoisiksi, niiden avulla voitaisiin muun muassa parantaa ennusteiden laatua ja korjata säätutkien mittaaman sateen virhettä suhteessa maan pinnalla mitattuun sademäärään.

Kesä valikoitui tutkittavaksi vuodenajaksi, sillä Netatmo-sademittari ei pysty mittaamaan kiinteässä olomuodossa tulevaa sadetta, kuten lunta [Netatmo, 2012]. Tämä rajaa Netatmo-sadehavaintojen mahdollista käyttöä, sillä Suomessa talvi on pitkä. Tiheämpi havaintoverkko olisi kuitenkin juuri kesällä hyödyllisin, koska kesäsateet ovat usein hyvin lyhytkestoisia ja paikallisia. Suomessa esiintyvät sateet ovat pääasiassa rintamasateita ja konvektiivisia sateita. Säärintamien yhteydessä esiintyvät rintamasateet ovat laaja-alaisia ja pitkäkestoisia. Konvektiiviset sateet puolestaan ovat usein kuuroluonteisia, joten ne voivat olla hyvin lyhytkestoisia ja paikallisia. Konvektiosateita esiintyy Suomessa lähin-

nä kesäisin, kun taas rintamasateita esiintyy kaikkina vuodenaikoina. Täten Netatmo-sadehavainnot voisivat parantaa kesäsateiden havainnointia, sillä Ilmatieteen laitoksen viralliset sademittarit ovat liian harvassa pystyäkseen mittaamaan kaikkia sadekuuroja.

Luvussa 2 kerrotaan kansalaisten Netatmo-sadeasemien ja vastaavasti luvussa 3 Ilmatieteen laitoksen virallisten automaattisadeasemien toiminnasta ja mittausten mahdollisista virhelähteistä. Luvussa 4 esitellään tutkielman aineisto ja kerrotaan, miten se on koottu ja mitä rajauksia aineistoon on tehty. Tämän jälkeen käydään läpi tutkimusmenetelmät luvussa 5. Luvussa 6 tarkastellaan tuloksia. Ensin tuloksia tarkastellaan erilaisten tilastollisten suureiden kautta ja kerrotaan, minkälaisia virheitä Netatmo-mittauksista löytyi. Tämän jälkeen tarkastellaan sadehavaintojen suuruusjakaumaa ja suuria sadehavaintoja. Tulosten loppuun esitetään vielä lyhyt tapaustutkimus. Yhteenveto ja johtopäätökset tuloksista tehdään luvussa 7.

2. Kansalaisten Netatmo-sadeasemat

Kuva 2.1: Netatmo-sademittari [Netatmo, 2023].

Kuvassa 2.1 näkyvä Netatmo-sademittari on Netatmo-nimisen yrityksen valmistama kaupallinen tuote, jonka kuluttaja voi ostaa osaksi Netatmo-sääasemaa (Netatmo Weather Station) [Netatmo, 2012]. Netatmo-sääasema sisältää vakiovarusteena ulko- ja sisämittarit. Ulkomittari mittaa lämpötilaa ja suhteellista kosteutta [Chapman et al., 2016]. Sisämittari mittaa näiden lisäksi ilmanpainetta, hiilidioksidipitoisuutta ja äänenvoimakkuutta. Netatmo-sademittari on sääaseman erillinen lisäosa, joka ei toimi ilman Netatmo-sääasemaa [Netatmo, 2012]. Täten kansalaisten Netatmo-sadeasemat ovat Netatmo-sääasemia, joihin on hankittu lisäosana sademittari. Yksittäiseen Netatmo-sääasemaan pystyy yhdistämään vain yhden Netatmo-sademittarin.

Netatmo-sademittari on muovinen kiikkuperiaatteella toimiva sademittari, jossa on halkaisijaltaan 130 mm keräyssuppilo [Netatmo, 2023]. Netatmo-sademittarin keräimen suuaukon pinta-ala eli alue, jolta sadetta kerätään, on noin 125 cm^2 [Bárdossy et al., 2021]. Kiikkusademittarin sisällä on kaksi kääntöpisteen erottamaa kippikauhan kaltaista asti-

aa, joihin sadevesi vuorotellen valuu [Haapanala, 2013]. Sadevesi kulkee keräyssuppilon kautta sademittarin sisällä olevaan keinulaudan kaltaisen kiikun ylempään astiaan. Kun vettä on astiassa sademittarin valmistajan tai omistajan asettama määrä, kiikku kippaa veden täyttyneestä astiasta ja vesi valuu ritilän kautta pois sademittarista. Samalla toinen astioista nousee ylös ja alkaa kerätä sadevettä. Aina kipatessaan vettä sademittari lähettää pulssin. Näistä pulsseista sademittari laskee sademäärän yhtälöstä:

$$h = \frac{V}{A}P, \tag{2.1}$$

missä V on vesitilavuus, jonka täytyttyä kiikku kippaa, A on sademittarin keräimen suuaukon pinta-ala ja P on pulssien eli kippausten lukumäärä.

Valmistajan tuotetietojen mukaan Netatmo-sademittarin mittausväli on 5 minuuttia, mittausalue $0.2 - 150$ mm/h ja tarkkuus 1 mm/h [Netatmo, 2023]. Valmistaja on asettanut Netatmo-sademittarin resoluutioksi 0.1 mm. Sademittarin resoluutio riippuu asetetusta kippaustilavuudesta ja sadekeräimen suuaukon pinta-alasta [Haapanala, 2013]. Täten kiikku kippaa veden aina, kun on satanut 0.1 mm, joka saadaan kaavasta (2.1). Laitteen omistaja voi myös itse kalibroida Netatmo-sademittarin ja täten muuttaa laitteen mittausresoluutiota [Netatmo, 2012].

Netatmo-sademittari toimii, kun lämpötila on 0 °C ja 50 °C välillä [Netatmo, 2023]. Se ei siis toimi Suomen talvioloissa, sillä se ei pysty mittaamaan lumisadetta tai mitään muutakaan kiinteässä olomuodossa satavaa. Aina Netatmo-sademittari ei toimi edes lämpötilan ollessa nollan yläpuolella, sillä näissäkin lämpötiloissa voi sade olla kiinteässä olomuodossa. Tällainen tilanne on kesäinen raekuuro, jonka aikana lämpötila voi olla jopa hellelukemissa.

Data siirtyy langattomasti Bluetoothin ja Wi-Fi:n yhdistelmän avulla Netatmo-sademittarista pilveen, josta siihen päästään käsiksi älylaitteilla, kuten puhelimella tai tabletilla [Netatmo, 2012]. Netatmo-sääaseman omistaja näkee havainnot Netatmo Weather App -sovelluksesta. Lisäksi omistajan suostumuksella Netatmo-havainnot näkyvät reaaliajassa kartalla Netatmon nettisivuilla. Täten kuka tahansa voi seurata, mitä Netatmo-sääasemat eri puolilla Suomea ja maailmaa ovat juuri mitanneet. Suostumusta havaintojen jakamiselle voi muuttaa milloin vain, joten tietyn Netatmo-aseman mittaukset voivat ensin näkyä mutta myöhemmin ei. Toisaalta Netatmo-sademittarin virtalähteenä toimivat paristot [Netatmo, 2023]. Jos paristot unohdetaan vaihtaa, mittauksia ei enää tule. Kunkin Netatmo-aseman havaintotietojen yhteydessä näkyvät aseman sijainti- ja korkeustiedot. Netatmo-aseman sijainnin ja korkeuden määrittämisen voi tehdä Netatmo Weather App -sovelluksessa asemaan yhdistetyn puhelimen paikallistamisjärjestelmän eli GPS:n (Global Positioning System) avulla [Netatmo, 2012]. Sijainnin ja korkeuden voi määrittää myös manuaalisesti.

Kansalaisten sadeasemiin liittyy paljon mahdollisia virhelähteitä. Ensinnäkään ei

voida tietää, noudattavatko kansalaisten Netatmo-sadeasemat Maailman ilmatieteen järjestön WMO (World Meteorological Organisation) määrittämiä standardeja sademittarin asemapaikan suhteen [Bárdossy et al., 2021]. Sademittari on saatettu sijoittaa esimerkiksi suoraan auringonpaisteeseen, parvekkeelle, lähelle seinää tai jopa sisätiloihin. Kansalaisten Netatmo-sadeasemien asennuksesta ja ylläpidosta ei myöskään ole saatavilla tietoa. Tämän takia kansalaisten sademittauksiin voi aiheutua lukuisia virheitä esimerkiksi laitteen virheellisestä asentamisesta tai kalibroinnista, huonosta ylläpidosta ja datan siirron ongelmista.

Yksi Netatmo-sademittauksissa esiintyvistä ongelmista on virheelliset nolla-arvot [de Vos et al., 2019]. Tällöin sademääräksi mitataan 0 mm, vaikka asemalla on satanut. Syynä voi olla, että sademittari on vinossa tai tukkeutunut esimerkiksi lehtien tai hyönteisten takia. Laitteen vääristä asetuksistakin voi tulla virheitä mittauksiin. Netatmo-sademittarin kippikauhan kippaustilavuus saattaa olla virheellinen niin, ettei todellinen kipattu vesitilavuus vastaa raportoitua arvoa [de Vos et al., 2019]. Virhe voi aiheutua tehdasasetuksista valmistuksen vaihteluvuuden takia, mutta se voi syntyä myös laitetta kalibroitaessa. Netatmo-sademittari voidaan kalibroida kaatamalla hitaasti tunnettu määrä, esimerkiksi 100 ml, vettä sademittariin ja laskemalla kiikun kippaukset [Netatmo Help Center, 2023]. Netatmo-sademittariin kaadettu vesimäärä jaetaan kippikauhan kippausten lukumäärällä, jolloin tiedetään vesitilavuus, jonka saavutettuaan kiikku kippaa. Tämän jälkeen arvo voidaan korjata Netatmo Weather App -sovelluksessa. Jos vettä kaadetaan liian nopeasti, osa vedestä ohittaa mittausjärjestelmän, jolloin Netatmo-asema kalibroinnin seurauksena yliarvioi sademäärää [de Vos et al., 2019].

Netatmo-sademittarien huomattava puute on, ettei niillä ole tuulisuojaa, kuten kuvasta 2.1 nähdään. Tämän takia Netatmo-sadeasemilla esiintyy tuulesta aiheutuvaa sademäärän aliarviointia [de Vos et al., 2019]. Lisäksi kansalaisten Netatmo-sadeasemat voivat ilmoittaa säästä riippumattomia suuria sademääriä. Syynä voi olla esimerkiksi lähistöllä oleva sadetin. Toisaalta ihmiset voivat kaataa nesteitä sademittariin muun muassa puhdistaakseen sitä. Laitteen käsittely tai siihen osuminen vahingossa aiheuttaa myös kiikkuliikettä ja siten virheellisiä sademittauksia. Toisinaan taas sademittaukset eivät vastaa paikallisia sateita, jos aseman sijainti on ilmoitettu väärin.

3. Ilmatieteen laitoksen automaattisadeasemat

Kuva 3.1: Ilmatieteen laitoksen OTT Pluvio^2 -sademittari Vantaalla Helsinki–Vantaan lentoasemalla.

Kaikilla Ilmatieteen laitoksen automaattisilla sadeasemilla on käytössä kuvassa 3.1 näkyvä OTT Pluvio^2 tai vastaavanlainen OTT Pluvio^2 L -sademittari. OTT Pluvio^2 ja OTT Pluvio^2 L ovat tasapainoperiaatteella toimivia punnitsevia sadevaakoja [OTT Messtechnik, 2008, OTT Hydromet, 2016a]. Ilmatieteen laitoksen sademittarit OTT Pluvio^2 ja OTT Pluvio^2 L ovat mallia 400, mikä tarkoittaa, että niiden sadetta keräävän suuaukon pinta-ala on 400 cm^2. Sademittarin suuaukko sijaitsee 1.5 metrin korkeudella. Tällaisissa sademittareissa on sisällä sateen keruuastia, jonka kapasiteetti on 750 mm [OTT Messtechnik, 2008]. Punnitseva sadevaaka määrittää keruuastian massan sisältöineen 6 sekunnin välein 0.01 mm resoluutiolla. Vähentämällä tästä

tyhjän keräysastian massa sademittari laskee sademäärän massan, jonka se laskentakaavalla muuntaa millimetreiksi. Erittäin tarkka ruostumattomasta teräksestä valmistettu vaaka on eristetty ympäristön vaikutuksilta. Valmistajan ilmoittama tarkkuus sademittareille on 0.1 mm. Enimmillään nämä Pluvio-sademittarit pystyvät mittaaman sadetta jopa 50 mm/min. Suomen oloissa tällainen sademäärä on erittäin epätodennäköinen. Edes Yhdysvalloissa 1948 mitattu minuuttisateen maailmanennätys, 31.2 mm/min, ei ole niin suuri [World Meteorological Organization, 2023].

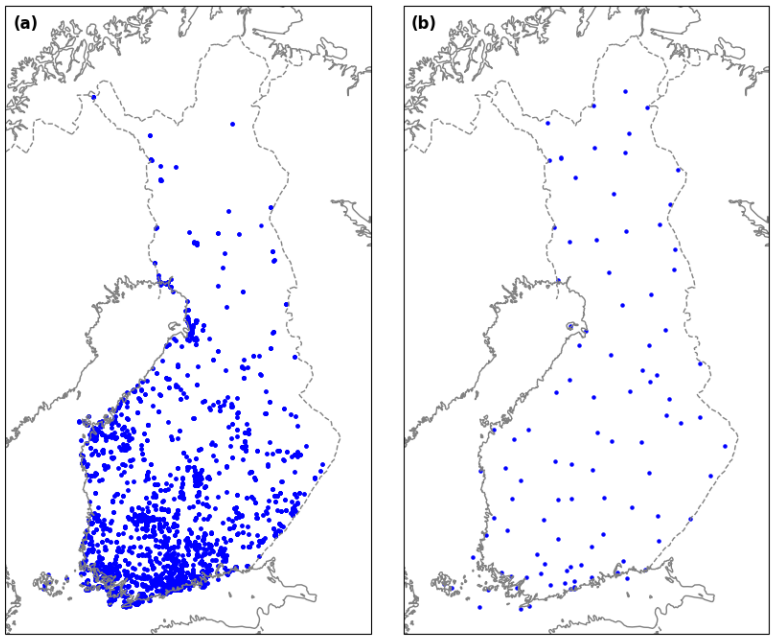
Pluvio-sademittarit toimivat lämpötilavälillä −40 °C ja +60 °C, joten ne mittaavat niin nestemäistä kuin kiinteääkin sadetta. Sademittareissa käytetään talvella sulatusnestettä [Ilmatieteen laitos, 2023a]. Lisäksi sademittareiden suuaukon kauluksessa on lämmitys, joka pitää sen sulana estäen lumen ja jään kertymisen. Kaikilla Ilmatieteen laitoksen sadeasemilla on myös tuulisuojus sademittarin ympärillä kuvan 3.1 näyttämällä tavalla [Ilmatieteen laitos, 2023a]. Tuulisuojaa käytetään vähentämään tuulen vaikutusta sademäärän mittaukseen. Se hidastaa ilman virtausta sadeastian suuaukolla, jolloin sade osuu todennäköisemmin keräysastiaan [Ilmatieteen laitos, 2023b].

Ilmatieteen laitoksen automaattisten sadeasemien paikat on valittu Maailman Ilmatieteen järjestön (WMO) ohjeiden mukaan, jotta ne olisivat mahdollisimman edustavia eikä asemapaikasta aiheutuisi suurta virhettä sademittauksiin. Lisäksi Pluvio-sademittarit täyttävät kaikki WMO:n vaatimukset sademittareille ja tarvitsevat hyvin vähän huoltoa tai muitakaan toimenpiteitä [OTT Hydromet, 2016b]. Tarvittavia toimenpiteitä ovat sateen keruuastian tyhjennys, satunnaiset silmämääräiset tarkistukset ja sulatusnesteen lisääminen talveksi. Edes kalibrointia ei tarvitse tehdä, sillä Pluvio-sademittarit on kalibroitu tehtaalla eikä muita kalibrointeja paikan päällä enää tarvita [OTT Messtechnik, 2008]. Haluttaessa Pluvio-sademittarille voidaan kuitenkin tehdä tarkkuustesti. Huoltotöiden ja tarkkuustestien aikana sademittarit eivät ilmoita sadetta, joten näistä ei synny virheellisiä mittauksia.

Ilmatieteen laitoksen automaattiset sademittauksetkaan eivät ole täysin virheettömiä. Yksi mahdollisista virhelähteistä on haihtumishäviö, sillä osa sataneesta vedestä saattaa haihtua keräysastiasta [Cecinati et al., 2018]. Tuulen vaikutuksesta taas sadeastiaan saattaa päätyä vähemmän sadetta kuin pitäisi tuulisuojuksesta huolimatta. Sähkökatkot voivat myös aiheuttaa katkoja mittauksiin, koska Pluvio-sademittarien virransaanti ja datan siirto tapahtuvat kaapelien kautta, [OTT Messtechnik, 2008].

4. Aineisto

Tutkielman aineistona on käytetty kansalaisten Netatmo-sadehavaintoja ja Ilmatieteen laitoksen virallisia automaattisia sadehavaintoja. Havaintoja on yhteensä neljän kesän ajalta käsittäen kesä-, heinä- ja elokuun vuosina 2019–2022. Kuvassa 4.1a on kaikki tutkimusaineiston 3431 Netatmo-sadeasemaa ja kuvassa 4.1b puolestaan kaikki Ilmatieteen laitoksen 101 automaattisadeasemaa Suomen kartalla. Ilmatieteen laitoksen asemia on tasaisesti eri puolilla Suomea, mutta Netatmo-asemien sijainnit painottuvat tiheästi asutuille alueille (kuva 4.1). Lisäksi kuvasta 4.1 nähdään, että Lapissa Netatmo-asemia on paikoin vähemmän kuin Ilmatieteen laitoksen asemia. Aineisto oli tutkielmaa varten muodostettu valmiiksi. Netatmo-havainnot on ostettu Ilmatieteen laitokselle, joten koko havaintoaineisto on peräisin Ilmatieteen laitoksen tietokannasta.



