

# **Otantaharha ekologisessa kansalaistiedeaineistossa – tarkastelussa Punkkilive**

Maantiede (soveltava paikkatietotutkimus)

Pro gradu -tutkielma

Laura Koivukoski

21.5.2026  
Turku

**Tutkinto-ohjelma, oppiaine:** Maantiede, Filosofian maisteri, Maantieteen linja

**Tekijä:** Laura Koivukoski

**Otsikko:** Otantaharha ekologisessa kansalaistiedeaineistossa – tarkastelussa Punkkilive

**Ohjaajat:** Niko Kulha, Jani Sormunen, Nora Fagerholm

**Sivumäärä:** 84 sivua

**Päivämäärä:** 21.5.2026

---

Kansalaistieteestä on tullut tärkeä aineistonkeruun menetelmä laajojen ympäristö- ja biodiversiteettiaineistojen keräämiseen. Erityisesti paikkatietoon perustuvat ekologiset kansalaistiedeaineistot mahdollistavat lajien levinneisyyden tarkastelun laajoilla alueilla ja pitkillä ajanjaksoilla. Kansalaistiedeaineistoihin liittyy kuitenkin otantaharhaa, joka ilmenee esimerkiksi spatiaalisena ja ajallisena sekä havainnointiin ja osallistujiin liittyvinä vinoumina. Vinoumat syntyvät, koska havaintojen kertymiseen vaikuttavat ekologisten tekijöiden lisäksi ihmisten liikkuminen, havaintokäyttäytyminen ja osallistumisaktiivisuus. Tämän tutkimuksen tavoitteena on tarkastella Punkkilive-kansalaistiedeaineiston puutiaishavaintojen alueellista ja ajallista vaihtelua Varsinais-Suomessa sekä analysoida puutiaisaineistoon liittyviä vinoumia.

Tutkimuksen aineistona hyödynsin Punkkilive-sovellukseen Varsinais-Suomesta vuosilta 2021–2023 ilmoitettuja puutiaishavaintoja. Tutkimuksen metodologinen kokonaisuus rakentui kolmesta vaiheesta. Ensin tarkastelin puutiaishavaintojen ajallista ja alueellista jakautumista. Tämän jälkeen mallinsin puutiaishavaintojen tiheyden ja väestömäärän välistä suhdetta maantieteellisesti painotetun regressiomallin (GWR) avulla, jonka perusteella muodostin jäännökset ja poikkeama-alueet. Lopuksi analysoin paikkatietoaineistoista muodostettujen habitaatti- ja ihmistoimintaa kuvaavien muuttujien avulla puutiaishavaintojen spatiaalisten vinoumien taustalla vaikuttavia tekijöitä. Toteutin analyysin yleistetyn additiivisen mallin (GAM) avulla.

Tulokset osoittivat, että puutiaishavainnot painoutuivat ajallisesti erityisesti touko-elokuulle. Havaintojen alueellinen vaihtelu oli voimakasta, ja havainnot keskittyivät erityisesti rannikolle, saaristoon ja kaupunkialueiden läheisyyteen. Väestömäärä selitti vain pienen osan puutiaishavaintojen alueellisesta vaihtelusta. Havaintojen poikkeama-alueet säilyivät alueellisesti samankaltaisina tarkasteluvuosien välillä. Erityisesti rannikko- ja saaristoalueilla puutiaishavaintoja kertyi väestömäärään nähden odotettua enemmän. Sen sijaan sisämaan taajama-alueilla havaintoja kertyi väestömäärään nähden odotettua vähemmän. GAM-mallin perusteella havaintojen poikkeamaa väestömäärästä selittivät erityisesti loma-asunnot, rantaviivan pituus, NDVI, liikuntapaikat, polut ja rakennetun ympäristön osuus maankäytöstä.

Tutkimustulokset osoittavat, miten kansalaistiedeaineistoon liittyvä spatiaalinen vinouma muodostuu puutiaishavaintoaineistossa. Tulosten perusteella havaintojen alueellinen jakauma ei määräydy pelkästään puutiaisten ekologisen esiintymisen perusteella, vaan siihen vaikuttavat myös alueiden saavutettavuus ja virkistyskäyttö. Tutkimuksen tuloksia voidaan hyödyntää puutiaisiin liittyvien alueellisten riskien tarkastelussa sekä kansalaistiedeaineistojen vinoumien ymmärtämisessä.

---

**Avainsanat:** kansalaistiede, otantaharha, spatiaalinen vinouma, puutiaiset

Master's thesis

**Subject:** Geography

**Author:** Laura Koivukoski

**Title:** Sampling bias in ecological citizen science data – the case of Punkkilive

**Supervisors:** Niko Kulha, Jani Sormunen, Nora Fagerholm

**Number of pages:** 84 pages

**Date:** 21.5.2026

---

Citizen science has become an important method for collecting extensive environmental and biodiversity data. Location-based ecological citizen science datasets provide opportunities to examine species distribution across large geographic areas and long time periods. However, citizen science datasets are affected by sampling bias, which may appear as spatial, temporal, detection, and observer-related biases. These biases arise because, in addition to ecological factors, the accumulation of observations is influenced by human mobility, observation behaviour and participation activity. The aim of this study is to examine the spatial and temporal variation of tick observations in the Punkkilive citizen science dataset in Southwest Finland and to analyse biases associated with the dataset.

I used tick observations reported to the Punkkilive application from Southwest Finland between 2021 and 2023. The methodological framework of the study consisted of three stages. First, I examined the temporal and spatial distribution of tick observations. Second, I modelled the relationship between tick observation density and population size using geographically weighted regression (GWR), from which I derived residuals and spatial deviation areas. Finally, I analysed factors underlying the spatial bias in tick observations using generalized additive model (GAM) with habitat and accessibility variables derived from spatial datasets.

The results showed that tick observations are temporally concentrated between May and August. Spatial variation was strong, and observations were concentrated particularly near coastal, archipelago, and urban areas. Number of inhabitants explained only a small proportion of the spatial variation in tick observations. The spatial pattern of deviations remained regionally similar throughout the study period. Coastal and archipelago areas had more tick observations than expected relative to number of inhabitants whereas inland urban areas had fewer observations than expected. According to the GAM model, deviations in tick observations relative to the number of inhabitants were explained particularly by summer cottages, shoreline length, NDVI, recreational facilities, trails, and the proportion of built environment.

This study provides information on how spatial bias related to citizen science data arises in tick observation data. The spatial distribution of observations was not determined only by the ecological occurrence of ticks but was also influenced by accessibility and recreational land use. The results of the study can be utilized in assessing spatial patterns related to tick occurrence as well as in improving the understanding of biases related with citizen science datasets.

---

**Key words:** citizen science, sampling bias, spatial bias, ticks

# Sisällysluettelo

|          |                                                                               |           |
|----------|-------------------------------------------------------------------------------|-----------|
| <b>1</b> | <b>Johdanto</b>                                                               | <b>6</b>  |
| <b>2</b> | <b>Teoreettinen viitekehys</b>                                                | <b>8</b>  |
| 2.1      | Kansalaistiede spatiaalisen tiedon tuotannossa                                | 8         |
| 2.2      | Kansalaistieteeseen pohjautuvat lajihavaintoaineistot ja niiden haasteet      | 11        |
| 2.2.1    | Lajihavaintoaineistot ekologisessa tutkimuksessa                              | 11        |
| 2.2.2    | Vinoumat ekologisissa kansalaistiedeaineistoissa                              | 12        |
| 2.2.3    | Spatiaalinen vinouma                                                          | 14        |
| 2.2.4    | Ajallinen vinouma                                                             | 16        |
| 2.2.5    | Havainnointiin liittyvä vinouma                                               | 18        |
| 2.2.6    | Osallistujiin liittyvä vinouma                                                | 20        |
| 2.3      | Puutiaisen ekologia ja puutiaishavaintojen vinoumat                           | 23        |
| 2.3.1    | Puutiaisen ekologia ja levinneisyys Suomessa                                  | 23        |
| 2.3.2    | Puutiaishavaintojen erityispiirteet kansalaistiedeaineistossa                 | 24        |
| <b>3</b> | <b>Aineistot ja menetelmät</b>                                                | <b>27</b> |
| 3.1      | Tutkimusalue                                                                  | 27        |
| 3.2      | Tutkimuksen metodologinen kokonaisuus                                         | 29        |
| 3.3      | Aineistot                                                                     | 31        |
| 3.3.1    | Punkkilive-kansalaistiedeaineisto                                             | 31        |
| 3.3.2    | Muut paikkatietoaineistot                                                     | 32        |
| 3.3.3    | Aineistojen esikäsittely ja standardointi                                     | 33        |
| 3.4      | Väestömäärään perustuva havaintomäärän mallintaminen                          | 37        |
| 3.5      | Poikkeama-alueiden tunnistaminen                                              | 40        |
| 3.6      | Poikkeama-alueiden paikkatietopohjainen selittäminen                          | 42        |
| <b>4</b> | <b>Tulokset</b>                                                               | <b>47</b> |
| 4.1      | Puutiaisten määrä ja jakautuminen Varsinais-Suomessa                          | 47        |
| 4.2      | Puutiaishavaintojen määrän ennustettavuus asukasluvun perusteella             | 50        |
| 4.3      | Puutiaishavaintojen poikkeamia selittävät ympäristö- ja saavutettavuustekijät | 56        |
| <b>5</b> | <b>Keskustelu</b>                                                             | <b>63</b> |
| 5.1      | Havaintojen spatiaalinen jakauma ja väestömäärään pohjautuva tarkastelu       | 63        |

|     |                                                           |    |
|-----|-----------------------------------------------------------|----|
| 5.2 | Puutiaisten havaintomäärän poikkeaminen asukasmäärästä    | 65 |
| 5.3 | Kansalaistiedeaineiston tulkinta ja menetelmien arviointi | 70 |
| 6   | Johtopäätökset                                            | 74 |
|     | Lähteet                                                   | 75 |

## 1 Johdanto

Ilmastonmuutos muuttaa lajien levinneisyyksiä ennennäkemättömällä tahdilla (IPCC 2023). Samalla myös zoonoosit eli eläimistä ihmisiin leviävät taudit leviävät uusille alueille. Muutosten seuranta edellyttää laajoja ja ajallisesti kattavia aineistoja, joiden kerääminen pelkästään perinteisen kenttätyön avulla on haastavaa (Dickinson ym. 2010). Kansalaistiede tarjoaa uudenlaisia mahdollisuuksia lajihavaintojen keräämiseen ekologisessa tutkimuksessa.

Suomessa puutiaiset ovat merkittävä zoonooseja levittävä eliöryhmä, ja kiinnostus niiden levinneisyyteen on kasvanut viime vuosina (Laaksonen ym. 2017; Sormunen ym. 2020). Perinteiseen kenttätyöhön perustuva seuranta on resurssivaatimustensa ja rajoitetun maantieteellisen kattavuuden vuoksi riittämätöntä kuvaamaan nopeasti muuttuvaa puutiaistilannetta (Laaksonen ym. 2017; Sormunen ym. 2023b). Suomessa otettiin käyttöön vuonna 2021 Punkkilive-sivusto, jonka avulla kansalaiset voivat ilmoittaa puutiaishavaintoja ja tarkastella riskialueita (Sormunen ym. 2023a). Kansalaistieteen keinoin kerätyn puutiaisaineiston hyödyntäminen tutkimuksessa edellyttää kuitenkin aineiston rajoitusten tunnistamista ja huomioon ottamista (Welvaert & Caley 2016).

Kansalaistiedepohjaisiin puutiaisaineistoihin liittyy otantaharhaa, sillä havaintojen kertymiseen vaikuttavat puutiaisten ekologisen esiintymisen lisäksi myös ihmisten liikkuminen, havaintokäyttäytyminen ja alueiden saavutettavuus (Geldmann ym. 2016). Havainnot voivat painottua alueille ja ajankohtiin, joilla ihmiset liikkuvat eniten, eikä aineisto siten välttämättä kuvaa puutiaisten todellista runsautta (Kulha ym. 2022). Lisäksi havaintoihin liittyy epävarmuutta sijaintitietojen osalta, sillä puutiaiset löytyvät usein isäntäeläimistä, kuten ihmisistä tai lemmikeistä, eivätkä välttämättä samasta paikasta, josta ne ovat peräisin (Laaksonen ym. 2017; Sormunen ym. 2023b). Näistä rajoitteista huolimatta kansalaistiedeaineistot tarjoavat mahdollisuuden tarkastella puutiaisten esiintymistä huomattavasti laajemmassa mittakaavassa kuin pelkän kenttätyön avulla olisi mahdollista (Laaksonen ym. 2017).

Tämä tutkimus täydentää aiempaa tutkimusta tarkastelemalla nimenomaan kansalaistiedeaineistoon liittyvän spatiaalisen vinouman muodostumista pelkän

puutiaisten levinneisyyden ja riskialueiden tutkimisen sijaan. Tutkimuksessa tarkastellaan, miten puutiaisten elinympäristöihin liittyvät tekijät ja ihmistoiminta yhdessä muokkaavat puutiaishavaintojen alueellista rakennetta. Näin tutkimus tuottaa tietoa siitä, miten puutiaisiin liittyviä kansalaistiedeaineistoja voidaan tulkita ekologisessa tutkimuksessa.

Tutkielman tavoitteena on tarkastella Punkkilive-kansalaistiedeaineistoon liittyvää otantaharhaa Varsinais-Suomessa. Otantaharhalla tarkoitetaan tässä tutkimuksessa poikkeamaa havaintomäärästä, joka olisi odotettavissa, jos havaintojen määrä riippuisi ainoastaan asukasluvusta. Tutkimuksessa tunnistetaan alueita, joilla puutiaishavaintoja esiintyy enemmän tai vähemmän kuin asukasluvun perusteella olisi odotettavissa. Lisäksi analysoidaan ympäristö- ja ihmistoimintaan liittyviä tekijöitä, jotka selittävät näitä poikkeamia. Tutkielma rakentuu seuraavien tutkimuskysymysten ympärille:

Tutkimuskysymykset:

1. Miten Punkkilive-kansalaistiedeaineiston puutiaishavainnot jakautuvat spatiotemporaalisesti Varsinais-Suomessa?
2. Kuinka hyvin puutiaishavaintojen tiheyttä voidaan ennustaa asukasluvun perusteella?
3. Mitkä tekijät selittävät havaintotiheyden poikkeamista asukasluvusta?

## 2 Teoreettinen viitekehys

### 2.1 Kansalaistiede spatiaalisen tiedon tuotannossa

Kansalaisten merkitys maantieteellisen tiedon tuottajina on kasvanut nopeasti teknologisen kehityksen myötä, ja samalla tiedon tuotanto on laajentunut asiantuntijoiden ulkopuolelle (See ym. 2016). Kansalaistieteellä tarkoitetaan tutkimustoimintaa, jossa kansalaiset osallistuvat tieteellisen tiedon keruuseen, analyysiin tai tulkintaan yhteistyössä tutkijoiden ja tutkimusorganisaatioiden kanssa (European Citizen Science Association 2015). Kansalaistieteen määritelmien keskeinen yhdistävä piirre on vapaaehtoisten osallistuminen tieteellisen tiedon tuottamiseen, vaikka osallistumisen laajuus ja muodot vaihtelevat huomattavasti eri hankkeiden välillä (Hecker ym. 2018). Kansalaistiede voidaan siten ymmärtää sekä tiedontuotannon tapana että laajempaa yhteiskunnallisena käytäntönä, joka kytkee yhteen tutkimuksen, osallistumisen ja tiedon avoimuuden. Samalla se haastaa perinteistä käsitystä siitä, kenellä on oikeus tuottaa tieteellisesti merkityksellistä tietoa ja millaisiin havaintoihin tutkimus voi perustua.

Vaikka kansalaistieteen juuret ulottuvat vuosisatojen taakse, itse käsite vakiintui vasta 1990-luvulla, minkä jälkeen sen käyttö ja tunnustus ovat lisääntyneet nopeasti (Vohland ym. 2021). Nykyisin kansalaistiede tunnistetaan laajasti tutkimuksen lähestymistapana tieteen, politiikan ja koulutuksen kentillä. Tämä kehitys on lisännyt tarvetta yhteisille käsitteille, standardeille ja ohjeistuksille, sillä kansalaistieteen monimuotoisuus vaikeuttaa hankkeiden, aineistojen ja osallistumismuotojen vertailua sekä niiden arviointia (Bonney ym. 2014; Vohland ym. 2021).

Useat kansalaistieteen määritelmät korostavat kansalaisten osallistumista tutkimukseen, mutta ovat avoimia sen suhteen, millä tavoin ja missä vaiheessa tutkimusprosessia osallistuminen tapahtuu (Haklay ym. 2021). Tämän vuoksi kansalaistiedettä on hyödyllistä tarkastella osallistumisen asteiden ja valtasuhteiden näkökulmasta. Haklayn ym. (2021) esittämä jaottelu havainnollistaa eroa tiedonkeruuseen painottuvan osallistavan kansalaistieteen ja demokraattisemman lähestymistavan välillä, jossa kansalaisilla on aito merkitys myös tutkimuskysymysten, menetelmien ja tulosten määrittelyssä (Haklay ym. 2018; Irwin ym. 1995). Tämä

näkökulma laajentaa käsitystä kansalaistieteestä pelkkänä aineistonkeruumenetelmänä kohti tutkimusprosessin eri vaiheisiin ulottuvaa osallistumista.

Kansalaistieteen keskeinen ominaisuus on kansalaisten osallistumisen laajentaminen tieteellisen tutkimusprosessin eri vaiheisiin (Haklay ym. 2018). Kansalaiset voivat osallistua aineiston keruun lisäksi tutkimuskysymysten muotoiluun, analyysiin, tulosten tulkintaan ja päätöksenteon tukemiseen (Hecker ym. 2018; Robinson ym. 2018). Tämä erottaa kansalaistieteen perinteisistä tutkimusasetelmista ja liittää sen osaksi tieteen demokratisoitumista. Osallistuminen ei kuitenkaan jakaudu tasaisesti. Esimerkiksi vähemmistö- ja pienituloiset ryhmät ovat usein aliedustettuina kansalaistiedehankkeissa, mikä asettaa haasteita kansalaistieteen inklusiivisuudelle (Robinson ym. 2018).

Kansalaistiede edustaa vaihtoehtoista tiedontuotannon mallia, jossa tieteellinen asiantuntemus yhdistyy paikalliseen, kokemukselliseen ja yhteisölliseen tietoon (Hecker ym. 2018). Tiedon yhteistuotanto voi rikastuttaa tutkimusta, laajentaa tarkasteltavaa ilmiötä ja vahvistaa tutkimuksen yhteiskunnallista merkittävyyttä (Hecker ym. 2018; Robinson ym. 2018). Samalla kansalaistiede korostaa avoimen tieteen periaatteita, joissa aineistot, metadata ja tulokset pyritään tekemään mahdollisimman avoimiksi ja läpinäkyviksi, mikä edistää niiden hyväksyttävyyttä ja uudelleenkäytettävyyttä (Robinson ym. 2018).

Tällainen tiedontuotannon muutos on ollut merkittävä maantieteellisessä tutkimuksessa, jossa teknologinen kehitys on laajentanut kansalaisten mahdollisuuksia osallistua paikkatiedon tuottamiseen. Aiemmin paikkatieto oli pääosin viranomaisten ja asiantuntijoiden tuottamaa, mutta mobiililaitteiden, GPS-paikannuksen ja verkkopohjaisten karttapalveluiden myötä kansalaiset osallistuvat yhä aktiivisemmin tiedon tuottamiseen (Goodchild 2007; Elwood ym. 2012; Bonney ym. 2014). Tämä on lisännyt mahdollisuuksia tuottaa ajantasaista, paikallisesti yksityiskohtaista ja maantieteellisesti kattavaa paikkatietoa.

Kansalaistieteen nopea kasvu kytkeytyy vahvasti teknologiseen ja yhteiskunnalliseen murrokseen (Hecker ym. 2018). Älypuhelimet, mobiilisovellukset,

paikkatietoteknologiat ja avoimet tietojärjestelmät ovat mahdollistaneet laajamittaisen ja ajantasaisen tiedonkeruun, jossa kansalaiset toimivat ikään kuin hajautettuina sensoreina (Goodchild 2007; Bonney ym. 2014; Zhu ym. 2015). Tämä on ollut erityisen merkittävää maantieteellisessä tutkimuksessa, jossa ilmiöiden spatiaalinen ja ajallinen kattavuus on keskeistä. Aiemmin paikkatieto oli pääosin viranomaisten ja asiantuntijoiden tuottamaa, mutta teknologisen kehityksen myötä kansalaiset osallistuvat yhä aktiivisemmin paikkatiedon tuottamiseen (Goodchild 2007; Elwood ym. 2012). Kansalaisten tuottama tieto voi täydentää perinteisiä havainto- ja kaukokartoitusaineistoja tarjoamalla paikallista ja ajallisesti tarkkaa informaatiota, jota ei olisi muuten saatavilla (Zhu ym. 2015). Tämä on lisännyt ekologisessa tutkimuksessa lajien levinneisyyttä koskevien aineistojen määrää ja ajallista kattavuutta (Bonney ym. 2014; Frigerio ym. 2021).

Kansalaistieteen vahvuudet liittyvät erityisesti sen mittakaavaan. Ne mahdollistavat sellaisten spatiotemporaalisten ilmiöiden tarkastelun, joita yksittäiset tutkimushankkeet tai perinteinen kenttätöyskentely eivät kykene kattamaan (Dickinson ym. 2010). Suuri määrä vapaaehtoisia voi tuottaa havaintoja samanaikaisesti eri alueilta, mikä lisää aineistojen maantieteellistä kattavuutta ja ajallista jatkuvuutta (See ym. 2016). Samalla kansalaisten tuottama tieto voi tarjota tutkimukselle paikallisesti merkityksellistä informaatiota, jota ei olisi mahdollista kerätä yksinomaan asiantuntijaresurssein (Goodchild 2007; See ym. 2016).

Kansalaistieteen käyttöön liittyy kuitenkin haasteita, jotka koskevat erityisesti aineiston laatua, kattavuutta ja edustavuutta (See ym. 2025). Havaintojen tarkkuus ja luotettavuus voivat vaihdella osallistujien osaamisen, motivaation ja raportointikäytäntöjen mukaan (See ym. 2016; Robinson ym. 2018). Lisäksi osallistuminen on usein epätasaista sekä alueellisesti että sosiaalisesti, mikä voi johtaa spatiaalisesti ja demografisesti vinoutuneisiin aineistoihin (Goodchild 2007; Elwood ym. 2012). Kansalaisten tuottamat aineistot eivät siten välttämättä kuvaa tasaisesti tutkittavaa ilmiötä, vaan myös havaitsijoiden toimintaa, liikkumista ja teknologisia valmiuksia.

## **2.2 Kansalaistieteeseen pohjautuvat lajihavaintoaineistot ja niiden haasteet**

### **2.2.1 Lajihavaintoaineistot ekologisessa tutkimuksessa**

Ekologisessa tutkimuksessa kansalaistiede on noussut keskeiseksi välineeksi lajihavaintojen, lajien levinneisyyden ja ympäristömuutosten seurannassa (Bonney ym. 2014). Valtaosa kansalaistiedehankkeista perustuu havaintojen tekemiseen ja aineistojen keruuseen, mikä mahdollistaa laajojen alueiden ja pitkien aikajaksojen tarkastelun esimerkiksi lajien esiintymisen ja runsauden vaihtelun kartoittamiseksi (Hecker ym. 2018; Frigerio ym. 2021: 84). Kansalaisten osallistuminen on lisännyt merkittävästi lajihavaintoja koskevien aineistojen määrää ja ajallista kattavuutta.

Kansalaistieteen merkitys ekologisessa tutkimuksessa liittyy erityisesti sen mahdollisuuteen tuottaa laajoja, ajallisesti jatkuvia ja maantieteellisesti kattavia aineistoja (Dickinson ym. 2010). Suuri vapaaehtoisten määrä mahdollistaa spatiotemporaalisten ilmiöiden tarkastelun laajassa mittakaavassa (See ym. 2016). Samalla kansalaisten tuottama tieto voi tarjota tutkimukselle paikallisesti merkityksellistä informaatiota, jota ei olisi mahdollista kerätä ainoastaan asiantuntijaresurssein (Goodchild 2007; See ym. 2016).

Teknologinen kehitys on vauhdittanut merkittävästi kansalaistieteen kasvua ekologisessa tutkimuksessa. Älypuhelimet, mobiilisovellukset ja GPS-paikannus ovat mahdollistaneet havaintojen keräämisen ja jakamisen lähes reaaliaikaisesti (Goodchild 2007; Bonney ym. 2014; Zhu ym. 2015). Kansalaiset tuottavat jatkuvasti paikallisesti tarkkaa tietoa ympäristöstä ja lajien esiintymisestä (Goodchild 2007; Zhu ym. 2015). Tämän seurauksena valtaosa nykyisestä lajihavaintoihin liittyvästä tiedosta perustuu kansalaistieteen tuottamiin havaintoihin.

Kansalaistieteen tuottamia lajihavaintoaineistoja hyödynnetään monipuolisesti ekologisessa tutkimuksessa. Niitä käytetään esimerkiksi lajien levinneisyyden mallintamiseen, muutosten seurantaan, vieraslajien leviämisen tarkasteluun sekä ympäristö- ja terveystarkkailuun (Bonney ym. 2014; Frigerio ym. 2021). Kansalaishavaintojen avulla voidaan myös havaita nopeasti ympäristössä tapahtuvia

muutoksia sekä täydentää perinteisiä havainto- ja kaukokartoitusaineistoja paikallisesti yksityiskohtaisella tiedolla (Zhu ym. 2015).

Vaikka kansalaistieteeseen perustuvat havaintoaineistot tarjoavat ekologiselle tutkimukselle merkittäviä mahdollisuuksia, ne eivät muodostu systemaattisesti tutkimusasetelman mukaisesti kerätystä aineistosta. Kansalaistiedeaineistot eivät ole satunnaisotoksia ympäristöstä, vaan ne heijastavat havaitsijoiden toimintaa, liikkumista, kiinnostuksen kohteita ja havaintohalukkuutta (Callaghan ym. 2020; Johnston ym. 2020). Tämän vuoksi kansalaishavainnot ovat lähtökohtaisesti vinoutuneita, eivätkä kuvaa ainoastaan tutkittavan ilmiön todellista esiintymistä (Dickinson ym. 2010; Johnston ym. 2023).

### 2.2.2 Vinoumat ekologisissa kansalaistiedeaineistoissa

Kansalaisten osallistuminen tiedonkeruuseen mahdollistaa laaja-alaiset ja ajallisesti tiheät havaintoaineistot, mutta niihin liittyy myös haasteita aineistojen edustavuuden ja luotettavuuden kannalta (Dickinson ym. 2010; Zhang 2020). Ekologiset kansalaistiedeaineistot perustuvat usein satunnaiseen näytteenottoon, jossa havainnot kerätään ilman ennalta määriteltyä otantasuunnitelmaa (Zhang & Zhu 2018). Tämän vuoksi ne eivät aina täytä hyvän empiirisen tutkimusasetelman peruseriaatteita, vaan havaintojen jakautuminen voi heijastaa niin havaitsijoiden toimintaa kuin tutkittavan ilmiön todellista esiintymistä (Geldmann ym. 2016; Zhang & Zhu 2018).

Ekologisissa kansalaistiedeaineistoissa havaintojen kertymiseen vaikuttavat samanaikaisesti sekä lajien ekologiset ominaisuudet että ihmistoiminta. Havaintojen määrä ei siten kuvaa lajien todellista esiintymistä, vaan myös sitä, missä ihmiset liikkuvat, mitä lajeja havainnoidaan ja kuinka aktiivisesti havaintoja raportoidaan (Dickinson ym. 2010; Zhu ym. 2015; Johnston ym. 2023). Tämä korostaa havaintoaineistoihin kytkeytyvien rajoitusten tunnistamista, kun aineistoja käytetään ekologisten ilmiöiden, lajien levinneisyyden ja ympäristöriskien tarkastelussa (Backstrom ym. 2025).

Yksi ekologisten kansalaistiedeaineistojen keskeisistä haasteista on otantaharha, joka voi heikentää aineistoista tehtävien johtopäätösten, kuten lajien levinneisyysmallien,

luotettavuutta (Backstrom ym. 2025). Otantaharha syntyy, kun havaintojen sijainti, ajankohta tai havaintopanos eivät jakaudu tasaisesti tutkimusalueella tai -ajanjaksolla (Isaac ym. 2014). Käytännössä otantaharha ilmenee erilaisina aineistoon kytkeytyvinä vinoumina, jotka liittyvät esimerkiksi havaintojen jakautumiseen ja havaitsijoiden toimintaan. Vinoumat voivat johtaa siihen, että aineistossa havaitaan muutoksia, jotka johtuvat todellisuudessa havainnointitoiminnan vaihtelusta eikä ekologisista prosesseista. Vastaavasti todelliset ekologiset muutokset voivat jäädä aineistossa havaitsematta.

Isaac ym. (2014) tunnistavat neljä keskeistä otantaharhan lähdettä ekologisissa kansalaistiedeaineistoissa: (1) havaintojen epätasainen alueellinen kattavuus, (2) ajallisesti vaihteleva kirjausaktiivisuus, (3) vaihteleva tiedonkeruupanos käyntiä kohden sekä (4) havaitsijoiden vaihteleva kyky havaita ja tunnistaa kohteita (taulukko 1). Mikäli näitä vinoumia ei huomioida, ne voivat johtaa harhaanjohtaviin arvioihin lajien runsaudesta, esiintyvyydestä tai niiden ajallisista muutoksista (Geldmann ym. 2016).

