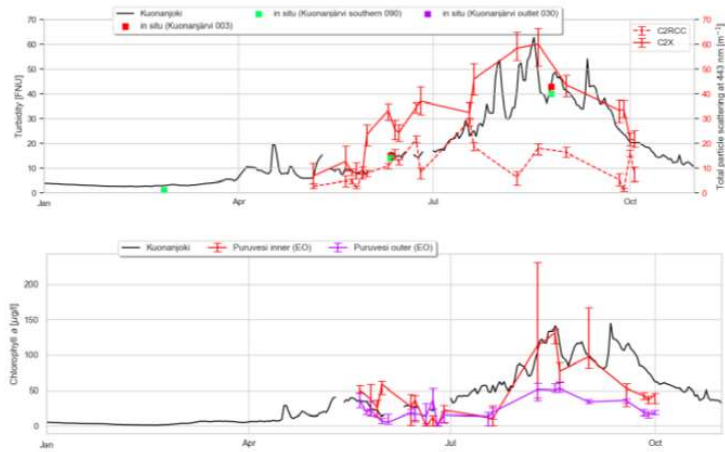
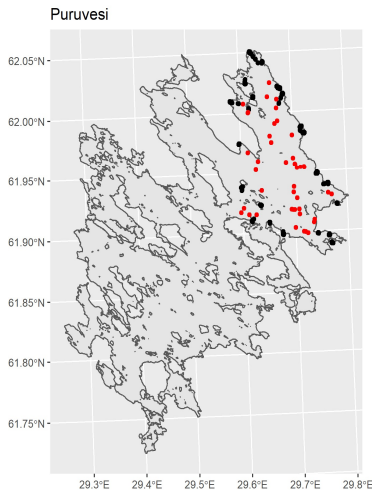




Action A6: Development of survey and assessment methods for lake habitats – Case study of Lake Puruvesi

Puruveden vedenlaadun ja vedenalaisten habitaattien mallintaminen kaukokartoituksen ja maastotutkimusten avulla



Jari Ilmonen, Metsähallitus

Seppo Hellsten, Sampsa Koponen, Sakari Väkevä, Jenni Attila, Ari-Pekka Jokinen, Antti Takolander, Janne Juntunen, Niko Kallio, Kimmo Tolonen, Elina Virtanen SYKE

Kristiina Nyholm, Anssi Lensu Jyu

Freshabit LIFE IP

Suomen ympäristökeskus

30.9.2022



Deliverable A6.3





Tiivistelmä

Raportti on osa FRESHABIT Life IP hankkeen laajaa A6 toimenpidettä, jossa on tehty kartoituksia ja mallinnuksia tiettyjen kirkasvetisten järvien vedenalaisesta luonnosta keskittyen suurelta osin Puruveden alueeseen. Satelliittitulkinnat ovat hyödyllinen menetelmä Puruveden tilan seurantaan. Maastohavaintoihin ja laboratorioanalyysiin verrattuna satelliittimenetelmät tuottavat havain-toja huomattavasti paremmalla alueellisella sekä ajallisella kattavuudella pitoisuuksien ollessa lähellä toisiaan. Satelliittimenetelmiä käytetään laajalti myös Suomen muilla sisävesialueilla. Satelliittikuvat yhdistettyinä maastotietoon tuottivat myös luotettavan syvyys ja pohjanlaatumallin, joita voitiin käyttää elinympäristöjen ennustamisessa.

Esimerkkinä rakennettiin tilastolliset habitaattimallit Puruvedelle neljälle pohjalehtiselle vesikasville; tummalahnaruoholle (*Isoëtes lacustris*), vaalealahnaruoholle (*Isoëtes echinospora*), raanille (*Littorella uniflora*) sekä nuottaruoholle (*Lobelia dortmanna*). Mallit rakennettiin käyttäen tehos-tettuja regressiopuita (Boosted Regression Trees, BRT) sekä yleistettyä additiivista mallia (Generalised Additive Model, GAM). Mallien avulla ennustettiin tutkittujen lajien esiintymistodennäköisyydet koko Puruvedelle sekä Puruveden koillisosassa sijaitsevalle Hummonselälle. Mallien ennustekyky oli yleisesti ottaen hyvä. Lajien esiintymistodennäköisyydet vaihtelivat lajikohtaisesti voimakkaasti syvyyden mukaan, joka oli tärkein esiintyvyyttä ennustava tekijä muilla lajeilla paitsi vaalealahnaruoholla, jonka esiintyvyys riippui voimakkaimmin veden virtausnopeudesta pohjan tuntumassa.



Summary in English

The report is part of the extensive A6 measure of the FRESHABIT Life IP project, where mapping and modeling of the underwater nature of clear-water lakes have been carried out, focusing largely on the Puruvesi area. Satellite interpretations are a useful method for monitoring the state of Puruvesi. Compared to field observations and laboratory analyses, satellite methods produce observations with significantly better spatial and temporal coverage when the concentrations are close to each other. Satellite methods are also widely used in other inland water areas in Finland. Satellite images combined with terrain data also produced a reliable depth and bottom quality model that could be used to predict habitats.

As an example, statistical habitat models were built in Puruvesi for four bottom-leaved aquatic plants; *Isoëtes lacustris*, *Isoëtes echinospora*, *Littorella uniflora* and *Lobelia dortmanna*. The models were built using Boosted Regression Trees (BRT) and the Generalized Additive Model (GAM). The models were used to predict the occurrence probabilities of the studied species for the whole of Puruvesi and Hummonselkä located in the northeastern part of Puruvesi. In general, the models' prediction capacity was good. The occurrence probabilities of the species varied strongly according to the depth, which was the most important factor predicting the occurrence of other species, with exception of *Isoetes echinospora* whose occurrence depended most strongly on the water flow rate near the bottom.

Hanke on saanut rahoitusta Euroopan unionin LIFE-ohjelmasta. Aineiston sisältö heijastelee sen tekijöiden näkemyksiä, eikä Euroopan komissio tai CINEA ole vastuussa aineiston sisältämien tietojen käytöstä.

The project has received funding from the LIFE Programme of the European Union. The material reflects the views by the authors, and the European Commission or the CINEA is not responsible for any use that may be made of the information it contains.



SISÄLLYS

Tiivistelmä.....	3
Summary in English.....	4
1. Satelliittihavaintojen tuottaminen.....	8
Summary	8
Taustaa	8
Menetelmät.....	9
Vedenlaatu	9
Ruovikot	9
Tulokset.....	10
Vedenlaatukartat	10
Vedenlaadun aikasarjat	12
Ruovikot	16
Johtopäätökset.....	16
2. Puruveden syvyys- ja pohjan laatumallin muodostaminen.....	18
Summary	18
Aineiston kuvaus.....	18
Satelliittikuvilta johdettu syvyystieto.....	18
Syvyysmalli.....	19
Syvyystietojen yhdistäminen.....	20
Kirjallisuus	22
3. Pohjanlaatu ja satelliittikuvaluokittelu	23
Summary	23
Aineiston kuvaus.....	23
Pohjatyypiluokittelu	23
Menetelmä ja tulokset	24
Luokittelutodennäköisyydet.....	26
Aineistojen yhdistäminen.....	27
Yhteenveto	27
Kirjallisuus	28
4. Pohjalehtisten vesikasvien lajimallien muodostaminen.....	30
Summary	30
Johdanto	30
Aineisto ja menetelmät.....	31
Tulokset.....	33
Mallien ennustekyky.....	33
Muuttujien tärkeys ja lajien vasteet ympäristömuuttujiin	35
Vasteet ympäristömuuttujiin: Isoetes lacustris	35
Vasteet ympäristömuuttujiin: Isoetes echinospora	36
Vasteet ympäristömuuttujiin: Lobelia dortmanna	37
Vasteet ympäristömuuttujiin: Littorella uniflora	38
Ennustettujen esiintymisten spatiaalinen jakautuminen Hummonselällä.....	39
Lajien ennustetut esiintymistodennäköisyydet Hummonselälle.....	41
Spatiaaliset ennusteet koko Puruvedellä.....	42



Pohjalehtisten lajien tärkeimmät esiintymisalueet Hummonselällä.....	46
Yhteenveto	49
Kirjallisuus	49

Deliverable A6.3





1. Satelliittihavaintojen tuottaminen

Sampsa Koponen, Sakari Väkevä & Jenni Attila

Summary

Satellite observations (Earth Observation, EO) provide information about the aquatic environment with good spatial and temporal coverage. In the FRESHABIT project the observations were used to support habitat modeling and to provide information about water quality parameters. SYKE provided (1) monthly map composites of turbidity and CDOM absorption from years 2017-2021 and months May to – October/November, (2) time series of turbidity and chlorophyll-a values from an area affected by eutrophication, and (3) a map of common reed areas. The demonstrations show the advantages and possibilities provided by satellite observations.

Taustaa

Satelliitteihin sijoitetut mittalaitteet kuvaavat maapallon pintaa havainnoimalla auringon valoa, joka heijastuu pinnalla olevasta kohteesta (ns. optiset instrumentit), kohteen itse lähettämää lämpösäteilyä (mm. mikroaaltoradiometrit) tai instrumentin itse lähettämää ja kohteen takaisin heijastamaa säteilyä (tutkat ja lidarit). Kohteen ominaisuuksia voi selvittää analysoimalla vastaanotettua signaalia.

Satelliittihavainnoinnin etuina ovat:

- Hyvä kattavuus: Havainto voi muodostaa kuvan, joka kattaa laajan alueen kokonaan
- Hyvä toistettavuus: Havainto voidaan toistaa jopa päivittäin.

Esim. EU:n Copernicus-ohjelman Sentinel-3 satelliitit (2 kpl) kuvaavat koko Suomen päivittäin 300 m maastoerotuskyvyllä, kun taas Sentinel-2 satelliitit (2 kpl) kuvaavat Suomen 2-3 päivän välein 10-60 m maastoerotuskyvyllä. Sentinel-2 satelliittien hyvän maastoerotuskyvyn edut korostuvat Suomen pienipiirteisissä vesistöissä, vaikka mittaus ei toistukaan päivittäin. Sentinel-satelliittien lisäksi SYKE hyödyntää USGS Landsat-ohjelman satelliitteja, jotka tarjoavat havaintoja 30 m maastoerotuskyvyllä hieman Sentinel-2:sta harvemmin.

FRESHABIT-projektissa satelliittihavaintoihin liittyviä tavoitteita olivat:

- Hyödyntää satelliittihavaintoja mallinnuksen lähtöaineistona. Lisätietoja mallinnustuloksista dokumentissa: Freshabit_A6_vesikasvimallit_työraportti.
- Seurata vedenlaatua satelliittihavaintojen avulla tietyllä kohdealueella

Satelliittihavaintojen avulla voi myös arvioida pohjan laatua ja veden syvyyttä. Näitä kuvataan erillisessä kappaleessa Freshabit syvyys- ja pohjanlaatumallit työraportti.



Menetelmät

Vedenlaatu

Työssä keskityttiin optisten satelliitti-instrumenttien hyödyntämiseen, sillä niillä on sopivat ominaisuudet vedenlaadun havainnointiin. Optinen satelliitti-instrumentti havaitsee näkyvän valon ja lähi-infrapuna-alueen säteilyä useilla aallonpituuksilla. Kohteen ominaisuudet vaikuttavat eri tavalla eri aallonpituuksilla, joten mitatun säteilyn spektrin muotoa ja voimakkuutta analysoimalla voi tehdä päätelmiä kohteen ominaisuuksista. Menetelmät voivat olla yksinkertaisia kanavasuhteeseen perustuvia, mutta vedenlaadun osalta nykyiset menetelmät perustuvat neuroverkkoihin. Neuroverkko on koulutettu suurella määrällä opetusdataa ja ne pystyvät muuntamaan mitatun spektrin nopeasti esim. klorofylli-, sameus- ja humusarvoiksi.

SYKE hyödyntää aineistotuotannossaan yllä mainittuja Sentinel-satelliitteja ja niille kehitettyjä prosessoreita. Tällaisia ovat esim. C2RCC ja C2X. Suomen vesistöt eroavat näiden prosessoreiden opetusdatasta esim. suuren humuspitoisuuden osalta. Näin ollen on usein tarpeen kalibroida prosessorin tulos vastaamaan paremmin maasto-olosuhteita. Kalibrointi suoritetaan hyödyntämällä maastohavaintoja.

Yksittäisessä kuvassa voi olla pilvien ja niiden varjojen aiheuttamia häiriöitä. Mallit tarvitsevat aukotonta lähtöaineistoa, joten muodostimme yksittäisistä kuvista ns. kuukausikomposiitteja, joissa yhden pikselin arvo muodostuu kuukauden aikana samasta pisteestä tehtyjen havaintojen keskiarvosta tai mediaanista. Pilvien vaikutuksen lisäksi keskiarvoistaminen vähentää yksittäisissä kuvissa esiintyvää kohinaa.

Matalassa ja kirkkaassa vedessä pohjan heijastus voi vaikuttaa vedenlaatutulkintaan. Nämä alueet usein poistetaan tulokartoista epäluotettavien arvojen takia. Tässä tapauksessa alueet kuitenkin jätettiin mukaan mallintajien toiveesta ja aineistot rajattiin rantaviivan ja bufferin avulla.

Tuotetut aineistot sisältävät sameus ja humus (CDOM absorptio) kuukausikeskiarvoja vuosilta 2017-2021 kuukausilta touko – loka/marras. Talven aikana jääpeite ja pimeys estävät kuvien prosessoinnin ja joiltakin yksittäisiltä kuukausilta riittävän laadukasta aineistoa ei syntynyt.

Kuukausikarttojen lisäksi on mahdollista seurata erillisten kohdealueiden vedenlaatua aikasarjana. Tässä tapauksessa SYKE prosessoi halutut alueet jokaisesta kuvasta ja laskee alueen pikseliarvoista keskiarvon, mediaanin ja erilaisia persentiilejä. Alueen koko on halkaisijaltaan yleensä noin 960 m ja sen sisällä olevat maa-alueet poistetaan. Myös pilvien peittämien alueet karsitaan pois tietovirrasta. Tilastollinen aikasarjatuotanto sisältää sameuden lisäksi klorofylli-a:n laskennan. Näissä aikasarjoissa on havaintoja parhaimmillaan 2-3 päivän välein, mutta pilvisellä säällä voi olla pitkiäkin taukoja. Myös veden pintalämpötilan seuranta on mahdollista 100 m pikseli koolla (Landsat-8 & 9). Nämä mittaukset perustuvat veden lähettämän lämpösäteilyn mittaukseen.

Ruovikot

Ruovikoiden havainnointi perustuu kasvillisuuden erottamiseen vedestä. Kasvillisuus heijastaa lähi-infrapuna-alueen säteilyä voimakkaasti, kun taas vesialueet ovat tällä aallonpituudella tummia. Ns. NDVI (eli Normalized Differential Vegetation Index) on hyvä menetelmä kasvillisuuden määrän arviointiin: Korkea NDVI merkitsee kasvillisuutta ja matala vettä. Maanpäällisen kasvit erotetaan vedessä kasvavista [Ranta10](#) rantaviiva-aineistoa hyödyntämällä. Tässäkin hyödynnettiin kuukausikomposiitteja poistamaan yksittäisissä kuvissa olleita pilvihäiriöitä. Lisätietoja menetelmästä löytyy täältä: www.syke.fi/saltmarshstorymap.



Tulokset

Vedenlaatukartat

Mallintajat olivat eniten kiinnostuneita vedenlaatumuuttujista, jotka vaikuttavat valon tunkeutumissyvyyteen, koska tämä vaikuttaa siihen, miten syvällä kasvit voivat elää. Näitä ovat sameus ja humuksen määrä.

Tulokset löytyvät SYKEN verkkolevyltä:

\\fs-seamboth\seamboth\freshabit\Puruvesi_cdom_turb_month

Kuvassa 1 näkyy esimerkki vuodelta 2021. Puruveden sameusarvot ovat pääasiassa alhaisia. Korkeampia arvoja näkyy rantojen läheisyydessä. Osa näistä on pohjan vaikutusta.

Yksittäisen päivien sameuskuvia on katsella TARKKA palvelun kautta. Yksi esimerkki näkyy Kuvassa 2. Suora-linkki palveluun:

https://wwwi4.ymparisto.fi/i4/fin/tarkka/index.html?type=TURB&date=2021-08-29&lang=fi&zoom=12&lat=61.90566&lon=29.63035&valuerange=1:6&colormap=EO_WQ_turb_7steps_adjutable&op=0

TARKKassa olevissa aineistoissa ranta-alueet on maskattu (tummempi harmaa), jotta ranta- ja pohjavaikutukset eivät näy.

