

Tartu Ülikooli Eesti Mereinstituut

Optilise kaugseire kasutamine merepõhja elupaikade kaardistamisel

(Deliverable 2.3.1. Optical remote sensing for mapping seabed habitats)

Teostatud projekti "Eesti mereala keskkonna ja loodusväärtuste hindamise ja seire innovaatilised lahendused" raames

Versioon 2, 15.04.2020

Leping: RITA1/02-60-01

Vastutav täitja: Georg Martin

Aruande koostajad: Kristjan Herkül, Ele Vahtmäe, Tiit Kutser

Tallinn 2020



RITA

The logo for mereRITA features three blue wavy lines above the text "mereRITA" in a serif font.

mereRITA

Aruande versioonide info

Versiooni number	Kuupäev	Info
1	06.04.2020	Esimene tellijale esitatud versioon
2	15.04.2020	Vastavalt juhtkomisjoni soovile lisati ingliskeelne kokkuvõte (Executive Summary)

Sisukord

Executive Summary	4
1. Sissejuhatus	6
2. Materjal ja metoodika	9
2.1. Uuringualad	9
2.2. Optilise kaugseire andmed	11
2.3. Merepõhja kohtvaatlused	12
2.4. Andmetöötlus ja matemaatiline modelleerimine	15
3. Tulemused	21
3.1. Maa-ameti ortofotod	21
3.2. Hypsrex	27
3.3. Satelliidipildid	30
3.4. Erinevate kaugseire meetodite võrdlus	31
4. Kokkuvõte	39
Kasutatud allikad	42
Lisa 1	44
Lisa 2	45
Lisa 3	46

Executive Summary

The aim of the study was to test the feasibility of using different optical remote sensing methods for mapping shallow water (maximum depth about 6 m) seabed habitats. Field study was carried out at four study areas in Western Saaremaa Island, Estonia, in May and June 2019.

The feasibility of the following optical remote sensing methods was tested:

- Airborne orthophoto (OP) mosaics produced by the Estonian Land Board (pixel size 0.2 m);
- Airborne hyperspectral imagery acquired by hyperspectral imager Hypspx (pixel size 0.8 m);
- Sentinel-2 (S2) satellite imagery (pixel size 10-60 depending on channel)

On-site sampling of seabed was done using underwater video. Substrate and biotic variables were recorded at sampling sites: percentage coverage of seabed substrate types, percentage coverage of macrophytes and selected macroinvertebrates. These variables are used in common seabed mapping studies in Estonia.

Machine learning methods – boosted regression trees, random forest, artificial neural networks – were used to model relationships between remote sensing variables and seabed habitat variables. Predictor variables that are commonly used in mapping studies – water depth, wave exposure, geographical coordinates – were also used in a separate set of models in order to compare the predictive performance of models with different sets of predictor variables (i.e. models with and without remote sensing data). Using the trained models, spatial predictions were made in 2 m grid covering the full extent of study areas. The predictive performance of the models was estimated based on mathematical validation and visual expert judgement.

Based on the mathematical validation and visual expert judgement the following conclusion could be made:

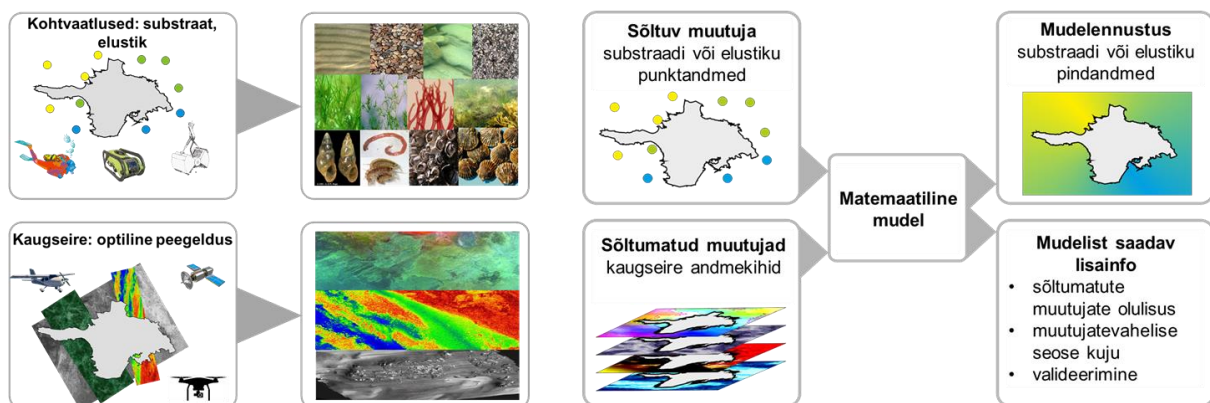
- Data derived from all three remote sensing methods (OP, Hypspx, S2) enabled to improve the prediction accuracy of models compared to models without remote sensing data.
- The modeling method random forest proved to be superior to other methods in terms of predictive accuracy and ease of use.
- Regardless that the mathematical validation showed that the numerical improvement of prediction accuracy attributable to remote sensing data was rather limited, the visual assessment of the map layers indicated significant qualitative improvements in model predictions. Models with remote sensing data produced predictions that exhibited more natural patterns compared to those without remote sensing data. The patterns of borderlines between patches and transitions between patches of communities and habitats followed the patterns that were also visually distinguishable in the remote sensing imagery.
- Sentinel-2 satellite imagery has too low spatial resolution for high resolution mapping of seabed habitat features in case there is a need to detect features with size < 100 m. However, Sentinel-2 is a good method for mapping in larger spatial scales.
- The mathematical model validation showed that:
 - Among benthic habitat variables, the total cover of vegetation, cover of brown algae and cover of eelgrass (*Zostera marina*) exhibited the highest gain due to the inclusion of remote sensing variables.

- Hypspx models showed higher predictive accuracy compared to the other remote sensing methods. This can be explained by the much wider spectral coverage and higher spectral resolution of Hypspx compared to orthophotos and satellite imagery.
- Models with satellite imagery had somewhat higher predictive accuracy compared to models based on orthophotos.
- The predictive accuracy of models with orthophoto data was the lowest and this can be attributed to the lowest spectral resolution among the methods and to the fact that orthophotos were from year 2017 while the other remote sensing material was recorded in 2019.

This study was financially supported by the European Regional Development Fund within National Programme for Addressing Socio Economic Challenges through R&D (RITA).

1. Sissejuhatus

Madala rannikuvee ja sisevee põhjasubstraadi ja -elustiku kaardistustööde läbiviimine kasutades ainult tavapäraseid kohtvaatluseid on kulukas ja aeganõudev, sest mootorpaatidega proovipunktide külastamine võib olla ohtlik või võimatu liiga madala vee, kivide ja rohke põhja- ja kaldaveetaimestiku tõttu. Sageli on sellistes piirkondades võimalik kohtvaatlusi teostada ainult jalgsi vees liikudes. Kui tugineda ainult kohtvaatlustele, mille arv raskeid liikumisolusid arvestades on madal, jääb proovivõtuvõrgustik liiga hõredaks, et sellest saadud andmete abil oleks võimalik luua interpoleerimise või matemaatilise modelleerimise abil rahuldava kvaliteediga kaardikihte. Sellises olukorras on praeguste tehniliste võimaluste juures ainuke mõistlik lahendus kasutada optilist kaugseiret koos kohtvaatlustega. See kombinatsioon võimaldab optiliste andmetega katta kogu uuringuala ning leida matemaatiliste mudelite abil seoseid optiliste (valgusspektrid) ja kohtvaatlustega saadud muutujate (substraat, elustik) vahel. Seejärel saab leitud mudelite abil ennustada substraadi ja elustiku muutujate väärtused kogu uuringu alal. Kohtvaatlustel, optilisel kaugseirel ja matemaatilisel modelleerimisel põhineva kaardistamise tööpõhimõtte skeem on toodud joonisel 1.1.



Joonis 1.1. Põhjasubstraadi ja elustiku andmete kogumise ja ruumilise leviku modelleerimise tööpõhimõtte skeem.

Kaugseire rakendamine madalaveeliste veekogude kaardistamisel

Viimaste aastakümnete tehnoloogiline areng on võimaldanud laiendada optilise kaugseire kasutatavust madalaveeliste elupaikade kaardistamiseks (Kutser et al 2020). Käesoleval ajal on kaugseire aparatuur kasutatav väga erinevatelt platvormidelt - satelliitidelt, lennukitelt, laevadelt, mehitamata õhu- ja veesõidukitelt. Vastavalt on väga erinevad ka kaugseire kasutamise hind ning iga konkreetse probleemi lahendamiseks vajaliku sensori valik. Näiteks on Eesti laiuskraadil iga 2-3 päeva järel võimalik saada Euroopa Kosmoseagentuuri (ESA) Sentinel-2 satelliitide pilte, mis on 10 m ruumilise lahutusega. Need pildid katavad suure osa Eestis korruga ja on tasuta. Vaja on ainult välja töötada Eesti oludesse sobivad kaugseire meetodid (S2 on maismaa, mitte veekogude seireks mõeldud sensor) nende andmete interpreteerimiseks. Satelliitidel paikneb ka kommerts-sensoreid, mille ruumiline lahutus on alla 1 m. Sellistelt sensoritelt saab palju detailsemat informatsiooni madalate veealade kohta. Teisalt ei võimalda andmete hind (suurusjärgus 35 €/km²) ning pildi suurus (100-200 km²) seirata suuri merealasid tiheda ajalise sammuga.

TÜ Eesti Mereinstituudil on olemas hüperspektraalne lennukil paiknev kaugseire seade HySpex, mida kasutatakse Maa-ameti lennukilt. Selle ruumiline lahutus sõltub lennu kõrgusest ning on seega muudetav. Kui lennata väga hea ruumilise lahutusega (alla 1 m), siis saab sama ajaga katta oluliselt väiksema mereala kui lennates näiteks 5 m lahutusega. Kuna üheks suuremaks kuluks on lennutunni maksumus, siis peab hoolega hindama millist ruumilist lahutust vajab iga konkreetne uurimisprobleem. Ka ei ole lennukiga võimalik katta väga suuri merealasi ja/või teostada mõõtmisi suure ajalise sagedusega (kas või juba ilmastiku tõttu).

Maa-ameti lennukil on ortofotokaamera, millega kaardistatakse kogu Eesti iga 2-3 aasta tagant ruumilise lahutusega 10-40 cm. Need andmed on tasuta saadaval. Merealad ei ole Maa-ameti haldusala. Seepärast mere kohal lende ei tehta. Teisalt on madalad rannikualad maismaale piisavalt lähedal ning jäävad paratamatult piltidele. Maa-ameti ortofotode kasutamisega madalate vete uurimiseks ja seireks on küll ka mõningad tehnoloogilised probleemid. Kaardistamiseks lennatakse ida-lääne suunaliselt. Vee kohal ei ole see hea variant, sest suurema osa lennukilt mõõdetud signaalist võib sel juhul moodustada veepinnalt peegeldunud päikesekiirgus ehk me ei saa nendelt piltidelt mingit infot vees või põhjas olevast. Ka teostatakse maapinna kaardistuselnnud varakevadel (enne taimestiku tärkamist). See ei ole optimaalne aeg paljude veekogu põhja seisundiga seotud probleemide uurimiseks ja seireks. Mõnel muul ajal ja vee kaugseire jaoks sobivate lennusuundadega lendamine tuleb eraldi tellida ning selle eest ka eraldi maksta.

Ruumiliselt veelgi paremad, tavaliselt paarisentimeetrise lahutusega, on droonidelt kogutavad andmed. Tavalistel odavatel droonidel paiknevad ainult tavapärased digitaalkaamerad, mitte spetsiaalsed kaugseire instrumendid. See tähendab, et kogutav informatsioon koosneb punasest, rohelisest ja sinisest valgusest (RGB) nagu Maa-ameti ortofotodegi puhul. Drooniga uuritavad alad on veel palju väiksemad kui lennukiga uuritavad alad ning sõltuvad veelgi rohkem ilmast, näiteks tuulest.

Lisaks lendavatele droonidele on võimalik kasutada ka ujuvaid "droone". TÜ Eesti Mereinstituut soetas hiljuti firma SurfBee ujuvroboti, millel on paradal nii kaugseire seadmed (Ramses spektromeetrid), vee kvaliteedi sensorid (klorofüll-a, lahustunud orgaanika ning fükotsüaniini fluoromeetrid) ning lisaks sonar sügavuse 3D kaardistamiseks ning video põhjasubstraadi ning taimestiku tuvastamiseks. Põhimõtteliselt saaks seda ujuvrobotit kasutada ka iseseisva (lähi)kaugseire informatsiooni kogumiseks. Käesoleva projekti raames kasutati seda kohtvaatluste videoülesvõtete tegemisel, et saada rohkem andmeid satelliitidelt, lennukilt ja droonilt saadud informatsiooni interpreteerimiseks kui seda võimaldaks jalgsi või paadiga madalas vees üksikutes punktides andmete kogumine.

Viimastel aastakümnetel on oluliselt paranenud ka kaugseire sensorite spektraalne lahutusvõime. Video, ortofotod ning suure ruumilise lahutusega satelliidid mõõdavad veest tulevat kiirgust tavaliselt vaid punases, rohelises ja sinises spektriosas (RGB). See tähendab, et informatsiooni hulk, mille põhjal tuvastada näiteks erinevaid põhjataimestiku rühmi, on väga piiratud. Palju detailsemat informatsiooni on võimalik koguda hüperspektraalsete sensoritega, mis 3-4 spektrikanali asemel mõõdavad vees tulevat kiirgust kümnetel ja sadadel lainepikkustel. Selline andmestik võimaldab palju detailsemalt tuvastada erinevaid veekogu põhja tüüpe. Näiteks Eesti Mereinstituudi HySpex mõõdab nähtavat valgust ligemale sajas spektrikanalis samas kui ortofotod ja videod on sisuliselt kolme spektrikanaliga ning mõnedel satelliitidel on nähtava spektriosa kirjeldamiseks kuni viis spektrikanalit. Hüperspektraalseid sensoreid tehakse juba ka droonidele. Kui veel mõni aasta tagasi kaalusid sellised riistad mitmeid kilosid, maksid üle 100 000€ ning vajasisid väga suurt ja kallist drooni, mis jaksas sellist lasti tõsta, siis käesoleval ajal on välja töötatud hüperspektraalseid sensoreid, mis kaaluvad vaid

paarsada grammi ning maksavad mõnikümme tuhat. Käesoleva projekti tarbeks, oli meil kasutada Maa-ameti lennukil paiknev HySpex kuid ei olnud droonil paiknevat sensorit.

Ruumilisest ja spektraalsest ning ajalisest vaatevinklist võetuna on seega saadaval väga palju erinevaid kaugseire sensoreid, millede seast optimaalseima (informatsiooni ja hinna suhe) valimine on tihtipeale keeruline protsess. Käesoleva töö üheks eesmärgiks oligi katsetada erinevate sensorite ja platvormide võimekust ja sobivust madalate merealade kaardistamisel.

Töö eesmärk ja uurimisrühm

Käesoleva töö eesmärgiks oli testida erinevate kaugseire meetodite rakendatavust madala rannikumere merepõhja substraadi ja elustiku leviku kaardistamisel. Uuringu fookuses olid sellised merepõhja substraadi ja elustiku muutujad, mida kasutatakse tavapärasel kaardistustöödel sh Euroopa Liidu loodusdirektiivi (92/43/EMÜ) elupaigatüüpide kaardistamisel.

Töö teostasid Tartu Ülikooli Eesti Mereinstituudi töötajad Kristjan Herkül, Ele Vahtmäe, Tiit Kutser. Välitöödel osalesid täiendavalt Lauri London, Kaire Kaljurand, Tiia Möller, Arno Põllumäe, Martin Teeveer, Maria Põllupüü, Laura Lõugas ja Külli Kutser. Videopovide analüüsid viis läbi Tiia Möller.

Uuringu tellis ja uuringut rahastab Eesti Teadusagentuur Euroopa Regionaalarengu Fondist toetatava programmi „Valdkondliku teadus- ja arendustegevuse tugevdamine“ (RITA) tegevuse 1 „Strateegilise TA tegevuse toetamine“ kaudu. Uuring valmis Keskkonnaministeeriumi eesmärkide elluviimiseks.

