



HELSINGIN YLIOPISTO  
MAATALOUS-METSÄTIETEELLINEN TIEDEKUNTA

# **Metsähanhen alalajien tunnistaminen nokan morfologian ja värityksen avulla**

Erika Jumppanen  
Maisterintutkielma  
Helsingin yliopisto  
Metsätieteiden maisteriohjelma  
Metsien ekologia ja käyttö  
Toukokuu 2022

Tiedekunta - Fakultet – Faculty		Laitos - Institution - Department	
Maatalous-metsätieteellinen tiedekunta		Metsätieteiden osasto, metsätieteiden maisteriohjelma	
Tekijä - Författare – Author			
Jumppanen, Erika Aune Maria			
Työn nimi - Arbetets titel – title			
Metsähanhen alalajien tunnistaminen nokan morfologian ja värityksen avulla			
Oppiaine /Läroämne – Subject			
Metsien ekologia ja käyttö			
Työn laji/ Ohjaaja - Arbetets art/Handledare - Level/Instructor		Aika - Datum - Month and year	Sivumäärä - Sidoantal - Number of pages
Maisterintutkielma / Veli-Matti Väänänen		Toukokuu 2022	30s + liitteet 6s
Tiivistelmä - Referat – Abstract			
<p>Euroopassa esiintyy kahta metsähanhen alalajia: tundrametsähänhi <i>Anser fabalis rossicus</i> sekä taigametsähänhi <i>Anser fabalis fabalis</i>. Taigametsähänhiä pesii Suomessa, mutta myös tundrametsähänhiä esiintyy kevät- ja syysmuuttojen aikana. Tundrametsähanhen muuttoreitti kulkee kuitenkin idempää kuin taigametsähanhen, ja ne myös muuttavat myöhemmin syksyllä kuin taigametsähanhet. Metsähanhen metsästys on ollut Suomessa viime vuosina tiukasti säädeltyä ja metsästystä on pyritty ajallisilla sekä paikallisilla rajoituksilla kohdentamaan elinvoimaiseen tundrametsähanheen. Saaliissa esiintyy kuitenkin edelleen molempia alalajeja. Alalajien erottaminen toisistaan silmänvaraisesti on vaikeaa, ja siksi olisi tärkeää saada tietoa saaliin alalajikoostumuksesta kestävänsä metsästyksen pohjaksi.</p> <p>Tutkimuksen tavoitteena oli löytää taiga- sekä tundrametsähanhen nokan muodoista tilastollisesti merkitseviä eroavaisuuksia morfometriikan ja nokan värityksen avulla. Tuloksista on tarkoitus luoda työkalu, jonka avulla alalaji voidaan tunnistaa tietyllä todennäköisyydellä valokuvasta. Lajintunnistus on tärkeää hanhien kannanhoidon kannalta, jotta metsähanhisaaliin alalajijakaumaa on mahdollista seurata. Saalislintujen tunnistaminen on välttämätöntä taigametsähanhen kannan elinvoimaisuuden turvaamisessa.</p> <p>Tutkimuksen aineisto koostuu noin yhdeksästä sadasta hanhen päästä. Aineisto kerättiin Ruotsista vuosina 2017–2018 sekä Suomesta ja Virossa vuosina 2018–2019. Päistä otettiin DNA-näyte sekä standardisoitu kuva. DNA-tulosten avulla oli mahdollista tutkia alalajien eroja, jotka voisivat auttaa tunnistuksessa. Kuvista otettiin myös absoluuttisia mittoja sekä mitattiin keltaisen värin määrää ylänokassa. Mallinnuksessa selvitettiin, onko mitoilla mahdollista selittää metsähanhen alalajia.</p> <p>Tutkielmassa esitellyn tilastollisen mallin avulla oli mahdollista neljän muuttujan: keltaisen värin määrän nokassa, nokan pituuden ja ala- ja ylänokan paksuuden perusteella ennustaa alalaji noin 93 prosentin tarkkuudella. Tulokset tukivat aiempia taksonomisia tutkimuksia, mutta eivät ole suoraan verrannollisia. Tulosten mukaan tundrametsähanhiä tulee yleisesti metsähanhisaaliiseen Suomessa sekä Virossa ja taigametsähanhiä on valtaosa metsähanhisaaliista Ruotsissa.</p>			
Avainsanat – Nyckelord			
Metsähänhi, taigametsähänhi, tundrametsähänhi, morfologia			
Säilytyspaikka - Förvaringsställe - Where deposited			
Helsingin yliopiston kirjasto – Helda / E-thesis (opinnäytteet) <a href="https://ethesis.helsinki.fi">ethesis.helsinki.fi</a>			
Muita tietoja - Övriga uppgifter - Additional information			

## Sisällysluettelo

1. Johdanto .....	3
1.1 Metsähanhen alalajit .....	3
1.2 Metsähanhen alalajien muuttoreitit sekä pesimä- ja talvehtimisalueet.....	3
1.3 Metsähanhen eri alalajien kannan kehitys ja koko .....	5
1.4 Metsähanhen alalajien tunnistaminen .....	6
1.5 Metsähanhien eri alalajien metsästys ja esiintyminen Suomessa metsästyskauden aikana .....	8
1.6 Tutkimuksen tavoitteet.....	10
2. Aineisto ja menetelmät .....	11
2.1 Näytteet ja DNA-analyysit.....	11
2.2 Näytteiden kuvaaminen ja mittaaminen.....	12
2.3 Tilastollinen analyysi ja muuttujien valinta.....	14
2.3.1 Mallin valinta ja ennustetarkkuuden testaus .....	16
3. Tulokset.....	16
3.1 Mallinnus .....	16
3.2 Metsähanhisaaliin alalajijakauma eri maissa .....	20
4. Tulosten tarkastelu .....	21
4.1 Metsähanhen alalajin määrittäminen .....	21
4.2 Ennustavat muuttujat .....	21
4.3 Alalajijakauma .....	22
5. Johtopäätökset.....	23
6. Kiitokset .....	23
7. Lähteet.....	24
8. Liitteet: .....	29

# 1. Johdanto

## 1.1 Metsähanhen alalajit

Metsähanhi *Anser fabalis* on sorsalintuihin kuuluva lintulaji, joka esiintyy Aasiassa sekä Euroopassa. Euroopassa esiintyvät metsähanhen alalajit pesivät Norjasta Uralille ulottuvalla alueella, ja talvehtii Brittein saarilta aina Ruotsiin ja Serbiaan asti ulottuvalla alueella (Nilsson ym. 1999, Heinicke 2018, Heinicke ym. 2018). Metsähanhen taksonomiasta on kiistelty pitkään (Johansen 1945, Delacour 1951, Vaurie 1965, Cramp ja Simmons 1977, Sangster ja Oreel 1996, Ruokonen ym. 2008, Heinicke 2009, Heinicke 2010), ja se on vieläkin osittain epäselvä (Ruokonen ja Aarvak 2011, Ottenburghs ym. 2016, Ottenburghs ym. 2017, Ottenburghs ym. 2020). Tässä tutkielmassa käytetään Ruokosen ym. (2008) esittämää, molekyyligeneettisen aineistoon perustuvaa jakoa, jossa metsähanhesta *Anser fabalis* erotetaan neljä alalajia: Läntinen taigametsähanhi *Anser fabalis fabalis*, läntinen tundrametsähanhi *A. f. rossicus*, itäinen taigametsähanhi *A. f. middendorffii* ja itäinen tundrametsähanhi *A. f. serrirostris*. Tämä jako on nykyään laajimmin hyväksytty (esim. Fox ja Leafloor 2018), vaikka läntistä tundrametsähanhea on ehdotettu omaksi lajikseen *Anser serrirostris rossicus* (Johansen 1945). Tuorein molekyyligeneettinen tutkimus (Ottenburghs ym. 2020) suosittelee näiden kahden kohtelemista alalajeina, joten tässä tutkielmassa noudatetaan edellä mainittua luokittelua ja tarkastellaan läntistä taiga- ja tundrametsähanhea alalajeina. Tämän työn kannalta ei kuitenkaan ole ratkaisevaa, pidetäänkö tundrametsähanhea omana lajinaan vai metsähanhen alalajina. Keskeistä on, että taiga- ja tundrametsähanhi eroavat geneettisesti (Ruokonen ym. 2008) ja niitä voidaan myös kannanhoidon ja suojelun näkökulmasta pitää erillisinä populaatioina (Heinicke 2018, Heinicke ym. 2018).

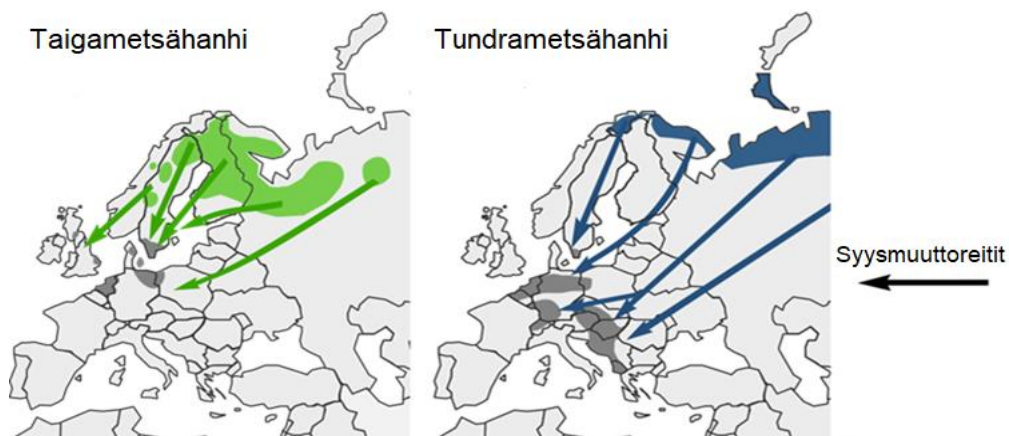
## 1.2 Metsähanhen alalajien muuttoreitit sekä pesimä- ja talvehtimisalueet

Taigametsähanhi pesii boreaalisilla metsäalueilla, tyypillisesti märillä soilla tai soistuneiden vesistöjen rannoilla (Nilsson ym. 1999). Taigametsähanhia pesii Pohjois-Venäjällä, Suomessa, Ruotsissa sekä Norjassa (Kuva 1) (Heinicke ym. 2018). Tundrametsähanhi pesii tundravyyhykkeellä, vesistöjen läheisyydessä (van den Bergh 1999) pesimäalueet ulottuvat Kuolan niemimaalta aina

Taimyrin niemimaalle saakka (Heinicke 2018). Alalajien pesimäalueet ovat kuitenkin osittain päällekkäisiä (Ruokonen ym. 2008).