Kuva 4.1: Karttakuva kaikista (a) tutkimusaineiston Netatmo-sadeasemista (3431) ja (b) Ilmatieteen laitoksen automaattisadeasemista (101) merkittynä sinisillä pisteillä.

Havaintoaineisto sisältää vuorokausisadekertymän (mm/24 h) ja tuntisadekertymän (mm/h). Netatmo-sadeasemat mittaavat edellisen 24 tunnin sadekertymän kerran päivässä 21 UTC eli 00 Suomen kesäaikaa. Koska Ilmatieteen laitoksen vuorokausisadehavaintojen havaintohetki on 06 UTC eli eri kuin Netatmo-asemien, näitä ei vertailussa voitu käyttää suoraan. Tämän takia aineistoon on määritetty Ilmatieteen laitoksen automaattisadeasemien vuorokausisadekertymä 21 UTC summaamalla tasatuntien tuntisadekertymät aikavälillä 22–21 UTC. Näin muodostettu Ilmatieteen laitoksen asemien vuorokausisadekertymä vaatii, että havaintoja on joka tunnilta. Jos yksikin tuntisadehavainto puuttuu, myös vuorokausisadehavainto puuttuu. Lopulliseen vuorokausisadekertymän havaintoaineistoon on yhdistetty kunkin Netatmo-aseman vuorokausisadehavainnon kanssa kahden lähimmän Ilmatieteen laitoksen automaattisadeaseman vuorokausisadehavainnot ja etäisyys Netatmo-asemasta. Aineistossa on myös havaintoaika ja kunkin aseman asematunnus. Tässä vuorokausisadekertymän aineistossa on yhteensä 3426 eri Netatmo-sadeasemaa ja 96 eri Ilmatieteen laitoksen automaattisadeasemaa.

Netatmo-asemat mittaavat edellisen tunnin sadekertymän noin 10 minuutin välein eli mittausväli ei ole aina tasan 10 minuuttia. Eri Netatmo-asemat mittaavat sadetta eri aikaan riippuen siitä, milloin laite on käynnistetty. Havaintoaineistoon otettiin tasatunneille osuvat havainnot, jotka on todellisuudessa mitattu välillä viittä minuuttia vaille ja viittä minuuttia yli tasatunnin. Lopulliseen tuntisadekertymän aineistoon on yhdistetty Netatmo-havaintojen kanssa kahden lähimmän Ilmatieteen laitoksen automaattisadeaseman mittaamat tuntisadekertymät tasatunneilta ja etäisyys Netatmo-asemasta. Lisäksi aineistossa on havaintoaika ja asematunnukset. Tässä tuntisadekertymän aineistossa on yhteensä 3313 eri Netatmo-sadeasemaa ja 96 eri Ilmatieteen laitoksen automaattisadeasemaa.

Vertailun vuoksi vuorokausi- ja tuntisateiden aineistot tehtiin myös Ilmatieteen laitoksen asemille keskenään. Näissä vertailuaineistoissa on yhdistetty sadehavainnot kahdelta eri Ilmatieteen laitoksen automaattisadeasemalta, joiden välinen etäisyys on enintään 60 km. Aineistoissa on sadehavaintojen lisäksi havaintoaika, kunkin aseman asematunnukset ja asemien välinen etäisyys. Tässä Ilmatieteen laitoksen automaattisadeasemien keskinäisen vertailun aineistossa on yhteensä 84 eri asemaa.

Näiden havaintotiedostojen lisäksi aineistossa on tiedostot Netatmo-asemien ja Ilmatieteen laitoksen sadeasemien sijainneista. Kyseisissä tiedostoissa on kunkin aseman asematunnus, koordinaatit ja korkeus merenpinnasta. Lisäksi Ilmatieteen laitoksen asemalistassa on havaintoasemien nimet. Netatmo-asemille on luotu Ilmatieteen laitoksen tietokantaan omat asematunnukset asemakoodien perusteella. Asemakoodi koostuu laitetunnuksesta ja aseman koordinaateista. Laitetunnus on laitteelle yksilöity eikä muutu missään tilanteessa, mutta Netatmo-sadeaseman koordinaatit voivat muuttua. Jos laitteen käyttäjä muuttaa aseman sijaintitietoa, muuttuu myös Netatmo-aseman asematun-

nus Ilmatieteen laitoksen tietokannassa.

Vuorokausi- ja tuntisateiden aineistoissa havaintojen lukumäärä keskimäärin kasvaa ajan funktiona tutkimusjakson aikana. Datan toimituksessa Netatmolta Ilmatieteen laitokselle on kuitenkin ollut ajoittain katkoja. Tämän takia tuntisateiden aineistossa on ajanhetkiä, jolloin havaintojen lukumäärä putoaa rajusti useasta sadasta alle kymmeneen ja joiltain ajanhetkiltä havainnot puuttuvat kokonaan.

4.1 Rajaukset

Ennen varsinaista analyysia Netatmo-asemien vuorokausi- ja tuntisateiden aineistoihin tehtiin rajauksia. Rajauksilla pyrittiin selkeyttämään aineistoa sekä poistamaan siitä virheellisiä havaintoja ja Netatmo-sadeasemia. Taulukossa 4.1. on tilastollisia tunnuslukuja vuorokausisateiden aineistolle ja taulukossa 4.2 tuntisateiden aineistolle ennen ja jälkeen rajausten.

Ensin vuorokausisateiden aineistosta poistettiin kaikki hyvin suuret yksittäiset Netatmo-asemien havainnot, jotka olivat vähintään 200 mm. Nämä havainnot oletettiin virheellisiksi, koska Suomessa ei tilastojen valossa ole koskaan satanut yli 200 mm vuorokaudessa [Ilmatieteen laitos, 2023c]. Suomen mittaushistorian suurin vuorokausisademäärä (198.4 mm) on mitattu Espoon Lahnuksessa 21.7.1944. Yli 200 millimetrin vuorokausisadehavaintoja oli mitannut 14 eri Netatmo-asemaa yhteensä 16 kertaa. Tuntisateiden aineistosta puolestaan poistettiin virheellisinä kaikki Netatmo-havainnot, jotka olivat yli 150 mm. Tämä rajaus tehtiin, sillä Netatmo-sademittari pystyy enimmillään mittaamaan sadetta 150 mm/h [Netatmo, 2023]. Yhteensä 6 eri Netatmo-asemaa oli kerran mitannut tuntisateeksi yli 150 mm. Netatmo-asemien maksimiarvot ennen rajauksia taulukoissa 4.1 ja 4.2 ovat Suomen oloissa täysin epärealistisia. Nämä hyvin suuret yksittäiset sademäärät, varsinkin tunnin aikana mitatut, ovat luultavasti syntyneet, kun Netatmo-sademittariin on kaadettu vettä esimerkiksi laitteen toimintaa testattaessa tai laitetta kalibroitaessa.

Tämän jälkeen molemmista aineistoista poistettiin ne rivit, joilta puuttui Ilmatieteen laitoksen automaattisadeasemien havaintoja. Ennen rajauksia havaintoja oli vähemmän Ilmatieteen laitoksen vertailuasemilta kuin Netatmo-asemilta, kuten nähdään taulukoista 4.1 ja 4.2. Rajausten jälkeen havaintojen lukumäärä on sama kaikilla asemilla, kun puuttuvia havaintoja ei ole.

Sitten aineistosta poistettiin Netatmo-asemat, jotka eivät olleet kertaakaan mitanneet sadetta. Tällaisten Netatmo-asemien lukumäärä vuorokausisadeaineistossa oli 239 ja tuntisadeaineistossa 196. Suurimmalla osalla näistä asemista havaintoja oli vain yhdeltä tai muutamalta päivältä. Joukossa oli kuitenkin asemia, joilta oli havaintoja yli kuukauden ajalta, mutta havainnot olivat aina nolla eli asemat eivät olleet mitanneet sadetta. Vuoro-

kausisateiden aineistossa yhteensä 44:ltä eri Netatmo-asemalta on pelkkiä nollahavaintoja yli 30 päivän ajalta. Vastaavasti tuntisateiden aineistossa on yli 720 nollahavaintoa, mikä vastaa yli 30 vuorokautta, yhtä lailla 44:ltä eri Netatmo-asemalta. Jos asemalta on vain vähän havaintoja, mutta asema ei ole kertaakaan mitannut sadetta, ei voida tietää, mittaako asema oikeasti sadetta vai ei. Jos taas asemalta on paljon pelkkiä nollahavaintoja, voi olettaa, ettei asema mittaa sadetta ja tuottaa siten virheellisiä havaintoja. Netatmo-asemat, joiden sadehavainto on aina nolla eivät myöskään anna aineistolle lisäarvoa.

Jos asemalta on vain vähän havaintoja, havaintojen laatua ei pysty kunnolla arvioimaan. Tämän takia asemat eivät ole mielekkäitä tilastollisen analyysin kannalta. Sen tähden aineistosta poistettiin Netatmo-asemat, joilta oli vähän havaintoja. Vuorokausisadekertymän aineistoa rajattiin niin, että sinne jäi vain asemat, joilta oli yli 30 havaintoa. Vastaavasti tuntisateiden aineiston Netatmo-asemilta täytyi olla yli 720 havaintoa. Näillä rajauksilla kaikilla Ilmatieteen laitoksen vertailuasemilla on satanut mittausjakson aikana. Vuorokausisateiden aineistossa oli 756 Netatmo-asemaa, joilta oli alle 31 havaintoa. Tuntisateiden aineistossa oli puolestaan 702 Netatmo-asemaa, joilta oli alle 721 havaintoa. Molemmissa aineistossa oli monia asemia, joilta oli vain yksi havainto, vuorokausisateiden aineistossa 221 ja tuntisateiden 70.

Edellä mainittujen rajausten lisäksi määritettiin vielä, että Netatmo-aseman ja lähimmän Ilmatieteen laitoksen vertailuaseman välinen etäisyys saa olla korkeintaan 60 km. Tähän rajaukseen päädyttiin, sillä suurilla etäisyyksillä sateiden vertailtavuus heikkenee. Ilmatieteen laitoksen asemien keskinäisessä vertailuaineistossa asemien väliseksi etäisyydeksi määritettiin myös enintään 60 km.

Kaikkiaan rajatussa aineistossa on Netatmo-asemilta yli 30 havaintoa vuorokausisateiden ja yli 720 havaintoa tuntisateiden tapauksessa. Aseman havaintojen keskiarvo on yli 0 mm, lähimmän vertailuaseman etäisyys on alle 60 km ja puuttuvien havaintoarvojen rivit on poistettu datasta. Lisäksi Netatmo-havaintojen vuorokausisademäärä on alle 200 mm ja tuntisademäärä alle 150 mm.

Taulukoista 4.1 ja 4.2 nähdään, että rajaukset karsivat aineistosta hieman alle 1000 Netatmo-asemaa. Kumpaankin aineistoon jäi silti yli 2500 Netatmo-asemaa, mikä on paljon verrattuna Ilmatieteen laitoksen automaattisadeasemien lukumäärään (kuva 4.1). Taulukoista nähdään, että myös muutama Ilmatieteen laitoksen asema karsiutui aineistosta rajauksia tehtäessä. Koska Ilmatieteen laitoksen automaattisadeasemia on paljon vähemmän kuin Netatmo-asemia, sama Ilmatieteen laitoksen asema on lähin ja toiseksi lähin asema monelle eri Netatmo-asemalle. Lisäksi taulukot 4.1 ja 4.2 osoittavat, että Netatmo-asemilta ja Ilmatieteen laitoksen asemilta on sadehavaintoja, eli nollaa suurempia havaintoja, suunnilleen saman verran rajausten jälkeen. Taulukot kertovat, että keskiarvot eivät rajauksissa juuri muutu. Tämä selittyy havaintojen suurella määrällä. Netatmo-asemilla havaintojen keskiarvo on kuitenkin hieman, vajaat 10 %, pienempi kuin

Ilmatieteen laitoksen asemilla sekä vuorokausi- että tuntisateiden tapauksessa (taulukot 4.1 ja 4.2). Syynä voi olla, että Netatmo-asemat aliarvioivat sademäärää. Rajausten myötä Netatmo-asemien keskihajonta puolestaan pienenee ja samalla lähenee Ilmatieteen laitoksen asemien keskihajontaa (taulukot 4.1 ja 4.2). Tämä keskihajonnan pieneneminen on pääasiassa hyvin suurten Netatmo-sadehavaintojen poistamisen ansiota. Samasta syystä myös Netatmo-asemien maksimisadehavainnot ovat pienemmät rajausten jälkeen. Rajausten jälkeen Netatmo-asemien suurimmat sadehavainnot ovat silti vielä yli kaksinkertaisia Ilmatieteen laitoksen asemien suurimpiin sadehavaintoihin verrattuna.

Taulukko 4.1: Tilastollisia tunnuslukuja koko aineiston vuorokausisateille (24 h) ilman rajauksia ja rajauksilla Netatmo-sadeasemilla ja niiden kahdella lähimmällä Ilmatieteen laitoksen (FMI) sadeasemalla.

	Netatmo-sadeasemat		**Lähimmät FMI-sadeasemat**		**Toiseksi lähimmät FMI-sadeasemat**	
	Kaikki	Rajaus	Kaikki	Rajaus	Kaikki	Rajaus
Asemat	3416	2556	90	88	90	86
Havainnot	379 029	342 455	366 867	342 455	366 777	342 455
Sadehavainnot	143 809	131 854	140 785	131 365	141 178	131 790
Keskiarvo [mm]	1.87	1.88	2.08	2.08	2.09	2.10
Keskihajonta [mm]	7.54	5.46	5.25	5.26	5.31	5.33
Minimi [mm]	0.0	0.0	0.0	0.0	0.0	0.0
Maksimi [mm]	1661.3	196.2	87.6	87.6	87.6	87.6

Taulukko 4.2: Tilastollisia tunnuslukuja koko aineiston tuntisateille (1 h) ilman rajauksia ja rajauksilla Netatmo-sadeasemilla ja niiden kahdella lähimmällä Ilmatieteen laitoksen (FMI) sadeasemalla.

	Netatmo-sadeasemat		**Lähimmät FMI-sadeasemat**		**Toiseksi lähimmät FMI-sadeasemat**	
	Kaikki	Rajaus	Kaikki	Rajaus	Kaikki	Rajaus
Asemat	3313	2549	89	88	89	86
Havainnot	8 376 631	8 026 592	8 357 177	8 026 592	8 356 148	8 026 592
Sadehavainnot	608 871	590 065	612 135	588 334	615 617	591 197
Keskiarvo [mm]	0.079	0.079	0.085	0.085	0.086	0.086
Keskihajonta [mm]	1.230	0.615	0.589	0.590	0.602	0.602
Minimi [mm]	0.0	0.0	0.0	0.0	0.0	0.0
Maksimi [mm]	1661.3	129.3	48.3	48.3	48.3	48.3

5. Menetelmät

Netatmo-sadehavaintojen laatua tutkittiin vertailemalla niitä Ilmatieteen laitoksen automaattisadeasemien havaintoihin. Vertailu tehtiin Python-ohjelmointikielellä tilastollisia tunnuslukuja, kuten keskiarvoa, korrelaatiota ja absoluuttista keskivirhettä, hyödyntäen. Tunnusluvut laskettiin kunkin Netatmo-sadeaseman havainnoille ja niitä verrattiin lähimpien Ilmatieteen laitoksen asemien tunnuslukuihin. Samoin tehtiin kahden Ilmatieteen laitoksen aseman välillä ja tuloksia verrattiin keskenään. Pääasiassa vertailuun käytettiin erilaisia kuvaajia, kuten hajontakuvioita, laatikko-janakuvioita, histogrammeja, aikasarjoja ja karttoja. Lisäksi aineiston tilastollisia tunnuslukuja vertailtiin taulukoiden avulla.

Netatmo-asemien ja Ilmatieteen laitoksen asemien havaintojen vastaavuutta tutkittiin vertailemalla asemien keskiarvoja. Vertailun apuna käytettiin keskiarvojen erotusta ja absoluuttista eroa. Keskiarvot kertovat, ovatko asemat tutkittavan mittausjakson aikana mitanneet sadetta keskimäärin saman verran.

Koska keskiarvojen perusteella ei voida päätellä, mittaavatko asemat sateita samaan aikaan ja samassa suhteessa, havaintoja vertailtiin myös Pearsonin korrelaation avulla. Pearsonin korrelaatio kuvaa havaintojen lineaarista riippuvuutta ja se saa arvoja välillä -1 ja 1. Korrelaation ollessa -1, toisen muuttujan arvojen kasvaessa toisen muuttujan arvot pienenevät samassa suhteessa. Vastaavasti korrelaation ollessa 1, toisen muuttujan arvojen kasvaessa myös toisen muuttujan arvot kasvavat samassa suhteessa. Jos korrelaatio on nolla, muuttujien välillä ei ole ainakaan lineaarista riippuvuutta. Korrelaatio ei kuitenkaan kerro, ovatko mitatut sateet saman suuruisia.

Keskiarvon ja korrelaation puutteiden täydentämiseksi määritettiin myös absoluuttinen keskivirhe MAE (Mean Absolute Error). Absoluuttinen keskivirhe saadaan kaavasta

$$MAE = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n}, \tag{5.1}$$

missä n on havaintojen lukumäärä, i on havainnon järjestysluku, y on tarkasteltavan aseman havainto ja x vertailuaseman havainto. Täten absoluuttinen keskivirhe on asemien havaintojen poikkeamien itseisarvojen keskiarvo, joka on sitä suurempi, mitä enemmän asemien havainnot eroavat toisistaan. Kutakin tunnuslukua, keskiarvoa, korrelaatiota ja absoluuttista keskivirheitä, tutkittiin asemien välisen etäisyyden ja Netatmo-aseman havaintojen lukumäärän suhteen.