2. Materjal ja meetoodika

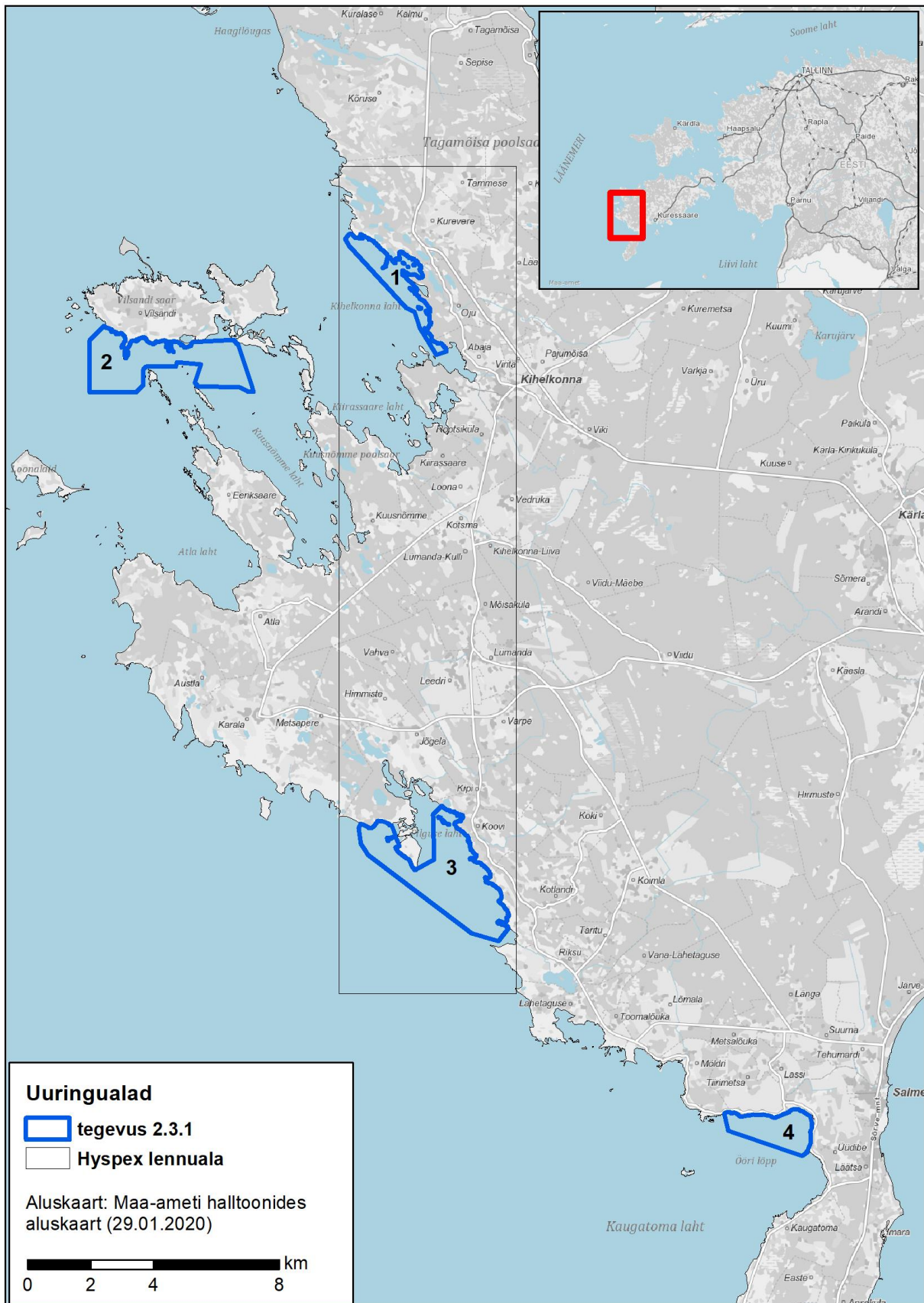
2.1. Uuringualad

Uuringud teostati neljal alal Saaremaa lääneosas, mille numeratsioon, nimetused ja ligikaudsed pindalad on järgnevad (joonis 2.1.1):

- Ala 1 – Kihelkonna laht Jaagarahu sadama piirkonnas (2,8 km²);
- Ala 2 – Vilsandi saare lõunarannik (6,5 km²);
- Ala 3 – Pilguse laht (8,3 km²);
- Ala 4 – Ööri lõpp Kaugatoma lahes (2,5 km²).

Uuringualad valiti välja Maa-ameti ortofotode alusel. Valitud aladel on ulatuslikud madalaveelised (alla 3 m) merealad ja ortofotodel oli visuaalselt tuvastatav erineva värvuse ja struktuuriga merepõhja alade esinemine. Lisaks sellele arvestati alade paigutamisel ka võimalikult sirgete lennutrajektoorde moodustamise võimalusega hüperspektraalse materjali kogumiseks lennukilt.

Uuringualade keskmine sügavus Veeteede Ameti 2019. a. andmete järgi oli 1,7 m. Visuaalsel hinnangul olid merepõhja mustrid jälgitavad Maa-ameti ortofotodel aladel 1 kuni 3 kuni ligikaudu 5-6 m sügavuseni, alal 4 kuni 4 m sügavuseni.



Joonis 2.1.1. Uuringualad. Näidatud on Maa-ameti lennuki lennuala, kus salvestati andmed hüperspektraalse instrumendiga Hypsax.

2.2. Optilise kaugseire andmed

Kasutati kolme erinevat liiki optilise kaugseire andmeid:

- Maa-ameti ortofotod (edaspidi lühendina „OF“);
- Hüperspektraalse spektromeetriga HySpex lennukilt kogutud andmed (edaspidi lühendina „HySpex“);
- Sentinel-2 satelliidipildid (edaspidi lühendina „S2“).

Maa-ameti ortofotod saadi Maa-ameti kodulehelt¹ kaardiruutude kaupa alla laadides. All laeti järgmiste Maa-ameti kaardiruutude ortofotod: 51493, 51581, 51582, 51583, 51771, 51772, 51781, 51783. Iga kaardilehe ortofotodest laeti alla kõigi saadaolevate kaardistusaastate (2005, 2008, 2011, 2012, 2015, 2017) tavalised RGB (*Red, Green, Blue* – punane, roheline, sinine) värvikanalitega fotod ja valemvärvi CIR (*color infrared*) fotod GeoTIFF formaadis. CIR fotodest kasutati ainult lähisinfrapuna kanalit. Ortofotode piksli suurus oli 0,2 m.

Päikese peegelduse (*sun glint*) mõju eemaldamiseks rakendati Hedley et al (2005) meetodit kasutades lähisinfrapuna kanalit. Alal 4 puhul rakendati eraldi päikese peegelduse ja lainete vahuga alade väljamaskimist kasutades sinise kanali intensiivsust ja punase-sinise kanali intensiivsuse suhet; mõlema näitaja puhul leiti lävendväärtused, mis iseloomustasid tugeva päikese peegelduse ja lainetuse vahuga alasid.

Modelleerimise sisendandmetena kasutamiseks arvutati täiendavalt kõigi kolme värvikanali omavahelised suhted ja seeläbi saadi sisendiks kokku kuus muutujat: R, G, B, R/G, G/B, R/B.

Hüperspektraalne spektromeeter Hypspx VNIR-1024 (tootja Norsk Elektro Optikk, NEO) kuulub Tartu Ülikooli Eesti Mereinstituudile ja see on paigaldatud Maa-ameti lennukile Cessna 208 Grand Caravan. Hypspx võimaldab koguda kaugseire andmeid nähtava valguse ja infrapuna valguse spektrivahemikus (409-988 nanomeetri) 2,7 nanomeetrise spektraalse lahutusega kokku 216-l spektrikanalil.

Hypspx andmed saadi kaugseire RITA (Kaugseireandmete kasutuselevõtt avalike teenuste väljatöötamisel ja arendamisel) projekti raames. Lendamise uurimisaladel toimus 19.07.2019.a. kella 09:00 ja 12:00 vahel. Kuna lennujooni tuli optimeerida arvestades Maa-ameti lennukiga tellitud töö ajalisi ja rahalisi piiranguid kaugseire RITA projektis, siis on Hypspx andmetega kaetud ainult alad 1 ja 3, sest need oli võimalik katta põhja-lõuna suunaliste lennujoonte ühe plokinä (joonis 2.1.1).

Piltide täpne ruumiline lahutus sõltub lennukõrgusest. Antud juhul lennati 1700 m kõrgusel, mis võimaldas koguda andmeid 0,8 m piksli suurusega. Kaugseire andmed koguti eraldi lennujoontena, mis hiljem pilditöötlusprotsessis tervikpildiks kokku mosaiigiti. Hypspx andmete eeltöötlus hõlmab kõigepealt kaugseirepiltide radiomeetrilist korrektsiooni kasutades NEO poolt välja arendatud tarkvara. Järgneb geomeetriline korrektsioon, mis teostati PARGE (PARAmetric Geocoding, ReSe Applications Schöpfler, University of Zurich) geokorrektsiooni tarkvaraga kasutades lennu vältel kogutud GPS ja INS logiandmeid. Atmosfääri korrektsioon teostati ENVI pilditöötlusprogrammis kasutades FLAASH (Fast Line-of-sight Atmospheric Analysis of Spectral Hypercubes) tarkvara.

¹ https://geoportaal.maaamet.ee/index.php?lang_id=1&page_id=610

Päikesegliindi eemaldamiseks kasutati Hedley et al (2005) välja töötatud gliindikorrektsooni meetodit. Maismaa alad eemaldati edasisest pilditötlusest kasutades lähisinfra puna spektripiirkonda, kus kontrast veest ja maismaalt tuleva optilise signaali tugevuse vahel on väga suur. Hypsrex andmete eeltöötlus teostati kaugseire RITA projekti raames.

Modelleerimise sisendandmeteks valmistati ette kaks komplekti muutujaid:

- 12 valitud kanali (470, 500, 520, 575, 600, 610, 625, 635, 650, 675, 705, 720 nm) peegeldustegurid (edaspidi „Hypsrex-kanalid“). Kanalite valikul on arvestatud sellega, et need kajastaksid erinevatele taimerühmadele omaste pigmentide (klorofüll-a, peridiniin jt) neeldumise maksimume (lohk peegelduse spektris) võis siis erinevatel põhjustel (tagasihajumine, kahe pigmendist tingitud lohu vahe, kus neeldumist peaaegu ei ole) tekkivaid haripunkte spektris (Vahtmäe et al 2006, Kotta et al 2014, Kutser et al 2020).
- 12 esimest peakomponendi arväärtused, mis saadi kõikide 490-700 nm vahemiku kanalite peakomponentanalüüsi tulemusel (edaspidi „Hypsrex-PCA“). 12 esimest peakomponenti kirjeldasid aladel 1 ja 3 kumulatiivselt kokku vastavalt 95% ja 94,3% varieeruvusest.

Sentinel-2 satelliidipildid saadi Copernicuse andmebaasist (<https://scihub.copernicus.eu/>). Materjal valiti välja sellisest ajavahemikust, mis kattuks võimalikult hästi välitööde läbiviimise ajaga (juuni - juuli 2019). Pildimaterjali valikul oli oluline ka pilvkatte puudumine uuringualadel ja vee läbipaistvus. Töös kasutamiseks valiti satelliidipilt 7. juunist 2019.

Modelleerimise sisendandmeteks valiti järgmised kuus kanalit: 443, 490, 560, 665, 705, 740 nm. 443 nm kanal oli 60 m piksliga, kanalid 490–665 nm 10 m piksliga ja kanalid 705–740 nm 20 m piksliga. 60 m ja 20 m piksliga kanalid arvutati ümber 10 m peale, et saavutada teiste kanalitega tehniliselt sama resolutsioon. Sentinel-2 puhul kaaluti nii atmosfäärikorrektsooniga kui ilma korrektsoonita materjali kasutamist. Visuaalsel hinnangul osutus ilma korrektsoonita materjal informatiivsemaks, sest merepõhja mustrid olid oluliselt paremini nähtavad ja piltides esines vähem müra. Seetõttu kasutati ilma korrektsoonita pilte. Atmosfäärikorrektsoon on oluline siis kui kasutatakse palju pilte ja/või kasutatakse analüütilisi (objektide füüsilistel omadustel põhinevaid) meetodeid, näiteks modelleeritud spektri kujusid (*spectral library*). Kuna antud töös mahtusid uuringualad kõik ühele pildile ja rakendati masinõpet, siis ei olnud atmosfäärikorrektsooni kasutamine ka sisuliselt vajalik.

Kõik ettevalmistatud kaugseire andme viidi rasterkihtidena ESRI ArcGIS failipõhisesse geoandmebaasi (*file geodatabase*).

2.3. Merepõhja kohtvaatlused

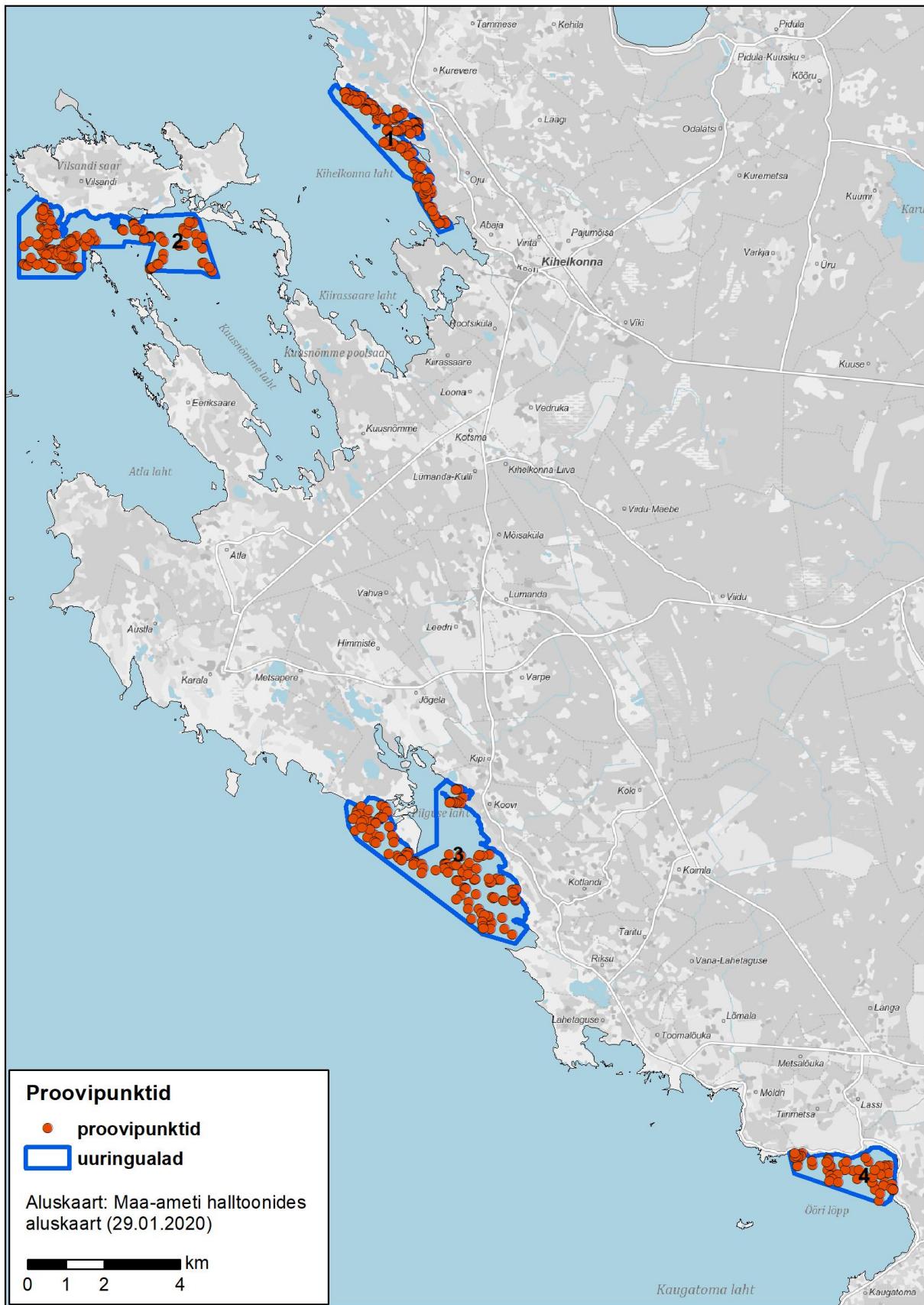
Merepõhja substraadi ja elustiku katvuste kohtvaatlused toimusid vahemikus 08.05.2019 kuni 06.06.2019. Kokku teostati kohtvaatlused 498 proovipunktis. Ligikaudu neljandik kohtvaatlusi teostati allveevideo transektina, mille puhul allveevideo salvestati mööda ortofotode järgi paika pandud joont. Kuna joonena teostatud vaatlustel eristati üks kuni neli erinevat punktvaatlust, siis moodustus kokku 535 unikaalset proovipunkti, mille paiknemine on näidatud joonisel 2.3.1.

Merepõhja substraaditüüpide ning põhjataimestiku ja –loomastiku katvuse hindamiseks kasutati paadis allveevideosüsteemi (nn *drop*-kaamera), mis koosneb veealuselt videokaamerast ning paadis

olevast salvestusseadmest ja ekraanist. Madalas kaldaäärses vees teostati videoülesvõtteid ka pinnaltpäästeülirkonnas jalgsi liikudes puidust varre külge kinnitatud GoPro kaameraga või protokolliti vaatlusandmeid otse kasutades ArcGIS Collector mobiilirakendust. Proovipunktid ja -jooned georefereeriti kasutades diferentsiaal-GNSS-seadmeid Trimble GeoExplorer 6000 ja Trimble R1. Madalas vees jalgsi proovipunktide külastamisel kasutati asukohamääranguks ka mobiiltelefoni Huawei P30 Pro. Hiljem kontoris hinnati videopildilt põhjataimestiku üldkatvus, põhjataimestiku ja -loomastiku liikide/rühmade katvused ja põhjasubstraadi tüüpide katvused protsentuaalselt.

Valdava osa proovipunktide ja -joonte paiknemisi määrati enne välitööd geoinfosüsteemis arvestades Maa-ameti ortofotodest nähtuvat merepõhja muustrilisust. Kuna enne välitööde toimumist oli kasutada ortofotosid kuni aastani 2017 (vt peatükk 2.2), siis punktide paigutamisel vaadeldi erinevate aastate ortofotosid ja märgiti proovipunktid merepõhjas visuaalselt eristuvatele aladele, mille suurus ja piirid olid aastate vahel sarnased. Sellise valiku vajaduse tingis asjaolu, et proovipunktidega on vaja katta ajaliselt püsivamad merepõhja alad, sest vastasel juhul võib tekkida oluline ebakõla kaks aastat varem salvestatud ortofoto ja kohtvaatlusel tuvastatud merepõhja substraadi ja elustiku vahel. Joone tüüpi vaatlused lisati aladel, kus esines merepõhja laigulisus väikeses ruumiskaalas (alla 10 m) ja nende laikude tabamine punktvaatlustega oleks raskendatud. Joonvaatlusi lisati ka kohtadesse, kus merepõhja iseloom muutus väga kiiresti ja punktipõhine vaatlus ei võimaldaks sellest muutlikkusest ülevaadet saada. Mõningad proovipunktid, eriti nende hulgast, mille külastused teostati jalgsi, lisati ka välitööde käigus vastavalt kohapeal tuvastatud merepõhja iseloomule.

Kohtvaatluste andmed sisestati Tartu Ülikooli Eesti mereinstituudi põhjaelustiku andmebaasi (Microsoft Access).