Taigametsähanhien muuttoreitit pesimäalueilta talvehtimisalueille jaetaan neljään eri muuttoreittiin: läntiseen, keskiseen ja kahteen itäiseen reittiin (Marjakangas ym. 2015, Heinicke ym. 2018). Läntistä reittiä käyttävät linnut pesivät Keski-Ruotsissa ja -Norjassa sekä Ruotsin Etelä-Lapissa ja talvehtivat Luoteis-Jyllannissa sekä Iso-Britanniassa. Keskistä reittiä käyttävät linnut pesivät Pohjois-Lapissa; Ruotsin ja Norjan puolella, sekä Suomessa ja Kuolan niemimaalla. Talvehtimaan linnut päätyvät Etelä-Ruotsiin sekä Kaakkois-Tanskaan. Ensimmäistä itäistä reittiä käyttävät linnut pesivät läntisessä Länsi-Siperiassa ja Pohjois-Petšorassa ja talvehtivat Koillis-Saksassa ja Luoteis-Puolassa. Toisen itäisen reitin linnut pesivät pohjoisessa Länsi-Siperiassa ja talvehtivat Keski-Aasiassa. Aasian tarkat talvehtimisalueet eivät kuitenkaan ole tiedossa (Heinicke ym. 2018)

Tundrametsähanhen muuttoreitti voidaan myös jakaa kahteen. Läntistä reittiä itäisen Ruotsin yli käyttävät Pohjois-Skandinaviassa pesivät linnut (Heinicke 2010). Itäistä reittiä käyttävät Pohjois-Venäjällä aina Uralilla asti pesivät linnut. Talvehtimisalueet jaetaan usein Pohjanmeren ja Baltian alueeseen (Saksa, Alankomaat ja Puola) sekä Keski-Euroopan alueeseen (Unkari, Itävalta ja Slovakia). Talvehtimisalueet eivät kuitenkaan rajoitu valtioiden rajojen mukaan, vaan yksilöitä löytyy myös lähivaltioiden alueilta Ruotsia ja Serbiaa myöten (Heinicke 2018).



Kuva 1 Taigametsähanhen sekä tundrametsähanhen pesimisalueet sekä syysmuuttoreitit. Nuolet kuvaavat populaatioiden syysmuuttoreittejä. Kuva piirretty Marjakangas ym. (2015) sekä Fox ja Leafloor (2018) perusteella. Muokattu Honka ym. (2022) ja Piironen ym. (2022) mukaan.

### 1.3 Metsähanhen eri alalajien kannan kehitys ja koko

Metsähanhen molempien alalajien populaatioiden koon ja kehityksen seuranta Euroopassa perustuu lintujen laskemiseen talvehtimisalueilla sekä muutonaikaisilla levähdysalueilla (Fox ja Leafloor 2018). Pesimäkauden ulkopuolella laskenta on helpompaa kuin laajoilla ja vaikeakulkuisilla pesimäalueilla. Kevät- ja syysmuuttojen aikana linnut kerääntyvät levähtämään ja talvehtimaan suppeammille ja helpommin saavutettaville alueille (Piironen ym. 2022b). Populaation koon arviointia talvisin vaikeuttaa kuitenkin talvehtimisalueen laajuus ja se, ettei laskennoissa alalajeja ole aina eritelty toisistaan (Heinicke ja de Jong 2013, Heinicke ym. 2018). Levähdysalueilla tehtävää arvioinnin tarkkuutta vaikeuttaa mahdollinen yksilön laskeminen useampaan otteeseen tai kokonaan laskematta jättäminen. Talvehtimisalueilla suoritettavat laskennat ovat myös mahdollisesti aliarvioineet metsähanhen alalajien kannankokoa (Heinicke ym. 2018, Piironen ym. 2022b).

Taigametsähanhien määrä Euroopassa on laskenut 1990-luvulta, jolloin lintuja arvioitiin olevan noin 100 000 yksilöä (Nilsson ym. 1999, Fox ja Madsen 1999). Populaatiokoon arvioitiin vuonna 2005 olevan 70 000–90 000 yksilöä ja vuonna 2009 yhteensä 63 000 yksilöä (Delany ja Scott 2006, Fox ym. 2010). Viimeisimmän arvion (Johnson ym. 2020) mukaan taigametsähanhia oli Euroopassa noin 75 000. Arvioiden laatua on kyseenalaistettu monesta syystä (Fox ym. 2010): laskennat ovat kattaneet vain osan taigametsähanhen esiintymisalueesta, arviot ovat perustuneet asiantuntija-arvioon ja alalajeja ei välttämättä ole laskennoissa eritelty toisistaan. Koko taigametsähanhikannan koon arviointia on vaikeuttanut myös muun muassa tiedon vähäisyys Keski-Aasian populaation koosta (Heinicke 2010). Lintujen määrän on kuitenkin arvioitu laskeneen 1990-luvulta niin talvehtimisalueilla Euroopassa kuin pesimisalueilla Siperiassa (Marjakangas ym. 2015).

Tundrametsähanhet ovat Euroopassa taigametsähanhia runsaslukuisempia (Fox ym. 2010). Euroopassa talvehtivien tundrametsähanhien määrää on seurattu 1980-luvulta lähtien, jolloin arviolta Euroopassa talvehti noin 250 000 yksilöä. Viimeisimmän julkaistun arvion mukaan tundrametsähanhia on noin 600 000 (Heinicke 2018). Keski-Euroopassa talvehtivien yksilöiden lukumäärän on huomattu laskevan ja Pohjanmeren ja

Baltian alueen populaatioiden kasvavan. Ilmiön taustalla saattaa yksinkertaisesti olla yksilöiden siirtyminen toisille talvehtimisalueille (Heinicke 2018).

#### 1.4 Metsähanhen alalajien tunnistaminen

Taigametsähanhi ja tundrametsähanhi ovat hyvin samannäköisiä (ks. Kuva 2), eikä alalajeja voida varmasti tunnistaa pelkän ulkonäön perusteella, sillä tuntomerkeissä on päällekkäisyyksiä (Delacour 1951). Burgers ym. (1991) tutkivat Alankomaissa tavattujen metsähanhien alalajien morfologisia eroja. Alankomaissa rengastuksen ohella on vuosien 1954–1986 aikana tunnistettu 12 176 hanhea, jotka onnistuttiin jakamaan kahteen eri alalajityyppiin pään ja nokan muodon perusteella. Otannasta erotetut alalajityypit tunnistettiin kirjallisuuden perusteella olevan tundrametsähanhi ja taigametsähanhi. Taigametsähanhien kuvaillaan tutkimuksessa olevan pitkänokkaisia sekä kapeamman alanokan omaavia. Tundrametsähanhien alanokassa irvistys oli suurempi ja nokan hampaat sahalaitaisemmat, lisäksi nokan todettiin olevan väritykseltään tummempi kuin taigametsähanhen. On olemassa myös poikkeustapauksia, kuten täysin mustanokkaisia taigametsähanhia. Taigametsähanhet olivat keskimäärin suurempia kuin tundrametsähanhet. Nokan pituus oli tutkimuksen mukaan yksinään yhtä hyvä alalajien eron selittäjä kuin lintujen paino sekä siiven pituus yhdessä (Burgers ym 1991).



Kuva 2 Taigametsähanhi (oikealla) ja tundrametsähanhi (vasemmalla). Kuvat: Antti Piironen

Heinicken (2010) tutkimuksessa ottamat mitat erosivat hieman vertailukohteesta (Cramp & Simmons 1977), mutta tulokset vastasivat muita aiempia tutkimuksia. Taigametsähanhen nokat olivat pidempiä, ja väritykseltään keltaisempia. Tundrametsähanhilla havaittiin olevan selkeämpi irvistys eli rako ylä- ja alanokan välissä. Tundrametsähanhien nokan väritys vaihteli samalla tavalla kuin Burgersin ym. (1991) tutkimuksessa.

Uusimmassa metsähanhen alalajien tunnistamista käsittelevässä tutkimuksessa de Jong (2019) esitti, millä tundrametsähanhen ja taigametsähanhien pään mitoilla olisi mahdollista tunnistaa alalajit toisistaan. Tavoitteena oli tunnistaa alalaji mahdollisimman vähillä muuttujilla. Paras yksittäinen alalajia selittävä muuttuja oli hampaiden lukumäärä. Myös alanokan paksuus ja kynnen pituus mainittiin sopivina muuttujina selittämään alalajien eroja. Kynnen pituudella tarkoitetaan nokan kärjen kuperaa osaa, joka voi olla edestä katsottuna joko pyöreä tai ovaali. Hampaiden lukumäärä on kuitenkin ongelmallinen suure, koska hanhen hampaat pienenevät vähitellen nokan molemmissa päissä, eikä voida yksiselitteisesti määrittellä, milloin hampaat loppuvat. Tämä luo vaihtelua mitaajien välille ja tekee eri henkilöiden laskemien hampaiden lukumäärien vertaamisen vaikeaksi (Piironen ym., julkaisematon aineisto). Lisäksi de Jongin (2019) ongelma on, että vastemuuttujan (metsähanhen alalaji) arvoa ei ole määritetty selittävästä muuttujista riippumattomasti. Vastemuuttujan arvo on yhden asiantuntijan, Thomas Heinicken näkemys kunkin linnun alalajista, joten se voi perustua tutkimuksessa myös selittävinä muuttujina oleviin morfologisiin ominaisuuksiin. Vaikka de Jongin (2019) tulokset eivät ole ristiriidassa aiempien tutkimusten kanssa ja lienevät siten oikeansuuntaisia, on kyseisiä tuloksia sovellettaessa huomioitava riski kehäpäätelmästä. Kaikkien edeltä mainittujen tutkimuksen alalajien silmänvarainen tunnistaminen tarkoittaa kuitenkin, että tulokset eivät ole suoraan verrattavissa tämän tutkimuksen tuloksiin.

Aiemmat tutkimukset ovat yksimielisiä muun muassa nokan värin olevan erottava tekijä taiga- ja tundrametsähanhen välillä, mutta kuinka luotettavasti alalaji olisi mahdollista ennustaa käyttämällä yhtä tai useampaa piirrettä ei ole esitetty. Aiemmin ei ole myöskään tarkasteltu vain piirteitä, jotka ovat helposti kuvista mitattavissa.