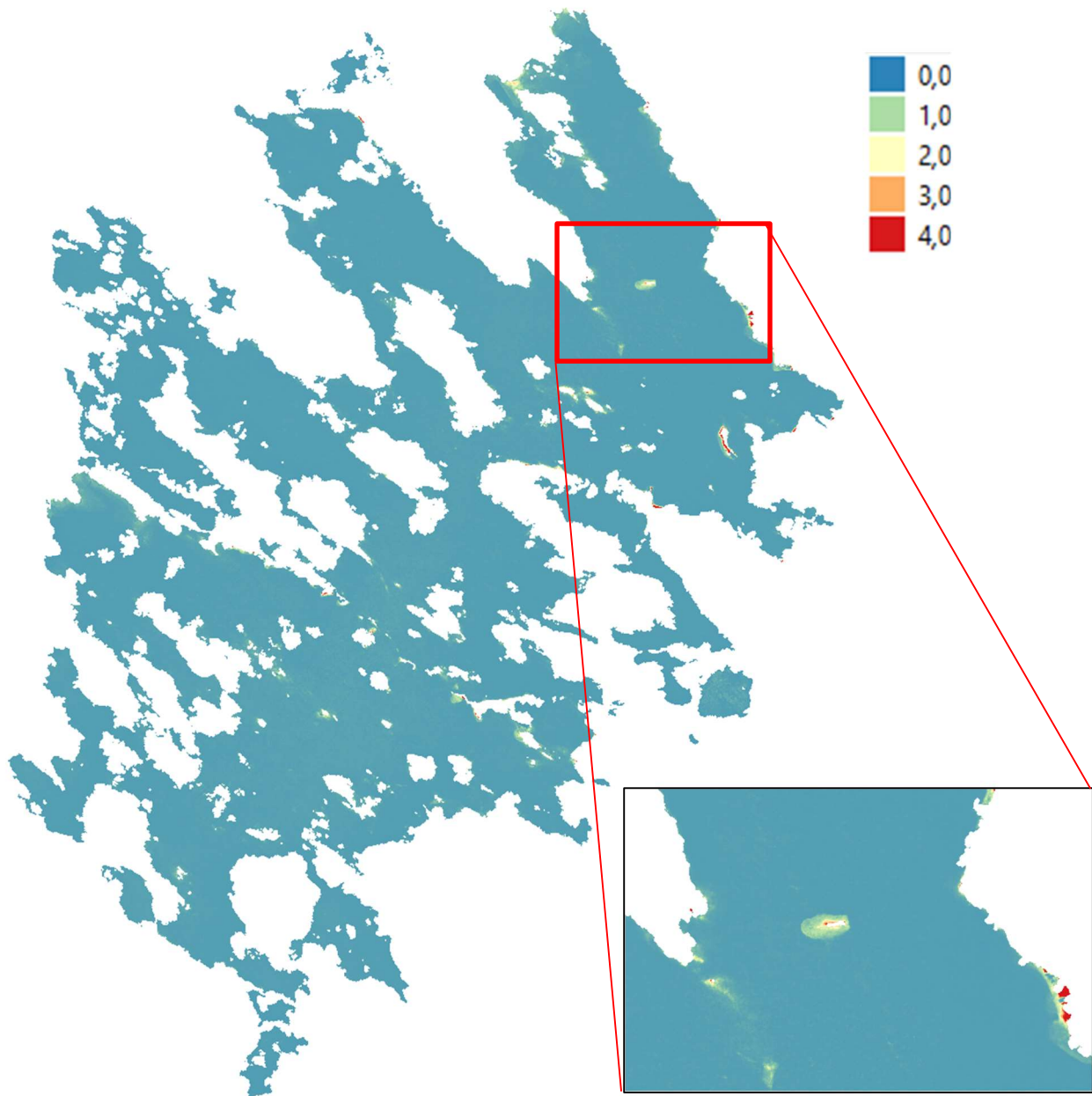


NATURA 2000

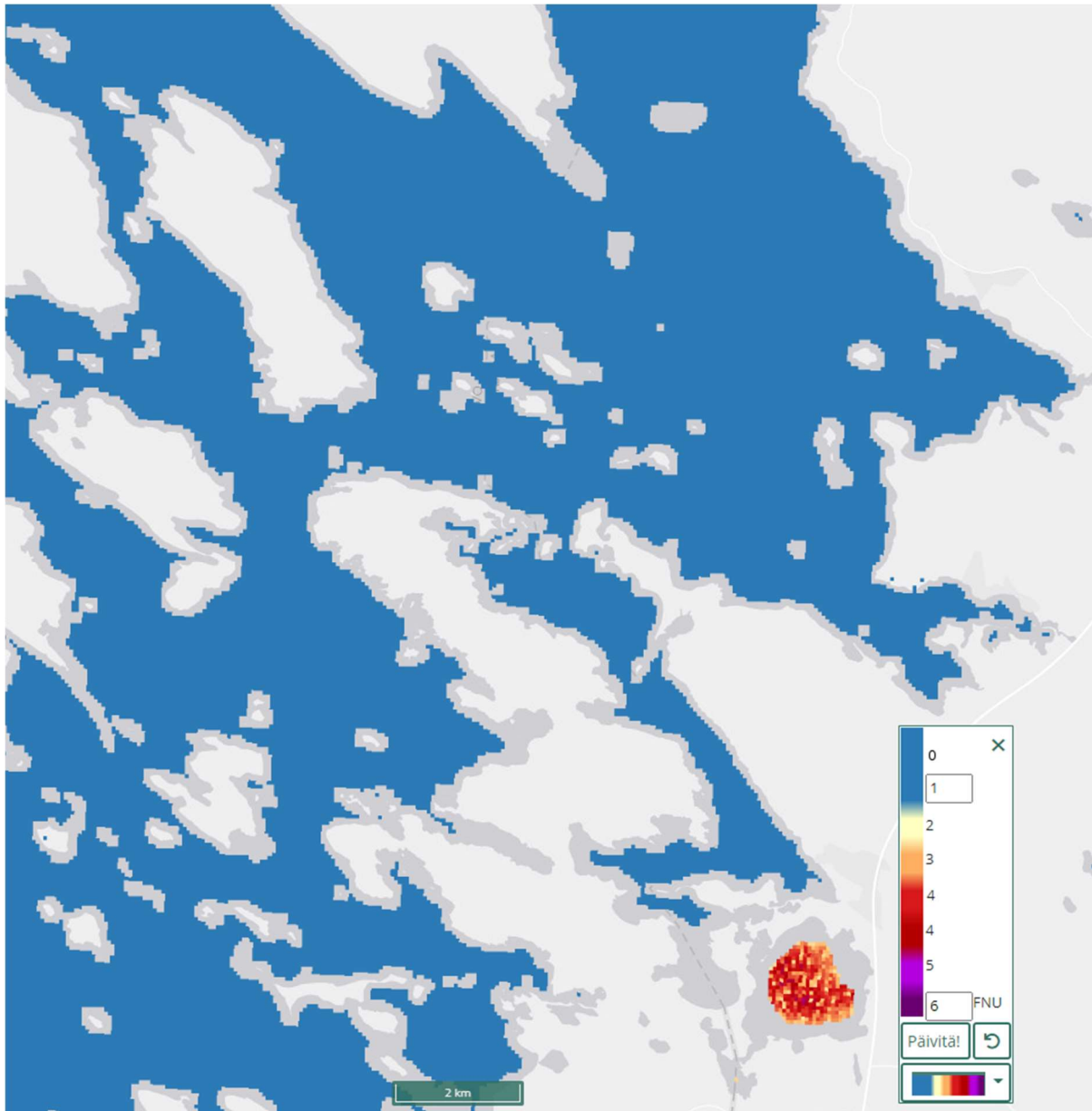
FRESHABIT

S Y K E

METSÄHALLITUS
FORSTSTYRELSEN



Kuva 1. Sentinel-2 satelliittihavaintoihin perustuvia Puruvesen sameusarvoja heinäkuussa 2021 (kuukausikomposiitti). Asteikko yksikössä FNU. Suurennos näyttää miten Hummonselällä oleva matalikko vaikuttaa tulokseen.



Kuva 2. Sentinel-2 satelliitin havaintoon perustuva sameuskartta Puruvedeltä 29.08.2021 [TARKKA](#)-palvelusta katseltuna. Näissä kartoissa käytetään laajempaa bufferia matalien alueiden poistamiseen, jolloin Puruvedellä näkyy vain matalia sameusarvoja. Vain Ristilahdella (oikea alakulma) näkyy kohonneita arvoja.

Vedenlaadun aikasarjat

Yksi satelliittihavainnoinnin kannalta mielenkiintoinen paikka Puruvedellä on Savonlahti. Hyvin rehevä Kuonanjärvi laskee siihen Kuonanjokea pitkin ja huonontaa vedenlaatua alueella. Tilanne näkyy hyvin tosivärikuvien avulla (Kuva 3) sekä sameuskartoista. Loppukesällä levän määrä Kuonanjärvessä kasvaa ja tämä samentaa vettä myös Puruvedellä.

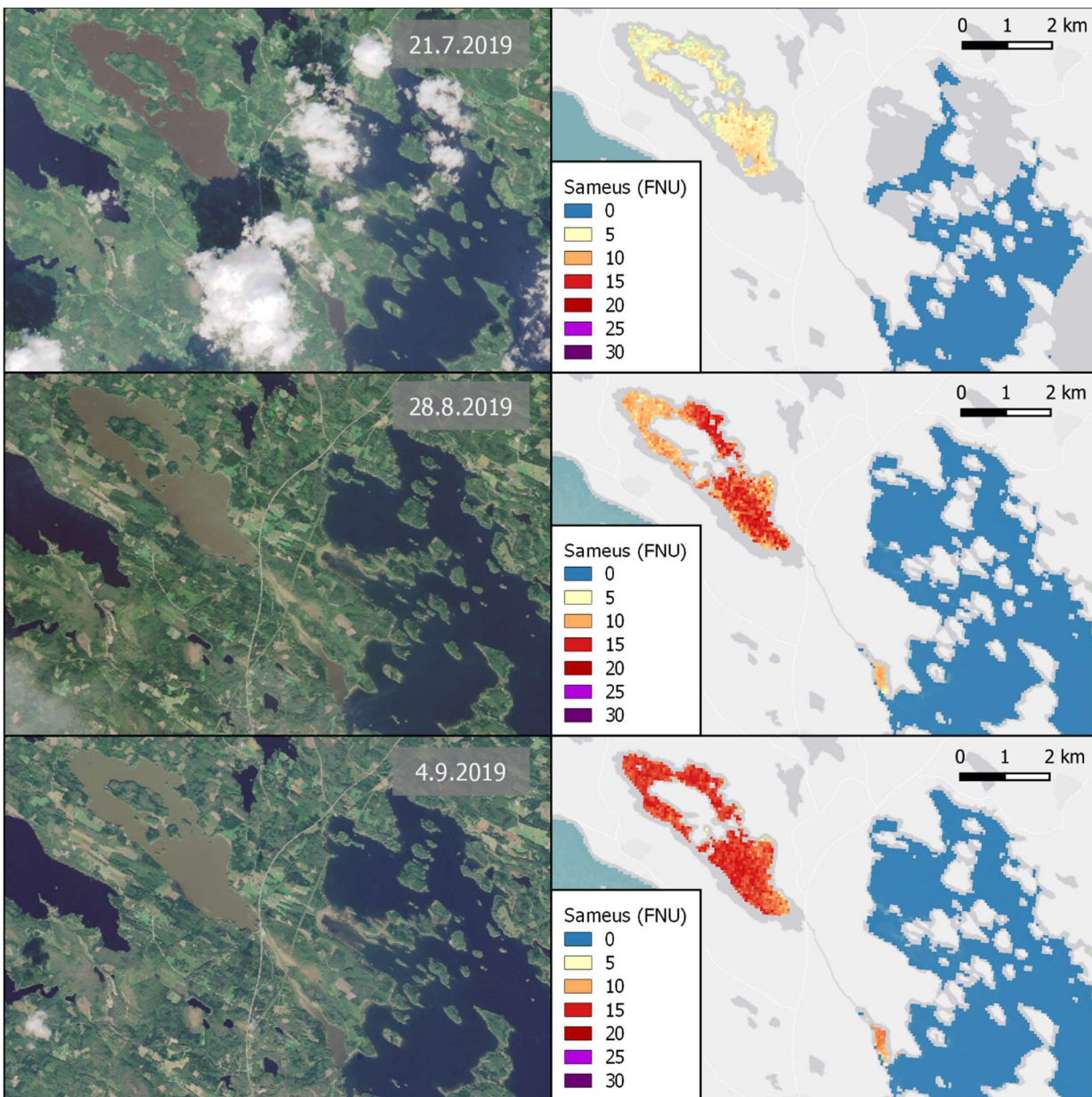
Seurasimme samentumisen alueellista vaikutusta määrittämällä joen purkupisteen edustalle, Savonlahdelle ja sen ulkopuolelle seuranta-alueita (Kuva 4 kartta), joiden sisältä poimimme kaikki saatavilla olevat Sentinel-2 sameusarvot. Tulosten mukaan (Kuva 4 aikasarjat) samentuminen toistuu Savonlahdella vuosittain



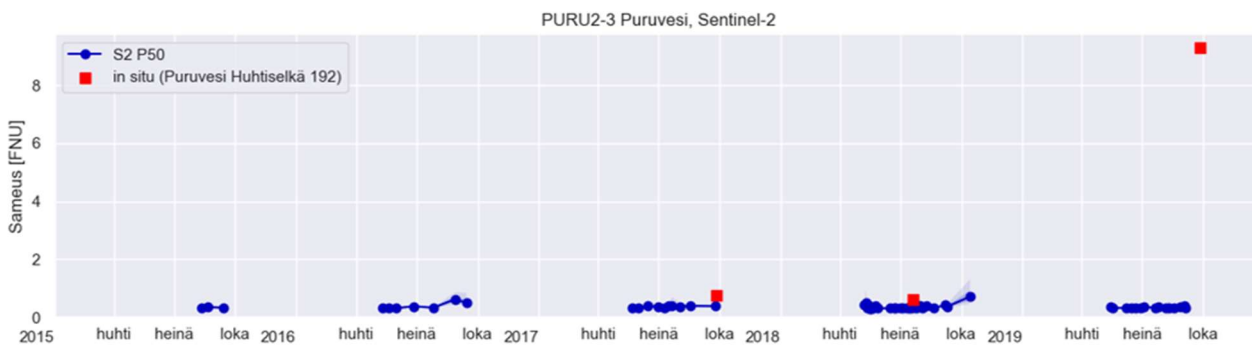
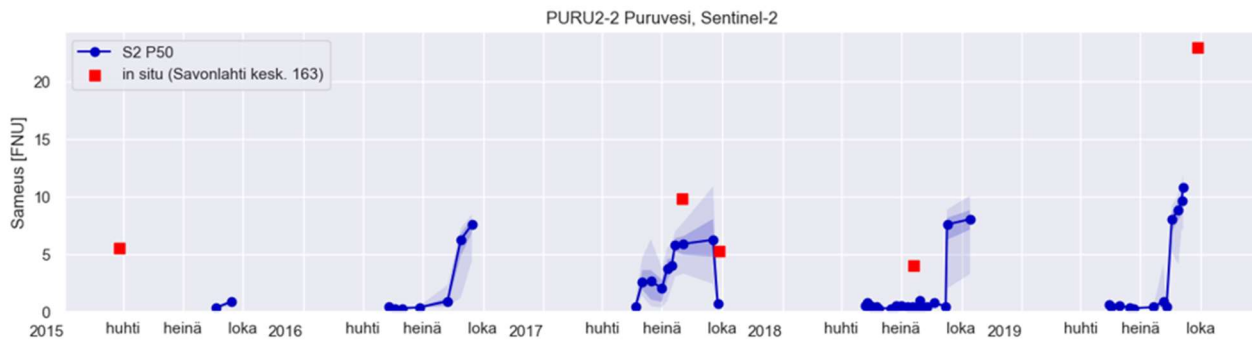
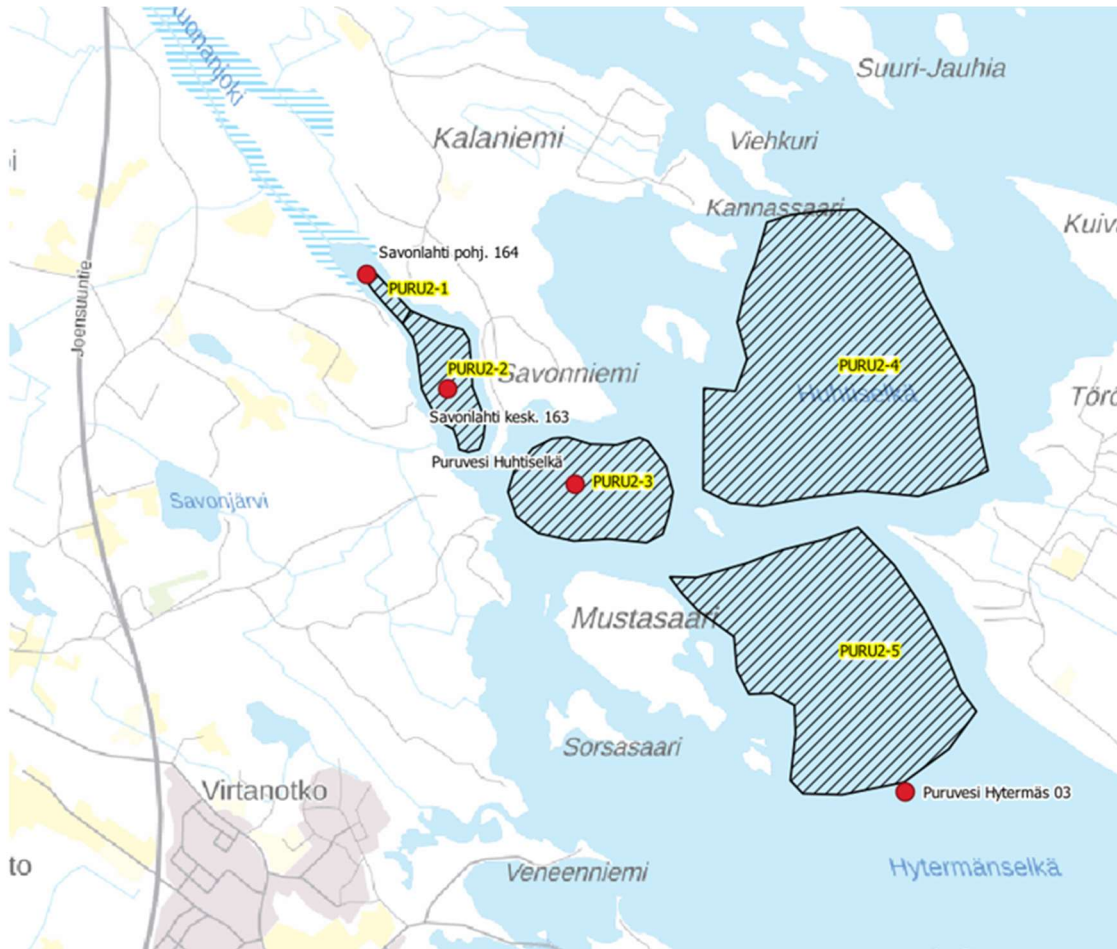
loppukesästä, mutta ei ulotu sen ulkopuolelle. Vuoden 2019 lokakuussa on yksi korkea maastohavainto molemmilta alueilta, mutta ajankohdalta ei ole pilvetöntä satelliittikuvaa.

Vuonna 2021 Kuonanjoessa oli automaattimittausasema, joka tuotti päivittäisiä klorofylli- ja sameushavaintoja. Näitä verrattiin satelliitti havaintoihin ja vastaavuus näiden kahden datalähteen välillä on varsin hyvä (Kuva 5). Näistä tuloksista on työn alla tieteellinen lehtiartikkeli (Operandum yhteistyö).

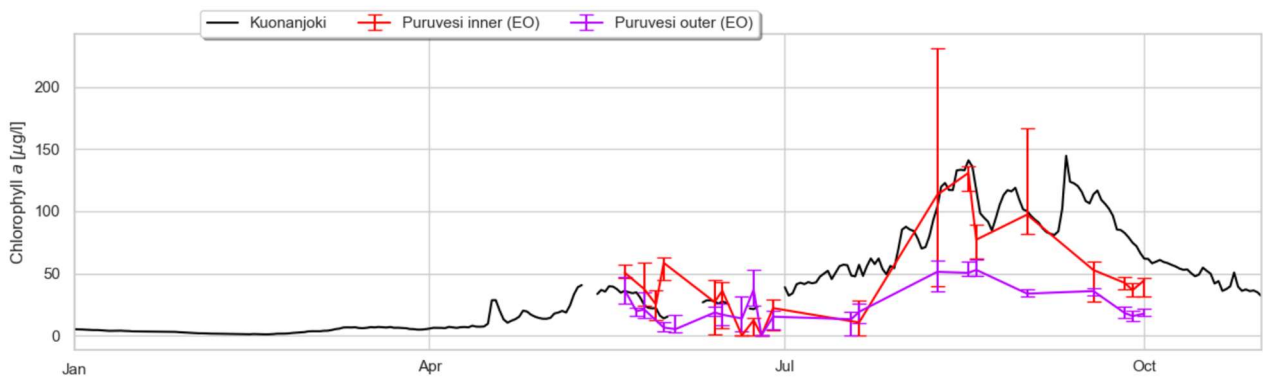
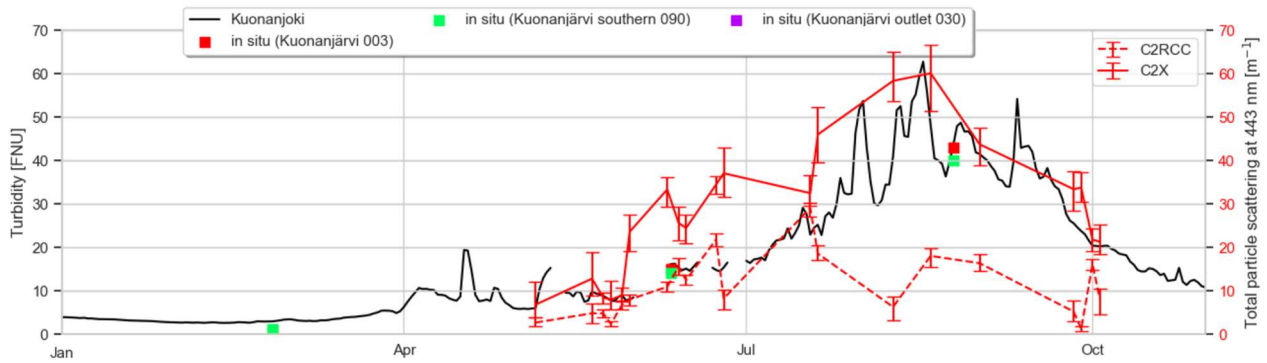
Kuva 6 esittää veden pintalämpötilan käyttäytymisen Hummonselällä vuosien 2021 ja 2022 aikana (aika-sarja päättyy raportin kirjoitushetkeen elokuun alussa) maasto- ja satelliittihavaintoihin perustuen. Kuvassa näkyy hyvin lämpötilan vuodenaikaisvaihtelu sekä satelliittihavaintojen tarkkuus maastomittauksiin verrattuna (datapisteet ovat hyvin lähellä toisiaan).



Kuva 3. Tosivärikuvia ja sameuskarttoja Kuonanjärven ja Puruveden Savonlahden alueella.



Kuva 4. Poiminta-alueet Savonlahden alueella (yläkuva) ja sameusaikasarjoja kahdelta poiminta-alueelta (alakuva). Saatteliittihavaintojen (sininen) lisäksi kuvassa näkyy maastohavaintoja (punainen).



Kuva 5. Klorofylli-a aikasarja Kuonanjärvessä olevalta automaattiasemalta (musta) sekä kahdelta Puruvesen puolella olevalta poiminta-alueelta (inner = lähempänä Kuonanjokea, outer = kauempana) vuonna 2021.

Rannikon ja järvien vertailuasemat

Paikka nro	Aseman nimi	Vesimuodostuman nimi	VPD-tunnus	VPD-tyyppi	ELY	Syvyys [m]
22065	Puruvesi 1 Hummonselkä	"Puruvesi (Saimaa), keskusallas"	04.181.1.001_001	SVh	ESA	60

Pintalämpötila Sameus a-klorofylli



Vie EO csv-tiedostoon

Koordinaatit: N 61.95648, E 29.70454

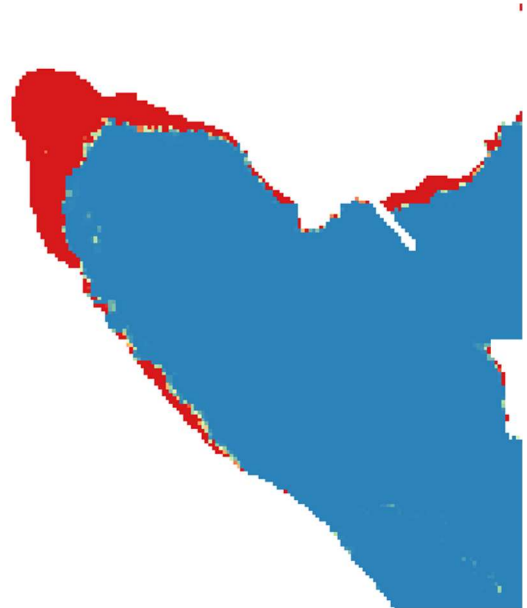


Kuva 6. Hummonselkä-aseman pintaveden lämpötila satelliittihavaintojen (EO, sininen) ja maastomittausten (MS, punainen) perusteella.