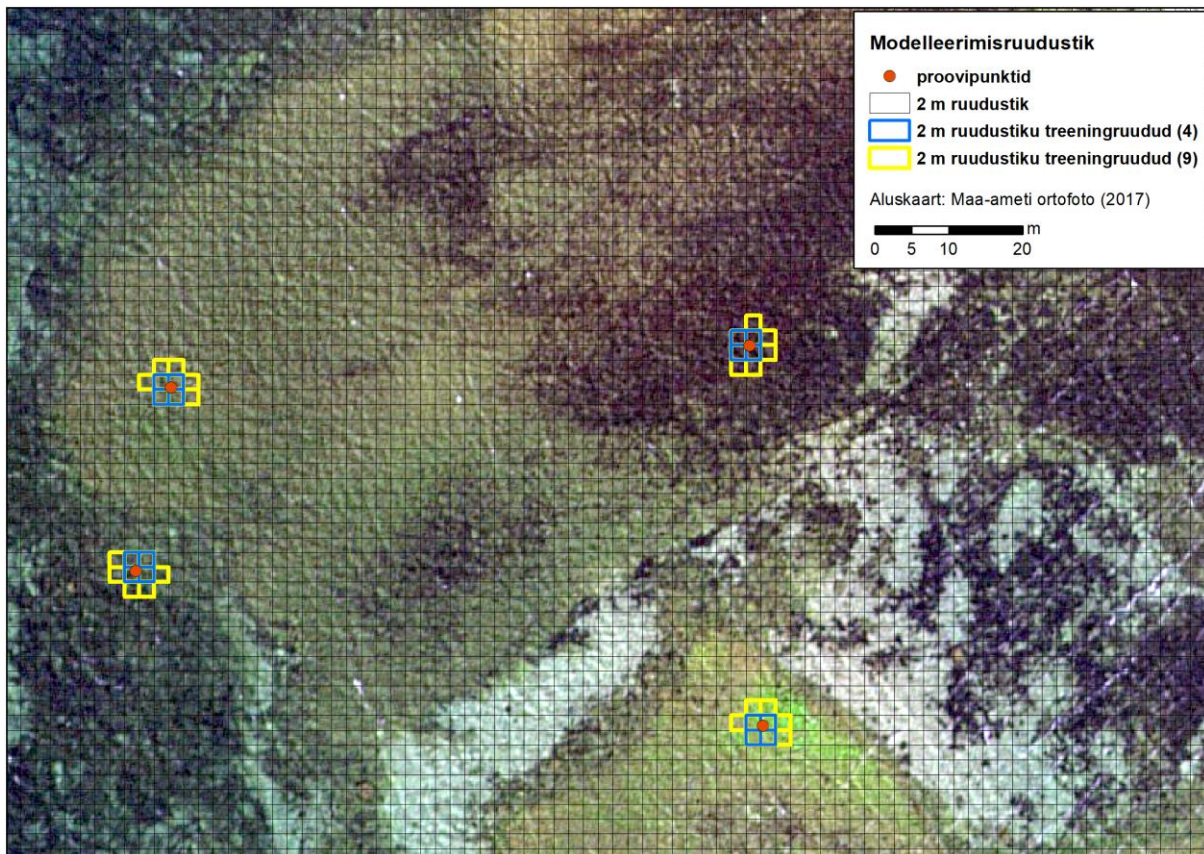
Joonis 2.3.1. Proovipunktide paiknemine.

2.4. Andmetöötlus ja matemaatiline modelleerimine

Modelleerimisruudustik

Uuringualadele loodi 2 m küljepikkusega ruudustik ja loodud 2×2 m segmendid moodustasid järgnevas modelleerimistöös vähima modelleeritava ühiku. 2 m suurus valiti kuna see oli optimaalne arvestades merepõhja kohtvaatluste ruumilise ulatusega, optilise kaugseire piltidel eristuvate merepõhja laikude suurusega ja modelleerimiseks moodustuvate andmeridade hulgaga, millest sõltub matemaatiliste arvutuste ajakulu. Katsetati ka 2 m suurema ja väiksema ruudu kasutamist. 2 m suurema ruudu kasutamine ei olnud antud aladel otstarbekas kuna see skaala oleks jäänud liiga suureks arvestades loodusliku merepõhja laigulisusega, mis oli tuvastatav kaugseire piltidel ja kohtvaatlustega. Siinjuures tuli täiendavalt arvestada sellega, et kohtvaatluste sidumiseks ruudustikuga seoti proovipunkti geograafiline asukoht talle lähima nelja või üheksa ruuduga ja seetõttu suurenes mudeli treeningandmetena kasutatava treeningala geograafiline ulatus (vt joonis 2.4.1). Neli ja üheksa naabrit valiti seetõttu, et sellise naabrite arvuga saab moodustada vastavalt ligikaudu 2×2 ja 3×3 proovipunktile lähima ruudu plokki (vt joonis 2.4.1). Üheksast naabrist suurema arvu valimine (4×4) ei oleks olnud otstarbekas, sest treeningala plokk oleks omandanud liiga suured mõõtmed võrreldes ortofoto ja Hypsax andmetest nähtuvate eriilmeliste merepõhja laikude suurusega. Sellisel viisil saadud treeningandmeid kasutati Maa-ameti ortofotode ja Hypsax andmete puhul. 2 m oluliselt väiksema ruudu (alla 1 m) kasutamine ei oleks olnud otstarbekas arvutusmahu olulise kasvu tõttu ja ka põhjusel, et merepõhja videovaatluste ruumiline ulatus jääb tavaliselt vahemikku 1–4 m.

4. alal, kus katsetati päikesepeegelduse ja lainevahu väljamaskimist, teostati modelleerimine täiendavalt 5 m ruuduga, sest väljamaskimise tõttu jäi informatiivseid pikseleid mõnes piirkonnas 2 m ruuduga kasutamiseks liiga väheseks.



Joonis 2.4.1. Näide 2×2 m modelleerimisruudustikust ja treeningruutude (4 ruutu, 9 ruutu) paiknemisest proovipunktide juures. Taustaks on Maa-ameti ortofoto (2017) 2. uuringuala (Vilsandi) piirkonnast.

Kaugseire andmed

Igas 2 m modelleerimisruudustiku ruudus arvutati:

- Maa-ameti ortofotode (OF) muutujate R, G, B, R/G, G/B, R/B keskmised ja standardhälbed;
- Hypspx muutujate (12 kanalit ja 12 esimest peakomponenti) keskmised väärtused ja standardhälbed.

Optilise signaali keskmine intensiivsus ja selle varieeruvus 2 m ruudus iseloomustavad elupaikade erinevaid aspekte. Optilise signaali keskmine intensiivsus annab infot substraadi tüübi ja taimestiku ohtruse kohta. Optilise signaali varieeruvus ruudus aga iseloomustab pigem taimestiku leviku väiksemastaapseid mustreid ning kuna erinevatele liikidele on iseloomulikud erinevad kasvumustrid, siis potentsiaalselt võimaldab sellise statistiku kasutamine edukamalt eristada erinevaid taimerühmi/liike.

Kogu uurimisala katvat 2 m ruudustiku andmestikku kasutati ennustusandmestikuna ja proovipunktidega seotud ruutude andmeid treeningandmetena.

Kuna Sentinel-2 satelliidiandmete (S2) piksel oli suurem kui 2 m modelleerimisruudustiku ruut, siis ei olnud võimalik S2 puhul keskmise ja standardhälbe arvutamine ja selle asemel kasutati kõigi kanalite pikslite väärtusi 2 m ruudustiku ruutude keskpunktidest. S2 puhul saadi treeningandmete väärtused nendest satelliidipildi pikslitest, kuhu langesid proovipunktide koordinaadid.

Merepõhja kohtvaatlused

Kohtvaatluste andmed sisestati Tartu Ülikooli Eesti mereinstituudi põhjaelustiku andmebaasi (Microsoft Access) ja modelleerimise jaoks andmete saamiseks tehti andmepäring sellest andmebaasist. Modelleerimiseks valiti ja loodi eelkõige sellised merepõhja substraadi ja elustiku muutujad, mida kasutatakse tavapärasel kaardistustöödel sh Euroopa Liidu loodusdirektiivi (92/43/EMÜ) elupaigatüüpide kaardistamiseks. Loodusdirektiivi elupaigatüüpidest on antud töö kontekstis fookuses karide elupaigatüüp (kood 1170) ja liivamadalate (1110) taimestikuvöönd. Liivamadalate tunnusliigiks on ka sette sees elavad karbid kuid kuna neid ei ole võimalik merepõhja visuaalse vaatlusega tuvastada ja nad ei anna optilist signaali, siis seda tunnusrühma selles töös ei käsitletud. Täiendavalt lisati mõned muutujad, mis võivad olla olulised just optilise kaugseire seisukohalt. Modelleeritavad muutujad olid järgmised:

- Kõva põhjasubstraadi katvus (%) – vajalik merepõhja elupaikade, sh loodusdirektiivi karide elupaigatüübi, eristamiseks;
- Liiva katvus (%) – vajalik merepõhja elupaikade, sh loodusdirektiivi liivamadalate elupaigatüübi, eristamiseks;
- HUB substraaditüüp (nominaaltunnus) – põhjasubstraadi tüüp HELCOM HUB elupaikade klassifikatsiooni 3. taseme järgi (HELCOM 2013). HUB substraaditüüp on merestrategia raamdirektiivi (2008/56/EÜ) merepõhja elupaikade põhitüüpide (*benthic broad habitat types*) Läänemere analoogiks.
- Põisadru katvus (%) – põhjataimestiku üks võtmeliike, sh loodusdirektiivi karide elupaigatüübi tunnusliik;
- Agariku katvus (%) – põhjataimestiku üks võtmeliike, sh loodusdirektiivi karide elupaigatüübi tunnusliik;
- Niitjate vetikate katvus (%) – loodusdirektiivi karide elupaigatüübi tunnusrühm;
- Kõrgemate taimede (va pikk merihein) katvus (%) – loodusdirektiivi liivamadalate elupaigatüübi tunnusrühm;
- Pika meriheina katvus (%) – põhjataimestiku üks võtmeliike, sh loodusdirektiivi liivamadalate elupaigatüübi tunnusliik;
- Mändvetikate katvus (%) – loodusdirektiivi liivamadalate elupaigatüübi tunnusrühm;
- Söödava rannakarbi katvus (%) – põhjaloomastiku üks võtmeliike, sh loodusdirektiivi karide elupaigatüübi tunnusliik;
- Põhjataimestiku üldkatvus (%) – põhjataimestiku koosluste üldine kirjeldav tunnus; kuna iseloomustab põhjataimestiku ohtrust horisontaalsel tasapinnal, siis võib olla sobilik muutuja, mida saab hinnata kaugseire abil;
- Rohevetikate katvus (%) – suur vetikate rühm, mille pigmentide koosseisu optilised omadused peaksid võimaldama neid kaugseire abil eristada;
- Pruunvetikate katvus (%) – suur vetikate rühm, mille pigmentide koosseisu optilised omadused peaksid võimaldama neid kaugseire abil eristada;
- Punavetikate katvus (%) – suur vetikate rühm, mille pigmentide koosseisu optilised omadused peaksid võimaldama neid kaugseire abil eristada;
- Lahtiste niitjate vetikate katvus (%) – substraadilt lahti tulnud niitjad vetikad võivad moodustada ulatuslikke liikuvaid matte ja katta erinevaid substraate ja taimekooslusi;
- Kogu taimestiku summaarne katvus (%) – kinnitunud ja lahtise taimestiku summaarne katvus; kuna visuaalselt ei ole lahtised ja kinnitunud vetikad hästi eristatavad, siis võib kaugseire

seisukohalt olla mõistlik nad summeerida. Väärtuse maksimum piiratud 100 isegi kui summaarne epifüütide või rinnetena kasvava taimestiku katvus ületab 100;

- Kinnitunud taimestiku summaarne katvus (%) – kõvale substraadile kinnitunud vetikate ja pehmele substraadile kinnitunud kõrgemate taimede summaarne katvus. Muutuja on loodud kaugseire seisukohalt, et hinnata taimede summaarset ohtrust. Väärtuse maksimum piiratud 100 isegi kui summaarne epifüütide või rinnetena kasvava taimestiku katvus ületab 100.

Treeningandmete loomisel seoti proovipunktide merepõhja elustiku ja substraadi andmed modelleerimisruudustikuga (vt seksioon Modelleerimisruudustik) ja läbi selle kaugseire andmetega.

Modelleerimisalgoritmid

Matemaatiliseks modelleerimiseks testiti kolme erinevat algoritmi:

- Juhumets (RF, *random forest*);
- Võimendatud regressioonipuud (BRT, *boosted regression trees*);
- Tehisnärvivõrgud (ANN, *artificial neural networks*).

RF on ansamblimeetod, mille puhul luuakse suur hulk otsuste puid kasutades iga puu loomisel tagasipanekuga juhuvalimit (*bootstrap*) tunnustest ja vaatlustest (Remm et al 2012). Seejärel kombineeritakse lõpptulemuseks parima ennustusvõimega otsuste puud. Ennustamisel saadakse lõplik ennustatava muutuja väärtus üksikute puude tulemuste keskmistamisel, kui tegemist on pideva tunnusega või hääletamise teel, kui tegemist on faktortunnusega. Kirjanduse andmeil saavutatakse stabiilseid tulemusi vähemalt 500 puuga (Liaw & Wiener 2002). Antud töös oli puude arv 500-1000.

BRT on ansamblimeetod, mis ühendab endas regressioonipuude ja võimendamise tugevusi. Meetod kujutab endast suure hulga regressioonipuude loomist, kus iga järgnev puu üritab sisendandmeruumis kirjeldada seda osa, mis eelmistel puudel kirjeldamata jäi. Mudelite loomisel seati puu keerukuseks (*tree complexity*) kolm, mis tähendab, et lubatud on sõltumatute muutujate kuni kolmekaupaa interaktsioonid. Õppimiskiirus (*learning rate*) määrati esialgu 0,1, aga kui see ei andnud stabiliseerunud tulemust (ennustuse hajuvuse ühtlane ja asümptootilisele sarnane vähenemine puude lisandumisel), siis vähendati edasi kuni 0,01.

ANN on meetod, mis imiteerib bioloogiliste olendite aju toimimist – õppimisvõimet ja võimet vähese lähteinfo järgi üldistusi teha (Remm et al 2012). ANN koosneb omavahel seotud sõlmedest (*node*), mis on aju neuronite analoogid ja mis on kihilise ülesehitused. Sisendkihi, mis saab sisendi otse õpetusandmetest ja väljundkihi vahel võib olla üks või mitu varjatud sõlmede kihti. ANN õpetamine ehk treenimine on regressioonimudeli parameetrite sobitamise analoog, mille käigus üritatakse minimeerida väljundi viga. Esmalt antakse ANN sõlmi ühendavatele kaaludele (regressioonikordajata analoogid) juhuslikud väärtused ja iteratiivse treenimise käigus täpsustatakse kaalusid, et saavutada võimalikult väike ennustuse viga (Remm et al 2012).

Kõik mudelarvutused viidi läbi vabavaralises statistikatarkvaras R 3.6.1 (R Core Team 2019). RF meetodi jaoks kasutati R-i paketti *randomForest* (Liaw & Wiener 2002, Breiman et al 2018), BRT jaoks pakette *gbm* (Greenwell et al 2019) ja *dismo* (Hijmans et al 2017) ning ANN jaoks paketti *neuralnet* (Fritsch et al 2019).

Sõltumatute muutujate olulisust RF mudelites hinnati paketi *randomForest* permutatsioonitesti abil, mille mõõdetavaks parameetrikas regressioonimudeli puhul oli mudeli ruutvea suurenemise protsent konkreetse muutuja väljajätmise korral mudelist. Klassifikatsioonimudeli puhul oli mõõdetavaks

parameetriks klassifitseerimise edukuse vähenemise protsent konkreetse muutuja väljajätmise korral mudelist.

Täiendavad muutujad ja mudelite komplektid

Merepõhja elustiku ja elupaikade kaardistustöödel, kus ei ole võimalik kasutada optilise või akustilise kaugseire andmeid, tuleb proovipunktipõhistest andmetest pindandmete tekitamiseks kasutada matemaatilisel modelleerimisel olemasolevaid keskkonnaandmeid pindandmete (raster) kujul. Peamisteks kasutatavateks muutujateks on vee sügavus ja avatus lainetusele (Isæus 2004), mis on saadaval piisavalt kõrge ruumilise resolutsiooniga, et neid saaks kasutada ka väiksemate alade kaardistamiseks. Eestis läbi viidud uuringutes on korduvalt leidnud tõendamist, et sügavus ja avatus lainetusele seonduvad väga hästi merepõhja substraadi ja elustiku levikumustritega (nt TÜ Eesti Mereinstituut 2014, Herkül et al 2016, Peterson & Herkül 2019). Üle-Eestilistes modelleerimistöodes on lisaks nendele muutujatele kasutatud ka suurt hulka muid muutujaid (vt nt TÜ Eesti Mereinstituut 2014, TÜ Eesti Mereinstituut 2018) kuid enamik nendest muutujatest ei sobi oma loodusliku varieeruvuse või andmekihi tehnilise ruumilahutuse tõttu väiksematel aladel (kaardistatava ala ruumiline ulatus kuni ca 20 km) kasutamiseks. Erinevatele keskkonnaandmetele lisaks on täiendavalt mudelisse võimalik sõltumatute muutujatena panna geograafilised koordinaadid meetermõõdukus. Geograafilised koordinaadid võivad parandada modelleerimistulemusi kui uuringuala merepõhja substraadi ja elustiku levikumustrites on selgelt väljenduvaid ruumigradiente, mida olemasolevad keskkonnamuutujad ei kajasta. Käesolevas töös kasutati täiendavate muutujatena järgmisi muutujaid:

- Mere sügavus (põhineb Veeteede Ameti sügavusandmetel);
- Avatus lainetusele lihtsustatud lainemudeli põhised (Nikolopoulos & Isæus 2008);
- Geograafilised koordinaadid Eesti riikliku koordinaatsüsteemi (1997) meetermõõdukus;
- Ala number faktortunnusena aladeülestes mudelites.

Täiendavate muutujate kasutamise peamiseks eesmärgiks oli selgitada välja kas ja mil määral optilise kaugseire andmete kasutamine parandab matemaatiliste mudelite ennustusvõimet ja kuidas kaugseire andmete kasutamine mõjuta mudelennustuste tulemusel valminud merepõhja substraadi ja elustiku levikukaarte. Treeniti erinevaid mudeleid, kus sisendiks olid optilise kaugseire andmed

Lisaks erinevatele treeningruutude arvule (4 ja 9) proovipunkti ümbruses ja modelleerimisalgoritmidele testiti OF ja Hypsax puhul eraldi mudelikomplekte, mis olid:

- Treenitud korruga üle kõigi alade andmete;
- Treenitud eraldi iga ala kohta.