## **1.5 Metsähanhien eri alalajien metsästys ja esiintyminen Suomessa metsästyskauden aikana**

Metsähanhen alalajeista vain taigametsähanhia pesii Suomessa, joskin on havaittu myös yksittäisiä tundrametsähanhien pesimistapauksia Käsivarren Lapissa. Syysmuuton aikana molempia lajeja esiintyy Suomessa (Kuva 1). Muuton aikana esiintyy kuitenkin ajallista ja alueellista vaihtelua (Honka ym. 2017, Piironen ym. 2022a, Piironen ym. 2022b). Taigametsähanhia tavataan syysmuuton aikana useimmiten Pohjois- ja Länsi-Suomessa ja tundrametsähanhia Koillis-Suomessa (Piironen ym. 2022a). Suomessa pesivien taigametsähanhien liikkeet muuton aikaan ovat paremmin tiedossa kuin tundrametsähanhien (Nilsson 2011, Honka ym. 2017). Ajallisilla vaihteluilla tarkoitetaan syysmuuton ajankohtaa. Jotta metsästystä on mahdollista rajoittaa taikka kohdistaa toiseen alalajeista, on tärkeää tietää, milloin muutto tapahtuu. Erityisen tärkeäksi tämä tulee ilmastomuutoksen mahdollisesti vaikuttaessa muuttojen ajankohtiin (Piironen ym. 2022b).

Metsähanhi on Suomessa riistaeläin. Metsästyslaki (615/1993) ei tunne metsähanhen alalajeja, eikä sen pohjalta voida säädellä suoraan kummankaan alalajin metsästystä. Sen sijaan metsästystä on pyritty kohdistamaan vuodesta 2017 lähtien tundrametsähanheen rajoittamalla metsähanhen metsästystä alueellisesti Kaakkois-Suomeen ja ajallisesti loppusyksyyn (MMM 2017). Lisäksi vuonna 2021 Lapissa metsähanhea sai metsästää 20–27.8. Lapissa on käytössä yhden linnun kausikiintiö ja molemmilla alueilla metsästäjällä on myös saalisilmoitusvelvollisuus. Alueellisten ja ajallisten rajoitusten tarkoituksena on ollut kohdistaa metsästystä elinvoimaisempaan tundrametsähanheen (MMM 2021).



Kuva 3 Metsähanhia pellolla. Kuva: Veli-Matti Väänänen.

Nopeassa metsästystilanteessa taiga- ja tundrametsähänhea on lähes mahdotonta erottaa toisistaan (Kuva 3). Edes saaliiksi saadun linnun alalajinmääritys ei onnistu ilman runsasta perehtyneisyyttä. Metsähanhisaaliin alalajikoostumuksen tunteminen ja seuraaminen on kuitenkin tärkeää taigametsähänhen suotuisan kannankehityksen turvaamiseksi ja tundrametsähänhen pyynnin mahdollistamiseksi taigametsähänhea vaarantamatta. Geneettiset analyysit saalisnäytteistä ovat luotettavia, mutta näytteiden kerääminen on työlästä ja analyysit kalliita. Saaliista on viime vuosina kerätty valokuvia, joiden perusteella asiantuntijat ovat määrittäneet saaliin alalajikoostumuksen (Piironen ja Alhainen 2019). Menetelmä on kustannustehokas, mutta sen ongelma on arvioiden subjektiivisuus ja erot eri asiantuntijoiden välillä. Lisätutkimukselle on tarvetta, jotta voidaan kehittää kustannustehokas, tarkka ja objektiivinen menetelmä metsähanhisaaliin alalajikoostumuksen seurantaan. Lisäksi menetelmä tulisi dokumentoida avoimesti ja sen luotettavuus tulisi arvioida.

## 1.6 Tutkimuksen tavoitteet

Tässä työssä tutkitaan, voidaanko taiga- ja tundrametsähanhi tunnistaa linnun päästä otetusta, standardisoidusta valokuvasta, ja kuinka luotettavaa valokuvaan perustuva tunnistus on. Lisäksi tutkitaan, mitkä piirteet (muuttujat) metsähanhen päässä parhaiten ennustavat alalajia, sekä arvioidaan metsähanhisaaliin alalajikoostumusta Suomessa, Ruotsissa ja Virossa.

Tarkasteltavat tutkimuskysymykset ovat:

- 1) Voidaanko metsähanhen alalaji tunnistaa linnun päästä otetusta, standardisoidusta valokuvasta? Hypoteesi: Kyllä.
- 2) Mitkä muuttujat metsähanhen päässä ja nokassa ennustavat linnun alalajia?  
Hypoteesi: Tundrametsähanhien nokassa on vähemmän keltaista väriä ja nokka on lyhyempi kuin taigametsähanhilla (Burgers ym. 1991). Lisäksi alanokka on paksumpi taiga- kuin tundrametsähanhella (de Jong 2019).
- 3) Missä suhteessa metsähanhen eri alalajeja on Suomen, Ruotsin ja Viron metsähanhisaaliissa? Hypoteesi: Taigametsähanhia on enemmän Ruotsin metsästyssaaliissa (Nilsson ja Månsson 2012), ja tundrametsähanhia itäisen Suomen ja Viron metsähanhisaaliissa (Honka ym. 2017, Piironen ym. 2022).

Vaikka metsähanhen alalajin määrittystä on aiemmin tutkittu, tämä tutkimus on ensimmäinen, jossa morfologiaa tai pukutuntemerkkejä verrataan geneettisesti määritettyyn alalajiin. Metsähanhen alalajien tunnistaminen on tärkeää hanhien kannanhoidon kannalta, jotta metsähanhisaaliin alalajijakaumaa on mahdollista seurata kustannustehokkaasti, esimerkiksi keräämällä metsästäjiltä valokuvia saaliiksi saaduista linnuista. Tutkimuksen tulosten perusteella voidaan luoda työkalu (tilastotieteellinen malli), jonka avulla alalaji voidaan tunnistaa tunnetulla todennäköisyydellä valokuvasta.

## 2. Aineisto ja menetelmät

### 2.1 Näytteet ja DNA-analyysit

Tutkimuksen aineistona käytetään Ruotsista, Suomesta sekä Virossa saadusta metsähanhisaaliista kerättyjä taiga- ja tundrametsähanhien päitä. Aineiston ovat keränneet Svenska Jägareförbundet, Suomen riistakeskus ja Antti Piironen. Aineisto on kerätty vuosina 2017–2019 ja yhteensä näytteitä kerättiin 901 kappaletta. Aineisto sisältää yksilöitä satunnaisesti molemmista sukupuolista sekä eri ikäluokista. Ruotsista saadun aineiston koko oli 740 näytettä ja Suomesta sekä Virossa kerätyssä aineistossa oli 161 näytettä. Suomessa näytteet on kerätty yksinomaan Kaakkois-Suomesta, jossa tundrametsähanhien osuus on suuri (Honka ym. 2017, Piironen ym. 2022). Valtaosa Ruotsin näytteistä on kerätty vain muutamasta kohteesta, eivätkä näytteet välttämättä ole maantieteellisesti satunnaisotos koko Ruotsin metsähanhisaaliista. Näin ollen näytteet eivät tällöin edusta koko metsähanhisaalista edellä mainituissa maissa. Viron näytteet edustavat tasaisimmin maan metsähanhisaalista. Näytteiden alueellinen painottuminen ei kuitenkaan vaikuta alalajin määrittämiseen.

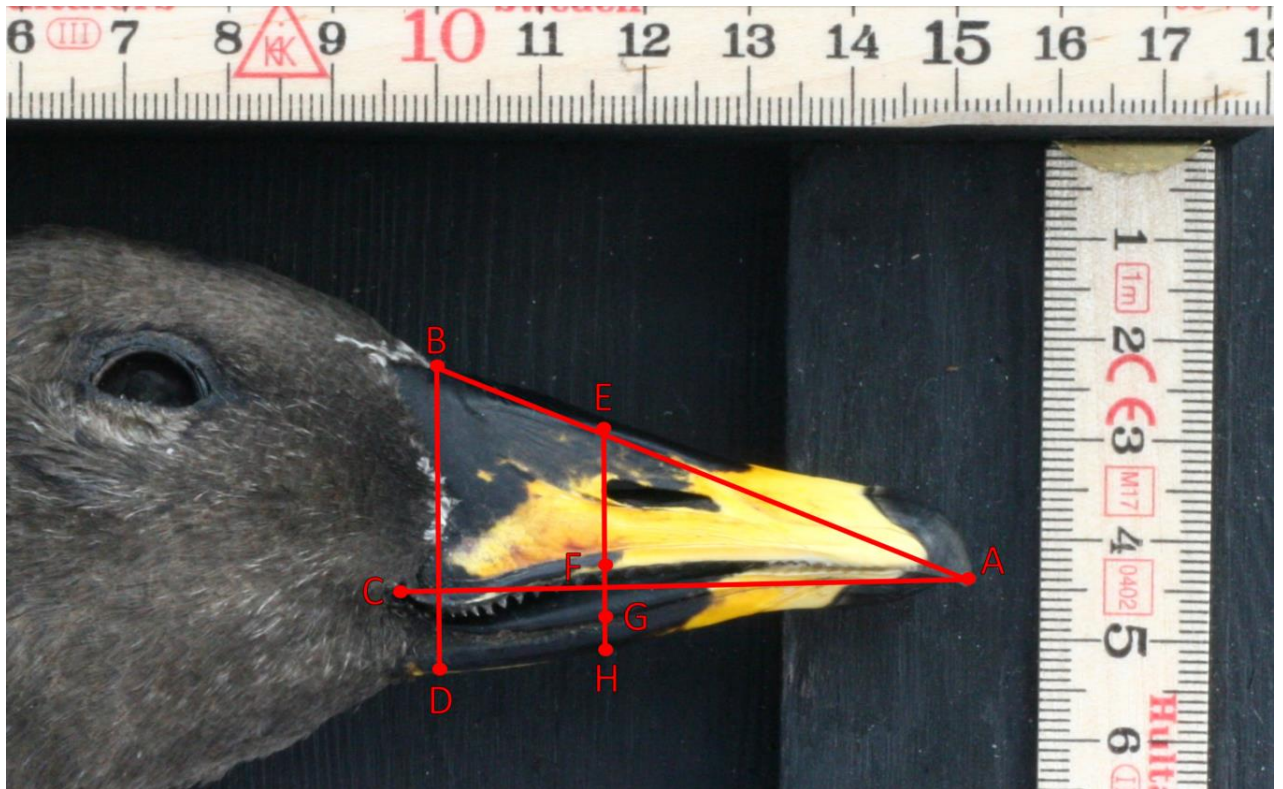
Päistä otettujen kudospätkien (koepala) avulla määritettiin jokaiselle näytteelle alalaji mitokondriaalisen DNA:n avulla. DNA-analyysin teki Johanna Honka Oulun Yliopistossa. Mitokondrio on ainoa soluelin, josta on mahdollista eristää geneettistä materiaalia ja mitokondriosta saadusta DNA:sta on mahdollista määrittää yksilön haplotyyppi. DNA:n eri variaatiot, jotka periytyvät naaraalta jälkeläiselle, määrittelevät yksilön haplotyyppin (Butler 2005). DNA-analyysin perusteella aineistossa oli 524 taigametsähanhea ja 168 tundrametsähanhea (Taulukko 1). Tundrametsähanhista 37 linnun kohdalla yksilöiltä löytyi sekä läntisen (*Anser fabalis rossicus*) että itäisen (*Anser fabalis serrirostris*) haplotyyppiä. Nämä yksilöt yhdistettiin tutkimuksen aineistossa tundrametsähanhiin. Kokonaan ilman DNA-tulosta jäi 203 yksilöä. DNA-analyysin epäonnistumisen syynä saattoi olla esimerkiksi epäonnistunut koepalan säilöntä tai liian pieni koepala.