Ruovikot

Tuotettu ruovikkoaineisto kattaa koko Suomen ja on saatavilla pyydettäessä. Esimerkki Puruvedeltä näkyy Kuva 7. Ruovikkotulkinta vastaa hyvin tosivärikuvassa esiintyviä vesikasvillisuusalueita.

Ruovikkotulkinnasta laskettu ruovikon kokonaispinta-ala Puruvedellä on 18,1 km² (koko Puruveden pinta-ala on 420,9 km²).



Kuva 7. Vasen: Tosivärikuva Puruveden Ketolanlahdelta (pohjoisosa). Sininen viiva on rantaviiva. Oikea: Ruovikkotodennäköisyys, jossa punainen on ruovikkoa ja sininen vettä (vuosi 2019).

Johtopäätökset

Karttamuotoisten satelliittitulosten mukaan suurimmassa osassa Puruvedtä vesi on kirkasta (sameusarvot ovat alhaisia). Kohonneita arvoja on yksittäisissä lahdissa, joissa järveen tuleva jokivesi ja/tai sulkeutuneisuus ovat aiheuttaneet rehevöitymistä.

Aikasarjaverailut osoittavat, että satelliittitulokset vastaavat hyvin maastohavaintoja, joten niiden käyttöä Puruveden vesiympäristön seurannassa on mahdollista laajentaa. Havaintojen muuttaminen aikasarjamuotoon tiivistää tietoa ja siten helpottaa vedenlaadun seuraamista. SYKE:n TARKKA palvelussa (www.syke.fi/tarkka) on tällä hetkellä 10 virtuaalista vertailuasemaa Puruvedellä (valittujen VESLA-asemien kohdalla), joista kerätään satelliittihavaintoja automaattisella prosessilla. Määrää on mahdollista kasvattaa sijoittamalla Puruvedelle uusia asemia.

Ruovikkotulkinnat vastaavat hyvin satelliittikuvilta ihmissilmän tehtyä arviota. Yllä kuva kuvattu menetelmä on toistettavissa, jolloin ruovikoiden (vesikasvien) peittämää alaa voi seurata vuosittain tai jopa kuukausittain kasvukauden aikana. Menetelmää ei ole automatisoitu, joten prosessointi ja tulosten käsittely vaatii vielä työaikaresursseja.

Deliverable A6.3



Yllä esitetyt tulokset osoittavat, että satelliittitulkinnat ovat hyödyllinen menetelmä Puruveden tilan seurantaan. Maastohavaintoihin ja laboratorioanalyysiin verrattuna satelliittimenetelmät tuottavat havain-
toja huomattavasti paremmalla alueellisella sekä ajallisella kattavuudella pitoisuuksien ollessa lähellä toisi-
aan. Satelliittimenetelmiä käytetään laajalti myös Suomen muilla sisävesialueilla. Esimerkiksi vertailuasemia
on Suomen järvissä ja rannikolla tällä hetkellä jo 362 kpl.



2. Puruveden syvyys- ja pohjan laatumallin muodostaminen

Ari-Pekka Jokinen

Summary

Depth model for the Puruvesi area was created as a combination of echo soundings and satellite derived bathymetry (SDB). SDB was derived from Sentinel-2 MSI imagery using the logarithmic band ratio of blue (492 nm) and red (665 nm) band. The model was calibrated with dive recorded depth points and the model accuracy assessed against echo soundings. Accuracy assessment showed mean absolute error of 0,46 meters and root mean squared error of 0,66 meters. Comparison to echo sounding revealed that uncertainty in SDB increased after 3 meters. Thus, the final depth layer was created by interpolating deep areas from echo soundings and using the SDB in shallow areas (less than 3 meters) where echo sounding was absent.

Aineiston kuvaus

Puruveden, ja erityisesti Hummonselän alueen syvyysaineisto on yhdistelmä kaikuluotauksista sekä satelliittikuvilta johdetusta syvyystiedosta. Luotausaineisto koostuu Väyläviraston ja Geologisen tutkimuskeskuksen (GTK) aineistosta sekä Metsähallituksen (MH) suorittamista Biobase-luotauksista. Syvyysmalli luotiin käytettäväksi yhtenä muuttujana kasvillisuusmalleissa. Kasvillisuusmallit on kuvattu erillisraportissa Freshabit vesikasvimallit.

Satelliittikuvilta johdettu syvyystieto

Satelliittikuvilta johdettu syvyystieto (satellite derived bathymetry, SDB) perustuu pohjasta takaisinheijastuvan näkyvän valon aallonpituuksien havaitsemiseen ja signaalin heikentymiseen sen kulkiessa vesipatsaan läpi. Valon absorboituminen vedessä vaihtelee eri valon aallonpituuksilla, joten kahden eri aallonpituuskanavan suhde muuttuu syvyyden lisääntyessä. Pohjatyyppi ja sen heijastavuus vaikuttaa molempiin kanaviin, mutta veden syvyys vaikuttaa herkemmin absorboituvaan kanavaan enemmän. Näin ollen, syvyyden vaikutus aallonpituuskanavien suhteeseen on pohjatyyppin vaikutusta suurempi, ja kanavasuhde tietyllä syvyydellä säilyy muuttumattomana pohjatyyppistä huolimatta (Stumpf et al. 2003). Syvyys (Z) voidaan johtaa satelliittikuvan kanavasuhteista kaavalla (1), jossa

$$Z = m_1 \frac{\ln(nR_w(\lambda_i))}{\ln(nR_w(\lambda_j))} - m_0$$

(1)

m_1 on sovittava vakio kanavasuhteen skaalaamiseksi syvyyksiksi, n vakio varmistamaan positiiviset arvot logaritmeista, $R_w(\lambda_{i,j})$ ilmakehäkorjattu reflektanssi veden pinnan tasossa kanavilla i, j ja m_0 poikkeama 0 metrin syvyyteen (Stumpf, Holderied et al. 2003). Tutkimusalueelta tarvitaan syvyystietoa mallin kalibrointiin sekä m_1 ja m_0 ratkaisemiseen.

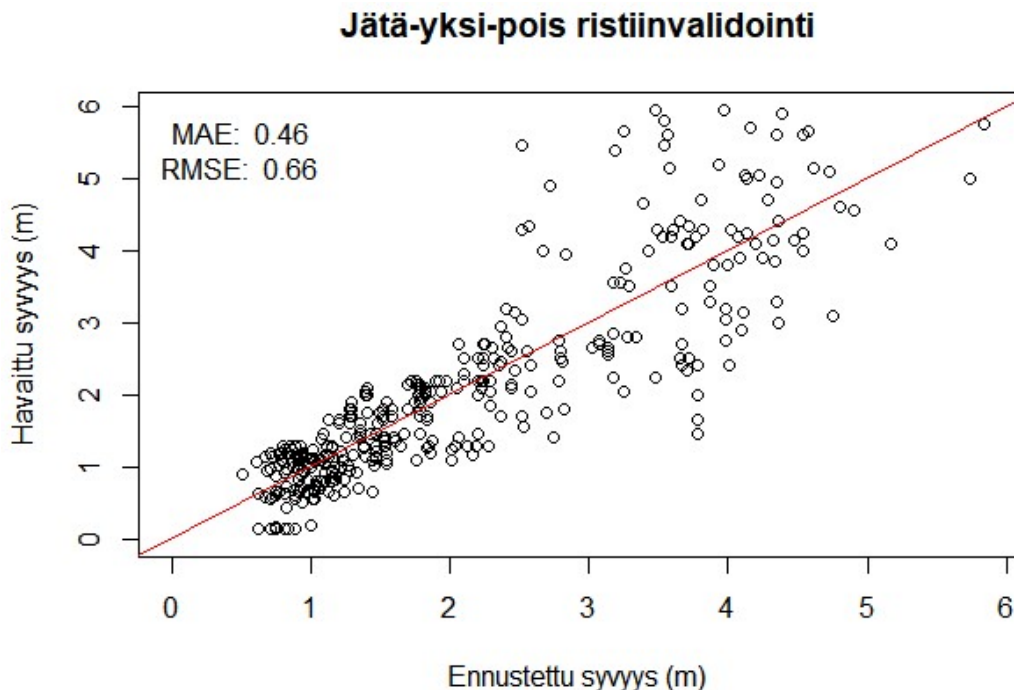


Satelliittipohjaisen syvyystiedon tuottamiseen käytettiin 31.7.2018 otettua Sentinel 2 satelliittikuvaa. Ilmakehäkorjaukseen käytettiin ACOLITE prosessoria (Vanhellemont 2019), jonka lopputuotteena oli Remote sensing reflectance (Rrs) veden pinnan tasosta. Sentinel 2 kuvan aallonpituuskanavista käytettiin sinisen (492 nm) ja punaisen (665 nm) kanavan suhdetta syvyystiedon tuottamisessa. Puruveden vesialue rajattiin satelliittikuvasta Maanmittauslaitoksen (MML) maastotietokannan järvipolygonilla, minkä lisäksi normalisoitua kasvillisuusindeksiä (NDVI) käytettiin maskina poistamaan rasterisolut, joissa NDVI arvo oli suurempi kuin 0,1. NDVI-maskia käytettiin, jotta osittain vedessä kasvava kasvillisuus (esim. ruovikot) saatiin erotettua aineistosta.

Syvyysmalli

Syvyysmallin kalibrointiin käytettiin Freshabit-lajihavaintojen ohessa maastossa kerättyä syvyystietoa. Pisteistä valittiin syvyydeltään 0,1–6,0 metrin pisteet. Syvempiä pisteitä ei valittu sillä oletuksella, että pohjasta heijastuva valo heikenee huomattavasti muutaman metrin jälkeen. Lisäksi pisteitä poistettiin samalla NDVI-maskilla, jota käytettiin logaritmisin kanavasuhdetason maskina. Pisteiden suodattamisen jälkeen 382 pistettä jäi käytettäväksi mallin kalibrointiin.

Mallina kanavasuhteen ja syvyyden välillä käytettiin Generalized Additive Modelia (GAM) mgcv-paketista R-Studioissa. Mallin selitysaste oli 76 % ja 'jätä-yksi-pois' ristiinvalidoinnin keskimääräinen absoluuttinen virhe 0,46 m (Kuva 8).

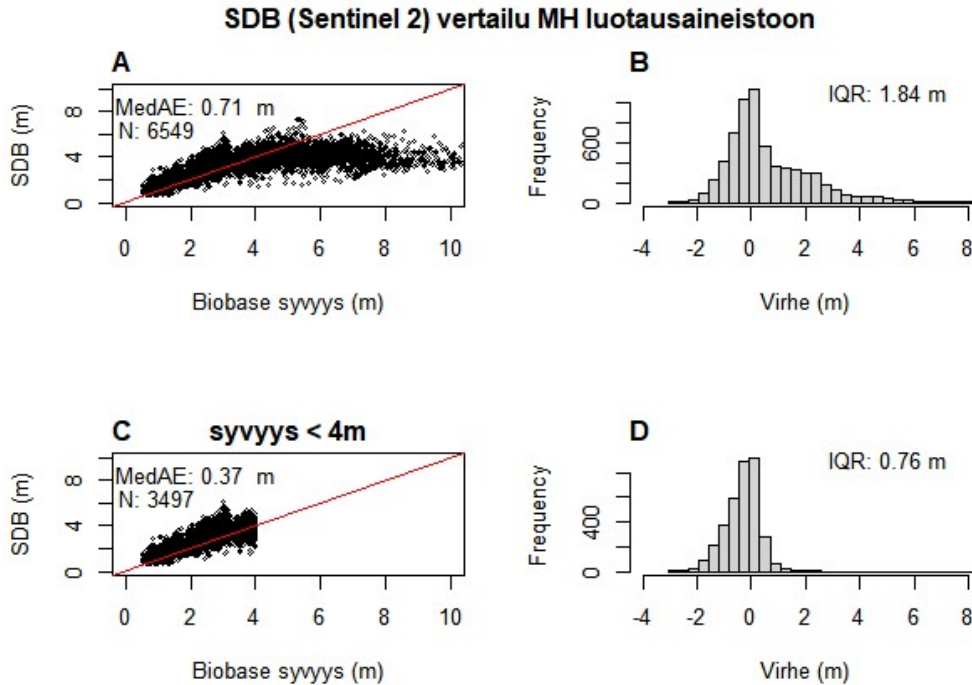


Kuva 8. Mallin ristiinvalidointi 'jätä-yksi-pois' menetelmällä.

Mallilla luotu syvyyssennuste validoitiin vertaamalla satelliittikuvapohjaista syvyystietoa Metsähallituksen Biobase-luotausten syvyystietoon (Kuva 9). Vertailusta nähdään kuinka 3–4 metrin jälkeen satelliittikuvasta tuotettu syvyystieto ei enää vastaa luotaamalla havaittuja syvyyksiä (Kuva 9 A). Mediaani absoluuttinen virhe koko vertailuaineistolle oli 0,71 metriä kun vertailemalla alle 4 metrin syvyyksiä virhe oli 0,37 metriä (Kuva 9



A ja C). Virheiden keskimäinen 50 % (IQR) laski 1,84 metristä 0,76 metriin kun yli 4 metrin syvyydet poissuljettiin (Kuva 9 B ja D).



Kuva 9. Syvyysmallin vertailu Metsähallituksen Biobase-luotausaineistoon. Satelliittipohjainen syvyystieto (SDB) ulottuu noin 3-4 metriin (A), jonka jälkeen syvyys ei vastaa luodattua syvyystietoa. Aineiston rajaaminen alle 4 metrin syvyyksiin (C) näyttää absoluuttisen virheen mediaanin lähes puolittuvan sekä virheiden keskimäisen 50 % pienenevän (B, D).

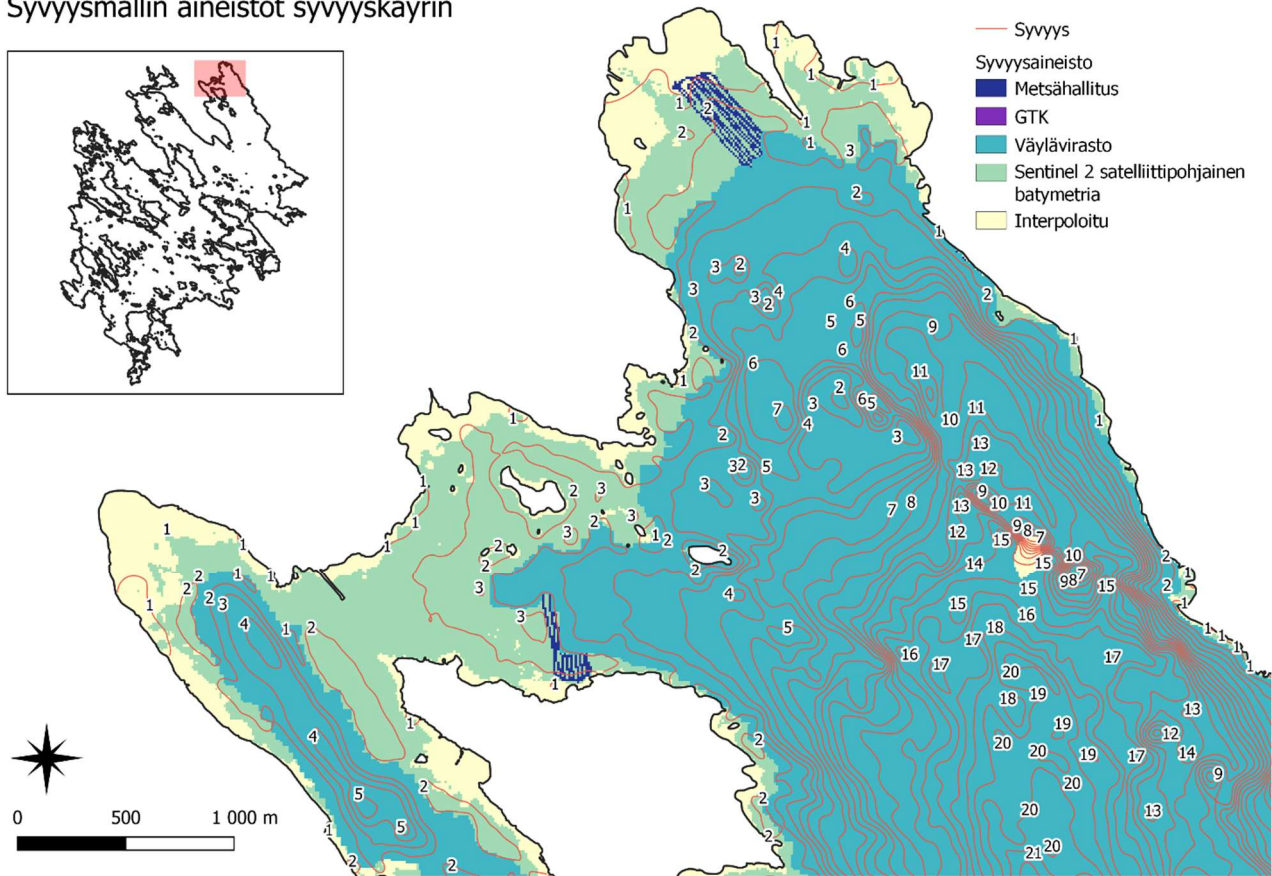
Hieman 0 alapuolelle keskittyvä virhejakauma (mediaani: -0,26 m) viittaa satelliittikuvapohjaisen syvyystiedon tuottaneen luotautietoa suurempia syvyysarvoja matalilla syvyyksillä kun erotus laskettiin vähentämällä satelliittikuvapohjainen syvyys luotausaineistosta (Kuva 2 D).

Syvyystietojen yhdistäminen

Yhtenäisen syvyystason luomiseksi satelliittipohjainen syvyystieto yhdistettiin luotausaineistojen kanssa. Luotausaineistot ovat pistetiheydeltään Sentinel 2 satelliittikuvan 10 metrin resoluutiota tiheämpiä, joten ensin pisteaineistoista luotiin 10 metrin resoluutiolla rasteritasot SAGA rasterize -työkalulla. Tämän jälkeen syvyysrasterit yhdistettiin seuraavassa järjestyksessä: 1. Väyläviraston avoin aineisto, 2. MH biobase-luotaus, 3. GTK:n aineisto, 4. satelliittikuvapohjainen syvyys siten, että ensin mainitun tason syvyystietoa käytettiin päällekkäisillä alueilla.



Syvyysmallin aineistot syvyyssäyrin



Kuva 10. Esimerkki syvyysmallista sekä käytettyjen aineistojen peittävydestä Hummonselän pohjoisosassa. Luotausaineisto rajoittui Hummonselän alueelle ja kattoi senkin vain osittain, joten aineistojen yhdistämisen jälkeen jääneet aukot interpoloitiin. Interpolointia varten yhdistetty syvyysrasteri muutettiin pisteiksi ja interpolointi tehtiin SAGA Multilevel B-spline-työkälulla. Koska Hummonselän ulkopuolella ainoastaan satelliittikuvapohjaista syvyystietoa oli saatavilla, yli 3 metrin syvyydet rajattiin pois aineistosta.

Yhdistetty syvyysaineisto kattaa koko Hummonselän alueen sekä matalia alueita muun Puruveden osalta. Aineistolitteeksi on indeksikartta, josta selviää mistä aineistolähteestä milläkin alueella syvyystieto on peräisin Kuva 11.



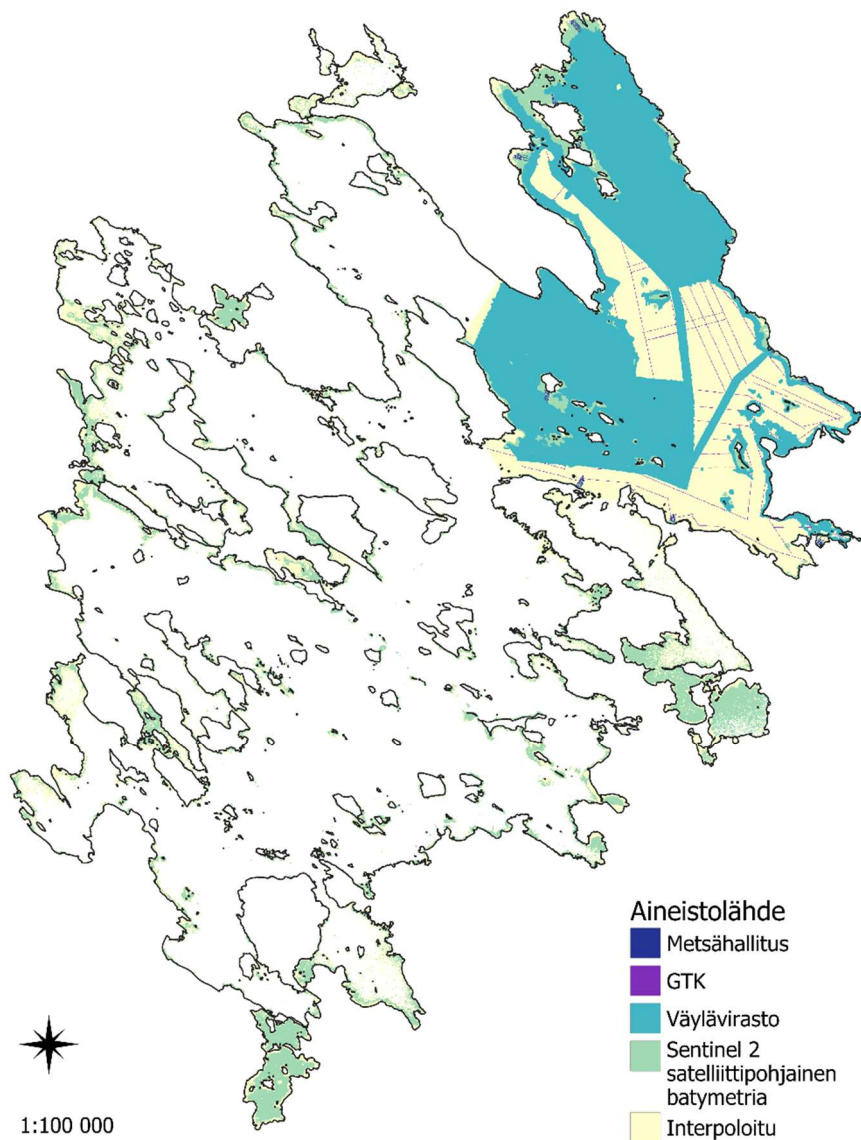
NATURA 2000

FRESHABIT

SYKE

METSÄHALLITUS
FORSTSTYRELSEN

Freshabit - Puruveden syvyyssaineiston indeksikartta



Kuva 11. Puruveden syvyyssaineiston indeksikartta. Eri väreillä on kuvattu eri aineisojen alueellinen kattavuus. Valkoiselta alueelta ei ole syvyyssaineistoa.