Tilastollisten suureiden avulla aineistosta etsittiin virheellisiä havaintoja tuottavia Netatmo-asemia. Etsintää tehtiin tapauksissa, joissa vähintään yksi tilastollinen suure kertoi aseman havaintojen mahdollisesta virheellisyydestä esimerkiksi heikon korrelaation tai suuren absoluuttisen keskivirheen kautta. Näin valikoituneiden Netatmo-asemien sadehavaintojen aikasarjoja verrattiin lähimpien Ilmatieteen laitoksen asemien aikasarjoihin. Löytyneiden virheellisiä havaintoja tuottavien Netatmo-asemien avulla selvitettiin asemien erilaisia mahdollisia virheitä.

Sitten sadehavaintojen laatua tutkittiin suuruusjakaumien ja suurten sadehavaintojen esiintyvyyden kautta. Netatmo-asemien nollasta poikkeaville sadehavainnoille määritettiin suuruusjakauma, jota verrattiin Ilmatieteen laitoksen vertailuasemien sadehavaintojen suuruusjakaumaan. Lisäksi Netatmo-sadeasemille laskettiin suuruusjakauma, jossa oli vain ne nollasta poikkeavat havainnot, joiden aikaan myös Ilmatieteen laitoksen vertaluasemalla oli satanut. Suuria sadehavaintoja puolestaan tutkittiin tarkastelmalla, kuinka monessa prosenttissa tapauksista Netatmo-asema mittaa suuria sateita Ilmatieteen laitoksen aseman niitä mitatessa. Vertailun vuoksi prosenttiosuudet laskettiin myös kahden Ilmatieteen laitoksen aseman välille. Lisäksi prosenttiosuudet määritettiin etäisyyden funktiona.

Lopuksi tapaustutkimuksen avulla selvitettiin, miten Netatmo-sademittarit mittaavat sadetta todellisessa sadetilanteessa. Netatmo-asemien tuntisadehavaintoja verrattiin lähimmän Ilmatieteen laitoksen aseman havaintoihin 30.6.2020. Tällöin laaja rintamasadealue oli Suomen yllä.

6. Tulokset

Tässä luvussa esitetään analyysin tulokset. Kaikki esitettävät tulokset on saatu käyttämällä luvussa 4.1 mainittuja rajauksia. Ensin alaluvussa 6.1 tarkastellaan havaintojen vastaavuutta tilastollisten suureiden kautta. Alaluvun 6.1 alkuun esitellään koko aineiston tilastollisten suureiden tunnuslukuja, jonka jälkeen perehdytään lähemmin keskiarvoon ja sen eroihin (6.1.1), korrelaatioihin (6.1.2), ja absoluuttisiin keskivirheisiin (6.1.3). Seuraavaksi alaluvussa 6.2 käydään läpi tilastollisten suureiden avulla löytyneitä Netatmo-asemien mittausten erilaisia virheitä. Sitten alaluvussa 6.3 vertaillaan sadehavaintojen suuruusjakaumia ja alaluvussa 6.4 tarkastellaan suuria sadehavaintoja. Lopuksi alaluvussa 6.5 esitetään lyhyt tapaustutkimus yksittäisestä sadetilanteesta.

6.1 Tilastolliset suureet

Pääasiassa Netatmo-sadeasemien havaintojen laatua tutkittiin tilastollisten suureiden avulla. Taulukossa 6.1 on vuorokausisateille ja taulukossa 6.2 tuntisateille kaikkien Netatmo-asemien ja niiden lähimpien Ilmatieteen laitoksen asemien välisten korrelaatioiden, absoluuttisten keskiarvojen erojen ja absoluuttisten keskivirheiden keskiarvot, mediaanit, keskihajonnat, minimit ja maksimit. Taulukoissa 6.1 ja 6.2 on samat arvot myös Ilmatieteen laitoksen asemien keskinäiselle vertailulle.

Vuorokausisateiden korrelaatioiden keskiarvo Netatmo-aseman ja lähimmän Ilmatieteen laitoksen aseman välillä on hieman pienempi mutta hyvin lähellä Ilmatieteen laitoksen asemien välisten korrelaatioiden keskiarvoa (taulukko 6.1). Tuntisateiden aineistossa puolestaan Netatmo-aseman ja lähimmän Ilmatieteen laitoksen aseman välisten korrelaatioden keskiarvo on suurempi kuin Ilmatieteen laitoksen asemien välinen (taulukko 6.2). Syynä tähän voi olla, että Ilmatieteen laitoksen asemien väliset etäisyydet ovat keskimäärin suurempia kuin Netatmo-aseman ja lähimmän Ilmatieteen laitoksen aseman väliset etäisyydet. Sateet nimittäin osuvat samoihin aikoihin sitä harvemmin, mitä kauempana sademittarit ovat toisistaan. Tämä selittää myös sitä, että absoluuttisten keskivirheiden keskiarvot ja mediaanit ovat pienempiä ja korrelaatioiden mediaanit suurempia Netatmo-asemien ja lähimpien vertailuasemien välillä kuin Ilmatieteen laitoksen asemien välillä sekä vuorokausi- että tuntisadeaineistossa (taulukot 6.1 ja 6.2). Keskiarvojen absoluutti-

sen eron keskiarvot ja mediaanit ovat kuitenkin suurempia Netatmo-asemien ja lähimpien vertailuasemien välillä kuin Ilmatieteen laitoksen asemien välillä sekä vuorokausi- että tuntisadeaineistossa (taulukot 6.1 ja 6.2). Tätä selittää se, että keskiarvot on laskettu kunkin asemaparin tapauksessa koko mittausjaksolle, jolloin sateiden eriaikaisuus ei vaikuta. Näin ollen systemaattiset erot Netatmo-asemien ja Ilmatieteen laitoksen asemien välillä ovat suurempia kuin vastaavat systemaattiset erot Ilmatieteen laitoksen asemien välillä.

Taulukko 6.1: Vuorokausisateiden tilastollisia tunnuslukuja Netatmo-aseman ja sen lähimmän Ilmatieteen laitoksen vertailuaseman (Netatmo) sekä kahden Ilmatieteen laitoksen aseman (FMI) välisille korrelaatioille, absoluuttisille keskiarvojen eroille ja absoluuttisille keskivirheille.

	Korrelaatio		**Absoluuttinen keskiarvojen ero** [mm]		**Absoluuttinen keskivirhe** [mm]	
	Netatmo	FMI	Netatmo	FMI	Netatmo	FMI
Keskiarvo	0.69	0.71	0.55	0.20	1.44	1.48
Mediaani	0.75	0.71	0.37	0.17	1.31	1.50
Keskihajonta	0.22	0.093	0.81	0.16	0.89	0.30
Minimi	-0.16	0.49	0.00038	0.00029	0.19	0.45
Maksimi	0.995	0.960	15.5	0.6	18.4	2.4

Taulukko 6.2: Tuntisateiden tilastollisia tunnuslukuja Netatmo-aseman ja sen lähimmän Ilmatieteen laitoksen vertailuaseman (Netatmo) sekä kahden Ilmatieteen laitoksen aseman (FMI) välisille korrelaatioille, absoluuttisille keskiarvojen eroille ja absoluuttisille keskivirheille.

	Korrelaatio		**Absoluuttinen keskiarvojen ero** [mm]		**Absoluuttinen keskivirhe** [mm]	
	Netatmo	FMI	Netatmo	FMI	Netatmo	FMI
Keskiarvo	0.46	0.32	0.022	0.008	0.09	0.11
Mediaani	0.46	0.30	0.014	0.007	0.08	0.11
Keskihajonta	0.22	0.12	0.035	0.006	0.046	0.020
Minimi	-0.03	0.13	0.00001	0.00013	0.010	0.034
Maksimi	0.98	0.91	0.69	0.03	0.87	0.16

Taulukoista 6.1 ja 6.2 nähdään, että korrelaatioiden, absoluuttisten keskiarvon erojen ja absoluuttisten keskivirheiden keskihajonta ja maksimiarvot ovat suurempia ja minimiarvot yleensä pienempiä Netatmo-asemien ja lähimpien Ilmatieteen laitoksen asemien

kuin Ilmatieteen laitoksen asemien välillä sekä vuorokausi- että tuntisadeaineistossa. Koska Netatmo-asemia on paljon enemmän kuin Ilmatieteen laitoksen sadeasemia, selittää se Netatmo-asemien ja lähimpien vertailuasemien välisiä pienempiä minimi- ja suurempia maksimiarvoja. Netatmo-asemien ja niiden lähimpien Ilmatieteen laitoksen asemien muodostamien vertailuparien suurempi lukumäärä nimittäin systemaattisesti laskee minimiarvoja ja kasvattaa maksimiarvoja verrattuna tilanteeseen, jossa vertailtavia asemapareja olisi yhtä vähän kuin Ilmatieteen laitoksen asemien keskinäisessä vertailuaineistossa. Keskihajonnat taulukoissa 6.1 ja 6.2 kuitenkin kertovat, että Netatmo-asemien ja lähimpien vertailuasemien välillä on mittausten vastaavuudessa enemmän eroja kuin Ilmatieteen laitoksen asemien välillä, sillä keskihajontoihin otoskoko ei vaikuta systemaattisesti. Tästä voidaan päätellä, että osa Netatmo-asemista tuottaa selvästi virheellisiä havaintoja.

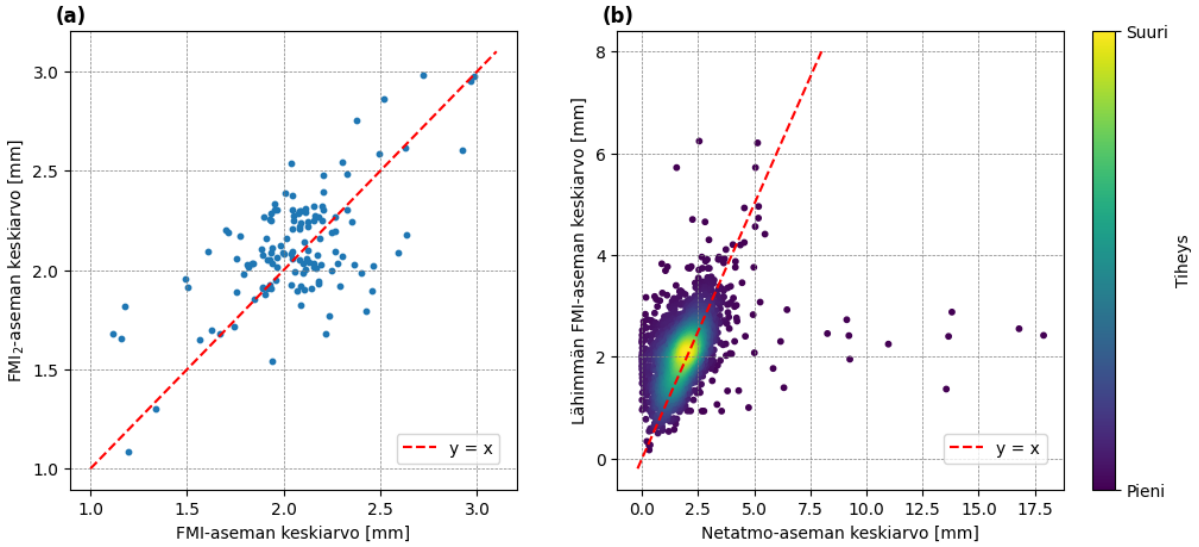
Vaikka yksittäisiin pieniin keskiarvojen absoluuttisiin eroihin Netatmo- ja vertailuasemien välillä on todennäköisesti syynä vertailuparien suuri lukumäärä, on muitakin mahdollisia syitä taulukoista 6.1 ja 6.2 havaittaviin Netatmo-asemien ja vertailuasemien välisiin parempiin tilastollisiin tunnuslukuihin. Netatmo-asemien ja vertailuasemien välinen suurempi korrelaation maksimiarvo sekä pienemmät absoluuttiset keskivirheet selittyvät osaksi sillä, että etäisyys Ilmatieteen laitoksen asemien välillä on keskimäärin suurempi kuin Netatmo-aseman ja sen lähimmän Ilmatieteen laitoksen aseman välillä. Syynä voi myös olla, että Netatmo-asemien havaintojen aikasarjat ovat hyvin eripituisia. Osalta asemista on havaintoja koko tarkasteltavan jakson ajalta eli neljän kesän ajalta, kun taas osalta asemista on havaintoja yhdeltä kesältä tai vain yhdeltä kuukaudelta. Ilmatieteen laitoksen asemien aineistossa havaintoja on joka asemaparin tapauksessa neljän kesän ajalta. Tämän takia lyhyemmät Netatmo-asemien mittausten aikasarjat ovat voineet aiheuttaa osan hyvin pienistä keskivirheistä sekä hyvin suuret korrelaatiot.

Kaikkiaan taulukoiden 6.1 ja 6.2 perusteella Netatmo-asemien sadehavainnot näyttävät keskimäärin varsin hyviltä. Havaittavissa on silti tapauksia, joissa Netatmo-asemat tuottavat selvästi Ilmatieteen laitoksen vertailuasemista poikkeavia havaintoja. Seuraavaksi tarkastellaan lähemmin keskiarvoja eroineen, korrelaatioita ja absoluuttisia keskivirheitä.

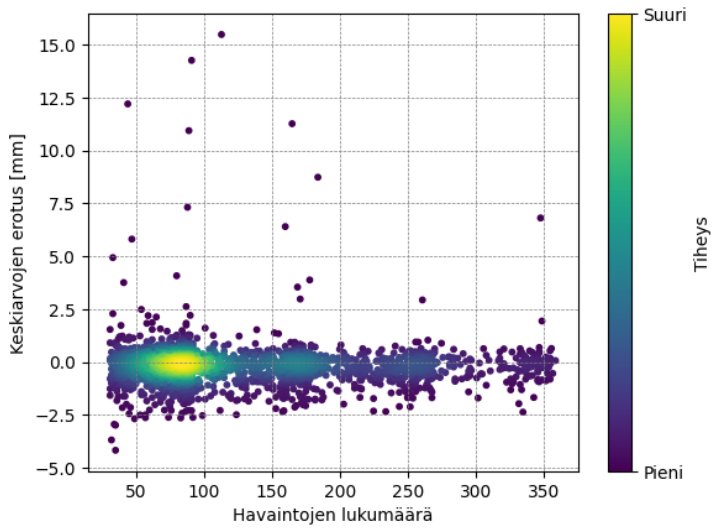
6.1.1 Keskiarvot ja niiden erot

Vuorokausisadehavaintojen keskiarvot on esitetty Ilmatieteen laitoksen asemien välillä kuvassa 6.1a ja Netatmo-asemien ja lähimpien Ilmatieteen laitoksen vertailuasemien välillä kuvassa 6.1b, missä väreillä kuvataan pisteiden tiheyttä eli lukumäärää samassa kohdassa. Punaiset katkoviivat kuvassa 6.1 havainnollistavat keskiarvojen vastaavuutta, sillä pisteen osuessa viivalle asemien keskiarvot ovat samat. Ilmatieteen laitoksen asemien väliset keskiarvot vastaavat kuvan 6.1a perusteella hyvin toisiaan, koska erot sadeha-

vaintojen keskiarvoissa ovat hyvin pieniä (alle 1 mm/vrk). Kuvasta 6.1b puolestaan nähdään, että Netatmo-asemien ja niiden lähimpien Ilmatieteen laitoksen vertailuasemien keskiarvot ovat suurimmaksi osaksi lähellä toisiaan. Joukossa on kuitenkin tapauksia, joissa Netatmo-aseman keskiarvo on paljon suurempi kuin vertailuaseman. Kuvasta 6.1b nähdään myös monia tapauksia, joissa vertailuaseman keskiarvo on huomattavasti suurempi kuin Netatmo-aseman. Osassa tapauksista Netatmo-aseman keskiarvo on lähellä nollaa eli Netatmo-asema on mitannut hyvin vähän sateita.



Kuva 6.1: Vuorokausisadehavaintojen keskiarvojen vertailu (a) Ilmatieteen laitoksen (FMI) automaattisadeasemien välillä ja (b) Netatmo-asemien ja lähimpien Ilmatieteen laitoksen vertailuasemien välillä.



Kuva 6.2: Netatmo-aseman ja lähimmän vertailuaseman vuorokausisateiden keskiarvojen erotus Netatmo-aseman vuorokausisadehavaintojen lukumäärän funktiona.

Netatmo-asemien ja lähimpien Ilmatieteen laitoksen asemien vuorokausisateiden

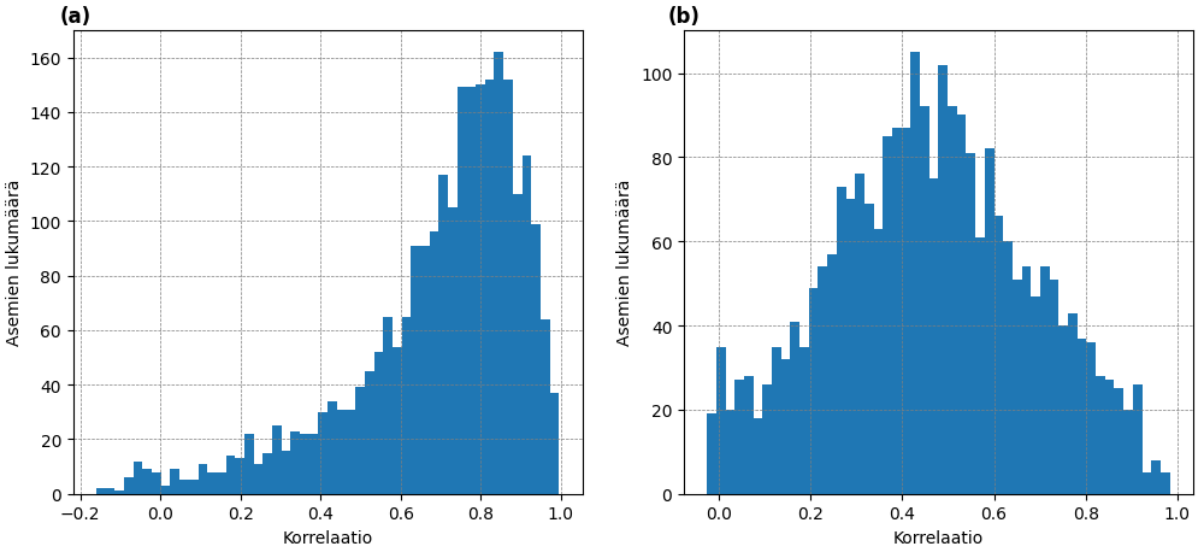
keskiarvojen erotus on esitetty Netatmo-aseman havaintojen lukumäärän funktiona kuvassa 6.2. Pääsääntöisesti keskiarvojen erot ovat pieniä. Suurimmat keskiarvojen erot ovat asemilta, joilla Netatmo-aseman keskiarvo on paljon suurempi kuin vertailuaseman. Nämä hyvin suuret Netatmo-asemien keskiarvot kertovat asemien mittausten virheellisyydestä. Monissa tapauksissa vertailuasemien keskiarvo puolestaan on paljon suurempi kuin Netatmo-aseman. Näissä tapauksissa havaintoja on usein vähän, mikä selittää keskiarvojen suurempaa eroa. Joissakin tapauksissa havaintoja on paljon, mutta Netatmo-aseman havaintojen keskiarvo on hyvin pieni, kun taas vertailuasemien havaintojen keskiarvo on huomattavasti suurempi. Suuria Netatmo-asemien ja vertailuasemien keskiarvojen välisiä eroja esiintyy siis riippumatta havaintojen lukumäärästä.