Taulukko 1. Ekologisiin kansalaistiedeaineistoihin liittyvät keskeisimmät vinoumat.

| Vinouma                                | Määritelmä                                                   | Vaikutus havaintoaineistoon                                              |
|----------------------------------------|--------------------------------------------------------------|--------------------------------------------------------------------------|
| <b>Spatiaalinen vinouma</b>            | Havaintojen epätasainen alueellinen jakautuminen             | Havaintoja kertyy enemmän asutuilta ja helposti saavutettavilta alueilta |
| <b>Ajallinen vinouma</b>               | Havaintojen epätasainen jakautuminen ajassa                  | Havaintoja kertyy enemmän viikonloppuisin tai suotuisina vuodenaikoina   |
| <b>Havainnointiin liittyvä vinouma</b> | Lajin havaittavuus ja tunnistettavuus                        | Näkyvät ja helposti tunnistettavat lajit havaitaan paremmin              |
| <b>Osallistujiin liittyvä vinouma</b>  | Havaitsijoiden motivaatio, osaaminen ja raportointikäytännöt | Aktiiviset havaitsijat tuottavat valtaosan havainnoista                  |

Ekologisten kansalaistiedeaineistojen luotettavuuteen vaikuttavat havainnoijan taustan lisäksi tutkimusasetelma, laadunvarmistus ja analyysimenetelmät sekä aineistonkeruuseen liittyvät tekijät, kuten osallistujien motivaatio, aineistojen hallinta ja resurssirajoitteet (Bonney ym. 2014; Robinson ym. 2018; Frigerio ym. 2021: 89). Huolellisesti suunniteltuina kansalaistieteen tuottamat aineistot voivat olla vertailukelpoisia asiantuntijoiden keräämien aineistojen kanssa. Opportunististen

aineistojen analysointi edellyttää kuitenkin systemaattisten virheiden ja vinoumien tunnistamista, jotta havaintojen käyttäytymisestä, havainnointiprosessista ja raportointikäytännöistä johtuvat rakenteet voidaan erottaa ekologisista ilmiöistä (Johnston ym. 2023).

### 2.2.3 Spatiaalinen vinouma

Spatiaalinen vinouma eli havaintojen epätasainen maantieteellinen jakautuminen on yksi ekologisten kansalaistiedeaineistojen keskeisistä vinoumatyypeistä. Se vaikuttaa siihen, kuinka hyvin kansalaisten keräämä aineisto kuvaa tutkittavan ilmiön todellista esiintymistä tilassa (Zhang & Zhu 2018). Spatiaalinen vinouma syntyy, kun havaintojen keruu perustuu havaintojen liikkumiseen eikä ennalta suunniteltuun, systemaattiseen otantaan (Zhu ym. 2015). Havaintajat tekevät havaintoja yleensä muun toiminnan ohessa, eivätkä valitse havaintopaikkoja systemaattisesti otantasuunnitelman perusteella. Näin ollen havaintojen sijainti heijastaa ennen kaikkea ihmisten liikkumista, ei ilmiön todellista levinneisyyttä. Tämän seurauksena havaintojen jakautuminen on epätasaista. Tietyiltä alueilta kertyy runsaasti havaintoja, kun taas monet alueet jäävät aliedustetuiksi tai kokonaan ilman havaintoja (Zhang 2020).

Havaintojen määrä kuvastaa havaintojen aktiivisuutta ja alueiden saavutettavuutta, eikä pelkästään ilmiön todellista esiintymistä (Zhang 2020). Kaupunkialueet, tiheästi asutut alueet ja kohteet, joissa liikkuminen on helppoa, ovat usein yliedustettuja, kun taas syrjäiset, vaikeakulkuiset ja harvaan asutut alueet jäävät aliedustetuiksi (Johnston ym. 2020; Zhang 2020). Tämä ei heijasta pelkästään ilmiön todellista esiintymistä, vaan myös havaintojen toiminnan ja saavutettavuuden ohjaamia eroja. Koska havaintoja tehdään havaintojen liikkumisreittien ja tuttujen ympäristöjen varrella, havaintotiheys painottuu usein alueille, joilla käyttäjäaktiivisuus ja saavutettavuus ovat korkeita (Zhu ym. 2015). Kansalaistieteeseen nojaava aineisto ei myöskään kata tarkasteltavan alueen ympäristötekijöiden vaihtelua tasaisesti, mikä heikentää näytteiden ekologista kattavuutta (Zhang 2020).

Spatiaalinen vinouma voi voimistua myös havaintojen kollektiivisen käyttäytymisen seurauksena. Kansalaistieteen havaintoalustat, kuten lintujen havainnointiin tarkoitettu eBird, tarjoavat käyttäjille karttapohjaisia näkymiä havaintojen tiheydestä ja lajien

runsaudesta (Sullivan ym. 2014). Esimerkiksi hotspot-alueiden näkyvyys voi kannustaa vapaaehtoisia havainnoimaan samoilla alueilla, mikä näkyy eBird-aineistossa lintuhavaintojen voimakkaana alueellisena kasaantumisenä (Callaghan ym. 2020). Tämä lisää havaintojen yliedustusta tietyillä alueilla ja voi edelleen kaventaa aineiston alueellista edustavuutta (Zhang 2020). Tämän seurauksena havainnot voivat painottua alueille, joilla havaintopanos on korkea, vaikka tarkasteltavan ilmiön todellinen esiintyvyys painottuisi toisaalle (Fisher-Phelps ym. 2017).

Kaikki kansalaistieteen havaintokohteet eivät kuitenkaan käyttäydy samalla tavoin havaitsijoiden näkökulmasta. Linnut ovat tyypillisesti haluttuja havaintokohteita, jolloin havaitsijat suuntaavat alueille, joilla havaintotodennäköisyys on korkea (Callaghan ym. 2020). Sen sijaan puutiaisets ovat esimerkki ”ei-toivotusta” havaintokohteesta, joiden havaitseminen liittyy satunnaisiin kohtaamisiin ja altistumiseen arkiympäristöissä. Tämä ero vaikuttaa spatiaalisen vinouman muodostumiseen.

Tällaisten ei-toivottujen eliöiden havaintojen kasaantuminen ei välttämättä johdu pelkästään havaitsijoiden tietoisesta toiminnasta, vaan kuvastaa ensisijaisesti ihmisten liikkumista ja altistumista puutiaisille eri ympäristöissä (Zhu ym. 2015; Sormunen ym. 2023a). Lisäksi havaintotiedon jakaminen voi vaikuttaa käyttäytymiseen eri tavoin kuin haluttujen lajien tapauksessa. Sen sijaan, että ihmiset hakeutuisivat korkean altistumisriskin alueille, he saattavat pyrkiä välttämään niitä. Tämä saattaa muuttaa tai heikentää ilmiötä, mutta ei kuitenkaan kokonaan poista vinoumaa.

Spatiaalista vinoumaa voidaan selittää useilla ympäristöä ja ihmistoimintaa kuvaavilla tekijöillä. Esimerkiksi Tanskassa neljästä kansalaistiedehankkeesta tehdyt analyysit osoittavat, että maankäyttö, väestötiheys ja tieverkon tiheys ovat merkittävimpiä havaintojen alueellista jakautumista ohjaavia muuttujia (Geldmann ym. 2016). Myös Zhang (2020) osoittaa eBird-aineistoon perustuvassa globaalin tason analyysissä, että havaintotiheyttä selittävät tieverkoston tiheys, väestömäärä, inhimillisen kehityksen taso (HDI) sekä paikalliset kieli- ja kulttuuritekijät. Näiden tekijöiden tunnistaminen on välttämätöntä, jotta aineistoon liittyviä vinoumia voidaan ymmärtää ja niiden mahdolliset vaikutukset havaittuihin alueellisiin rakenteisiin voidaan ottaa huomioon tulkinnessa.

Kansalaistiedeaineistoja on perinteisesti parannettu suodattamalla tai standardoimalla havaintoja niin, että mukaan jätetään vain luotettaviksi katsotut havainnot (Carvalho ym. 2013). Tämä lähestymistapa voi kuitenkin pienentää aineiston kokoa huomattavasti, eikä havaintojen luotettavuutta ole aina mahdollista arvioida jälkikäteen erityisesti silloin, kun saatavilla on niukasti metadataa. Menetelmällinen kehitys tarjoaa kuitenkin vaihtoehtoja, joiden avulla vinoumia voidaan huomioida esimerkiksi mallinnuksen keinoin. Näihin kuuluvat muun muassa lajikertymäkäyrään perustuvat otantakattavuuden arviointimenetelmät, käyntikohtaiseen lajimäärään perustuva havaintointensiteetin arviointi sekä esiintymismallinnus, jossa erotetaan havaittavuus ja lajin todellinen esiintyminen toisistaan (Warton ym. 2013).

Menetelmät perustuvat kuitenkin oletuksiin havaitsijoiden käyttäytymisestä, eivätkä aina huomioi ulkopuolisia taustatekijöitä, jotka vaikuttavat havaintojen spatiaaliseen jakautumiseen (Geldmann ym. 2016). Ihmiset tekevät havaintoja usein lähellä asuinpaikkaansa sekä alueilla, joilla he viettävät aikaa tai paikoissa, jotka tunnetaan luonnon monimuotoisuudesta. Lisäksi tieverkoston on osoitettu vaikuttavan kasvilajien havaintojen määrään siten, että havaintojen lukumäärä kasvaa, kun etäisyys teihin pienenee. Näin ollen kansalaistiedeaineistojen spatiaalinen vinouma muodostuu sekä ympäristötekijöiden että ihmisten liikkumisen ja saavutettavuuden yhteisvaikutuksesta, jota voidaan pyrkiä hillitsemään tilastollisen mallinnuksen keinoin.

#### 2.2.4 Ajallinen vinouma

Ajallinen vinouma tarkoittaa havaintojen epätasaista jakautumista siten, että ne painottuvat tiettyihin vuorokauden-, viikon- tai vuodenaikojen ajankohtiin. Lisäksi ajallista vinoumaa vahvistaa havaintomäärien nopea kasvu, jota on havaittu kansalaistiedehankkeissa viime vuosikymmenten aikana (Di Cecco ym. 2021).

Kansalaisten tuottamissa aineistoissa havaintojen määrä vaihtelee voimakkaasti eri ajankohtien välillä, mikä heijastaa ensisijaisesti havainnoijien aktiivisuutta, eikä niinkään ekologisten ilmiöiden todellista ajallista vaihtelua (Sauermann ym. 2015; Zhang 2020).

Useat tutkimukset osoittavat, että kansalaishavainnointi painottuu viikonloppuihin, leudompiin sääolosuhteisiin ja tiettyihin vuodenaikoihin, erityisesti kevääseen (Rosário

ym. 2024; Díaz-Calafat ym. 2024). Esimerkiksi Rosário ym. (2024) havaitsivat iNaturalist-aineistoon perustuvassa analyysissään, että viikonpäivä, kuukausi ja lämpötila olivat keskeisimpiä havainnointia selittäviä tekijöitä. Havaintoja tehtiin selvästi enemmän viikonloppuisin ja suotuisissa sääolosuhteissa, kun taas äärimmäiset lämpötilat, runsas sade ja voimakas tuuli vähensivät havaitsijoiden aktiivisuutta (Rosário ym. 2024). Lomapäivien ja lumipeitteen vaikutus oli sen sijaan vähäinen. Samankaltaisia viikonpäiviin ja vuodenkiertoon liittyviä ajallisia vinoumia on havaittu myös tieteellisissä aineistoissa ja kansalaistiedeaineistoissa (Díaz-Calafat ym. 2024).

Ajalliset vinoumat ovat ongelmallisia silloin, kun aineistoja käytetään muutosten tarkasteluun, kuten populaatiotrendien, levinneisyyden muutosten tai fenologisten ilmiöiden analysointiin (Johnston ym. 2023). Havaintojen ajallinen epätasaisuus voi johtaa siihen, että havaitut muutokset heijastavat ensisijaisesti havaintorytmin vaihtelua, eivätkä todellisia ekologisia muutoksia (Backstrom ym. 2025; Rosário ym. 2024). Esimerkiksi muuttolintujen ensihavaintoja raportoidaan todennäköisemmin viikonloppuisin, mikä voi vääristää fenologisten ajoitusten tulkintaa, ellei viikonpäivien vaikutusta huomioida mallissa (Díaz-Calafat ym. 2024).

Myös havaintomäärien yleinen kasvu ajan kuluessa vahvistaa ajallista vinoumaa. Kansalaistiedeaineistojen laajentuminen ja käyttäjien määrän lisääntyminen johtavat siihen, että raportoitujen havaintojen määrä kasvaa merkittävästi vuositasonalla (Di Cecco ym. 2021; Backstrom ym. 2025). Tämä voi vääristää populaatio- ja fenologisia analyysejä, jos muutoksia ei suhteuteta aineiston kokonaismäärään. Ajallista vinoumaa voidaan kuitenkin hallita esimerkiksi normalisoimalla havaintomäärän kunkin vuoden tai ajanjakson kokonaisaineistoon, mikä parantaa eri vuosien vertailtavuutta ja ekologisten prosessin arvioinnin luotettavuutta (Backstrom ym. 2025). Myös Fink ym. (2023) ovat tutkimuksessaan erottaneet ekologisia muutoksia havaintomäärien vuosittaisesta vaihtelusta. Aineiston ajallinen epätasaisuus voidaan huomioida tilastollisesti ennen trendien arviointia, jotta havaintomäärien vuosittainen vaihtelu ei vääristäisi tuloksia (Fink ym. 2023).

Ajallista vinoumaa vahvistaa myös havaitsijoiden osallistumisen epätasaisuus. Valtaosa vapaaehtoisista osallistuu aineiston keruuseen vain satunnaisesti, kun taas pieni joukko aktiivisia käyttäjiä voi tuottaa suuren osan havainnoista (Sauermann ym. 2015). Lisäksi havainnoijien kiinnostuksen kohteet ja raportointikäytännöt voivat muuttua ajan myötä, mikä voi synnyttää uusia ajallisia vinoumia ja vaikeuttaa ekologisten muutosten tulkintaa (Johnston ym. 2023).

Ajalliset vinoumat kytkeytyvät usein myös spatiaalisiin vinoumiin, jolloin havainnointiajankohdat ja -paikat eivät ole toisistaan riippumattomia. Havaintomäärien kasvu tiettyinä ajanjaksoina edellyttää havaintojen suhteuttamista ajankohdan kokonaisaineistoon eri elinympäristöissä (Backstrom ym. 2025). Ajallinen vinouma voi ilmetä useilla eri aikaskaaloilla, kuten vuorokaudenajan, vuodenaikojen ja pitkän aikavälin havaintokertymän vaihteluna. Nämä vinoumat voivat kytkeytyä toisiinsa ja yhdessä spatiaalisen vinoumien kanssa vahvistaa havaintoaineiston epätasaisuutta. Tämän vuoksi ajallisten vinoumien tunnistaminen ja huomioiminen on keskeistä kansalaistiedeaineistojen hyödyntämisessä (Rosário ym. 2024).

#### 2.2.5 Havainnointiin liittyvä vinouma

Havaitsemisvinouma (*engl.* detection bias) tarkoittaa, että kaikkia lajeja ei havaita tasaisesti kaikissa tilanteissa, jolloin havaintoaineisto voi painottua tiettyihin lajeihin tai kuvata lajistoa puutteellisesti (Millar ym. 2018). Kansalaistiedeaineistossa tämä vinouma ilmenee, koska vapaaehtoisten havainnointikyky, kiinnostuksen kohteet ja havaintotiheys vaihtelevat merkittävästi (Dickinson 2010; Gutierrez-Velez & Wiese 2020). Lisäksi lajien esiintyminen voi olla ajallisesti vaihtelevaa, mikä aiheuttaa kausittaisia havaintovinoumia. Opportunistisessa kansalaistiedeaineistossa, jossa havainnot kerätään ilman otantasuunnitelmaa, havaitsemisvinouma esiintyy yhdessä raportointi- ja otantavinoumien kanssa (van Strien ym. 2013). Näiden vinoumien yhteisvaikutus voi johtaa ekologisesti harhaanjohtaviin levinneisyys- ja populaatiotrendeihin.

Kaikkia lajeja ei havaita yhtä helposti. Näkyvät ja suuret lajit, kuten runsaslukuiset ja ruokintapaikoilla vierailevat lajit, havaitaan todennäköisemmin kuin piilossa olevat, harvinaisemmat lajit (Kelling ym. 2015). Havaintotodennäköisyydellä tarkoitetaan

todennäköisyyttä havaita laji silloin, kun se on paikalla. Se vaihtelee sekä lajin ominaisuuksien että havainnoijan mukaan (Dickinson ym. 2010; Royle ym. 2007). Vähemmän kokeneet havainnoijat voivat tehdä sekä virheellisiä poissaolohavaintoja (*engl.* false absences) että virheellisiä läsnäolohavaintoja (*engl.* false positives). Tämä korostuu erityisesti ekologisissa kansalaistiedeaineistoissa, jossa havaintoja raportoidaan yleensä vain läsnäolotietoina (*engl.* presence-only). Tällaisilla sivustoilla, kuten iNaturalist tai Tiira, poissaoloja ei kirjata, mikä voi aiheuttaa aineistoon omia haasteita. Kokeneemmat havainnoijat sen sijaan havaitsevat enemmän lajeja ja tunnistavat ne tarkemmin (Kelling ym. 2015).

Havaitsemisvinoumaa voidaan arvioida tarkastelemalla havaintotiheyttä ja havainnoinnin ajankohtaa suhteessa lajin havaitsemistodennäköisyyteen (Gutierrez-Velez & Wiese 2020). Jos havaintoja kertyy pääasiassa alueilta tai ajanjaksoilta, joissa laji on helposti havaittavissa, aineisto voi heijastaa havaittavuuden vaihtelua todellisen esiintymisen sijaan. Myös havaitsijoiden kokemus ja koulutus vaikuttavat havaintojen tarkkuuteen (Dickinson ym. 2010; Kelling ym. 2015). Lisäksi Gutierrez-Velez ja Wiese (2020) ovat soveltaneet havaintojen luotettavuuteen perustuvaa painotusta, jonka avulla voidaan vähentää havaitsemisvinoumien vaikutusta ja parantaa mallien ennustetarkkuutta. Menetelmä huomioi tilanteet, joissa lajia ei havaita sen esiintymisestä huolimatta, ja ottaa tämän vaikutukset huomioon mallinnuksessa. Tämä korostaa havaitsijoiden taitotason ja lajien havaittavuuden huomioimisen merkitystä kansalaistiedeaineistojen analyysissä.

Havaintovirheiden huomioiminen on tärkeää myös analyysien luotettavuuden kannalta. Virheellisten poissaolo- tai läsnäolohavaintojen sivuuttaminen voi vääristää mallien arvioita ja johtaa liian suureen luottamukseen tulosten tarkkuudesta (Guillera-Arroita 2016). Vinouman voimakkuuteen vaikuttavat sekä lajin havaittavuus että havaintotodennäköisyyden yhteys ympäristötekijöihin (Guillera-Arroita 2016; Gutierrez-Velez & Wiese 2020). Arvioiden tarkkuutta voidaan parantaa lisäämällä havaintopanosta esimerkiksi toistamalla havaintoja useammin, pidentämällä havaintoaikaa tai käyttämällä tarkempia havainnointimenetelmiä (Guillera-Arroita 2016).

Suurin osa vakiintuneista ekologisista kansalaistiedehankkeista keskittyy tällä hetkellä lintulajeihin, mutta myös nisäkkäisiin ja muihin eliöryhmiin painottuvat hankkeet ovat lisääntyneet viime vuosina (Ditmer ym. 2020). Koska nisäkkäät ovat lintuja vaikeammin havaittavissa, niihin liittyvä kansalaistiedeaineisto on erityisen altis havainnointivinoumille. Tämän vuoksi ihmistoiminnan muuttuvia vaikutuksia on arvioitava huolellisesti, jotta aineistosta voidaan tehdä ekologisesti mielekkäitä johtopäätöksiä. Havaittavuuteen liittyvien tekijöiden lisäksi kansalaistiedeaineistojen rakenteeseen vaikuttavat myös havaitsijoiden motivaatio, kokemus ja raportointikäytännöt. Siten osa vinoumista kytkeytyy itse havaintoprosessin sijaan siihen, millaisia valintoja osallistujat tekevät havaintoihin liittyen.

## 2.2.6 Osallistujiin liittyvä vinouma

Osallistujiin liittyvillä vinoumilla (*engl.* observer bias) tarkoitetaan kansalaistiedeaineistoissa havaitsijoiden välisestä vaihtelusta johtuvia eroja siinä, mitä havaintoja kerätään ja raportoidaan sekä missä, milloin ja miten tämä tapahtuu (Arazy & Malkinson ym. 2021). Toisin kuin havainnointivinouma, joka liittyy lajien vaihtelevaan havaittavuuteen, havaitsijavinouma kytkeytyy havaitsijoiden taustatekijöihin, päätöksentekoon, motivaatioon ja raportointikäytäntöihin (Dickinson ym. 2010; Arazy & Malkinson ym. 2021). Näitä vinoumia voidaan jäsentää havaintoprosessin eri vaiheiden kautta. Havaintoprosessi etenee vaiheittain: ensin laji havaitaan, sitten havaitsija arvioi havainnon merkityksellisyyttä raportoinnin kannalta ja lopuksi päättää, raportoidaanko se.

Kansalaistiedeaineistojen laadun tarkastelu on perinteisesti painottunut havaitsijoiden taitotasoon ja siihen, miten vapaaehtoisten tuottama aineisto vertautuu ammattilaisten keräämään aineistoon (Dickinson ym. 2010). Kuten ekologian kenttätöissä yleensäkin, myös kansalaistieteessä havaitsijoiden kokemus, koulutus ja asiantuntemus vaihtelevat merkittävästi. Puutteellinen koulutus tai ohjeistus voi lisätä virheiden ja vinoumien määrää, vaikka toisaalta itseohjautuva oppiminen ja pitkäaikainen osallistuminen voivat parantaa havaitsijoiden suoriutumista ajan kuluessa (Dickinson ym. 2010; Kelling ym. 2015). Vapaaehtoisten ja ammattilaisten vertailu osoittaa, että vapaaehtoiset voivat tuottaa ammattilaistasoista aineistoa tietyissä tehtävissä ja

eliöryhmissä. Sen sijaan vaativissa taksonomisissa ryhmissä havaittajien väliset erot korostuvat (Brown & Williams 2018).

Keskeinen osallistujiin liittyvän vinouman lähde on havaittajien vaihteleva osallistumisaktiivisuus. Kansalaistiedehankkeissa osallistumisen kynnys on usein matala, minkä seurauksena mukana on havaittajia, joiden sitoutuminen, osaaminen ja tavoitteet vaihtelevat huomattavasti (Kelling ym. 2015). Valtaosa osallistujista tuottaa vain vähän havaintoja, kun taas pieni joukko aktiivisia havainnoijia vastaa usein suuresta osasta havaintoja. Tämä epätasainen osallistujajakauma voi johtaa siihen, että aineisto heijastaa voimakkaasti aktiivisten havaittajien kiinnostuksen kohteita ja toimintamalleja kaikkien havaittajien keskimääräisen käyttäytymisen sijaan (Johnston ym. 2023)

Arazy ja Malkinson (2021) kuvaavat havaittajavinoumaa päätöksentekoprosessina, jossa havaittajan toiminta voidaan nähdä sarjana valintoja. Ennen varsinaista havaitsemista havaittaja tekee päätöksen siitä, mitä lajeja seurataan, missä ja milloin havainnointi tehdään sekä kuinka paljon siihen panostetaan. Nämä valinnat eivät ole satunnaisia, vaan ne heijastavat havaittajan mieltymyksiä, tavoitteita, käytettävissä olevaa aikaa ja välineistöä sekä arkisia liikkumismalleja (Arazy & Malkinson 2021; Johnston ym. 2023). Tämän seurauksena havaintoaineistoon voi syntyä systemaattisia vinoumia, jotka liittyvät esimerkiksi tiettyjen lajien suosimiseen tai havaintojen keskittymiseen helposti saavutettaville alueille.

Havaittajien mieltymykset ja motivaatio muodostavat keskeisen, mutta usein hankalasti mitattavan vinouman lähteen. Osa lajeista herättää erityistä kiinnostusta esteettisten ja suojelullisten syiden sekä niiden harvinaisuuden vuoksi, jolloin niitä raportoidaan tavanomaista useammin (Scher & Clark 2023). Tällainen valikoiva raportointi ei liity siihen, havaitaanko laji, vaan siihen, koetaanko havainto raportoinnin arvoiseksi (Johnston ym. 2023). Tämä erottaa havaittajien mieltymyksiin ja motivaatioon liittyvät vinoumat havainnointivinoumasta, sillä kyse ei ole lajien havaittavuudesta, vaan havaittajien päätöksistä siitä, mitä havaintoja raportoidaan (Arazy & Malkinson 2023).

Raportointiprosessin seuraavassa vaiheessa voi syntyä niin sanottu kirjaamisvinouma, joka tarkoittaa vaihtelua siinä, kirjataanko havainto sen jälkeen, kun laji on havaittu ja tunnistettu (Arazy & Malkinson 2021). Kirjaamispäätökseen vaikuttavat havaitsijan omat käsitykset havaintojen merkityksestä, käytettävissä oleva teknologia sekä käytännön rajoitteet, kuten ajankäyttö tai internetyhteys (Johnston ym. 2023; Arazy & Malkinson 2021). Esimerkiksi älypuhelinsovellukset mahdollistavat havaintojen välittömän jakamisen, kun taas vaativampi välineistö, kuten järjestelmäkamerat, voivat viivästyttää tai estää havaintojen raportoimista (Dickinson ym. 2012; Arazy & Malkinson 2021). Näin aineistoon voi syntyä vinoumaa, joka heijastaa enemmän havaitsijoiden raportointikäytäntöjä kuin tarkasteltavan ilmiön todellista esiintymistä.

Osallistujiin liittyvien vinoumien voimakkuuteen vaikuttavat myös kansalaistiedehankkeiden rakenteelliset erot. Jäsennellyissä kansalaistiedehankkeissa havaitsijoiden toimintaa ohjaavat selkeät toimintaohjeet, jotka rajaavat valinnanvapautta ja vähentävät yksilöllisestä käyttäytymisestä aiheutuvia vinoumia (Brown & Williams 2018; Arazy & Malkinson 2021). Ei-systemaattisissa hankkeissa havaitsijoiden valinnat korostuvat, mikä heikentää mahdollisuuksia tehdä tilastollisesti vahvoja päätelmiä ilman erillisiä korjausmenetelmiä (Kelling ym. 2015; Arazy & Malkinson 2021). Esimerkiksi vakioiduissa lajiseurannoissa havaintojen keruu perustuu ennalta määritettyihin reitteihin ja aikatauluihin, kun taas opportunistisissa kansalaistiedehankkeissa tarkkailijat päättävät itse havainnointipaikoista- ja ajankohdista. Mitä suurempi vaikutus havaitsijoiden omilla päätöksillä on, sitä merkittävämmäksi vinouma voi muodostua.

Havaitsijavinoumien merkitys korostuu pitkissä aikasarjoissa, joissa havaitsijajoukko muuttuu ajan myötä. Osallistujarakenteen muutokset ja teknologian kehittyminen voivat vaikuttaa havaintoaineiston rakenteeseen ilman, että taustalla oleva ekologinen prosessi on muuttunut (Gonsamo & D’Odorico 2014; Johnston ym. 2023). Siten havaitsijoiden välisen vaihtelun tunnistaminen ja huomioiminen on tärkeää, kun kansalaistiedeaineistoja käytetään ajallisten trendien ja laajamittaisten ekologisten muutosten tarkasteluun (Brown & Williams 2018). Kokonaisuudessaan vinoumien ymmärtäminen on välttämätöntä, jotta kansalaistiedeaineistoista voidaan tehdä

luotettavia ekologisia johtopäätöksiä ja erottaa havaitsijoiden toiminnasta johtuva vaihtelu varsinaisista biologisista ilmiöistä (Ditmer ym. 2020; Arazy & Malkinson 2021).

## 2.3 Puutiaisen ekologia ja puutiaishavaintojen vinoumat

### 2.3.1 Puutiaisen ekologia ja levinneisyys Suomessa

Suomessa tavataan kahta ihmisiä purevaa puutiaislajia: Etelä- ja Lounais-Suomessa yleistä puutiaista (*Ixodes ricinus* L.) (tästä eteenpäin puutiainen) sekä idän suunnalta levinnyttä taigapunkkia (*Ixodes persulcatus* Schulze), jota esiintyy erityisesti Itä- ja Pohjois-Suomessa mutta myös länsirannikolla (Kulha ym. 2022; Sormunen ym. 2023a). Molempien lajien esiintymisalueet ovat laajentuneet viime vuosikymmeninä pohjoiseen sekä taigapunkin myös kohti länsi- ja luodesuuntaa. Tätä kehitystä selittävät muun muassa ilmastonmuutoksen myötä pidentynyt kasvukausi ja leudommat talvet (Sormunen ym. 2025). Lisäksi ympäristö- ja maankäyttömuutokset voivat tarjota puutiaisille uusia elinympäristöjä ja lisätä sopivien isäntäeläinten runsautta, mikä edesauttaa niiden levittäytymistä uusille alueille. Varsinais-Suomen alueella havaintojen voidaan olettaa koskevan ainoastaan puutiaista (Laaksonen ym. 2017).

Puutiaisella on nelivaiheinen elinkaari, joka koostuu muna-, toukka-, nymfi- ja aikuisvaiheesta (Randolph 2004). Toukkavaiheesta eteenpäin puutiaiset tarvitsevat veriaterian kehittyäkseen seuraavaan vaiheeseen. Aikuinen naaras munii tyypillisesti 2000–3000 munaa, jotka se laskee maahan pudonneille lehdille tai pintamaahan (Tanski 2023). Munista kuoriutuvat toukat kiinnittyvät isäntäeläimiin, jotka voivat olla esimerkiksi lintuja, jyrsijöitä tai hirvieläimiä. Myös nymfi- ja aikuisvaiheen puutiaiset loisivat tasa- tai vaihtolämpöisissä eläimissä, suosien keskimäärin kookkaampia isäntiä kuin toukat.

Yhden kehitysvaiheen aikana puutiainen pysyy isäntäeläimeen kiinnittyneenä useita päiviä, minkä jälkeen se pudottautuu maahan ja siirtyy seuraavaan kehitysvaiheeseen. Koko elinkierto kestää puutiaisella tyypillisesti kahdesta kuuteen vuotta ympäristöolosuhteista ja isäntäeläinten saatavuudesta riippuen (Tanski 2023). Puutiaisten aktiivisuus ja kehitysvaiheiden ajoittuminen määräytyvät erityisesti lämpötilan ja suhteellisen ilmankosteuden mukaan. Suotuisissa olosuhteissa

puutiaiset aktivoituvat ja siirtyvät kasvillisuuden alemmista osista ylöspäin etsimään isäntäeläimiä (Estrada-Peña ym. 2013). Epäsuotuisissa olosuhteissa puutiaiset voivat puolestaan viettää pitkiä jaksoja passiivisina karikkekerroksessa (Randolph 2004; Estrada-Peña ym. 2013). Elinkierron hitaus ja vaihteittainen riippuvuus isäntäeläimistä tekevät puutiaisista herkkiä elinympäristön muutoksille, mutta samalla mahdollistavat niiden selviytymisen vaihtelevissa olosuhteissa.