Mudelite valideerimine

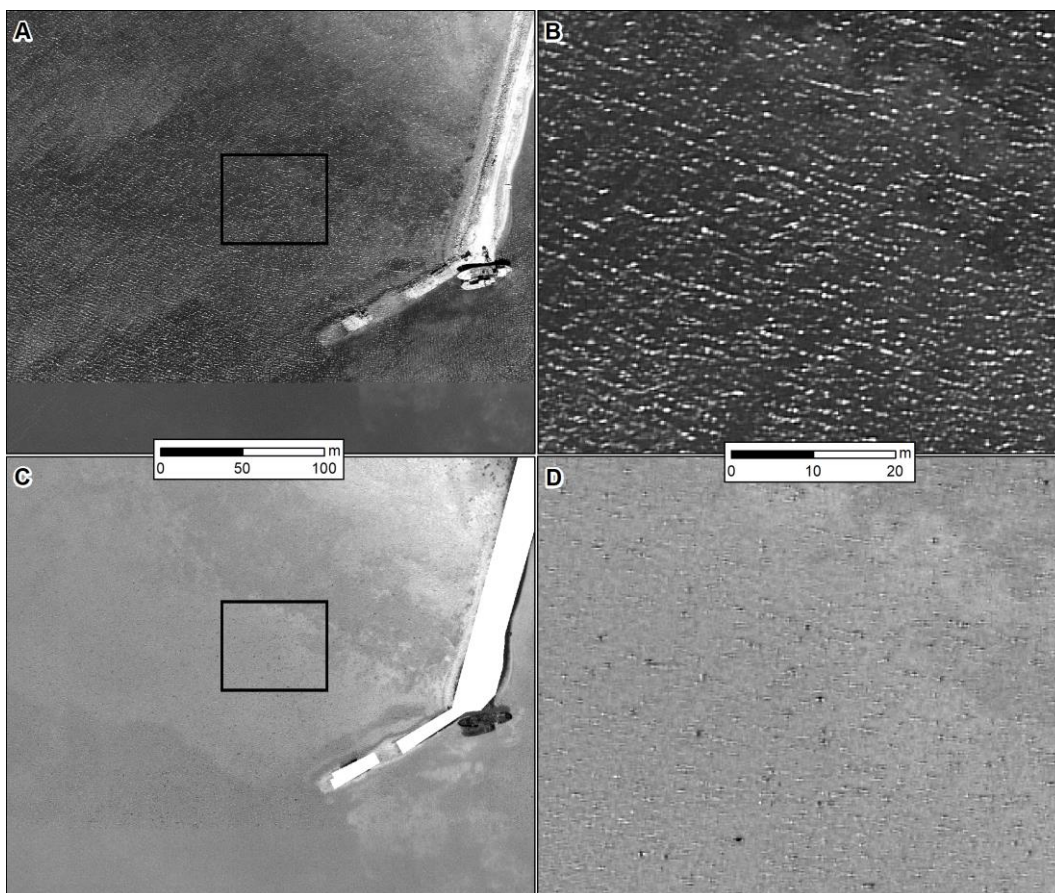
Mudelennustuste valideerimiseks kasutati nii sisemist valideerimist, välimist valideerimist kui visuaalset eksperthinnangut. Sisemisel valideerimisel teostati ennustused samadele andmepunktidele, mida kasutati mudeli treenimiseks. Välimisel valideerimisel treeniti spetsiaalselt eraldi mudelid, milles kasutati 75% juhuslikult valitud sisendandmeid ja ülejäänud 25% andmete abil valideeriti mudelite ennustusvõimet. Väline valideerimine teostati omakorda kahes eraldi variandis: (1) juhuvalim modelleerimisruudu põhised, kus treeninguks ja valideerimiseks valiti andmepunkti kõigi modelleerimisruutude seast, (2) juhuvalim proovipunktide põhised, kus kõik ühe konkreetse proovipunktiga seotud ruudud läksid juhuvalimi alusel kas treening või valideerimisandmeteks. Pidevtunnuseliste muutujate mudelite headust mõõdeti lineaarse korrelatsioonikordaja (r) ja keskmise absoluutvea (MEA, *mean*

absolute error) abil. Nominaaltunnuste juures kasutati Coheni vastavuskordajat ehk kapa kordajat ja klassifitseerimise edukust (E ; õigesti klassifitseeritud vaatluste osa kõigist vaatlustest). r , K ja E varieeruvad 0 ja 1 vahel ja kõrgem väärtus näitab paremat ennustusvõimet. MEA on samades mõõtühikutes, mis mõõdetav parameeter ja kõrgem väärtus näitab suuremat viga ehk halvemat ennustusvõimet. Lisaks matemaatilisele valideerimisele anti mudelennustustele ka visuaalne eksperthinnang. Visuaalne eksperthinnang on mudelite valideerimisel oluline komponent, sest matemaatiline valideerimine ei kajasta adekvaatselt nende piirkondade ruumimustreid, kust puuduvad proovipunktid. Mitmesugused modelleerimisega seotud vead, näiteks mudelite ülesobitamisest tingitud väga järsud üleminekud või muud looduses mitte-esinevad mustrid on tuvastatavad ainult mudelennustuse visuaalse ülevaate käigus.

3. Tulemused

3.1. Maa-ameti ortofotod

Uuringualade valikul oli lähtunud sellest, et nende kohta on olemas hea kvaliteediga OF ja et merepõhjas esineks elupaigalist varieeruvust. Vaatamata eelvalikule esines siiski mõningaid probleeme seoses päikesepeegeldusega (*sun glint*). Hedley et al (2005) meetodil lähisinfrapuna kanali abil peegelduse efekti vähendamine oli üsna tõhus, kuid jättis pildile siiski müra, mis suurendas oluliselt väärtuste varieeruvust (standardhälve) modelleerimisruutudes (joonis 3.1.1). Müra tõttu ei olnud otstarbekas standardhälbe kasutamine mudelites, sest mudelennustuste visuaalsel vaatlusel oli näha, et see tekitas vigaseid ennustusi kohtades, mis olid päikesepeegeldusest tugevasti mõjutatud. 4. alal katsetati päikesehelgiga ja lainetuse vahuga pikslite väljamaskimist sinise kanali intensiivsuse ja punase-sinise kanali intensiivsuse suhte alusel ja selle tulemuslikkus oli suhteliselt hea. Väljamaskimise korral oli võimalik mudelites kasutada ka modelleerimisruudu OF muutujate standardhälvet ja mudelennustuste visuaalse hinnangu alusel andis see lähenemine mõnevõrra paremaid tulemusi kui päikesepeegelduse korrigeerimine Hedley et al (2005) meetodil ja standardhälbe mudelist väljajätmine.



Joonis 3.1.1. Näide Maa-ameti ortofoto rohelisest kanalist enne (A-B) ja pärast päikesepeegelduse korrigeerimist (C-D) 1. alal Jaagarahu sadama lähedal. Pildi A allosas on näha kahe erineva lennujoone üleminekut, mille erinevused on tunduvalt vähenenud korrigeeritud pildis (C). Pildil D on näha, et vaatamata korrigeerimisele on pildis siiski palju müra.

Kui modelleerimisalgoritme juhumeets (RF) ja võimendatud regressioonipuud (BRT) on varasemates merepõhja substraadi ja elustiku leviku kaardistamise töödes kasutatud laialdaselt (nt TÜ Eesti Mereinstituut 2014, TÜ Eesti Mereinstituut 2016, TÜ Eesti Mereinstituut 2018), siis tehisnärvivõrkude (ANN) meetodit varem kasutatud ei ole. Arvestades käesoleva töö väga suure analüütilise töö mahuga ja asjaoluga, et ANN puhul ei ole universaalseid reegleid varjatud kihtide arvu ja nendes paiknevate sõlmede arvu valimiseks ja parima tulemuse leidmiseks tuleb seetõttu proovida erinevate kombinatsioonidega, katsetati ANN meetodit ainult 4. alal põhjataimestiku üldkatvuse juures nelja naabriga treeningandmetega. Ühe varjatud kihi puhul katsetati 2, 3, 4 ja 5 sõlmega mudeleid ning kahe varjatud kihi korral vastavalt 2 + 1 ja 3 + 2 sõlmega mudeleid. Erinevate üldkatvuse mudelite puhul jäi ANN mudelite korrelatsioonikordaja (10-kordne sisemine ristvalideerimine) vahemikku ligikaudu 0,35–0,4 samas kui RF saavutas korrelatsioonikordaja üle 0,6. Seetõttu ei kasutatud ANN meetodit ülejäänud alade ja ülejäänud muutujate modelleerimisel.

RF ja BRT, 4 ja 9 naabriga treeningruutude, aladekaupa ja aladeüleste ning erinevate sõltumatute muutujate komplektidega mudelite matemaatilise valideerimise keskmistatud tulemused on esitatud lisa 1. Olulisemad leiud on järgmised:

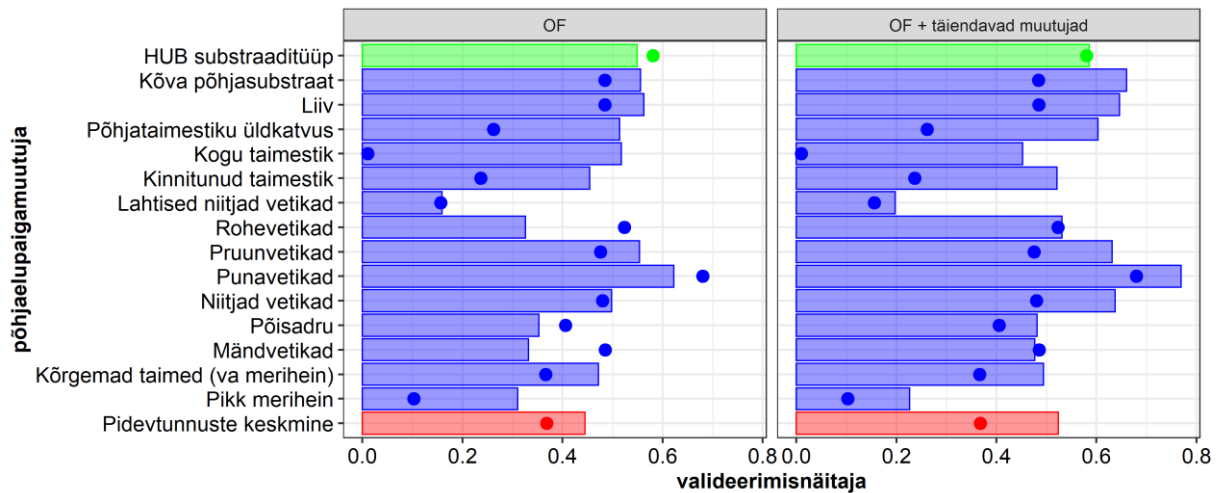
- OF kasutamine parandas mudelennustuste täpsust võrreldes mudeliga, kus ortofotosid polnud kasutatud;
- 9 naabriga treeningandmed andsid paremaid tulemusi kui 4 naabriga treeningandmed;
- Alade kaupa mudelid andsid paremaid tulemusi kui aladeüleste mudelid;
- RF andis paremaid tulemusi kui BRT;
- Täiendavate muutujate (sügavus, avatus lainetusele, geograafilised koordinaadid) lisamine parandas mudelite ennustusvõimet matemaatilisel valideerimisel kuid need muutujad võisid lisada ebaloomulikke mustreid mudelennustustesse (joonis 3.1.4).

Matemaatilise valideerimise tulemuste lihtsamaks ja ülevaatlikumaks esitamiseks ning kaugseire andmete mõju hindamiseks valiti välja kolm mudelit, mis kõik olid teostatud RF meetodil 9 naabriga ja alade kaupa:

- Ainult täiendavate muutujatega (sügavus, lainetus, geograafilised koordinaadid) mudel ehk mudel, kus kaugseire andmeid ei ole kasutatud (lisa 1 mudel nr 1) („täiendavad muutujad“);
- Ainult kaugseire andmetega mudel, kus sõltumatuteks muutujateks olid kanalite ja kanalisuhete keskmised (lisa 1 mudel nr 9) („OF“);
- Kaugseire andmeid ja täiendavaid muutujaid sisaldav mudel kus sõltumatuteks muutujateks olid mõlema ülaltoodud mudeli sõltumatud muutujad (lisa 1 mudel nr 10) („OF + täiendavad muutujad“).

Modelleeritavate põhjaelupaigamuutujate mudelite omavaheline võrdlus näitas, et OF-muutujad parandasid kõige rohkem taimestiku üldist ohtrust näitavate muutujate (põhjataimestiku üldkatvus, kogu taimestiku katvus, kinnitunud taimestiku katvus) ennustusi (joonis 3.1.2). OF oli täiendavate muutujate mudelist oluliselt parem ka pika meriheina korral. Köva põhjasubstraadi ja liiva ennustus olid OF korral mõnevõrra paremad kui täiendavate muutujate korral. Mõningate põhjaelupaigamuutujate korral (eelkõige rohevetikad, mändvetikad) andis täiendavate muutujate mudel paremaid tulemusi kui OF mudel. OF + täiendavate muutujate mudelid andsid paremaid tulemusi kui ainult täiendavate muutujate mudelid kõigil juhtudel välja arvatud HUB substraat, rohevetikad ja

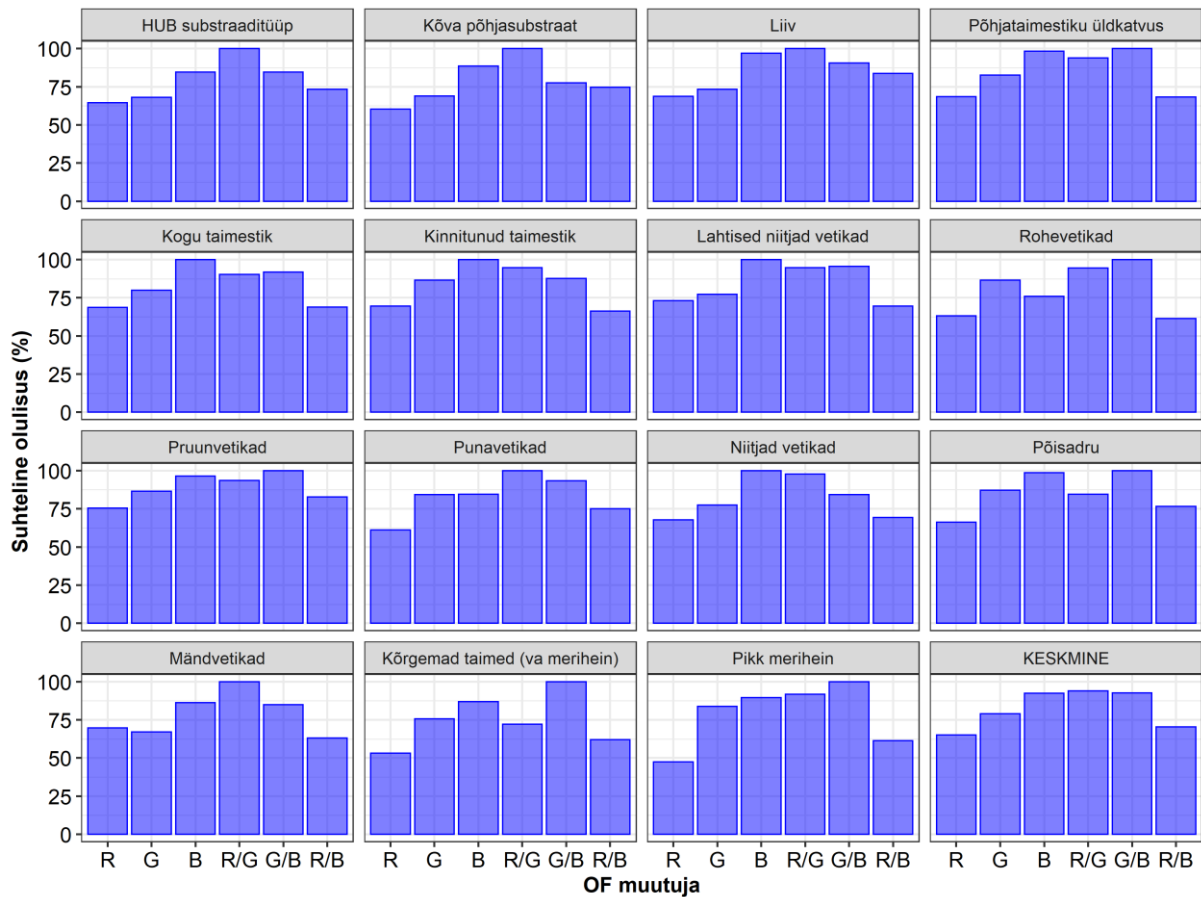
mändvetikad, mille korral tulemused olid samad (joonis 3.1.2). Kõige halvema tulemuslikkusega oli lahtiste niitjate vetikate leviku modelleerimine, mis on ka ootuspärane tulemus kuna tegemist on ajutise ja väga liikuva nähtusega. Lahtiste vetikate modelleerimise halva tulemuse juures tuleb kindlasti arvestada, et käesolevas töös kasutati 2017. aasta Maa-ameti ortofotosid. Lahtised vetikamatid on väga probleemsed ka tavapärase kaardistustööde kohtvaatluste puhul, sest kuna nad katavad ja varjavad merepõhja, siis on videovaatlustes äärmiselt keeruline tuvastada tegelikke merepõhja omadusi; sageli on probleemiks ka kinnitunud ja lahtiste vetikate eristamine.