Kirjallisuus

Stumpf, R. P., et al. (2003). "Determination of water depth with high-resolution satellite imagery over variable bottom types." *Limnology and Oceanography* **48**(1part2): 547-556.

Vanhellemont, Q. (2019). "Adaptation of the dark spectrum fitting atmospheric correction for aquatic applications of the Landsat and Sentinel-2 archives." *Remote Sensing of Environment* **225**: 175-192.



3. Pohjanlaatu ja satelliittikuvaluokittelu

Ari-Pekka Jokinen

Summary

Sentinel-2 MSI satellite image was used to derive generic bottom types in shallow water areas for modelling purposes. The main target was to classify sandy bottoms which appear brighter compared to other bottom types due to higher reflectance. The other bottom type classes were soft dark bottoms and rocky hard bottoms. Using the field observations and echo sounding based bottom map, training and validation data was created by drawing polygons in visually representative areas for each bottom type. Additionally, depth model was used to ensure that polygons were within shallow water and NDVI to detect vegetation growing above water surface. The image was classified using random forest machine learning model. Accuracy metrics showed 84 % total accuracy for the classification. Based on visual comparison with bottom type maps the final product was limited to areas shallower than 2 meters. For modelling purposes, the classification results were exported also as classifier probabilities.

Aineiston kuvaus

Pohjatyypiluokittelussa tehtiin pohjanlaatutaso käytettäväksi kasvillisuusmalleissa, jotka on kuvattu erillisraportissa Freshabit vesikasvimallit. Lopullinen pohjanlaatutaso oli yhdistelmä luotausperusteisesta ja satelliittikuvapohjaisesta pohjanlaatutulkinnasta, jossa satelliittikuvapohjaista tulkintaa käytettiin vain matalilla alueilla. Tarkoituksena oli erityisesti erottaa hiekkapohjaiset alueet muista pohjatyypeistä näkösyvyyden alueilla.

Pohjatyypiluokittelu

Tässä raportissa kuvattu luokittelu perustuu hiekkapohjien erottamiseen niiden korkeamman albedon ja siten suuremman heijastuneen auringonsäteilyn määrään viereisiin pohjatyyppeihin verrattuna. Lähtöaineistona luokittelulle käytettiin samaa ACOLITE prosessorilla (Vanhellemont 2019) ilmakehäkorjattua Sentinel 2 satelliittikuvaa kuin syvyysmallin tuottamisessa, mistä valittiin näkyvän valon aallonpituuskanavat. Pohjatyypialueiden referenssinä käytettiin pääasiallisesti GTK:n pohjatyypiluokittelua, joka perustuu kaikuluotausaineistosta johdettuihin pohjatyyppeihin (Hämäläinen 2017). Aineistoa tarkennettiin Freshabit-lajihavaintopisteiden pohjatyypihavainnoilla, missä sukellushavainnot eivät täsmänneet GTK:n pohjaluokituksen kanssa. Pohjatyypit sellaisenaan olivat liian yksityiskohtaisia satelliittikuvapohjaiseen luokitteluun, joten luokkia sulautettiin yhteen muodostaen kolme luokkaa: hiekka, pehmeä pohja ja kova pohja (Taulukko 1).



NATURA 2000

FRESHABIT

S Y K E

METSÄHALLITUS
FORSTSTYRELSEN

Taulukko 1. Pohjatyypiluokat ja niihin sisällytetyt pohjatyypit

Luokka	Pohjatyypit
Hiekka	Hiekka (0,06–2 mm), sora (2–16 mm), pienet kivet (16–64 mm)
Pehmeä pohja	Lieju, turve, hienojakoinen aines (<0,06 mm)
Kova pohja	Isot kivet (64–250 mm), lohkaaret (250–4000 mm), kallio, moreeni (GTK)

Freshabit-pisteistä pohjatyypit määräytyi havaintopisteen suurimman kattavuuden mukaan. 50/50 tilanteessa pohjatyypit määrättiin ensimmäisen sarakkeen mukaan. Hiekan lisäksi sora ja pienet kivet (< 64 mm) luettiin samaan luokkaan hiekan kanssa. Pehmeiden pohjien luokkaan sisältyivät hienojakoinen alle 0,06 mm aines, lieju sekä turvepohjat, kun taas koviin pohjiin luettiin jäljelle jääneet moreeni, isot kivet (> 64 mm), lohkaaret sekä kallio.

Menetelmä ja tulokset

Opetus- ja testiaineisto luotiin digitoimalla polygoneja kultakin pohjatyypialueelta eri puolilta Hummonsella, mistä GTK:n pohjaluokitteluaineistoa oli saatavilla. Alueilla, joilla GTK:n pohjatyypit ei ollut yhtenevä Freshabit-maastohavaintojen kanssa, sukellettua pohjatyypitietoa käytettiin opetus- ja testiaineiston luonnissa. Lisäksi normalisoitua kasvillisuusindeksiä (NDVI) käytettiin vedenpinnan yläpuolisen kasvillisuuden tunnistamiseen ja huomioimiseen opetus- ja testiaineistossa. Syvyysmallia käytettiin rajaamaan opetus- ja testiaineisto matalille alle 4 metrin syvyyksille alueille. Opetus- ja testiaineisto käsitti kaikkiaan 27797 pikseliä, jotka jaettiin opetus- ja testipikselihin vastaavasti 70/30 suhteessa.

Luokitteluun käytettiin random forest (RF) menetelmää, jossa käytetään suurta joukkoa yksittäisiä luokittelupuualgoritmeja (decision tree, DT) määrittämään luokka pikselille ja lopullinen luokitus pikselille määräytyy näiden arvojen enemmistöstä (Breiman 2001). DT otokset suoritettiin bootstrap menetelmällä opetusaineistosta ja RF käsitti 500 luokittelupuuta. Toistetulla k-fold ristiinvalidoinnilla, jossa aineisto ositettiin 10 kertaa ja toistettiin kolmesti, RF mallin tarkkuuden keskiarvo oli 84,2 %. Konfuusiomatriisista nähdään RF mallilla luokiteltu testiaineisto sekä vertailusta lasketut tunnusluvut (Taulukko 2). Luokittelun kokonaistarkkuus oli 84 % ja tuottajan sekä käyttäjän tarkkuus eri luokissa yli 80 %.

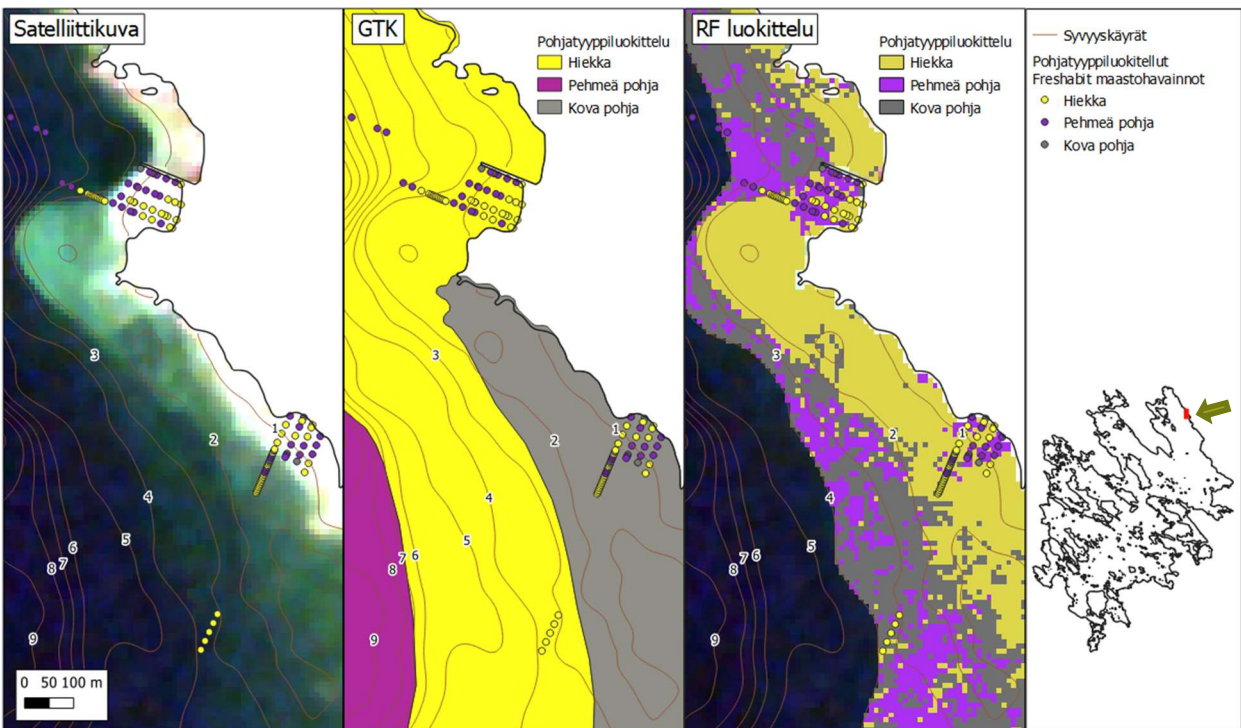
Taulukko 2. Random forest -luokittelun konfuusiomatriisi testiaineistosta sekä siitä lasketut luokittelun tunnusluvut.

	Hiekka	Pehmeä pohja	Kova pohja	Yhteensä	Käyttäjän tarkkuus
Hiekka	2409	161	337	2907	0,83
Pehmeä pohja	107	1758	311	2176	0,81
Kova pohja	205	199	2853	3257	0,88
Yhteensä	2721	2118	3501	8340	
Tuottajan tarkkuus	0,89	0,83	0,81		Kokonaistarkkuus 0,84



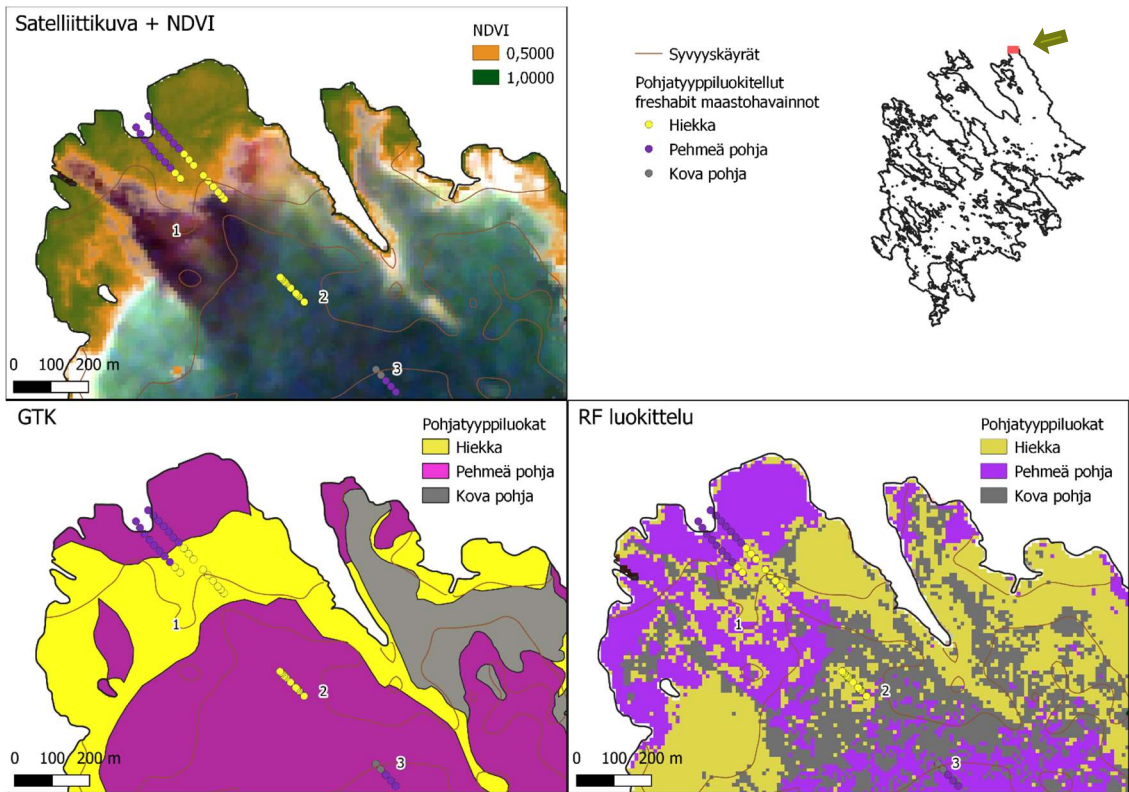
Satelliittikuvalta tehdyn pohjatyypiluokittelun (alle 4 m syvyydet) visuaalisessa tarkastelussa nähdään, että hiekkapohjiksi luokitellut alueet seuraavat satelliittikuvassa kirkkaampina näkyviä alueita (Kuva 12). Hiekkapohjien luokittelu seuraa kutakuinkin 2 metrin syvyyssävyä, mikä viittaa valon absorption vesipatsaassa kasvavan yli 2 metrin syvyyksillä niin suureksi, että pikselit muistuttavat enemmän tummasävyisiä pohjatyyppejä ja hiekkapohjan luokittelu vaikeutuu. Matalilla alueilla pehmeän pohjan luokittelu osittain sekoittuu hiekan kanssa, sillä hienojakoinen alle 0,06 mm aines (hieno siltti) luokiteltiin pehmeisiin pohjiin kuuluvaksi ja esiintyy vaihtelevissa suhteissa hiekan kanssa useilla Freshabit-havaintopisteillä.

Kuva 12 nähdään myös eroja pohjatyypiluokittelussa luotausaineistosta ja maastohavainnoista johdetun luokittelun välillä. Keskimmaisessä sarakkeessa (Kuva 12 GTK) harmaalla kuvattu alue on GTK:n luotausperusteisessa aineistossa moreenia ja luokiteltu siten kovaksi pohjaksi, kun taas Freshabit maastohavaintopisteissä pohjatyypin on määritelty havaintoruudun peittävyuden mukaan. Maastohavainnoilla täydennetty RF malli luokittelee matalan ranta-alueen tässä tapauksessa hiekkavaltaiseksi (Kuva 12 RF luokittelu).



Kuva 12. Pohjatyypiluokittelun vertailu GTK:n maalajiluokitukseen, Freshabit-havaintoihin sekä lähtöaineistona olleeseen satelliittikuvaan. Mustat ääriiviivat kuvaavat Puruveden rantaviivaa ja nuolella osoitettu punainen alue kuvien esittämän alueen.

Vedenpinnan yläpuolisesta kasvillisuudesta (esim. ruovikot) voimakkaammin heijastunut säteily voidaan virheellisesti tulkita luokittelussa hiekkapohjaksi, jos kasvillisuus jätetään huomioimatta. Kasvillisuusindekseillä (esim. NDVI) alueet voidaan kuitenkin huomioida luokittelussa, jolloin pohjanlaatu-tulkinta on sidoksissa kasvillisuustyyppille ominaiseen kasvupohjaan. Ruovikoiden tunnistamista on kuvattu tarkemmin Freshabit EO-erillisraportissa sekä Koponen et al. (2022). Freshabit-maastohavainnoissa turvepohjat ovat vallitsevia ruovikkoalueilla ja RF-luokittelun visuaalisessa tarkastelussa pehmeiksi luokitellut pohjat ranta-alueilla seuraavat melko hyvin korkeampia NDVI arvoja (Kuva 13).

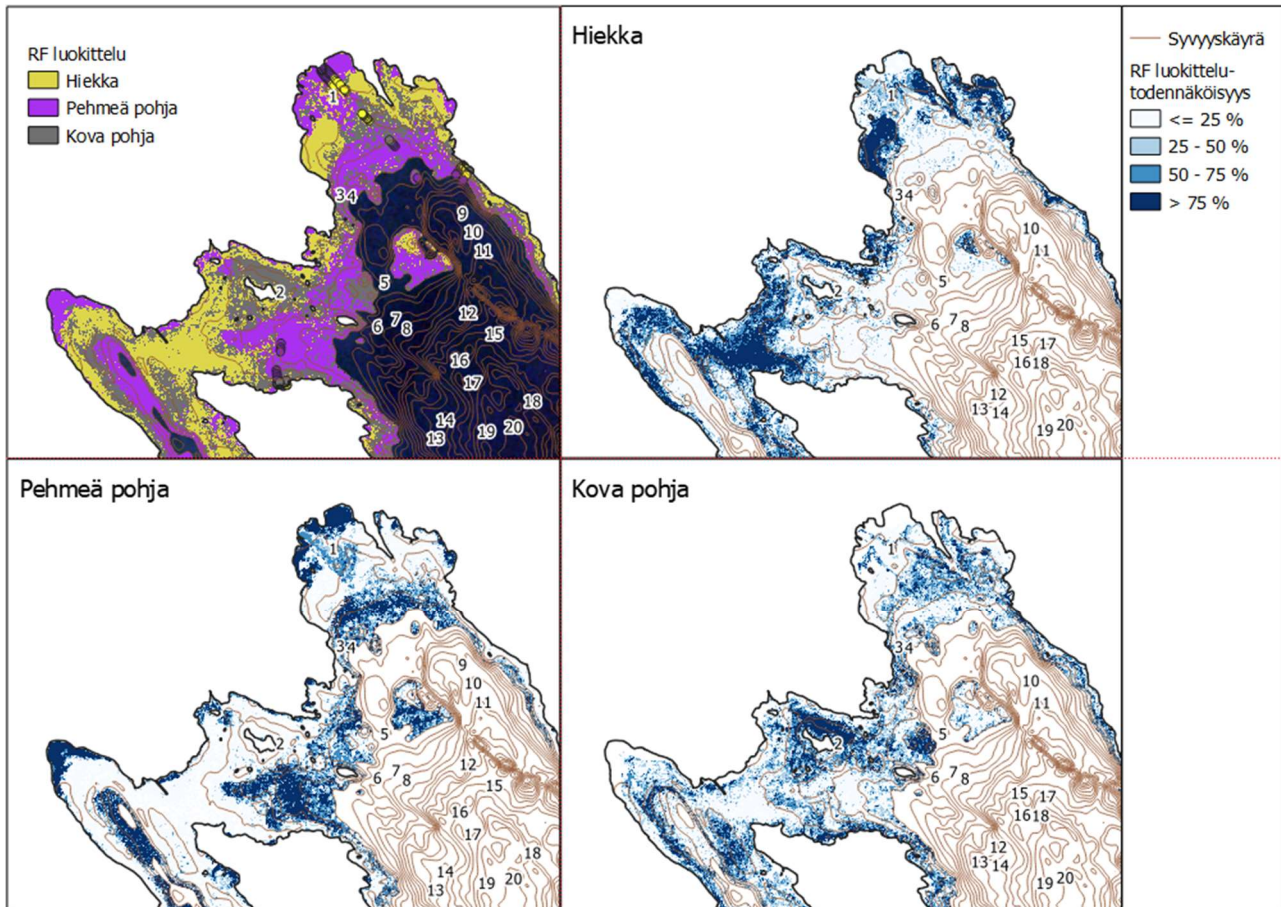


Kuva 13. Pohjatyyppiluokittelun vertailu Hummonselän pohjoisosassa. Satelliittikuvan päälle on lisätty NDVI-taso, josta nähdään ruovikkoalueet. Mustat ääriiviivat kuvaavat Puruveden rantaviivaa ja nuolella osoitettu punainen alue kuvien esittämän alueen.