Kuvissa 6.1b ja 6.2 korostuvat Netatmo-asemien suuret keskiarvot suhteessa vertailuasemien keskiarvoihin, sillä näissä tapauksissa keskiarvojen erot ovat hyvin suuria. Koska Ilmatieteen laitoksen asemilla havaintojen keskiarvot eivät ole epärealistisen suuria, erot keskiarvoissa ovat toisin päin pienempiä. Kuvien 6.1b ja 6.2 perusteella näyttää kuitenkin siltä, että Ilmatieteen laitoksen vertailuaseman havaintojen keskiarvo on yleisemmin suurempi kuin Netatmo-aseman. Netatmo-asemat saattavat siis aliarvioida sademäärää, kuten jo taulukoista 4.1 ja 4.2 oli pääteltävissä.

Keskiarvoja etäisyyden suhteen tarkasteltaessa huomattiin, että keskimäärin keskiarvojen erot hieman kasvavat etäisyyden kasvaessa. Koska suuria keskiarvojen eroja esiintyy silti kaikilla etäisyyksillä, etäisyys Netatmo-aseman ja vertailuaseman välillä ei ole pääsyy keskiarvojen suuriin eroihin. Oikeastaan suurimpien keskiarvojen erojen tapauksissa asemien välinen etäisyys on lyhyt. Tämä on oletettavasti seurausta siitä, että suurin osa vertailuasemista sijaitsee lyhyillä etäisyyksillä Netatmo-asemista.

6.1.2 Korrelaatio

Jotta saadaan keskiarvojen lisäksi tietoa siitä, kuinka samoihin aikoihin ja samassa suhteessa Netatmo-asemat ja niiden lähimmät vertailuasemat mittaavat sateita, tarkastellaan seuraavaksi korrelaatioita. Netatmo-asemien ja lähimpien Ilmatieteen laitoksen vertailuasemien välisten korrelaatioiden frekvenssijakauma on esitetty kuvassa 6.3a vuorokausisateille ja kuvassa 6.3b tuntisateille. Kuvasta 6.3 nähdään, että vuorokausisateiden korrelaatio on keskimäärin parempi kuin tuntisateiden. Tämä on luonnollista, sillä esimerkiksi laajan saderintaman kulkiessa alueen yli sateet eivät välttämättä osu samalla tunnilla vertailuasemalle, vaikka vuorokauden aikana asemien sademäärät voivat vastata toisiaan. Vuorokausisateiden histogrammin huippu kuvassa 6.3a sijoittuu korrelaation 0.8 läheisyyteen ja 38 % vuorokausisateiden korrelaatioista on vähintään 0.8. Tuntisateiden histogrammin huippu kuvassa 6.3b taas sijoittuu korrelaation 0.5 läheisyyteen ja 43 % tuntisateiden korrelaatioista on vähintään 0.5.



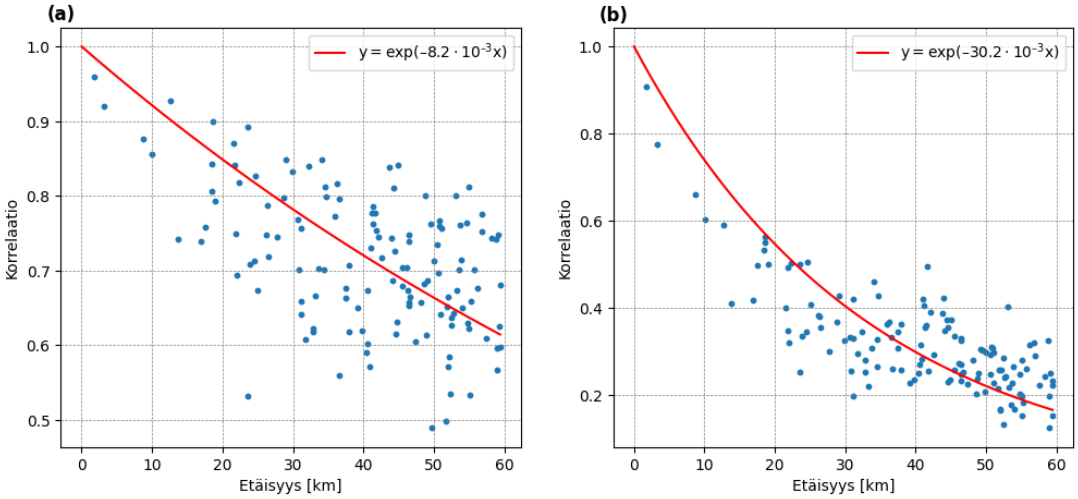
Kuva 6.3: Histogrammit Netatmo-sadeasemien ja lähimpien Ilmatieteen laitoksen vertailuasemien välisten korrelaatioiden jakaumasta (a) vuorokausisateiden ja (b) tuntisateiden tapauksessa.

Yleisesti korrelaatiota 0.7 pidetään hyvänä. Korrelaation 0.7 saavuttaa 60 % vuorokausisateiden ja 16 % tuntisateiden tapauksista. Hyvälle ja huonolle korrelaatiolle ei kuitenkaan ole tarkkoja yleisiä raja-arvoja, sillä hyvän korrelaation raja riippuu tapauksesta. Hyvän korrelaation arvioimiseksi määritettiin korrelaatiot Ilmatieteen laitoksen asemien havaintojen välille (kuva 6.4). Suurin osa Ilmatieteen laitoksen asemien välisistä korrelaatioista on vähintään 0.6 vuorokausisateiden tapauksessa kuvassa 6.4a ja noin puolet vähintään 0.3 tuntisateiden tapauksessa kuvassa 6.4b. Näihin lukuihin nähden voidaan korrelaatioita Netatmo-asemien ja lähimpien vertailuasemien välillä pitää keskimäärin hyvinä. Netatmo-aseman ja lähimmän vertailuaseman välisistä korrelaatioista 75 % on vähintään 0.6 vuorokausisateiden ja 0.3 tuntisateiden tapauksessa.

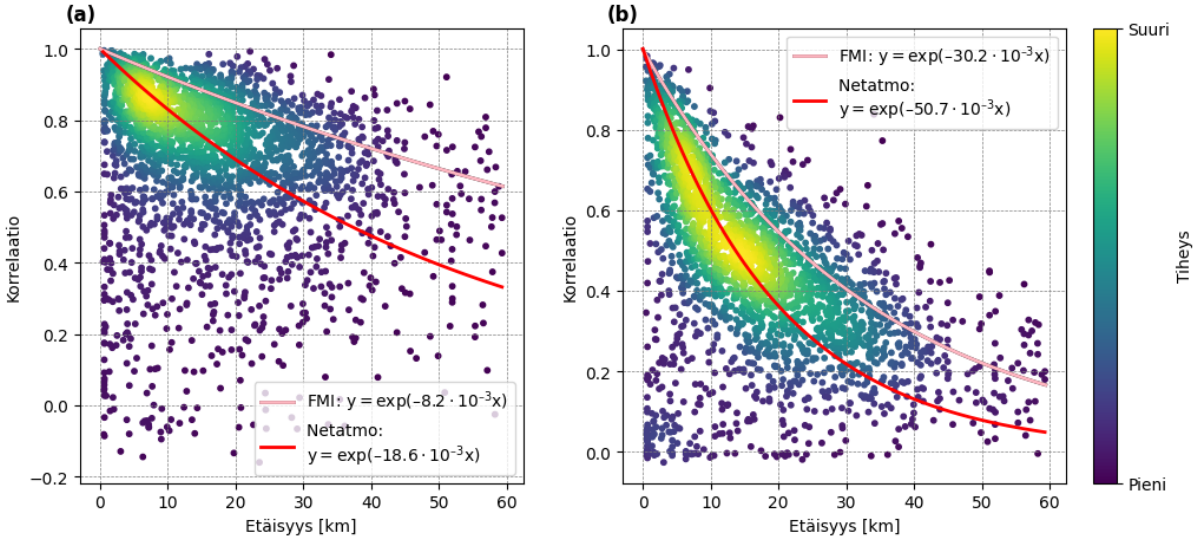
Voimakkaita ja heikkoja korrelaatioita esiintyy Netatmo-asemien ja vertailuasemien välillä riippumatta Netatmo-aseman havaintojen lukumäärästä. Etäisyys kuitenkin vaikuttaa korrelaatioon. Kuvan 6.4 pisteisiin sovitetuista punaisista käyristä nähdään, että Ilmatieteen laitoksen sadeasemien välisen etäisyyden kasvaessa korrelaatio heikkenee. Lisäksi kuvan 6.4 perusteella korrelaation heikkeneminen on nopeampaa tuntisateiden kuin vuorokausisateiden tapauksessa. Netatmo-asemien ja lähimpien Ilmatieteen laitoksen vertailuasemien korrelaatiot etäisyyden suhteen on esitetty kuvassa 6.5a vuorokausisateille ja 6.5b tuntisateille. Vertailun vuoksi punaisten sovituskäyrien lisäksi kuvaan 6.5 on lisätty kuvan 6.4 sovitusfunktiot vaaleanpunaisella. Kuvan 6.5 punaisista käyristä nähdään, että myös Netatmo-asemien ja vertailuasemien korrelaatio heikkenee keskimäärin etäisyyden funktiona. Koska punaiset käyrät kulkevat vaaleanpunaisten alapuolella, ovat nämä korrelaatiot kuitenkin keskimäärin heikompia kuin korrelaatiot samalla etäisyydellä toisistaan olevien Ilmatieteen laitoksen asemien välillä. Tämä voi kertoa Netatmo-asemien

havaintojen virheellisyydestä.

Pääsääntöisesti Netatmo-asemien ja Ilmatieteen laitoksen asemien väliset korrelaatiot vaikuttavat silti hyviltä, koska suurin osa korrelaatioista sijoittuu vaaleanpunaisen käyrän läheisyyteen, kuten vihreän ja keltaisen sävyt kuvassa 6.5 kertovat. Huonoja korrelaatioita esiintyy kumminkin kaikilla etäisyyksillä (kuva 6.5). Joukossa on nimittäin monia tapauksia, joissa korrelaatio on hyvin heikko tai sitä ei ole lainkaan, vaikka Netatmo-asema ja vertailuasema sijaitsevat hyvin lähellä toisiaan.

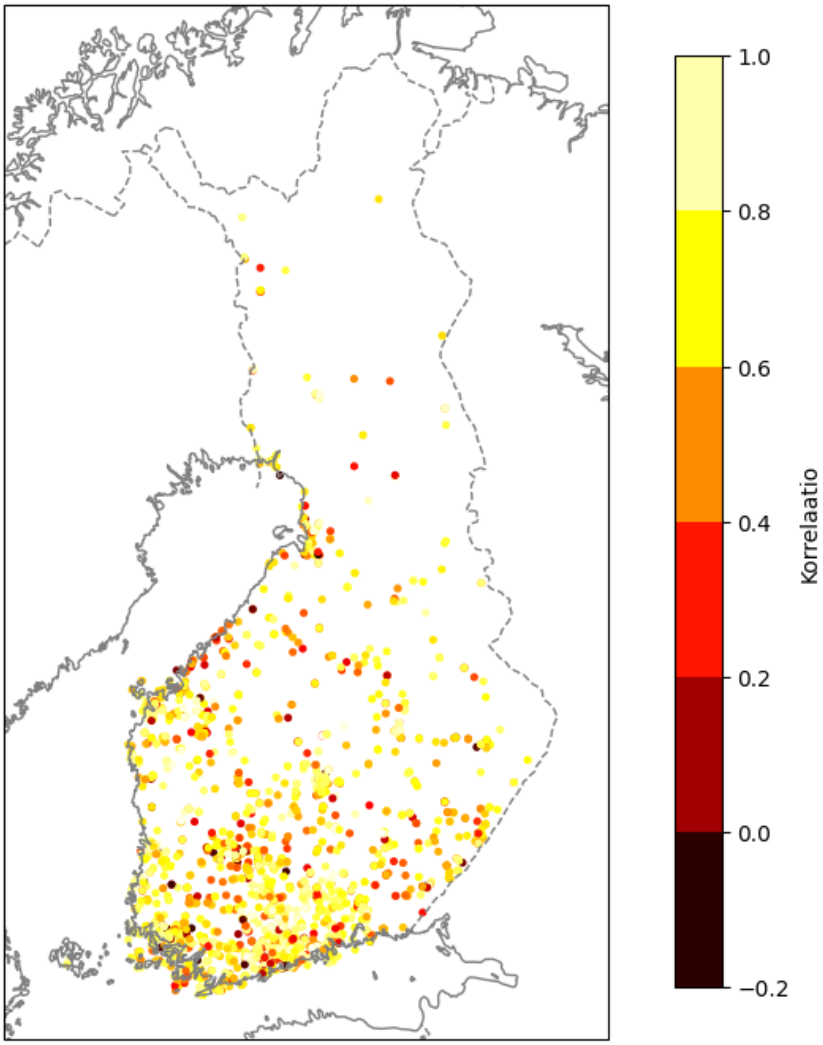


Kuva 6.4: Ilmatieteen laitoksen automaattisadeasemien väliset korrelaatiot etäisyyden funktiona (a) vuorokausisateille ja (b) tuntisateille. Punaiset käyrät on sovitettu pisteisiin.



Kuva 6.5: Netatmo-sadeasemien ja lähimpien Ilmatieteen laitoksen automaattisadeasemien väliset korrelaatiot etäisyyden funktiona pistein esitettynä (a) vuorokausisateille ja (b) tuntisateille. Punaiset käyrät on sovitettu pisteisiin. Vaaleanpunaiset käyrät ovat kuvan 6.4 sovitusfunktiot Ilmatieteen laitoksen automaattisadeasemien välisiin korrelaatioihin.

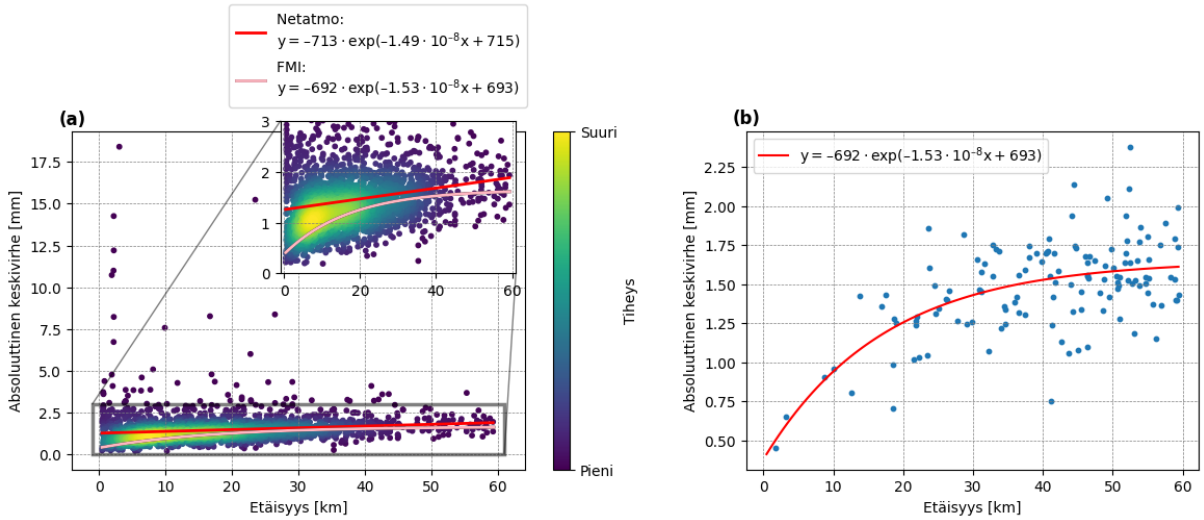
Kuvassa 6.6 on värein esitettynä Netatmo-asemien ja lähimpien Ilmatieteen laitoksen vertailuasemien välisten vuorokausisateiden korrelaatioiden alueellinen jakauma Suomen kartalla. Hyvän vuorokausisateiden korrelaation lähimmän vertailuaseman kanssa omaavia Netatmo-asemia näyttäisi kuvan 6.6 keltaisten pisteiden perusteella olevan tasaisesti ympäri Suomen. Vastaavasti myös huonon korrelaation tapaukset, kuvan punaiset pisteet, ovat hajallaan eri puolilla Suomea. Tämän perusteella Netatmo-sadehavainnot voisivat olla käyttökelpoisia, sillä käyttökohteiden kannalta hyvien ja huonojen asemien jakauma on tärkeä. Kun korrelaation perusteella hyviä Netatmo-asemia on alueella enemmän, eivät huonot asemat painotu niin paljon. Vastaavasti, kun huonot Netatmo-asemat ovat siellä täällä, niiden vaikutus käyttökelpoisuuteen ei ole kovin suuri.



Kuva 6.6: Netatmo-asemien ja lähimpien Ilmatieteen laitoksen asemien välisten vuorokausisateiden korrelaatioiden jakauma kartalla.