Puutiaiset ovat sidoksissa elinympäristöihin, jotka tarjoavat niille suotuisat ilmasto-olosuhteet ja riittävän määrän isäntäeläimiä (Estrada-Peña ym. 2013; Bourdin ym. 2022). Suomessa puutiainen suosii erityisesti metsäisiä ympäristöjä, joissa maanpintaa peittävä karike- tai sammalkerros ylläpitää korkeaa ilmankosteutta (Tanski 2023). Tällaisia elinympäristöjä ovat esimerkiksi lehtimetsät, mustikkatyyppin kangasmetsät ja suurten vesistöjen läheiset metsäalueet. Puutiaisia havaitaan myös avoimissa ja ihmisen muokkaamissa ympäristöissä, kuten kaupunkien puistoissa ja pihoilla, joissa viheralueet voivat tarjota niille suojaa esimerkiksi kuivuudelta (Janzén ym. 2024; Sormunen ym. 2025). Lisäksi alueella elävien ihmisten määrä sekä luonnossa liikkuminen ja virkistyskäyttö vaikuttavat siihen, missä ja milloin puutiaisia havaitaan (Sormunen ym. 2020; Sormunen ym. 2023a). Runsaat puutiaishavainnot voivat siten heijastaa korkeita puutiaistiheyksiä, ihmisten aktiivisuutta tai näiden tekijöiden yhteisvaikutusta (Sormunen ym. 2016; Sormunen ym. 2023a).

### 2.3.2 Puutiaishavaintojen erityispiirteet kansalaistiedeaineistossa

Toisin kuin esimerkiksi lintuhavainnot, ihmisten tekemät puutiaishavainnot eivät yleensä perustu aktiiviseen havainnointiin tai tavoitteelliseen etsintään. Puutiaiset havaitaan useimmiten sattumanvaraisesti ihmisen tai muun isäntäeläimen kohtaamisen yhteydessä. Havaintojen ajallinen ja spatiaalinen jakauma heijastaa samanaikaisesti puutiaisten ekologista esiintymistä, isäntäeläinten liikkeitä sekä ihmisen toimintaa (Laaksonen ym. 2017; Johnston ym. 2023). Kansalaishavainnot voivat lisäksi kuvata alueellisia eroja ihmisten altistumisessa puutiaiskontakteille, jolloin osa aineistossa havaittavista vinoumista voi liittyä todellisiin altistumisriskeihin pelkän havaintosijatoiminnan epätasaisuuden sijaan (Sormunen 2026). Esimerkiksi havaintojen painottuminen tiheästi asutuille alueille tai viikonlopuille voi liittyä siihen, että ihmiset

liikkuvat tällöin enemmän luonnossa ja altistuvat siten todennäköisemmin puutiaisille. Tällöin osa havaintoaineiston spatiaalisista ja ajallisista vinoumista voi tarjota myös epidemiologisesti tärkeää tietoa puutiaiskontaktien riskialueista ja -ajankohdista.

Puutiaisten tutkimus kansalaishavainnoinnin kohteena kytkeytyy niiden elinkiertoon ja ajoittaiseen loissuhteeseen. Puutiainen viettävät suurimman osan elinkaarestaan vapaana elinympäristössään, missä ne altistuvat voimakkaasti mikroilmaston ja elinympäristön vaihteluille (Randolph 2004; Ehrmann ym. 2017). Verenimun aikana ne voivat kuitenkin kulkeutua isäntäeläinten mukana huomattavia matkoja, mikä voi irrottaa yksittäisen havainnon sen alkuperäisestä ympäristöstään (Tanski 2023). Tämän vuoksi puutiaisia havaitaan harvoin suoraan luonnossa, vaan useimmiten ihmisten tai lemmikkieläinten iholta.

Isaac ym. (2014) tunnistavat ekologisissa kansalaistiedeaineistoissa neljä keskeistä vinumatyyppiä: spatiaalisen, ajallisen, havainnointiin liittyvän sekä osallistujiin liittyvän vinouman, jotka voidaan havaita myös puutiaisaineistossa. Ajallinen vinouma kytkeytyy erityisesti puutiaisten biologiseen aktiivisuuteen, joka vaihtelee vuodenaikojen ja sääolosuhteiden mukaan (Estrada-Peña ym. 2013). Havaintoja kertyy eniten ajanjaksoina, jolloin puutiaiset ovat aktiivisia ja ihmiset liikkuvat luonnossa, mikä korostaa puutiaisten aktiivisuuskausia (Sormunen ym. 2023). Puutiaishavaintojen tutkimuksessa korostuvat kuitenkin erityisesti havaintojen spatiaaliseen jakautumiseen liittyvät tekijät, sillä havainnot kytkeytyvät voimakkaasti ihmisten liikkumiseen ja altistumiseen eri ympäristöissä.

Spatiaalinen vinouma on keskeinen puutiaishavaintoaineistoissa. Havaintoja kertyy erityisesti asutuilta alueilta, vapaa-ajan ympäristöistä, kulkuyhteyksien läheisyydestä sekä paikoista, joissa ihmiset liikkuvat aktiivisesti (Sormunen ym. 2023a; Zhang 2020). Tämän vuoksi havaintotiheys ei kuvaa suoraan eliöiden todellista esiintymistiheyttä, vaan pikemmin ihmistoiminnan alueellista jakautumista. Sormunen ym. (2023a) osoittavat, että vaikka väestömäärä selittää yleisesti hyvin puutiaishavaintojen määrää Suomessa, havaintomäärissä esiintyy alueellisia poikkeamia, jotka voivat liittyä puutiaisten runsauteen, ympäristötekijöihin tai eroihin havaintokäyttäytymisessä.

Lisäksi puutiaisten kulkeutuminen isäntäeläinten mukana vaikeuttaa havaintojen alkuperäisen paikan määrittämistä (Sormunen ym. 2023b).

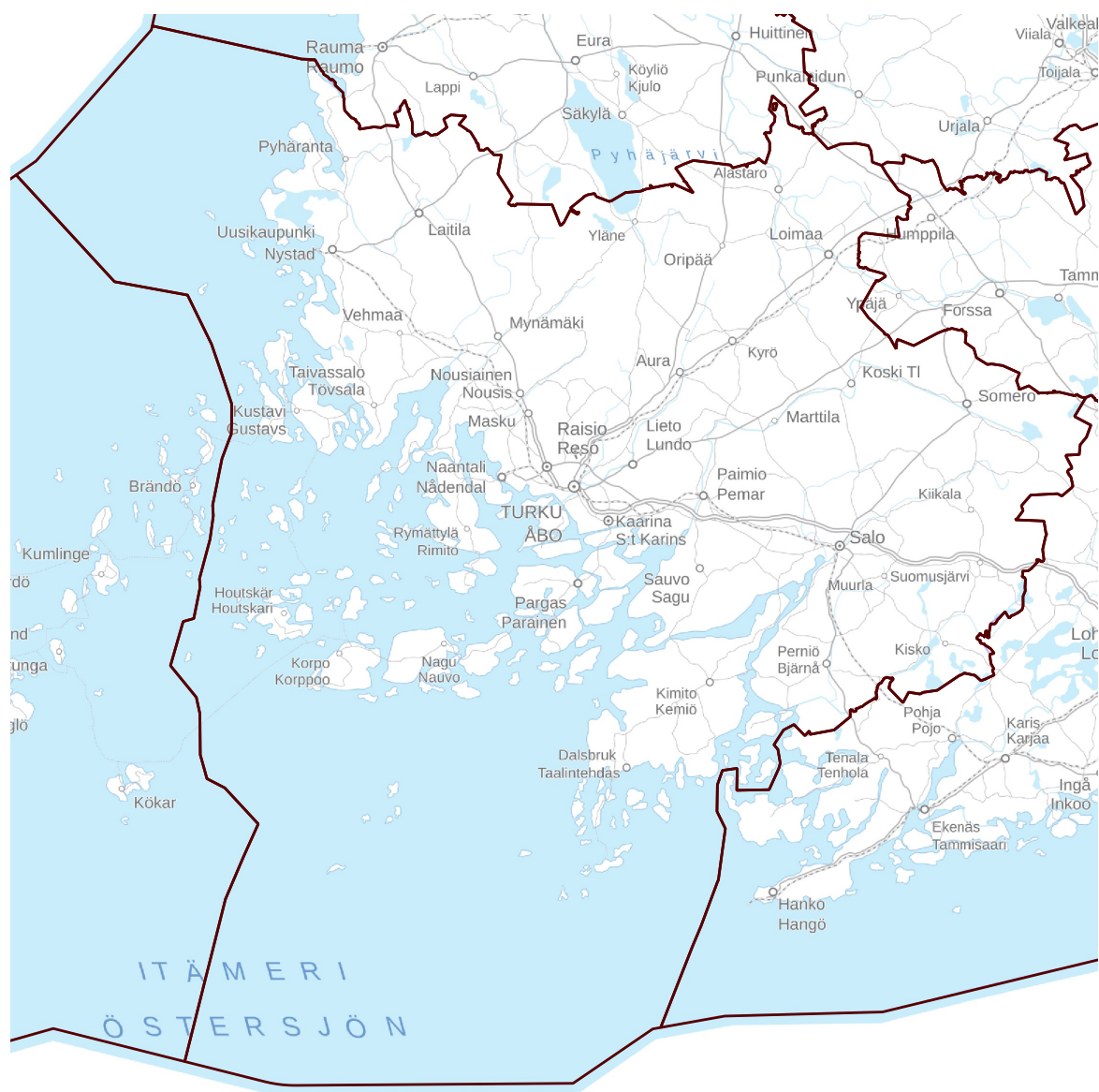
Myös havainnointi- ja raportointivinoumat vaikuttavat puutiaisaineistoihin. Osa kansalaisista raportoi havaintoja aktiivisesti, kun taas toiset ilmoittavat niitä harvemmin (Arazy & Malkinson 2021). Tunnistusvinoumaa on havaittu puutiaisaineistoissa suhteellisen vähän verrattuna moniin muihin eliöryhmiin (Sormunen ym. 2023b; Gorleri & Areta 2022). Suomessa kansalaiset tunnistavat puutiaiset hyvin niiden tunnusomaisen ulkonäön ja verenimukäyttäytymisen sekä laajan mediahuomion vuoksi (Sormunen ym. 2023b).

Näistä rajoitteista huolimatta kansalaistieteeseen perustuvat puutiaisaineistot tarjoavat mahdollisuuden tarkastella puutiaisten esiintymistä laajoilla alueilla ja pitkillä aikajaksoilla, jota perinteisellä kenttätutkimuksella ei ole mahdollista saavuttaa (Dickinson ym. 2010). Tämä korostuu erityisesti tauteja levittävien eliöiden tutkimuksessa, jossa riskialueiden ja aktiivisuuskausien tunnistaminen on tärkeää altistumisriskin arvioinnin kannalta (Sormunen ym. 2023a). Tilastollinen ja spatiotemporaalinen mallinnus mahdollistaa havaintojen kertymiseen vaikuttavien tekijöiden ja ympäristöön liittyvien ekologisten prosessien erottamisen toisistaan (Dickinson ym. 2010; Tanski 2023).

### 3 Aineistot ja menetelmät

#### 3.1 Tutkimusalue

Tämän tutkimuksen tutkimusalue on Varsinais-Suomen maakunta Lounais-Suomessa. Varsinais-Suomen kokonaispinta-ala on 20 537,55 km<sup>2</sup>, josta lähes puolet (9600 km<sup>2</sup>) on merialueita (Kalliolevo ym. 2023). Alue ulottuu Turunmaan ulkosaaristosta sisämaahan Someron-Loimaan alueelle saakka (kuva 1). Maastonmuodoiltaan alue on pääosin tasaista, ja itäisimmät osat yhtyvät läntisen Uudenmaan ylänköalueeseen.



Kuva 1. Tutkimusalueen raja (Maakuntaraja © Syke 2025 & Taustakartta © Maanmittauslaitos 2025).

Varsinais-Suomi kuuluu Köppen-Geigerin ilmastoluokituksessa pääosin Dfb-ilmastotyyppiin, jolle on ominaista kostea mannerilmasto ja lämpimät kesät (Peel ym. 2007). Merellinen sijainti tasaa lämpötilavaihteluita ja lieventää talven pakkasia verrattuna sisämaan alueisiin. Vuoden keskilämpötila vaihtelee ulkosaariston noin +7 °C:sta sisämaan +6 °C:een, ja vuotuinen sademäärä vaihtelee ulkosaariston 500–550 millimetristä sisämaan 600–750 millimetriin (Kalliolevo ym. 2023).

Maankäytöltään Varsinais-Suomen maapinta-alasta suurin osa on metsiä ja maatalousalueita (Kalliolevo ym. 2023). Havumetsät kattavat alueen metsäpeitteestä suurimman osuuden, mutta alueella esiintyy myös sekametsiä ja lehtimetsiä. Metsätyypit vaihtelevat boreaalisista havumetsistä sekametsiin ja rannikon hemiboreaaliin lehtometsiin (Käyhkö & Skånes 2008). Erityisesti rannikko- ja saaristoalueilla esiintyy runsaasti lehtipuuvaltaisia metsiä ja pensaikkoisia alueita.

Varsinais-Suomessa väestö keskittyy erityisesti Turun kaupunkiseudulle ja pääasiassa rannikolla sijaitseviin taajamiin, kun taas sisämaan alueet ovat harvemmin asuttuja (Väestöruutuaineisto 2023). Alueella sijaitsee myös runsaasti vapaa-ajan asuntoja, jotka keskittyvät erityisesti saaristo- ja rannikkoalueille (Maastotietokanta, rakennukset 2026). Lisäksi virkistysalueet, retkeilyreitit ja saaristo houkuttelevat ulkoilijoita ja matkailijoita ympärivuotisesti (Matkailu ja virkistys Varsinais-Suomessa 2026).

Puutiaisten kannalta keskeisiä ympäristötekijöitä ovat riittävä ilmankosteus, leudot talviolosuhteet, suojaava aluskasvillisuus sekä sopiva isäntäeläinten saatavuus (Ehrmann ym. 2017; Tanski 2023). Varsinais-Suomen merellinen ilmasto, pitkä kasvukausi ja suhteellisen korkea sademäärä tarjoavat suotuisat olosuhteet puutiaisten talvehtimiselle ja kehitykselle. Lisäksi alueella esiintyy runsaasti hirvieläimiä, jotka toimivat puutiaisten isäntäeläiminä yksilöiden eri kehitysvaiheissa. Myös ihmisten liikkuminen, luonnon virkistyskäyttö ja asutuksen sijoittuminen vaikuttavat siihen, missä puutiaisia kohdataan ja missä havaintoja syntyy. Edellä kuvattujen ilmastollisten, ekologisten ja saavutettavuuteen kytkeytyvien tekijöiden yhteisvaikutus tekee Varsinais-Suomesta puutiaisen esiintymiselle suotuisan ja ympäristöolosuhteiltaan vaihtelevan alueen, mikä tarjoaa hyvän lähtökohdan puutiaishavaintojen spatiaalisen vaihtelun tarkasteluun.

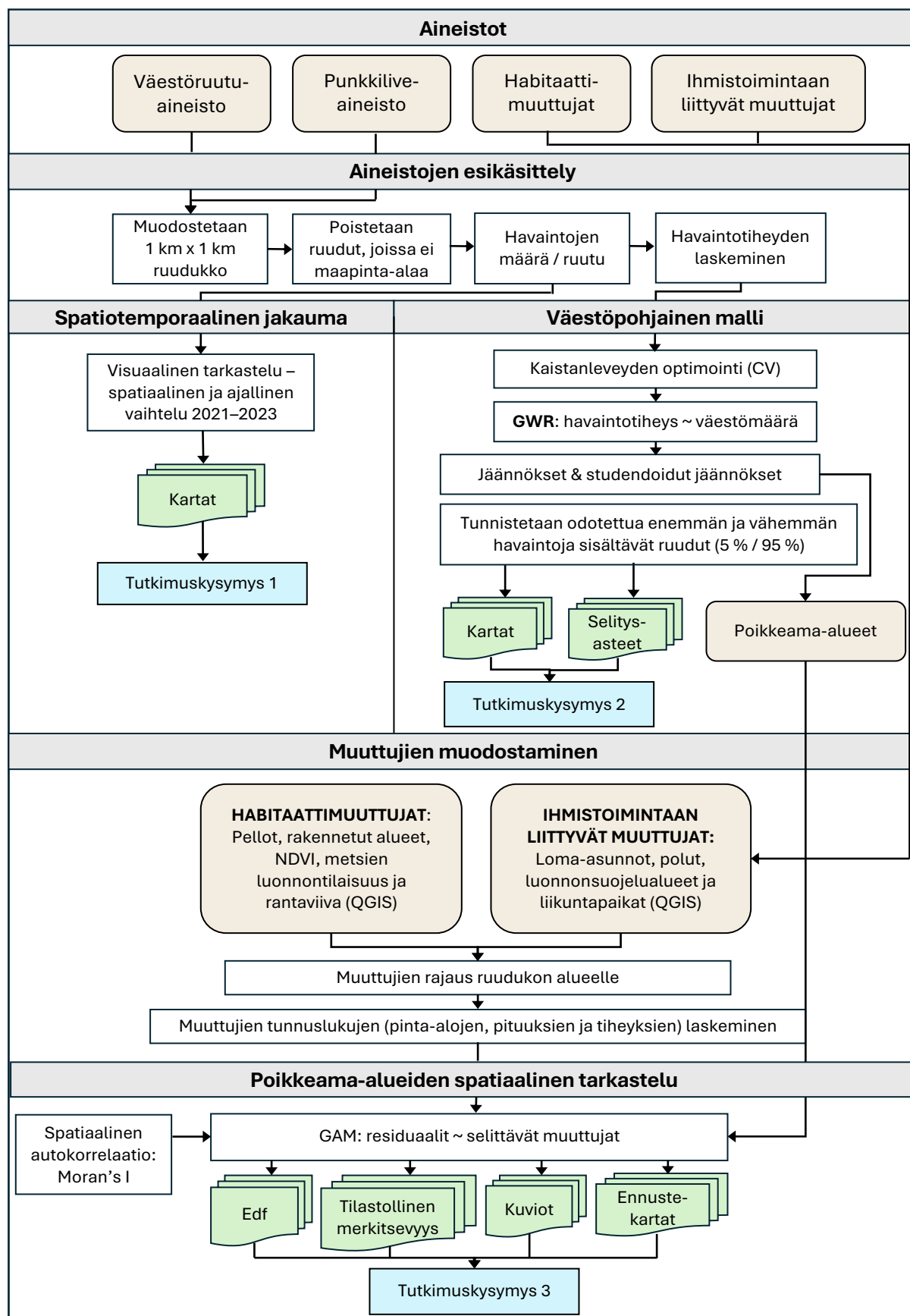
### 3.2 Tutkimuksen metodologinen kokonaisuus

Tutkimukseni metodologinen kokonaisuus rakentuu kolmesta osa-alueesta: (1) puutiaishavaintojen spatiotemporaalisen jakautumisen tarkastelusta, (2) väestömäärään perustuvasta havaintotiheyden mallinnuksesta sekä (3) puutiaishavaintoihin liittyvien alueellisten vinoumien selittämisestä ympäristö- ja ihmistoimintamuuttujien avulla (kuva 2). Menetelmällinen kokonaisuus etenee siten kansalaistieteeseen pohjautuvan puutiaishavaintoaineiston kuvailevasta tarkastelusta mallipohjaiseen analyysiin ja lopulta spatiaalisen vinouman tulkintaan. Ensimmäiseen tutkimuskysymykseen vastatakseni hyödynsin havaintojen alueelliseen visualisointiin myös 3 km x 3 km ruutuaineistoa, jonka avulla mallinsin yleispiirteisesti koko tutkimusalueen havaintojen jakautumista. Varsinaisen mallinnuksen ja poikkeama-alueiden tarkastelun toteutin 1 km x 1 km ruututasolla.

Tarkastelun lähtökohtana on oletus siitä, että kansalaistieteeseen perustuvat puutiaishavainnot eivät kuvaa yksinomaan puutiaisten ekologista esiintymistä, vaan myös ihmisten määrää, liikkumista ja havaintokäyttäytymistä. Tämän vuoksi käytin väestömäärää ensin havaintojen kertymistä selittävänä muuttujana.

Puutiaishavaintojen ja väestömäärän välistä suhdetta tarkastelin maantieteellisesti painotetun regressiomallin avulla, joka mahdollistaa muuttujien välisten suhteiden alueellisen vaihtelun tarkastelun (Brunsdon ym. 1996; Madetoja 2018). Mallin jäännösten avulla tarkastelin puutiaishavaintoihin liittyvää spatiaalista vinoumaa tunnistamalla poikkeama-alueet, joilla havaintoja oli enemmän tai vähemmän kuin väestömäärän perusteella olisi ollut odotettavissa.

Seuraavaksi selitin näitä poikkeama-alueita paikkatietoaineistoista muodostetuilla ympäristö- ja ihmistoimintamuuttujilla. Yleistetty additiivinen malli soveltui tähän tarkasteluun, koska se mahdollistaa epälineaaristen yhteyksien mallintamisen (Wood 2017; Zuur ym. 2009). Tulkitsin mallin tuloksia efektiivisten vapausasteiden, osittaisvaikutuskuvioiden ja ennustekartan avulla. Näin tutkimuksen metodologisen kokonaisuuden avulla voidaan tarkastella, missä puutiaishavainnot poikkeavat väestöpohjaisesta odotuksesta ja mitkä ympäristöön tai ihmistoimintaan liittyvät tekijät selittävät näitä poikkeamia (kuva 2).



Kuva 2. Tutkimuksen metodologinen kokonaisuus.

### 3.3 Aineistot

#### 3.3.1 Punkkilive-kansalaistiedeaineisto

Punkkilive-verkkosivusto (<https://www.punkkilive.fi/>) on suunniteltu Turun yliopiston puutiaisprojektin ja Pfizer Oy Suomen yhteistyönä vuosina 2020–2021 (Sormunen ym. 2023a). Sivuston tavoitteena on tarjota kansalaisille työkaluja puutiaisten ajallisten ja alueellisten riskitekijöiden itsenäiseen seurantaan ja arviointiin eri puolilla Suomea. Sivusto sisältää sekä tekstimuotoista tietoa puutiaisista että reaaliaikaisesti päivittyvän karttanäkymän. Punkkiliven ”Punkkihavainnot kartalla”-sivulla käyttäjät voivat tarkastella karttaa, joka esittää havaintoja eri ajanjaksoina (Punkkilive 2026). ”Ilmoita puutiaishavainto” -sivulla käyttäjät voivat raportoida havainnon joko etsimällä sijainnin osoitehaun avulla, selaamalla karttaa ja sijoittamalla havaintokuvakkeen manuaalisesti tai näiden yhdistelmänä. Havainnon koordinaatit tallentuvat automaattisesti käyttäjän antaman sijaintitiedon perusteella.

Puutiaishavaintoa ilmoitettaessa käyttäjiltä kysytään neljä kysymystä:

- 1) havainnon ajankohta
- 2) havaittujen puutiaisten lukumäärä (kolme luokkaa: 1, 2–5 tai yli 5)
- 3) isäntä tai ympäristö, josta puutiainen löytyi (eläin, ihminen tai luonto; eläinluokassa alaluokat: koira, kissa ja muu) sekä
- 4) onko käyttäjä havainnut puutiaisia alueella aiemmin

Näistä havainnon ajankohta ja havaittujen puutiaisten lukumäärä ovat pakollisia tietoja, kun taas muihin kysymyksiin vastaaminen on vapaaehtoista. Lisäksi käyttäjillä on mahdollisuus liittää havaintoon valokuvia.

Tässä tutkielmassa hyödynsin Punkkilive-aineistoa vuosilta 2021, 2022 ja 2023.

Aineistoa kerätään puutiaisten esiintyvyyden mahdollisimman ajantasaisen ja tarkan kartoituksen tueksi. Koko vuoden kattavaa aineistoa ei ole julkisesti saatavilla, vaan olen saanut sen koottuna vuosikohtaisina CSV-muotoisina tiedostoina. Vuoden 2021

tarkastelujakso kattoi ajanjakson 15.4.–31.12., kun taas vuosien 2022 ja 2023 aineistojen ajanjakso kattoi koko kalenterivuoden 1.1.–31.12.

Hyödynsin aineistosta havainnon ajankohtaa vuositasolla sekä sijaintitietoja. Lopulta toin havainnot analyysia varten R-ohjelmistoon CSV-muodossa. Punkkilive-aineiston havainnot muodostavat perustan ensimmäiseen tutkimuskysymykseen vastaamiselle, sillä sen avulla tarkastelin puutiaishavaintojen alueellista ja ajallista jakautumista Varsinais-Suomessa vuosina 2021–2023.

### 3.3.2 Muut paikkatietoaineistot

Paikkatietoaineistojen valinta perustui tutkimuskirjallisuuteen, jonka pohjalta tunnistin puutiaisten esiintymiseen vaikuttavia keskeisiä ympäristöön tai ihmistoimintaan liittyviä tekijöitä. Näiden perusteella määrittelin tutkimuksessa käytettävät muuttujat, jotka kuvaavat joko habitaattiolosuhteita tai ihmisten aktiivisuutta (taulukko 2). Tämän jälkeen kokosin muuttujia vastaavat paikkatietoaineistot analyysia varten.

Habitaattimuuttujat kuvaavat ympäristön soveltuvuutta puutiaisille. Ihmisten liikkumista kuvaavat muuttujat ovat puolestaan tärkeitä kansalaistiedeaineiston näkökulmasta, sillä ihmisten toiminta ja liikkuminen ohjaavat vahvasti havaintojen kertymistä. Näin huomioin aineistovalinnassa sekä puutiaisten ekologiset elinympäristötekijät että havaintojen syntymiseen vaikuttavat prosessit. Hyödynsin muuttujia kolmanteen tutkimuskysymykseen vastaamisessa. Tarkastelin niiden avulla, mitkä ympäristöön tai ihmistoimintaan liittyvät tekijät selittävät havaintotiheyden poikkeamaa väestömäärään perustuvasta odotuksesta.

Taulukko 2. Tutkimuksessa käytettyjen paikkatietoaineistojen tiedot.

| MUUTTUJA                                       | AINEISTO<br>(aineistomuoto)                                      | AINEISTON<br>TUOTTAJA       | JULKAISU-<br>VUOSI | TARKKUUS |
|------------------------------------------------|------------------------------------------------------------------|-----------------------------|--------------------|----------|
| <b>HABITAATTIMUUTTUJAT</b>                     |                                                                  |                             |                    |          |
| Rakennettu alue                                | Corine Land Cover<br>2018<br>(rasteri)                           | Syke                        | 2018               | 20 m     |
| Pellot                                         | Corine Land Cover<br>2018<br>(rasteri)                           | Syke                        | 2018               | 20 m     |
| Metsien luonnon-<br>tilaisuus                  | Monimuotoi-<br>suudelle tärkeät<br>metsäalueet 2018<br>(rasteri) | Syke                        | 2018               | 96 m     |
| Kasvillisuus-indeksi<br>(NDVI)                 | NDVI:n<br>maksimiarvo<br>(rasteri)                               | Syke                        | 2021               | 10 m     |
| Rantaviiva                                     | Corine Land Cover<br>2018<br>(rasteri)                           | Syke                        | 2018               | 20 m     |
| <b>IHMISTEN LIIKKUMISTA KUVAAVAT MUUTTUJAT</b> |                                                                  |                             |                    |          |
| Väestömäärä                                    | Väestöruutu-<br>aineisto 1 km x 1<br>km<br>(vektori)             | Tilastokeskus               | 2023               | 1:1000   |
| Loma-asunnot                                   | Maastotietokanta<br>(vektori)                                    | Maanmittaus-<br>laitos      | 2025               | 1:10 000 |
| Tiet ja polut                                  | OpenStreetMap<br>(vektori)                                       | OpenStreetMap<br>Foundation | 2025               | 10 m     |
| Luonnon-<br>suojelualueet                      | Luonnonsuojelu-<br>ja erämaa-alueet<br>(vektori)                 | Metsähallitus               | 2021               | 1:20 000 |
| Liikuntapaikat                                 | Lipas-aineisto<br>(vektori)                                      | Jyväskylän<br>yliopisto     | 2025               | 1:20 000 |

### 3.3.3 Aineistojen esikäsittely ja standardointi

**Rakennetut alueet:** Laskin rakennettujen alueiden osuuden Corine maanpeiteaineistosta, joka kuvaa maakäyttöä ja maanpeitettä Suomessa (Corine maanpeite 2018). Erotin aineistosta rakennetut alueet (Luokka 1: Rakennetut alueet).

Ennen analyysia poistin aineistosta luokat 141 (taajamien viheralueet ja puistot) sekä 142 (urheilu- ja vapaa-ajan toiminta-alueet), sillä ne voivat olla puutiaisen kannalta potentiaalisia elinympäristöjä, eivätkä siten kuvaa puutiaisille epäsuotuisaa rakennettua ympäristöä yhtä yksiselitteisesti. Rakennettujen alueiden pinta-alan laskin leikkaamalla aineiston ruutukohtaisesti ja summaamalla rakennettujen alueiden pinta-alan. Tämän jälkeen suhteutin rakennetun alueen osuuden ruudun absoluuttiseen maapinta-alaan. Rakennetun maankäytön osuuden voidaan olettaa vaikuttavan puutiaishavaintojen tiheyteen eri tavoin. Rakennetun maankäytön osuuden kasvaessa myös puutiaishavaintojen määrä voi lisääntyä tiettyyn pisteeseen asti, sillä rakennetut ympäristöt keskittyvät usein alueille, joilla ihmisten määrä on suuri (Janzén ym. 2024). Hyvin voimakkaasti rakennetuissa ympäristöissä havaintotiheyden voidaan olettaa vähenevän, koska lopulta puutiaisille sopivien elinympäristöjen määrä vähenee.