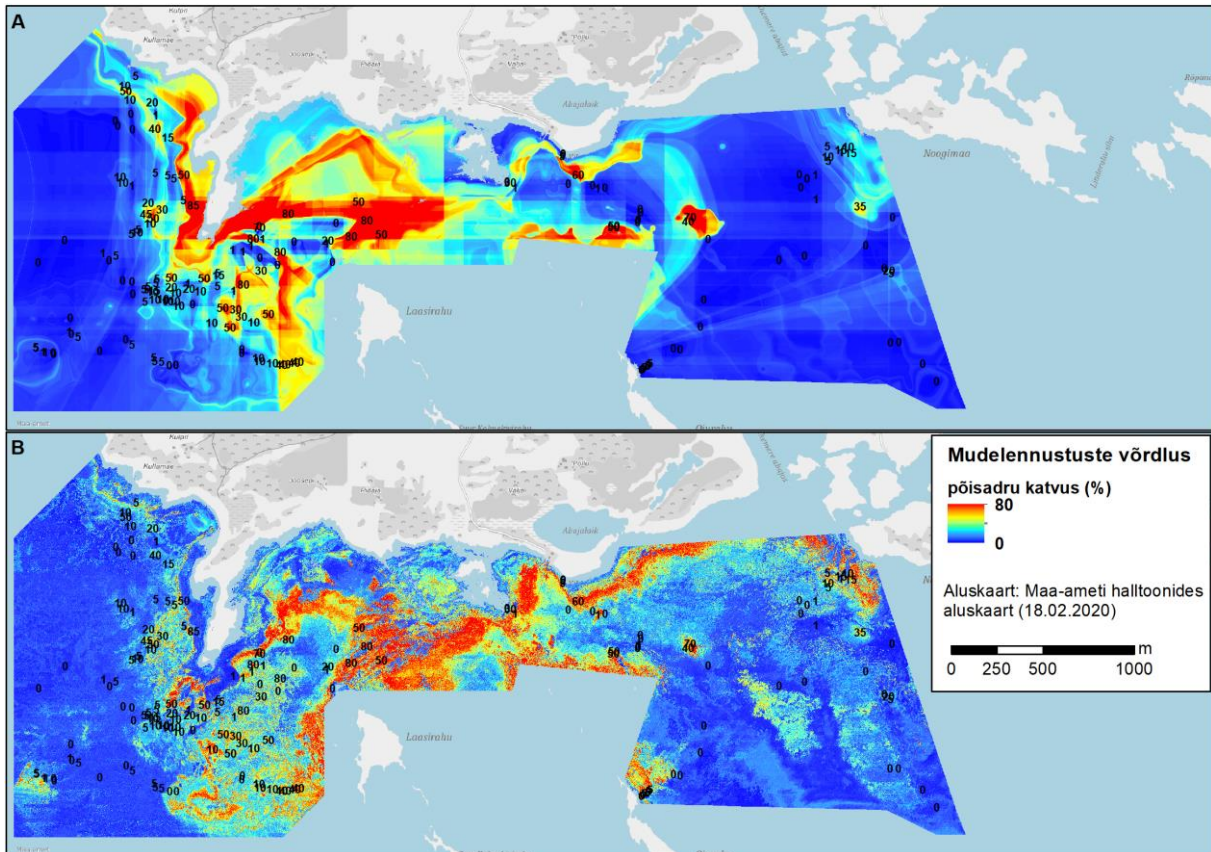
Joonis 3.1.2. Põhjaelupaigamuutujate mudelite, kus sõltumatuteks muutujateks olid ainult täiendavad muutujad (punktid), ainult OF muutujad (tulbad) ning OF ja täiendavad muutujaid (tulbad), valideerimisnäitajate ülevaade. Rohelisega on nominaaltunnused (valideermisnäitajaks E), sinisega pidevtunnused (r), punasega pidevtunnuste keskmine (r).

OF muutujate olulisuse analüüs näitas, et väga suuri erinevusi muutujate vahel ei esinenud (joonis 3.1.3). Keskmisena oli kõige olulisem muutuja punase ja rohelse kanali suhe (R/G), aga sellele väga sarnase olulisusega olid ka sinine kanal (B) ja rohelse ja sinise suhe (G/B).

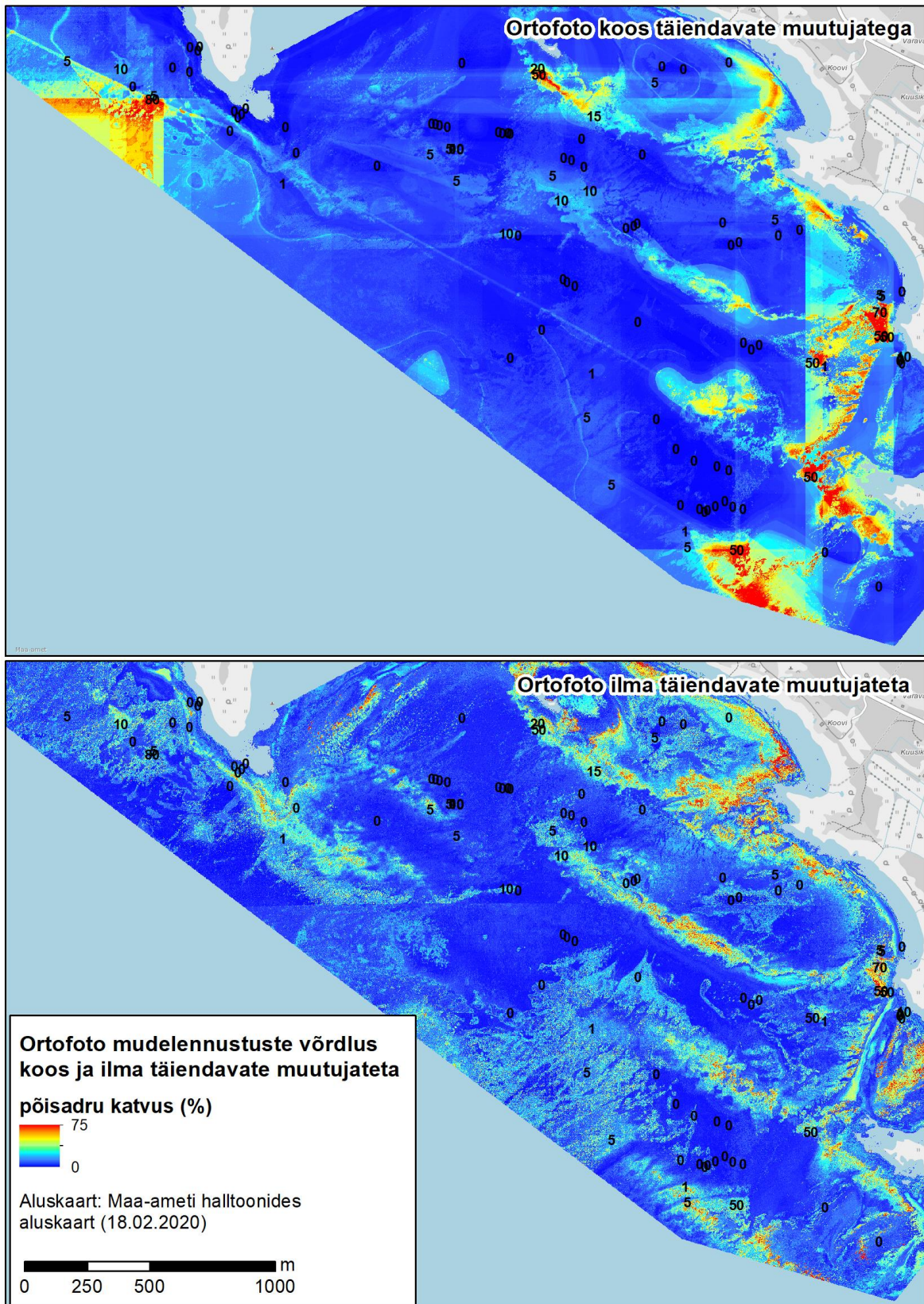
Täiendavate muutujate (sügavus, avatus lainetusele, geograafilised koordinaadid) lisamine mudelisse koos OF andmetega valdavalt parandas mudelite ennustusvõimet (joonis 3.1.2) kuid samas väga sageli lisas ruumiennustusse looduses mitte-esinevaid mustreid. Joonisel 3.1.5 on näidatud põisadru mudelennustused 3. alal koos ja ilma täiendavate muutujateta. Antud näite puhul on täiendavata muutujatega mudeli matemaatiline ennustusvõime kõrgem, aga samas on mudelennustuse kaardikihil näha geograafiliste koordinaatidega ja sügavusega seotud modelleerimisvigasid – järsud põhja-lõuna- ja ida-lääne-suunalised üleminekud ja sügavuse samajooni järgivad väga kitsad kõrgema väärtusega vööndid (joonis 3.1.5). Sarnaseid näiteid võib tuua kõigilt aladelt ja erinevate modelleeritavate muutujate puhul. Seetõttu ei saa usaldada ainult matemaatilise valideerimise tulemusi vaid kõik tulemused peavad läbima visuaalse eksperthinnangu, mis oluliselt suurendab modelleerimistöde ajakulu.



Joonis 3.1.3. Maa-ameti ortofotopõhiste (OF) muutujate olulisus põhjasubstraadi ja –elustiku muutujate mudelites. Modelleeritavate muutujate hulgast jäeti välja söödava rannakarbi ja agariku katvus, sest nende esinemissagedus oli liiga madal adekvaatsete mudelite loomiseks. Olulisuse hindamisel on rakendatud mudelit, kus olid kasutusel OF muutujate keskväärtsused (mudel nr 9 lisas 1). Suhteline olulisus on võrreldavuse jaoks skaleeritud kõige olulisema muutuja järgi (kõige olulisem = 100).



Joonis 3.1.4. Põisadru katvuse mudelennustuste võrdlus 2. alal. A – mudel, kus ei ole kasutatud Maa-ameti ortofotosid (mudel nr 1 lisas 1). B – mudel, kus on kasutatud Maa-ameti ortofotosid (mudel nr 5 lisas 1). Rasterpind on saadud modelleerimise tulemusel. Numbrid näitavad väärtusi proovipunktides.

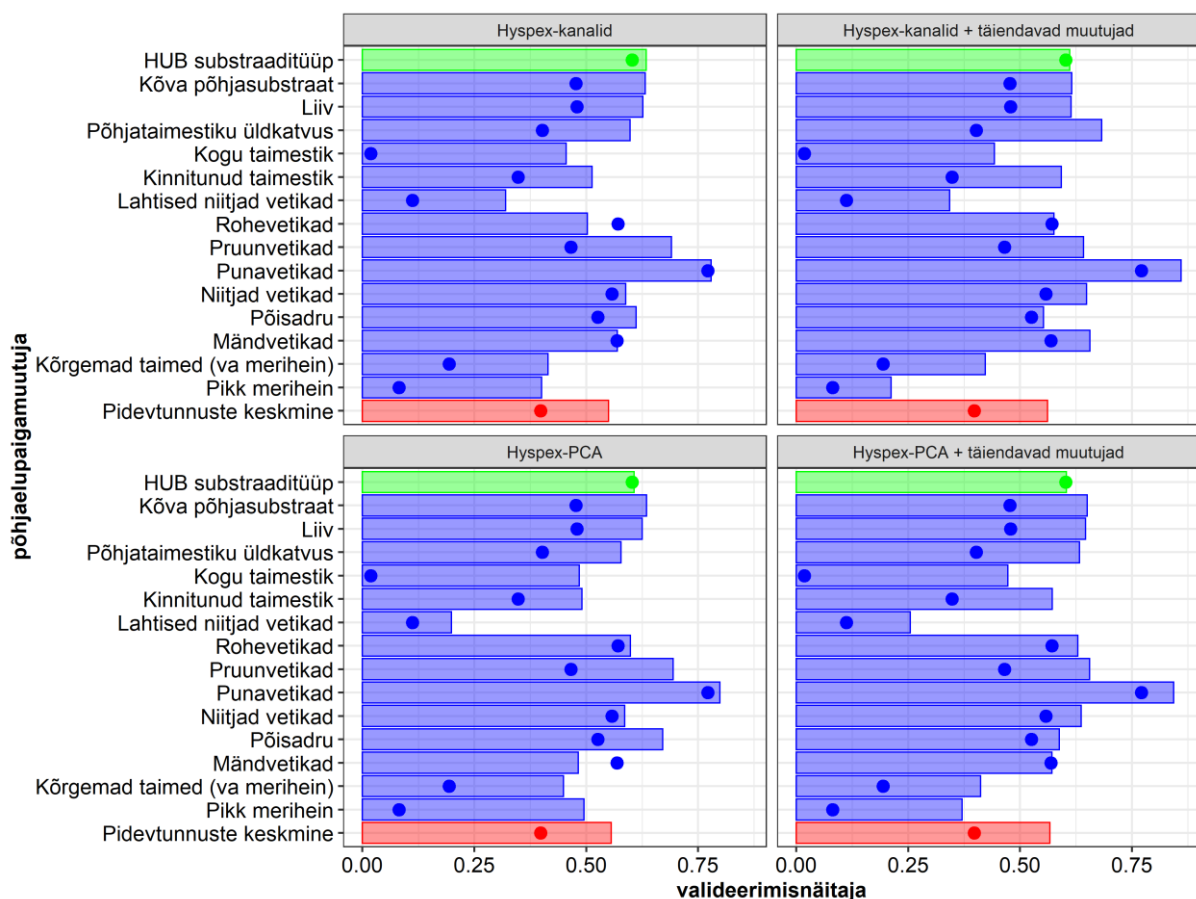


Joonis 3.1.5. Põisadru katvuse mudelennustuste võrdlus 3. alal: koos täiendavate muutujatega ja ilma täiendavate muutujateta. Rasterpind on saadud modelleerimise tulemusel. Numbrid näitavad väärtusi proovipunktides.

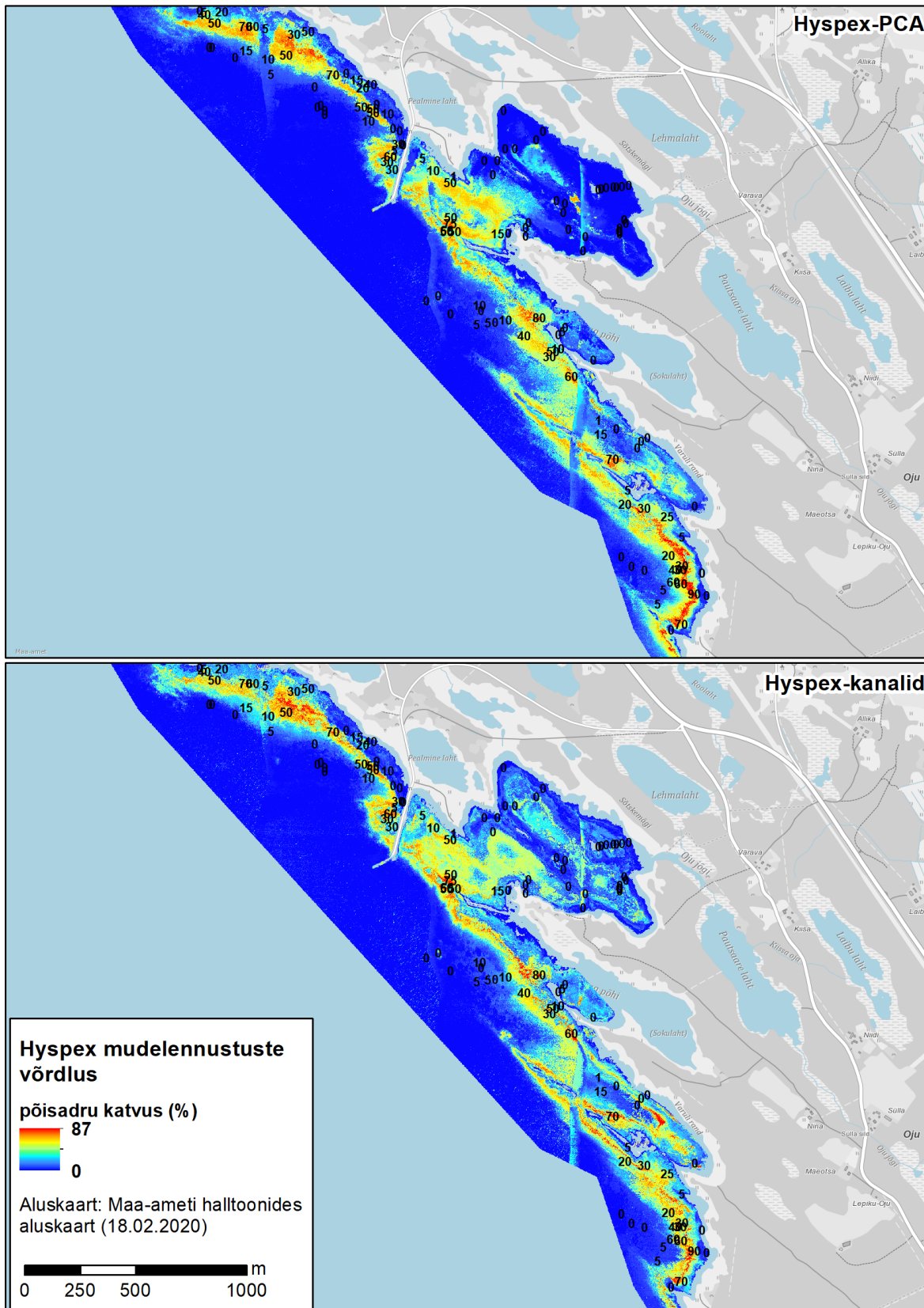
3.2. Hypsrex

Kuna OF materjali põhjal tehtud mudelid näitasid, et RF, 9 naabriga treeningandmed ja alakaupa mudelid andsid paremaid tulemusi kui BRT, 4 naabriga treeningandmed ja aladeülesed mudelid, siis teostati Hypsrex materjaliga ainult RF 9 naabriga alakaupa mudelid. Sarnaselt OF materjalile, parandas ka Hypsrex kasutamine mudelite ennustusvõimet (lisa 2, joonis 3.2.1). Sarnaselt OF mudelitele, oli ka Hypsrex mudelite puhul kõige halvema valideerimistulemusega lahtiste vetikate mudel (joonis 3.2.1). Kanalipõhine ja PCA-põhine Hypsrex andmestik andsid matemaatilisel valideerimisel väga sarnaseid tulemusi aga mudelennustuste visuaalses pildis oli siiski näha teatavaid erinevusi (joonis 3.2.2). Visuaalse eksperthinnangu järgi võib mõnevõrra paremaks pidada Hypsrex-PCA versiooni (joonis 3.2.2). Hypsrex andmete kasutamise probleemseks aspektiks oli mudelennustuste rasterkihtides selgesti eristatavad lennujoonte servad, mis on nähtavad põhja-lõunasuunaliste ümbritsevast selgesti eristuvate kõrgemate või madalamate väärtustena (joonis 3.2.2).