6.1.3 Absoluuttinen keskivirhe

Tarkastellaan seuraavaksi absoluuttista keskivirhettä, jotta saadaan tietoa siitä, kuinka paljon Netatmo-asemien havainnot eroavat Ilmatieteen laitoksen vertailuasemien havainnoista. Vuorokausisateiden absoluuttinen keskivirhe etäisyyden funktiona on esitetty kuvassa 6.7a Netatmo-asemien ja lähimpien Ilmatieteen laitoksen vertailuasemien välillä ja kuvassa 6.7b Ilmatieteen laitoksen asemien välillä. Punaiset käyrät kuvassa 6.7 ovat sovitusfunktioita ja vaaleanpunainen käyrä kuvassa 6.7a on kuvan 6.7b punainen käyrä. Kuvasta 6.7 nähdään, että absoluuttinen keskivirhe kasvaa asemien välisen etäisyyden kasvaessa. Ilmatieteen laitoksen asemien välillä absoluuttinen keskivirhe on kaikissa tapauksissa alle 2.5 mm. Suurin osa Netatmo-aseman ja lähimmän vertailuaseman välisistä absoluuttisista keskivirheistä on myös alle 2.5 mm. Kaikilla etäisyyksillä esiintyy kuitenkin suuria yli 3 mm absoluuttisia keskivirheitä Netatmo-asemien ja vertailuasemien välillä. Samoin kuin keskiarvojen erojen tapauksessa monet suurimmista absoluuttisista keskivirheistä esiintyvät, kun Netatmo-aseman ja vertailuaseman välinen etäisyys on lyhyt. Absoluuttisten keskivirheidenkin kohdalla syy tähän on luultavasti se, että suurin osa asemista sijaitsee näillä lyhyillä etäisyyksillä, kuten kuvan 6.7a keltaisen ja vihreän sävystä nähdään.

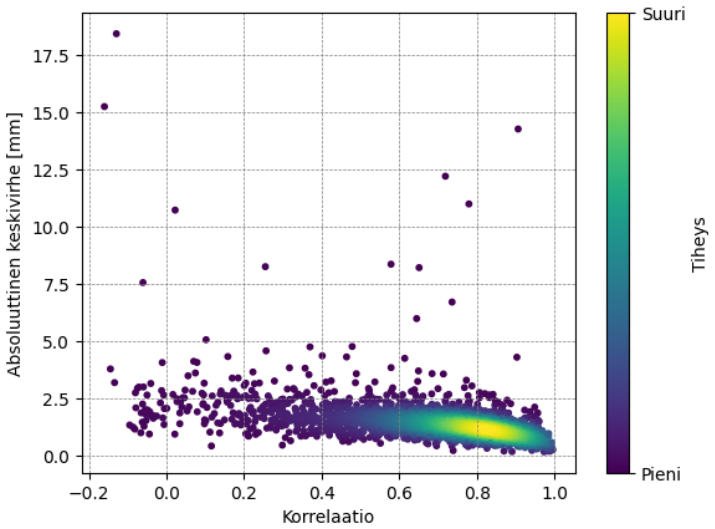


Kuva 6.7: Absoluuttinen keskivirhe etäisyyden funktiona vuorokausisateille (a) Netatmo-aseman ja lähimmän Ilmatieteen laitoksen aseman välillä ja (b) kahden Ilmatieteen laitoksen aseman välillä. Punaiset käyrät on sovitettu pisteisiin. Lisäksi kuvaan (a) on lisätty vaaleanpunaisella kuvan (b) punainen käyrä.

Kuvan 6.7a vaaleanpunaisen käyrän perusteella Netatmo-asemien ja vertailuasemien välinen absoluuttinen keskivirhe on useimmissa tapauksissa suurempi kuin Ilmatieteen laitoksen asemien välillä. Ero ei kuitenkaan ole niin suuri kuin korrelaation tapauksessa kuvassa 6.5a, sillä vaaleanpunaisen käyrän alapuolelle mahtuu paljon suurempi osa tapauksista kuvassa 6.7a kuin sen yläpuolella on kuvassa 6.5. Lisäksi kuvan 6.7a punainen käyrä

on varsin lähellä vaaleanpunaista käyrää. Punainen käyrä myös kulkee hieman keltaisen alueen eli asemien välisten absoluuttisten keskivirheiden suurimman keskittymän yläpuolelta. Tähän ovat syynä useat suuret absoluuttiset keskivirheet lähellä toisiaan sijaitsevien asemien välillä. Ilman näitä selkeästi virheellisiä tapauksia punainen käyrä kulkisi lähempänä keltaisen alueen keskikohtaa ja olisi siten lähempänä myös vaaleanpunaista käyrää.

Absoluuttisen keskivirheen ja havaintojen lukumäärän välillä ei ole havaittavissa selvää yhteyttä. Hieman yleisemmin absoluuttinen keskivirhe on suuri, jos havaintoja on vähän. Tämä on seurausta siitä, että pienellä otoskoolla sadehavaintojen suuret yksittäiset erot korostuvat. Absoluuttinen keskivirhe on kuitenkin keskimäärin saman suuruinen ja suuria absoluuttisia keskivirheitä esiintyy riippumatta havaintojen lukumäärästä.



Kuva 6.8: Netatmo-asemien ja lähimpien Ilmatieteen laitoksen vertailuasemien väliset vuorokausisadehavaintojen absoluuttiset keskivirheet korrelaation funktiona.

Koska keskiarvot ja absoluuttiset keskivirheet ovat läheisessä yhteydessä toisiinsa, absoluuttinen keskivirhe on suurempi tai yhtä suuri kuin keskiarvojen absoluuttinen ero. Keskimäärin absoluuttinen keskivirhe on sitä suurempi, mitä suurempi on keskiarvojen absoluuttinen ero. Korrelaatioiden ja absoluuttisten keskivirheiden välillä ei ole samankaltaista riippuvuussuhdetta. Täten korrelaatiota ja absoluuttista keskivirhettä vertaamalla saadaan hyvä käsitys havaintojen laadusta. Jos siis korrelaatio on suuri ja absoluuttinen keskivirhe pieni, mittaukset vastaavat hyvin toisiaan ja Netatmo-aseman havaintojen voidaan olettaa olevan hyviä. Netatmo-asemien ja lähimpien Ilmatieteen laitoksen vertailuasemien väliset absoluuttiset keskivirheet ja korrelaatiot vuorokausisateille on esitetty toistensa suhteen kuvassa 6.8. Keskiarvojen absoluuttisen eron ja korrelaation välinen riippuvuus on kuvan 6.8 kaltainen, joten sitä ei ole erikseen tässä esitetty. Suurin osa tapauksista sijoittuu kuvan 6.8 keltaisen ja vihreän värin perusteella hyvän korrelaation ja pienen absoluuttisen keskivirheen alueelle. Lisäksi kuvasta 6.8 nähdään, että absoluutti-

nen keskivirhe odotetusti hieman pienenee, kun korrelaatio paranee.

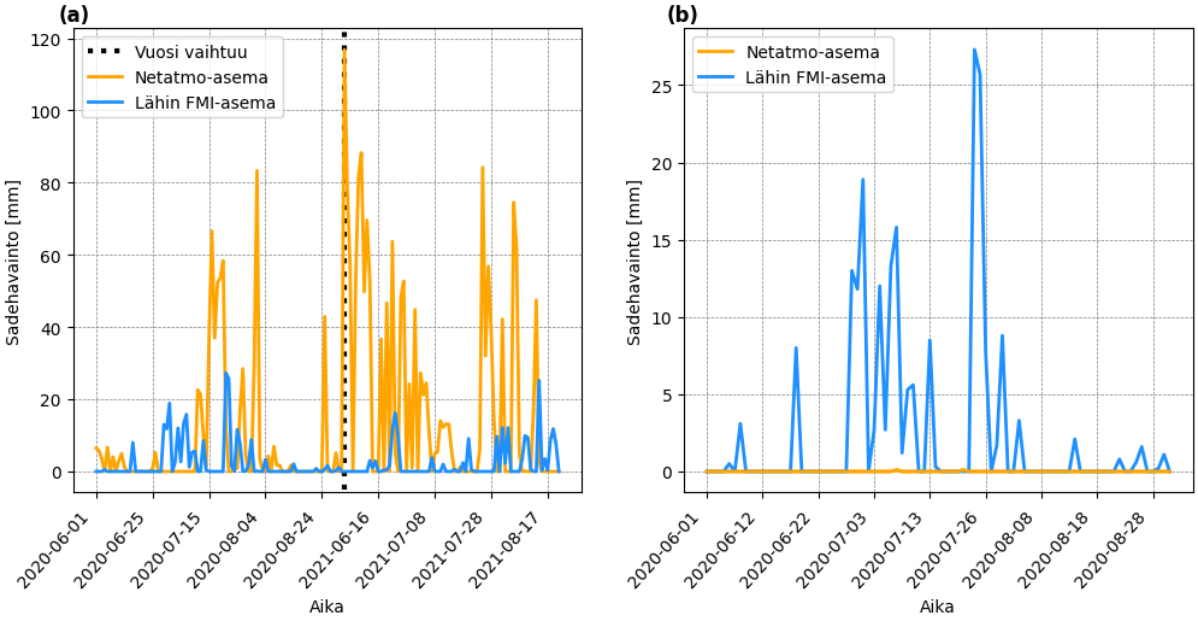
Selvästi virheellisinä tapauksina kuvasta 6.8 erottuvat pisteet, joissa absoluuttinen keskivirhe on hyvin suuri ja korrelaatio olematon. Kuvasta 6.8 erottuu kuitenkin myös useita tapauksia, joissa korrelaatio on hyvä mutta absoluuttinen keskivirhe suuri. Näissä tapauksissa mittaukset näyttävät korrelaation valossa hyviltä, mutta absoluuttinen keskivirhe paljastaa niiden virheellisyyden. Keskiarvojen eroa ja korrelaatiota verrattaessa havaittiin keskiarvojen eron ollessa suuri ja korrelaation hyvä, että osassa tapauksista Netatmo-aseman keskiarvo oli paljon suurempi kuin vertailuaseman ja toisissa tapauksissa taas vertailuaseman keskiarvo paljon suurempi kuin Netatmo-aseman. Tilanne voi olla toisinkin päin. Kuvan 6.8 vasemmassa alalaidassa on tapauksia, joissa absoluuttinen keskivirhe on pieni mutta korrelaatio on erittäin huono. Nämäkin mittaukset ovat siis selvästi virheellisiä. Täten yhdessä keskiarvot eroineen, korrelaatiot ja absoluuttiset keskivirheet antavat Netatmo-havaintojen laadusta parhaan kuvan. Seuraavaksi tarkastellaan näiden tilastollisten suureiden kautta Netatmo-asemien mittausten virheitä lähemmin ja pohditaan syitä niihin.

6.2 Netatmo-asemien mittausten erilaisia virheitä

Seuraavaksi käydään läpi Netatmo-asemien ja niiden lähimpien Ilmatieteen laitoksen vertailuasemien mittausten aikasarjojen avulla erilaisia Netamo-asemien virhemittauksia. Netatmo-mittauksista selvimmin virheellisiä ovat tapaukset, joissa kaikki tarkastellut tilastolliset suureet kertovat virheellisyydestä. Kaksi tällaista erilaista tapausta on esitetty kuvassa 6.9. Molemmissa tapauksissa Netatmo-aseman ja lähimmän Ilmatieteen laitoksen vertailuaseman keskiarvojen ero on suuri, korrelaatio huono ja absoluuttinen keskivirhe suuri.

Kuvassa 6.9a Netatmo-asema on mitannut sateita huomattavasti enemmän kuin vertailuasemat, minkä lisäksi sateet eivät osu samoille päiville ainakaan samassa suhteessa. Koska Netatmo-aseman sateet kuvassa 6.9a ovat keskimäärin hyvin suuria, asema ei todennäköisesti mittaa oikeaa sadetta. Netatmo-sademittari on saatettu esimerkiksi sijoittaa kukkapenkkiin tai kasvimaalle, jota kasteltaessa vettä on valunut myös Netatmo-sademittariin. Kuvasta 6.9a nimittäin nähdään, että Netatmo-asema on mitannut sadetta monesti silloin, kun vertailuasemalla ei ole satanut. Esimerkiksi Verkkokauppa mainostaa Netatmo-sademittaria kehottamalla mittaamaan sateen määrää ja sen perusteella päättämään, paljonko lisäkastelua kasvit tarvitsevat [Verkkokauppa, 2023]. Lisäksi kuvassa 2.1, joka on peräisin Netatmon omilta sivuilta, Netatmo-sademittari on kukkapenkissä [Netatmo, 2023]. Jos Netatmo-sademittari on mitannut kuvan 6.9a aikasarjan ajan oikeaa sadetta, se ei todennäköisesti sijaitse ilmoitetuissa koordinaateissa ja samalle se myös yliarvioi sademäärää. Sademäärän yliarviointi voi olla seurausta esimerkiksi virheellises-

tä kalibroinnista. Kuvan 6.9b tapauksessa Netatmo-sadeasema ei ole mitannut sadetta. Netatmo-sademittari kuvassa 6.9b on mitannut muutaman kerran 0.1 mm sadehavainnon, mutta kyseessä ei ole oikea sade. Pienet virhehavainnot ovat voineet syntyä esimerkiksi laitteen heilahtaessa. Netatmo-sademittari voi kuvan 6.9b tapauksessa olla tukossa tai vinossa niin, ettei sade pääse sinne sisälle.

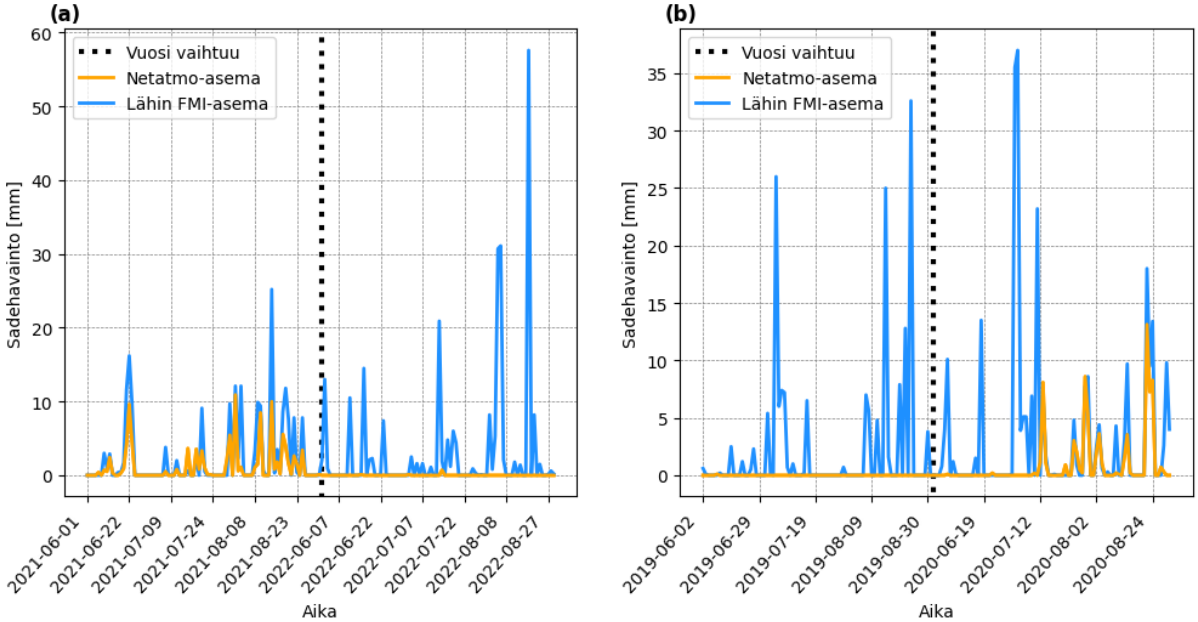


Kuva 6.9: Aikasarjat Netatmo-aseman ja lähimmän Ilmatieteen laitoksen aseman vuorokausisadehavainnoista, kun (a) Netatmo-asema on mitannut epämääräisiä hyvin suuria sateita ja (b) Netatmo-asema ei ole mitannut käytännössä lainkaan sadetta, vaikka vertailuasemalla on satanut.

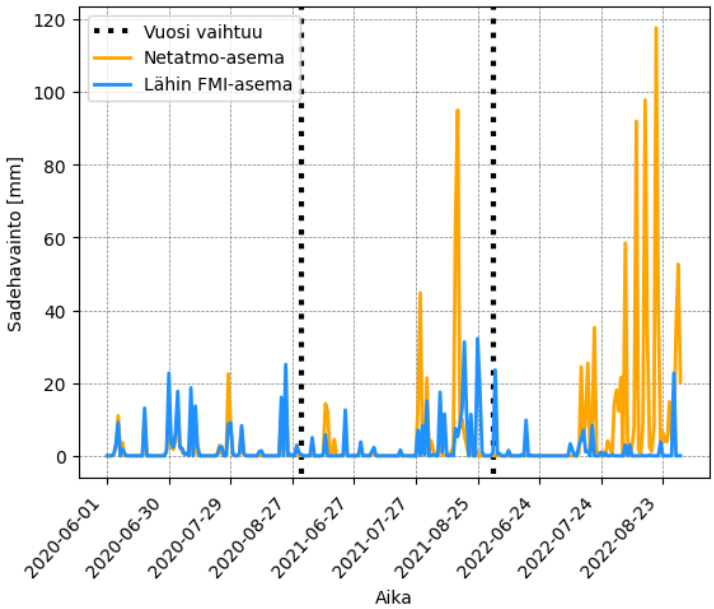
Aina suuri keskiarvojen ero, huono korrelaatio ja suuri absoluuttinen keskivirhe eivät tarkoita, että Netatmo-aseman tuottamat havainnot olisivat jatkuvasti huonoja. Kuvissa 6.10 ja 6.11 on esimerkkejä tällaisista tapauksista. Kuvan 6.10a tapauksessa Netatmo-sadeasema on mitannut ensimmäisen kesän ajan sateita hyvin, mutta seuraavana kesänä se ei enää ole mitannut sadetta. Syynä voi olla, että Netatmo-sademittari on otettu talveksi sisään ja unohdettu seuraavana kesänä laittaa takaisin ulos. Toinen mahdollinen syy on, että sademittari on tukkeutunut esimerkiksi lehdistä. Vastaavasti kuvassa 6.10b Netatmo-sadeasema ei ensin ole mitannut sadetta, mutta noin toisen kesän puolivälissä se alkoi tuottamaan hyviä sadehavaintoja. Tässä tapauksessa Netatmo-asema on saatettu asentaa ensin vinoon tai unohdettu sisälle. Sitten aseman omistaja on oletettavasti tajunnut, ettei se mittaa lainkaan sadetta ja korjannut tilanteen.