Taulukko 1 Metsähanhinäytteiden alalaji mitokondrio-DNA:n perusteella.

DNA haplotyyppi	Ruotsi	Suomi	Viro
<i>A. f. fabalis</i>	513	4	7
<i>A. f. rossicus</i>	26	36	106
<i>A. albifrons</i>	1	1	1
<i>A. f. middendorffii</i>	0	1	0
<i>A. brachyrhynchus</i>	3	0	0
Ei DNA-tulosta	197 (n.27 %)	0	6 (n.5 %)
yht.	740	42	120

## 2.2 Näytteiden kuvaaminen ja mittaaminen

Jokainen näyte kuvattiin järjestelmäkameralla (Canon 5D) tummalla alustalla, johon oli kiinnitetty viivoitin mittakaavaksi (Kuva 4.). Kuvassa olevan mittakaavan (viivoittimen) avulla on mahdollista mitata nokan mittoja myös valokuvista. Mittaamiseen käytettiin imageJ-ohjelmistoa (Schneider ym. 2012). Jokaisesta päästä otetut mitat on esitetty kuvassa 4. Kuvista mitattiin nokan korkeus (B–D), nokan pituus kärjestä yläreunaan (A–B), nokan pituus kärjestä suupieleen (A–C), ylänokan (E–F) ja alanokan korkeus (G–H), ylä- ja alanokan väliin jäävä rako eli irvistys (F–G), nokan pinta-ala sekä nokan keltaisen/oranssin värin pinta-ala, ja värillisen alueen osuus koko nokasta (Kuva 4). Mittayksikkönä käytettiin millimetriä, joka pyöristettiin yhden desimaalin tarkkuuteen. Ylä- ja alanokan korkeus otettiin jokaisella linnulla sieraimen vasemman reunan kohdalta, jotta mittaukset ovat helpommin toistettavissa eivätkä niin riippuvaisia mittaajan omasta arvioinnista. Värillisen alueen osuus muunnettiin prosenttiluvuiksi jakamalla värillisen alueen pinta-ala ylänokan alueen pinta-alalla. ImageJ-ohjelmalla ei otettu mittoja pään höyhenpeitteisistä osista, sillä höyhenpeite helposti vääristää mittaustuloksia. Kuvista mittaukset suoritti yksi ja sama henkilö: minä, jotta eri mittaajien välistä mittausvirhettä ei syntyisi.



Kuva 4 Metsähänhinäytteistä otetut mitat. Kuva: Antti Piironen

Hanhen pää on kolmiulotteinen ja epäsymmetrinen objekti, joten valokuvasta otetut mitat eivät ole täsmälleen kohtisuorassa mitattavaan pintaan nähden. Kohtisuoruutta on parannettu asettamalla kuvaustasulle puupala, joka nostaa nokan kärkeä suoristaen päätä. Nokka ei tästä huolimatta ole välttämättä täysin kohtisuorassa, eivätkä mittaukset tällöin täysin vastaa käsin otettuja mittoja. Tämä voi lisätä vaihtelua päistä otettuihin mittoihin, mutta vaihtelun voidaan olettaa olevan satunnaista, ja mittausten keskenään vertailukelpoisia. Ennen tilastollisia analyysejä aineistosta poistettiin epäonnistuneiden kuvien tai kuvassa selkeästi auki olleen nokan takia 28 yksilöä sekä höyhenpuvun perusteella nuoret, alle vuoden vanhat yksilöt, joita oli 194 kappaletta. Nuoret yksilöt poistettiin, koska niiden nokan väritys muuttuu ensimmäisen syksyn aikana (Mouronval ym. 2019). Tilastolliseen analyysiin soveltuvia näytteitä oli 467, joista 341 oli taigametsähänhia ja 126 tundrametsähänhia.

## 2.3 Tilastollinen analyysi ja muuttujien valinta

Metsähanhen alalajin ennustamiseksi nokan mittojen perusteella aineistoon sovitettiin yleistetty lineaarinen malli (*engl. generalised linear model, GLM*), jonka avulla analysoitiin, mitkä muuttujat ennustavat parhaiten selitettävää muuttujaa (metsähanhen alalajia). Aineisto koostui kahdeksasta selittävästä muuttujasta (kaikki nokasta otetut mitat) sekä yhdestä selitettävästä muuttujasta (alalaji, ks. Taulukko 2). Vastemuuttuja (alalaji) noudattaa Bernoulli-jakaumaa ja saa arvoksi joko 0 (tundrametsähanhi) tai 1 (taigametsähanhi). Vastemuuttujan ollessa binäärinen, yleistettyä lineaarista mallia kutsutaan logistiseksi regressioksi, jossa käytettiin logit-linkkifunktiota.

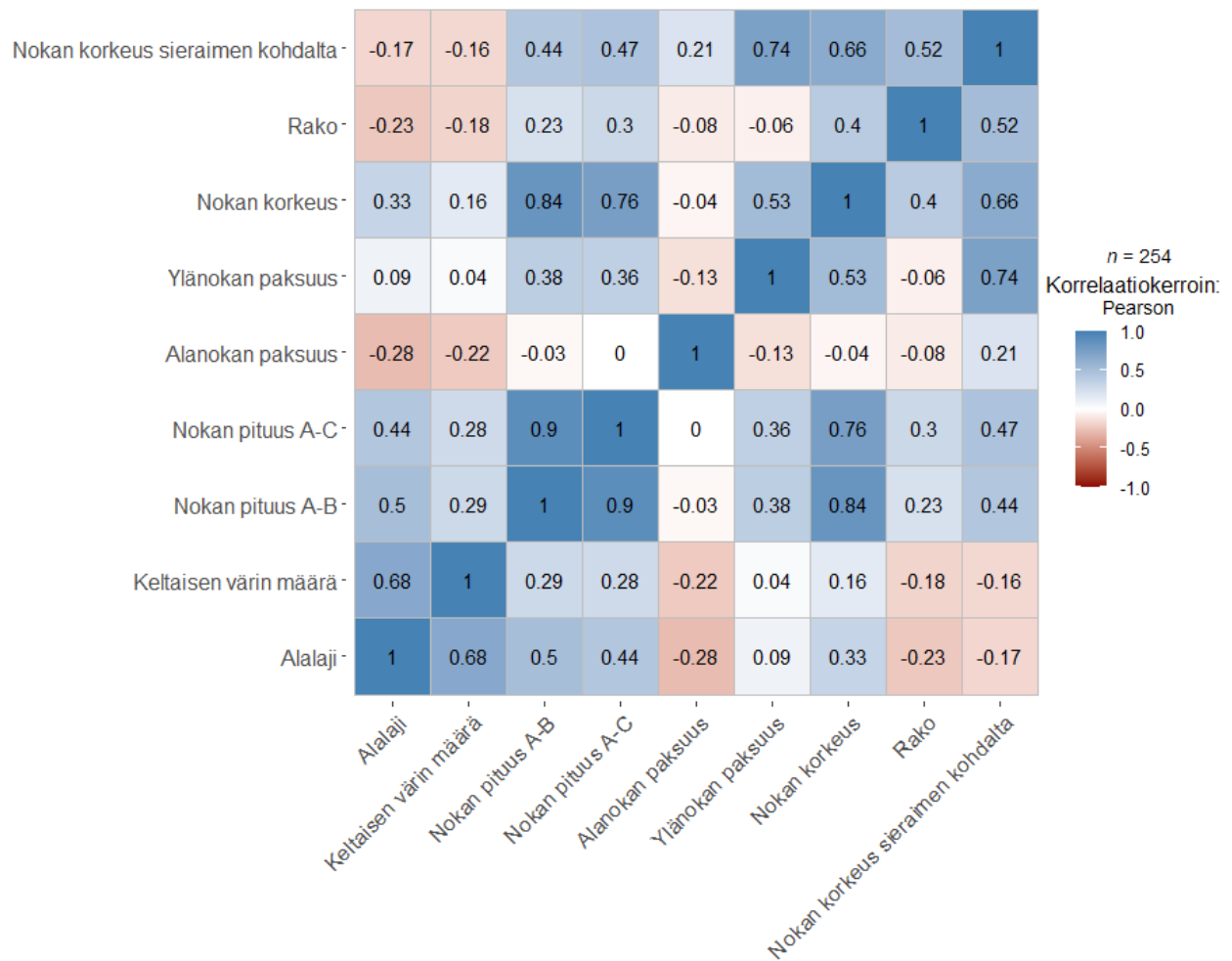
Taulukko 2 Mahdolliset selittävät muuttujat

---

Muuttujat:
Keltaisen värin määrä
Nokan pituus A-B
Nokan pituus A-C
Nokan korkeus
Nokan korkeus sieraimen kohdalta
Ylänokan paksuus
Alanokan paksuus
Irvistys eli ylä- ja alanokan väliin jäävä rako

---

Korrelogrammin avulla on mahdollista tarkastella eri muuttujien välisiä korrelaatioita sekä yksittäisen muuttujan korrelaatiota selitettävään muuttujaan. Ennen mallinnusta karsittiin pois selittävät muuttujat, joiden korrelaatio (Pearsonin korrelaatiokertoimen arvo) muiden selittävien muuttujien kanssa oli yli 0,6 (ks. Kuva 5). Keskenään korreloivista muuttujista aina toinen jätettiin siis pois mallinnuksesta. Mallinnukseen valikoitui neljä muuttujaa: keltaisen määrä nokassa, alanokan paksuus, nokan pituus A-B ja ylänokan paksuus.