Vedenalainen kasvillisuus voidaan erottaa tummempina alueina kirrkaammilta pohjilta (Vahtmäe and Kutser 2013, Wilson et al. 2022), mutta tummemmilla pohjilla vedenalainen kasvillisuus sekoittuu pohjan signaalin kanssa, ja niiden erottaminen vaikeutuu (Wilson et al. 2020). Näin ollen vedenalaisen kasvillisuuden peittävyys vaikeuttaa pohjatyyppiluokittelua. Vedenalaista kasvillisuutta ei kuitenkaan tarkasteltu tässä pohjatyyppiluokittelussa, sillä 95 % maastohavaintopisteistä kasvillisuuden peittävyys oli alle 20 % ja suurimmat peittävydet olivat ruovikoilla, missä kasvillisuus huomioitiin NDVI:llä.

Luokittelutodennäköisyydet

RF luokittelusta laskettiin kullekin pohjatyyppiluokalle tasot, jotka esittävät kuinka suuri osuus DT algoritmien tuloksista luokitteli pikselin kyseiseen luokkaan. Kuva 14 eri luokkien todennäköisyydet on visualisoitu yhdessä RF luokittelun kanssa.



Kuva 14. RF luokittelu eri pohjatyyppeihin ja kunkin luokan luokittelutodennäköisyydet.

Aineistojen yhdistäminen

Aineistojen yhdistämisessä satelliittikuvan pohjatyypiluokittelua käytettiin 2 metrin syvyyteen asti ja sitä syvemmille alueille käytettiin GTK:n luokittelua. Satelliittikuvalta luokiteltu pohjatyypipiennuste yhdistettiin GTK:n aineistosta yleistetyn tason kanssa, jotta saatiin pohjatyypitaso koko Puruveden alueelle, jossa Hummonselän alueen pohjatyypiluokittelu käsitti myös syvemmät alueet. Muualla Puruvedellä pohjatyypiluokittelu kattoi vain matalat alle 2 m syvyiset alueet.

Yhteenveto

Satelliittikuvalta luokiteltiin matalat vesialueet yleistettyihin pohjatyyppeihin: hiekkaiset, pehmeät ja kovat pohjat. Hiekkapohjien erottuvuus perustuu niiden korkeampaan valon heijastavuuteen, kun taas muut pohjatyypit näkyvät satelliittikuvalta tummempina tai vaihtelevampina. Vedenpinnan yläpuoliset kasvillisuusalueet voidaan tunnistaa kasvillisuusindekseillä ja huomioida siten luokittelussa. Tässä työssä vedenalaisen kasvillisuuden peittävyys oli pientä eikä sitä arvioitu, mutta menetelmää kehittäessä tai laajemmin vedenalaisen kasvillisuuden peittämällä pohjilla sen vaikutus olisi syytä huomioida. Syvyyden kasvaessa valoa absorboituu enemmän vesipatsaaseen ja luokittelu vaikeutuu, minkä takia luokittelutulokset rajattiin alle 2 metrin syvyyksille visuaalisen tarkastelun jälkeen. Lopullisessa pohjanlaatutasossa yhdistettiin satelliittikuvatulkinta GTK:n pohjanlaatutason kanssa, jolloin matalat ranta-alueet Hummonselän ulkopuolelta saatiin myös katettua.



Kirjallisuus

Breiman, L. (2001). "Random forests." *Machine learning* 45(1): 5-32.

Hämäläinen, J. (2017). *Konneveden ja Puruveden pohjanlaatukartat*, GTK.

Koponen, S., et al. (2022). Blue Carbon Habitats:—a comprehensive mapping of Nordic salt marshes for estimating Blue Carbon storage potential, Nordisk Ministerråd.

Vahtmäe, E. and T. Kutser (2013). "Classifying the Baltic Sea Shallow Water Habitats Using Image-Based and Spectral Library Methods." *Remote Sensing* 5(5): 2451-2474.

Vanhellemont, Q. (2019). "Adaptation of the dark spectrum fitting atmospheric correction for aquatic applications of the Landsat and Sentinel-2 archives." *Remote Sensing of Environment* 225: 175-192.

Wilson, K. L., et al. (2020). "Branching Algorithm to Identify Bottom Habitat in the Optically Complex Coastal Waters of Atlantic Canada Using Sentinel-2 Satellite Imagery." *Frontiers in Environmental Science* 8.

Wilson, K. L., et al. (2022). "Comparing Sentinel-2 and Worldview-3 imagery for coastal bottom habitat mapping in Atlantic Canada." *Remote Sensing* 14(5): 1254.

Deliverable A6.3





METSÄHALLITUS
FORSTSTYRELSEN

4. Pohjalehtisten vesikasvien lajimallien muodostaminen

Antti Takolander

Summary

Statistical species distribution models were built in Lake Puruvesi for four aquatic macrophytes, which are characteristic for oligotrophic lakes: the lake quillwort (*Isoëtes lacustris*), spiny quillwort (*Isoëtes echinospora*), shoreweed (*Littorella uniflora*) and water lobelia (*Lobelia dortmanna*). The models were built using Boosted Regression Trees (BRT) and Generalized Additive Model (GAM). The models were used to predict the probabilities of occurrence of the species studied for the entire Lake Puruvesi and Hummonselkä, which is a sub-basin located in northeast part of Lake Puruvesi. Environmental data used as explanatory variables in the models consisted of modelled environmental variables as well as variables interpreted from the satellite images as part of FRESHABIT project.

In general, the predictive capacity of the models was very good. Depth was the most important variable explaining the occurrence of the species studied, with the exception of *Isoëtes echinospora*, for which the most important explanatory variable was bottom current velocity. Based on the results of model predictions, the most important area for the occurrence of species studied is the northernmost part of the Hummonselkä sub-basin.

Johdanto

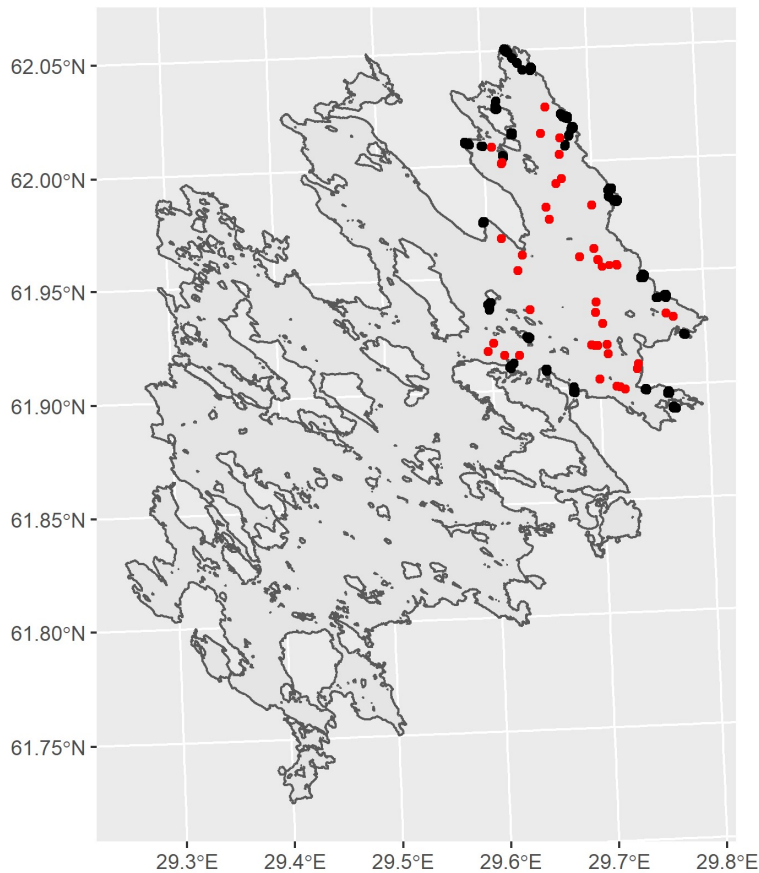
Tilastolliset habitaattimallit ovat ekologisessa tutkimuksessa laajalti käytetty menetelmä, jonka avulla voidaan saada täydentävä kuva eri lajien levinneisyyksistä kenttähavaintojen ja ympäristömuuttujien suhteen perusteella. Mallinnuksessa rakennetaan tilastollinen yhteys lajin levinneisyshavaintojen ja sitä selittävien ympäristömuuttujien välille, minkä jälkeen mallilla voidaan ennustaa lajien esiintymistodennäköisyyksiä uusilla tai kartoittamattomilla alueilla. Mallien ennusteita voidaan käyttää esimerkiksi uusien luontokartoitusten kohdentamisessa, suojelualuesuunnittelussa tai ihmistoimintojen- ja paineiden ekologisten vaikutusten arvioinnissa (Verbruggen et al. 2013, Virtanen et al. 2018).

Freshabit LIFE -hankkeen A6-työpaketissa ennustettiin tilastollisia habitaattimalleja käyttäen neljän pohjalehtisen vesikasvin levinneisyyttä Puruvedelle ja Puruveden Hummonselälle. Tutkitut lajit olivat tummalahnaruoho (*Isoëtes lacustris*), vaalealahnaruoho (*Isoëtes echinospora*), raani (*Littorella uniflora*) sekä nuottaruoho (*Lobelia dortmanna*). Lajit ovat hitaasti kasvavia, ja tunnusomaisia oligotrofisille järville (Smolders et al. 2002).

Tutkimusalueena oli Puruveden koillisosassa sijaitseva Hummonselkä (kuva 1).

METSÄHALLITUS
FORSTSTYRELSEN

Puruvesi



Kuva 1. Tutkimusalue ja kasvillisuuskartoituspisteet (mustalla). Mallinnuksessa käytettiin nollahavaintoina pisteitä, joista oli otettu pohjaeläinnäytteitä (punaisella).

Aineisto ja menetelmät

Lajien esiintymistä oli kartoitettu Metsähallituksen vuoden 2017 kesällä toteuttamilla kasvillisuuskartoituksilla. Aineistoa kerättiin kahlaamalla sekä laitesukeltamalla, ja se kattoi 22 tutkimusala, joista jokainen koostui neljästä kasvillisuuslinjasta, jotka oli satunnaistettu rannalta ulospäin 100 metrin leveyiselle alueelle siten, että jokaisen linjan väliin jäi vähintään 20 metriä. Pohjakasvillisuutta kartoitettiin linjoille sijoitetuilla 2*2m kasvillisuusruuduilla. Tarkka kuvaus aineiston keruumenetelmistä löytyy Metsähallituksen maastotyöohjeesta (Laaksonen et al. 2017).

Habitaattimallien rakentamisessa käytettiin selittävinä muuttujina syvyyttä, hiekkapohjan todennäköisyyttä, sameutta, aallokon keskimääräistä pyyhkäisymatkaa järven pinnalla (fetch), kokonaistyyppi- ja fosforipitoisuutta, etäisyyttä lähimpään ruovikkoon, veden virtausnopeutta pohjan läheisyydessä (current velocity) sekä aallokon voimaa järven pohjalla (bottom shear stress) (taulukko 1). Syvyys-, sameus- ja pohjanlaatuaineistot tuotettiin satelliittikuvista (ks. osareportit Syvyysmalli ja Pohjanlaatu- ja satelliittikuvaluokittelu). Sameudesta laskettiin keskiarvoistettu taso kesäkuukausille (kesä – syyskuu) vuosien 2017 - 2021 kuukausikomposiittien perusteella.

Milestone A6



METSÄHALLITUS
FORSTSTYRELSEN

Satelliittiaineistosta ei voitu tulkita luotettavasti syvyyttä kuin kolmeen metriin saakka. Tämän vuoksi Hummonselällä syvyystietoa täydennettiin interpoloimalla Väyläviraston tekemiä syvyysluotauksia, sekä Metsähallituksen kesällä 2017 suorittamia BioBase-tutkimusluotauksia. Näin toimimalla pystyttiin Hummonselälle tuottamaan melko luotettava spatiaalinen syvyystaso.

Mallien kalibrointivaiheessa käytettiin sukeltajan kasvillisuusruudulle kirjaamaa syvyyttä, kun taas spatiaaliset ennusteet Hummonselälle ja Puruvedelle tehtiin käyttäen satelliittitulkittua syvyystietoa. Ympäristömuuttujien arvot poimittiin spatiaalisista tasoista käyttäen R:n ”raster”-pakettia. Kaikki ympäristömuuttujatasot eivät ulottuneet aivan rantaan saakka, ja erityisesti bottom shear stressin ja current velocityn kohdalla ennusteet eivät ajoittain peittäneet kaikkia lajihavaintoja, jotka keskittyivät usein varsin lähelle rantaa. Mikäli kasvillisuusruutu ei osunut mihinkään ympäristömuuttujatason rasterisolun, annettiin sille lähimmän dataa sisältävän rasterisolun arvo. Habitaattimallien spatiaalisia ennusteita varten kaikki ympäristömuuttujatasot skaalattiin 10 x 10 metrin resoluutioon.

Taulukko 1. Habitaattimalleissa käytetyt ympäristömuuttujat ja niiden tietolähteet ja resoluutio.

Muuttuja	Lähde	Alkuperäinen resoluutio
Syvyys	Satelliittiaineisto + luotaus	10 x 10 m
Hiekkapohjan todennäköisyys	Satelliittitulkinta	10 x 10 m
Sameus	Satelliittitulkinta (Sentinel-2 MSI)	10 x 10 m
Fetch	Jyväskylän yliopisto	5 x 5 m
Kokonaistyyppi & Kokonaisfosfori	VESLA-mittaukset, interpoloitu	5 x 5 m (interpolointi)
Curent velocity	3D-COHERENS-malli	50 x 50 m
Bottom shear stress	3D-COHERENS-malli	50 x 50 m
Etäisyys lähimpään ruovikkoon	Satelliittitulkinta (Sentinel-2 MSI)	10x 10 m

Habitaattimallit rakennettiin käyttäen Boosted Regression Trees (BRT) -algoritmia (Elith et al. 2008). Lähtöaineistoina käytettiin kasvillisuuskartoituksen tuloksia, joista neljän tutkimuslajin erilliset peittävydet muutettiin presence-absence-havainnoiksi. Kaikki yli nollan peittävydet tulkittiin lajin esiintymäksi (presence). Lisäksi havaintoaineistoa täydennettiin 44 pohjaeläinnäytepisteellä, jotka sijaitsivat yli 8 metrin syvyydessä (kuva 1, punaisella) luotettavan



syvyysvasteen tuottamiseksi lajimalleihin. Koska kasvillisuusaineiston perusteella havaittiin, että suurin tutkittavien lajien esiintymissyvyys oli 6 metriä, kaikki pohjaeläinpisteet merkittiin puuttuviksi havainnoiksi (absence).

Tummalahnaruohosta oli niin vähän havaintoja ($N = 43$), että toimivan BRT-mallin rakentaminen osoittautui hankalaksi. Tämän vuoksi tummalahnaruohon osalta päädyttiin käyttämään yksinkertaista, kahden selittävän muuttujan GAM-mallia (Generalized Additive Model (Hastie and Tibshirani 1990)). Samaa mallia käytettiin levinneisyyden ennustamiseen koko Puruvedelle.

Ympäristömuuttujien arvot Hummonselän ulkopuolella poikkesivat joidenkin muuttujien osalta melko paljon Hummonselän arvoista. Habitaattimallien ennustekyky saattaa heikentyä huomattavasti, mikäli malleja käytetään ekstrapoloimaan kauas alkuperäisen ympäristömuuttuja-avaruuden ulkopuolelle (esim. (Yates et al. 2018, Charney et al. 2021)). Tämän vuoksi tutkimuksessa rakennettiin 2 erillistä mallisarjaa: yksi Hummonselälle käyttäen pääasiallisesti kaikkia käytettävissä olevia ympäristömuuttujia, ja toinen, jolla ennustettiin lajien levinneisyyttä koko Puruvedelle. Jälkimmäisistä malleista pudotettiin pois kokonaistyyppi ja -fosfori sekä current velocity ja bottom shear stress lukuun ottamatta vaalealahnaruohon GA-mallia, jossa samaa kahden muuttujan mallia käytettiin sekä Hummonselälle että koko Puruvedelle. Molemmissa tapauksissa muuttujien tärkeyttä arvioitiin tarkastelemalla mallien ennustekykyä (predictive deviance) takaisinpäin askeltavalla valinta-algoritmilla. Mikäli mallin ennustekyky ei oleellisesti heikentynyt muuttujan poisjättämisen seurauksena, se pudotettiin pois mallista. Vaalealahnaruohon GAM-mallissa muuttujien valinta tehtiin samoin takaisinpäin askeltavana, ja muuttujien merkittävyyttä arvioitiin AIC:n (Akaike Information Criteria) avulla.

Mallin hyperparametrit (tree complexity, bag fraction, interaction depth) valittiin 10-kertaisen ristiinvalidoinnin tulosten perusteella. Näin saadun mallin osuvuutta arvioitiin AUC-arvolla 10-kertaisella ristiinvalidoinnilla, jossa mallit rakennettiin käyttäen satunnaisesti valittua 70% osuutta alkuperäisestä datasta ja validoitiin 30% osuudella. Kaikkea kasvillisuuskartoitusaineistoa käytettiin lopullisen ennustemallin tuottamiseen.

Tulokset

Mallien ennustekyky

Ristiinvalidoinnin tulosten perusteella mallien ennustekyky oli hyvä (taulukko 2 ja 3). Ennustekyky oli keskimäärin hieman alhaisempi koko Puruveden malleille, mikäli oli odotettavissa, koska selittäviä muuttujia oli näistä malleista pudotettu pois. Vaalealahnaruoho (*I. echinospora*) sai muita lajeja alhaisempia AUC-arvoja (8.75, taulukko 2), mikä johtui todennäköisesti tämän lajin kohdalla käytetystä yksinkertaisemmasta mallista. GAM-malli oli kuitenkin ennustekyvyltään tarkempi kuin *I. echinosporalle* ensin testattu BRT-malli (AUC 8.75 vs AUC 8.21).

Milestone A6



METSÄHALLITUS
FORSTSTYRELSEN

Taulukko 2. Ristiinvalidoinnin tulokset (AUC) lajeittain Hummonselälle sekä koko Puruvedelle rakennetuille malleille.

Alue	Malli	Laji	AUC	sdAUC
Hummonselkä	GAM	<i>I.echinospora</i>	0.875	0.018
Hummonselkä	BRT	<i>I.lacustris</i>	0.915	0.011
Hummonselkä	BRT	<i>L.dortmanna</i>	0.934	0.011
Hummonselkä	BRT	<i>L.uniflora</i>	0.933	0.014
Koko Puruvesi	GAM	<i>I.echinospora</i>	0.875	0.018
Koko Puruvesi	BRT	<i>I.lacustris</i>	0.903	0.023
Koko Puruvesi	BRT	<i>L.dortmanna</i>	0.924	0.014
Koko Puruvesi	BRT	<i>L.uniflora</i>	0.925	0.021

Mallien ennustekykyä arvioitiin myös ristiintaulukoimalla ennustetut esiintyvyydet (presence-absence) sekä lajien esiintyvyydet (taulukko 3). Tätä varten mallien ennustamat esiintymistodennäköisyydet muutettiin presence-absence-ennusteiksi määrittämällä raja-arvo, joka maksimoi Cohenin Kappa -suureen. Mallien pääasiallinen ennustekyky oli hyvä. Vaalealahnaruohon osalta GAM-malli ennusti lajin esiintyvyyttä useisiin paikkoihin, joissa havaintoja ei ollut (taulukko 3, kuva 7a), mikä johtui osaltaan siitä, että alkuperäisiä havaintoja lajista oli suhteellisen vähän.