Kuvassa 6.11 puolestaan on tilanne, jossa Netatmo-asema mittasi aluksi hyvin mutta alkoi sitten mittaamaan paljon enemmän sateita kuin vertailuasema. Netatmo-sademittari on tässä tapauksessa saatettu kalibroida väärin niin, että se alkoi yliarvioimaan sateita. Koska mitatut sateet eivät lopuksi edes osu samoille päiville, on mahdollista, että sa-

deaseman sijainti on muuttunut, mutta aseman koordinaatteja ei ole vaihdettu. Kuvan 6.11 tapauksessa kyse voi myös olla virhemittauksista, jolloin suurten sadehavaintojen aiheuttaja ei ole sade.

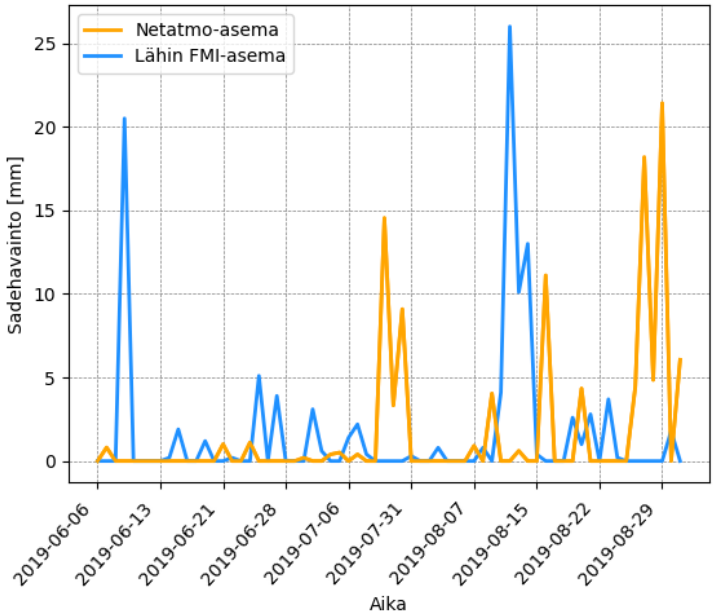


Kuva 6.10: Aikasarjat Netatmo-aseman ja lähimmän Ilmatieteen laitoksen aseman vuorokausisadehavainnoista, kun Netatmo-asema (a) mittasi aluksi hyvin mutta sitten ei käytännössä mitään ja (b) ei aluksi mitannut sadetta mutta alkoi sitten mitata hyvin.



Kuva 6.11: Aikasarja Netatmo-aseman ja lähimmän Ilmatieteen laitoksen aseman vuorokausisadehavainnoista, kun Netatmo-asema mittasi aluksi hyvin mutta alkoi sitten mitata epärealistisen suuria sadehavaintoja.

Välillä joku tilastollisista suureista saattaa näyttää hyvältä, vaikka toiset suureet kertovat Netatmo-aseman havaintojen virheellisyydestä. Kuvassa 6.12 Netatmo-aseman ja lähimmän vertailuaseman keskiarvojen ero on pieni, mutta korrelaatio huono ja absoluuttinen keskivirhe suuri. Tarkasteltavan mittausjakson aikana kuvan 6.12 Netatmo-asema ja sen lähin Ilmatieteen laitoksen asema ovat mitanneet keskimäärin saman verran sadetta mutta aivan eri aikoihin. Luultavasti tässä tapauksessa Netatmo-aseman koordinaatit eivät pidä paikkaansa. Netatmo-aseman sijainniksi on saatettu ilmoittaa Jyväskylä, vaikka todellisuudessa asema sijaitsee jossain aivan muualla. Pitkien aikojen sademäärät eri puolilla Suomea voivat olla samankaltaisia, mikä selittää keskiarvojen pientä eroa. Syynä virheellisiin koordinaatteihin voi olla väärin vahingossa tai osaamattomuutta katsotut ja ilmoitetut koordinaatit. Käyttäjä on myös voinut tahallaan ilmoittaa Netatmo-asemansa koordinaatit väärin.

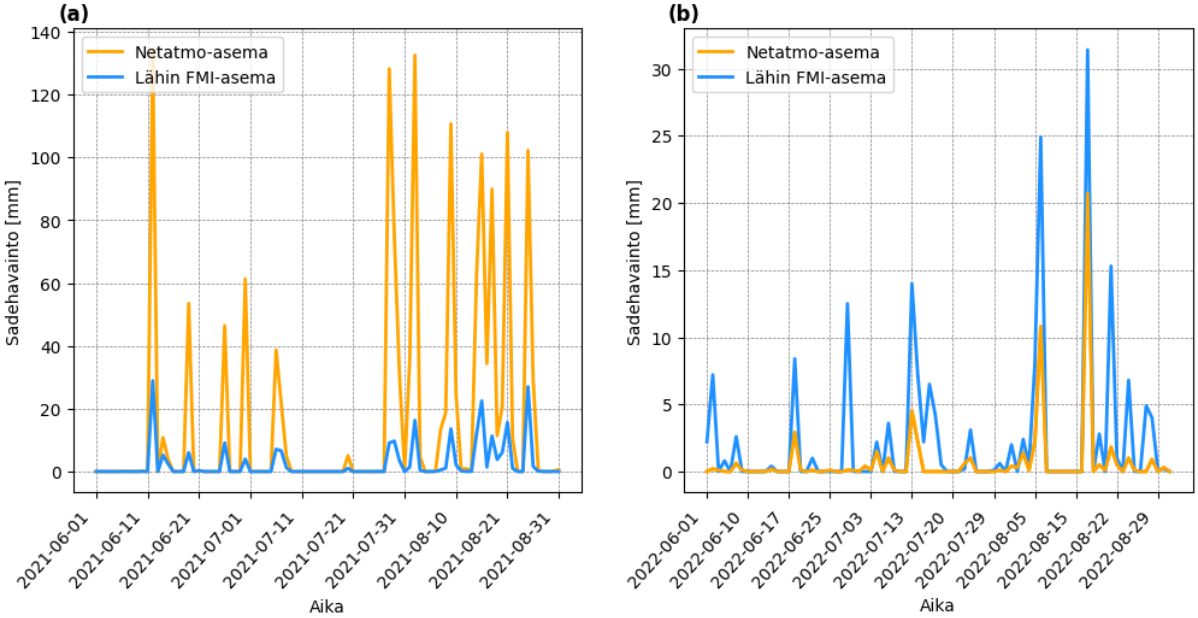


Kuva 6.12: Aikasarja Netatmo-aseman ja lähimmän Ilmatieteen laitoksen vertailuaseman vuorokausisadehavainnoista tapauksessa, jossa asemat ovat mitanneet sateita eri aikoihin mutta keskimäärin saman verran mittausjakson aikana.

Aineistosta löytyi useita Netatmo-sadeasemia, joilla oli samat koordinaatit. Rajatussa aineistossa oli yhteensä 138 Netatmo-sadeasemaa, joilla oli samat koordinaatit vähintään jonkun toisen Netatmo-sadeaseman kanssa. Netatmo-asemien sama sijainti voi johtua epätarkasta koordinaatin määrittämisestä. Esimerkiksi naapurit ovat voineet merkitä asemilleen samat koordinaatit. Saman koordinaatin tapaukset sijoittuvat pääsääntöisesti isoihin kaupunkeihin. Moni on voinut vain hakea asuinkaupunkinsa koordinaatteja ja merkitä ne sitten asemansa koordinaateiksi. Toki aseman omistaja on voinut tahallaan laittaa asemalleen Helsingin koordinaatit, vaikka asema sijaitsisi oikeasti Oulussa. Esi-

merkiksi yli kymmenellä Netatmo-sadeasemalla on samat koordinaatit, jotka osoittavat Kamppiin Helsingissä. Joidenkin Netatmo-sadeasemien koordinaatit taas osuvat merelle tai muihin vesistöihin, joten ne eivät ainakaan voi täysin pitää paikkaansa. Netatmo-asema saattaa silti sijaita vesistön lähellä.

Kuvan 6.13 tapauksissa puolestaan korrelaatio on hyvä, mutta keskiarvojen ero ja absoluuttinen keskivirhe ovat suuria. Näissä tapauksissa Netatmo-asema on mitannut sateita samaan aikaan ja samassa suhteessa mutta kuvan 6.13a tapauksessa paljon enemmän ja kuvan 6.13b tapauksessa paljon vähemmän kuin lähin Ilmatieteen laitoksen asema. Kuvan 6.13a tapauksessa kyse voi olla siitä, että Netatmo-asema yliarvioi sademäärää virheellisen kalibroinnin seurauksena. Toinen vaihtoehto on, että Netatmo-asema sijaitsee paikassa, kuten räystäskourun alla tai muuten talon seinän vieressä, jonne valuu suuremman alueen sademäärä. Kuvan 6.13b tapauksessa Netatmo-asema puolestaan aliarvioi sadetta. Tämä voi aiheutua esimerkiksi laitteen virheellisistä tehdasasetuksista. Toisaalta Netatmo-sademittari voi sijaita paikassa, jonne kaikki satava vesi ei pääse, kuten puiden suojassa.



Kuva 6.13: Aikasarjat Netatmo-aseman ja lähimmän Ilmatieteen laitoksen vertailuaseman vuorokausisadehavainnoista tapauksissa, joissa asemat ovat mitanneet sateita samoihin aikoihin ja samassa suhteessa mutta (a) Netatmo-asema on mitannut sateita paljon enemmän ja (b) Netatmo-asema on mitannut sateita paljon vähemmän kuin vertailuasema.

Joissain tapauksissa keskiarvojen eron ja absoluuttisen keskivirheen ollessa pieniä myös korrelaatio on pieni. Tämä voi aiheutua siitä, että sekä Netatmo- että vertailuasemalla on satanut vähän tarkasteltavan ajanjakson aikana. Lisäksi tarkasteltava ajanjakso voi olla lyhyt. Koska tässä tapauksessa sademäärät ovat vähäisiä, keskiarvojen erot ja absoluuttiset keskivirheet ovat pieniä. Osassa näistä tapauksista Netatmo-asema voi tuottaa

virheellisiä mittauksia, mutta voi myös olla, että mittaukset ovat hyviä, sillä vähäisistä sademääristä ja lyhyistä aikasarjoista on vaikea arvioida mittausten laatua.

Netatmo-sademittauksiin voi siis aiheutua virheitä monella eri tavalla. Näiden virheiden havaitsemiseen tarvitaan useampaa tilastollista suuretta, sillä joku suure saattaa näyttää hyvältä, vaikka toiset suureet kertoisivat Netatmo-aseman havaintojen virheellisyydestä. Koska kyse on kansalaisten sademittauksista, aseman sadehavaintojen laatu voi myös muuttua ajan myötä.

6.3 Sadehavaintojen suuruusjakauma

Tarkastellaan seuraavaksi sadehavaintojen suuruusjakaumaa. Taulukon 6.3 ”Erikseen”-sarakkeissa on esitetty vuorokausisateiden ja tuntisateiden jakauma Netatmo-asemilla ja lähimmillä Ilmatieteen laitoksen vertailuasemilla, kun sadetta on mitattu eli havainnot ovat yli 0 mm. Vertailun vuoksi taulukon 6.3 ”Molemmat”-sarakkeissa on vuorokausisateiden ja tuntisateiden jakauma, kun Netatmo-asemalla ja lähimmällä Ilmatieteen laitoksen asemalla on mitattu sadetta samaan aikaan. Netatmo-asemien vuorokausisadetapauksista 78 % ja tuntisadetapauksista 62 % oli sellaisia, joiden aikaan myös lähimmällä Ilmatieteen laitoksen asemalla satoi. Nämä prosenttiosuudet vaikuttavat hyviltä, sillä aivan kaikissa tapauksissa ei ole voinut sataa molemmilla asemilla sateiden paikallisuuden takia. Ilmatieteen laitoksen asemien keskinäisessä vertailussa taas vuorokausisadetapauksista 84 % ja tuntisadetapauksista 57 % oli sellaisia, joissa molemmilla asemilla satoi, kun toisella asemalla oli satanut. Koska prosenttiosuudet ovat lähellä Netatmo-asemien ja vertailuasemien prosenttiosuuksia, Netatmo-asemat näyttävät saavan sateita kiinni hyvin. Ilmatieteen laitoksen asemien sadetapausten prosenttisosuus on pienempi tuntisateiden tapauksessa, koska Ilmatieteen laitoksen asemien välinen etäisyys on keskimäärin suurempi kuin Netatmo-asemien ja vertailuasemien etäisyys. Koska vuorokausisateiden tapauksessa sadetapausten vastaavuus on Ilmatieteen laitoksen asemien välillä parempi, osa Netatmo-asemista ei välttämättä mittaa sadetta oikeassa sadetilanteessa.

Taulukosta 6.3 nähdään, että Netatmo-asemien mediaanit ovat pienempiä kuin niiden lähimpien Ilmatieteen laitoksen vertailuasemien. Sadehavaintojen mediaanit ovat taulukon 6.3 perusteella suurempia, kun molemmilla asemilla on satanut samaan aikaan. Tämä on luonnollista, sillä todennäköisesti sataa enemmän, kun sadetta tulee muuallakin kuin vain yhdessä pisteessä. Lisäksi vuorokausisateiden tapauksessa Netatmo-asemien ja lähimpien vertailuasemien mediaanit ovat lähempänä toisiaan, kun sekä Netatmo- että lähimmällä Ilmatieteen laitoksen asemalla on satanut samaan aikaan (taulukko 6.3). Luultavasti monissa tapauksissa, joissa sadetta on ollut vain Netatmo-asemalla tai lähimmällä vertailuasemalla, sademäärät ovat olleet pieniä.

Mediaanien lisäksi ala- ja yläkvartiilit sekä 90. persentiilit eli prosenttipisteet tau-

lukossa 6.3 ovat pienempiä Netatmo-sadehavaintojen tapauksessa kuin vertailuasemilla. Tämän perusteella Netatmo-asemat yleisesti aliarvioivat sademäärää, kuten jo aiemmin keskiarvoista nähtiin taulukoissa 4.1 ja 4.2 sekä kuvissa 6.1b ja 6.2. Lisäksi vuorokausisateiden tapauksessa mediaanien tavoin ala- ja yläkvartiili sekä 90. persentiili ovat suurempia ja lähempänä toisiaan, kun molemmilla asemilla on satanut samaan aikaan (taulukko 6.3). Tämä voi olla seurausta siitä, että sademäärät ovat vuorokauden aikana suurempia tapauksissa, joissa on satanut laajemmalla alueella käsittäen sekä Netatmo-aseman että vertailuaseman, sillä tällöin sade on todennäköisesti ollut pitkäkestoisempaa. Toisaalta on myös mahdollista, että Netatmo-asemat ovat mitanneet olemattomia sateita, kuten 0.1 mm pieniä virhemittauksia, vaikkei todellisuudessa eikä myöskään Ilmatieteen laitoksen vertailuasemalla ole satanut.

Taulukko 6.3: Sadehavaintojen suuruusjakauman tunnuslukuja Netatmo-asemille (Netatmo) ja lähimmille Ilmatieteen laitoksen sadeasemille (FMI), kun asemalla on satanut ("Erikseen"-sarakkeet) ja molemmilla asemilla on satanut samaan aikaan ("Molemmat"-sarakkeet).

	Vuorokausisateet (24 h)				**Tuntisateet (1 h)**			
	Erikseen		**Molemmat**		**Erikseen**		**Molemmat**	
	Netatmo	FMI	Netatmo	FMI	Netatmo	FMI	Netatmo	FMI
Minimi [mm]	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Alakvartiili (25 %) [mm]	0.4	0.7	0.8	0.9	0.1	0.2	0.2	0.3
Mediaani (50 %) [mm]	1.9	2.5	2.9	3.2	0.4	0.5	0.6	0.7
Yläkvartiili (75 %) [mm]	6.3	7.4	7.8	8.3	1.2	1.4	1.5	1.7
90. persentiili (90 %) [mm]	13.2	14.4	14.8	15.5	2.7	2.9	3.1	3.3
99. persentiili (99 %) [mm]	34.7	32.6	36.7	34.2	9.1	8.9	9.5	9.6
99.9. persentiili (99.9 %) [mm]	77.2	62.5	76.2	64.0	20.8	19.0	21.0	19.3
Maksimi [mm]	196.2	87.6	196.2	87.6	129.3	48.3	84.1	48.3

Tuntisateiden tapauksessa 0.1 mm/h sadehavaintoja oli 26 % kaikista Netatmo-sadehavainnoista ja 21 % kaikista lähimpien Ilmatieteen laitoksen asemien sadehavainnoista. Kun molemmilla asemilla oli satanut, tuntisadehavainnoista 0.1 mm/h oli Netatmo-asemien tapauksessa enää 16 % ja lähimpien vertailuasemien tapauksessa 14 %. Vuorokausisateiden tapauksessa ero on vielä suurempi. Kaikista Netatmo-asemien vuoro-

kausisateista 13 % oli 0.1 mm/vrk mutta lähimmän Ilmatieteen laitoksen asemien vuorokausisateita vain 7 %. Vastaavasti vuorokausisadehavainnoista 0.1 mm/vrk oli enää 6 % Netatmo-asemien ja 5 % vertailuasemien tapauksessa, kun molemmilla asemilla oli satanut. Täten kaikista Netatmo-asemien sadehavainnoista huomattavasti suurempi osa on 0.1 mm pieniä sadehavaintoja kuin kaikista vertailuasemien sadehavainnoista. Toisaalta prosenttiosuuksien ero tasoittuu, kun molemmilla asemilla on satanut samaan aikaan. Jotkut Netatmo-asemat ovat siis todennäköisesti mitanneet pieniä 0.1 mm virhemittauksia. Osa 0.1 mm havaintojen suuremmasta lukumäärästä Netatmo-asemilla selittyy kuitenkin niiden Netatmo-asemien havainnoilla, jotka systemaattisesti aliarvioivat sademäärää.