**Peltomaat:** Peltujen osuuden laskin Corine maanpeiteaineiston maatalousalueista, josta erotin ensin peltomaat, ja sitten pellot (Luokka 3: Pellot). Luokka sisältää intensiivisesti viljeltyt pellot, joissa viljellään esimerkiksi vilja-, palko-, juurikas- ja rehukasveja. Luokittelun mukaan näillä alueilla maatalousmaan osuus on yli 75 %. Heterogeenisemmilla alueilla, joissa metsää tai rakennettua ympäristöä esiintyy enemmän (yli 25 %), luokitellaan aineistossa pienipiirteiseksi maatalousmosaiikiksi. Jätin kyseisen luokan analyysistä pois, sillä se voi tarjota puutiaisille suotuisia elinympäristöjä sekä sisältää enemmän ihmistoimintaa verrattuna yhtenäisiin peltoalueisiin. Lopulta laskin peltomaiden pinta-alan ruutukohtaisesti ja suhteutin sen ruudun maapinta-alaan. Peltujen osuuden voidaan olettaa olevan negatiivisesti yhteydessä puutiaishavaintojen tiheyteen, koska laajat peltoalueet tarjoavat puutiaisille niukasti suojaa ja suotuisia ilmasto-olosuhteita sekä houkuttelevat yleensä vähemmän ihmistoimintaa (Geldmann ym. 2016).

**Metsien luonnontilaisuus:** Metsien luonnontilaisuutta kuvaavan muuttujan johdin Zonation-menetelmällä tuotetusta rasteriaineistosta, joka kuvaa metsien monimuotoisuuden prioriteettiarvoja asteikolla 0–1 (Monimuotoisuudelle tärkeät metsäalueet 2018). Aineisto perustuu muun muassa lahoppopotetiaaliin, kasvillisuusluokkaan ja puuston rakenteeseen (Mikkonen ym. 2018). Mitä suurempia rasteriarvot ovat, sitä korkeampi on metsien monimuotoisuuden prioriteetti. Hyödynsin

lahopuupotentiaalia kuvaavaa aineistoa, sillä lahopuun määrä toimii hyvänä metsien luonnontilaisuuden indikaattorina (Kunttu ym. 2015). Yhdistin rasteriaineiston ruudukkoon laskemalla kunkin 1 km x 1 km ruudun pikseliarvojen keskiarvon. Metsien luonnontilaisuuden oletetaan olevan positiivisesti yhteydessä puutiaishavaintojen tiheyteen siten, että luonnontilaisuuden kasvaessa myös puutiaishavaintojen tiheyden oletetaan kasvavan.

**Kasvillisuusindeksi (NDVI):** Kasvillisuuden määrää kuvaavan muuttujan muodostin Suomen ympäristökeskuksen tuottamasta NDVI-aineistosta (Normalized Difference Vegetation Index). NDVI-arvot kuvaavat kasvillisuuden määrää, ja ne perustuvat vuoden 2021 kasvukauden maksimiarvoihin (NDVI:n maksimiarvo 2021). Alkuperäiset arvot vaihtelevat välillä -1 ja 1, jotka on aineistossa skaalattu välille 0–200. Korkeat arvot viittaavat tiheään ja elinvoimaiseen kasvillisuuteen, kun taas matalat arvot kuvaavat vähäistä kasvillisuutta tai esimerkiksi rakennettua ympäristöä. Ennen analyysiä rajasin aineistosta vesialueet pois, jotta ruutukohtaiset NDVI-arvot kuvaisivat ainoastaan maa-alueiden kasvillisuutta. Yhdistin aineiston analyysiruudukkoon laskemalla kunkin ruudun pikseliarvojen keskiarvon. Tämän jälkeen skaalasin arvot edelleen välille 0–1. Hyödynsin muuttujaa selittämään kasvillisuuden määrän vaikutusta puutiaishavaintojen tiheyteen. NDVI-arvon voidaan olettaa olevan positiivisesti yhteydessä puutiaishavaintojen tiheyteen siten, että kasvillisuuden määrän ja peittävyuden lisääntyessä myös puutiaisten elinolosuhteet muuttuvat suotuisammiksi (Ehrmann ym. 2017).

**Rantaviiva:** Muodostin vesistöjen vaikutusta kuvaavan muuttujan laskemalla maa- ja vesialueiden välisen reunavyöhykkeen määrän ruutukohtaisesti. Muuttuja perustuu Corine maanpeiteaineistoon, josta erotin vesialueita kuvaavan luokan. Kullekin 1 km x 1 km ruudulle laskin rantaviivan kokonaispituuden leikkaamalla vesialueita kuvaavan geometrian ruudukon kanssa ja laskemalle yhteen syntyvien viivojen pituudet ruuduittain. Tämän jälkeen suhteutin arvon ruudun maapinta-alaan, jolloin tuloksena on reunatiheys (m/ha). Käytin muuttujaa kuvaamaan vesistöjen ja maa-alueiden rajavyöhykkeiden määrää ja niiden mahdollista vaikutusta puutiaishavaintojen tiheyteen. Rantaviivan määrän oletetaan olevan positiivisesti yhteydessä

puutiaishavaintojen tiheyteen siten, että rantaviivan määrän kasvaessa myös puutiaishavaintojen määrän oletetaan kasvavan.

**Loma-asunnot:** Muodostin loma-asutusta kuvaavan muuttujan Maanmittauslaitoksen maastotietokannan rakennusaineistosta, joka sisältää rakennusten sijaintitiedot ja käyttötarkoitukseluokitukset (Maanmittauslaitoksen maastotietokohteet 2025). Erotin aineistosta loma-asunnot kohdeluokan (42231) perusteella, minkä jälkeen yhdistin Varsinais-Suomen kuntakohtaiset aineistot yhdeksi tasoksi. Laskin loma-asuntojen määrän ruutukohtaisesti kunkin 1 x 1 km ruudun sisälle sijoittuvien rakennusten perusteella. Tämän jälkeen muunsin muuttujan tiheydeksi jakamalla loma-asuntojen lukumäärän ruudun maapinta-alalla. Lopputuloksena on loma-asuntojen määrä hehtaaria kohden (kpl/ha). Hyödynsin muuttujaa kuvaamaan vapaa-ajan asutuksen vaikutusta puutiaishavaintojen tiheyteen pysyvän asutuksen ohella.

Loma-asuntoja kuvaavan muuttujan jakauma oli voimakkaasti oikealle vino ja sisälsi runsaasti nolla-arvoja, minkä vuoksi tein muuttujalle logaritmisen muunnoksen ( $\log_{10}$ ). Muunnos vähentää suurten arvojen vaikutusta malliin ja parantaa mallin vakautta, kun havaintojen jakauma on epätasainen. Siten muunnos mahdollistaa muuttujan vaikutuksen tulkinnan suhteellisena muutoksena, jolloin suurempien arvojen vaikutus ei korostu liikaa verrattuna pienempiin arvoihin. Loma-asuntojen määrän oletetaan olevan positiivisesti yhteydessä puutiaishavaintojen tiheyteen, koska loma-asutus lisää ihmisten oleskelua ja liikkumista luonnonläheisissä ympäristöissä.

**Polut:** Hain polkuverkostoaineiston OpenStreetMap-tietokannasta QGIS QuickOSM-lisäosan avulla. Aineisto sisältää tie- ja katuverkon, josta erotin polut (path). Lisäksi yhdistin OSM-aineistoa täydentääkseni Lipas-aineistosta kaikki reitit. Yhdistin nämä yhdeksi polkuverkostoa kuvaavaksi viivamuotoiseksi tasoksi. Laskin polkujen kokonaispituuden ruutukohtaisesti ruutujen sisälle sijoittuvien polkujen pituuksista. Tämän jälkeen suhteutin pituuden ruudun maapinta-alaan, jolloin tuloksena on polkuverkoston tiheys (m/ha). Polkumuuttuja kuvaa pääasiassa sellaisia liikkumisympäristöjä, joissa ihmisten ja puutiaisten kohtaaminen on todennäköistä. Siksi sisällytin aineistoon vain polkuja ja ulkoilureittejä kuvaavat kohdeluokat. Polkuverkoston tiheyden oletetaan olevan positiivisesti yhteydessä puutiaishavaintojen

tiheyteen, koska polut ohjaavat ihmisten liikkumista alueille, joilla voi esiintyä myös puutiaisille suotuisia elinympäristöjä.

**Luonnonsuojelualueet:** Luonnonsuojelualueita kuvaava aineisto sisältää luonnonsuojelulain ja asetuksen perusteella valtion ja yksityisille maille perustetut luonnonsuojelualueet sekä erämaalain perusteella perustetut erämaa-alueet (Luonnonsuojelu- ja erämaa-alueet 2025). Aineisto koostuu kolmesta kohdeluokasta eli yksityisistä suojelualueista, valtion maiden suojelualueista ja erämaa-alueista. Rajasin aineiston ulkopuolelle yleisöltä suljetut suojelualueet sekä vesialueilla sijaitsevat suojelualueet, jotta muuttuja kuvaisi ensisijaisesti virkistykseen soveltuvia ympäristöjä. Laskin suojelualueiden pinta-alan ruutukohtaisesti, ja lopuksi suhteutin sen ruudun maapinta-alaan. Luonnonsuojelualueiden osuuden oletetaan olevan positiivisesti yhteydessä puutiaishavaintojen tiheyteen siten, että luonnonsuojelualueiden osuuden kasvaessa myös puutiaishavaintojen määrän oletetaan kasvavan. Suojelualueet sisältävät usein luonnontilaisia ympäristöjä sekä houkuttelevat virkistyskäyttöä, vaikka joukossa on myös puutiaisille heikommin soveltuvia elinympäristöjä, kuten soita.

**Liikuntapaikat:** Lipas-tietokanta sisältää tiedot Suomen julkisista liikuntapaikoista, ulkoilureiteistä ja virkistysalueista (Lipas-data 2025). Valitsin aineistosta ulkoliikuntaan ja virkistykseen soveltuvat kohteet, kuten urheilukentät, uimarannat ja lähiliikuntapaikat. Rajasin analyysin ulkopuolelle sisäliikuntapaikat ja muut sisätiloissa sijaitsevat kohteet, kuten kuntosalit, liikuntahallit, tanssi- ja ryhmäliikuntatilat sekä sisäämpumaradat ja -palloiluhallit. Laskin liikuntapaikkojen määrän ruutukohtaisesti, minkä jälkeen muunsin muuttujan liikuntapaikkojen tiheydeksi jakamalla kohteiden lukumäärän ruudun maapinta-alalla. Hyödynsin muuttujaa kuvaamaan nimenomaan ulkona tapahtuvaa liikkumista, jossa ihmisten ja puutiaisten kohtaaminen on mahdollista. Liikuntapaikkojen määrän oletetaan olevan positiivisesti yhteydessä puutiaishavaintojen tiheyteen siten, että liikuntapaikkojen määrän kasvaessa myös puutiaishavaintojen määrän oletetaan kasvavan.

### 3.4 Väestömäärään perustuva havaintomäärän mallintaminen

Toiseen tutkimuskysymykseen vastatakseni tarkastelin puutiaishavaintojen tiheyttä suhteessa väestömäärään regressioanalyysin avulla. Regressioanalyysin tavoitteena on

muodostaa malli, jossa selitettävän muuttujan vaihtelua selitetään yhden tai useamman selittävän muuttujan avulla (Madetoja 2018). Lähtökohtana on oletus, että kansalaistiedeaineistoon perustuvien havaintojen kertymistä ohjaa voimakkaasti ihmisten määrä ja liikkuminen, mutta tämä yhteys ei ole välttämättä spatiaalisesti yhtenäinen koko tutkimusalueella (Johnston ym. 2020). Siksi mallinsin puutiaishavaintojen tiheyden ja väestömäärän välistä suhdetta maantieteellisesti painotetun regressiomallin (Geographically Weighted Regression, GWR) avulla. Menetelmä huomioi alueellisen vaihtelun siten, että muuttujien väliset suhteet voivat vaihdella tutkimusalueen eri osissa (Madetoja 2018).

Saaristo- ja rannikkoalueiden huomioimiseksi poistin ruudut, joissa ei ollut lainkaan maapinta-alaa. Tämän rajauksen avulla säilytin mahdollisimman tarkan rannikon geometrian samalla, kun jätin pelkästään merialueella sijaitsevat ruudut tarkastelun ulkopuolelle. Tämän jälkeen poistin analyysistä ruudut, joissa sekä puutiaishavaintojen määrä että väestömäärä olivat nolla. Nämä ruudut eivät sisällä tietoa mallinnettavasta ilmiöstä, jolloin niiden poistaminen parantaa mallin kykyä kuvata havaintojen ja väestömäärän välistä yhteyttä.

Toteutin analyysin 1 x 1 km ruutuaineistolle Varsinais-Suomen alueella. Laskin jokaiselle ruudulle puutiaishavaintojen määrän sekä ruudun asukasluvun. Väestötiedot perustuvat Tilastokeskuksen 1 x 1 km väestöruutuaineistoon. Koska analyysiruudukko ei ollut täysin yhtenevä väestöruudukon kanssa, laskin väestömäärän kullekin ruudulle sen alueelle sijoittuvista väestöarvoista. Tarkastelussa hyödynsin sekä kaikkien vuosien (2021–2023) yhteenlaskettuja havaintoja että vuosikohtaisia havaintoja erikseen.

Koska puutiaishavaintoja ei voi syntyä vesialueilta, huomioin analyysissä ruudun todellisen maapinta-alan. Muodostin selitettäväksi muuttujaksi havaintotiheyden, jonka suhteutin ruudun maapinta-alaan:

$$Havaintotiheys = \frac{\text{havaintojen määrä}}{\text{maapinta-ala (ha)}} \quad (1)$$

Näin ollen aineisto sisältää jokaiselle ruudulle havaintotiheyden sekä väestömäärän, joita käytin mallinnuksessa selitettävänä ja selittävänä muuttujana. Tiheyden

hyödyntäminen mahdollistaa havaintojen määrän suhteuttamisen todelliseen maapinta-alaan, jolloin ruudut ovat keskenään vertailukelpoisia.

Puutiaishavaintojen tiheyden ja väestömäärän välistä yhteyttä mallinsin maantieteellisesti painotetun regressiomallin (GWR) avulla. Menetelmä mahdollistaa regressiokertoimien alueellisen vaihtelun tarkastelun, jolloin voidaan analysoida, miten väestömäärän selitysvoima puutiaishavainnoille vaihtelee tutkimusalueen eri osissa (Madetoja 2018). Mallin selitettävänä muuttujana käytin puutiaishavaintojen tiheyttä ja selittävänä muuttujana ruudun väestömäärää. Mallin kaava voidaan esittää muodossa:

$$y_i = \beta_0(u_i, v_i) + \sum_m \beta_m(u_i, v_i) x_{i,m} + \varepsilon_i \quad (2)$$

jossa  $y_i$  on selitettävän muuttujan arvo havainnolle  $i$ ,  $(u_i, v_i)$  ovat havainnon  $i$  koordinaatit ja  $x_{i,m}$  kuvaa  $m$ :nnen selittävän muuttujan arvo havainnolle  $i$  (Brunsdon ym. 1996; Madetoja 2018). Lisäksi  $\varepsilon$  on satunnainen häiriötermi, jonka odotetaan noudattavan normaalijakaumaa  $N(0, \sigma^2)$ .

Tässä tutkimuksessa  $y_i$  kuvaa puutiaishavaintojen tiheyttä ruudussa  $i$ . Kokonaisuudessaan malli tarkastelee, miten puutiaishavaintojen tiheys suhteutuu väestömäärään paikallisesti ja miten tämä suhde vaihtelee alueellisesti tutkimusalueen sisällä.

Sovitin GWR:n hyödyntäen ruutujen keskipisteiden koordinaatteja, jolloin mallin parametrit voidaan estimoida erikseen kullekin sijainnille. Määritin mallille optimaalisen kaistanleveyden (*engl.* bandwidth) ristiinvalidointiin perustuvan pisteytyksen (leave-one-out cross-validation, CV) avulla. Se mahdollistaa analyysiin sopivan spatiaalisen mittakaavan huomioimisen ja minimoi mallin ennustevirhettä (Comber ym. 2018). Optimaalinen kaistanleveys määritetään vertaamalla eri vaihtoehdoilla estimoitujen paikallisten mallien sopivuutta ja valitsemalla arvo, joka tuottaa pienimmän CV-pisteytyksen. GWR-mallissa voidaan estimoida paikallinen regressiomalli jokaisessa pisteessä niiden spatiaalisten etäisyyksien perusteella. Kaistanleveys taas määrittää, kuinka laajalta alueelta havaintoja otetaan mukaan kunkin paikallisen mallin estimointiin.

### 3.5 Poikkeama-alueiden tunnistaminen

Toiseen tutkimuskysymykseen liittyen tarkastelin GWR-mallin selitysvoimaa ja jäännöksiä, joiden avulla arvioin, missä havaintotiheys poikkeaa väestömäärään perustuvasta odotuksesta. GWR-mallista tulkitsin erityisesti mallin globaalia selitystasetta ( $R^2$ ), joka kuvaa selitysvoimaa koko tutkimusalueella. Globaalilla mallilla tarkoitetaan esimerkiksi regressiomallia, jossa muuttujien välinen suhde oletetaan samaksi koko tutkimusalueella (Madetoja 2018). Tässä tutkimuksessa se ilmaisee, kuinka suuri osa puutiaishavaintojen alueellisesta vaihtelusta on selitettävissä väestömäärällä. Lisäksi tarkastelin paikallisia selitystasetta (local  $R^2$ ), jotka kuvaavat, kuinka hyvin väestömäärä selittää puutiaishavaintojen määrää kussakin 1 km x 1 km ruudussa. Paikallisessa mallissa muuttujien välinen suhde voi vaihdella tutkimusalueella, kuten GWR:ssä (Madetoja 2018).

Mallin selitysvoimaa tarkastelin jäännösten (*engl.* residual) avulla. Jäännökset kuvaavat havaittujen ja mallin ennustamien havaintomäärien välistä erotusta (Madetoja 2018; Nahhas 2024). Jäännös ilmaisee, kuinka paljon puutiaishavaintojen tiheys poikkeaa väestömäärään perustuvan mallin ennusteesta tietyssä ruudussa. Positiiviset jäännökset kuvaavat ruutuja, joissa havaintoja on väestömäärään nähden odotettua enemmän, kun taas negatiiviset jäännökset kuvaavat ruutuja, joissa havaintoja on odotettua vähemmän. Jäännösten tarkastelu on keskeinen osa regressiomallin arviointia, sillä niiden avulla voidaan tunnistaa alueita, joilla havaintojen määrään vaikuttavat väestömäärän ohella merkittävästi myös muut tekijät (Madetoja 2018).

Jäännösten tulkinnan tueksi laskin studentoidut jäännökset, joiden avulla tunnistin poikkeamaruudut suhteessa koko aineiston jakaumaan. Studentoitu jäännös saadaan jakamalla jäännös sen arvioidulla keskihajonnalla (Nahhas 2024). Studentoidut jäännökset ovat keskenään vertailukelpoisia, sillä niiden keskiarvo on lähellä nollaa. Tämä helpottaa poikkeavien jäännösarvojen tunnistamisessa. Studentoitujen jäännösten avulla tunnistin poikkeamaruudut hyödyntäen kvantiililuokittelua. Määrittelin aineistosta studentoitujen jäännösten 5 ja 95 % kvantiilit. Niiden avulla luokittelin ruudut kolmeen ryhmään: odotettua vähemmän sisältävät ruudut (alin 5 %), odotettua enemmän havaintoja sisältävät ruudut (ylin 5 %) ja muut ruudut. Tämä

mahdollistaa sellaisten alueiden tunnistamisen, joissa havaintotiheys poikkeaa selvästi väestömäärään perustuvan mallin ennusteesta.

Tarkastelin jäännösten jakaumaa histogrammin avulla arvioidakseni niiden soveltuvuutta mallinnukseen. Jakauma poikkesi normaalijakaumasta ja oli selvästi oikealle vino. Tavanomaisessa mallinnustilanteessa tällainen voisi heikentää mallin luotettavuutta erityisesti virhe-estimaattien osalta. Tässä tutkimuksessa huomion kohteena olivat kuitenkin erityisesti jakauman ääripäät eli alueet, joilla havaintoja esiintyi selvästi odotettua enemmän tai vähemmän. Tämän vuoksi pystyin hyödyntämään aineistoa mallin sovituksessa sellaisenaan.

Toteutin analyysit R-ohjelmistossa hyödyntäen paikkatiedon käsittelyyn, mallinnukseen ja visualisointiin tarkoitettuja paketteja (taulukko 3). Spatiaalisen vektoriaineiston käsittelyssä käytin sf- ja dplyr-paketteja, kun taas rasteriaineistojen analysoinnissa hyödynsin etenkin terra- ja exactextractr-paketteja. GWR:n sovittamiseen käytin spgwr-pakettia, ja tilastollisessa mallinnuksessa sekä jatkoanalyysseissa (GAM) hyödynsin mgcv- ja spdep-paketteja. Tulosten visualisoinnissa hyödynsin ggplot2-pakettia sekä karttaesitysten tukena ggspatial-, ggmap- ja prettymapr-paketteja.

Taulukko 3. Tutkimuksessa käytetyt R-paketit ja niiden versiot.

| Nimi          | Versio | Lähde                                 |
|---------------|--------|---------------------------------------|
| dplyr         | 1.1.4  | Wickham ym. 2023                      |
| lubridate     | 1.9.4  | Grolemund & Wickham 2011              |
| sf            | 1.0–19 | Pebesma & Bivand 2023<br>Pebesma 2018 |
| terra         | 1.8–29 | Hijmans 2025                          |
| ggmap         | 4.0.0  | Kahle & Wickham 2013                  |
| ggplot2       | 3.5.1  | Wickham 2016                          |
| exactextractr | 0.10.0 | Baston 2023                           |
| ggspatial     | 1.1.9  | Dunnington 2023                       |
| prettymapr    | 0.2.5  | Dunnington 2024                       |
| spgwr         | 0.6–37 | Bivand & Yu 2024                      |
| mgcv          | 1.8–42 | Wood 2011                             |
| spdep         | 1.3–13 | Pebesma & Bivand 2023                 |

### 3.6 Poikkeama-alueiden paikkatietopohjainen selittäminen

Kolmanteen tutkimuskysymykseen vastatakseni tarkastelin, mitkä ihmistoimintaan tai puutiaisen habitaattiin liittyvät tekijät selittävät GWR-mallissa havaittuja poikkeamia väestömäärään perustuvasta odotuksesta. Tässä analyysivaiheessa hyödynsin GWR-mallin jäännöksiä vasteena ja selitin niiden vaihtelua paikkatietoaineistoista johdetuilla muuttujilla. Poikkeama-alueiden jatkoanalyysissä hyödynsin koko jäännösjakaumaa jatkuvana muuttujana, jotta säilytin enemmän informaatiota ilmiön vaihtelusta ja pystyin tarkastelemaan havaintojen ali- ja yliedustusta koko tutkimusalueella.

Koska aineisto on spatiaalista, testasin GWR:n jäännösten spatiaalista riippumattomuutta Moranin I-testillä ennen GAM-mallin sovittamista. Spatiaalinen autokorrelaatio kuvaa lähekkäisten havaintojen välistä riippuvuutta eli sitä, missä määrin samankaltaiset arvot esiintyvät toistensa läheisyydessä (Getis 2010: 256). Moranin I-testin avulla voidaan tarkastella tätä ilmiötä tilastollisesti. Positiivinen arvo viittaa siihen, että samankaltaiset arvot sijaitsevat lähellä toisiaan, kun taas arvo lähellä nollaa viittaa satunnaiseen spatiaaliseen rakenteeseen.

Tarkastelin spatiaalista autokorrelaatiota ennen GWR-mallia, jolloin se oli 0.197 ( $p < 0.001$ ). Tässä tutkimuksessa käytetty GWR poisti kuitenkin spatiaalisen autokorrelaation käytännössä kokonaan (Moran's I = 0.026,  $p < 0,001$ ). Tämä viittaa siihen, että jäännökset ovat spatiaalisen rakenteen osalta pääosin satunnaisia, vaikka tulos onkin suuren havaintomäärän vuoksi edelleen tilastollisesti merkitsevä.

Sisällytin malliin ympäristöön tai ihmistoimintaan liittyviä muuttujia, jotka johdin paikkatietoaineistoista ruututasolle (taulukko 4). Muuttujien valinnassa tavoitteena oli kattaa keskeiset puutiaisten esiintymiseen ja havaintojen kertymiseen vaikuttavat tekijät. Muuttujien tarkastelu erillisinä selittäjinä mahdollisti niiden vaikutusten tulkinnan.

Taulukko 4. GAM-mallinnuksessa käytetyt habitaatti- tai ihmistoimintaa kuvaavat muuttujat.

| Muuttujaryhmä                     | Muuttuja                    | Muuttujan tyyppi | Muuttujan laskutapa                      | Muuttujan vaihteluväli (1 x 1 km) | Muuttujan yksikkö |
|-----------------------------------|-----------------------------|------------------|------------------------------------------|-----------------------------------|-------------------|
| Habitaatti-muuttujat              | Pellot                      | polygoni         | peltojen pinta-ala / ruudun maapinta-ala | 0–1                               | m/ha              |
|                                   | Rakennetut alueet           | polygoni         | rakennetun pinta-ala / maapinta-ala      | 0–1                               | m/ha              |
|                                   | Metsien luonnontilaisuus    | rasteri          | keskiarvo / ruutu                        | 0–1                               | %                 |
|                                   | Kasvillisuus-indeksi (NDVI) | rasteri          | keskiarvo / ruutu                        | 0–1                               | %                 |
|                                   | Rantaviiva                  | viiva            | pituus / maapinta-ala                    | 0–922.7                           | m/ha              |
| Ihmistoimintaa kuvaavat muuttujat | Loma-asunnot                | piste            | lukumäärä / maapinta-ala                 | 0–2.6                             | kpl / ha          |
|                                   | Polut                       | viiva            | pituus / maapinta-ala                    | 0–616.4                           | m / ha            |
|                                   | Luonnonsuojelu-alueet       | polygoni         | pinta-ala / maapinta-ala                 | 0–1                               | %                 |
|                                   | Liikuntapaikat              | piste            | lukumäärä / maapinta-ala                 | 0–0.43                            | kpl / ha          |

Analysoin aineistoa yleistetyllä additiivisella mallilla (Generalized Additive Model, GAM). GAM on yleistetyn lineaarisen mallin laajennus, jossa selittävien muuttujien välistä yhteyttä voidaan kuvata epälineaaristen silotefunktioiden avulla (Wood 2017: 119; Zuur ym. 2009). Malli arvioi kullekin selittävälle muuttujalle erillisen vaikutuskuvion, joka kuvaa muuttujan vaikutusta vasteeseen ilman ennalta määriteltä muotoa. Näin voidaan välttää tarve oletuksille lineaarisuudesta tai tietystä polynomimuodosta, mikä tekee mallista joustavan ja aineistolähtöisen (Zuur ym. 2009). Käytin GAM-mallia kolmanteen tutkimuskysymykseen vastatakseni, jossa tarkastelin, miten paikkatietoaineistoista johdetut muuttujat selittävät GWR-mallin jäännösten vaihtelua.

GAM-mallin etuna on sen kyky kuvata monimutkaisia ja epälineaarisia riippuvuuksia vasteen ja selittävien muuttujien välillä ilman, että mallista tulee rakenteeltaan monimutkainen (Wood ym. 2017: 119). Toteutin analyysit käyttäen mgcv-pakettia, joka on yksi keskeisimmistä yleistettyjen additiivisten mallien estimointiin kehitetyistä työkaluista (Wood 2017: 217). Paketin gam-funktio vastaa rakenteeltaan pitkälti yleistettyjen lineaaristen mallien (glm) sovittamista, mutta mahdollistaa silotefunktioiden avulla epälineaaristen yhteyksien mallintamisen. Muodostin mallin niin sanottuna täytenä mallina, johon sisällytin kaikki valitut selittävät muuttujat samanaikaisesti, vaikka kaikki niistä eivät osoittautuneet tilastollisesti merkitseviksi. Ratkaisu perustuu tavoitteeseen tarkastella ilmiötä kokonaisvaltaisesti yksittäisen optimaalisen mallin rakentamisen sijaan.

Mallinsin osaa muuttujista lineaarisina parametrisinä termeinä ja osaa epälineaarisina silotefunktioina. Jaottelu perustui mallin alustavaan tarkasteluun, jossa arvioin kunkin muuttujan vaikutuksen muotoa efektiivisten vapausasteiden (edf) avulla. Muuttujat, joiden edf-arvo oli alle 1,5, tulkitsin lineaarisiksi, jolloin niiden mallintaminen silotefunktiolla ei olisi tuonut lisää informaatiota. Tällaiset muuttujat sovitin lineaarisina termeinä, jotta malli olisi mahdollisimman parsimoninen. Sen sijaan muuttujia, joiden edf-arvo oli suurempi kuin 1,5, mallinsin epälineaarisina silotefunktioina, jotta niiden vaikutusten muotoja voitaisiin kuvata joustavasti.

Mallin alustavassa tarkastelussa hyödynsin silotefunktioiden kuvaajia, joiden avulla arvioin muuttujien vaikutusten muotoja ja havaintojen jakautumista eri arvoilla. Tarkastelun pohjalta havaitsin, että ruudun maapinta-alan osuutta kuvaavan muuttujan vaikutus korostui voimakkaasti erityisesti pienillä arvoilla. Näillä alueilla jäännökset olivat voimakkaasti positiivisia, mikä viittaa siihen, että malli aliarvioi havaintotiheyttä ruuduissa, joilla maapinta-alaa on hyvin vähän.

Tarkastelun perusteella havaitsin, että vaikutus tasoittuu, kun ruudun maapinta-ala ylittää noin 10 % ruudun kokonaisalasta. Tämän vuoksi rajasin lopullisen analyysin koskemaan vain ruutuja, joissa maapinta-alaa on vähintään 10 %. Rajauksen tavoitteena oli varmistaa, että analyysi kohdistuu lopulta alueisiin, joissa sekä

puutiaishavainnot että ihmistoiminta ovat mahdollisia ja joissa mallin tuottamat tulokset ovat luotettavampia.