Kuva 5 Kaikki tarkasteltavat muuttujat sisältävä korrelogrammi, jossa käytetty Pearsonin korrelaatiokerrointa.

Näytteiden lukumäärän ollessa epätasainen alalajien välillä valittiin satunnaisella otannalla taigametsähänhiyksilöitä saman verran kuin tundrametsähänhia. Lopulta aineisto koostui 126 taigametsähänhesta ja 126 tundrametsähänhesta.



### **2.3.1 Mallin valinta ja ennustetarkkuuden testaus**

Ennen mallien sovittamista aineisto jaettiin opetusaineistoon (80 % koko aineistosta), jota käytettiin mallin parametrien estimointiin, sekä testiaineistoon (20 % koko aineistosta), jota käytettiin mallin ennustetarkkuuden testaukseen. Aineistoon sovitettiin mallit kaikilla muuttujien yhdistelmillä. Mallin valinnassa käytettiin Akaiken informaatiokriteeriä (*engl. Akaike's information criteria*, Akaike 1973) ja parhaaksi malliksi valittiin malli, jonka pienille otoksille korjatun informaatiokriteerin (AICc) arvo oli pienin. Informaatioteoreettinen mallinvalintamenetelmä tasapainottaa mallin ennustekykä ja yksinkertaisuutta (Burnham ja Anderson 2002). Mallin ennustetarkkuutta testattiin ennustamalla mallilla testiaineistoa, ja tarkastelemalla luokitteluvirhettä.

Kaikki analyysit tehtiin R-ohjelmointiympäristössä (R Core Team 2021) käyttäen glmulti (Calgagno 2020), jtools (Long J.A. 2020), ggfortify (Horikoshi ja Tang 2016), devtools (Wickham ym. 2021), caret (Kuhn 2022), InformationValue (Prabhakaran 2016) ja ISLR (James ym. 2021) kirjastoja.

## **3. Tulokset**

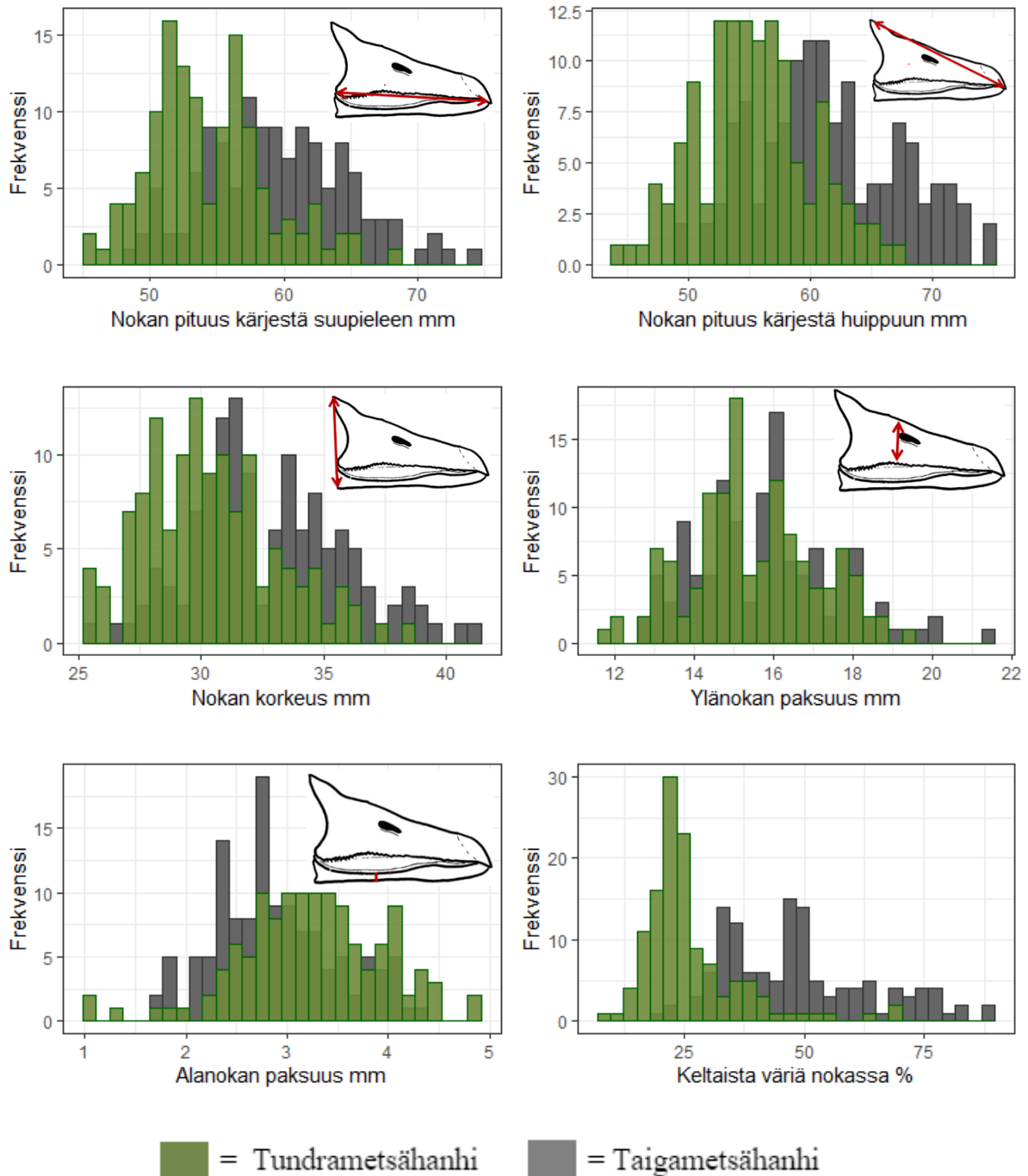
### **3.1 Mallinnus**

Neljää muuttujaa tarkasteltiin glmulti-pakettia hyödyntäen ja paketin avulla aineistoon sovitetun yleistetyn lineaarisen mallinnuksen perusteella parhaaksi alalajia selittäväksi malliksi nousi kolmen muuttujan malli, joka sisälsi muuttujat: keltaisen värin määrä nokassa, nokan pituus A-B ja alanokan paksuus (Taulukko 5), vain ylänokan paksuus jäi pois. Ensimmäiseksi ja toiseksi tulleen mallien välillä AICc-arvojen ero oli alle 2, mikä tarkoittaa mallien välisen eron olevan merkityksetön. Parhaaksi malliksi valittiin siis neljän muuttujan malli: keltaisen määrä nokassa - alanokan paksuus - nokan pituus A-B - ylänokan paksuus. Mallin AICc-arvo on 134,31.

Taulukko 5. Aineistoon sovitettujen mallien AIC, AICC sekä AIC painoarvot.

Malli	AIC	AICC	Painot
Keltaisen määrä nokassa - Alanokan paksuus - Nokan pituus A-B	139,37	135,35	3,7
Keltaisen määrä nokassa - Alanokan paksuus - Nokan pituus A-B - Ylänokan paksuus	140,43	134,31	6,2
Keltaisen määrä nokassa - Nokan pituus A-B	150,08	145,22	2,7
Keltaisen määrä nokassa - Nokan pituus A-B - Ylänokan paksuus	151,99	146,06	1,7
Keltaisen määrä nokassa - Alanokan paksuus - Ylänokan paksuus	178,87	164,93	1,4
Keltaisen värin määrä - Alanokan paksuus	179,09	162,92	3,8
Keltaisen värin määrä - Ylänokan paksuus	184,15	171,65	6,1
Keltaisen värin määrä	185,81	169,61	1,35
Alanokan paksuus - Nokan pituus A-B	216,47	217,25	6,2
Alanokan paksuus - Ylänokan paksuus - Nokan pituus A-B	216,88	221,83	6,1
Nokan pituus A-B	232,51	241,27	3,7
Nokan pituus A-B - Ylänokan paksuus	234,02	240,14	6,5
Alanokan paksuus - Ylänokan paksuus	277,89	276,16	9,9
Alanokan paksuus	278,07	274,10	2,6
Ylänokan paksuus	289,89	293,29	1,9

Kaikkien mallinnukseen valittujen muuttujien frekvenssijakaumat (histogrammit) on esitetty alalajeittain kuvassa 6. Alalajien nokkien mitoissa on paljon päällekkäisyyttä, mutta selkeimmät erot näytävät histogrammien perusteella olevan keltaisen värin määrässä, alanokan paksuudessa ja nokan pituudessa.



Kuva 6 Tundrametsähanhen ja taigametsähanhen nokan mittoja esitettyä histogrammeissa.

Muuttujien tilastollista vaikutusta on mahdollista tarkastella p-arvojen avulla. Selittävät muuttujat ovat tilastollisesti merkitseviä p-arvon ollessa  $<0,5$  ja erittäin merkitseviä p-arvon ollessa  $<0,001$ . Tutkimuksissa yleensä käytetään 5 %:n riskitasoa, mikä tarkoittaa tuloksien olevan 95 % varmuudella pätevä ja sisältävän 5 % virheen mahdollisuuden. Kaikista muuttujista vain taulukossa 3 esitetyt muuttujat olivat tilastollisesti merkitseviä. Vaikka muuttuja ylänokan paksuus ei ole tilastollisesti merkitsevää, otettiin se mukaan malliin, sen tullessa joka tapauksessa mitatuksi.

Taulukko 3 Valittujen muuttujien estimaatti, keskivirhe sekä z- ja p-arvot.