99. ja 99.9. persentiilit taulukossa 6.3 ovat puolestaan Netatmo-asemien sadehavaintojen tapauksessa suurempia kuin lähimpien Ilmatieteen laitoksen asemien tapauksessa. Erot 99. persentiileissä ovat vielä varsin pieniä, mutta 99.9. persentiilien kohdalla jo suurempia. Netatmo-asemien epärealistisen suuret sadetapaukset näkyvät siten kunnolla jakaumassa vasta 99.9 prosenttipisteen tienoilla, joten niiden osuus kaikista havainnoista on pieni. Muista prosenttipisteistä poiketen vuorokausisateiden 99.9. persentiili on Netatmo-asemien tapauksessa pienempi, kun molemmilla asemilla on satanut, verratuna tilanteeseen, jossa on kaikki Netatmo-asemien sadetapaukset. Näin ollen osa Netatmo-asemista on oletettavasti mitannut virheellisiä suuria sadehavaintoja, kun asemalla ei ole satanut.

Vuorokausisateiden tapauksessa Netatmo-asemien suurin sadehavainto, 196.2 mm, on mitattu samaan aikaan kuin myös vertailuasemalla on satanut, koska suurimmat Netatmo-vuorokausisadehavainnot ovat taulukon 6.3 ”Erikseen”- ja ”Molemmat”-sarakkeissa samat. Tuntisateiden tapauksessa taulukon 6.3 ”Erikseen”-sarakkeen suurin sadehavainto, 129.3 mm, on suurempi kuin ”Molemmat”-sarakkeen sadehavainto, 84.1 mm. Toisin sanoen Netatmo-aseman tuntisateiden suurin sadehavainto on mitattu, kun lähimmällä vertailuasemalla ei ole satanut. Täten kyseessä tuskin on oikea sadehavainto. Netatmo-sademittariin on saatettu esimerkiksi kaataa vettä. Lähimpien Ilmatieteen laitoksen asemien suurin vuorokausisadehavainto, 87.6 mm, ja tuntisadehavainto, 48.3 mm, on odotetusti mitattu samaan aikaan, kun vähintään jollain Netatmo-asemalla on satanut (taulukko 6.3). Netatmo-asemat ovat siis mitanneet sadetta ainakin näiden suurten sadetapausten aikaan. Seuraavaksi tarkastellaan lähemmin, kuinka hyvin Netatmo-asemat mittaavat suuria sateita.

6.4 Suuret sadehavainnot

Netatmo-sadehavaintojen käyttökelpoisuutta selvitettäessä suuret sadehavainnot ovat mielenkiintoisia, sillä osa suurista paikallisista kesäisistä konvektiosateista jää mittaamatta harvan havaintoverkon tähden. Taulukon 6.4 vasemmalla puolella on esitetty prosenttiosuudet Netatmo-asemista, joilla on satanut vähintään 5 mm tai 10 mm vuorokau-

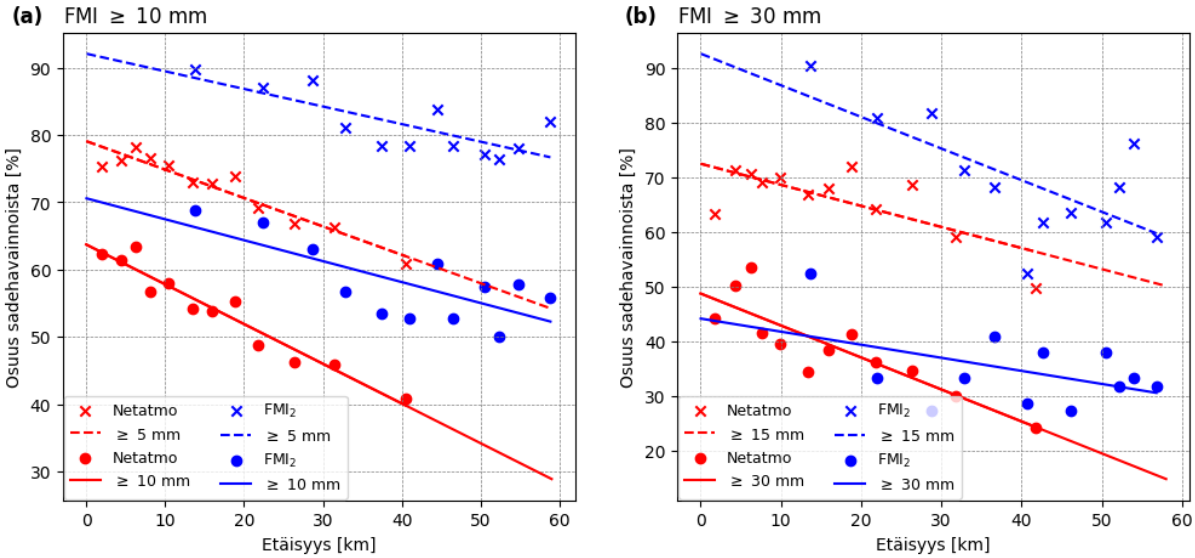
dessa, kun lähimmällä Ilmatieteen laitoksen asemalla on satanut vähintään 10 mm samana aikana. Lisäksi taulukon 6.4 vasemmalla puolella on vastaavat vuorokausisateiden prosenttiosuudet Ilmatieteen laitoksen asemille keskenään. Kun lähimmällä Ilmatieteen laitoksen asemalla on satanut vähintään 10 mm vuorokaudessa, Netatmo-asemista 54 % on mitannut sadehavainnoksi vähintään 10 mm ja 72 % vähintään 5 mm (taulukko 6.4). Ilmatieteen laitoksen asemille keskenään vastaavat prosenttiosuudet ovat hieman suuremmat. Taulukosta 6.4 nähdään, että toisista Ilmatieteen laitoksen asemista 58 % on mitannut vähintään 10 mm ja 82 % vähintään 5 mm vuorokausisateen, kun ensimmäisillä Ilmatieteen laitoksen asemilla on satanut vähintään 10 mm.

Taulukko 6.4: Taulukossa on vuorokausisateiden prosenttiosuudet Netatmo-asemilla, joilla on satanut vähintään 10 mm tai 5 mm, kun lähimmällä Ilmatieteen laitoksen asemalla (FMI) on satanut vähintään 10 mm. Netatmo-asemille on myös prosenttiosuudet vähintään 15 mm tai 30 mm sateille, kun lähimmällä Ilmatieteen laitoksen asemalla on satanut vähintään 30 mm. Vertailun vuoksi taulukossa on lisäksi vastaavat prosenttiosuudet Ilmatieteen laitoksen asemille (FMI_2), kun toisella Ilmatieteen laitoksen asemalla (FMI) on satanut vähintään 10 mm tai 30 mm. Luku N kertoo FMI-tapausten, vähintään 10 mm tai 30 mm, kokonaismäärän, josta prosenttiosuus on laskettu.

	FMI $\geq$ 10 mm		**FMI $\geq$ 30 mm**
Netatmo $\geq$ 10 mm	54 % (N = 23 106)	**Netatmo $\geq$ 30 mm**	39 % (N = 1 880)
FMI_2 $\geq$ 10 mm	58 % (N = 3 046)	**FMI_2 $\geq$ 30 mm**	35 % (N = 257)
Netatmo $\geq$ 5 mm	72 % (N = 23 106)	**Netatmo $\geq$ 15 mm**	66 % (N = 1 880)
FMI_2 $\geq$ 5 mm	82 % (N = 3 046)	**FMI_2 $\geq$ 15 mm**	70 % (N = 257)

Vastaavasti taulukon 6.4 oikealla puolella on vuorokausisateiden prosenttiosuudet Netatmo-asemista, joilla on satanut vähintään 15 mm tai 30 mm, kun lähimmällä Ilmatieteen laitoksen asemalla on satanut vähintään 30 mm. Taulukon 6.4 oikealla puolella on myös vastaavat vuorokausisateiden prosenttiosuudet Ilmatieteen laitoksen asemille keskenään. Kun lähimmällä Ilmatieteen laitoksen asemalla on satanut vähintään 30 mm vuorokaudessa, Netatmo-asemista 39 % oli myös mitannut vähintään 30 mm ja 66 % vähintään 15 mm sadetta vuorokaudessa. Ilmatieteen laitoksen asemille prosenttiosuudet ovat samankaltaisia. Kun Ilmatieteen laitoksen asemilla oli satanut vähintään 30 mm vuorokaudessa, toisista Ilmatieteen laitoksen asemista 35 % oli mitannut sademääräksi vähintään 30 mm ja 70 % vähintään 15 mm. Näin ollen suurten sadetapausten prosenttiosuus Netatmo-asemilla on suurempi kuin toisilla Ilmatieteen laitoksen asemilla, kun lähtöehtona on vähintään 30 mm vuorokausisade Ilmatieteen laitoksen asemilla. Tähän ja muutenkin varsin pieniin prosenttiosuuksien eroihin on oletettavasti syynä Ilmatieteen laitoksen automaattisadeasemien vähäisempi lukumäärä, jonka seurauksena suurten sateiden tapauksia on taulukon 6.4 perusteella Ilmatieteen laitoksen asemien keskinäisessä vertailuaineistossa huomattavasti vähemmän kuin Netatmo-asemien ja lähimpien Ilma-

tieteen laitoksen asemien aineistossa. Toinen syy on, että Ilmatieteen laitoksen asemien välinen etäisyys on keskimäärin suurempi kuin Netatmo-asemien ja lähimpien vertailuasemien.



Kuva 6.14: Kuvaajissa on x-akselilla sadeasemien välisen etäisyyden mediaani kahdessatoista yhtä suuressa etäisyyden mukaan järjestetyssä luokassa ja y-akselilla prosenttiosuus Netatmo-asemien ja Ilmatieteen laitoksen sadeasemien (FMI_2) sadehavainnoista, jotka (a) ovat vähintään 5 mm ja 10 mm, kun Ilmatieteen laitoksen vertailuasemalla (FMI) on satanut vähintään 10 mm ja (b) ovat vähintään 15 mm ja 30 mm, kun Ilmatieteen laitoksen vertailuasemalla (FMI) on satanut vähintään 30 mm saman vuorokauden aikana.

Ilmatieteen laitoksen asemien välinen suurempi etäisyys nähdään kuvan 6.14 havaintopisteistä. Kuvassa 6.14 on esitetty suurten sadehavaintojen tapausten prosenttiosuuksia etäisyyden funktiona taulukon 6.4 arvoille. Sadehavainnot on kuvassa 6.14 järjestetty kahteentoista yhtä suureen luokkaan etäisyyden perusteella ja kunkin pisteen etäisyydeksi on valittu kyseisen luokan etäisyyden mediaani. Näistä pisteistä nähdään, että punaisella merkityistä Netatmo-asemista suurin osa sijaitsee paljon lähempänä vertailuasemaa kuin sinisellä merkityt Ilmatieteen laitoksen asemat toista Ilmatieteen laitoksen asemaa. Tämän takia varsinkin kuvassa 6.14b, jossa on huomattavasti vähemmän havaintoja kuin kuvassa 6.14a, Ilmatieteen laitoksen asemien prosenttiosuuksissa on paljon hajontaa. Tosin havaintojen vähyyden takia myös Netatmo-asemien tapauksessa kuvassa 6.14b on enemmän hajontaa kuin kuvassa 6.14a. Taulukosta 6.4 nähdään, että oikeastaan vain kuvan 6.14a Netatmo-asemien ja lähimpien vertailuasemien tapauksessa tarkasteltavana on kohtuullisen suuri havaintojoukko, kun taas muissa tapauksissa havaintojoukko on varsin pieni. Tämän takia yksittäiset tapaukset korostuvat. Hajonnasta huolimatta toinen Ilmatieteen laitoksen asema mittaa suuria sateita, kun toisella asemalla on niitä mitattu, pääsääntöisesti useammin kuin Netatmo-asema vertailuaseman niitä mitatessa

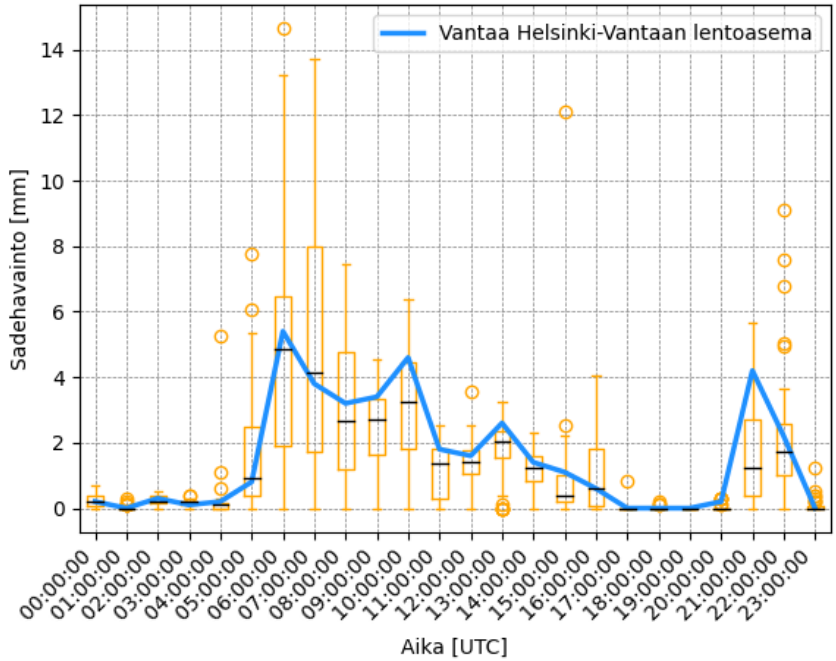
(kuva 6.14).

Kuvasta 6.14 nähdään myös, että suurien sateiden mittaamisen prosenttiosuus toisella asemalla pienenee asemien välisen etäisyyden kasvaessa. Tämä on luonnollista. Pisteisiin sovitettujen suorien perusteella prosenttiosuus vähenee vaihtelevalla nopeudella etäisyyden kasvaessa (kuva 6.14). Netatmo-asemien suurten sadehavaintojen mittaamisen prosenttiosuus vähenee kuvassa 6.14a nopeammin kuin Ilmatieteen laitoksen asemien. Kuvassa 6.14b pisteiden hajonta on niin suurta ja tapausten määrä niin pieni, ettei sovitussuorien jyrkkyydestä voi päätellä mitään varmaa. Taulukon 6.4 ja kuvan 6.14 perusteella Netatmo-asemat näyttävät saavan suuria sateita hieman huonommin kiinni kuin Ilmatieteen laitoksen asemat. Kokonaisuudessaan erot ovat silti niin pieniä, että ne voivat johtua luonnollisista syistä, kuten sateiden paikallisuudesta. Toisaalta taas huonompaa tulosta Netatmo-asemille voi aiheuttaa myös Netatmo-havaintoaineistossa vielä olevat virheellisiä sadehavaintoja tuottavat asemat. Kaikkiaan Netatmo-asemat mittaavat silti suuria sateita varsin hyvin.

6.5 Tapaustutkimus 30.6.2020

Tarkastellaan vielä lopuksi tapaustutkimuksen kautta Netatmo-sadeasemien mittauksia todellisessa sadetilanteessa. Tapaustutkimuksen ajankohdaksi valikoitui 30.6.2020, jolloin laaja sadealue liikkui Suomen yli ja monin paikoin mitattiin suuria sademääriä. Tarkasteltavaksi valittiin Ilmatieteen laitoksen Vantaa Helsinki–Vantaan lentoaseman automaattisadeasema ja sen lähistöllä olevat Netatmo-sadeasemat. Ilmatieteen laitoksen sadeasema Helsinki–Vantaan lentokentällä on noin 40 aineiston Netatmo-sadeaseman lähin vertailuasema ja kaikki Netatmo-sadeasemat ovat alle 20 km etäisyydellä Ilmatieteen laitoksen Helsinki–Vantaan lentoaseman havaintopisteestä.

Kuvassa 6.15 on esitettynä Netatmo-asemien tuntisadehavaintojen jakauma ja niiden lähimmän Ilmatieteen laitoksen vertailuaseman Vantaa Helsinki–Vantaan lentoaseman tuntisadehavaintojen aikasarja 30.6.2020. Tutkahavaintojen perusteella esimerkiksi välillä 09:00–10:00 UTC satoi varmasti kaikkialla 20 km säteellä Helsinki–Vantaan lentoaseman havaintoasemasta, mutta kuvan 6.15 perusteella kaikki Netatmo-asemat tällä etäisyydellä eivät ole mitanneet sadetta. Edes koko vuorokauden 30.6.2020 aikana kaikki Netatmo-asemat, joille Helsinki–Vantaan lentoaseman sadeasema on lähin vertailuasema, eivät olleet mitanneet sadetta. Laajan rintamasadealueen liikkuessa alueen yli näillä Netatmo-asemilla olisi pitänyt sataa. Netatmo-asemat, jotka eivät ole mitanneet sadetta todellisessa sadetilanteessa voivat olla esimerkiksi tukossa, vinossa tai sisätiloissa. Mahdollista on myös, että Netatmo-aseman sijainti on ilmoitettu väärin eikä Netatmo-aseman oikeassa sijaintipaikassa ole satanut. Tämä on kuitenkin epätodennäköistä, sillä sadetta tuli laajasti melkein koko Suomessa vuorokauden 30.6.2020 aikana.