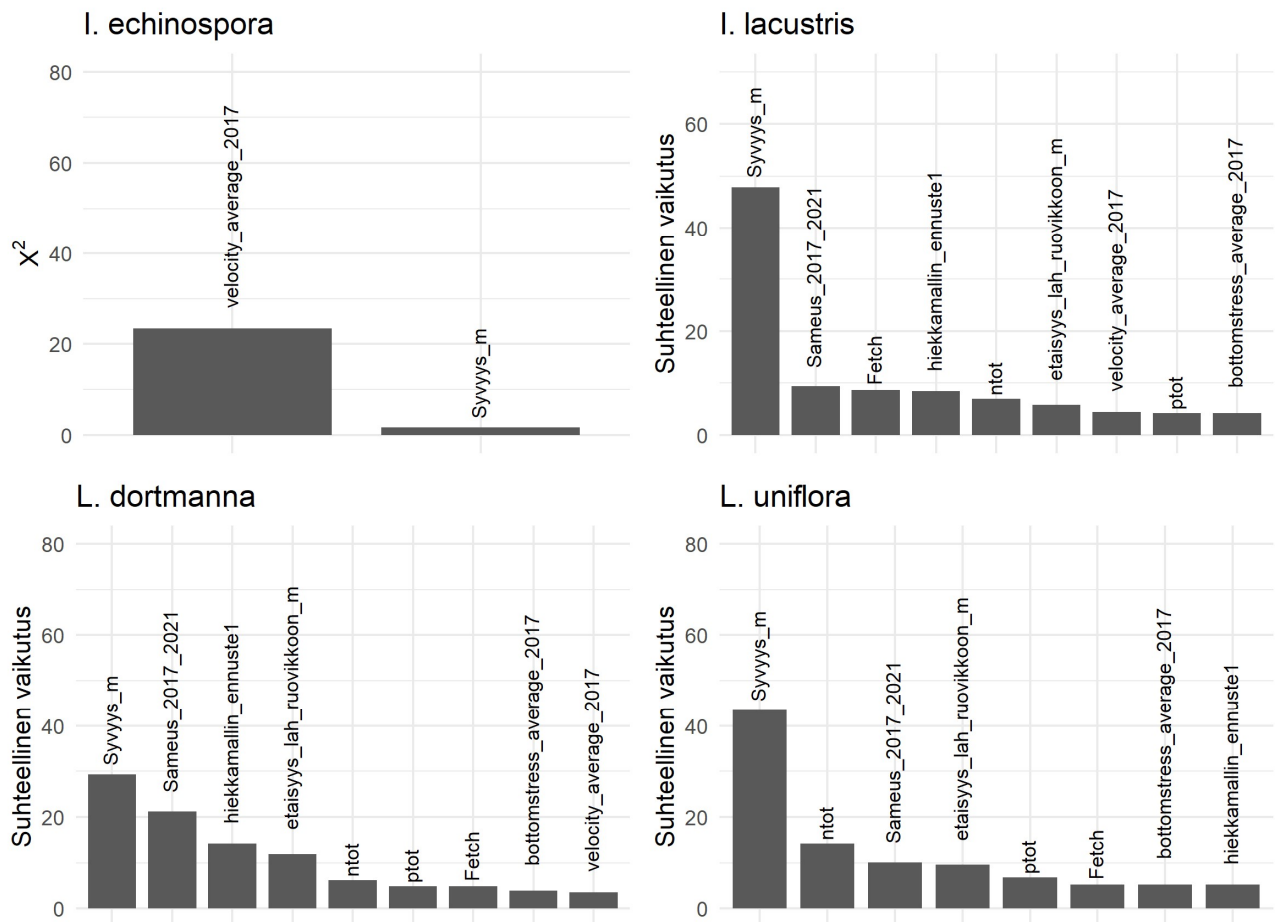
Taulukko 3. Ennustettu ja havaittu esiintyminen eri lajeilla Hummonselän alueella. N on keskiarvoistettu tulos 10-kertaisesta ristiinvalidoinnista, jossa alkuperäinen data on jaettu kalibrointi- ja validointiaineistoon suhteessa 70/30, ja rakennetulla mallilla ennustettu esiintyvyyttä 30 % validointiaineistoon (N = 255 / kasvillisuusruutua / laji).

Havaittu	Ennustettu	N Laji
0	0	178 <i>L.uniflora</i>
0	1	21 <i>L.uniflora</i>
1	0	7 <i>L.uniflora</i>
1	1	49 <i>L.uniflora</i>
0	0	151 <i>I.lacustris</i>
0	1	20 <i>I.lacustris</i>
1	0	14 <i>I.lacustris</i>
1	1	70 <i>I.lacustris</i>
0	0	152 <i>L.dortmanna</i>
0	1	18 <i>L.dortmanna</i>
1	0	11 <i>L.dortmanna</i>
1	1	74 <i>L.dortmanna</i>
0	0	173 <i>I.echinospora</i>
0	1	68 <i>I.echinospora</i>
1	0	1 <i>I.echinospora</i>
1	1	12 <i>I.echinospora</i>



Muuttujien tärkeys ja lajien vasteet ympäristömuuttujiin

Tärkeimmät lajien esiintyvyyttä selittävät muuttujat vaihtelivat lajeittain, mutta syvyys oli tärkeä selittävä tekijä lähes kaikkien lajien esiintyvyydelle (kuva 2). Muita tärkeitä muuttujia olivat aallokon pyyhkäisymatka (fetch) sekä sameus, kun taas pohjatyyppin vaikutus oli vastoin ennako-odotuksia suhteellisen pieni.



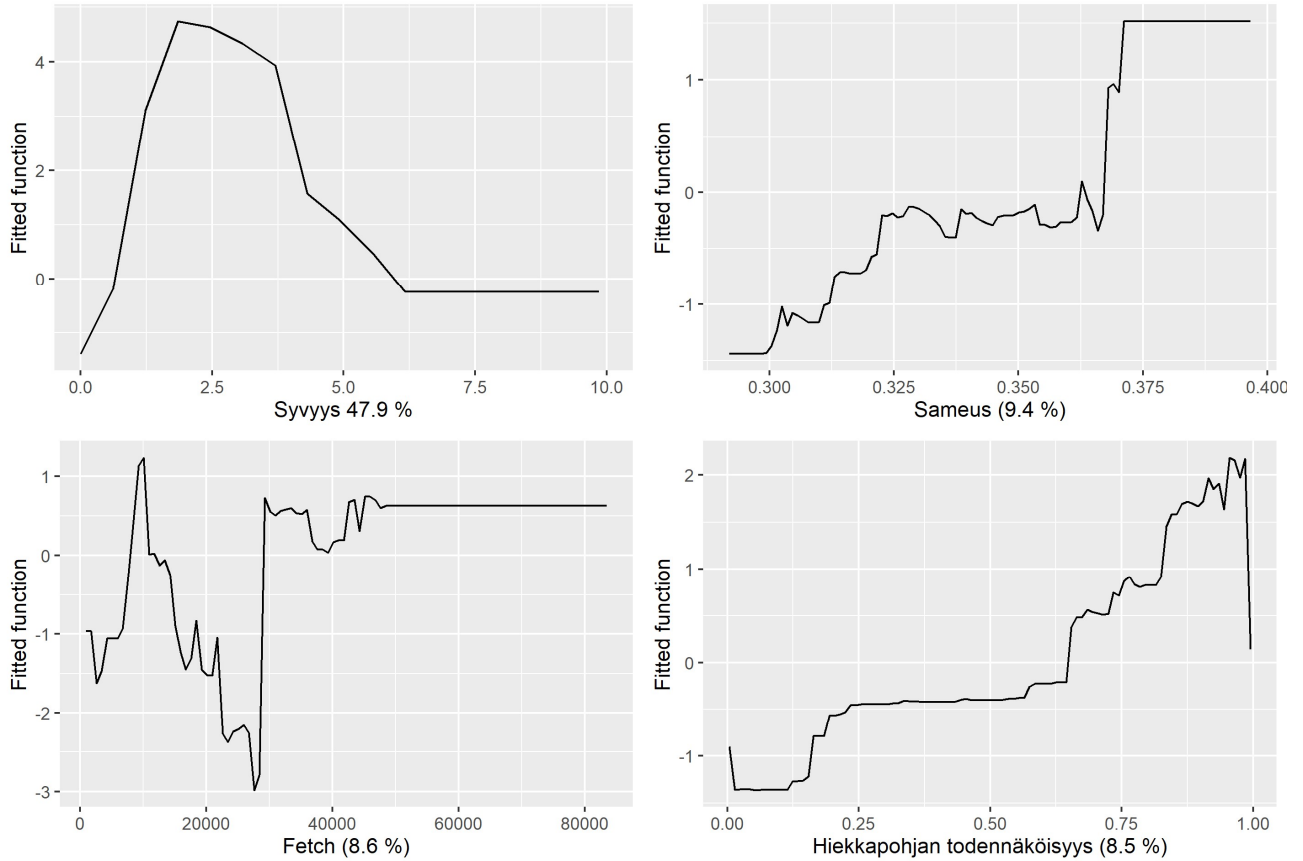
Kuva 2. Muuttujien suhteellinen tärkeys eri lajien levinneisyyksille Hummonselällä. Vaalealahnaruoholle tehdyssä GAM-mallissa muuttujien tärkeyttä kuvataan X²-testisuurella.

Vasteet ympäristömuuttujiin: Isoetes lacustris

Syvyys oli tärkein tummalahnaruohon esiintyvyyttä selittävä tekijä, ja esiintymistodennäköisyys oli suurimmillaan 1 – 5 metrin syvyydellä (kuva 3). Seuraavaksi tärkeimmät oli satelliittikuvista tulkittu hiekkapohjien esiintymistodennäköisyys sekä sameus. Tummalahnaruohoa esiintyi todennäköisemmin ruuduilla, joilla hiekkapohjien esiintymistodennäköisyys oli suurempi kuin 0.6 Sameuden lisääntyessä tummalahnaruohon esiintymistodennäköisyys nousi. Tämä saattaa johtua siitä, että satelliittikuvatulkinnassa suuret sameusarvot ovat pohjasta tulleita heijastusta, jolloin kyseiset ruudut ovat ensisijaisesti melko matalia. Esiintymistodennäköisyyden vaste tuulen pyyhkäisymatkaan (fetch) oli monitulkintaisempi, mutta



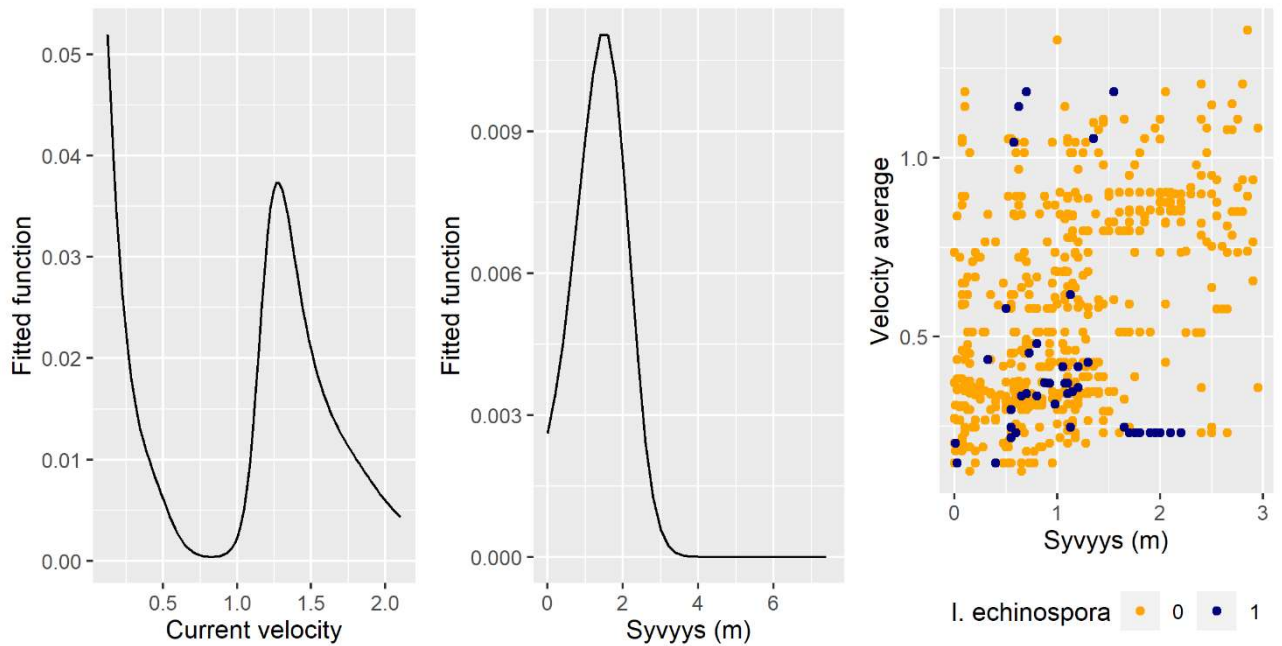
pääsääntöisesti mallin perusteella laji vaikuttaa esiintyvän matalilla, melko avoimilla rannoilla, joissa on hiekkapohja.



Kuva 3. Neljän tärkeimmän ympäristömuuttujan vaikutus tummalahnaruohon (*I. lacustris*) esiintymistodennäköisyyteen Hummonselälle rakennetussa BRT-mallissa.

Vasteet ympäristömuuttujiin: *Isoetes echinospora*

Vaalealahnaruohon esiintymiseen eniten vaikuttaviksi muuttujiksi valikoitui GAM-malliin veden virtausnopeus pohjan lähellä (current velocity) sekä syvyys (kuva 4). Mallin ennusteet perusteella vaalealahnaruoho esiintyy 0 – 2,5 metrin syvyydessä. Vaalealahnaruohon vaste veden virtaukseen oli kaksijakoinen. Pääosa havainnoista keskittyi paikoille, jossa virtausnopeus oli alhainen, mutta lajista tehtiin havaintoja myös paikoilla, joissa virtausnopeus oli suhteellisen suuri.

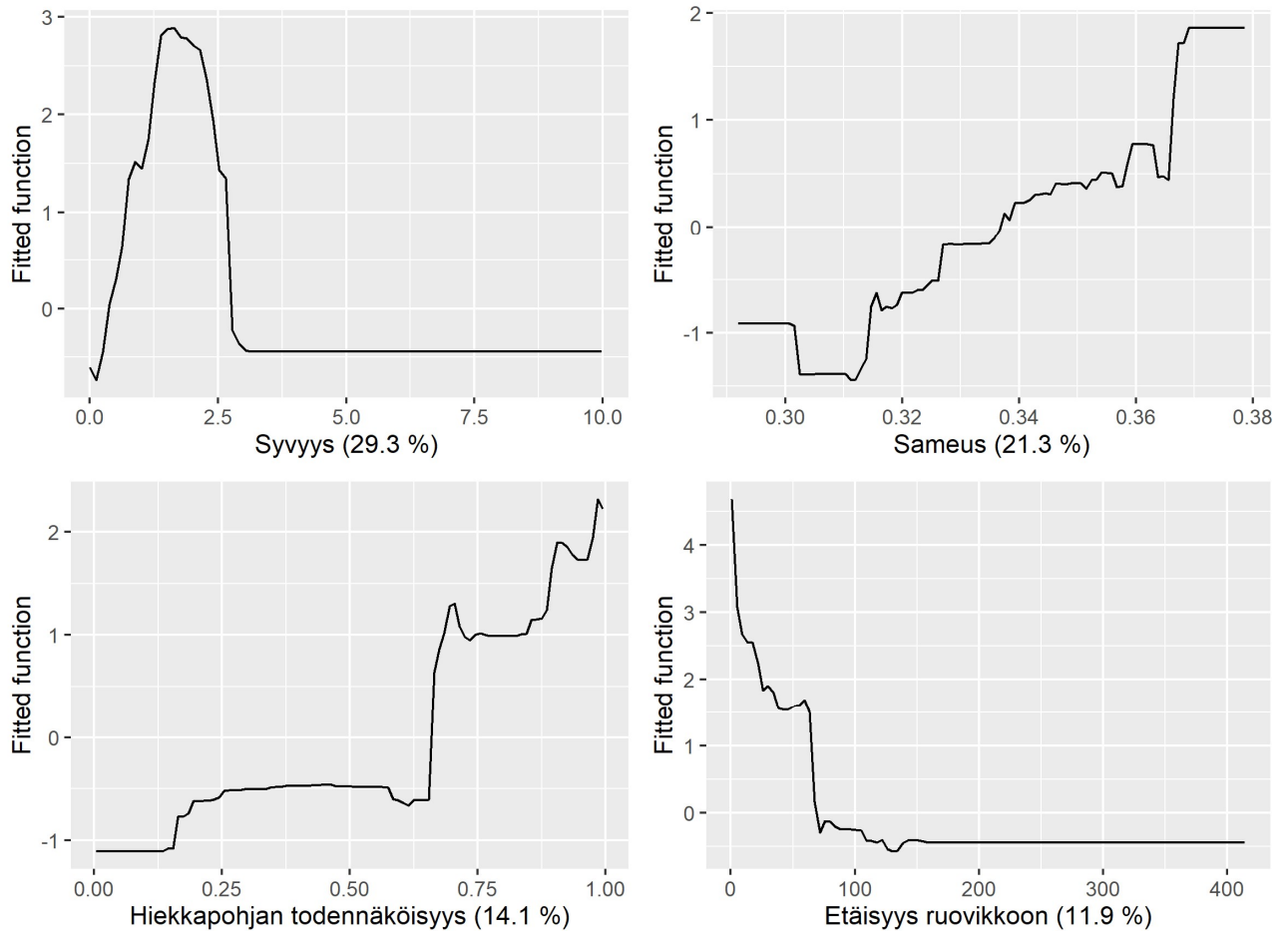


Kuva 4. Ympäristömuuttujien vaikutus vaalealahnaruohon (*I. echinospora*) esiintymistodennäköisyyteen GAM-mallissa sekä eri syvyyksillä ja pohjan virtausnopeuksilla tehnyt havainnot vaalealahnaruohon esiintymisestä (arvot 1).

Vasteet ympäristömuuttujiin: *Lobelia dortmanna*

Tärkeimmät nuottaruohon esiintymistodennäköisyyteen vaikuttavat tekijät BRT-mallissa olivat syvyys ja sameus, ja näiden lisäksi hiekkapohjan esiintymistodennäköisyys sekä etäisyys lähimpään ruovikkoon (kuva 5). Mallin perusteella nuottaruohoa esiintyy matalilla (alle 2,5 m) hiekkapohjilla ruovikoiden laidoilla. Kuten edellä todettiin, satelliittikuvista tulkittu sameusrasteri sai joillakin matalilla alueilla melko korkeita arvoja johtuen hiekkapohjista tulevasta heijastuksesta. Sameuden suhteellisen suuri vaikutus BRT-mallissa on tulkittavissa tästä johtuvaksi tekijäksi, eli hiekkapohjat, jotka ovat matalilla syvyyksillä (alle 3 m) saattavat saada korkeita sameusarvoja satelliittitulkinnassa. Toisaalta matalat, lähellä rantaa ja ruovikoita olevat alueet ovat keskimäärin sameampia kuin monet muut. BRT-malli ennustaa korkeita esiintymistodennäköisyyksiä näille alueille, mutta tämä johtunee todennäköisesti siitä, että nämä pohjat ovat nuottaruohon luonnollista habitaattia, kuten muiden kuvan 5 muuttujien ja esiintymistodennäköisyyden suhteesta voidaan päätellä.

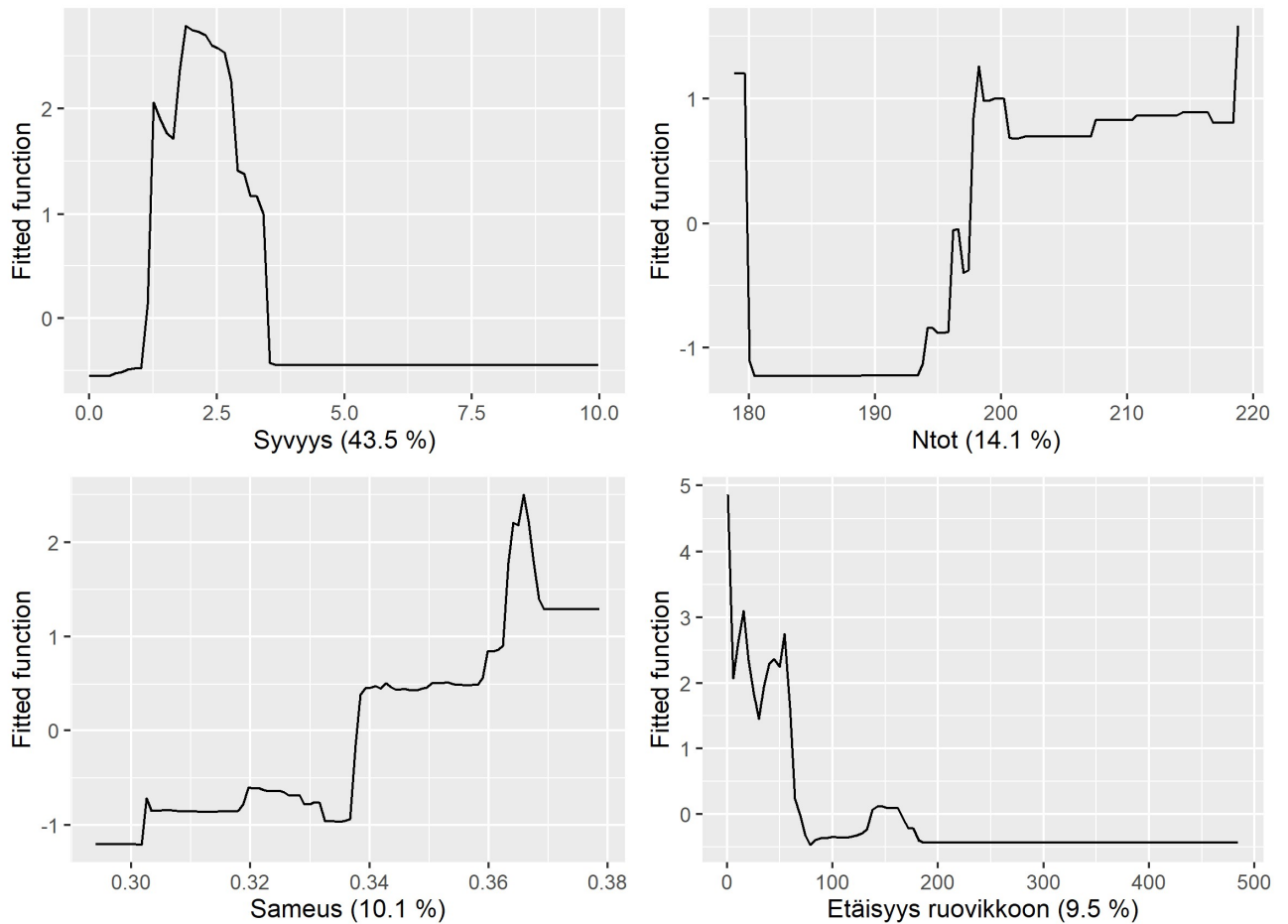
Milestone A6



Kuva 5. Neljän tärkeimmän ympäristömuuttujan vaikutus nuottaruohon (*L. dortmanni*) esiintymistodennäköisyyteen Hummonselälle rakennetussa BRT-mallissa.