GAM-mallissa muuttujien vaikutusten muotoa kuvaavat efektiiviset vapausasteet (edf). Arvo  $\text{edf} = 1$  viittaa lineaariseen yhteyteen, kun taas suuremmat arvot kuvaavat epälineaarista vaikutusta. Mitä suurempi edf-arvo on, sitä epälineaarisempi yhteys on muuttujan ja vasteen välillä (Zuur ym. 2009). Lisäksi mallin kuvioissa on 95 %:n luottamusväli. Laajempi luottamusväli viittaa suurempaan epävarmuuteen, kun taas kapeampi luottamusväli kertoo varmemmasta estimaatista.

Efektiivisiä vapausasteita hyödynsin myös mallin joustavuuden ja mahdollisen ylisovittamisen riskin arviointiin. Korkeat edf-arvot voivat viitata siihen, että malli mukautuu liikaa aineiston vaihteluun erityisesti tilanteissa, joissa havaintoja on vähän tietyillä arvoilla. Tämän vuoksi tarkastelin mallin silotefunktioita ja rajoitin osan silotefunktioiden joustavuutta asettamalla niille arvon  $k = 4$ . Silotefunktioiden joustavuuden rajaaminen kohdistui seuraaviin muuttujiin: NDVI, rantaviiva, loma-asunnot ja liikuntapaikat. Näiden muuttujien alustavat tarkastelut viittasivat voimakkaasti epälineaarisiiin ja hyvin monimutkaisiin yhteyksiin. Näin estin mallia sovittamasta aineistoon tarpeettoman monimutkaisia vaikutusmuotoja. Mallin soveltuvuutta arvioin tarkastelemalla mallin sovitetta sekä silotefunktioiden muotoa ja niihin liittyviä luottamusvälejä.

Lisäksi tarkastelin selittävien muuttujien välistä epälineaarista riippuvuutta (*engl. concavity*). Se kuvaa, missä määrin yhden silotefunktion selittämää vaihtelua voidaan ennustaa muiden mallissa olevien silotefunktioiden avulla (Kovács 2022). Kohdistin tarkastelun ainoastaan epälineaarisiiin silotefunktioihin, eikä lineaarisina parametrisina termeinä mallinnettuja muuttujia sisällytetty arviointiin. Korkeat arvot voivat viitata siihen, että muuttujat kuvaavat osittain samaa ilmiötä, mikä vaikeuttaa yksittäisten funktioiden vaikutusten tulkintaa. Analyysissä suurin osa concavity-arvoista jäi mataliksi, mutta erityisesti rantaviivan silotefunktiossa havaittiin melko korkeita arvoja. Tämä viittaa siihen, että osa muuttujista sisältää osittain samankaltaista spatiaalista informaatiota toisten muuttujien kanssa. Luvut eivät kuitenkaan ylittäneet arvoa 0,8, joten ne eivät muodostaneet merkittävää ongelmaa mallin tulkinnan kannalta.

Arvioin mallin toimivuutta tarkastelemalla selityssastetta, selittävien muuttujien merkitsevyyttä, vaikutusten muotoa sekä jäännöksiä. Parametristen muuttujien osalta tulkitsin positiivisen vaikutuksen siten, että muuttujan arvon kasvaessa ruudun puutiaishavaintojen tiheys on suurempi kuin väestömäärän perusteella olisi odotettavissa. Negatiivinen vaikutus puolestaan viittaa siihen, että havaintoja esiintyy odotettua vähemmän suhteessa väestömäärään. Epälineaaristen silotefunktioiden osalta vaikutus ei ole yksiselitteisesti positiivinen tai negatiivinen, vaan arvioin vaikutusta koko muuttujan arvojen vaihteluvälillä sen muodon perusteella.

Selittävien muuttujien vaikutusten tulkintaa varten muodostin GAM-mallin avulla osittaisvaikutuskuvioita (partial effect plots). Kuvioissa vaihtelin tarkasteltavan muuttujan arvoa sen koko havaintovälillä, kun taas muiden mallissa mukana olevien muuttujien arvot pidin vakioituina suhteessa niiden keskiarvoihin. Näin voitiin tutkia yksittäisen muuttujan vaikutusta vasteeseen ilman muiden muuttujien vaikutusta. Estimoin vaikutukset mallin termikohtaisina ennusteina, jolloin vaikutukset voitiin esittää mallin funktion asteikolla nollan molemmiin puoliin. Lisäksi sisällytin kuvioihin 95 %:n luottamusvälit, jotka kuvaavat estimaattien epävarmuutta eri arvojen alueilla.

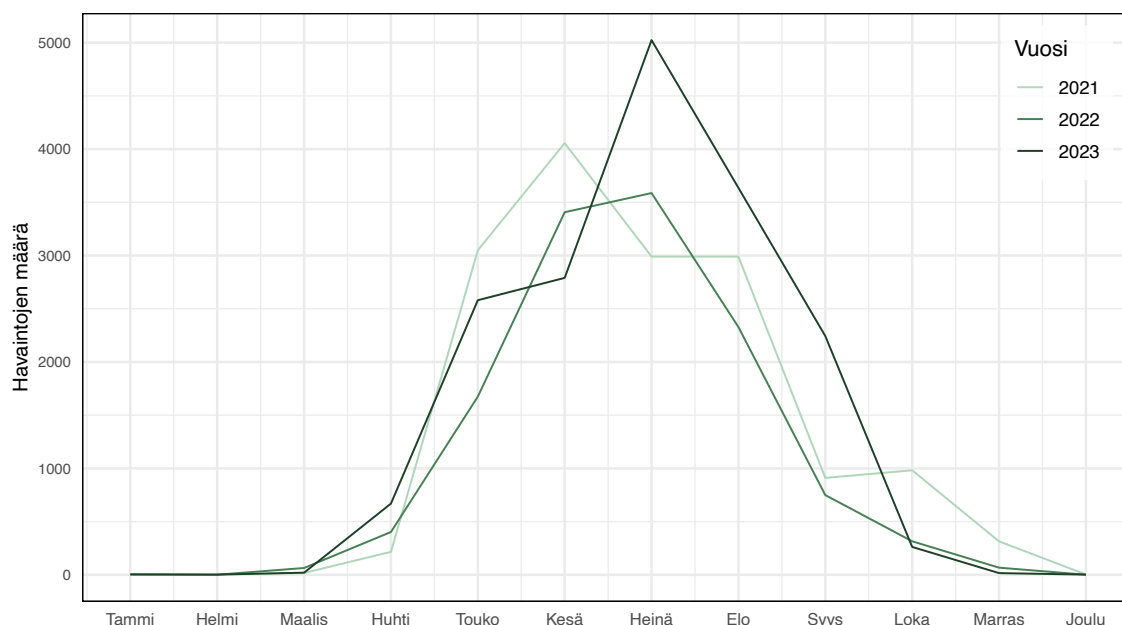
Sovitin GAM-mallin koko aineistoon, ja muodostin karttamuotoisen esityksen jäännösten alueellisesta vaihtelusta ja poikkeama-alueista. GAM-mallin spatiaalinen ennuste kuvaa, missä ympäristö- ja ihmistoimintamuuttujien yhteisvaikutus ennustaa puutiaishavaintojen ylittävän tai alittavan väestömäärään perustuvan odotuksen. Lopuksi muodostin jäännösten ennustearvot GAM-mallin predict-funktiolla jokaiselle ruudulle erikseen.

## 4 Tulokset

### 4.1 Puutiaisten määrä ja jakautuminen Varsinais-Suomessa

Varsinais-Suomessa on vuosina 2021–2023 raportoitu Punkkilive-kansalaistiedesovelluksen kautta 45 379 havaintoa. Näistä 15 532 havaintoa on tehty vuonna 2021 ja 12 600 vuonna 2022. Vuonna 2023 havaintoja kertyi tarkasteluvuosista eniten, yhteensä 17 247.

Vuosittaiset puutiaishavainnot painottuvat jokaisena tarkasteluvuotena loppukeväästä alkusyksyyn (kuva 3). Havainnot lisääntyvät toukokuussa ja saavuttavat vuosittaisen huippunsa kesä–heinäkuussa. Vuonna 2021 puutiaishavaintoja on kirjattu muita vuosia pidempään syksyyn asti. Sen sijaan vuonna 2023 puutiaishavaintojen huippu painottui muita vuosia voimakkaammin heinäkuuhun, jolloin havaintoja tehtiin eniten kuukausitasolla.

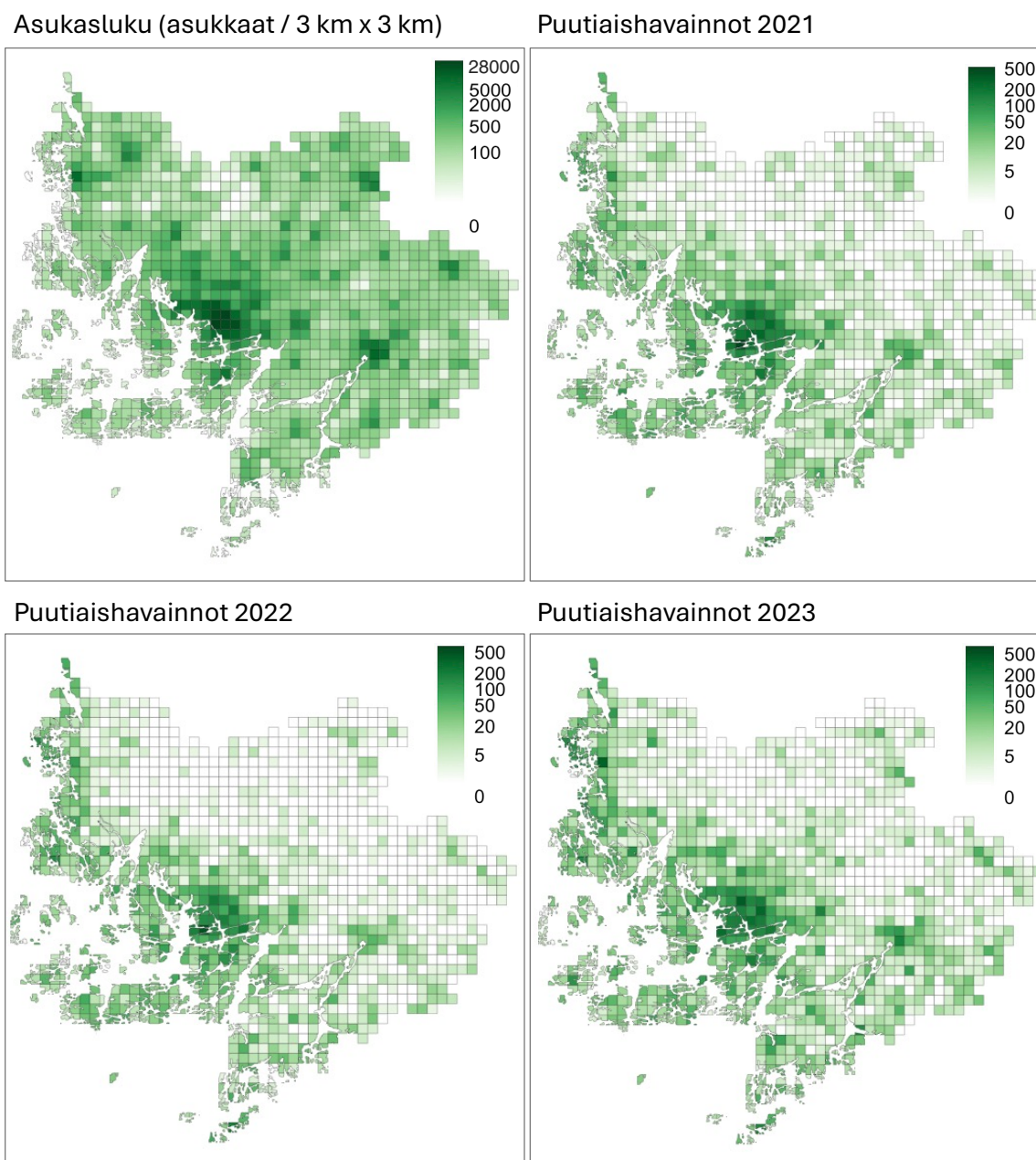


Kuva 3. Puutiaishavaintojen määrä Varsinais-Suomessa Punkkilive-aineistossa tarkasteluvuosittain.

Puutiaishavainnot eivät jakaudu Varsinais-Suomessa maantieteellisesti tasaisesti, vaan ne painottuvat rannikolle ja kaupunkiseutujen läheisyyteen. Sisämaassa havaintoja on kirjattu huomattavasti vähemmän kuin rannikkoalueilla, ja useissa harvaan asutuissa 3 km x 3 km ruuduilla havaintoja ei ole kirjattu lainkaan. Nämä ruudut sijoittuvat pääosin

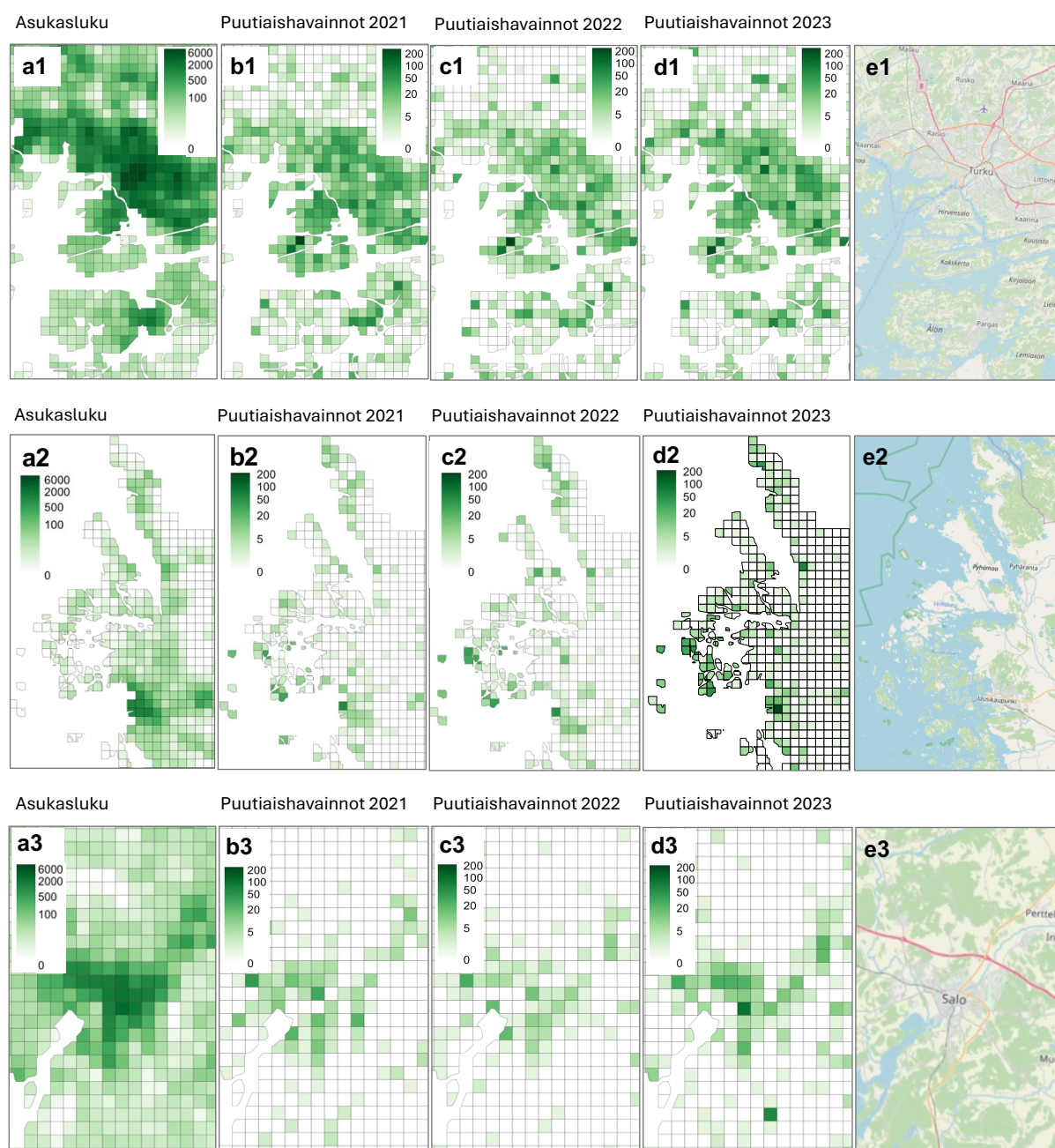
harvaan asutuille alueille, mikä viittaa ihmistoiminnan vaikutukseen havaintojen kertymisessä.

Ruutukohtainen tarkastelu osoittaa, että puutiaishavainnot ja väestö jakautuvat samankaltaisesti. Puutiaishavaintoja tehdään runsaasti erityisesti alueilla, joille väestö on keskittynyt, kuten kaupunkiseuduilla ja rannikon läheisyydessä. Saaristoalueet kuitenkin erottuvat alueina, joilla puutiaishavaintoja kertyy runsaasti myös vähäisestä asukasmäärästä huolimatta.



Kuva 4. Väestömäärä (A) ja puutiaishavaintojen alueellinen jakautuminen Varsinais-Suomessa vuosina 2021–2023 (B–D) 3 km x 3 km ruuduissa.

Puutiaishavaintojen ja väestön alueellinen jakautuminen ovat pääosin samankaltaisia, mutta tarkasteluvuosien välillä esiintyy myös eroja havaintomäärissä ja alueellisessa kattavuudessa. Havaintojen painopiste säilyy koko tarkastelujakson ajan pitkälti samoilla rannikko- ja kaupunkialueilla, vaikka havaintomäärät vaihtelevat vuosittain. Tarkastelu Turun, Uudenkaupungin ja Salon seuduilla osoittaa, että puutiaishavaintojen ja väestön välinen suhde vaihtelee erityisesti rannikko-, saaristo- ja sisämaan alueiden välillä (kuva 5).



Kuva 5. Asukkaiden määrä (a1-a3) ja puutiaishavaintojen määrä (b1-d3) kolmella alueella 1 km x 1 km ruutukoossa Varsinais-Suomessa vuosina 2021–2023. Ensimmäisellä rivillä (a1–d1) on Turun seutu,

toisella rivillä (a2–d2) Varsinais-Suomen luoteisosa eli Uudenkaupungin seutu. Taustakartta (e1–e3) © OpenStreetMap 2026.

Turun kaupunkiseudulla puutiaishavaintojen alueellinen jakautuminen noudattaa pääosin väestötiheyden rakennetta (kuva 5 a1–d1). Naantalin, Raision, Turun ja Kaarinan tiheimmin asutut alueet muodostavat yhtenäisen vyöhykkeen, jossa puutiaishavaintoja tehdään runsaasti. Erityisesti Turun eteläpuolella Hirvensalon ja Kaksikerran saaret erottuvat alueina, joilla puutiaishavaintoja kertyy runsaasti suhteessa väestötiheyteen. Suurempien kaupunkialueiden ulkopuolella puutiaishavaintoja tehdään vähemmän erityisesti sisämaassa ja ulkosaaristoon päin siirryttäessä.

Uudenkaupungin seudulla puutiaishavainnot muodostavat rannikkoa seuraavan vyöhykkeen, joka erottuu selvästi ympäröivistä sisämaan alueista (kuva 5 a2–d2). Havaintoja tehdään runsaasti erityisesti Uudenkaupungin edustalla saaristossa, vaikka alueen väestötiheys on verrattain alhainen. Sisämaassa puutiaishavaintoja esiintyy vähän, ja useissa ruuduissa havaintoja ei kirjata lainkaan. Siten puutiaishavaintojen alueellinen jakautuminen poikkeaa siten monin paikoin asukastiheyden rakenteesta.

Salon seudulla puutiaishavaintojen alueellinen jakautuminen seuraa pääosin väestön sijoittumista, ja havaintoja tehdään eniten Salon kaupunkikeskustan ympäristössä (kuva 5 a3–d3). Vuonna 2022 puutiaishavaintoja tehdään alueella vähemmän kuin muina tarkasteluvuosina, kun taas vuonna 2023 havaintoja kertyy eniten. Vaikka havaintojen määrä vähenee Salosta sisämaata kohti siirryttäessä, myös siellä on kirjattu puutiaishavaintoja erityisesti väestöltään tiheimmissä ruuduissa.

#### **4.2 Puutiaishavaintojen määrän ennustettavuus asukasluvun perusteella**

GWR-mallin perusteella asukasluku selittää kaikkien tarkasteluvuosien (2021–2023) yhteenlasketussa aineistossa 10,6 % puutiaishavaintojen alueellisesta vaihtelusta ( $R^2 = 0,106$ ,  $p < 0,001$ ). Asukasluvun ja havaintotiheyden välinen yhteys on positiivinen, eli puutiaishavaintoja tehdään keskimäärin enemmän alueilla, joilla asuu enemmän ihmisiä. Selitysaste jää kuitenkin melko matalaksi, mikä osoittaa, että väestömäärä yksin ei riitä selittämään puutiaishavaintojen alueellista jakautumista.

Vuosittaisissa tarkasteluissa globaalin mallin selitysaste jää matalaksi, sillä asukasluku selittää puutiaishavaintojen vaihtelusta vuosittain vain noin 0,05–0,2 % (taulukko 5).

Selitysaste on korkein vuonna 2021 ja matalin vuonna 2023. Tämä osoittaa, että väestömäärän ja puutiaishavaintojen välinen yhteys on koko tutkimusalueen tasolla heikko.

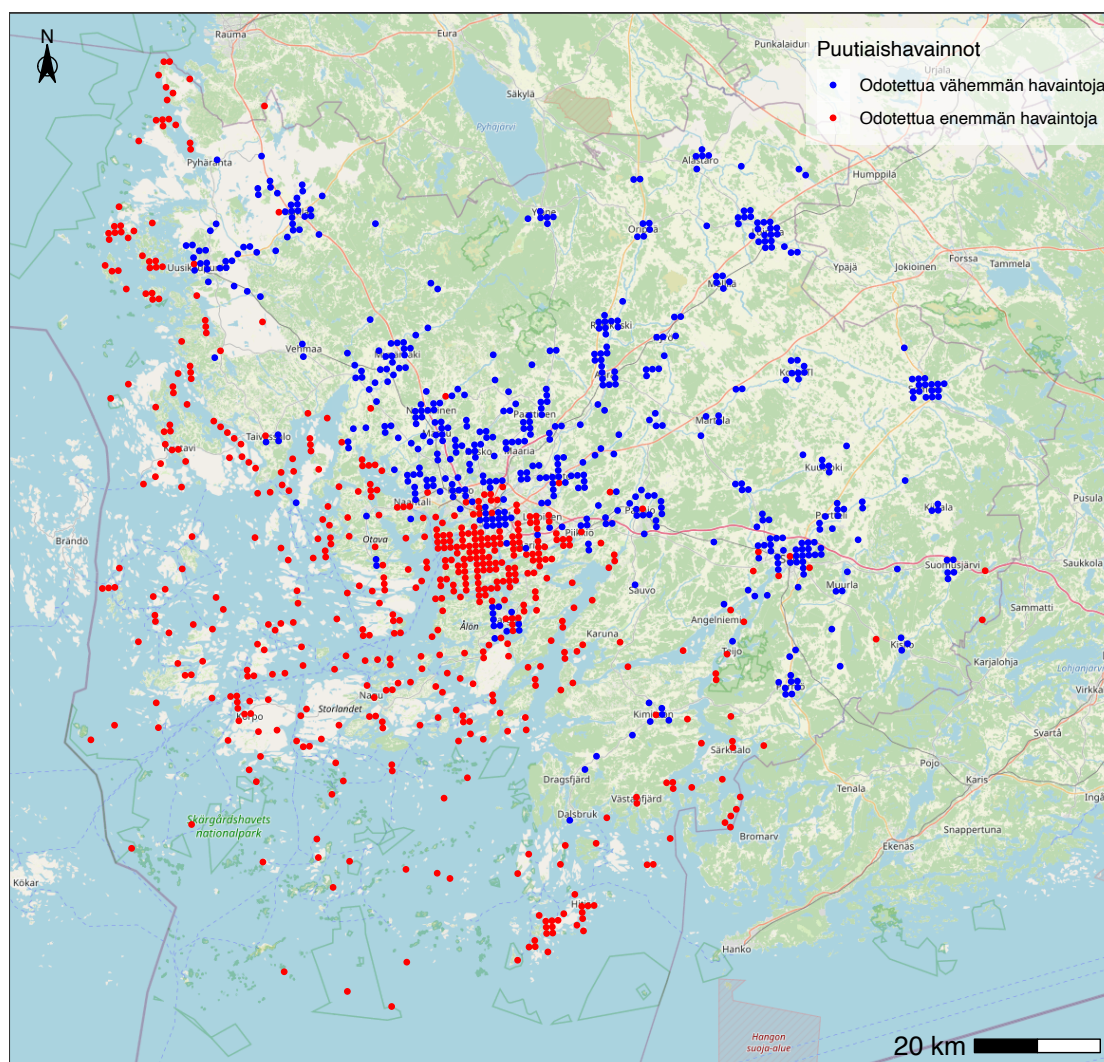
Paikallinen selitysaste (local  $R^2$ ) vaihtelee kuitenkin huomattavasti tutkimusalueella. Tarkasteluvuosina paikalliset  $R^2$ -arvot vaihtelevat välillä –0,12–0,56, kun taas mediaani jää jokaisena vuotena melko alhaiseksi (0,02–0,12). Tämä tarkoittaa, että asukasluku selittää useimmilla alueilla vain pienen osan puutiaishavaintojen vaihtelusta.

Taulukko 5. GWR-mallin selitysasteet vuosina 2021–2023. Paikallinen  $R^2$  kuvaa selitysvoiman spatiaalista vaihtelua 1 x 1 km ruututasolla.

| Vuosi | Globaalin mallin $R^2$ | Paikallinen $R^2$ min | Paikallinen $R^2$ mediaani | Paikallinen $R^2$ max |
|-------|------------------------|-----------------------|----------------------------|-----------------------|
| 2021  | 0.0022                 | -0.0520               | 0.1226                     | 0.5566                |
| 2022  | 0.0011                 | -0.1220               | 0.0238                     | 0.2100                |
| 2023  | 0.00048                | -0.00074              | 0.0228                     | 0.2512                |

Mallin poikkeama-alueet eivät jakaudu tasaisesti Varsinais-Suomessa, vaan muodostavat alueellisia keskittymiä. Tämä osoittaa, että väestömäärään perustuva malli selittää puutiaishavaintoja joillakin alueilla selvästi paremmin kuin toisilla. Osalla alueista havaintomäärät vastaavat melko hyvin väestömäärän perusteella odotettua tasoa, kun taas toisilla alueilla havaintoja esiintyy systemaattisesti enemmän tai vähemmän kuin malli ennustaa.

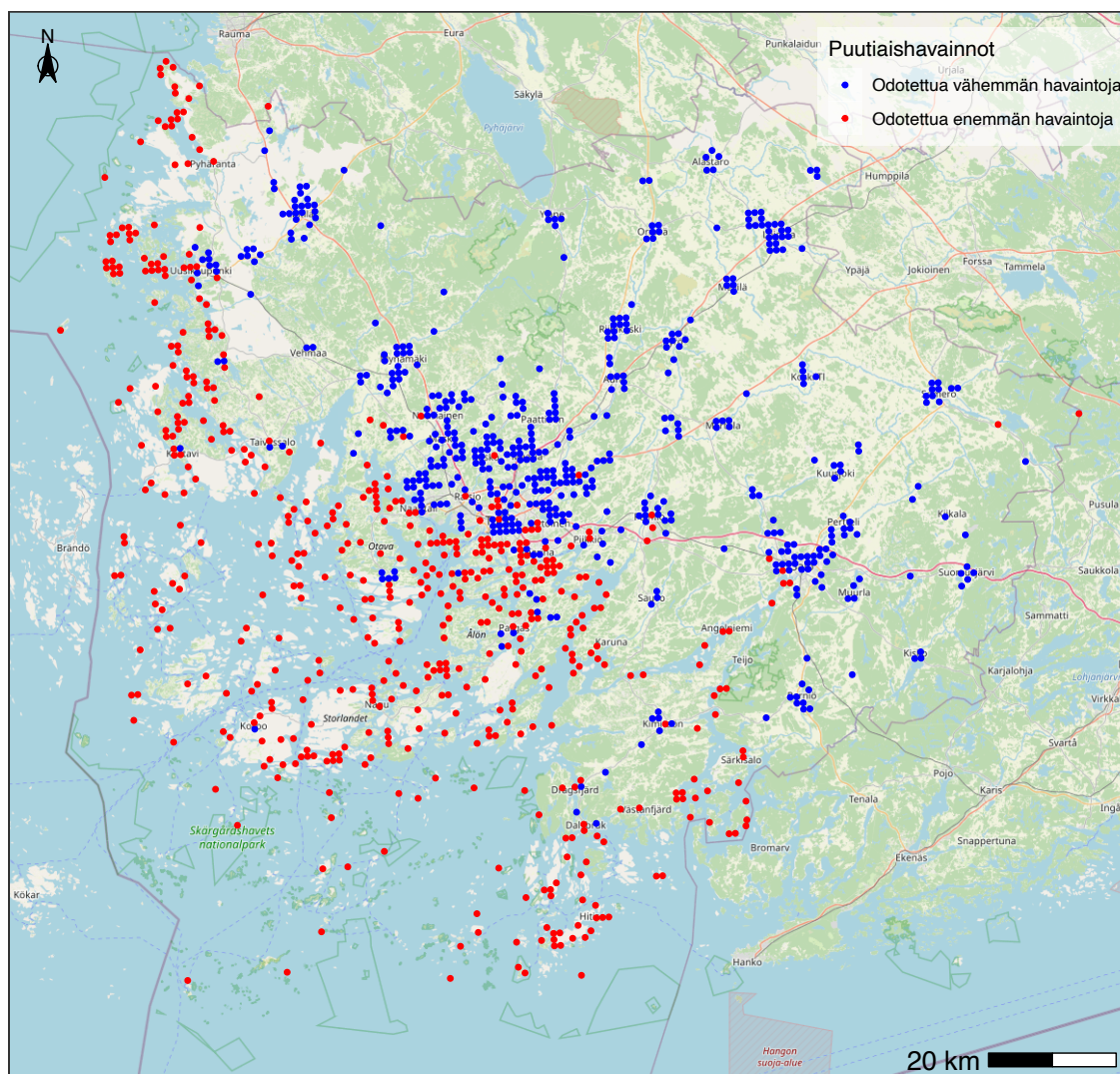
Odotettua enemmän puutiaishavaintoja sisältävät alueet painottuvat erityisesti rannikko- ja saaristoalueille (kuva 6). Näillä alueilla havaintoja tehdään enemmän kuin pelkän asukasluvun perusteella olisi odotettavissa. Odotettua vähemmän havaintoja sisältävät alueet sijoittuvat puolestaan laajasti sisämaan kuntakeskuksiin sekä Turun pohjoispuolelle. Rannikkoa kohti siirryttäessä poikkeamat muuttuvat pääosin positiivisiksi. Suurimpien kaupunkien alueilla positiiviset ja negatiiviset poikkeamat muodostavat myös tiiviitä vyöhykkeitä, joissa odotettua enemmän ja vähemmän havaintoja sisältävät alueet sijaitsevat välittömästi toistensa läheisyydessä.