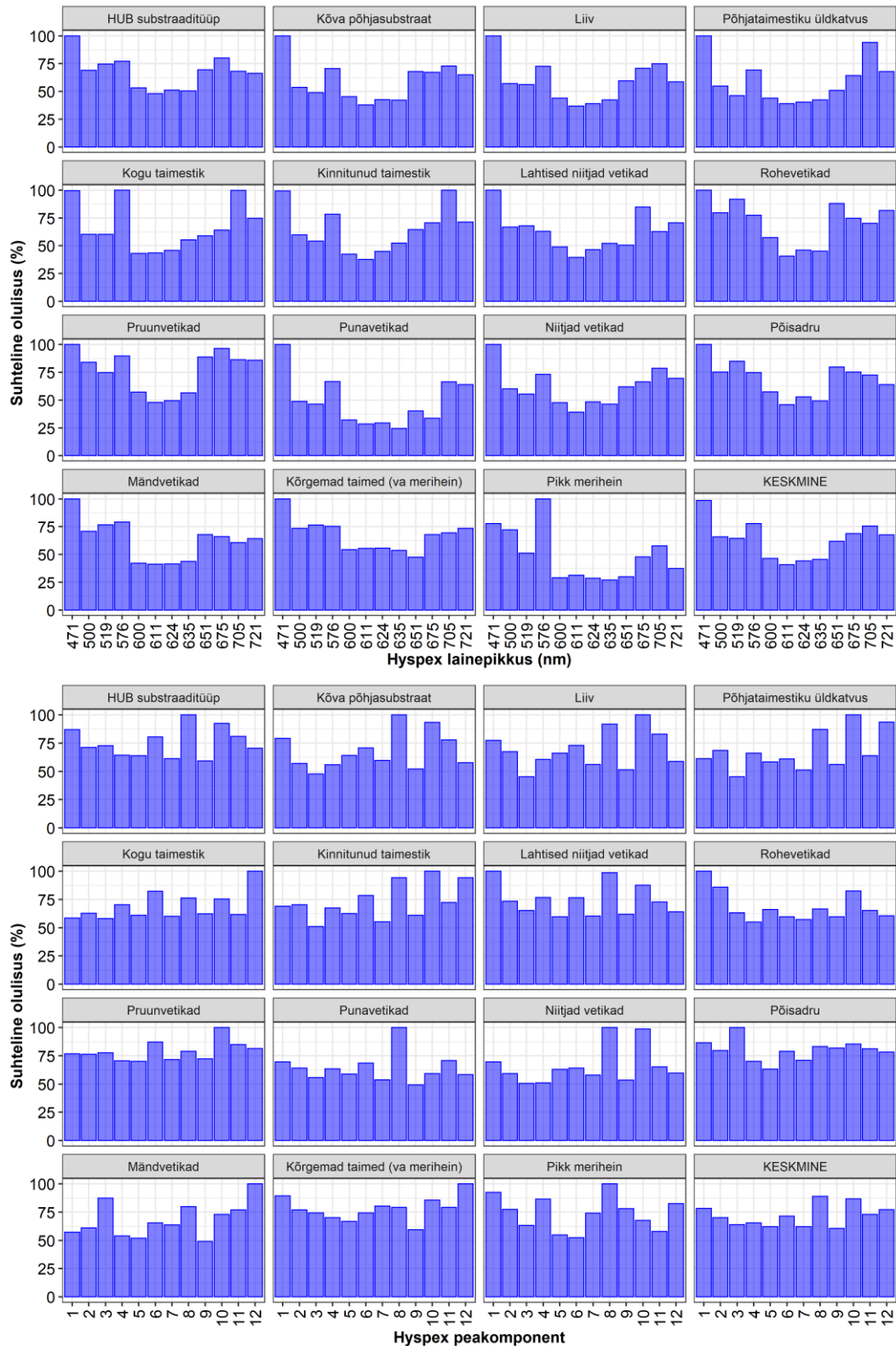
Hypsrex kanalite olulisuse analüüs näitas, et keskmiselt oli kõige olulisem kõige lühem lainepikkus (471 nm), mis oli kõige olulisem kanal kõigi põhjaelupaigamuutujate puhul ja pikk merihein ja kõige vähemolulised olid keskmised lainepikkused (600–635 nm) (joonis 3.2.3). PCA-põhiste andmete juures selget domineerimisstruktuuri välja tuua ei saa (joonis 3.2.3).



Joonis 3.2.1. Põhjaelupaigamuutujate mudelite, kus sõltumatuteks muutujateks olid ainult täiendavad muutujad (punktid), ainult Hypsrex muutujad (tulbad) ning Hypsrex ja täiendavad muutujaid (tulbad), valideerimismäärtajate ülevaade. Rohelisega on nominaaltunnused (valideerimismäärtajaks E), sinisega pidevtunnused (r), punasega pidevtunnuste keskmine (r).



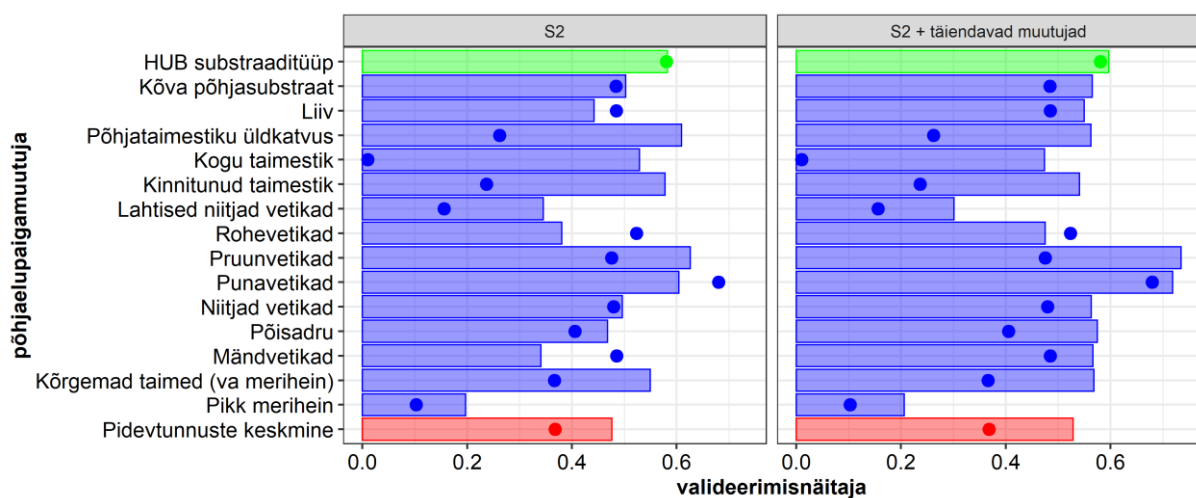
Joonis 3.2.2. Põisadru katvuse mudelennustuste võrdlus 1. alal, kus sisendina on kasutatud Hypsax-PCA ja Hypsax-kanalid andmeid. Rasterpind on saadud modelleerimise tulemusel. Numbrid näitavad väärtusi proovipunktides.



Joonis 3.2.3. Hyspex muutujate olulisus põhjasubstraadi ja –elustiku muutujate mudelites. Modelleeritavate muutujate hulgast jäeti välja söödava rannakarbi ja agariku katvus, sest nende esinemissagedus oli liiga madal adekvaatsete mudelite loomiseks. Olulisuse hindamisel on rakendatud mudeleid, kus kasutusel olid Hyspex muutujate keskvaartused (mudelid nr 5 ja 7 lisas 2). Suhteline olulisus on võrreldavuse jaoks skaleeritud kõige olulisema muutuja järgi (kõige olulisem = 100).

3.3. Satelliidipildid

Sarnaselt OF ja Hypsrex kaugseire andmetele parandas ka Sentinel-2 satelliidipiltide (S2) kasutamine mudelite ennustusvõimet (lisa 3, joonis 3.3.1). Kõige suurem ennustusvõime paranemine S2 andmetega mudelites võrreldes ilma kaugseire andmeteta mudelitega, oli seotud üldise taimeistiku ohtrusega (joonis 3.3.1). Erinevalt Hypsrex kanalitest, kus selgelt kõige olulisem oli kõige lühema lainepikkusega kanal, oli S2 puhul olulisuse muster erinevate põhjaelupaigamuutujate vahel üsna varieeruv (joonis 3.3.2). Näited satelliidipiltide andmete abil tehtud mudelennustustest on toodud erinevate kaugseire meetodite võrdluse peatükis 3.4.



Joonis 3.3.1. Põhjaelupaigamuutujate mudelite, kus sõltumatuteks muutujateks olid ainult täiendavad muutujad (punktid), ainult S2 muutujad (tulbad) ning S2 ja täiendavad muutujaid (tulbad), valideerimisnäitajate ülevaade. Rohelisega on nominaaltunnused (valideerimisnäitajaks E), sinisega pidevtunnused (r), punasega pidevtunnuste keskmine (r).

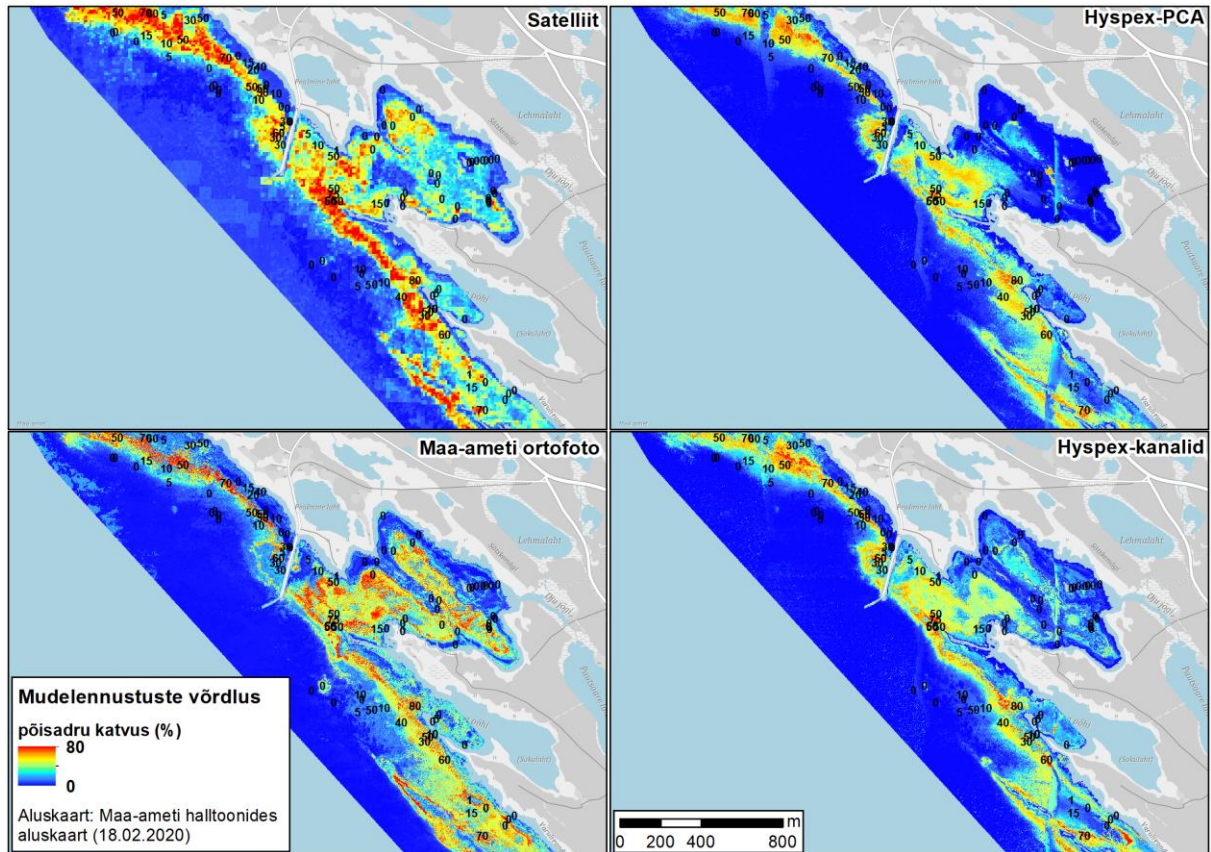


Joonis 3.3.2. Sentinel-2 (S2) muutujate olulisus põhjasubstraadi ja –elustiku muutujate mudelites. Modelleeritavate muutujate hulgast jäeti välja söödava rannakarbi ja agariku katvus, sest nende esinemissagedus oli liiga madal adekvaatsete mudelite loomiseks. Suhteline olulisus on võrreldavuse jaoks skaleeritud kõige olulisema muutuja järgi (kõige olulisem = 100).

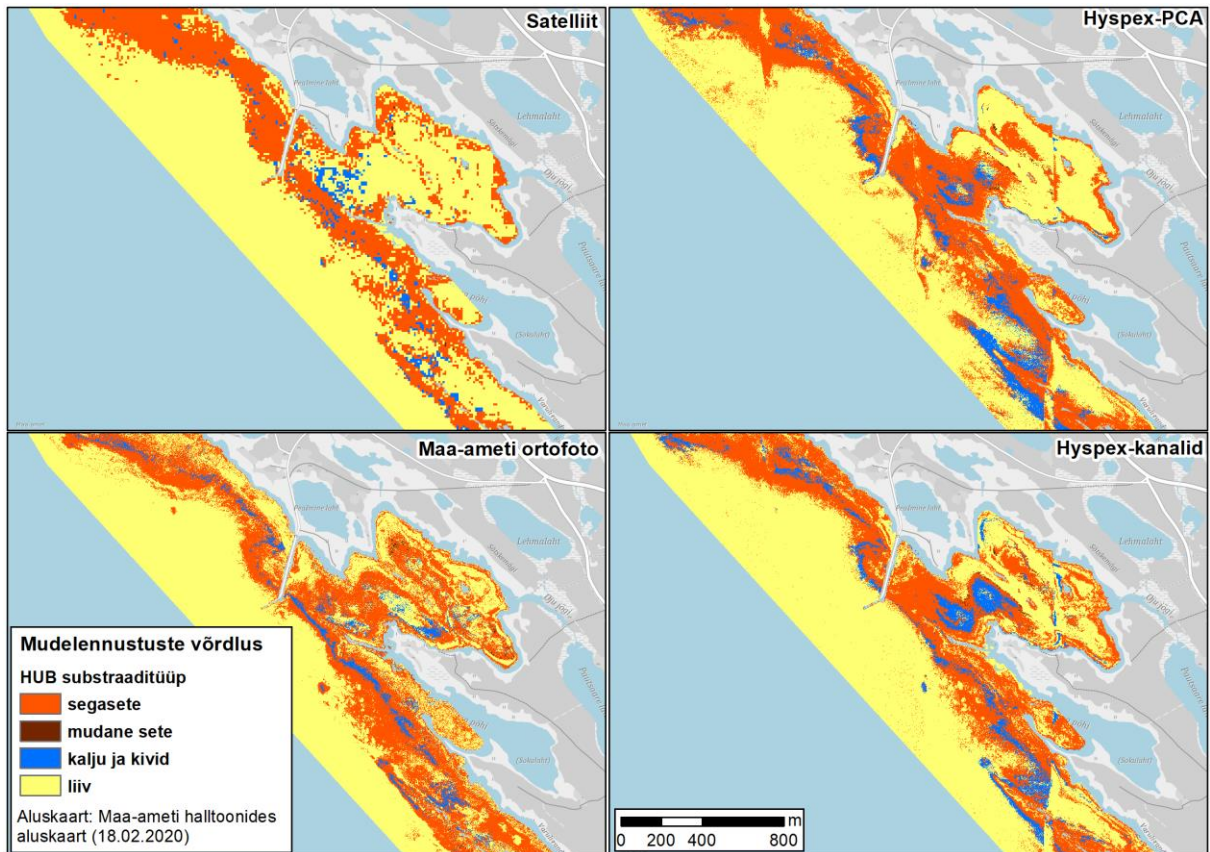
3.4. Erinevate kaugseire meetodite võrdlus

Erinevate kaugseire andmete põhjal loodud mudelennustuste võrdlusest nähtub, et nad peegeldavad üldjoontes sarnaseid merepõhja elustiku ja elupaikade ruumimustreid (vt joonised 3.4.1–3.4.3). Lähemal vaatlusel on siiski tuvastatavad erinevused. Joonisel 3.4.1 on toodud näide põisadru katvuse mudelennustuste tulemustest 1. ala põhjaosas. Sellest võrdlusest on ühe olulise erinevusena näha, et satelliidi- (S2) ja ortofotopõhised (OF) mudelid prognoosisid kõrget põisadru katvust ka näidatud ala idaosa lahesopis samas kui Hypsax mudelid, eriti PCA versioon, näitas väga madalat katvust või põisadru puudumist. Kuna kohtvaatlused selles lahesopis põisadru ei tuvastanud, siis tuleb adekvaatsemaks pidada Hypsax tulemust, eriti Hypsax PCA-põhist tulemust. Hypsaxi parem tulemus on ilmselt seotud sellega, et sellel instrumendil on teiste kaugseire andmetega võrreldes väga palju kõrgem spektraalne lahutus, mis võimaldab täpsemalt taimerühmasid omavahel eristada. Samas võib Hypsaxipõhise ennustuse probleemkohana välja tuua ebaadekvaatsed mustrid lennujoonte servades (vt vertikaalsed mustrid joonistel 3.4.1–3.4.3).