Muuttuja	Estimaatti	Keskivirhe	z-arvo	p-arvo
(vakiotermi)	-11,911	3,067	-3,883	0,001
Keltaisen värin määrä nokassa	13,366	2,273	6,32	2,61E-10*
Nokan pituus A-B	0,465	0,102	4,538	5,67E-06*
Alanokan paksuus	-1,534	0,361	-4,245	2,18E-05*
Ylänokan paksuus	-0,076	0,162	-0,471	0,638

Parhaan mallin ennustetarkkuus testiaineistossa (20 % koko aineistosta) on havainnollistettu kuvan 6 sekaannusmatriisissa. Sekaannusmatriisi kertoo ennustavan mallin tarkkuudesta, mutta näyttää myös mahdolliset erot ennustettavien luokkien välillä. Testiaineistossa oli 23 yksilöä molemmista alalajeista. Valitun mallin luokitteluvirhe oli 0,0667 eli malli luokitteli alalajit oikein noin 93 % varmuudella. Valittu malli ennusti tundrametsähanhista 21 yksilöä oikein ja 3 yksilöä väärin. Taigametsähanhista 19 yksilöä oikein ja 2 yksilöä väärin, joten mallin tekemä virhe näyttää olevan yhtä suuri molempien alalajien kohdalla.

	Tundrametsähanhi (Tosi)	Taigametsähanhi (Tosi)
Tundrametsähanhi (Mallin ennustama)	21	2
Taigametsähanhi (Mallin ennustama)	3	19

Kuva 7 Valitun mallin sekaannusmatriisi. Vihreät ruudut = malli ennusti alalajin oikein, punaiset ruudut = malli ennusti alalajin väärin. Malli ennusti metsähanhen alalajin oikein 93 % tapauksista.

### 3.2 Metsähanhisaaliin alalajijakauma eri maissa

Koko aineiston alalajijakauma eri maissa on esitetty taulukossa 6. Ruotsissa metsästyskauden 2017–2018 aikana metsästetyistä näytteistä 4,8 % oli tundrametsähanhia ja 95,2 % taigametsähanhia. Suomen ja Viron näytteet ovat osa 2017–2018 ja 2018–2019 metsästyskausien metsähanhisaalista. Suomessa näytteistä 90 % oli tundrametsähanhia ja Virossa 93,8 %. Taigametsähanhia Suomen näytteistä oli 10 % ja Virossa 6,2 %.

Taulukko 6. Metsähanhisaaliista kerättyjen näytteiden alalajijakauma Ruotsissa, Suomessa ja Virossa metsästyskausilla 2017–18 ja 2018–19.

Maa \ Alalaji	Tundrametsähanhi	Taigametsähanhi
Ruotsi	26 (4,8 %)	513 (95,2 %)
Suomi	36 (90 %)	4 (10 %)
Viro	106 (93,8 %)	7 (6,2 %)
yht.	168 (24,3 %)	524 (75,5 %)

## **4. Tulosten tarkastelu**

### **4.1 Metsähanhen alalajin määrittäminen**

Tuloksien perusteella voidaan todeta olevan mahdollista tunnistaa metsähanhen alalaji linnun päästä otetusta standardisoidusta kuvasta 93 % tarkkuudella (Kuva 6). De Jong 2019 tutkimuksessaan onnistui mallintamaan kahden muuttujan, hampaiden lukumäärän ja alanokan paksuuden, avulla alalajin oikein 99,6 % tarkkuudella. De Jongin tutkimuksen alalajien määrittäminen perustuu yhden asiantuntijan arvioihin ei DNA-tuloksiin. Hampaiden lukumäärän voidaan todeta olevan huono alalajit erottava muuttuja, hampaiden lukumäärän vaihdellessa eri mittajien välillä. De Jongin tutkimus altistuikin kehäpäätelemille, tutkimuksen perustuessa arvioihin ja tuloksien tukiessa näitä ennalta sovittuja arvioita. Tämän tutkimuksen tuloksien tarkkuus riittää hyvin kannanhoidon tarpeisiin ja mahdollistaa tulevaisuudessa asiantuntija-arviosta riippumattoman metsähanhisaaliin alalajijakaumien vertailun. Tässä tutkimuksessa käytettyä menetelmää ei ole aiemmin sovellettu vesilinnuilla, mutta menetelmän voidaan todeta soveltuvan erittäin hyvin alalajien tunnistusta käsittelevään tutkimukseen.

### **4.2 Ennustavat muuttujat**

Tuloksia on vaikeaa verrata aiempiin tutkimuksiin, sillä vastaavia aiempia tutkimuksia ei metsähanhen alalajeille ole tehty ja niihin täytyy suhtautua kriittisesti. Kuitenkin, kuten aiemmissa tutkimuksissa (Cramp ja Simmons 1977, Burgers ym. 1991) keltaisen värin määrä nokassa on kuitenkin selkein erottava tekijä alalajien välillä, ja taigametsähanhilla oli selkeästi enemmän keltaista väriä nokassaan verrattuna tundrametsähanheen. Taigametsähanhien nokat olivat myös tässäkin tutkimuksessa keskimäärin pidempiä kuin tundrametsähanhien (Cramp ja Simmons 1977, Burgers ym. 1991). Alanokan paksuus, joka aiemmin on todettu eroavan alalajien välillä (de Jong 2019). Niinpä muuttujat ala- ja ylänokan paksuus, nokan pituus A-B sekä keltaisen värin määrä nokassa muodostivat hyvän mallin ennustamaan alalajia. Jokainen muuttujista on helposti ja nopeasti mitattavissa valokuvista, mikä mahdollistaa mallin käytön alalajin tunnistuksessa

suuremmallakin mittakaavalla ilman DNA-analyysyjä. Mutta kuten aiemmin todettiin, tämän tutkimuksen mittauksia ei voida suoraan vertailla aiempiin taksonomisiin tutkimuksiin; tämän tutkimuksen mittojen ollessa kuvista eikä fyysisestä linnusta mitattuja ja alalajin määrittämisen perustuessa DNA-analyysiin eikä asiantuntija-arvioihin.

### **4.3 Alalajijakauma**

Nilssonin ja Månssonin tutkimuksessa (2012), Ruotsin metsähanhisaaliista 95 % oli taigametsähanhia (Taulukko 6). Sen sijaan Suomen metsähanhisaalista 90 % ja Viron saaliista 93,8 % oli määritetty visuaalisten tuntomerkkien perusteella tundrametsähanhiksi (Honka ym. 2017, Piironen ym. 2022a). Ruotsissa tehdyt tutkimukset eivät kuitenkaan perustu DNA-testaukseen, eivätkä ole tällöin suoraan vertailukelpoisia tähän tutkimukseen. Viron näytteet on kerätty muunnaikaiselta levähdysalueelta ja edustavat tällöin kohtuullisen hyvin alueen kautta muuttavaa alalajijakaumaa. Suomen ja Ruotsin näytteiden kohdalla metsästys on painottunut pienelle alueelle, eivätkä ne välttämättä kuvaa maan koko alueella esiintyvää alalajijakaumaa. Tutkimuksen tulos ei siis edusta maantieteellisesti täydellisesti näiden maiden alalajijakaumaa, mutta todistaa ainakin sen, että molempia alalajeja esiintyy jokaisen valtion mailla. Näytteiden keräämisen ajankohta on mahdollisesti vaikuttanut alalajijakaumaan erityisesti Suomessa, jossa myöhäinen metsästyskausi kohdistaa metsästystä enemmän tundrametsähanheen kuin aiemmin syksyllä muuttavaan taigametsähanheen.

## **5. Johtopäätökset**

Metsähanhen kaksi Suomessa esiintyvää alalajia taigametsähanhi ja tundrametsähanhi oli mahdollista tunnistaa toisistaan valokuvasta otettujen mittojen perusteella. Tulosten määritystarkkuus on hyvä ja luotettava. Erityisesti hyötyä tuloksista on Suomen riistahallinnolle Suomen metsähanhisaaliin alalajikoostumuksen seurannassa, mutta tuloksia voidaan käyttää myös laajemmin Euroopassa. Tutkimuksen tuloksia voidaan soveltaa metsähanhisaaliin alalajijakauman seuraamiseen, sekä mahdolliseen alalajien saalisosuuksien säätelyyn, keräämällä muun muassa metsästäjältä kuvia saaliiksi jääneistä linnuista. Kuvista alalajin tunnistaminen tunnetulla todennäköisyydellä on edullisempaa ja nopeampaa kuin DNA-analyysien vuosittainen teettäminen, eikä sisällä samanlaista virheen mahdollisuutta kuin asiantuntija-arviot. Alalajien erottelua tarvitaan myös tarkemmassa kannankoon arvioinnissa, jotta arviot eivät vääristy suuntaan tai toiseen. Alalajin määrittäminen saaliin alalajijakaumassa on tärkeää myös metsästyksen kohdistamisessa niin ajallisesti kuin alueellisestikin enemmän tundrametsähanheen kuin taigametsähanheen. Tutkimuksen tulokset ovat jatkossa tärkeä osa metsähanhien metsästyksen suunnittelua.

## **6. Kiitokset**

Suuret kiitokset Oulun yliopiston Johanna Hongalle kallisarvoisista DNA-analyyseistä ja Suomen ja Ruotsin riistakeskuksille kiitokset tutkimuksen aineistosta. Erityiset kiitokset ohjauksesta Veli-Matti Väänäselle sekä isoimmat kiitokset Antti Piroselle aineiston keruusta, tärkeän aiheen jakamisesta sekä erinomaisesta ohjauksesta. Kiitokset myös Suomen Riistanhoito-Säätiölle pro gradun stipendistä.



## 7. Lähteet

Akaike, H. 1973. Information Theory and an Extension of the Maximum Likelihood Principle. *Breakthroughs in Statistics*, Springer New York, s. 610–24, [https://doi.org/10.1007/978-1-4612-0919-5\\_38](https://doi.org/10.1007/978-1-4612-0919-5_38).

Butler, J.M. 2005. *Forensic DNA typing: biology, technology, and genetics of STR markers*. 2. painos. Elsevier Academic Press, Amsterdam. 302 s.

Burgers, J., Smit, J.J. & van der Voet, H. 1991. Origins and systematics of two types of the Bean Goose *Anser fabalis* (Latham, 1787) wintering in The Netherlands. *Ardea* 79: 307–16.

Burnham, K. P. & Anderson, D. R. 2002. *Model Selection and Multimodel Inference*. Toinen painos. Springer. New York. 514 s.

Calgagno, V. 2020. Glmulti: Model selection and multimodel interference made easy. Versio 1.0.8. Saatavissa: <https://cran.r-project.org/web/packages/glmulti/index.html>

Cramp, S. & Simmons, K. E. L. 1977. *Handbook of the Birds of Europe, the Middle East and North Africa. The Birds of the Western Palearctic. Vol 1, Ostrich to Ducks*. 1. painos. Oxford University Press, New York. 722 s.