Kuva 6. Vuoden 2021 poikkeama-alueet.

Kuvassa on sinisellä ruudut, joiden studentoidut jäännökset sijoittuvat alimpaan (5 %) ja punaisella ruudut, jotka sijoittuvat ylimpään (95 %) kvantiiliin. Siten kartta kuvaa alueita, joilla puutiaishavaintoja esiintyy väestömäärään nähden selvästi odotettua vähemmän tai enemmän.

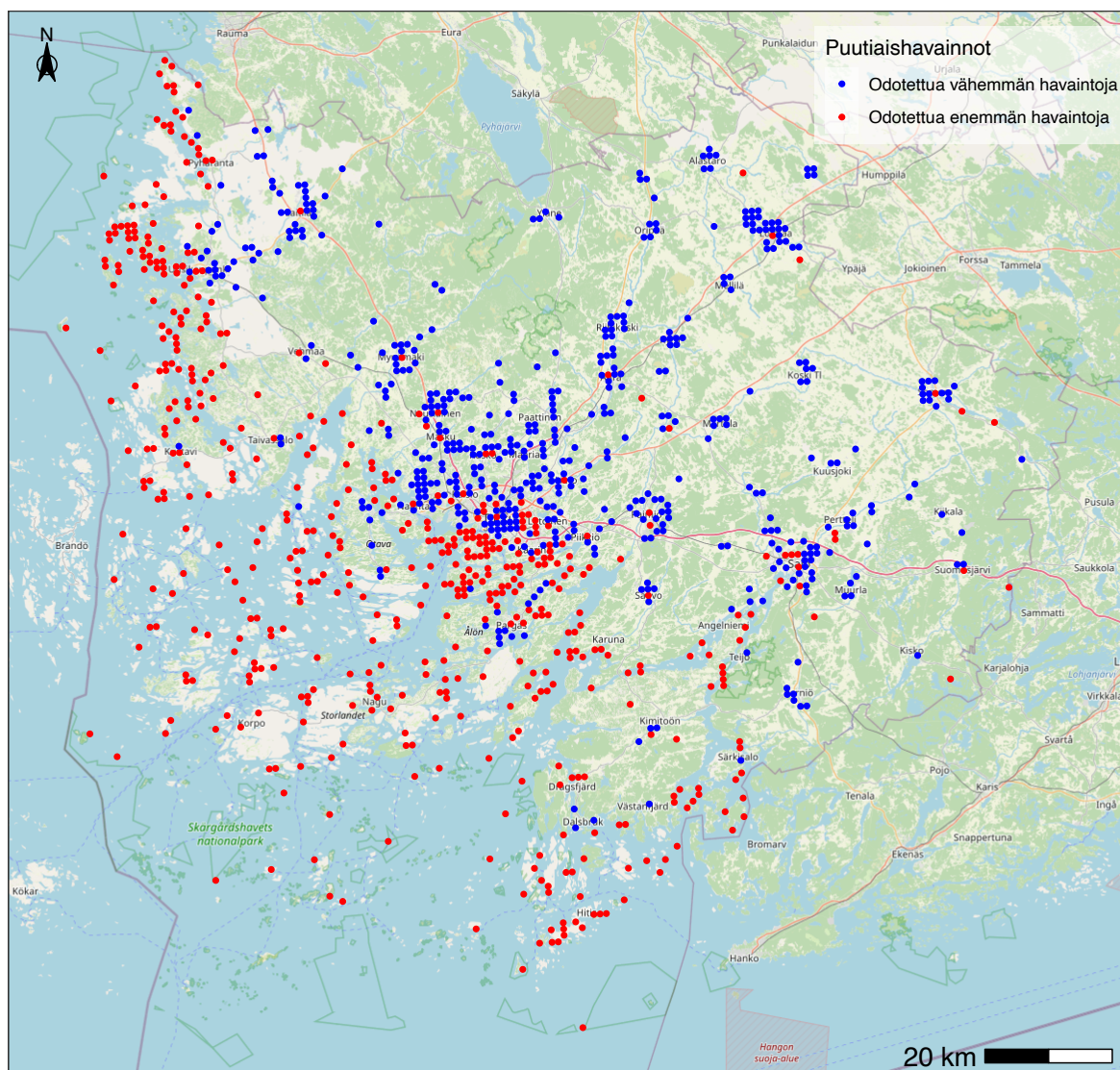
Vuonna 2022 puutiaishavaintojen poikkeama-alueiden alueellinen rakenne säilyi pääosin samankaltaisena kuin edellisenä vuonna. Rannikko- ja saaristoalueet erottuvat edelleen alueina, joilla puutiaishavaintoja tehdään odotettua enemmän suhteessa asukaslukuun, kun taas tiiviimmissä asutuskeskittymissä havaintoja kertyy odotettua vähemmän (kuva 7). Turun ja Uudenkaupungin edustan saaristoalueilla positiivisia poikkeamia esiintyy kuitenkin enemmän kuin vuonna 2021.



Kuva 7. Vuoden 2022 poikkeama-alueet.

Kuvassa on sinisellä ruudut, joiden studentoidut jäännökset sijoittuvat alimpaan (5 %) ja punaisella ruudut, jotka sijoittuvat ylimpään (95 %) kvantiiliin.

Vuonna 2023 poikkeama-alueiden alueellinen rakenne säilyy pääosin samankaltaisena kuin aiempina vuosina (kuva 8). Rannikko- ja saaristoalueilla puutiaishavaintoja tehdään edelleen odotettua enemmän suhteessa asukaslukuun, kun taas sisämaan tiheimmin asutuilla alueilla havaintoja kertyy odotettua vähemmän. Turun saaristossa positiiviset poikkeamat ovat kuitenkin hieman vähäisempiä kuin vuonna 2022. Kuten aiempinakin vuosina, rannikko- ja saaristoalueet muodostavat melko yhtenäisen vyöhykkeen, jolle positiiviset poikkeamat keskittyvät. Naantalın, Turun ja Kaarinan muodostamalla kaupunkialueella esiintyy puolestaan sekä positiivisia että negatiivisia poikkeamia, jotka sijoittuvat paikoin hyvin lähelle toisiaan.

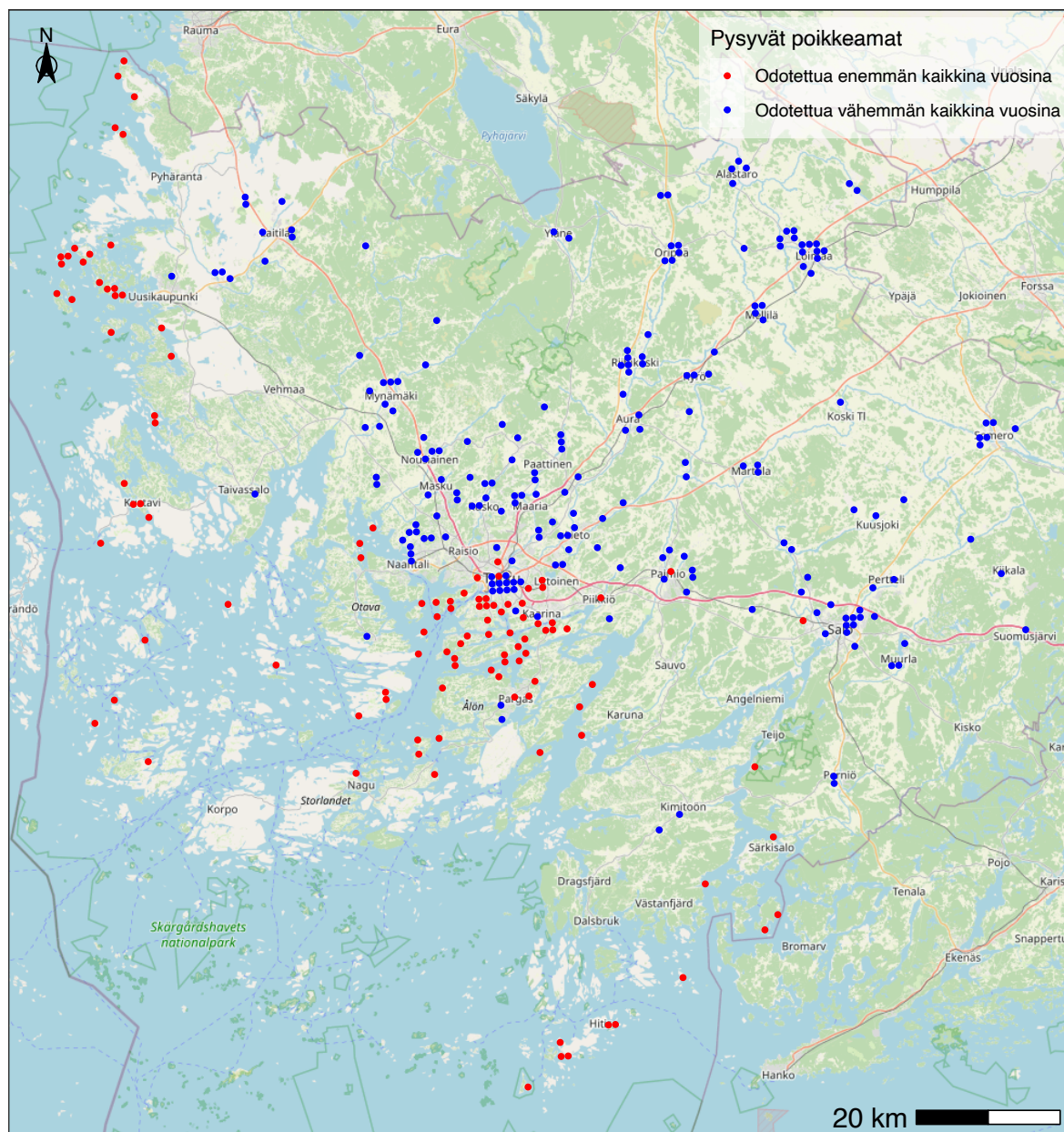


Kuva 8. Vuoden 2023 poikkeama-alueet.

Kuvassa on sinisellä ruudut, joiden studentoidut jäännökset sijoittuvat alimpaan (5 %) ja punaisella ruudut, jotka sijoittuvat ylimpään (95 %) kvantiliin.

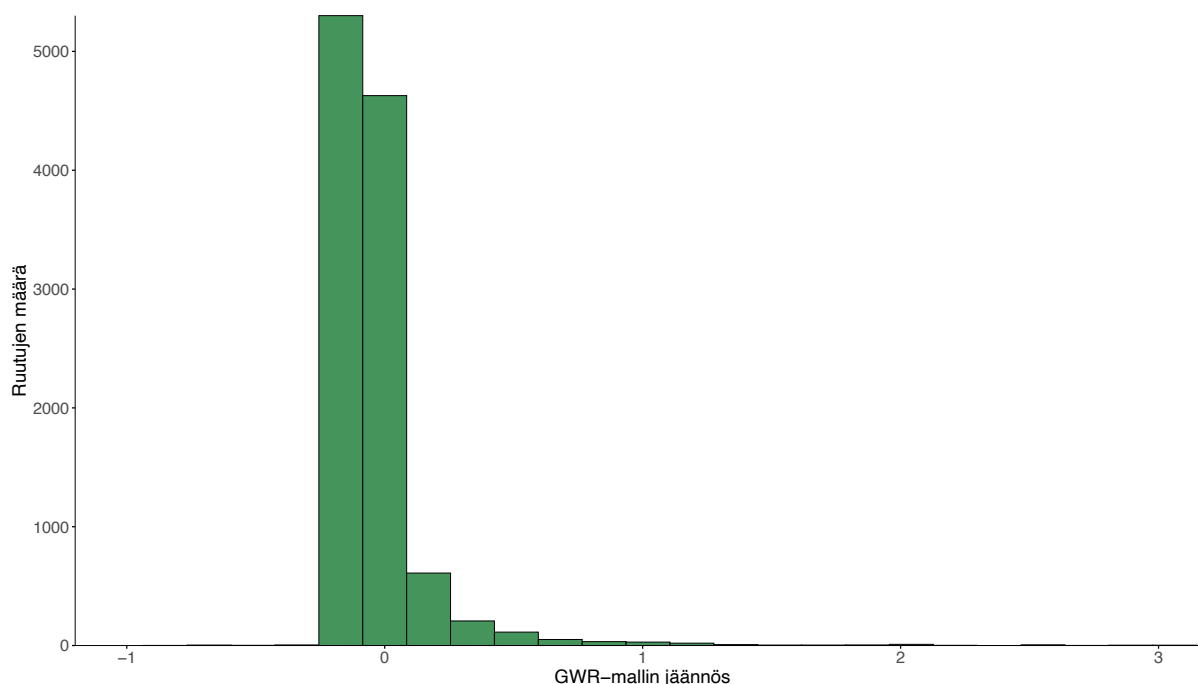
Pysyvät poikkeama-alueet jakautuvat rannikon ja sisämaan välillä (kuva 9). Odotettua enemmän puutiaishavaintoja sisältävät pysyvät poikkeama-alueet painottuvat erityisesti rannikko- ja saaristoalueille, etenkin Uudenkaupungin edustalle ja Turun sisäsaaristoon. Positiivisten poikkeama-alueiden sijainti vaihtelee kuitenkin saaristoalueilla huomattavasti vuosien välillä, eikä täysin yhtenäistä pysyvää rakennetta muodostu. Odotettua vähemmän havaintoja sisältävät pysyvät poikkeamat sijoittuvat puolestaan pääosin sisämaan taajamiin ja asutuskeskuksiin, erityisesti Loimaan seudulle sekä Turun kaupunkiseudun pohjoispuolelle. Lisäksi Turun kaupunkikeskustaan muodostuu yhtenäinen negatiivisten poikkeamien alue.

Kokonaisuudessaan negatiiviset poikkeama-alueet säilyivät vuosien välillä sijainniltaan yhtenäisempinä kuin positiiviset poikkeama-alueet.



Kuva 9. Pysyvät äärevimmät poikkeama-alueet vuosina 2021–2023. Punaiset alueet kuvaavat ruutuja, joissa puutiaishavaintoja esiintyi kaikkina vuosina odotettua enemmän suhteessa väestömäärään ja siniset alueet ruutuja, joissa havaintoja esiintyi odotettua vähemmän.

Jäännösten jakauma painottuu lähelle nollaa, mutta hieman negatiivisten arvojen puolelle (kuva 10). Tämä osoittaa, että väestömäärään perustuva malli ennustaa useimmille ruuduille hieman enemmän havaintoja kuin niissä todellisuudessa esiintyy. Jakauma on kuitenkin oikealle vino, sillä aineistossa on myös yksittäisiä suuria positiivisia poikkeamia. Tämä taas viittaa siihen, että tietyillä alueilla puutiaishavaintoja kertyy enemmän kuin väestömäärään perusteella olisi odotettavissa.



Kuva 10. GWR-mallin jäännösten jakauma.

Mallin toimivuus vaihtelee myös alueellisesti, sillä se yli- ja aliarvioi puutiaishavaintoja erityisesti rannikko-, saaristo- ja taajama-alueilla. Tulokset viittaavat siihen, että puutiaishavaintojen alueelliseen jakautumiseen vaikuttavat väestömäärän lisäksi myös elinympäristöön ja ihmistoimintaan liittyvät tekijät.

#### 4.3 Puutiaishavaintojen poikkeamia selittävät ympäristö- ja saavutettavuustekijät

Yleistetty additiivinen malli (GAM) selittää puutiaishavaintojen tiheyden jäännösten vaihtelusta 16,3 % (selitysaste = 0,161). Tämä osoittaa, että tarkastellut habitaatti- ja ihmisten liikkumiseen liittyvät muuttujat selittävät osan GWR-mallin yli- ja aliarvioinneista. Kuitenkin valtaosa vaihtelusta jää edelleen selittämättä.

Useat mallin selittävistä muuttujista ovat tilastollisesti merkitseviä ( $p < 0,05$ ). Rantaviivan pituus, loma-asuntojen määrä, rakennetun ympäristön osuus, NDVI, liikuntapaikkojen määrä sekä maapinta-alan osuus selittävät tilastollisesti merkitsevästi puutiaishavaintojen väestöön suhteutettuja poikkeamia (taulukko 6). Lisäksi metsien luonnontilaisuus ja polkujen tiheys ovat tilastollisesti merkitseviä lineaarisina vaikutuksina. Sen sijaan peltojen ja luonnonsuojelualueiden osuudet eivät ole tilastollisesti merkitseviä.

Taulukko 6. GAM-mallissa käytettyjen selittävien muuttujien vaikutukset puutiaishavaintojen poikkeamiin suhteessa väestömäärään.

| Muuttuja                    | Tyyppi      | Estimaatti $\beta$ / edf | p-arvo      |
|-----------------------------|-------------|--------------------------|-------------|
| Pellot                      | lineaarinen | $\beta = 0.015$          | 0.147       |
| Metsien luonnontilaisuus    | lineaarinen | $\beta = 0.074$          | < 0.001 *** |
| Polut                       | lineaarinen | $\beta = 0.00022$        | 0.006 **    |
| NDVI                        | silote      | edf = 2.30               | 0.00035 *** |
| Rakennetun ympäristön osuus | silote      | edf = 2.43               | < 0.001 *** |
| Rantaviiva                  | silote      | edf = 2.95               | < 0.001 *** |
| Loma-asunnot                | silote      | edf = 2.75               | < 0.001 *** |
| Luonnonsuojelualueet        | silote      | edf = 2.85               | 0.52        |
| Liikuntapaikat              | silote      | edf = 1.91               | < 0.001 *** |
| Maan osuus                  | silote      | edf = 7.45               | < 0.001 *** |

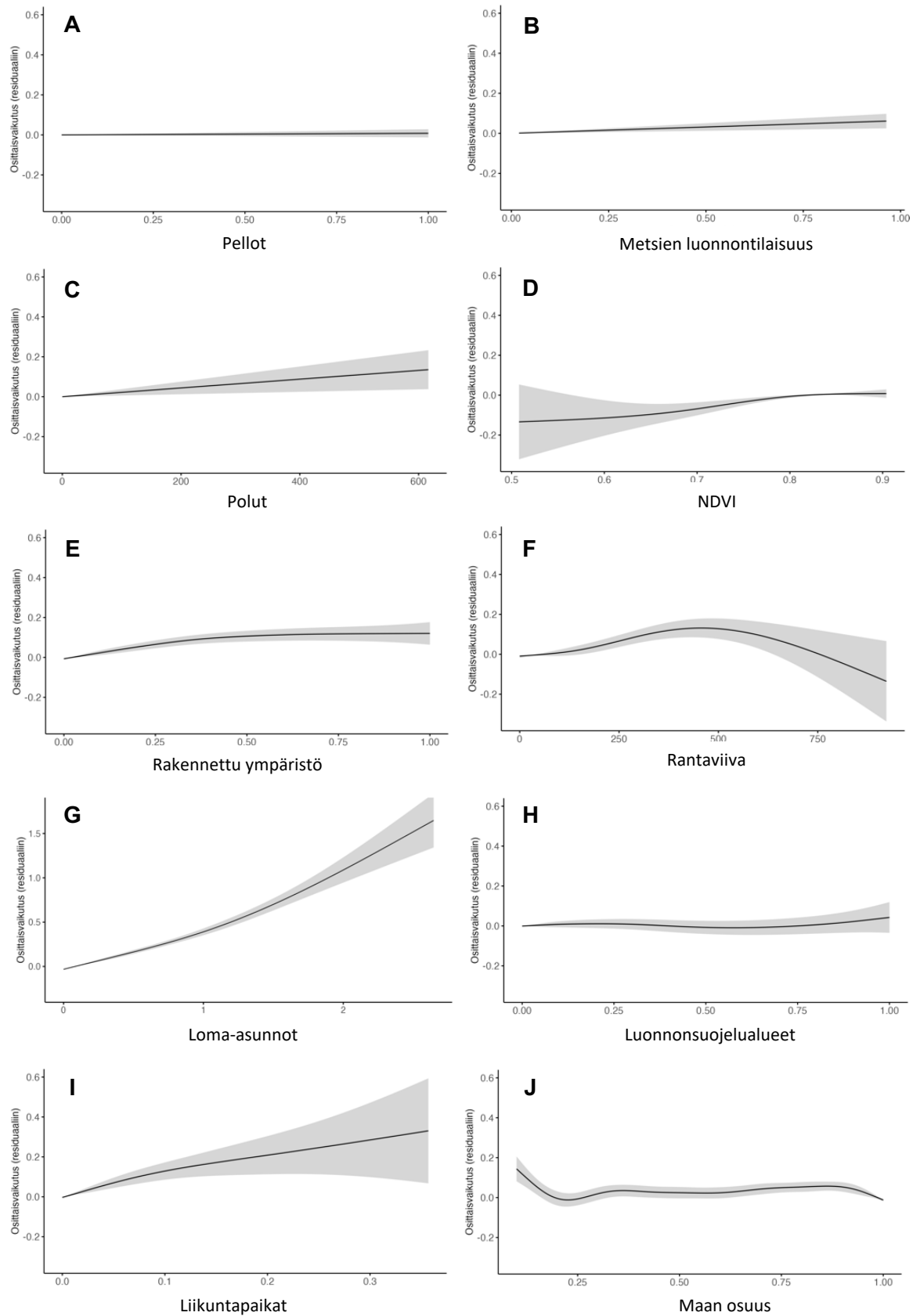
Lineaarisista muuttujista metsien luonnontilaisuus ja polkujen osuus ovat tilastollisesti merkitseviä, kun taas peltojen osuus ei ollut merkitsevä. Metsien luonnontilaisuuden positiivinen parametriestimaatti viittaa siihen, että luonnontilaisemmilta alueilta kirjataan enemmän puutiaishavaintoja kuin väestömäärän perusteella olisi odotettavissa. Myös polkujen määrän vaikutus on positiivinen, mutta vaikutus on vähäinen.

Silotteina mallinnettujen muuttujien efektiiviset vapausasteet ( $\text{edf} > 1,5$ ) osoittavat, että osa yhteyksistä on epälineaarisia. Selvimät epälineaariset yhteydet havaitaan rantaviivan pituuden, loma-asuntojen määrän ja maapinta-alan osuuden kohdalla. Myös rakennetun ympäristön osuus, NDVI ja liikuntapaikkojen määrä ovat tilastollisesti merkitseviä ja lievästi epälineaarisia.

Silotteina mallinnettujen muuttujien efektiiviset vapausasteet ( $\text{edf} > 1,5$ ) viittaavat epälineaariin yhteyksiin. Erityisesti rantaviivan pituus, loma-asuntojen määrä ja maapinta-alan osuus osoittavat selvästi epälineaarisen vaikutuksen jäännöksiin. Myös rakennetun ympäristön osuus, NDVI ja liikuntapaikkojen määrä ovat tilastollisesti merkitseviä ja lievästi epälineaarisia.

GAM-mallin osittaisvaikutuskuviot osoittavat, että muuttujien vaikutukset puutiaishavaintojen väestöön suhteutettuihin poikkeamiin vaihtelevat sekä

voimakkuudeltaan että muodoltaan (kuva 11). Ihmistoimintaa ja saavutettavuutta kuvaavista muuttujista voimakkain yhteys havaitaan loma-asuntojen määrällä, jonka vaikutus on selvästi positiivinen ja epälineaarinen (kuva 11 G). Vaikutus voimistuu erityisesti suurilla arvoilla, mikä viittaa siihen, että runsaasti loma-asuntoja sisältävissä ruuduissa puutiaishavaintoja tehdään huomattavasti enemmän kuin väestömäärän perusteella olisi odotettavissa.



Kuva 11. GAM-mallin osittaisvaikutuskuviot lineaaristen (pellot, metsien luonnontilaisuus ja polut) ja epälineaaristen muuttujien vaikutuksista puutiaishavaintojen tiheyden jäännöksiin.

Myös liikuntapaikkojen määrällä on positiivinen ja lievästi epälineaarinen vaikutus jäännöksiin (kuva 11 I). Vaikutus voimistuu liikuntapaikkojen määrän kasvaessa, mikä viittaa siihen, että liikuntapaikkojen suuri määrä on yhteydessä odotettua suurempiin puutiaishavaintomääriin. Suurilla arvoilla luottamusvälit kuitenkin levenevät, mikä lisää mallin ennusteeseen liittyvää epävarmuutta alueilla, joilla liikuntapaikkoja on runsaasti.

Polkujen tiheyden vaikutus jäännöksiin on lineaarinen ja positiivinen, mutta melko vähäinen. Tämä viittaa siihen, että kulkureittien lisääntyminen kasvattaa havaintomääriä vain maltillisesti (kuva 11 C). Rakennetun ympäristön osuudella havaitaan puolestaan lievästi epälineaarinen positiivinen vaikutus (kuva 11 E), jossa vaikutus voimistuu erityisesti, kun rakennettua ympäristöä on yli 25 % ruudun maapinta-alasta.

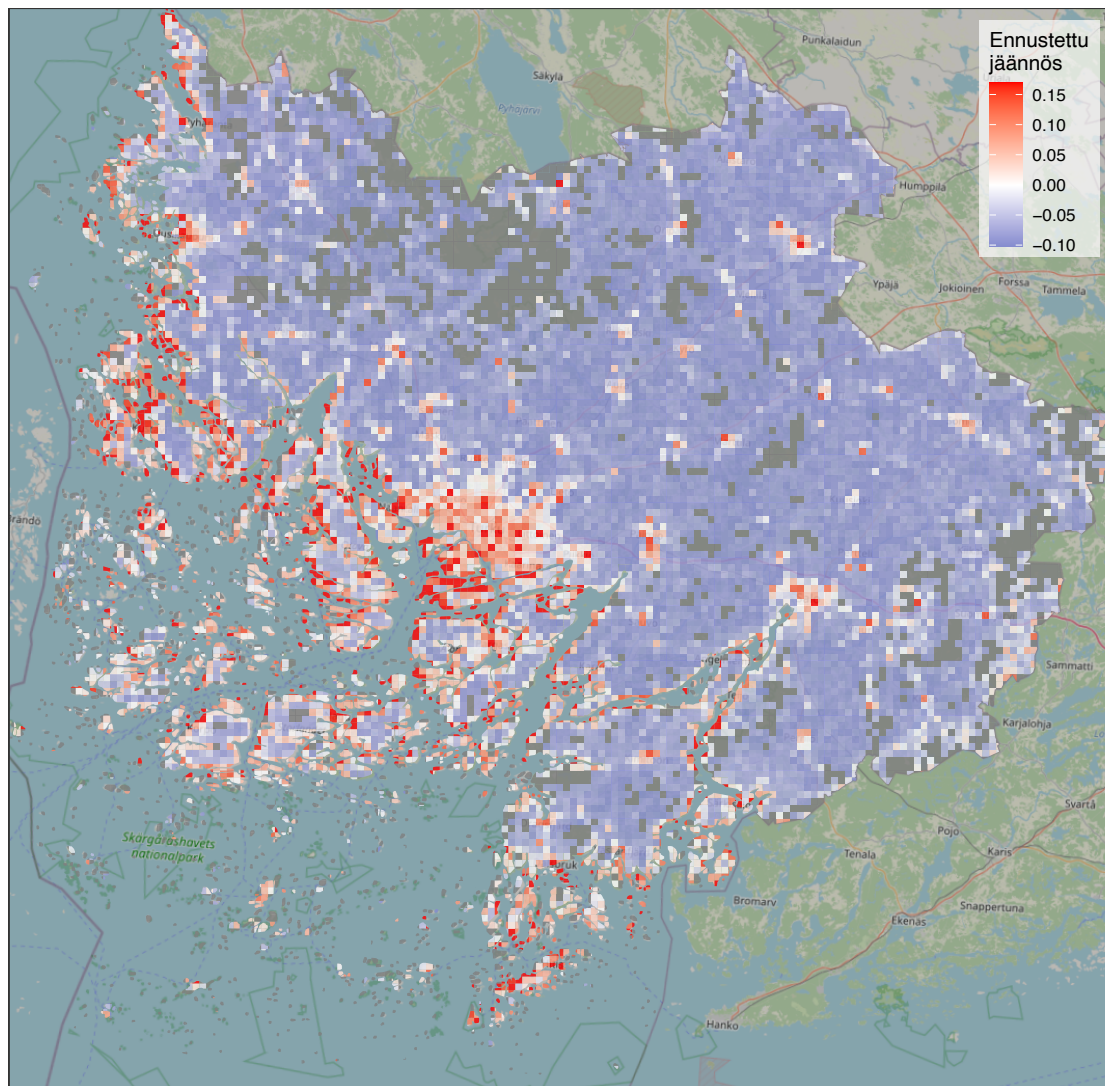
Habitaattimuuttujista rantaviivan määrällä on selvin epälineaarinen yhteys jäännöksiin (kuva 11 F). Vaikutus kasvaa keskisuurille arvoille asti (400 m/ha), mutta heikkenee suurimmilla arvoilla. Tämä viittaa siihen, että rantaympäristöt lisäävät havaintojen määrää erityisesti keskisuurilla arvoilla, mutta ruudut, joissa rantaviivaa on paljon, eivät enää lisää havaintoja samassa suhteessa.

NDVI:n vaikutus on tilastollisesti merkitsevä ja lievästi epälineaarinen (kuva 11 D). Vaikutus on negatiivinen erityisesti matalilla NDVI-arvoilla, mutta lähestyy nollaa kasvillisuuden määrän kasvaessa. Tämä viittaa siihen, että vähäisen kasvillisuuden alueilla puutiaishavaintoja tehdään vähemmän kuin väestömäärän perusteella olisi odotettavissa. Kasvillisuuden lisääntyessä vaikutus kuitenkin heikkenee ja lähestyy nollaa. Metsien luonnontilaisuuden vaikutus on lineaarinen ja positiivinen, mutta melko pieni (kuva 11 B). Tämä osoittaa, että luonnontilaisemmilta alueilta kirjataan hieman enemmän havaintoja kuin väestömäärän perusteella olisi odotettavissa.