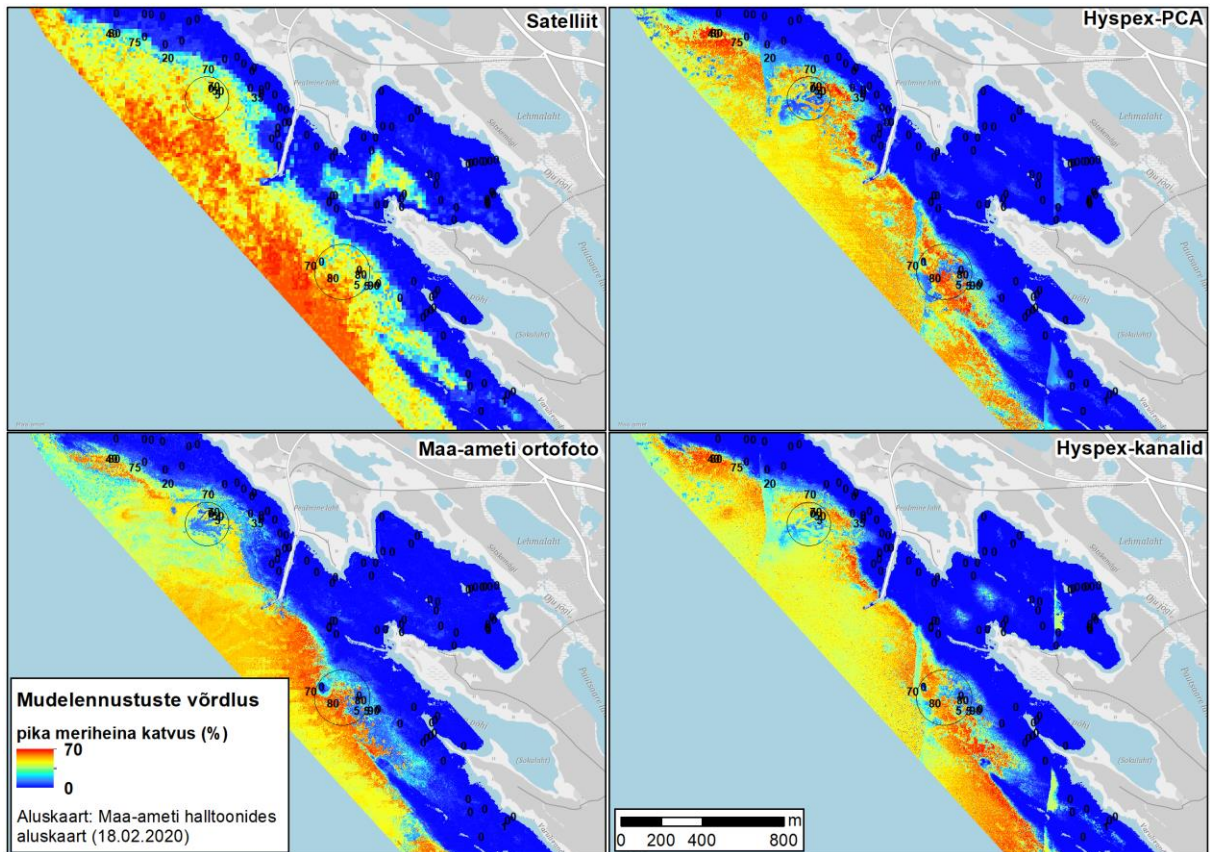
Kõrgema ruumilise lahutusega meetodite eelis satelliidiandmete ees tuleb selgelt välja pika meriheina mudelennustustes, kus S2-põhises mudelennustuses ei ole võimalik eristada pika meriheina laikusid (vt joonised 3.4.3–3.4.5).



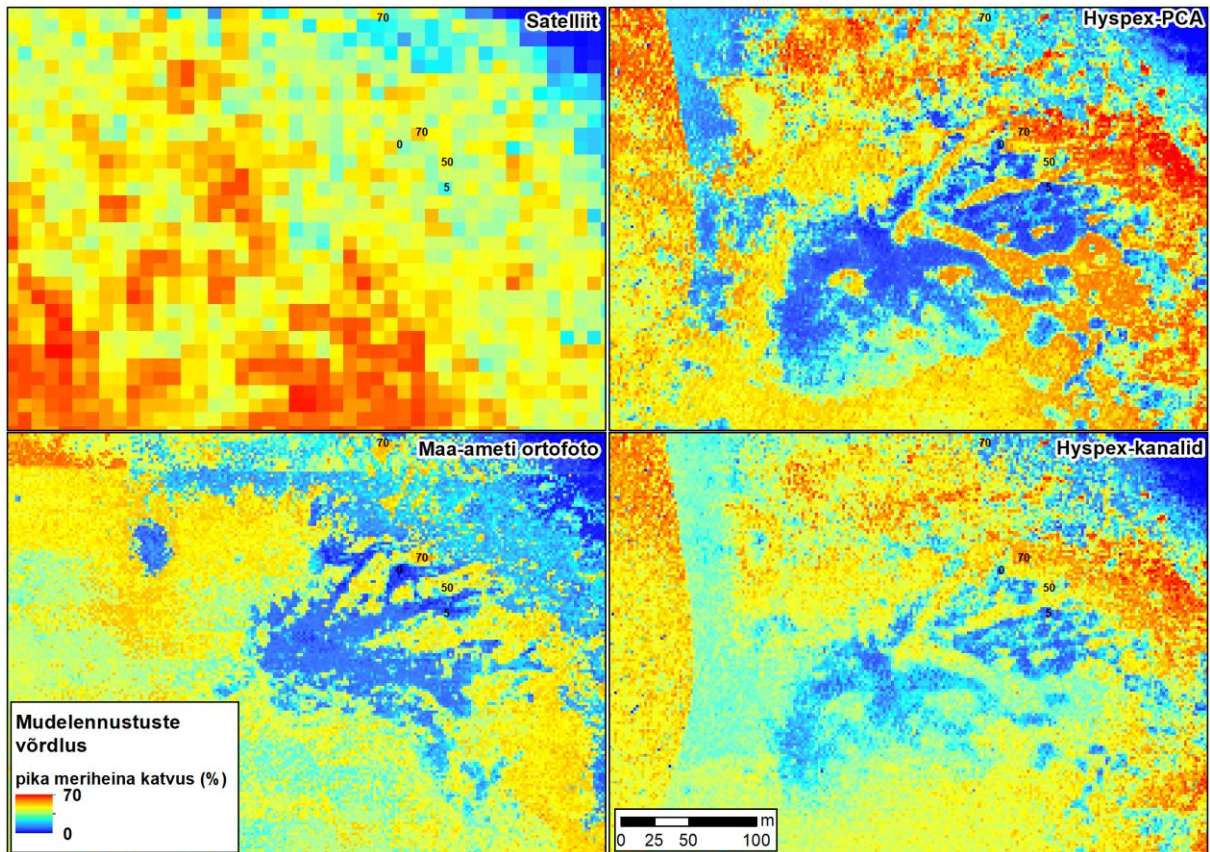
Joonis 3.4.1. Põisadru katvuse mudelennustuste võrdlus 1. ala põhjaosas. Rasterpind on saadud modelleerimise tulemusel. Numbrid näitavad väärtusi proovipunktides.



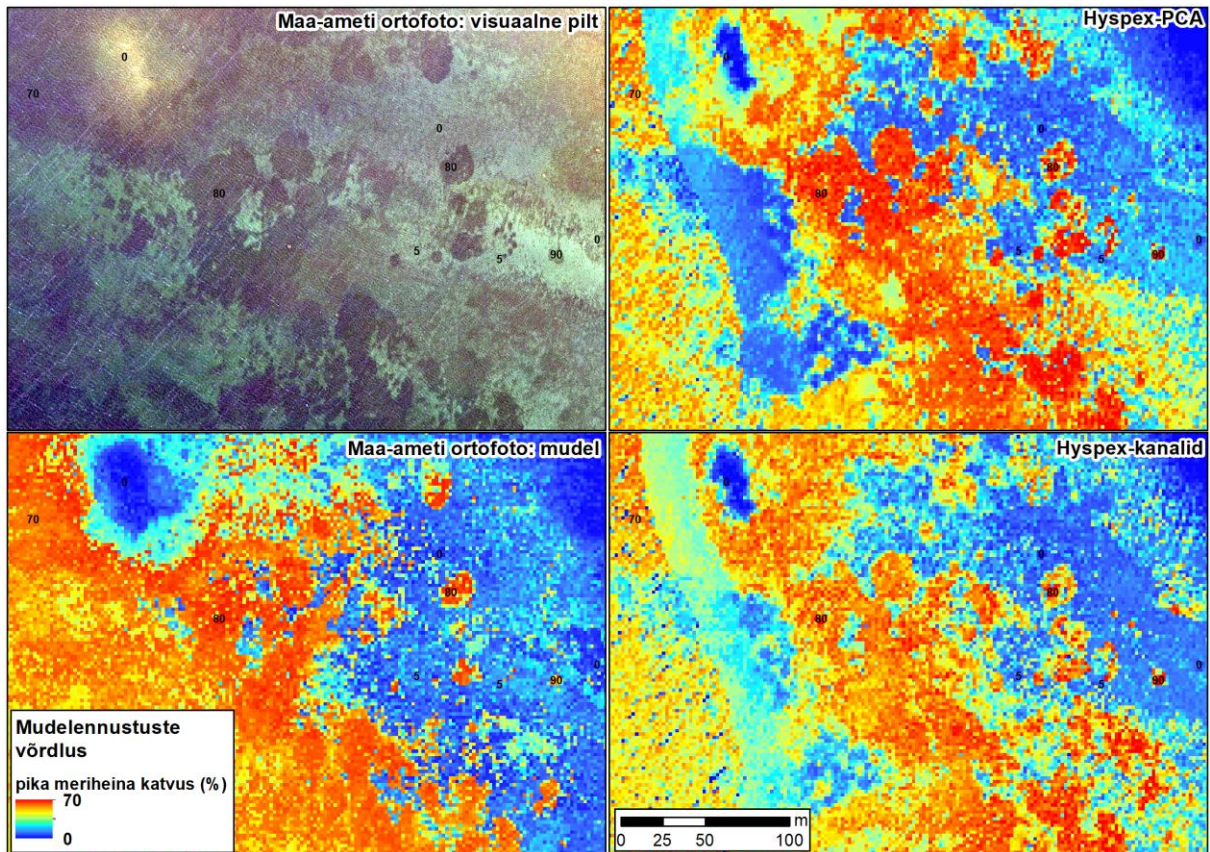
Joonis 3.4.2. HELCOM HUB elupaigaklassifikatsiooni 3. taseme substraaditüübi mudelenustuste võrdlus 1. ala põhjaosas. Rasterpind on saadud modelleerimise tulemusel.



Joonis 3.4.3. Pika meriheina katvuse mudelennustuste võrdlus 1. ala põhjaosas. Rasterpind on saadud modelleerimise tulemusel. Numbrid näitavad väärtusi proovipunktides. Ringiga on näidatud alad, millest on suurendused järgmistel joonistel.



Joonis 3.4.4. Pika meriheina katvuse mudelennustuste võrdlus 1. ala põhjaosas – suurendatud piirkond, mis on näidatud põhjapoolse ringiga joonisel 3.4.3. Rasterpind on saadud modelleerimise tulemusel. Numbrid näitavad väärtusi proovipunktides.

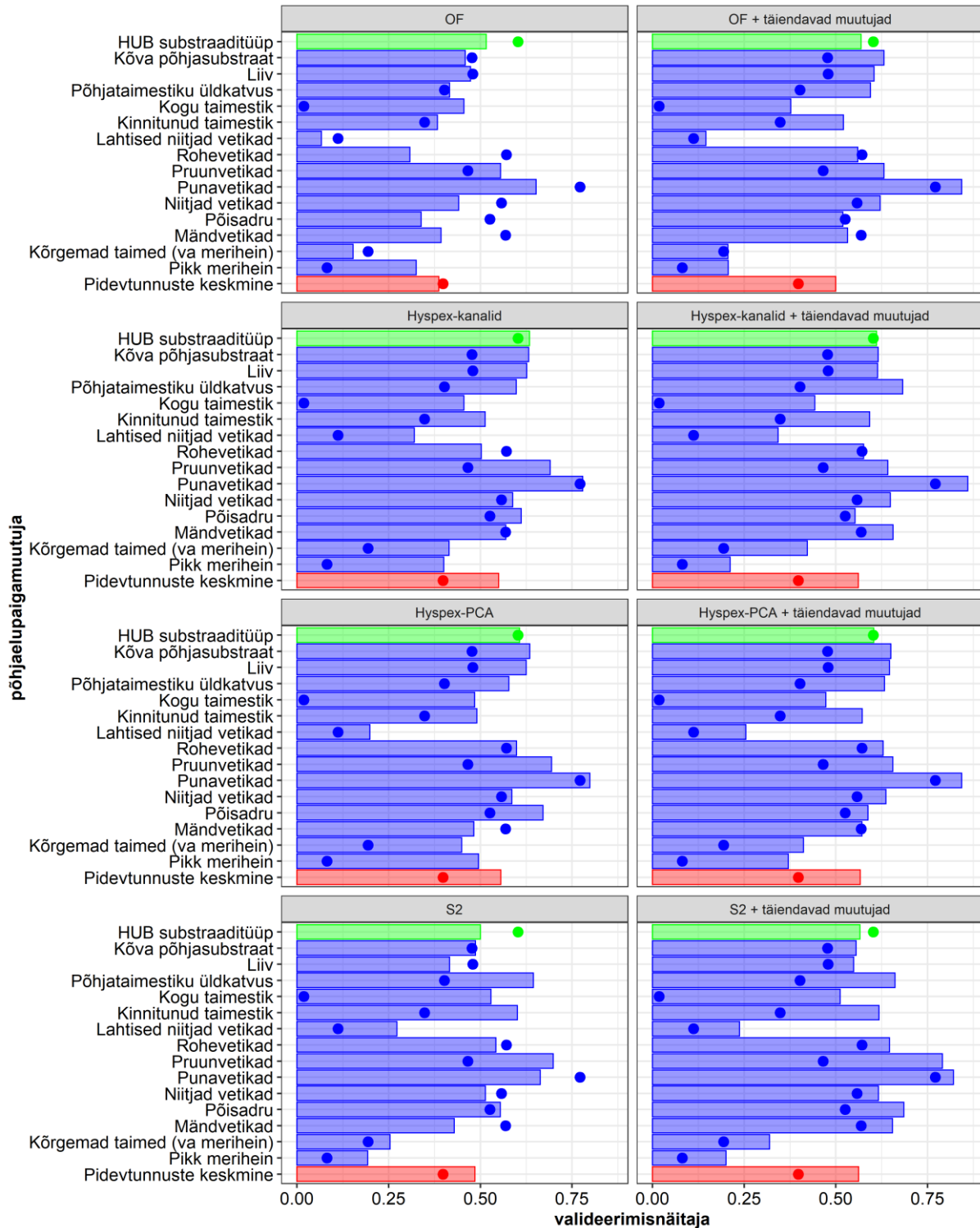


Joonis 3.4.5. Pika meriheina katvuse mudelennustuste võrdlus 1. ala põhjaosas – suurendatud piirkond, mis on näidatud lõunapoolse ringiga joonisel 3.4.3. Rasterpind on saadud modelleerimise tulemusel. Numbrid näitavad väärtusi proovipunktides. Eraldi on üleval vasakul toodud Maa-ameti ortofoto visuaalne pilt.

Erinevatel kaugseire andmetel põhinevate mudelite matemaatilise võrdlemise jaoks koondati valideerimisandmed ainult aladelt 1 ja 3, sest Hypsrex andmed ei olnud saadaval aladelt 2 ja 4 (joonis 2.1.1). Nende alade mudelite matemaatilise valideerimise tulemuste võrdlusest (joonis 3.1.5) saab teha järgmised järeldused:

- Hypsrex andmetel põhinevad mudelennustused ühtisid kõige paremini kohtvaatluste andmetega. Selle põhjuseks on ilmselt Hypsrex andmete väga palju kõrgem spektraalne lahutus võrreldes teiste meetoditega.
- Sentinel-2 satelliidipildid (S2) tagasid mõnevõrra nõrgema ennustusvõime kui Hypsrex aga kõrgema ennustusvõime kui ortofoto (OF).
- OF mudelite ennustusvõime oli kõige madalam. See on seletatav kahe asjaoluga: OF spektraalne resolutsioon oli kõige madalam (3 kanalit) ja ortofotod pärinesid 2017. aastast samas kui teiste meetodite andmed pärinesid 2019. aastast. OF puhul tuleb antud võrdluses märkida, et kõiki alasid sisaldavas valideerimises andis OF mõnevõrra paremaid tulemusi (vt joonis 3.1.2).
- Kaugseire andmete lisamine parandas ilma kaugseire andmeteta mudelitega võrreldes enim taimestiku üldist ohtrust kirjeldavate muutujate, pruunvetikate ja pika meriheina valideerimistulemusi.

- Üle kõigi kaugseire andmete oli kõige halvema valideerimistulemusega lahtiste niitjate vetikate katvus. See on ootuspärane tulemus kuna tegemist on ajutise ja väga liikuva nähtusega.
- Substraadimuutujatest ei parandanud kaugseire kasutamine HUB substraaditüübi (nominaaltunnus) ennustust kuid parandas mõnevõrra kõva põhjasubstraadi ja liiva katvuse ennustusi. HUB substraaditüübi puhul võib kaugseire olla probleemne seetõttu, et taimestik katab sageli merepõhja substraati ja seetõttu ei ole alati võimalik saada otsest signaali substraadilt.



Joonis 3.1.5. Põhjaelupaigamuutujate katvushinnangute (ainult pidevtunnused) mudelennustuste täpsus väljendatuna välise valideerimise lineaarse korrelatsioonikordajaga (r) ainult OF andmetega mudelites (mudel nr 9 tabelis 3.1.1) ning OF ja täiendavaid muutujaid sisaldavates mudelites (mudel nr 10 tabelis 3.1.1). Välja on jäetud södava rannakarbi ja agariku katvus, sest nende esinemissagedus oli liiga madal adekvaatsete mudelite loomiseks.

4. Kokkuvõte

Uuringu eesmärgiks oli testida erinevate kaugseire meetodite rakendatavust madalate merealade merepõhja elupaikade (põhjasubstraadi ja põhjaelustiku) kaardistamisel. Uuring viidi läbi neljal uuringualal Saaremaa läänerannikul 2019. aasta mais ja juunis.

Testitavateks kaugseire meetoditeks olid:

- Maa-ameti ortofotod (OF), piksli suurus 0,2 m;
- hüperspektraalse spektromeetriga Hypslex lennukilt kogutud andmed, piksli suurus (0,8 m);
- Sentinel-2 satelliidipildid (S2), piksli suurus olenevalt spektraalsest kanalist 10-60 m.

Elupaikade kohtvaatlused teostati peamiselt allveevideo abil. Kohtvaatlustel registreeriti põhjaelupaikade muutujate väärtused, mida hinnatakse ka tavapäraste põhjaelupaikade kaardistamise ja seire käigus: põhjataimestiku üldkatvus, põhjataimestiku ja loomastiku liikide/rühmade katvused ja põhjasubstraadi tüüpide katvused protsentuaalselt.

Matemaatiliste seoste loomiseks kohtvaatlustel tuvastatud elupaigamuutujate ja kaugseirest saadud andmete vahel testiti juhumetsa (*random forest*), võimendatud regressioonipuude (*boosted regressioon trees*) ja tehisnärvivõrkude (*artificial neural networks*) modelleerimismeetodeid. Lisaks kaugseiremuutujatele kasutati ka täiendavaid muutujaid (sügavus, avatus lainetusele, geograafilised koordinaadid), et võrrelda kas ja milles avaldub kaugseire muutujate mudelitesse lisamise mõju. Loodud mudelite abil teostati mudelennustused 2 m ruudustikus kogu uuringuala piires. Mudelennustuste headuse hindamiseks kasutati matemaatilist valideerimist ja visuaalset eksperthinnangut.