De Jong, A. 2019. Less Is Better. Avoiding Redundant Measurements in Studies on Wild Birds in Accordance to the Principles of the 3Rs. *Frontiers in Veterinary Science* 6: 195–195. <https://doi.org/10.3389/fvets.2019.00195>

Delacour, J. 1951. Taxonomic notes on the bean goose *Anser fabalis* Lath. *Ardea* 39: 135–42. Saatavissa: <https://ardea.nou.nu/ardeapdf/a39-135-142.pdf>

Delany, S. & Scott, D. 2006. *Waterbird population estimates*. Wetlands International, Wagenigen, The Netherlands. Saatavissa: <https://www.wetlands.org/publications/waterbird-populations-estimates-fifth-edition/>

Dementiev, G.P. & Gladkov, N.A. 1967. *Birds of the Soviet Union. Vol IV*. Israel Program for Scientific Translations, Jerusalem. 683 s.

Fox, A.D. & Leafloor, J.O. 2018. A global audit of the status and trends of Arctic and Northern Hemisphere goose populations. *Conservation of Arctic Flora and Fauna*

International Secretariat. Akureyri, Iceland. 32 s. Saatavissa:  
<http://hdl.handle.net/11374/2364>

Fox, A.D., Ebbinge, B.S., Mitchell, C., Heinicke, T., Aarvak, T., Colhoun, K. Clausen, P., Dereliev, S., Faragó, S., Koffijberg, K., Kruckenberg, H., Loonen, M. J. J. E., Madsen, J., Mooij, J., Musil, P., Nilsson, L., Pihl, S. & van der Jeugd, H. 2010. Current estimates of goose population sizes in western Europe, a gap analysis and an assessment of trends. *Ornis Svecica* 20(3–4): 115–127. <https://doi.org/10.34080/os.v20.19922>

Heinicke, T. 2010. Tundra bean goose *Anser fabalis rossicus* during spring migration in northern Sweden – Rare visitor or regular passage migrant? *Ornis Svecica* 20: 174–183. <https://doi.org/10.34080/os.v20.22618>

Heinicke, T. 2009. Status of the Bean Goose *Anser fabalis* wintering in central Asia. *Wildfowl* 59: 77–99. <https://wildfowl.wwt.org.uk/index.php/wildfowl/article/view/1215>

Heinicke, T. 2018. A2 Western Tundra Bean Goose *Anser fabalis rossicus*. *Julkaisussa: Fox, A.D. & Leafloor, J.O. 2018. A global audit of the status and trends of Arctic and northern hemisphere goose populations. Conservation of Arctic Flora and Fauna International Secretariat. Akureyri, Iceland. Saatavissa: https://pure.au.dk/portal/files/143585881/Goose\_report\_p2\_Final\_Feb1.pdf*

Heinicke, T. & de Jong, A. 2013. Tundra bean geese *Anser fabalis rossicus* in central and southern Sweden autumn 2009–spring 2012. *Ornis Norvegica* 36(1): 32–37. Saatavissa: <https://boap.uib.no/index.php/ornis/article/view/438>

Heinicke, T., Fox, A.D. & de Jong, A. 2018. A1 Western Taiga Bean Goose *Anser fabalis fabalis*. *Julkaisussa: Fox, A.D. & Leafloor, J.O. 2018. A global audit of the status and trends of Arctic and northern hemisphere goose populations. Conservation of Arctic Flora and Fauna International Secretariat. Akureyri, Iceland. Saatavissa: https://pure.au.dk/portal/files/143585881/Goose\_report\_p2\_Final\_Feb1.pdf*

Hirschfeld A. & Heyd A. 2005. Mortality of migratory birds caused by hunting in Europe: bag statistics and proposals for the conservation of birds and animal welfare. *Ber zum Vogelsch* 42: 47–74. Saatavissa: [https://www.komitee.de/media/mortality\\_of\\_migratory\\_birds\\_caused\\_by\\_hunting\\_in\\_europe.pdf](https://www.komitee.de/media/mortality_of_migratory_birds_caused_by_hunting_in_europe.pdf)

- Honka, J., Kvist, L., Heikkinen, M. E., Helle, P., Searle, J.B., & Aspi, J. 2017. Determining the subspecies composition of bean goose harvests in Finland using genetic methods. *European Journal of Wildlife Research* 63(1): 1–14. <https://doi.org/10.1007/s10344-017-1077-6>
- Horikoshi, M., Tang, Y. 2018. Ggfortify: data visualization tools for statistical analysis results. Versio 0.4.14. Saatavissa: <https://CRAN.R-project.org/package=ggfortify>.
- James, G., Witten, D., Hastie, T., Tibshirani, R. 2021. ISLR: Data for an introduction to statistical learning with applications in R. Versio 1.4. Saatavissa: <https://cran.r-project.org/web/packages/ISLR/index.html>
- Johansen, H. 1945. About races of the bean goose. *Dansk Ornitologisk Foreningens Tidsskrift* 39(1): 106–127.
- Johnson, F.A., Jensen, G. & Heldbjerg, H. 2020. Harvest assessment for taiga bean geese in the central management unit. AEWAs European goose management platform data centre, Aarhus University, Denmark. 12 s.
- Kuhn, M. 2022. caret: Classification and regression training. Versio 6.0-92. Saatavissa: <https://cran.r-project.org/web/packages/caret/index.html>
- Long, J.A. 2020. jtools: Analysis and Presentation of Social Scientific Data. Versio 2.1.0, Saatavissa: <https://cran.r-project.org/package=jtools>.
- Mouronval, J-B, Väänänen, V-M., Piironen, A. & Alhainen, M. 2019. Vesilintujen iän ja sukupuolen määrittäminen. Suomen riistakeskus, Helsinki. 214 s.
- Nilsson, L. 2011. The migrations of Finnish Bean Geese *Anser fabalis* in 1978–2011. *Ornis Svecica* 21(1): 157–166. <https://doi.org/10.34080/os.v21.22599>
- Nilsson, L. 2008. Recent changes in numbers and distribution of the Taiga Bean Goose *Anser fabalis fabalis* in south Sweden during 1977–2005. *Die Vogelwelt* 129(3): 263–267. Saatavissa: [http://www.zoekol.biol.lu.se/waterfowl/GooseInv/Goose-Publ/2008a-VOGELWELT-%20129-%20263\\_%20267.pdf](http://www.zoekol.biol.lu.se/waterfowl/GooseInv/Goose-Publ/2008a-VOGELWELT-%20129-%20263_%20267.pdf)
- Nilsson, L., van den Bergh, L. & Madsen, J. 1999. Taiga Bean Goose *Anser fabalis fabalis*. Julkaisussa: Madsen, J., Cracknell, G. & Fox, T. (toim.). Goose populations of the Western Palearctic. A review of the status and distribution. Wetlands International, Wageningen, Netherlands. s. 20–36.

Nilsson, L. & Månsson, J. 2012. Inventeringar av sjöfåglar, gäss och tranor i Sverige. Årsrapport 2011/12. Biologiska institutet, Lund universitet, Sverige. 55 s. Saatavissa: [http://www.zoo.ekol.lu.se/waterfowl/andinv/ANDRAP/Waterfowl\\_Report\\_12-13.pdf](http://www.zoo.ekol.lu.se/waterfowl/andinv/ANDRAP/Waterfowl_Report_12-13.pdf)

Ottenburghs, J., Megens, H.-J., Kraus, R.H., Madsen, O., van Hooft, P., van Wieren, S. E., Crooijmans, R.P.M., Ydenberg, R.C., Groenen, M.A., & Prins, H.H. 2016. A tree of geese: A phylogenomic perspective on the evolutionary history of True Geese. *Molecular Phylogenetics and Evolution* 101: 303–313. <https://doi.org/10.1016/j.ympev.2016.05.021>

Ottenburghs, J., Megens, H. J., Kraus, R. H., Hooft, van, Wieren, van, Crooijmans, R. P.M., Ydenberg, R.C., Groenen, M.A., & Prins, H.H. 2017. A history of hybrids? Genomic patterns of introgression in the True Geese. *Evolutionary Biology* 17(1): 201–201. <https://doi.org/10.1186/s12862-017-1048-2>

Ottenburghs J., Honka, J., Müskens, G.J.D., & Ellegren, H. 2020. Recent introgression between Taiga Bean Goose and Tundra Bean Goose results in a largely homogeneous landscape of genetic differentiation. *Heredity* 125(1–2): 73–84. <https://doi.org/10.1038/s41437-020-0322-z>

Piironen, A. & Alhainen, M. 2019. Tundrametsähanhen metsästysalue laajenee. *Metsästäjä* 5: 12–13. Saatavissa: <https://www.lehtiluukku.fi/lehti/metsastaja/05-2019/221820.html>

Piironen, A., Paasivaara, A. & Laaksonen, T. 2021. Birds of three worlds: Moulting migration expands a boreal-temperate flyway to a third biome. *Movement Ecology* 9(47): 1–47. <https://doi.org/10.1186/s40462-021-00284-4>

Piironen, A. & Laaksonen, T. 2022a. Predicting spatio-temporal distributions of migratory populations using Gaussian process modelling. *Journal of Applied Ecology* 59(4): 1146–1156. <https://doi.org/10.1111/1365-2664.14127>

Piironen, A., Fox, A.D., Kampe-Persson, H., Skjellberg, U., Therkildsen, O.R. & Laaksonen, T. 2022b. When and where to count? Implications of migratory connectivity and non-breeding distribution to population censuses in a migratory bird population. <https://doi.org/10.1101/2022.02.26.482079>

Prabhakaran, S. 2016. InformationValue: Performance analysis and companion functions for binary classification models. Versio 1.2.3. Saatavissa: <https://cran.r-project.org/web/packages/InformationValue/index.html>

MMM, 2017. Maa- ja metsätalousministeriön asetus metsähanhen metsästyksen kieltämisestä metsästysvuosina 2017–2019. Maa- ja metsätalousministeriön asetus 1209/2017. Saatavissa: <https://mmm.fi/documents/1410837/1822248/Muistio-MMMa-mets%C3%A4hanhen-mets-kielto-metsvuosina+2017-2019.pdf/842e8bee-ba1f-4f3b-aabd-f25352510a08/Muistio-MMMa-mets%C3%A4hanhen-mets-kielto-metsvuosina+2017-2019.pdf?t=1502092129000>