Vasteet ympäristömuuttujiin: Littorella uniflora

BRT-mallissa raanin esiintymistodennäköisyyteen vaikutti enimmäkseen syvyys (43,5 %). Muiden ympäristömuuttujien vaikutus oli suhteellisen pieni (kuva 6). Vasteiden perusteella voidaan todeta, että mallin ennusteessa raani esiintyy melko matalilla rannoilla 1 – 3,5 metrin syvyydessä ruovikoiden läheisyydessä. Kokonaistyyppitaso tuotettiin interpoloimalla VESLA-tietokannasta ladatuista vedenlaatumittauksista. Koska mittauksia on varsin rajallisesti, interpoloinnin alueellinen tarkkuus on karkea, ja siten kokonaistyyppien vaikutuksen tulkinta hankalaa. Vaikuttaa kuitenkin siltä, että kokonaistyyppi oli interpoloidussa GIS-tasossa hieman korkeampi matalilla alueilla, joilla raania esiintyi. Kausaalista suhdetta näiden välillä ei kuitenkaan välttämättä ole.



Kuva 6. Neljän tärkeimmän ympäristömuuttujan vaikutus raanin (*L. uniflora*) esiintymistodennäköisyyteen Hummonselälle rakennetussa BRT-mallissa.

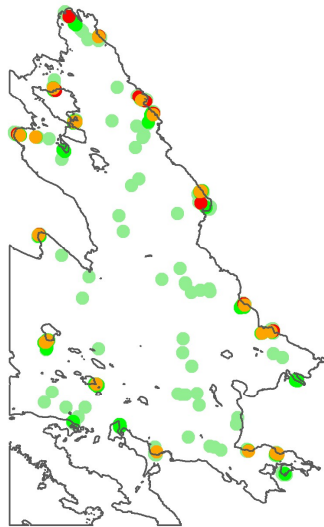
Ennustettujen esiintymisten spatiaalinen jakautuminen Hummonselällä

Ennustettua levinneisyyttä ja lajihavaintoja tarkasteltaessa ei esiin noussut merkittäviä alueellisia säännönmukaisuuksia (kuva 7a ja 7b) ennustekyvyyssä. Kaikki mallit ennustivat oikein Hummonselän keski-osan syvänteet (absence), mikä oli odotettavissa. Vaalealahnaruoholla GAM-malli ennusti hieman muita lajeja enemmän false positive -tilanteita, eli lajeja ei ollut havaittu ruudussa mutta malli ennusti kuitenkin sen löytyvän. Tämä saattaa johtua käytetystä Kappa-raja-arvosta, joka valittiin siksi, että mallin suorituskyky saatiin hyväksi true positive -havaintojen osalta.

Milestone A6

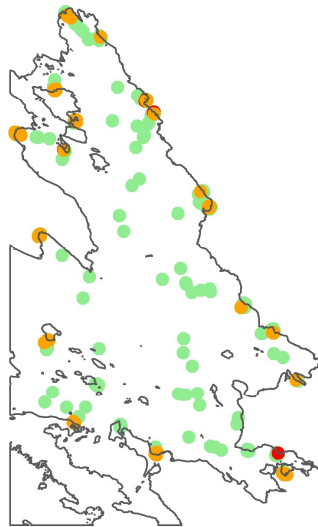


I. lacustris



- False negative
- False positive
- True negative
- True positive

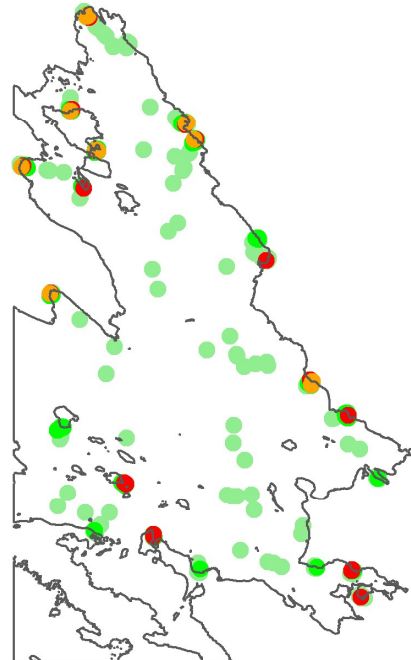
I. echinospora



- False negative
- False positive
- True negative
- True positive

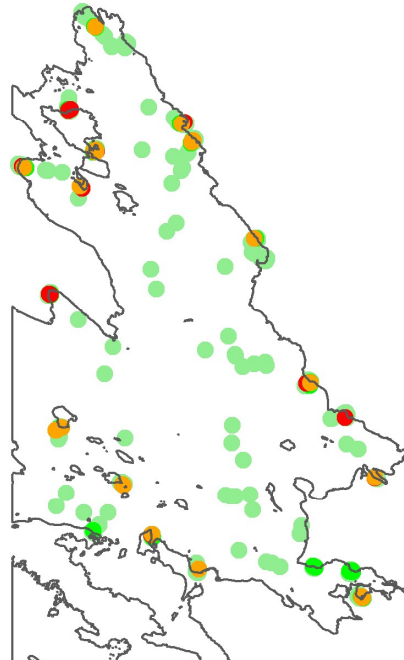
Kuva 7a. Lahnaruohojen ennustetut esiintyvyyssuokittelut tutkimusaloilla. False negative = laji on havaittu, mutta ei ennustettu, false positive = laji on ennustettu, mutta ei havaittu, true negative = malli ennustaa laji ei esiinny ruudussa, eikä sitä havaittu, true positive = malli ennustaa lajin esiintyvän ruudussa jossa se havaitaan.

L. dortmanna



- False negative
- False positive
- True negative
- True positive

L. uniflora



- False negative
- False positive
- True negative
- True positive

Kuva 7b. Nuottaruohon ja raanin ennustetut esiintyvyyssuokittelut tutkimusaloilla. Havaintojen värikoodit avattu kuvan 7a kuvatekstissä.

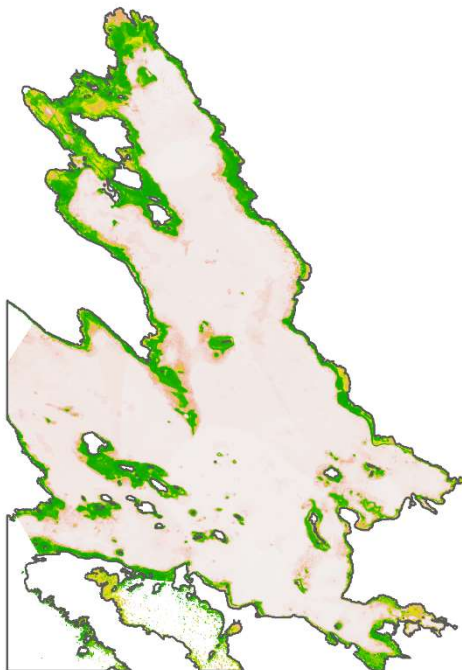


METSÄHALLITUS
FORSTSTYRELSEN

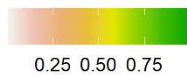
Lajien ennustetut esiintymistodennäköisyydet Hummonselälle

Mallit ennustivat kaikille lajeille melko suurta esiintymistodennäköisyyttä Hummonselän pohjois- ja luoteisosaan, jossa syvyys oli matala ja aallokon pyyhkäisymatka saarien suojassa maltillinen (kuva 8 ja 9). Tummalahnaruoho sai lähes kautta koko Hummonselän korkeita esiintymistodennäköisyyksiä ranta-alueille (kuva 8). Johtuen vaalealahnaruohon (*I. echinospora*) alhaisesta esiintyvyydestä lähtödatassa, sen ennustetut esiintymistodennäköisyydet olivat kauttaaltaan matalia (kuva 9). Vaalealahnaruohon ennustetut todennäköisimmät esiintymisalueet painottuivat Hummonselän pohjoisosaan mataliin, suojaisiin rantoihin (kuva 8).

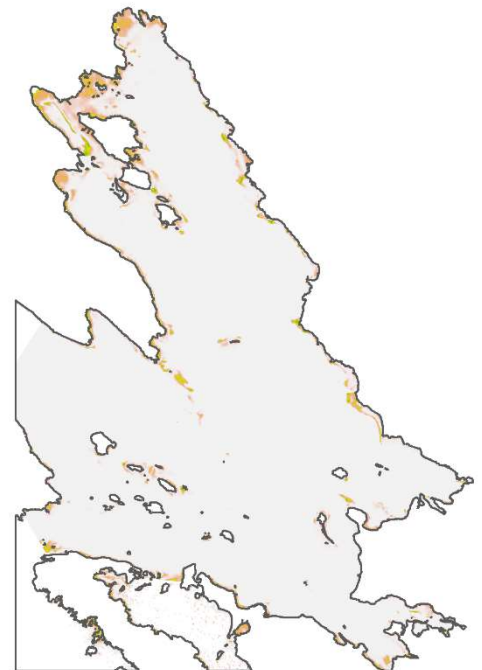
I. lacustris



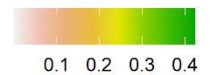
Todennäköisyys



I. echinospora



Todennäköisyys



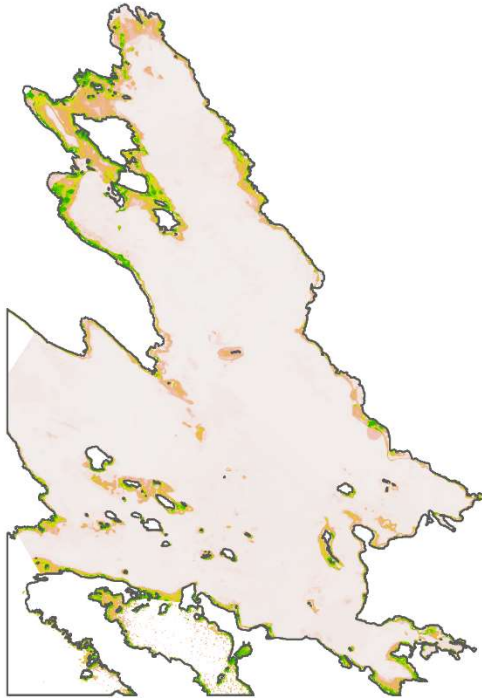
Kuva 8. Ennustetut tummalahnaruohon (*I. lacustris*) ja vaalealahnaruohon (*I. echinospora*) esiintymistodennäköisyydet Hummonselällä.

Milestone A6



METSÄHALLITUS
FORSTSTYRELSEN

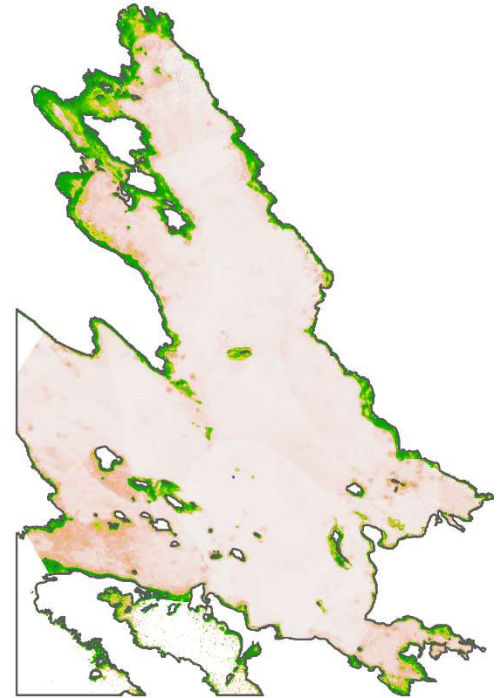
L. uniflora



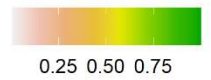
Todennäköisyys



L. dortmanna



Todennäköisyys



Kuva 9. Ennustetut nuottaruohon (*L. dortmanna*) ja raanin (*L. uniflora*) esiintymistodennäköisyydet Hummonselällä.

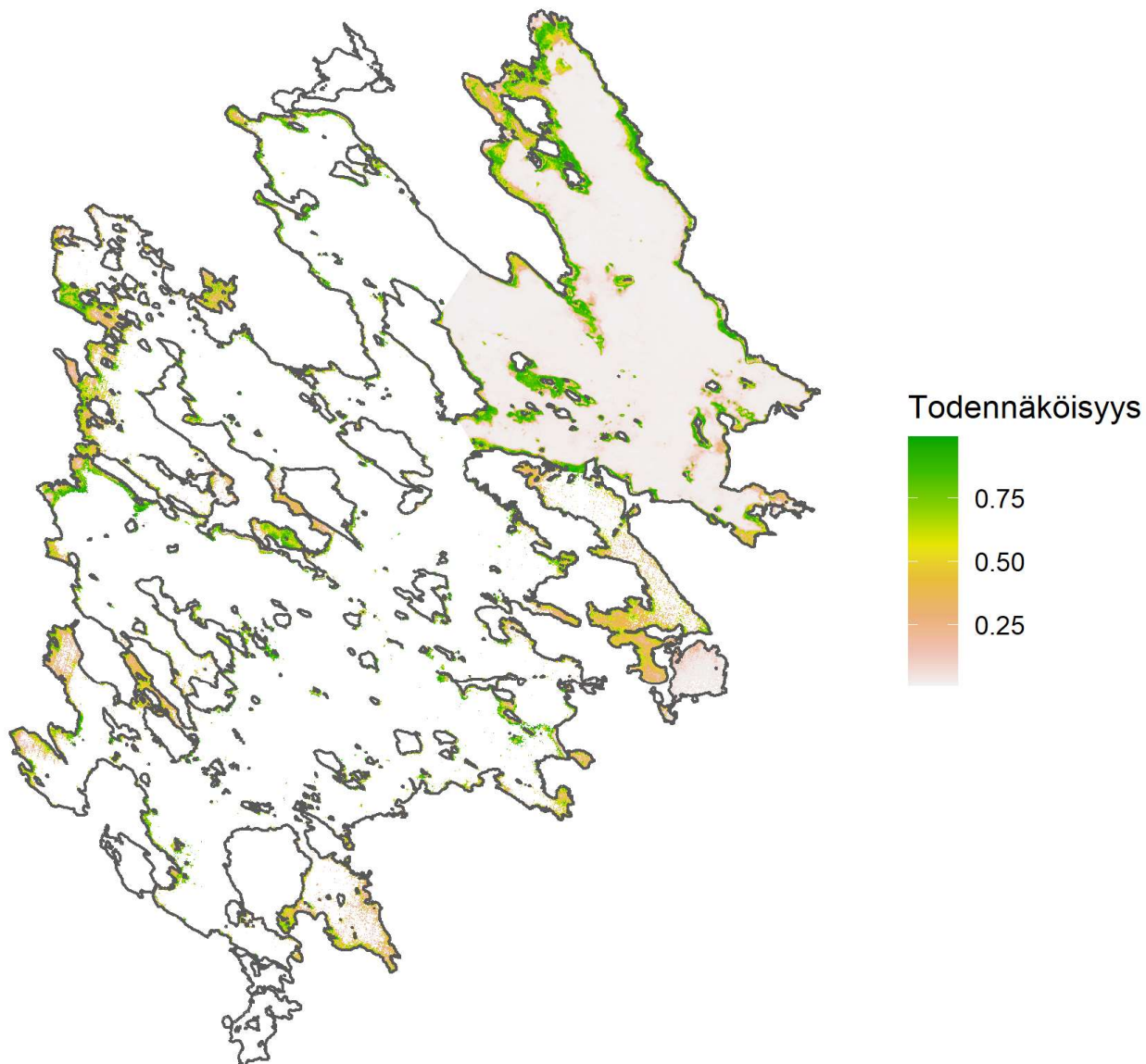
Spatiaaliset ennusteet koko Puruvedellä

Koska satelliittiaineistoista tulkitut syvyystiedot eivät olleet luotettavia kuin varsin matalilla (< 3m) alueilla, tehtiin ennusteet Hummonselän ulkopuolella ainoastaan alle 3m syvyisille alueille. Mallit ennustivat lajien esiintyvyyksiä pääpiirteittäin samoille matalille ja suojaisille alueille (kuvat 10 – 13), jotka ovat tulosten perusteella potentiaalisesti kiinnostavia kohteita tulevien vuosien luontokartoituksia silmälläpitäen.



METSÄHALLITUS
FORSTSTYRELSEN

I. lacustris

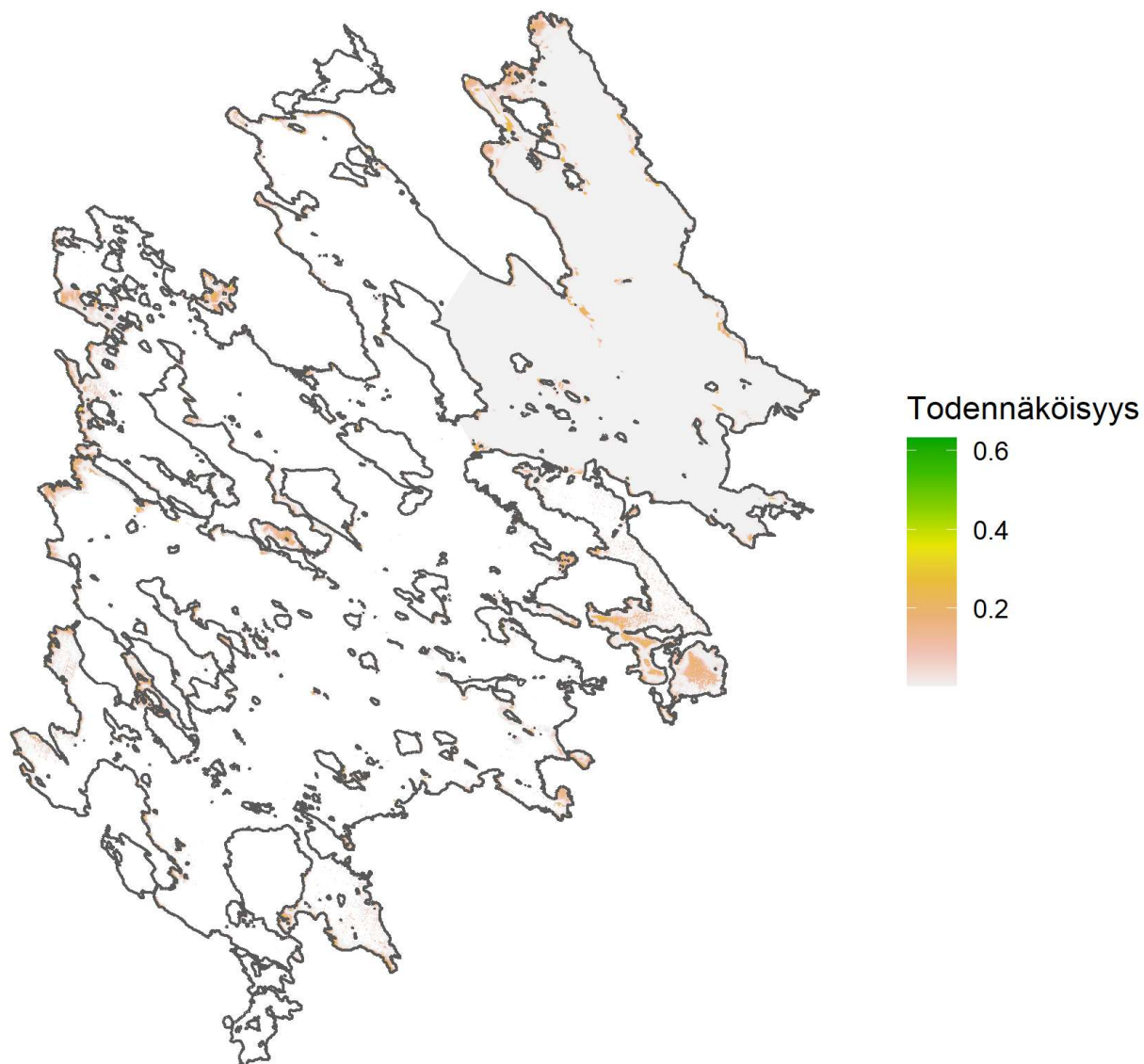


Kuva 10. Ennustetut tummalahnaruohon esiintymistodennäköisyydet Puruvedellä.