Maapinta-alan osuuden vaikutus jää kokonaisuudessaan vähäiseksi ja melko tasaiseksi (kuva 11 J). Ainoastaan ruuduissa, joissa maapinta-alaa on alle 20 %, havaintoja esiintyy hieman enemmän kuin väestömäärän perusteella voitaisiin odottaa. Peltojen ja luonnonsuojelualueiden osuuksilla ei sen sijaan havaita selkeää yhteyttä jäännöksiin, sillä niiden osittaisvaikutukset pysyvät koko tarkasteluvälillä lähellä nollaa (kuva 11 A ja H).

GAM-mallin ennustamat poikkeamat osoittavat selkeän alueellisen eron rannikon ja sisämaan välillä (kuva 12). Positiiviset ennustetut jäännökset painottuvat erityisesti rannikko- ja saaristoalueille, mikä tarkoittaa, että habitaatti- ja ihmistoimintamuuttujien yhteisvaikutus ennustaa näille alueille väestömäärään nähden odotettua enemmän puutiaishavaintoja. Negatiiviset ennustetut jäännökset sijoittuvat puolestaan pääosin sisämaan alueille, joissa havaintoja ennustetaan väestömäärään nähden odotettua vähemmän puutiaishavaintoja. Ennustetuissa jäännöksissä näkyy siten samankaltainen alueellinen rakenne kuin väestömäärään perustuvissa poikkeamissa.

Laajimmat yhtenäiset positiivisten ennusteiden alueet sijoittuvat erityisesti sisäsaaristoon, jossa positiiviset jäännökset muodostavat rantaviivaa seuraavan vyöhykkeen. Sisämaassa ennusteet ovat puolestaan pääosin negatiivisia. Ennusteen rakenne ei kuitenkaan ole täysin yhtenäinen, sillä myös sisämaan alueilla esiintyy paikallisia positiivisia jäännöksiä ja rannikkoalueilla yksittäisiä negatiivisia ruutuja. Kokonaisuudessaan mallissa korostuu rannikon ja saariston positiivinen sekä sisämaan negatiivinen ennusterakenne.



Kuva 12. GAM-mallin ennustamat jäännökset puutiaishavaintojen tiheyksien poikkeamille suhteessa väestömäärään.

Punaiset alueet kuvaavat ruutuja, joissa havaintoja ennustetaan esiintyvän väestömäärään nähden odotettua enemmän ja siniset alueet ruutuja, joissa havaintoja ennustetaan esiintyvän odotettua vähemmän. Harmaat ruudut ovat alueita, joissa ei joko ollut asukkaita eikä puutiaishavaintoja tai maapinta-ala oli alle 10 %. (Taustakartta © OpenStreetMap 2026)

## 5 Keskustelu

### 5.1 Havaintojen spatiaalinen jakauma ja väestömäärään pohjautuva tarkastelu

Puutiaishavaintojen spatiaalinen jakauma Varsinais-Suomessa osoittautui selvästi epätasaiseksi. Joillakin alueilla havaintoja kertyi runsaasti, kun taas osa tutkimusalueesta jäi kokonaan ilman havaintoja. Tämä on linjassa ekologisille kansalaistiedeaineistoille ominaisen spatiaalisen vinouman kanssa. Havaintojen alueellinen jakauma muodostuu sekä havaitsijoiden toiminnan että tutkittavan ilmiön todellisen esiintymisen yhteisvaikutuksesta (Zhang 2020; Zhu ym. 2015). Tässä tutkimuksessa havaintojen keskittyminen ei siten kuvannut suoraan puutiaisten levinneisyyttä, vaan pikemmin niitä ympäristöjä ja tilanteita, joissa ihmiset ja lemmikit altistuvat puutiaisille sekä tekevät niistä havaintoja.

Väestömäärään perustuva tarkastelu osoitti, että asukasluku selitti puutiaishavaintojen määrää melko heikosti. Joillakin alueilla havaintojen määrä kasvoi väestön mukana, mutta kokonaisuudessaan mallin selitysaste jää heikoksi. Tämä tarkoittaa, että vaikka havaintojen määrä oli osittain sidoksissa havaitsijoiden määrään, väestö ei yksin riittänyt selittämään havaintojen spatiaalista vaihtelua. Vastaava ilmiö on tunnistettu myös aiemmissa tutkimuksissa, jossa havaintotiheyden on osoitettu riippuvan väestön lisäksi myös esimerkiksi saavutettavuudesta ja maankäytöstä (Geldmann ym. 2016; Zhang 2020).

Tässä tutkimuksessa väestö ei kuvaa puutiaisten määrää tai ekologista esiintymistä, vaan on epäsuora mittari otannan laajuudelle (*engl.* sampling effort) eli sille, kuinka monta mahdollista havaitsijaa alueella on (Kelling ym. 2015; Zhang 2020). Havaintojen synty edellyttää paitsi puutiaisten esiintymistä myös ihmisen läsnäoloa ja altistumista alueella. Näin ollen väestö kuvaa todennäköisyyttä sille, että havainto voi ylipäätään syntyä. Tämä vastaa aiempaa kansalaistieteeseen pohjautuvaa tutkimusta, jossa havaintomäärien on osoitettu kuvaavan ennen kaikkea havaintointensiteettiä, eikä ilmiön todellista runsautta (Zhang 2020).

Tuloksissa osa alueista poikkesi selvästi väestöön perustuvasta odotuksesta. Poikkeamat eivät kuitenkaan jakaudu satunnaisesti, vaan muodostavat alueellisia

keskittymiä. Odotettua enemmän havaintoja sisältävät alueet painottuvat erityisesti rannikko- ja saaristoalueille, kun taas odotettua vähemmän havaintoja sisältävät alueet keskittyvät sisämaan asutuskeskittymiin ja kaupunkien keskusta-alueille. Tämä tukee aiempia havaintoja siitä, että puutiaisten esiintyminen vähenee nopeasti siirryttäessä rannikkoalueilta kohti sisämaata Varsinais-Suomessa (Sormunen ym. 2016).

Poikkeamien ajallinen pysyvyys vaihteli kuitenkin alueittain. Erityisesti negatiiviset poikkeama-alueet säilyivät vuosien välillä melko samoilla alueilla sisämaan taajamissa ja Turun kaupunkikeskustassa. Tämä viittaa siihen, että näillä alueilla havaintoja kertyi toistuvasti vähemmän kuin väestömäärän perusteella olisi odotettavissa. Sen sijaan positiiviset poikkeamat vaihtelivat saaristo- ja rannikkoalueilla enemmän vuosien välillä, mikä voi heijastaa vapaa-ajan asumisen, virkistyskäytön ja havaintojen satunnaisten kertymisen ajallista vaihtelua.

Havaintojen jakautumiseen vaikuttavat myös muut tekijät kuin pelkkä väestömäärä (Sormunen ym. 2023a). Kansalaistiedeaineistoille tyypillisesti havaintojen kertymiseen vaikuttavat voimakkaasti myös ihmisten liikkuminen, arkiympäristöt ja alueiden saavutettavuus (Geldmann ym. 2016; Johnston ym. 2020). Rannikko- ja saaristoalueilla korkeat havaintomäärät suhteessa väestöön voivat liittyä sekä puutiaisille suotuisiin elinympäristöihin että vapaa-ajan liikkumiseen, kuten mökkeilyyn ja virkistyskäyttöön, jotka lisäävät altistumisen mahdollisuutta. Lisäksi saaristoalueilla ruutujen maapinta-ala voi olla suhteellisen pieni, mikä korostaa havaintotiheyttä ja johtaa suuriin positiivisiin poikkeamiin.

Vastaavasti tiheästi asutuilla kaupunkialueilla havaintoja kertyy vähemmän kuin väestömäärän perusteella olisi odotettavissa. Tämä viittaa siihen, että vaikka potentiaalisia havaittsijoita on paljon, altistuminen puutiaisille on vähäisempää rakennetuissa ympäristöissä, vaikka puutiaisia on havaittu runsaasti myös kaupunkien puistoissa ja joutomailla (Sormunen ym. 2025). Tästä huolimatta havaintojen määrä ei kasva samassa suhteessa väestön kanssa, mikä näkyy negatiivisina poikkeamina. Havaintojen jakautuminen heijastaa siten sekä puutiaisten habitaattiedellytyksiä että ihmisten toimintaa eri ympäristöissä, joten pelkkä väestömäärä ei riitä kuvaamaan tätä ilmiötä kokonaisuudessaan.

Lisäksi havaintojen tarkastelu suhteessa maankäyttöön viittaa siihen, että havaintoja kertyy ympäristöistä, joissa yhdistyvät puutiaisille suotuisat elinolosuhteet ja ihmisten liikkuminen. Eniten havaintoja kertyy metsäisiltä ja kasvillisuudeltaan runsailta alueilta sekä rannikko- ja saaristoalueilta, jotka tarjoavat puutiaisille sopivia mikroilmastoja (Laaksonen ym. 2017; Tanski 2023). Samalla nämä alueet ovat usein myös ihmisten aktiivisessa virkistyskäytössä, mikä lisää altistumisen ja havaintojen syntymisen todennäköisyyttä (Janzén ym. 2024). Sen sijaan tiheästi rakennetuilta kaupunkialueilta havaintoja kertyy vähemmän suhteessa väestömäärään, mikä viittaa siihen, ettei havaintotiheys määräydy suoraan asukastiheyden perusteella (Zhang 2020).

Kokonaisuudessaan kansalaishavainnointiin pohjautuvat puutiaishavainnot osoittavat, että havaintojen määrä on osittain sidoksissa väestöön, mutta merkittävä osa vaihtelusta jää selittämättä. Tämä viittaa siihen, että kansalaisten havaintoihin perustuvassa puutiaisineistossa esiintyy otantaharhaa, joka liittyy havaitsijoiden toimintaan ja ympäristön ominaisuuksiin (Isaac ym. 2014; Zhu ym. 2015; Johnston ym. 2020). Pelkkä väestömäärä ei siten riitä kuvaamaan havaintojen spatiaalista jakautumista, vaan tarvitaan tarkempaa ymmärrystä tekijöistä, jotka ohjaavat havaintojen kertymistä.

## **5.2 Puutiaisten havaintomäärän poikkeaminen asukasmäärästä**

Väestöön perustuvan mallin jäännökset kuvaavat alueita, joilla puutiaishavaintoja esiintyy enemmän tai vähemmän kuin pelkän asukasluvun perusteella olisi odotettavissa. Positiiviset poikkeamat viittaavat alueisiin, joilla havaintoja kertyy odotettua enemmän, kun taas negatiiviset poikkeamat kuvaavat alueita, joilla havaintoja kertyy odotettua vähemmän. Nämä poikkeamat heijastavat erityisesti havaintojen kertymiseen vaikuttavia ympäristö- tai ihmistoimintaan liittyviä tekijöitä.

Tuloksissa korostui ihmisten liikkuminen ja aktiivisuus puutiaishavaintojen selittäjänä. Erityisesti rannikko-, saaristo- ja virkistyskäytössä olevilla alueilla puutiaishavaintoja oli väestömäärään nähden runsaasti, kun taas sisämaan taajama-alueet olivat usein aliedustettuja. Tämä tukee aiempaa tutkimusta, jonka mukaan kansalaistiedeaineistoissa havaintojen synty edellyttää paitsi lajin esiintymistä myös ihmisen läsnäoloa ja havaintohalukkuutta (Callaghan ym. 2020; Johnston ym. 2020).

Näin ollen puutiaishavaintojen spatiaalinen vaihtelu muodostuu ekologisten olosuhteiden ja ihmistoiminnan yhteisvaikutuksesta, mikä vaikeuttaa puutiaisten todellisen esiintymisen erottamista havaitisijatoiminnan aiheuttamasta vinoumasta.

Tässä tutkimuksessa loma-asuntojen määrä oli keskeinen selittäjä positiivisille poikkeamille. Alueilla, joilla on paljon vapaa-ajan asuntoja, puutiaishavaintoja kertyi odotettua enemmän suhteessa vakituiseen väestöön. Nämä alueet painottuivat pääasiassa rannikkoalueille ja saaristoon. Tulokset tukevat aiempaa tutkimusta siitä, että pelkkä asukasluku voi aliarvioida kansalaistiedeaineistossa alueen todellista havaintopotentiaalia (Zhang & Zhu 2018). Vastaavasti Sormunen ym. (2023b) havaitsivat, että kesäasukkaiden tekemät havainnot voivat kasvattaa puutiaishavaintojen määriä suhteessa viralliseen väestöön, koska vapaa-ajan asukkaat eivät sisälly väestötilastoihin. Myös tämän tutkimuksen tulokset osoittivat, että vapaa-ajan asuminen lisää altistumisten ja havaintojen syntymisen todennäköisyyttä.

Varsinais-Suomessa loma-asunnot sijoittuvat rannikko- ja saaristoalueille, joissa myös puutiaisille suotuisat olosuhteet usein toteutuvat, mikä tukee aiempia havaintoja puutiaisten yleisyydestä rannikkoalueilla ja vesistöjen läheisyydessä (Sormunen ym. 2016; Laaksonen ym. 2017). Samankaltaista alueellista keskittymistä on havaittu myös lintuihin perustuvissa kansalaistiedeaineistoissa, joissa havaintojen on todettu painottuvan luonnonympäristöihin ja rannikkoalueille aktiivisen virkistyskäytön seurauksena (Johnston ym. 2020). Siten myös puutiaishavaintojen spatiaalinen rakenne saattaa heijastaa paitsi lajin elinympäristövaatimuksia myös ihmisten luonnossa liikkumista.

Myös polkuverkoston ja liikuntapaikkojen yhteys havaintopoikkeamiin korosti ihmisten liikkumisen merkitystä puutiaishavaintojen kertymisessä. Nämä muuttujat kuvaavat alueiden saavutettavuutta ja sitä, missä luonnonympäristöissä ihmiset liikkuvat aktiivisesti. Aiemman tutkimuksen perusteella juuri helposti saavutettavien ja helppokulkuisten alueiden on havaittu keräävän suhteellisesti enemmän kansalaistiedehavaintoja, koska ne lisäävät havaintotodennäköisyyttä (Geldmann ym. 2016; Zhang 2020). Tulokset viittasivat siihen, että puutiaishavaintojen spatiaalinen

jakautuminen heijasti puutiaisten esiintymisen lisäksi ihmisten aktiivista liikkumista luonnonympäristöissä.

Ihmistoimintaan liittyvien tekijöiden lisäksi ympäristötekijät selittävät havaintojen poikkeamia. Tulokset ovat linjassa aiemman tutkimuksen kanssa, jonka mukaan puutiaisten esiintymiseen vaikuttavat erityisesti elinympäristön rakenne ja mikroilmasto (Ehrmann ym. 2017). Tässä tutkimuksessa kasvillisuuden vihreyttä kuvaavan NDVI:n perusteella väestömäärään suhteutetut havaintopoikkeamat ovat suurempia kasvillisuudeltaan runsailla alueilla. Tämä tukee aiempia havaintoja siitä, että puutiaiset suosivat kosteita ja kasvillisuudeltaan tiheäkasvuisia ympäristöjä, joissa kasvillisuus ylläpitää suotuisia olosuhteita, kuten kosteuden säilymistä ja tasaisempia lämpötiloja (Sormunen ym. 2016; Ehrmann ym. 2017; Tanski 2023). Matalilla NDVI-arvoilla luottamusvälit olivat selvästi suurempia, mikä viittaa suurempaan epävarmuuteen näillä alueilla. Aineiston tarkastelun perusteella nämä matalimmat NDVI-arvot painottuivat voimakkaasti rakennettuihin ympäristöihin, joissa havaintojen määrä voi vaihdella huomattavasti ihmistoiminnan, havaintointensiteetin sekä puutiaisille soveltuvien elinympäristöjen vähäisyyden vuoksi. Siten havaintoja voi kertyä suhteessa väestömäärään joko runsaasti tai hyvin vähän, mikä näkyy poikkeamien jakautumisena ääripäihin.

Rakennetun ympäristön osuus selitti osaltaan havaintojen poikkeamia väestöön perustuvasta odotuksesta. Tässä tutkimuksessa puutiaishavaintojen määrä kasvoi erityisesti silloin, kun rakennetun maankäytön osuus ylitti noin 25 %, mutta vaikutus tasaantui osuuden lähestyessä 50 %:a. Tulokset tukevat aiempia havaintoja rakennetun ympäristön ja puutiaistiheyden välisestä epälineaarista suhteesta. Puutiaistiheyden on havaittu olevan korkeimmillaan ympäristöissä, joissa rakennetun maankäytön osuus on kohtalainen, kun taas sekä hyvin vähäinen että korkea rakennetun ympäristön osuus vähentää puutiaisten esiintymistä (Sormunen ym. 2025). Ilmiö voi liittyä siihen, että rakennetun ja luonnonympäristön rajavyöhykkeet tarjoavat puutiaisille suotuisia mikroilmasto-olosuhteita sekä isäntäeläimille soveltuvia elinympäristöjä (Ehrmann ym. 2017). Samalla kaupunkiympäristöjen infrastruktuuri ohjaa isäntäeläinten liikkumista ja voi keskittää niitä tietyille viherympäristöille (Bourdin ym. 2022). Sen sijaan

voimakkaasti rakennetuilla alueilla sopivien elinympäristöjen ja isäntäeläinten määrä vähenee, mikä rajoittaa puutiaisten esiintymistä.

Tulosten perusteella havaintoja kertyi suhteellisesti enemmän luonnontilaisilta metsäalueilta, vaikka vaikutus jäi kokonaisuudessaan melko pieneksi. Luonnontilaiset metsät tarjoavat puutiaisille suotuisia elinympäristöjä, kuten kosteita ja kasvillisuudeltaan suojaisia alueita (Ehrmann ym. 2017). Ne eivät kuitenkaan usein ole yhtä intensiivisessä virkistyskäytössä tai yhtä helposti saavutettavia kuin rakennetummat alueet. Tulosten perusteella luonnontilaisimmat metsäalueet painottuivat erityisesti saaristoon alueille, joille tie- ja polkuverkosto ei ulotu. Tämä rajoittaa ihmisten liikkumista ja siten myös havaintojen syntymistä. Siten ekologisesti suotuisimmat alueet eivät välttämättä näyttäyty voimakkaasti yliedustettuina havaintoaineistossa. Yksittäiset poikkeukset, kuten Ruissalo, osoittivat kuitenkin, että saavutettavuuden parantuessa myös havaintojen määrä kasvoi luonnontilaisissa ympäristöissä.

Rantaviivan määrä oli keskeinen havaintopoikkeamia selittävä tekijä. Rannikko- ja saaristoalueilla yhdistyvät puutiaisille suotuisat ympäristöolosuhteet, kuten kosteus ja monipuolinen kasvillisuus sekä ihmisten aktiivisuus (Ehrmann ym. 2017; Laaksonen ym. 2017). Tässä tutkimuksessa rantaviivan vaikutus ei kuitenkaan ollut lineaarinen, vaan positiiviset poikkeamat korostuvat erityisesti keskisuurilla rantaviivan arvoilla. Tämä viittaa siihen, että havaintoja kertyy suhteellisesti eniten alueilla, joissa vesistöjen läheisyys lisää sekä puutiaisille soveltuvia elinympäristöjä että ihmisten liikkumista, mutta joissa saavutettavuus säilyy vielä hyvänä.

Sen sijaan kaikkein korkeimmilla rantaviivan arvoilla poikkeamat muuttuivat negatiivisiksi. Nämä alueet sijoittuivat Varsinais-Suomen ulkosaaristoon ja koostuivat useista pienistä saarista, jolloin ruudussa rantaviivaa oli paljon, mutta saavutettavuus oli heikko. Tämä saattaa vähentää ihmisten liikkumista ja havaintojen syntymistä. Samalla pienet ja hajanaiset saaret eivät välttämättä tarjoa puutiaisille yhtä suotuisia elinympäristöjä kuin yhtenäisemmät ja kasvillisuudeltaan runsaammat alueet, sillä niiden kasvillisuus voi olla niukkaa ja mikroilmasto kuiva (Ehrmann ym. 2017; Tanski

2023). Tulokset osoittivat, että rantaympäristöjen vaikutus puutiaishavaintojen kertymiseen määräytyy ympäristöolosuhteiden lisäksi saavutettavuuden kautta.

Tilastollisesti ei-merkitsevien muuttujien osalta tulokset voivat heijastaa sekä tarkastelun mittakaavaa että kansalaistiedeaineiston erityispiirteitä. Esimerkiksi luonnonsuojelualueiden osuus ei osoittautunut merkitseväksi selittäjäksi, vaikka ne tarjoavat usein puutiaisille suotuisia elinympäristöjä. Tämä voi johtua siitä, että havaintojen synty on yhteydessä myös alueiden saavutettavuuteen ja ihmisten liikkumiseen, jotka vaihtelevat suojelualueiden välillä. Lisäksi suojelustatus ei välttämättä sellaisenaan takaa puutiaisille suotuisten elinolosuhteiden esiintymistä. Tässä tutkimuksessa peltojen osuus ei ollut tilastollisesti merkitsevä selittäjä. Tulos voi liittyä siihen, että laajat peltoalueet ovat usein harvaan asuttuja ja puutiaisille epäsuotuisia ympäristöjä, jolloin myös havaintoja kertyy niiltä vähän.

Samansuuntaisesti aiemmissa kansalaistiedeaineistoja tarkastelevissa tutkimuksissa on havaittu, että lajihavaintoja kertyy suhteellisesti vähemmän maatalousvaltaisilta alueilta (Geldmann ym. 2016). Koska sekä havaintojen määrä että väestöön perustuva odotusarvo jäivät peltoalueilla alhaisiksi, ne eivät tuottaneet voimakkaita poikkeamia jäännöksiin eivätkä siten selittäneet niiden vaihtelua merkittävästi.

Tulosten perusteella väestöön perustuvat havaintopoikkeamat eivät jakautuneet satunnaisesti, vaan kytkeytyivät selvästi sekä ympäristötekijöihin että ihmistoimintaan. Positiiviset poikkeamat korostuivat erityisesti alueilla, joilla yhdistyivät puutiaisille suotuisat elinympäristöt ja ihmisten luonnossa liikkuminen. Tällaisia alueita olivat esimerkiksi rannikko- ja saaristoalueet sekä hyvin saavutettavat luonnonympäristöt. Negatiiviset poikkeamat puolestaan painottuvat tiheästi rakennettuihin kaupunkiympäristöihin, joissa väestömäärä on suuri, mutta puutiaishavaintoja kertyy suhteessa asukaslukuun vähemmän. Lisäksi on mahdollista, että tietoisuus Punkkilivestä on ollut vähäisempää alueilla, joilla puutiaisia esiintyi harvemmin, mikä on voinut vaikuttaa havaintojen raportointiin. Tulokset osoittivat, että havaintojen ylliedustus ei välttämättä tarkoita puutiaisten korkeaa tiheyttä eikä aliedustus niiden puuttumista. Poikkeamat heijastavat ennen kaikkea sitä, missä ekologisesti suotuisat ympäristöt ja ihmisten liikkuminen kohtaavat. Samalla tulokset osoittivat, että

kansalaistiedeaineistoihin liittyvät spatiaaliset vinoumat ovat osittain ennustettavissa ympäristö- ja saavutettavuustekijöiden perusteella.

### 5.3 Kansalaistiedeaineiston tulkinta ja menetelmien arviointi

Tulokset osoittivat, että puutiaishavaintojen spatiaalinen vinouma ei ollut satunnainen, vaan kytkeytyi sekä ympäristötekijöihin että ihmisten toimintaan. Tämä tukee aiempaa tutkimusta, jonka mukaan kansalaistiedeaineistojen keskeinen haaste on havaintojen epätasainen maantieteellinen kattavuus ja joka syntyy havaitsijoiden liikkumisen seurauksena (Zhu ym. 2015; Zhang 2020; Johnston ym. 2020).

Kansalaistiedeaineistossa puutiaishavainnot eivät siten kuvaa yksinomaan puutiaisten todellista esiintymistä, vaan myös sitä, missä ja miten ihmiset kohtaavat puutiaisia.

Väestöön perustuva lähestymistapa tarjosi yhden tavan kuvata havaintointensiteettiä, mutta sen selitysvoima jäi rajalliseksi. Tämä korosti sitä, että väestömäärä oli vain epäsuora mittari havaintopotentialille, eikä se yksin riittänyt kuvaamaan havaintojen syntyyn vaikuttavia tekijöitä, kuten ihmisten liikkumista, alueiden saavutettavuutta tai maankäyttöä (Kelling ym. 2015; Zhang 2020). Lisäksi tulokset osoittivat, että väestöön perustuva tarkastelu on herkkä tarkasteluyksikön ominaisuuksille. Erityisesti ruudut, joissa maapinta-alaa oli vähän, mutta havaintoja suhteessa maapinta-alaan paljon, muodostivat suurimmat poikkeamat. Tällaiset ruudut hallitsivat tuloksia ja peittivät alleen mantereisten alueiden vaihtelua, mikä vaikeutti poikkeamien yhteyksien tunnistamista puutiaisten elinympäristöihin ja ihmisten liikkumiseen liittyviin tekijöihin. Kun rajasin nämä ruudut analyysin ulkopuolelle, poikkeamien alueellinen rakenne selkeytyi, mikä paransi myös tulosten tulkittavuutta.

Tässä tutkimuksessa GAM-mallin tulokset osoittivat, että selittävien tekijöiden ja havaintopoikkeamien väliset suhteet olivat monin paikoin epälineaarisia. Tämä tukee tilastollisen mallin soveltuvuutta ekologisten kansalaistiedeaineistojen analysointiin (Dickinson ym. 2010; Isaac ym. 2014; Johnston ym. 2020). Kuitenkaan mallin tulkinta ei ole yksiselitteistä. Malli kuvaa tekijöitä, jotka selittävät poikkeamia väestöön perustuvasta odotuksesta, mutta se ei mahdollista suoraa erottelua sen välillä, johtuuko havaintojen suuri määrä puutiaisten todellisesta runsaudesta vai ihmisten aktiivisesta liikkumisesta ja havainnoinnista. Tulokset korostavat siten

kansalaistiedeaineistoihin liittyvää keskeistä haastetta: havaintojen alueellinen vaihtelu heijastaa samanaikaisesti sekä ekologisia prosesseja että havaitsijakäyttäytymistä.

Lisäksi kansalaistiedeaineistoihin liittyy muita vinoumia, kuten havaitsijakohtainen vaihtelu, niin sanotut superhavaitsijat, raportointikäyttäytymisen erot sekä havaintojen ajallinen epätasaisuus (Kelling ym. 2015; Johnston ym. 2020). Tässä tutkimuksessa keskityttiin erityisesti spatiaaliseen vinoumaan, mutta myös muut otantaharhan muodot voivat vaikuttaa havaintojen kertymiseen ja siten tuloksiin. Aktiiviset havaitsijat voivat tuottaa suuren osan aineistosta, mikä voi korostaa tiettyjä alueita riippumatta niiden ekologisesta soveltuvuudesta. Aiemman tutkimuksen perusteella kansalaistiedeaineistojen yhdistäminen maankäyttöä ja saavutettavuutta kuvaaviin selittäviin muuttujiin on parantanut mallien tulkittavuutta ja luotettavuutta (Geldmann ym. 2016). Samankaltainen ilmiö havaittiin myös tässä tutkimuksessa, jossa erityisesti ihmisten liikkumista ja alueiden saavutettavuutta kuvaavat muuttujat selittivät havaintopoikkeamien alueellista vaihtelua.

Yksi tutkimuksen keskeisistä epävarmuustekijöistä liittyy isäntäeläimiin. Niiden merkitys puutiaisten esiintymiselle ja leviämiselle on keskeinen, mutta tässä tutkimuksessa niiden vaikutusta huomioitiin vain epäsuorasti. Monet ympäristömuuttujat, kuten metsän tai rakennetun ympäristön osuus, vaikuttavat puutiaisten esiintymiseen osittain juuri isäntäeläinten kautta (Ehrmann ym. 2017). Lisäksi on havaittu, että erityisesti kaupunkiympäristöissä isäntäeläinten liikkuminen ja keskittyminen voivat selittää paikallisesti korkeita puutiaistiheyksiä (Sormunen ym. 2025). Tarkkaa tietoa isäntäeläinten runsaudesta ja liikkumisesta ei kuitenkaan ollut saatavilla, minkä vuoksi niiden vaikutusta ei voitu sisällyttää analyysiin omana selittävänä muuttujanaan. Tämän seurauksena isäntäeläinten vaikutus jäi mallissa epäsuorasti muiden ympäristömuuttujien kautta kuvatuksi.

Tässä tutkimuksessa ihmisten liikkumista ja saavutettavuutta kuvaavat muuttujat osoittautuivat keskeisiksi havaintojen alueellista vaihtelua selittäviksi tekijöiksi. Niiden yhteys havaintopoikkeamiin näyttöytyi selvempänä kuin habitaattimuuttujien tai pelkän väestömäärän. Tämä tukee käsitystä siitä, että kansalaistiedeaineistoissa havaintojen kertymistä ohjaa ennen kaikkea ihmisten liikkuminen ja alueiden saavutettavuus, eikä

pelkästään alueella asuvien potentiaalisten havaitsijoiden määrä (Geldmann ym. 2016). Tuloksia voidaan tarkastella laajemmin ihmistoiminnan ja habitaatin välisenä vuorovaikutuksena. Vaikka puutiaisten esiintyminen määräytyy ekologisten tekijöiden, kuten elinympäristön rakenteen ja mikroilmaston perusteella, havaintojen synty edellyttää aina myös ihmisen läsnäoloa ja havainnon kirjaamista.