MMM, 2021. Maa- ja metsätalousministeriön asetus metsähanhen metsästyksen kieltämisestä metsästysvuonna 2021–2022. Maa- ja metsätalousministeriön asetus 593/2021, Saatavissa: <https://www.finlex.fi/fi/laki/alkup/2021/20210593>

R Core Team, 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>

Ruokonen, M., Litvin, K. & Aarvak, T. 2008. Taxonomy of the bean goose–pink-footed goose. *Molecular Phylogenetics and Evolution* 48(2): 554–562. <https://doi.org/10.5253/078.099.0112>

Ruokonen, M. & Aarvak, T. 2011. Typology revisited: historical taxa of the bean goose–pink-footed goose complex. *Ardea* 99(1): 103–112. <https://doi.org/10.5253/078.099.0112>

Sangster, G. & Oreeel, G.J. 1996. Progress in taxonomy of Taiga and Tundra Bean Geese. *Dutch Birding* 18(6): 310–316. Saatavissa: [https://www.researchgate.net/publication/292709059\\_Progress\\_in\\_taxonomy\\_of\\_Taiga\\_and\\_Tundra\\_Bean\\_Geese](https://www.researchgate.net/publication/292709059_Progress_in_taxonomy_of_Taiga_and_Tundra_Bean_Geese)

Schneider, C.A., Rasband, W.S. & Eliceiri, K.W. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* 9(7): 671–675. <http://doi:10.1038/nmeth.2089>

Van den Bergh, L. 1999. Tundra Bean Goose *Anser fabalis rossicus*. Julkaisussa: Madsen, J., Cracknell, G. & Fox, T. (toim.). *Goose populations of the Western Palearctic. A review of the status and distribution*. Wetlands International, Wageningen, The Netherlands. s. 38–66. <https://doi.org/10.1093/auk/117.1.271>

Vaurie, C. 1965. *The Birds of the Palearctic Fauna. Non-Passeriformes*. H.F. & G., Witherby, London. 763 s.

Wickham, H., Hester, J., Chang, W., Bryan, J. 2021. Devtools: tools to make developing r packages easier. Versio 2.4.3. Saatavissa: <https://cran.r-project.org/web/packages/devtools/index.html>

## 8. Liitteet:

### 8.1 Käytetty malli

```
#Metsähanhen alalaji R
# Tyhjennä muisti ja aloita
rm(list = ls())
setwd("~/R/GEESE")

#Valitse tarvittavat paketit
library("devtools")
library("ggfortify")
library("jtools")
library("caret")
library("InformationValue")
library("ISLR")
library("glmulti")

# Lataa data ja tyhjää tyhjät ruudut
data <- read.table("graduaineisto-metsähanhi-erika-jumppanen-2021.txt", header =
TRUE, sep = ",", dec = ".")
data <- data[which(complete.cases(data$bill_col_ratio) &
complete.cases(data$culmen_length_tip_top) &
complete.cases(data$gap_rel) & complete.cases(data$bill_height) &
complete.cases(data$low_man_height)), ]
nrow(data[which(data$subspecies == "fabalis"), ]) #316
nrow(data[which(data$subspecies == "rossicus"), ]) #126

# jotta datassa fabaliksia yhtä paljon kuin rossicuksia
```

```

a <- data[which(data$subspecies == "fabalis"), ][sample(x =
nrow(data[which(data$subspecies == "fabalis"), ]),
size = nrow(data[which(data$subspecies ==
"rossicus"), ]), replace = FALSE), ]
b <- data[which(data$subspecies == "rossicus"), ]

data <- rbind(a, b)

# Korrelogrammi
library("ggstatsplot") #ggstatsplot::ggcorrmat(
library("ggcorrplot")

ggstatsplot::ggcorrmat(
  data = data[c("subspecies_bin",
               "bill_col_ratio",
               "culmen_length_tip_top",
               "bill_gape_bill_length_tip_corner",
               "low_man_height",
               "up_man_height",
               "bill_height",
               "gap",
               "height_where_meas")],
  cor.vars.names = c("Alalaji", "Keltaisen värin määrä", "Nokan pituus A",
                    "Nokan pituus B", "Alanokan paksuus",
                    "Ylänokan paksuus", "Nokan korkeus", "Rako", "Nokan korkeus
sieraimen kohdalta"),
  type = "parmetric", # parametric for Pearson, nonparametric for
Spearman's correlation
  ggcorrplot.args = list(method = "square"),
  matrix.type = "full",
  sig.level = 1,
  tl.col = c("black"),
  colors = c("darkred", "white", "steelblue"),
  caption = c("x = ei merkitsevä kun p < 0,5")
)

```

```

names(data)

# Luo fabalis & rossicus -muuttujat

which(data$subspecies == "fabalis")
data[which(data$subspecies == "fabalis"),]
fab=data[which(data$subspecies == "fabalis"),]
ros=data[which(data$subspecies == "rossicus"),]

#ylläolevilla muuttujilla, histogrammit

a=ggplot() +
  geom_histogram(data=fab, aes(x =
bill_gape_bill_length_tip_corner),col="gray20", fill="gray40", alpha=1.0, bins =
30)+
  geom_histogram(data=ros, aes(x =
bill_gape_bill_length_tip_corner),col="darkgreen", fill="darkolivegreen4",
alpha=0.9, bins = 30)+
  theme_bw()+
  labs(title="", y="Frekvenssi", x="Nokan pituus kärjestä suupieleen mm")

b=ggplot() +
  geom_histogram(data=fab, aes(x = culmen_length_tip_top),col="gray20",
fill="gray40", alpha=1.0, bins = 30)+
  geom_histogram(data=ros, aes(x = culmen_length_tip_top),col="darkgreen",
fill="darkolivegreen4", alpha=0.9, bins = 30)+
  theme_bw()+
  labs(title="", y="Frekvenssi", x="Nokan pituus kärjestä huippuun mm")

c=ggplot() +
  geom_histogram(data=fab, aes(x = bill_height),col="gray20", fill="gray40",
alpha=1.0, bins = 30)+
  geom_histogram(data=ros, aes(x = bill_height),col="darkgreen",
fill="darkolivegreen4", alpha=0.9, bins = 30)+
  theme_bw()+
  labs(title="", y="Frekvenssi", x="Nokan korkeus mm")

```



```
d=ggplot() +
  geom_histogram(data=fab, aes(x = up_man_height), col="gray20", fill="gray40",
alpha=1.0, bins = 30)+
  geom_histogram(data=ros, aes(x = up_man_height),col="darkgreen",
fill="darkolivegreen4", alpha=0.9, bins = 30)+
  theme_bw()+
  labs(title="", y="Frekvenssi", x="Ylänokan paksuus mm")
```

```
e=ggplot() +
  geom_histogram(data=fab, aes(x = low_man_height), col="gray20",
fill="gray40", alpha=1.0, bins = 30)+
  geom_histogram(data=ros, aes(x = low_man_height),col="darkgreen",
fill="darkolivegreen4", alpha=0.9, bins = 30)+
  theme_bw()+
  labs(title="", y="Frekvenssi", x="Alanokan paksuus mm")
```

```
f=ggplot() +
  geom_histogram(data=fab, aes(x = bill_col_ratio), col="gray20", fill="gray40",
alpha=1.0, bins = 30)+
  geom_histogram(data=ros, aes(x = bill_col_ratio),col="darkgreen",
fill="darkolivegreen4", alpha=0.9, bins = 30)+
  theme_bw()+
  labs(title="", y="Frekvenssi", x="Keltaista väriä nokassa %")
```

```
#Kuva
```

```
grid.arrange(a,b,c,d,e,f,ncol=2,nrow=3)
```

```
# Jaetaan data testi ja treeni dataan
```

```
set.seed(1)
```

```
sample <- sample(c(TRUE, FALSE), nrow(data), replace = TRUE, prob =
c(0.8,0.2))
```

```
train_data <- data[sample, ]
```

```
test_data <- data[!sample, ]
```

```
#Testataan malleja gmulti-paketin avulla
```

```
candidates1 <- glmulti(y = "subspecies_bin",
```

```

        xr = c("bill_col_ratio", "low_man_height",
"culmen_length_tip_top",
            "up_man_height", "bill_height",
"bill_gape_bill_length_tip_corner", "gap", "height_where_meas"),
        intercept = TRUE,
        level = 1,
        data = train_data,
        name = "glmulti.output",
        marginality = FALSE,
        method = "h", # try also other alternatives!
        crit = "aic",
        minsize = -1,
        maxsize = -1,
        minK = -1,
        maxK = -1,
        fitfunc = "glm",
        includeobjects = TRUE,
        family = "binomial"
)

```

```

weightable(candidates1)
summary.glmulti(candidates1)
coef.glmulti(candidates1)

```

#Valitaan paras malli

```

candidates <- glmulti(y = "subspecies_bin",
        xr = c("bill_col_ratio", "low_man_height",
"culmen_length_tip_top", "up_man_height"),
        intercept = TRUE,
        level = 1,
        data = train_data,
        name = "glmulti.output",
        marginality = FALSE,
        method = "h", # try also other alternatives!
        crit = "aicc",

```

```

        minsize = -1,
        maxsize = -1,
        minK = -1,
        maxK = -1,
        fitfunc = "glm",
        includeobjects = TRUE,
        family = "binomial"
    )
weightable(candidates)
summary.glmulti(candidates)
coef.glmulti(candidates)

# Testataan paras malli
hist(candidates@objects[[1]][["residuals"]])
newdata <- data.frame(bill_col_ratio = seq(min(train_data$bill_col_ratio),
max(train_data$bill_col_ratio),
                    len = length(train_data[,1])),
                    low_man_height = rep(mean(train_data$low_man_height),
length(train_data[,1])),
                    culmen_length_tip_top =
rep(mean(train_data$culmen_length_tip_top), length(train_data[,1])),
                    up_man_height = rep(mean(train_data$up_man_height),
length(train_data[,1])))
plot(train_data$bill_col_ratio, train_data$subspecies_bin)
lines(newdata$bill_col_ratio, predict(candidates@objects[[1]], type = "response",
newdata = newdata))
abline(h = 0.5, col = "red")
# Käytetään parasta mallia testidatan kanssa
pred <- predict(candidates@objects[[1]], test_data, type = "response")
# Optimaalinen leikkauskohta
optimal <- optimalCutoff(test_data$subspecies_bin, pred)[1]
optimal
# Sekaannusmatriisi
confusionMatrix(test_data$subspecies_bin, pred)
# Luokitteluvirhe
misClassError(test_data$subspecies_bin, pred, threshold = optimal)

```