METSÄHALLITUS
FORSTSTYRELSEN

I. echinospora

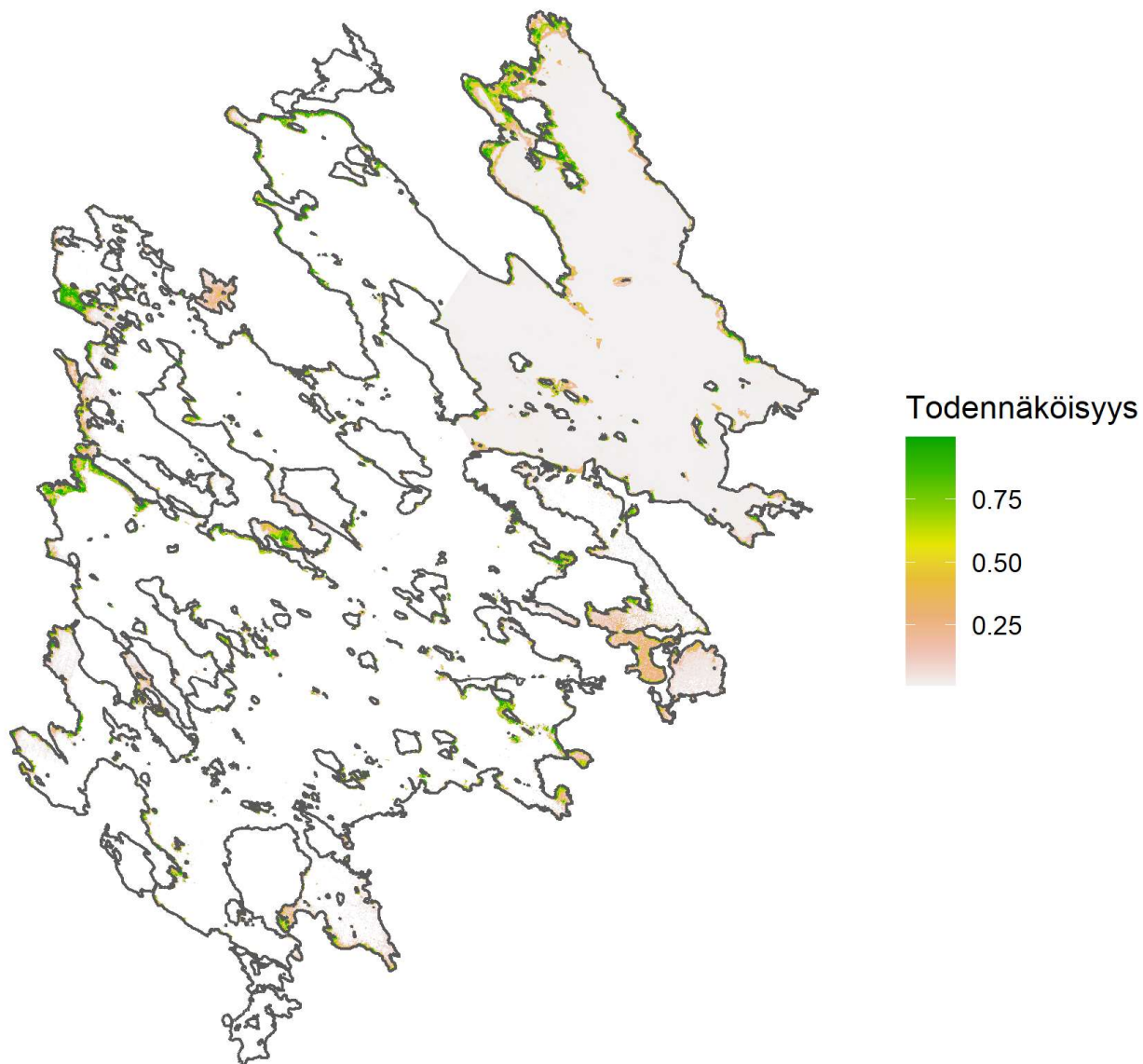


Kuva 11. Ennustetut vaalealahnaruohon esiintymistodennäköisyydet Puruvedellä.



METSÄHALLITUS
FORSTSTYRELSEN

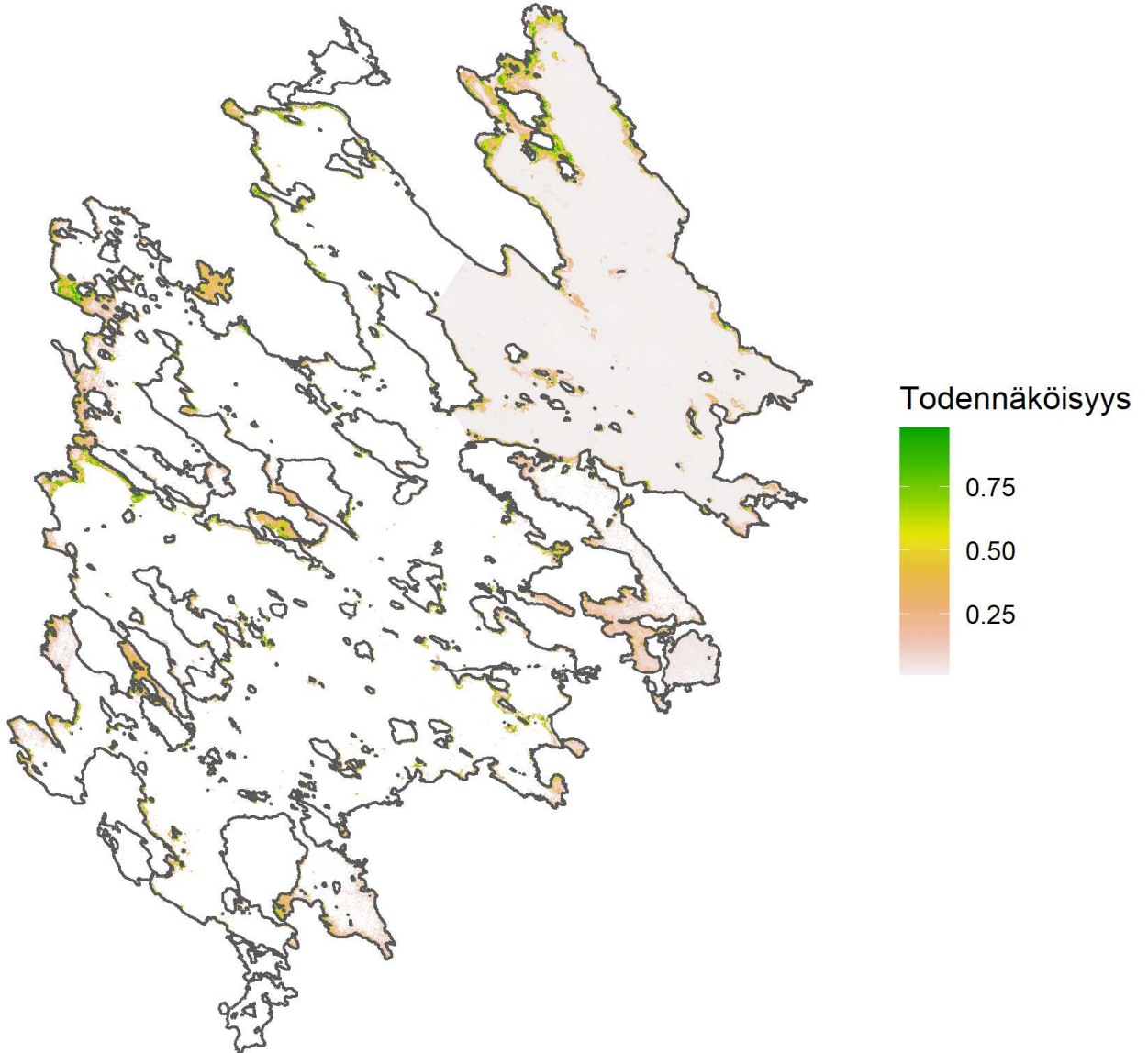
L. dortmanna



Kuva 12. Ennustetut nuottaruohon esiintymistodennäköisyydet Puruvedellä.



L. uniflora



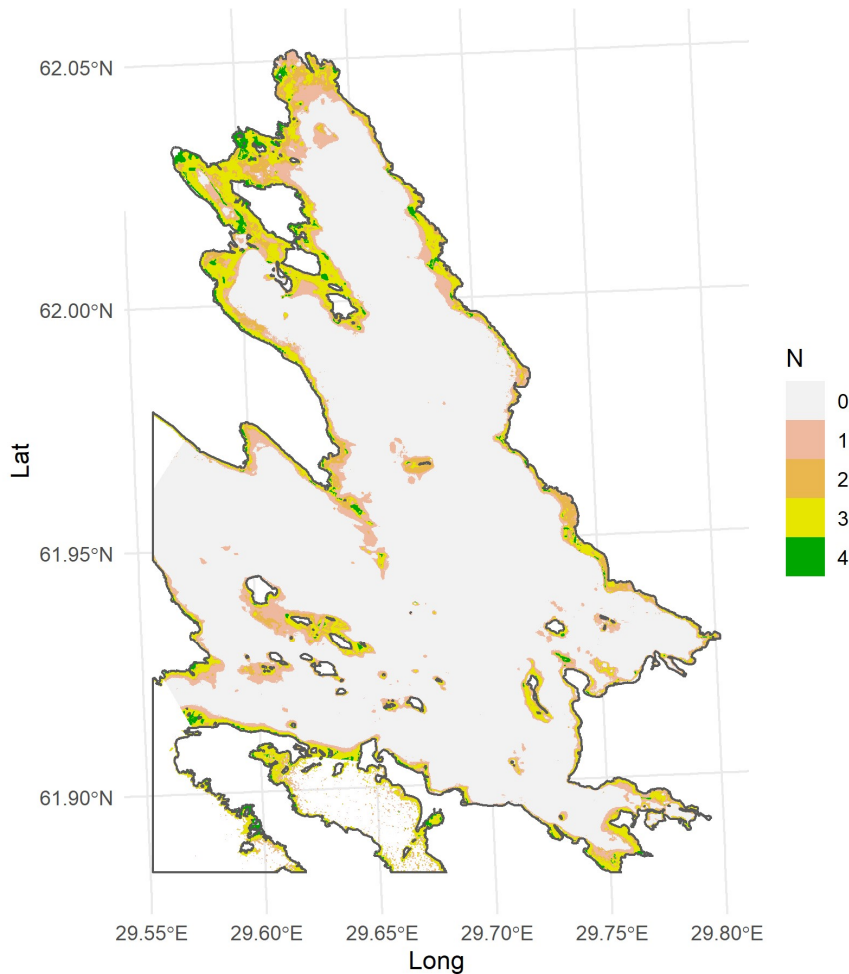
Kuva 13. Ennustetut raanin esiintymistodennäköisyydet Puruvedellä.

Pohjalehtisten lajien tärkeimmät esiintymisalueet Hummonselällä

Mallinnettujen lajien tärkeimpien esiintymisalueiden selvittämiseksi lajien ennustetut esiintymistodennäköisyydet muutettiin presence-absence-ennusteiksi laskemalla lajikohtainen kynnyisarvo, joka maksimoi oikein ennustettujen havaintojen määrän. Tämän jälkeen kaikkien lajien ennusteet summattiin yhteen. Hummonselällä pohjaruusukkeellisten tärkeimpiä ennustettuja esiintymisalueita olivat pohjoisosan matalat, suojaiset rannat (kuva 14).

METSÄHALLITUS
FORSTSTYRELSEN

Lajeja yhteensä



Kuva 14. Ennustettu tutkittujen lajien lukumäärä yhteensä Hummonselän alueella.

BRT-mallinnuksen tulosten perusteella neljän tutkitun lajin elinympäristövaatimukset ovat pääpiirteissään samanlaiset. Huomattavimmat erot liittyivät kasvusyvyyteen ja rannan avoimuuteen. Pääpiirteissään kuitenkin samat alueet sekä Hummonselällä että Puruvedellä nousivat esiin suotuisina elinympäristöinä kaikille lajeille. Tummalahnaruoho esiintyi tutkituista lajeista selkeästi syvimmällä, eikä rannan avoimuus asettanut sen esiintymiselle yksiselitteisiä esteitä. Tämä saattaa johtua siitä, että aallokon voima on pienempi alhaisemmillä kasvusyvyyksillä.

Vaalealahnaruoho esiintyi selkeästi matalammalla. GAM-mallinnuksen perusteella todennäköisimmät esiintymispaikat sijaitsivat suojaisilla alueilla, joilla virtausnopeus pohjan tuntumassa on suhteellisen pieni. Toisaalta laji voi esiintyä myös avoimemmilla paikoilla. Ensin mainittuja alueita esiintyi tutkimusalueella erityisesti Hummonselän pohjoisosassa. Vaalealahnaruohohavaintoja oli kasvillisuuskartoitusdatassa vähän ($N = 43 / 808$ kasvillisuusruutua), minkä vuoksi lajin ennustetut esiintymistodennäköisyydet olivat muihin lajeihin verrattuna matalia.



Nuottaruohon esiintymistodennäköisyys oli suurimmillaan 1 - 3 metrin syvyydessä ja se oli tiukasti sidoksissa hiekkapohjiin. Mallin tulosten perusteella laji esiintyy ruovikoiden laitamilla. Tämä saattaa johtua soveliaiden olosuhteiden esiintymisestä ruovikoiden läheisyydessä, tai vaihtoehtoisesti johtua alkuperäisten nuottaruohohabitaattien ruovikoitumisesta.

Raani esiintyi tulosten perusteella samalla syvyysvyöhykkeellä nuottaruohon kanssa, kuitenkin enimmillään hieman syvemmällä (3,5 m). Kuten nuottaruohon, myös raanin esiintyvyyteen vaikutti ruovikoiden läheisyys. Kokonaistyyppi oli syvyyden jälkeen merkittävin esiintyvyyttä selittävä tekijä. Hummonselän kokonaistyyppitaso saatiin spatiaalisella interpoloinnilla kesän aikaisista VESLAn ravinnemittauksista. Koska mittauspisteitä oli rajallinen määrä, eivät erot Hummonselän eri alueilla olleet kovin suuria. Korkeampia pitoisuuksia löytyi tyypillisesti pohjoisen matalilta alueilta, jotka olivat muilta ominaisuuksiltaan tutkimuslajeille sopivia, joten kokonaistyyppien vaikutus saattaa selittyä osittain näillä tekijöillä.

BRT-mallit ovat yleensä ennusteiltaan tarkkoja (Norberg et al. 2019) erityisesti interpolointitehtävissä, eli esiintymistodennäköisyyksien ennustamisessa ympäristömuuttuja-avaruuden sisäpuolelle (Charney et al. 2021). Tämänkaltaisen "opportunistisen" tilastollisen mallin kääntöpuolena kuitenkin on, että malli saattaa poimia selittäviksi tekijöiksi muuttujia, jotka eivät suoraan kausaalisesta selitä lajien esiintyvyyttä, mikä saattaa mutkistaa tulosten tulkintaa. Näin vaikuttaa tapahtuneen tässä tutkimuksessa rakennetuissa malleissa erityisesti sameuden kohdalla. Useilla lajeilla (tummalahnaruoho, nuottaruoho ja raani) sameus nousi yhdeksi tärkeimmäksi selittäväksi muuttujaksi. Vastoin ennako-odotuksia kaikissa tapauksissa sameuden kasvu vaikutti suotuisasti lajien esiintyvyyteen. Tämä todennäköisesti johtuu siitä, että matalat hiekkapohjat tuottivat heijastuksia satelliittidatasignaaliin, ja tunnistuivat korkean sameuden alueiksi, toisaalta voi myös olla kyse siitä, että alueet, jotka ovat esimerkiksi syvyydeltään tai pohjanlaadultaan lajeille sopivia, ovat keskimääräisesti sameampia. Tulosten tulkinnan kannalta sameuden vaikutus tarkoittanee siis lajien esiintymistä matalilla, hiekkaisilla ranta-alueilla.

Kun kaikkien lajien esiintymisennusteita tarkasteltiin yhdessä (kuva 14), lajien tärkeimmiksi esiintymisalueiksi Hummonselällä osoittautuivat pohjois- ja eteläosien matalat, suojaiset lahdet. Myös joidenkin Hummonselän eteläosien saarten ympäristöt vaikuttivat tärkeiltä elinympäristöiltä.

Koska syvyystietoja Hummonselän ulkopuolelta oli saatavilla niukasti, tehtiin koko Puruvedelle ennusteet ainoastaan matalille (alle 3m) alueille, joiden syvyys- ja pohjanlaatutiedot pystyttiin tulkitsemaan satelliittikuvista. Tämän vuoksi syvemmällä esiintyvien lajien (tummalahnaruoho) osalta ennusteiden tarkkuus on kyseenalainen. Myös muiden lajien osalta tulosten käytössä on syytä noudattaa varovaisuutta. Koska kasvillisuuskartoitusaineistoa oli saatavilla ainoastaan Hummonselän alueelta, jouduttiin mallien tuloksia ennustamaan (ekstrapoloimaan) alkuperäisen maantieteellisen alueen (ja osin ympäristömuuttujien) ulkopuolelle. Alueet, joissa lajien ennustettiin esiintyvän, voivat kuitenkin olla kiinnostavia pilottikohteita tulevien vuosien kasvillisuuskartoituksille.



Yhteenveto

Työssä rakennettiin tilastolliset habitaattimallit Puruvedelle neljälle pohjalehtiselle vesikasville: tummalahnaruoholle (*Isoëtes lacustris*), vaalealahnaruoholle (*Isoëtes echinospora*), raa-nille (*Littorella uniflora*) sekä nuottaruoholle (*Lobelia dortmanna*). Mallit rakennettiin käyttäen tehostettuja regressiopuita (Boosted Regression Trees, BRT) sekä yleistettyä additiivista mallia (Generalised Additive Model, GAM). Mallien avulla ennustettiin tutkittujen lajien esiintymistodennäköisyydet koko Puruvedelle sekä Puruveden koillisosassa sijaitsevalle Hummonselälle. Mallinnuksen lähtöaineistona käytettiin Hummonselällä suoritettuja kasvillisuuskartoituksia, sekä FRESHABIT-hankkeen puitteissa tuotettuja mallinnettuja ja satelliittihavainnoista tulkit-tuja ympäristömuuttujia.

Mallien ennustekyky oli yleisesti ottaen hyvä. Lajien esiintymistodennäköisyydet vaihtelivat lajikohtaisesti voimakkaasti syvyyden mukaan, joka oli tärkein esiintyvyyttä ennustava tekijä muilla lajeilla paitsi vaalealahnaruoholla, jonka esiintyvyys riippui voimakkaimmin veden vir-tausnopeudesta pohjan tuntumassa. Mallinnuksen tulosten perusteella tutkittujen pohjaleh-tisten kasvien tärkein esiintymisalue on Hummonselän pohjoisin osa.

Kirjallisuus

- Charney, N. D., S. Record, B. E. Gerstner, C. Merow, P. L. Zarnetske, and B. J. Enquist. 2021. A test of species distribution model transferability across environmental and geographic space for 108 western North American tree species. *Frontiers in Ecology and Evolution*:393.
- Elith, J., J. R. Leathwick, and T. Hastie. 2008. A working guide to boosted regression trees. *Journal of animal ecology* **77**:802-813.
- Hastie, T., and R. Tibshirani. 1990. *Generalized additive models*. CRC Monographs on Statistics & Applied Probability. New York: Chapman & Hall.
- Laaksonen, M., M. Suonio, L. Vertio, and J. Ilomnen. 2017. Järvien kartoitusmenetelmät ja maastotyömenetelmän kuvaus.
- Norberg, A., N. Abrego, F. G. Blanchet, F. R. Adler, B. J. Anderson, J. Anttila, M. B. Araújo, T. Dallas, D. Dunson, and J. Elith. 2019. A comprehensive evaluation of predictive performance of 33 species distribution models at species and community levels. *Ecological monographs* **89**:e01370.
- Smolders, A. J., E. Lucassen, and J. G. Roelofs. 2002. The isoetid environment: biogeochemistry and threats. *Aquatic Botany* **73**:325-350.
- Verbruggen, H., L. Tyberghein, G. S. Belton, F. Mineur, A. Jueterbock, G. Hoarau, C. F. D. Gurgel, and O. De Clerck. 2013. Improving transferability of introduced species' distribution models: new tools to forecast the spread of a highly invasive seaweed. *PloS one* **8**:e68337.

Milestone A6



Virtanen, E. A., M. Viitasalo, J. Lappalainen, and A. Moilanen. 2018. Evaluation, Gap Analysis, and Potential Expansion of the Finnish Marine Protected Area Network. *Frontiers in Marine Science* **5**.

Yates, K. L., P. J. Bouchet, M. J. Caley, K. Mengersen, C. F. Randin, S. Parnell, A. H. Fielding, A. J. Bamford, S. Ban, A. M. Barbosa, C. F. Dormann, J. Elith, C. B. Embling, G. N. Ervin, R. Fisher, S. Gould, R. F. Graf, E. J. GREGG, P. N. Halpin, R. K. Heikkinen, S. Heinänen, A. R. Jones, P. K. Krishnakumar, V. Lauria, H. Lozano-Montes, L. Mannocci, C. Mellin, M. B. Mesgaran, E. Moreno-Amat, S. Mormede, E. Novaczek, S. Opiel, G. Ortuño Crespo, A. T. Peterson, G. Rapacciuolo, J. J. Roberts, R. E. Ross, K. L. Scales, D. Schoeman, P. Snelgrove, G. Sundblad, W. Thuiller, L. G. Torres, H. Verbruggen, L. Wang, S. Wenger, M. J. Whittingham, Y. Zharikov, D. Zurell, and A. M. M. Sequeira. 2018. Outstanding Challenges in the Transferability of Ecological Models. *Trends in ecology & evolution* **33**:790-802.