Kuva 6.15: Sininen käyrä on Ilmatieteen laitoksen havaintoaseman Vantaa Helsinki–Vantaan lentoasema tuntisadehavaintojen aikasarja 30.6.2020. Oransseilla laatikko-janakuvioilla on esitettynä Netatmo-asemien, joiden lähin Ilmatieteen laitoksen asema on Vantaa Helsinki–Vantaan lentoasema, tuntisadehavaintojen jakauma joka tunnille 30.6.2020. Mustat viivat laatikoiden keskellä ovat mediaaneja. Kunkin laatikon alaraja on alakvartiili Q_1 ja yläraja yläkvartiili Q_3, joten laatikon sisällä on 50 %, alapuolella 25 % ja yläpuolella 25 % havainnoista. Laatikoiden ylä- ja alapuolella olevat oranssit janat ovat raja-arvoja. Alempi raja-arvo on enintään $Q_1 - 1.5 \times (Q_3 - Q_1)$ ja ylempi korkeintaan $Q_3 + 1.5 \times (Q_3 - Q_1)$. Raja-arvojen ulkopuolelle jäävät poikkeavat havainnot on merkitty oransseilla ympyröillä.

Ilmatieteen laitoksen sadehavaintojen sininen käyrä kuvassa 6.15 kulkee pääasiassa Netatmo-asemien sadehavaintojen mustien mediaaniviivojen kautta tai niiden yläpuolelta. Lisäksi 30.6.2020 Netatmo-asemien vuorokausisadehavaintojen mediaani on 32.8 mm, kun taas Helsinki–Vantaan lentoaseman vuorokausisadehavainto on 35.5 mm. Tämän perusteella Netatmo-asemat näyttävät keskimäärin mitanneen hieman vähemmän sadetta. Tähän ovat oletettavasti syynä Netatmo-asemat, jotka eivät ole sadetilanteessa mitanneet sadetta. Pääasiassa Netatmo-asemat ovat kuitenkin mitanneet sateita realistisesti. Kun Helsinki–Vantaan lentoasemalla on mitattu yli 3 mm tuntisademääriä välillä 06–10 UTC, kuvan 6.16 oranssien laatikoiden alarajoista eli alakvartiileista nähdään, että Netatmo-asemista yli 75 % on tällä aikavälillä mitannut yli 1 mm tuntisateita. Näin ollen ainakin 75 % tapaustutkimuksen Netatmo-asemista näyttää mitanneen sateita hyvin.

Asemien sademäärissä esiintyy kuvan 6.16 perusteella luonnollista hajontaa. Hajontaa on vuorokausisademäärissäkin, sillä vuorokausisateeksi Netatmo-asemista 75 % on mitannut vähintään 13.2 mm (alakvartiili) ja 25 % puolestaan vähintään 42.7 mm (yläkvartiili). Kuvasta 6.16 nähdään myös, että joillain Netatmo-asemilla on mitattu suuria yli

12 mm tuntisademääriä. Tämän tapaustutkimuksen Netatmo-asemien suurimmaksi vuorokausisateeksi on puolestaan mitattu 73.6 mm. Suurimmat Netatmo-asemien mittaamat tunti- ja vuorokausisademäärä ovat yli kaksinkertaisia Helsinki–Vantaan lentoaseman sadehavaintohin verrattuna. Näin suuret sademäärien paikalliset erot ovat tarkasteltavassa sadetilanteessa periaatteessa mahdollisia mutta varsin epätodennäköisiä, sillä kyseessä on pitkäkestoinen sadealue. Suuret sadehavainnot ja sademäärien hajonta kertovat, että Netatmo-asemat pystyvät mittaamaan suuria sateita todellisessa sadetilanteessa ja voivat antaa lisätietoa sademäärien vaihtelusta pienenkin alueen sisällä. Osa sademäärien hajonnasta voi silti johtua Netatmo-mittausten virheistä.

7. Johtopäätökset

Suomen virallisten sademittarien havaintoverkko on harva ja sen laajentaminen vaatisi enemmän resursseja. Yksi keino laajentaa havaintoverkkoa ja siten muun muassa parantaa sääennusteiden laatua olisi kansalaisten sadehavaintojen hyödyntäminen. Tässä tutkielmassa selvitettiin, olisivatko kansalaisten Netatmo-sadehavainnot käyttökelpoisia kesällä Suomen alueella. Vuosien 2019–2022 kesä-, heinä- ja elokuun Netatmo-sadehavaintojen laatua tutkittiin vertailemalla niitä Ilmatieteen laitoksen automaattisadeasemien havaintoihin.

Ennen varsinasta analyysiä aineistoa selkeytettiin ja siitä poistettiin selvästi virheellisiä havaintoja ja sadeasemia. Asema rajattiin pois analyysistä, jos jokin seuraavista kriteereistä täyttyi. Aseman vuorokausisadehavaintoja oli alle 31, tuntisadehavaintoja oli alle 721, havaintojen keskiarvo oli 0 mm ja lähimmän vertailuaseman etäisyys oli yli 60 km. Lisäksi yksittäinen havaintoarvojen rivi poistettiin tarkasteltavasta aineistosta, jos vuorokausisademäärä oli yli 200 mm, tuntisademäärä yli 150 mm tai jokin vertailuaseman havainto puuttui.

Vertailua tehtiin tilastollisia suureita, keskiarvoja, korrelaatioita ja absoluuttisia keskivirheitä, hyödyntäen. Keskimäärin Netatmo-sadehavainnot näyttävät näiden tilastollisten suureiden valossa hyviltä, mutta osa Netatmo-asemista tuottaa selvästi Ilmatieteen laitoksen vertailuasemista poikkeavia virheellisiä havaintoja. Tästä kertovat muun muassa pienet korrelaatiot ja suuret absoluuttiset keskivirheet lähellä toisiaan sijaitsevien asemien välillä.

Keskiarvojen ja sadehavaintojen suuruusjakauman perusteella osa Netatmo-asemista on mitannut paljon suurempia sademääriä kuin Ilmatieteen laitoksen vertailusemat, mutta useimmiten Netatmo-asemien sademäärät ovat pienempiä kuin Ilmatieteen laitoksen lähimpien vertailuasemien. Näin ollen Netatmo-sademittarit yleisesti aliarvioivat sademäärää. Netatmo-sademittareista puuttuva tuulisuojus voi olla yksi tekijä sademäärän aliarvion takana.

Korrelaatiot Netatmo-asemien ja lähimpien Ilmatieteen laitoksen asemien välillä taas vaikuttavat pääsääntöisesti hyviltä. Suurin osa näistä korrelaatioista on nimittäin lähellä vastaavia Ilmatieteen laitoksen asemien välisiä korrelaatioita. Vaikka korrelaatiot ovat silti enimmäkseen huonompia Netatmo-asemien ja Ilmatieteen laitoksen vertailua-

semien välillä kuin samalla etäisyydellä sijaitsevien Ilmatieteen laitoksen asemien välillä keskimäärin, näin voidaan olettaakin olevan. Ilmatieteen laitoksen automaattisademittarit nimittäin ovat lähtökohtaisesti tarkempia kuin Netatmo-sademittarit ja lisäksi näiden sademittarien toimintaperiaatteet ovat erilaiset. Täten ei voida olettaa, että Netatmo-asemien ja Ilmatieteen laitoksen asemien vertailun tulokset olisivat yhtä hyviä kuin Ilmatieteen laitoksen asemien keskinäisessä vertailussa. Samankaltaiset tulokset saatiin absoluuttisia keskivirheitäkin tarkastelemalla.

Netatmo-asemien ja lähimpien Ilmatieteen laitoksen vertailuasemien välisten vuorokausisateiden korrelaatioiden alueellinen jakauma vaikuttaa myös hyvältä, sillä hyvät ja huonot korrelaatiot ovat jakautuneet varsin tasaisesti eri puolille Suomea. Korrelaatioiden alueellisen jakauman perusteella Netatmo-sadehavainnot ovat käyttökelpoisia, sillä huonojen Netatmo-asemien havaintojen vaikutus ei ole suuri, kun ne ovat hajallaan ympäri Suomen ja kaikkialla on enemmän hyviä asemia. Mahdollisimman suuri osa huonoista Netatmo-asemista täytyisi kuitenkin pyrkiä karsimaan tilastollisen vertailun perusteella ennen Netatmo-havaintojen käyttämistä.

Netatmo-asemien sademittauksista löytyi tilastollisten suureiden avulla monia erilaisia virheitä. Näiden virheiden havaitsemiseen tarvittiin useampaa tilastollista suuretta, sillä joku suure saattoi näyttää hyvältä, vaikka toiset suureet kertoivat Netatmo-aseman havaintojen virheellisyydestä. Jotkin Netatmo-asemat olivat mitanneet epärealistisen suuria sateita ja välillä vielä eri aikoihin kuin vertailuasemilla oli satanut. Syynä voi olla Netatmo-sademittarin sijainti esimerkiksi kasvimaalla, jolloin siihen on päätynyt vettä kasveja kasteltaessa. Toisinaan taas Netatmo-sademittarit eivät olleet mitannet lainkaan sadetta pieniä virhemittauksia lukuun ottamatta. Tällöin Netatmo-sademittari on saattanut olla sisätiloissa, tukossa tai vinossa niin, ettei se ole pystynyt mittaamaan sadetta. Pienet virhemittaukset puolestaan ovat voineet syntyä esimerkiksi laitteen heilahtaessa. Sadehavaintojen suuruusjakaumaa tarkasteltaessa havaittiin, että Netatmo-sadeasemat ovat mitanneet selvästi enemmän pieniä 0.1 mm sadehavaintoja kuin niiden lähimmät vertailuasemat. Koska suurin osa näistä pienistä sadehavainnoista on mitattu, kun vertailuasemalla ei ole satanut, osa niistä on virhemittauksia, vaikka osa selittyykin Netatmo-sademittarien sademäärän aliarvioimisella.

Jotkut Netatmo-sademittarit olivat mitanneet sateita keskimäärin saman verran kuin vertailuasematkin, mutta aivan eri aikoihin. Syynä on luultavasti Netatmo-asemien virheelliset koordinaatit, joiden takia asemat ovat mitanneet sateita hyvin, mutta eri paikoissa kuin ilmoitetut sijantitiedot antavat ymmärtää. Osassa tapauksista taas Netatmo-asemat ovat mitanneet sateita samaan aikaan ja samassa suhteessa kuin vertailuasemat, mutta joko paljon enemmän tai vähemmän. Sademäärän yliarvioon Netatmo-asemilla voi tällaisessa tapauksessa olla syynä laitteen virheellinen kalibrointi tai sijainti esimerkiksi räystäskourun alla, jolloin mittariin valuu sademäärä suuremmalta aleelta. Sademäärän

aliarviointi voi puolestaan aiheutua virheellisistä tehdasasetuksista tai sijainnista paikassa, jossa kaikki satava vesi ei päädy sademittariin. Netatmo-aseman havaintojen laatu voi myös muuttua ajan myötä, koska kyse on kansalaisten sademittauksista. Joissain tapauksissa Netatmo-asema mittasi ensin sateita hyvin mutta sitten hyvin suuria epämääräisiä havaintoja tai ei sadetta lainkaan. Vastaavasti eräs asema ei aluksi mitannut mitään, mutta alkoi sitten tuottamaan hyviä havaintoja.

Suuria sateita Netatmo-asemat mittaavat varsin hyvin. Ilmatieteen laitoksen asemat saavat asemien keskinäisen vertailun perusteella suuria sateita kiinni vain hieman paremmin. Todennäköisesti ero olisi suurempi, jos Ilmatieteen laitoksen asemat sijaitsisivat lähempänä toisiaan, sillä etäisyys niiden välillä on keskimäärin suurempi kuin Netatmo-aseman ja lähimmän Ilmatieteen laitoksen vertailuaseman välillä.

Tapaustutkimuksen kautta puolestaan selvisi, että laajan rintamasadealueen ollessa Suomen yllä 30.6.2020, kaikki Netatmo-asemat eivät mitanneet sadetta, vaikka alueella varmasti satoi. Pääsääntöisesti Netatmo-asemat kuitenkin mittasivat sateita hyvin ja asemien sademäärissä esiintyi luonnollista hajontaa suhteessa niiden lähimmän Ilmatieteen laitoksen aseman havaintoihin. Koska tapaustutkimuksessa Netatmo-asemat mittasivat suuria sadehavaintoja ja sademäärissä esiintyi hajontaa, Netatmo-asemat pystyvät mittaamaan suuria sateita todellisessa sadetilanteessa ja voivat antaa lisätietoa sademäärien vaihtelusta pienenkin alueen sisällä. On silti huomioitava, että osa Netatmo-asemien välisestä vaihtelusta voi johtua Netatmo-mittausten virheistä.

Kaikkiaan Netatmo-kesäsadehavainnot vaikuttavat käyttökelpoisilta, sillä suurin osa havainnoista on hyviä. Joukossa on kuitenkin lukuisia virheellisiä Netatmo-asemia ja yksittäisiä havaintoja vielä alkurajaustenkin jälkeen. Tämän takia Netatmo-sademittaukset tarvitsevat laadun varmistusta ennen niiden käyttöönottoa. Laadun varmistuksessa Netatmo-havaintoja voisi Ilmatieteen laitoksen havaintoihin tehtävän tilastollisen vertailun lisäksi verrata toisiinsa. Tällöin tietyllä alueella olevan Netatmo-aseman huonot havainnot pystytään karsimaan pois, jos ne eroavat selvästi alueen muista Netatmo-havainnoista. Tämä on hyödyllinen menetelmä siinäkin mielessä, että Netatmo-asemia on enemmän ja monesti tiheämmässä kuin virallisia Ilmatieteen laitoksen sadeasemia. Tulevaisuudessa Netatmo-sadeasemien määrä oletettavasti myös kasvaa Suomessa, kun entistä useampi hankkii laitteen. Ilmatieteen laitos käyttää Netatmo-lämpötilahavaintojen laadun varmistukseen TITAN-ohjelmaa. TITAN-ohjelmassa laadun varmistus perustuu läheisten havaintojen vertailuun samalla ajanhetkellä, mutta tällä hetkellä se ei ota huomioon asemien aikasarjoista laskettuja tilastollisia suureita tai vertailuja [Båserud et al., 2020]. Tämän tutkimuksen tulosten perusteella Netatmo-sadehavaintojen kohdalla olisi kuitenkin tarpeellista ottaa myös aikasarjojen tilastot mukaan laadun varmistukseen. Netatmo-kesäsadehavainnot voivat siis olla hyvä lisä Ilmatieteen laitoksen virallisiin sadehavaintoihin, mutta vain kattavan laadun varmistuksen kautta.

Lähdeluettelo

[Bárdossy et al., 2021] Bárdossy, A., Seidel, J., ja Hachem, A. E. (2021). The use of personal weather station observations to improve precipitation estimation and interpolation. *Hydrology and Earth System Sciences*, 25:583–601.

[Båserud et al., 2020] Båserud, L., Lussana, C., Nipen, T. N., Seierstad, I. A., Oram, L., ja Aspelien, T. (2020). TITAN automatic spatial quality control of meteorological in-situ observations. *Advances in Science and Research*, 17:153–163.

[Cecinati et al., 2018] Cecinati, F., Moreno-Ródenas, A. M., Rico-Ramirez, M. A., claire ten Veldhuis, M., ja Langeveld, J. G. (2018). Considering Rain Gauge Uncertainty Using Kriging for Uncertain Data. *Atmosphere*, 9(11):446.

[Chapman et al., 2016] Chapman, L., Bell, C., ja Bell, S. (2016). Can the crowdsourcing data paradigm take atmospheric science to a new level? A case study of the urban heat island of London quantified using Netatmo weather stations. *International Journal of Climatology*, 37(9):3597–3605.

[de Vos et al., 2019] de Vos, L., Leijnse, H., Overeem, A., ja Uijlenhoet, R. (2019). Quality Control for Crowdsourced Personal Weather Stations to Enable Operational Rainfall Monitoring. *Geophysical Research Letter*, 46(15):8820–8829.

[Haapanala, 2013] Haapanala, S. (2013). Kojemeteorologia. Luento 8: Sateen mittaaminen. *Helsingin yliopisto*.

[Ilmatieteen laitos, 2023a] Ilmatieteen laitos (2023a). Havaintosuureet: mitä havainnoimme? https://www.ilmatieteenlaitos.fi/havaintosuureet. (Luettu: 21.10.2023).

[Ilmatieteen laitos, 2023b] Ilmatieteen laitos (2023b). Oman sääaseman perustaminen. https://www.ilmatieteenlaitos.fi/oma-saaasema. (Luettu: 21.10.2023).

[Ilmatieteen laitos, 2023c] Ilmatieteen laitos (2023c). Sade-ennätyksiä. https://www.ilmatieteenlaitos.fi/sade-ennatyksia. (Luettu: 28.10.2023).

[Netatmo, 2012] Netatmo (2012). *Netatmo User Manual: Netatmo Weather Station NWS01*.

[Netatmo, 2023] Netatmo (2023). Smart Rain Gauge. https://www.netatmo.com/en-gb/smart-rain-gauge. (Luettu: 03.05.2023).

[Netatmo Help Center, 2023] Netatmo Help Center (2023). How to calibrate the rain sensor? https://helpcenter.netatmo.com/hc/en-us/articles/360020908932-How-to-calibrate-the-rain-sensor-. (Luettu: 19.09.2023).

[OTT Hydromet, 2016a] OTT Hydromet (2016a). *Operating Instructions: Precipitation Gauge OTT Pluvio^{2}L*. Dokumentin numero: 70.040.000.B.E 01-0116. OTT Hydromet GmbH, Kempten, Saksa.

[OTT Hydromet, 2016b] OTT Hydromet (2016b). *OTT Pluvio^2 L: Precipitation gauge using the balance principle*. OTT Hydromet GmbH, Kempten, Saksa.

[OTT Messtechnik, 2008] OTT Messtechnik (2008). *Operating Instructions: Precipitation Gauge OTT Pluvio^2*. Dokumentin numero: 70.020.000.B.E 02-0708. OTT MESSTECHNIK GmbH Co. KG, Kempten, Saksa.

[Verkkokauppa, 2023] Verkkokauppa (2023). Netatmo Rain Gauge -langaton sademittari. https://www.verkkokauppa.com/fi/product/182526/Netatmo-Rain-Gauge-langaton-sademittari. (Luettu: 25.12.2023).

[World Meteorological Organization, 2023] World Meteorological Organization (2023). Records of Weather and Climate Extremes. https://wmo.int/sites/default/files/2023-11/Table_Records_6Nov2023.pdf. (Luettu: 09.01.2024).