Erinevate kaugseire meetodite abil saadud mudelennustuste visuaalse eksperthinnangu ja matemaatilise valideerimise tulemusel saab väita:

- Kõigi kolme kaugseire meetodi (OF, Hypslex, S2) kasutamine võimaldab tõsta mudelennustuste täpsust võrreldes mudeliga, kus kaugseire andmeid ei kasutata.
- Juhumetsa modelleerimismeetod oli parim oma ennustuste täpsuse ja praktilise rakendatavuse poolset.
- Erinevalt matemaatilise valideerimise tulemustest, mis näitasid, et kaugseire andmete kasutamisel paraneb mudelite ennustusvõime mõõdukalt, võib visuaalse eksperthinnangu põhjal järeldada, et mudelennustuste kvaliteedi tõus on märkimisväärne. Kaugseire andmete lisamine mudelisse muutis mudelennustustes nähtavad mustrid looduslikele sarnasemaks; kaugseire abil toodetud mudelennustustes olid jälgitavad piirid ja üleminekud, mis olid ka kaugseire pildidel visuaalselt tuvastatavad.
- Erinevate kaugseire meetoditega saadud andmetel põhinevate mudelennustuste matemaatiline valideerimine näitas:
 - Kaugseire andmete lisamine parandas ilma kaugseire andmeteta mudelitega võrreldes enim taimestiku üldist ohtrust kirjeldavate muutujate, pruunvetikate ja pika meriheina valideerimistulemusi.

- Hypspx andmed võimaldavad saada teiste meetoditega võrreldes kõrgema ennustusvõimega mudeleid. Selle põhjuseks võib pidada Hypspx andmete väga palju kõrgemat spektraalset lahutust võrreldes teiste meetoditega.
- Satelliidipildid (S2) tagasid mõnevõrra nõrgema ennustusvõime kui Hypspx aga kõrgema ennustusvõime kui ortofoto (OF).
- OF mudelite ennustusvõime oli kõige madalam, mis on seletatav kahe asjaoluga: OF spektraalne resolutsioon oli kõige madalam (3 kanalit) ja ortofotod pärinesid 2017. aastast samas kui teiste meetodite andmed pärinesid 2019. aastast.

Euroopa Liidu kohustustega seotud elupaikade kaardistamise vajadusi silmas pidades saab töö tulemuste põhjal teha järgmised järeldused:

- Optiline kaugseire sobib madalate merealade (antud uuringulade näitel kohati maksimaalselt 6 m) loodusdirektiivi (92/43/EMÜ) taimestikuga seotud elupaikade vööndite kaardistustöödel täpsema tulemuse saavutamiseks võrreldes ainult kohtvaatlustel ja modelleerimisel põhineva meetodikaga:
 - karide (1170) elupaigatüübi taimestiku tunnusliikidega seotud alad: põisadru, agarik, niitjad vetikad;
 - liivamadalate (1110) elupaigatüübi taimestiku tunnusliikidega seotud alad: pikk merihein, teised kõrgemad taimed, määndvetikad.
- Merestrategie raamdirektiivi (2008/56/EÜ) merepõhja elupaikade põhitüüpide aluseks oleva HELCOM HUB substraaditüübi kaardistamisel ei parandanud optilise kaugseire andmete kasutamine matemaatilise valideerimise põhjal mudelite ennustusvõimet. Visuaalse eksperthinnangu põhjal võib siiski väita, et optiliste andmete kasutamine võimaldab looduslikele sarnasemate ruumimustritega mudelennustuste loomist.

Saadud tulemuste puhul peab arvestama, et uuringud viidi läbi Saaremaa läänerannikul, kus vee läbipaistvus on parem kui suuremal osal ülejäänud Eesti merealal. Kuna kaugseire abil merepõhja kaardistamist limiteerib vee läbipaistvus, siis ei saa uuringus saadud tulemusi ja meetodite rakendatavust otse üle kanda kogu Eesti merealale.

Tabelis 4.1 on toodud analüüside tulemuste ja nende interpretatsiooni kokkuvõtte.

Tabel 4.1. Erinevate kaugseire meetodite rakendatavuse kokkuvõte. Värvitaust näitab meetodi rakendatavuse hinnangut: roheline – parim, valge – neutraalne või ebarelevantne, kollane – ebaotstarbekas või probleemne, punane – halvim või sobimatu.

Omadus	Maa-ameti ortofoto	Hypex lennukilt	Satelliit Sentinel-2
Kasutatavus merepõhja elupaigamuutujate kõrge ruumilise lahtusega (laigud < 100 m) kaardistamiseks	Võib olla sobilik, kui materjal on saadaval sobivast aastast, hooajast ja piisava kvaliteediga (materjali kogumine ei ole optimeeritud veealadele ja võib esineda palju päikesepeegeldust, lainetust, lainevahtu)	Sobilik, andis teiste meetoditega võrreldes kõige täpsemad mudelennustused	Sobimatu: piksli suurus ≥ 10 m ja seetõttu ei kajasta väikesemõõtmelist elupaikade varieeruvust
Kasutatavus merepõhja elupaigamuutujate madala ruumilise lahtusega (laigud > 100 m) kaardistamiseks	vt eelmine	vt eelmine	Sobilik
Kasutatavus merepõhja elupaigamuutujate kaardistamisel väikestel uuringualadel (kuni ca 100 km ²)	vt eelmine	vt eelmine	Sobilik, kui ei ole vajalik kõrge ruumilahutusega tulemus
Kasutatavus merepõhja elupaigamuutujate üle-Eestisel kaardistamisel	Ei sobi kuna üle-Eestiline OF materjal kombineeritakse mitmest erinevast aastast ja võimalik, et erinevatest hooaegadest (varakevad kuni kesksuvi).	Tehniliselt sobilik, aga ebaotstarbekas suure rahalise kulu tõttu	Meetod on olemusel sobilik, aga üle-Eestist kaardistamist antud töös ei testitud
Kasutatavus merepõhja elupaigamuutujate aastatevaheliste muutuste hindamiseks	Ajaseeriaid otseselt ei uuritud kuid uuringualade olemasolevate OF-de visuaalne vaatlus näitas, et materjali kvaliteet on liiga muutlik (päikesepeegeldus, lainevaht, halb vee läbipaistvus) ajaseeriade hindamiseks. Täiendavalt on probleemiks see, et kaardistamist ei toimu igal aastal ja erinevatel aastatel võib see olla teostatud erinevatel kuudel (aprillist augustini).	Ajaseeriaid ei uuritud kuid meetod on olemuselt sobilik	Ajaseeriaid ei uuritud kuid meetod on olemuselt sobilik
Rahaline kulukus	Tasuta	Väga kulukas	Tasuta
Operatiivsus, organisatoorne keerukus	Materjali kogumine toimub vastavalt Maa-ameti tööde plaanile	Mehitatud lennuki tellimine ei ole operatiivne ega organisatoorselt lihtne	Ajas ja ruumis fikseeritud ülelennud
Arendusvõimalused ja -vajadused		Vajalik vähendada lennujoonte vahelisi erinevusi, aga selle eelduseks on homogeensete sügavaveeliste alade olemasolu pildil	Vajalik testida kasutatavust kogu Eesti mereala ulatuses
Täiendavad kommentaarid		Meetod on parim kui on vaja saada võimalikult täpseid kaardistamise tulemusi	Satelliidipiltide eelis on töodes, kus hinnanguid on vaja suurte alade kohta ja vajalik on aastatevaheliste muutuste hindamine, sest satelliidipildid katavad väga suure ala ja materjal uueneb regulaarsete intervallidega

Kasutatud allikad

- Breiman L, Cutler A, Liaw A, Wiener M (2018) randomForest: Breiman and Cutler's random forests for classification and regression. R package version 4.6-14. <http://cran.r-project.org/web/packages/randomForest/>
- Fritsch S, Guenther F, Wright MN (2019) neuralnet: Training of Neural Networks. R package version 1.44.2. <http://cran.r-project.org/web/packages/neuralnet/>
- Greenwell B, Boehmke B, Cunningham J and GBM Developers (2019) gbm: Generalized Boosted Regression Models. R package version 2.1.5. <https://cran.r-project.org/web/packages/gbm/>
- Hedley JD, Harborne AR, Mumby PJ (2005) Simple and robust removal of sun glint for mapping shallow-water benthos. *International Journal of Remote Sensing* 26, 2107-2112
- HELCOM (2013) HELCOM HUB – Technical Report on the HELCOM Underwater Biotope and habitat classification. *Balt. Sea Environ. Proc. No. 139.*
- Herkül K, Torn K, Suursaar U, Alari V, Peterson A (2016) Variability of benthic communities in relation to hydrodynamic conditions in the North-Eastern Baltic Sea. *Journal of Coastal Research: Special Issue 75 - Proceedings of the 14th International Coastal Symposium, Sydney, 6-11 March 2016:* pp. 867 – 871.
- Hijmans RJ, Phillips S, Leathwick J, Elith J (2017) dismo: Species Distribution Modeling. R package version 1.1-4. <https://cran.r-project.org/web/packages/dismo/>
- Isæus M (2004) Factors structuring *Fucus* communities at open and complex coastlines in the Baltic Sea. Doctoral dissertation. Department of Botany, Stockholm University.
- Kotta J, Remm K, Vahtmäe E, Kutser T, Orav-Kotta H (2014) In-air spectral signatures of the Baltic Sea macrophytes and their statistical separability. *Journal of Applied Remote Sensing*, 8.
- Kutser T, Hedley J, Giardino C, Roelfsema C, Brando VE (2020) Remote sensing of shallow waters – A 50 year retrospective and future directions. *Remote Sensing of Environment*, 240.
- Liaw A, Wiener M (2002) Classification and Regression by randomForest. *R News* 2(3):18–22.
- Nikolopoulos A, Isæus M (2008) Wave Exposure Calculations for the Estonian Coast. AquaBiota Water Research AB, Stockholm, Sweden.
- Peterson A, Herkül K (2019) Mapping benthic biodiversity using georeferenced environmental data and predictive modeling. *Marine Biodiversity* 49 (131), 131-146.
- R Core Team (2019) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- Remm K, Remm J, Kaasik A (2012) Ruumiliste loodusandmete statistiline analüüs. Õpik-käsiraamat. Tartu Ülikooli Ökoloogia ja Maateaduste Instituut. Tartu.

TÜ Eesti Mereinstituut (2014) Eesti territoriaalmere merepõhja elupaikade ja liikide leviku modelleerimine. Teostatud projekti „Eesti merealade planeerimiseks looduskaitse teabe koondamine, sh. territoriaalmere mereelupaikade modelleerimine“ raames. Leping J/6/2013.

TÜ Eesti Mereinstituut (2016) Loodusdirektiivi mereliste elupaikade inventuur valitud aladel Eesti majandusvööndis. Teostatud projekti „Eesti merealade loodusväärtuste inventeerimine ja seiremetoodika väljatöötamine“ raames Leping 10-4.5.5/14/21.

TÜ Eesti Mereinstituut (2018) Eesti mereala elupaikade kaardiandmete kaasajastamine. Leping 4-1/17/82.

Vahtmäe E, Kutser T, Martin G, Kotta J (2006) Feasibility of hyperspectral remote sensing for mapping benthic macroalgal cover in turbid coastal waters — a Baltic Sea case study. Remote Sensing of Environment, 101, 342-351.

Lisa 1

Maa-ameti ortofotodel ja täiendavatel muutujatel põhinevate mudelite keskmised valideerimisnäitajad, mis on saadud keskmistades üle kõigi modelleeritavate muutujate kohtvaatlusel mõõdetud ja mudeli ennustatud väärtuse paaride. Modelleeritavate muutujate hulgast jäeti välja söödava rannakarbi ja agariku katvus, sest nende esinemissagedus oli liiga madal adekvaatsete mudelite loomiseks. Valideerimisnäitajad: r – lineaarne korrelatsioonikordaja (pidevtunnused), E – klassifitseerimise edukus ehk õigesti klassifitseeritud vaatluste osa kõigist vaatlustest (HUB substraaditüüp). Alade kaupa mudelid ei sisalda sõltumatut muutujat ala nr.

sõltumatud muutujad	mudeli nr.																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	
R, G, B keskmine		+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	
R, G, B standardhäve		+	+	+														
sügavus	+	+	+			+	+			+	+							
lainetus	+	+	+			+				+				+	+	+	+	
geograafilised koordinaadid	+	+	+			+				+		+	+				+	
ala nr.		+	+			+	+	+		+	+		+		+	+	+	
R/G, G/B, R/B		+	+						+	+	+	+	+			+	+	
NIR keskmine		+																
NIR standardhäve		+																
algoritm, naabrite arv, ruudu või proovipõhine valideerimine	valideerimis-																	
koik alad korraga mudelis	näitajad																	
RF, 4, proovipunkt	r	0.47	0.65	0.64	0.43	0.40	0.60	0.49	0.46	0.42	0.63	0.51	0.59	0.59	0.51	0.54	0.56	0.62
	E	0.58	0.68	0.68	0.55	0.50	0.63	0.53	0.56	0.52	0.62	0.55	0.61	0.61	0.59	0.61	0.59	0.60
RF, 4, ruut	r	0.96	0.90	0.89	0.60	0.62	0.94	0.77	0.67	0.70	0.93	0.83	0.90	0.91	0.82	0.81	0.86	0.92
	E	1.00	0.91	0.91	0.65	0.66	0.99	0.84	0.75	0.74	0.98	0.87	0.95	0.96	0.89	0.90	0.90	0.97
RF, 9, proovipunkt	r	0.44	0.66	0.64	0.45	0.41	0.60	0.53	0.49	0.40	0.64	0.55	0.61	0.62	0.51	0.57	0.58	0.63
	E	0.55	0.66	0.66	0.55	0.51	0.61	0.55	0.59	0.53	0.62	0.56	0.62	0.63	0.60	0.61	0.61	0.62
RF, 9, ruut	r	0.99	0.95	0.95	0.65	0.65	0.97	0.83	0.70	0.73	0.97	0.89	0.94	0.95	0.86	0.84	0.90	0.96
	E	1.00	0.96	0.96	0.69	0.71	0.99	0.90	0.78	0.76	0.99	0.91	0.96	0.97	0.92	0.93	0.93	0.98
BRT, 9, proovipunkt	r	0.36	0.56	0.54	0.42	0.40	0.51	0.38	0.43	0.39	0.53	0.45	0.52	0.53	0.39	0.45	0.46	0.52
BRT, 9, ruut	r	0.97	0.90	0.91	0.58	0.56	0.95	0.79	0.63	0.64	0.94	0.80	0.90	0.91	0.81	0.84	0.85	0.92
<i>alade kaupa mudelid</i>																		
RF, 9, proovipunkt	r	0.70	0.80	0.80	0.77	0.76	0.77	0.77		0.75	0.79	0.77	0.79		0.79		0.79	0.79
	E	0.58	0.64	0.63	0.58	0.54	0.59	0.50		0.55	0.59	0.54	0.60		0.62		0.60	0.60
RF, 9, ruut	r	1.00	0.97	0.97	0.86	0.87	0.99	0.93		0.90	0.98	0.94	0.97		0.94		0.95	0.98

E 0.99 0.95 0.95 0.74 0.75 0.98 0.88 0.81 0.98 0.90 0.96 0.92 0.94 0.98

Lisa 2

Hypex kaugseire andmetel ja täiendavatel muutujatel põhinevate mudelite keskmised valideerimisnäitajad. Valideerimisnäitajad on saadud keskmistades üle kõigi modelleeritavate muutujate. Valideerimisnäitajad: r – lineaarne korrelatsioonikordaja (pidevtunnused), E – klassifitseerimise edukus ehk õigesti klassifitseeritud vaatluste osa kõigist vaatlustest (HELCOM HUB substraaditüüp).

sõltumatud muutujad	mudeli nr.														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
kanalid, keskmine			+	+	+			+		+		+		+	
kanalid, standardhälve			+	+											
PCA, keskmine		+				+	+		+		+		+		+
PCA, standardhälve		+				+									
sügavus	+	+	+					+	+			+	+	+	+
lainetus	+	+	+					+	+						
geograafilised koordinaadid	+	+	+					+	+	+	+			+	+
valideerimisnäitajad															
<i>proovipunktipõhine valideerimine</i>															
r	0.68	0.79	0.78	0.78	0.78	0.78	0.78	0.78	0.78	0.78	0.78	0.79	0.79	0.78	0.79
E	0.60	0.60	0.62	0.61	0.64	0.58	0.61	0.61	0.61	0.67	0.63	0.59	0.58	0.64	0.62
<i>ruudupõhine valideerimine</i>															
r	0.99	0.96	0.96	0.88	0.90	0.90	0.92	0.97	0.97	0.95	0.96	0.92	0.94	0.97	0.96
E	0.99	0.96	0.96	0.81	0.85	0.85	0.90	0.98	0.98	0.94	0.96	0.89	0.92	0.96	0.96

Lisa 3

Satelliidiandmetel (Sentinel-2) ja täiendavatel muutujatel põhinevate mudelite keskmised valideerimisnäitajad. Valideerimisnäitajad on saadud keskmistades üle kõigi modelleeritavate muutujate. Valideerimisnäitajad: r – lineaarne korrelatsioonikordaja (pidevtunnused), E – klassifitseerimise edukus ehk õigesti klassifitseeritud vaatluste osa kõigist vaatlustest (HELCOM HUB substraaditüüp).

sõltumatud muutujad	mudeli nr.							
	1	2	3	4	5	6	7	8
kanalid 1-6		+	+	+	+	+	+	+
sügavus	+	+	+	+	+	+		
lainetus	+	+	+	+	+			+
geograafilised koordinaadid ala nr.	+	+	+					
		+		+				
valideerimisnäitajad								
<i>kõik alad korraga mudelis</i>								
r	0.73	0.79	0.79	0.78	0.78	0.75	0.78	0.74
E	0.58	0.60	0.62	0.63	0.62	0.55	0.64	0.59
<i>alade kaupa mudelid</i>								
r	0.73	0.79	0.79	0.79	0.79	0.77	0.79	0.77
E	0.59	0.60	0.60	0.60	0.57	0.54	0.58	0.58