Samalla on huomioitava, että habitaattien ominaisuuksia, saavutettavuutta ja ihmisten liikkumista kuvaavien muuttujien muodostamiseen liittyy epävarmuutta. Tuloksiin vaikuttaa esimerkiksi se, mitä muuttujia malliin valitaan, miten ne muodostetaan sekä missä määrin ne korreloivat keskenään. Vaikka ihmistoimintaan liittyvät tekijät näyttäytyivät tässä tutkimuksessa keskeisinä havaintojen vaihtelua selittävinä tekijöinä, ne eivät vähennä habitaatin merkitystä puutiaisten esiintymiselle. Habitaatti määrittää, missä puutiaisille on ekologisesti suotuisat olosuhteet, kun taas ihmisten liikkuminen vaikuttaa siihen, missä puutiaisia havaitaan ja raportoidaan. Siten tulokset tukevat aiempia tutkimuksia, joiden mukaan kansalaistiedeaineistojen alueellinen vaihtelu heijastaa samanaikaisesti sekä ympäristötekijöitä että havaintoprosessiin liittyviä vinoumia (Zhang 2020; Johnston ym. 2020).

Jatkossa puutiaishavaintojen tulkintaa voitaisiin tarkentaa sisällyttämällä analyyseihin tietoa isäntäeläimistä. Isäntäeläimet vaikuttavat keskeisesti puutiaisten elinkiertoön ja leviämiseen ylläpitämällä ja kuljettamalla puutiaisia eri ympäristöjen välillä, mutta tarkkaa spatiaalista tietoa niiden runsaudesta ja liikkumisesta ei ole helposti saatavilla (Ehrmann ym. 2017; Laaksonen ym. 2017; Sormunen ym. 2025). Lisäksi dynaamisen väestöaineiston hyödyntäminen voisi kuvata nykyistä täsmällisemmin ihmisten todellista liikkumista ja läsnäoloa luonnossa. Nykyiset väestömuuttujat perustuvat pääasiassa vakituiseen asukaslukuun, eivätkä ne tavoita esimerkiksi matkailua, vapaa-ajan asumista tai päivittäistä liikkumista, jotka voivat vaikuttaa merkittävästi havaintojen syntyyn.

Lisäksi jatkossa voisi olla hyödyllistä tarkastella kaupunki- ja maaseutualueita erikseen, sillä ihmisten liikkuminen, alueiden saavutettavuus sekä puutiaisten elinympäristöt eroavat näissä ympäristöissä merkittävästi toisistaan (Geldmann ym. 2016). Tämän vuoksi myös havaintopoikkeamia selittävien muuttujien vaikutukset voivat vaihdella

alueittain. Väestöön perustuvan mallin sovittaminen erikseen kaupunki- ja maaseutualueille voisi tuottaa paremmin paikallisia olosuhteita kuvaavia jäännöksiä, jolloin niiden yhteyttä esimerkiksi habitaattiin, saavutettavuuteen ja virkistyskäyttöön olisi mahdollista tulkita tarkemmin. Näiden tekijöiden tarkempi huomioiminen voisi auttaa erottamaan aiempaa paremmin puutiaisten todellista esiintymistä ja havaintoprosessiin liittyvää vaihtelua kansalaistiedeaineistoissa.

## 6 Johtopäätökset

Tutkimustulokset osoittavat, että Punkkilive-kansalaistiedeaineiston puutiaishavainnot jakautuvat Varsinais-Suomessa epätasaisesti erityisesti alueellisesti, mutta myös ajallisesti. Ajallinen vaihtelu noudatti pääosin puutiaisten aktiivisuudelle tyypillistä kausivaihtelua. Spatiaalinen vaihtelu oli huomattavasti voimakkaampaa, ja havainnot keskittyivät erityisesti rannikolle, saaristoon ja kaupunkialueiden läheisyyteen.

Väestömäärä vaikutti puutiaishavaintojen kertymiseen, mutta sen selitysvoima jäi heikoksi. Rannikkoalueilla havaintoja esiintyi systemaattisesti enemmän kuin väestömäärän perusteella olisi ollut odotettavissa, kun taas sisämaan taajama-alueilla havaintoja kertyi odotettua vähemmän.

Poikkeama-alueiden tarkastelu osoitti, että havaintojen yliedustusta selittivät erityisesti ihmistoimintaan ja saavutettavuuteen liittyvät tekijät. Loma-asuntojen määrä, rantaviivan määrä, liikuntapaikat, polut sekä rakennetun ympäristön osuus olivat yhteydessä siihen, että puutiaishavaintoja kertyi väestömäärään nähden odotettua enemmän. Tulokset osoittavat, että puutiaishavaintojen alueellinen jakautuminen kytkeytyy ihmisten liikkumiseen ja ympäristöjen saavutettavuuteen, vaikka taustalla vaikuttavat myös puutiaisten elinympäristöihin liittyvät tekijät.

Tutkimus osoittaa, että ekologista kansalaistiedeaineistoa voidaan hyödyntää puutiaisiin liittyvien alueellisten riskien tarkastelussa, mutta tulkinta edellyttää otantaharhan huomioimista. Jatkossa olisi tärkeää hyödyntää ihmisten liikkumista kuvaavia aineistoja, kuten dynaamista väestödataa. Tällöin ihmisten alueellista aktiivisuutta ja liikkumista voitaisiin tarkastella nykyistä tarkemmin. Myös kaupunki- ja maaseutualueiden erillinen tarkastelu voisi tarkentaa havaintopoikkeamien tulkintaa erilaisissa ympäristöissä. Lisäksi jatkossa olisi tärkeää sisällyttää analyyseihin enemmän tietoa puutiaisten isäntäeläimien alueellisesta esiintymisestä, sillä puutiaisten todellinen esiintyminen määräytyy myös isäntäeläinten liikkumisen ja elinympäristöjen perusteella. Näin voitaisiin erottaa tarkemmin puutiaisten todelliseen esiintymiseen liittyvät ekologiset tekijät havaintoprosessiin liittyvästä ihmistoiminnasta.

## Lähteet

- Arazy, O. & Malkinson, D. (2021) A framework of observer-based biases in citizen science biodiversity monitoring: Semi-structuring unstructured biodiversity monitoring protocols. *Frontiers in Ecology and Evolution* 9. <https://doi.org/10.3389/fevo.2021.693602>
- Backstrom, L. J., Callaghan, C. T., Worthington, H., Fuller, R. A. & Johnston, A. (2025) Estimating sampling biases in citizen science datasets. *Ibis* 167(1) 73–87. <https://doi.org/10.1111/ibi.13343>
- Baston, B. (2023) exactextractr: Fast Extraction from Raster Datasets using Polygons. R package version 0.10.0. <https://CRAN.R-project.org/package=exactextractr>
- Bivand, R. & Yu, D. (2024) spgwr: Geographically Weighted Regression. R package version 0.6-37. <https://CRAN.R-project.org/package=spgwr>
- Bonney, R., Cooper, C. B., Dickinson, J., Kelling, S., Phillips, T., Rosenberg, K. V. & Shirk, J. (2009) Citizen science: a developing tool for expanding science knowledge and scientific literacy. *BioScience* 59(11) 977–984. <https://doi.org/10.1525/bio.2009.59.11.9>
- Bonney, R., Shirk, J. L., Phillips, T. B., Wiggins, A., Ballard, H. L., Miller-Rushing, A. J. & Parrish, J. K. (2014) Next steps for citizen science. *Science* 343(6178) 1436–1437. <https://doi.org/10.1126/science.1251554>
- Bourdin, A., Bord, S., Durand, J., Galon, C., Moutailler, S., Scherer-Lorenzen, M. & Jactel, H. (2022) Forest diversity reduces the prevalence of pathogens transmitted by the tick *Ixodes ricinus*. *Frontiers in Ecology and Evolution*. <https://doi.org/10.3389/fevo.2022.891908>
- Brown, E. D. & Williams, B. K. (2019) The potential for citizen science to produce reliable and useful information in ecology. *Conservation Biology* 33(3) 561–569. <https://doi.org/10.1111/cobi.13223>
- Brunsdon, C., Fotheringham, A. S. & Charlton, M. E. (1996) Geographically weighted regression: a method for exploring spatial nonstationarity. *Geographical analysis* 28(4) 281–298. <https://doi.org/10.1111/j.1538-4632.1996.tb00936.x>

- Callaghan, C. T., Poore, A. G., Major, R. E., Rowley, J. J. & Cornwell, W. K. (2019) Optimizing future biodiversity sampling by citizen scientists. *Proceedings of the Royal Society B* 286(1912). <https://doi.org/10.1098/rspb.2019.1487>
- Carvalho, L. G., Kunin, W. E., Keil, P., Aguirre-Gutiérrez, J., Ellis, W. N., Fox, R., Groom, Q., Hennekens, S., Van Landuyt, W., Maes, D., Van de Meutter, F., Michez, D., Rasmont, P., Ode, B., Potts, S. G., Reemer, M., Roberts, S. P. M., Schaminée, J., WallisDeVries, M. F. & Biesmeijer, J. C. (2013) Species richness declines and biotic homogenisation have slowed down for NW-European pollinators and plants. *Ecology letters* 16(7) 870-878. <https://doi.org/10.1111/ele.12121>
- Comber, A., Wang, Y., Lu, Y., Zhang, X. & Harris, P. (2018) Hyper-local geographically weighted regression: extending GWR through local model selection and local bandwidth optimization. *Journal of Spatial Information Science* 17, 63-84.
- Corine maanpeiteaineisto (2018) Suomen ympäristökeskus, Helsinki. 12.3.2026. <https://ckan.ymparisto.fi/dataset/corine-maanpeite-2018>
- Di Cecco, G. J., Barve, V., Belitz, M. W., Stucky, B. J., Guralnick, R. P. & Hurlbert, A. H. (2021) Observing the observers: How participants contribute data to iNaturalist and implications for biodiversity science. *BioScience* 71(11) 1179–1188. <https://doi.org/10.1093/biosci/biab093>
- Díaz-Calafat, J., Jaume-Ramis, S., Soacha, K., Álvarez, A. & Piera, J. (2024) Revealing biases in insect observations: A comparative analysis between academic and citizen science data. *Plos one* 19(7). <https://doi.org/10.1371/journal.pone.0305757>
- Dickinson, J. L., Shirk, J., Bonter, D., Bonney, R., Crain, R. L., Martin, J., Phillips, T. & Purcell, K. (2012) The current state of citizen science as a tool for ecological research and public engagement. *Frontiers in Ecology and the Environment* 10(6) 291-297. <https://doi.org/10.1890/110236>
- Dickinson, J. L., Zuckerberg, B. & Bonter, D. N. (2010) Citizen science as an ecological research tool: challenges and benefits. *Annual review of ecology, evolution, and systematics* 41(1) 149-172. <https://doi.org/10.1146/annurev-ecolsys-102209-144636>

- Ditmer, M. A., Iannarilli, F., Tri, A. N., Garshelis, D. L. & Carter, N. H. (2021) Artificial night light helps account for observer bias in citizen science monitoring of an expanding large mammal population. *Journal of Animal Ecology* 90(2) 330-342. <https://doi.org/10.1111/1365-2656.13338>
- Dunnington, D. (2023) ggspatial: Spatial Data Framework for ggplot2. R package version 1.1.9. <https://CRAN.R-project.org/package=ggspatial>
- Dunnington, D. (2024) prettymapr: Scale Bar, North Arrow, and Pretty Margins in R. R package version 0.2.5. <https://CRAN.R-project.org/package=prettymapr>
- European Citizen Science Association (2015) Ten Principles of Citizen Science. Berlin. <http://doi.org/10.17605/OSF.IO/XPR2N>
- Elwood, S., Goodchild, M. F. & Sui, D. Z. (2012). Researching volunteered geographic information: Spatial data, geographic research, and new social practice. *Annals of the association of American geographers* 102(3) 571-590. <https://doi.org/10.1080/00045608.2011.595657>
- Ehrmann, S., Liira, J., Gärtner, S., Hansen, K., Brunet, J., Cousins, S. A., Deconchat, M., Decocq, G., De Frenne, P., De Smedt, P., Diekmann, M., Gallet-Moron, E., Kolb, A., Lenoir, J., Lindgren, J., Naaf, T., Paal, T., Valdés, A., Verheyen, K., Wulf, M. & Scherer-Lorenzen, M. (2017) Environmental drivers of *Ixodes ricinus* abundance in forest fragments of rural European landscapes. *BMC ecology* 17(1) 31. <https://doi.org/10.1186/s12898-017-0141-0>
- Fisher-Phelps, M., Cao, G., Wilson, R. M. & Kingston, T. (2017) Protecting bias: across time and ecology, open-source bat locality data are heavily biased by distance to protected area. *Ecological informatics* 40, 22-34. <https://doi.org/10.1016/j.ecoinf.2017.05.003>
- Frigerio, D., Richter, A., Per, E., Pruse, B. & Vohland, K. (2021) Citizen science in the natural sciences. Teoksessa Hecker, S., Haklay, M., Bowser, A., Makuch, Z., Vogel, J. & Bonn, A. (toim.) *In The science of citizen science*, 79-96. Springer International Publishing, Cham.
- Geldmann, J., Heilmann-Clausen, J., Holm, T. E., Levinsky, I., Markussen, B. O., Olsen, K., Rahbek, C. & Tøttrup, A. P. (2016) What determines spatial bias in citizen science? Exploring four recording schemes with different proficiency

- requirements. *Diversity and Distributions* 22(11) 1139-1149.  
<https://doi.org/10.1111/ddi.12477>
- Getis, A. (2009) Spatial autocorrelation. Teoksessa Fischer, M. M. & Getis, A. (toim.) *Handbook of applied spatial analysis: Software tools, methods and applications*, 255-278. Springer, Berlin & Heidelberg.
- Gonsamo, A. & D'Odorico, P. (2014) Citizen science: best practices to remove observer bias in trend analysis. *International journal of biometeorology* 58(10) 2159-2163.  
<https://doi.org/10.1007/s00484-014-0806-8>
- Goodchild, M. F. (2007) Citizens as sensors: the world of volunteered geography. *GeoJournal* 69(4) 211-221. <https://doi.org/10.1007/s10708-007-9111-y>
- Grolemund, G. & Wickham, H. (2011) Dates and Times Made Easy with lubridate. *Journal of Statistical Software* 40(3) 1-25. <https://www.jstatsoft.org/v40/i03/>
- Guillera-Arroita, G. (2017) Modelling of species distributions, range dynamics and communities under imperfect detection: advances, challenges and opportunities. *Ecography* 40(2) 281-295. <https://doi.org/10.1111/ecog.02445>
- Gutierrez-Velez, V. H. & Wiese, D. (2020) Sampling bias mitigation for species occurrence modeling using machine learning methods. *Ecological Informatics*.  
<https://doi.org/10.1016/j.ecoinf.2020.101091>
- Haklay, M., Dörler, D., Heigl, F., Manzoni, M., Hecker, S. & Vohland, K. (2021) What is citizen science? The challenges of definition. Teoksessa Vohland, K., Land-Zandstra, A., Ceccaroni, L., Lemmens, R., Perelló, J., Ponti, M., Samson, R. & Wagenknecht, K. (toim.) *In The science of citizen science*, 13-33. Springer International Publishing, Cham.
- Hecker, S., Haklay, M., Bowser, A., Makuch, Z., Vogel, J. & Bonn, A. (2018) Teoksessa Hecker, S., Haklay, M., Bowser, A., Makuch, Z., Vogel, J. & Bonn, A. (toim.) *Citizen science: Innovation in open science, society and policy*, 1–23. Ucl Press.
- Hijmans, R. (2025) terra: Spatial Data Analysis. R package version 1.8-29.  
<https://CRAN.R-project.org/package=terra>
- IPCC (2023) Climate Change 2023: Synthesis Report. Teoksessa Lee, H. & Romero, J. (toim.) *Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change IPCC*, 35-115. Geneva, Switzerland. <https://doi.org/10.59327/IPCC/AR6-9789291691647>

- Irwin, A. (1995) Citizen science. London: Routledge.
- Isaac, N. J., van Strien, A. J., August, T. A., de Zeeuw, M. P. & Roy, D. B. (2014) Statistics for citizen science: extracting signals of change from noisy ecological data. *Methods in Ecology and Evolution* 5(10) 1052–1060.  
<https://doi.org/10.1111/2041-210X.12254>
- Janzén, T., Choudhury, F., Hammer, M., Petersson, M., & Dinnézt, P. (2024) Ticks-public health risks in urban green spaces. *BMC Public Health* 24(1) 1031.  
<https://doi.org/10.1186/s12889-024-18540-8>
- Johnston, A., Matechou, E. & Dennis, E. B. (2023) Outstanding challenges and future directions for biodiversity monitoring using citizen science data. *Methods in Ecology and Evolution* 14(1) 103–116. <https://doi.org/10.1111/2041-210X.13834>
- Johnston, A., Moran, N., Musgrove, A., Fink, D. & Baillie, S. R. (2020) Estimating species distributions from spatially biased citizen science data. *Ecological Modelling* 422. <https://doi.org/10.1016/j.ecolmodel.2019.108927>
- Kahle, D. & Wickham, H. (2013) ggmap: Spatial Visualization with ggplot2.
- Kalliolevo, H., Koskinen, A. & Juvonen, T. (2023) Varsinais-Suomen luonnon monimuotoisuus – Nykytiedot, uhanalaisten luontotyyppien suojele ja ennallistaminen. Varsinais-Suomen liitto.
- Kelling, S., Johnston, A., Hochachka, W. M., Iliff, M., Fink, D., Gerbracht, J., Lagoze, C., La Sorte, F. A., Moore, T., Wiggins, A., Wong, W-K., Wood, C. & Yu, J. (2015) Can observation skills of citizen scientists be estimated using species accumulation curves? *PloS one* 10(10) e0139600.  
<https://doi.org/10.1371/journal.pone.0139600>
- Kovács, L. (2022) Feature selection algorithms in generalized additive models under concavity. *Computational Statistics*.
- Kulha, N., Ruokolainen, K., Vesterinen, E. J., Lamppu, M., Klemola, T. & Sormunen, J. J. (2022) Does environmental adaptation or dispersal history explain the geographical distribution of *Ixodes ricinus* and *Ixodes persulcatus* ticks in Finland? *Ecology and Evolution* 12(12) e9538. <https://doi.org/10.1111/2041-210X.13834>

- Kunttu, P., Junninen, K. & Kouki, J. (2015) Dead wood as an indicator of forest naturalness: A comparison of methods. *Forest Ecology and Management* 353, 30-40.
- Laaksonen, M., Sajanti, E., Sormunen, J., Penttinen, R., Hänninen, J., Ruohomäki, K., Sääksjärvi, I., Vesterinen E., Vuorinen, I., Hytönen, J. & Klemola, T. (2017) Crowdsourcing-based nationwide tick collection reveals the distribution of *Ixodes ricinus* and *I. persulcatus* and associated pathogens in Finland. *Emerging Microbes & Infections* 6, e31. <http://dx.doi.org/10.1038/emi.2017.17>
- Lipas-data 10/2025 (2025) Jyväskylän yliopisto, Jyväskylä. 12.3.2026. [https://jyx.jyu.fi/jyx/Record/jyx\\_123456789\\_106354](https://jyx.jyu.fi/jyx/Record/jyx_123456789_106354)
- Luonnonsuojelu- ja erämaa-alueet (2026) Metsähallitus, Helsinki. 10.3.2026. <https://ckan.ymparisto.fi/dataset/luonnonsuojelu-ja-eramaa-alueet>
- Maastotietokanta, rakennukset (2026) Maanmittauslaitos. 12.3.2026. <https://www.maanmittauslaitos.fi/kartat-ja-paikkatieto/aineistot-ja-rajapinnat/tuotekuvaukset/maastotietokanta>
- Madetoja, J. (2018) Error propagation in geographically weighted regression. Doctoral dissertations, Aalto yliopisto, 10/2018.
- Matkailu ja virkistys Varsinais-Suomessa (2026) Varsinais-Suomen liitto. 1.5.2026. <https://varsinais-suomi.fi/varsinais-suomi/matkailu/>
- Mikkonen, N., Leikola, N., Lahtinen, A., Lehtomäki, J. & Halme, P. (2018) *Monimuotoisuudelle tärkeät metsäalueet Suomessa: Puustoisten elinympäristöjen monimuotoisuusarvojen Zonation-analyysien loppuraportti*. Suomen ympäristökeskuksen raportteja 9/2018. <https://helda.helsinki.fi/handle/10138/234359>
- Millar, E. E., Hazell, E. C. & Melles, S. J. (2019) The ‘cottage effect’ in citizen science? Spatial bias in aquatic monitoring programs. *International Journal of Geographical Information Science* 33(8) 1612-1632. <https://doi.org/10.1080/13658816.2018.1423686>
- Monimuotoisuudelle tärkeät metsäalueet Suomessa (2018) Suomen ympäristökeskus, Helsinki. 18.3.2026. <https://ckan.ymparisto.fi/dataset/monimuotoisuudelle-tarkeat-metsaalueet-high-biodiversity-value-forests-2018-zonation-fin-eng-sw>

- Nahhas, R. W. (2024) Introduction to regression methods for public health using R. Chapman and Hall/CRC. <https://doi.org/10.1201/9781003263197>
- NDVI:n maksimiarvo (2021) Suomen ympäristökeskus, Helsinki. 12.3.2026. <https://ckan.ymparisto.fi/dataset/ndvi-n-maksimiarvo-v-2021>
- Pebesma, E. & Bivand, R. (2023). Spatial Data Science: With Applications in R. Chapman and Hall/CRC. <https://doi.org/10.1201/9780429459016>
- Pebesma, E. (2018) Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal* 10 (1) 439-446. <https://doi.org/10.32614/RJ-2018-009>
- Peel, M. C., Finlayson, B. L. & McMahon, T. A. (2007) Updated world map of the Köppen-Geiger climate classification. *Hydrology and earth system sciences* 11(5) 1633-1644. <https://doi.org/10.5194/hess-11-1633-2007>
- Punkkilive (2026) Tietoa Punkkilivestä. 7.1.2026. <https://www.punkkilive.fi/tietoa-punkkilivesta>
- Robinson, L. D., Cawthray, J. L., West, S. E., Bonn, A., & Ansine, J. (2018) Ten principles of citizen science. Teoksessa Hecker, S., Haklay, M., Bowser, A., Makuch, Z., Vogel, J. & Bonn, A. (toim.) *Citizen science: Innovation in open science, society and policy*. 27–40. Ucl Press.
- Rosário, I. T., Tiago, P., Chozas, S. & Capinha, C. (2025) When do citizen scientists record biodiversity? Non-random temporal patterns of recording effort and associated factors. *People and Nature* 7(4) 860-870. <https://doi.org/10.1002/pan3.70017>
- Royle, J. A., Kéry, M., Gautier, R., & Schmid, H. (2007) Hierarchical spatial models of abundance and occurrence from imperfect survey data. *Ecological Monographs* 77(3) 465-481. <https://doi.org/10.1890/06-0912.1>
- Sauermann, H. & Franzoni, C. (2015) Crowd science user contribution patterns and their implications. *Proceedings of the national academy of sciences* 112(3) 679–684. <https://doi.org/10.1073/pnas.1408907112>
- Scher, C. L. & Clark, J. S. (2023) Species traits and observer behaviors that bias data assimilation and how to accommodate them. *Ecological Applications* 33(3) e2815. <https://doi.org/10.1002/eap.2815>
- See, L., Mooney, P., Foody, G., Bastin, L., Comber, A., Estima, J., Fritz, S., Kerle, N., Jiang, B., Laakso, M., Liu, H-Y., Milčinski, G., Nikšič, M., Painho, M., Pöddör, A.,

- Olteanu-Raimond, A-M. & Rutzinger, M. (2016) Crowdsourcing, citizen science or volunteered geographic information? The current state of crowdsourced geographic information. *ISPRS International Journal of Geo-Information* 5(5) 55. <https://doi.org/10.3390/ijgi5050055>
- See, L., Olteanu-Raimond, A. M. & Fonte, C. C. (2025) Recent advances in Volunteered Geographic Information (VGI) and citizen sensing. *International Journal of Digital Earth* 18(1) 2480220. <https://doi.org/10.1080/17538947.2025.2480220>
- Shahi, K. (2023) Volunteered geographic information (VGI) in spatial data infrastructure (SDI) continuum. *EAI Endorsed Transactions on Internet of Things* 9(1) e3. <https://doi.org/10.4108/eetiot.v9i1.2979>
- Sormunen, J. J. (2026) Citizen Science Tick Observations Serve as an Early Warning System for Tick-Borne Diseases. *Zoonoses and Public Health* 73(3) 234-243. <https://doi.org/10.1111/zph.70045>
- Sormunen, J. J., Kulha, N., Alale, T. Y., Klemola, T., Sääksjärvi, I. E. & Vesterinen, E. J. (2023a) For the people by the people: Citizen science web interface for real-time monitoring of tick risk areas in Finland. *Ecological Solutions and Evidence* 4(4) e12294. <https://doi.org/10.1002/2688-8319.12294>
- Sormunen, J. J., Kulha, N., Klemola, T., Mäkelä, S., Vesilahti, E. M. & Vesterinen, E. J. (2020) Enhanced threat of tick-borne infections within cities? Assessing public health risks due to ticks in urban green spaces in Helsinki, Finland. *Zoonoses and Public Health* 67(7) 823-839. <https://doi.org/10.1111/zph.12767>
- Sormunen, J. J., Kylänpää, S., Sippola, E., Elo, R., Kiran, N., Pakanen, V-M., Kallio, E. R., Vesterinen, E. J. & Klemola, T. (2025) There Goes the Neighbourhood – A Multi-City Study Reveals Ticks and Tick-Borne Pathogens Commonly Occupy Urban Green Spaces. *Zoonoses and Public Health* 72(3) 313–323. <https://doi.org/10.1111/zph.13208>
- Sormunen, J. J., Penttinen, R., Klemola, T., Hänninen, J., Vuorinen, I., Laaksonen, M., Sääksjärvi, I. E., Ruohomäki, K. & Vesterinen, E. J. (2016) Tick-borne bacterial pathogens in southwestern Finland. *Parasites Vectors* 9(1) 1–10. <https://doi.org/10.1186/s13071-016-1449-x>
- Sormunen, J. J., Sääksjärvi, I. E., Vesterinen, E. J. & Klemola, T. (2023b) Crowdsourced tick observation data from across 60 years reveals major increases and

- northwards shifts in tick contact areas in Finland. *Scientific Reports* 13, 21274.  
<https://doi.org/10.1038/s41598-023-48744-8>
- Sullivan, B. L., Aycrigg, J. L., Barry, J. H., Bonney, R. E., Bruns, N., Cooper, C. B., Damoulas, T., Dhondt, A. A., Dietterich, T., Farnsworth, A., Fink, D., Fitzpatrick, J. W., Fredericks, T., Gerbracht, J., Gomes, C., Hochachka, W. M., Iliff, M. J., Lagoze, C., La Sorte, F. A., Merrifield, M., Morris, W., Phillips, T. B., Reynolds, M., Rodewald, A. D., Rosenberg, K. V., Trautmann, N. M., Wiggins, A., Winkler, D. W., Wong, W-K, Wood, C. L. & Kelling, S. (2014) The eBird enterprise: An integrated approach to development and application of citizen science. *Biological conservation* 169, 31-40. <https://doi.org/10.1016/j.biocon.2013.11.003>
- Tanski, N. (2023) Puutiaisten (*Ixodes ricinus* ja *Ixodes persulcatus*) levinneisyyteen ja runsauteen vaikuttavat tekijät Suomessa. *Terra* 135(1) 27-38.  
<https://doi.org/10.30677/terra.116130>
- Väestöruutuaineisto 1 km x 1 km (2023) Tilastokeskus, Helsinki. 14.12.2025.  
<https://stat.fi/fi/palvelut/tilastodatapalvelut/paikkatietoaineistot/vaestoruutuaineisto-1-km>
- Van Strien, A. J., Van Swaay, C. A. & Termaat, T. (2013) Opportunistic citizen science data of animal species produce reliable estimates of distribution trends if analysed with occupancy models. *Journal of Applied Ecology* 50(6) 1450–1458.  
<https://doi.org/10.1111/1365-2664.12158>
- Vohland, K., Land-Zandstra, A., Ceccaroni, L., Lemmens, R., Perelló, J., Ponti, M., Samson, R. & Wagenknecht, K. (2021) The science of citizen science evolves. Teoksessa Vohland, K., Land-Zandstra, A., Ceccaroni, L., Lemmens, R., Perelló, J., Ponti, M., Samson, R. & Wagenknecht, K. (toim.) *The Science of Citizen Science*, 1-12. Springer International Publishing, Cham.
- Warton, D. I., Renner, I. W. & Ramp, D. (2013) Model-based control of observer bias for the analysis of presence-only data in ecology. *PloS one* 8(11) e79168.  
<https://doi.org/10.1371/journal.pone.0079168>
- Welvaert, M. & Caley, P. (2016) Citizen surveillance for environmental monitoring: combining the efforts of citizen science and crowdsourcing in a quantitative data framework. *SpringerPlus* 5(1) 1890.

- Wickham, H. (2016) *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.
- Wickham, H., François, R., Henry, L., Müller, K. & Vaughan, D. (2023) *dplyr: A Grammar of Data Manipulation*. R package version 1.1.4. <https://CRAN.R-project.org/package=dplyr>
- Wood, S. N. (2011) Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society (B)* 73(1) 3-36.
- Wood, S. N. (2017) *Generalized additive models: an introduction with R*. Chapman and hall/CRC Press.
- Zhang, G. (2020) Spatial and temporal patterns in volunteer data contribution activities: A case study of eBird. *ISPRS International Journal of Geo-Information* 9(10) 597. <https://doi.org/10.1371/journal.pone.0079168>
- Zhang, G. & Zhu, A. X. (2018) The representativeness and spatial bias of volunteered geographic information: a review. *Annals of GIS* 24(3) 151–162. <https://doi.org/10.1080/19475683.2018.1501607>
- Zhu, A. X., Zhang, G., Wang, W., Xiao, W., Huang, Z. P., Dunzhu, G. S., Ren, G., Qin, C-Z., Yang, L., Pei, T. & Yang, S. (2015) A citizen data-based approach to predictive mapping of spatial variation of natural phenomena. *International Journal of Geographical Information Science* 29(10) 1864-1886. <https://doi.org/10.1080/13658816.2015.1058387>
- Zuur, A. F., Ieno, E. N., Walker, N. J., Saveliev, A. A. & Smith, G. M. (2009) Things are not always linear; additive modelling. Teoksessa Zuur, A. F., Ieno, E. N., Walker, N. J., Saveliev, A. A., & Smith, G. M. (toim.) *Mixed effects models and extensions in ecology with R*, 35-69. New York, Springer New York.