

TALLINNA TEHNIKAÜLIKOOL

Infotehnoloogia teaduskond

Tiia Sildam 193398IAIB

Mario Veelaid 193787IAIB

Andra Velve 193528IAIB

MERETUULE KIIRUSE JA SUUNA LEIDMINE SENTINEL-1 SATELLIIDIPILTIDELT

Bakalaureusetöö

Juhendaja:

Sander Rikka

PhD

Kaasjuhendaja:

Martin Simon

MsC

Autorideklaratsioon

Kinnitame, et oleme koostanud antud lõputöö iseseisvalt ning seda ei ole kellegi teise poolt varem kaitsmisele esitatud. Kõik töö koostamisel kasutatud teiste autorite tööd, olulised seisukohad, kirjandusallikatest ja mujalt pärinevad andmed on töös viidatud.

Autorid: Tiia Sildam
Mario Veelaid
Andra Velve

Kuupäev: 30.05.2022

Annotatsioon

Käesoleva töö eesmärgiks oli uurida kas ja kui täpselt on võimalik tehisavaradari andmetest määrata tuule suunda ja kiirust kasutades masinõppe meetodeid.

Hüpotees oli, et närvivõrgu põhisele mudelile ei ole vaja sobitada välist informatsiooni tuule kiiruse ja suuna kohta ning tuule kiiruse ja suuna saab kätte ainult mudelist. Töö käigus uuriti erinevate tegurite mõju treenimise tulemustele, valideeriti loodud mudeleid rannikujaamade mõõtmistega ja tehti järeldused saadud tulemustest.

Töös kasutatud pildid pärinesid Sentinel-1 satelliidi tehisavaradarilt. Piltidele vastavad tuule kiirused ja suunad pärinesid skatteromeetrilt, valideerimise andmeteks olevad *in situ* mõõtmised mõõtmisjaamadelt. Masinõppe rakendamiseks oli andmehulk, mis koosnes SAR andmetest tehtud alampiltidest suuruses 1024x1024 pikslit ja igale pildile vastavast kiirusest või suunast. Töö käigus loodi treenimise moodul, mis kogus kokku vastavuses olevad andmestikud, treenis mudeli ning valideeris mudelit eraldi valideerimise andmetega.

Töö käigus treeniti kolme erinevat mudelit erinevate sisenditega. Parima tulemuse andis 20-kihiline sügava närvivõrgu mudel, mille sisendiks olid kiirelangemisnurga kanaliga pildid. Mudeli ennustuste Pearsoni korrelatsioon *in situ* mõõtmistega oli 0.71, mudeli R^2 skoor oli 0.33 ja ruutkeskmise viga 2.95 m/s. Hüpoteesi osa, et närvivõrguga on võimalik SAR piltidelt tuule kiirust ennustada, osutus tõseks. Seevastu mudelite võime ennustada tuule suunda SAR piltidelt oli alla keskmise. Parima tulemuse saavutas mudel B - mudeli ennustuste Pearsoni korrelatsioon *in situ* mõõtmistega oli 0.23, mudeli R^2 skoor oli -0.52 ja ruutkeskmise viga 60.28°. Hüpoteesi osa, et suunda on võimalik SAR piltidelt ilma välise teabeta ennustada, osutus osaliselt tõseks. Kui sisendiks olid 0°-180° kraadi vahemikus andmed, siis mudelite ennustamiste täpsus ei olnud küll hea, aga ennustuste ja *in situ* mõõtmiste vahel oli seos.

Lõputöö on kirjutatud eesti keeles ning sisaldab teksti kahekümne neljal leheküljel, seitset peatükki, kuut joonist, viit tabelit.

Abstract

Retrieval of wind speed and direction from Sentinel-1 radar satellite images using AI

The aim of this work was to investigate whether and how accurately it is possible to determine the wind direction and speed from synthetic-aperture radar also known as SAR data using machine learning methods. The hypothesis was that the neural network-based model does not need to match external information with information from SAR images, and that wind speed and direction can only be obtained from the model. In the course of the work, the influence of different factors on the training results was investigated, the created models were validated on the basis of validation data and conclusions were drawn from the obtained results.

The images used in the work originate from the Sentinel-1 satellite radar. The wind speeds and directions corresponding to the images came from a scatterometer, the *in situ* measurements used as validation data originate from measuring stations. To implement machine learning, there was a data set consisting of sub-images with sizes 1024x1024 pixels originating from TIFF images and the speed or direction corresponding to each image. In the course of the work, a training module was created, which collected the matching data, trained the model and validated the model on the basis of separate validation data.

As a result of the work, three different models were trained with different inputs. The best results for speed predictions were obtained with a 20-layer deep neural network model A with input images having incidence angle as the third channel. This model's Pearson correlation of *in situ* measurements and predictions was 0.71, model's R^2 score was 0.33 and root mean square error 2.95 m/s. The part of the hypothesis that the neural network can predict wind speed from SAR images proved to be true. In contrast, the ability of the models to predict wind direction from SAR images was below average. Best result was obtained with model B with input images having incidence angle as third channel. This model's Pearson correlation of *in situ* measurements and predictions was 0.23, model's R^2

score was -0.52 and root mean square error 60.28°. The hypothesis that the direction can be predicted from SAR images without external information turned out to be partially true. When the input data was in range of 0°-180°, the accuracy of the model predictions was not good. But it should be noted that there was still a considerable connection and relation between the predictions and *in situ* measurements.

The thesis is in Estonian and contains twenty four pages of text, seven chapters, six figures, five tables.

Lühendite ja mõistete sõnastik

CMEMS	<i>Copernicus Marine Environment Monitoring Service</i> , Copernicuse merekeskonna seireteenus
EW	<i>Extra-Wide Swath</i> , Sentinel-1 töörežiim
GMF	<i>Geophysical Model Function</i> , geofüüsikalise mudeli funktsioon
IA	<i>incidence angle</i> , kiire langemisnurk
IW	<i>Interferometric Wide Swath</i> , Sentinel-1 töörežiim
NWP	<i>Numerical Weather Prediction</i> , numbriline ilmaennustus
OCN produkt	<i>Ocean product</i> , tehisavaradarilt saadav produkt
OWI komponent	<i>Ocean Wind field component</i> , OCN produkti ookeani tuulevälja komponent, mis sisaldab teavet tuule parameetrite kohta
RMSE	<i>Root mean squared error</i> , ruutkeskmine viga
SAR	<i>synthetic-aperture radar</i> , tehisavaradar
SCAT	<i>scatterometer</i> , skatteromeeter
VH polarisatsioon	polarimeetriline kanal, mille puhul radar saadab välja vertikaalselt polariseeritud kiirgust ja võtab vastu horisontaalselt polariseeritud kiirgust
VV polarisatsioon	polarimeetriline kanal, mille puhul radar saadab välja ja võtab vastu vertikaalselt polariseeritud kiirgust

Sisukord

Joonised	8
Tabelid	9
1 Sissejuhatus	10
2 Taust ja hüpotees	11
2.1 Varasemad katsed ja arengud	11
2.2 Läänemere tuuleolud	12
2.3 Eesmärk	12
3 Radarseire instrumendid	14
3.1 Pinnalt radarisignaali tagasihajumise mõjurid	14
3.2 SAR - tehisavaradar	14
3.3 SCAT - skatteromeeter	16
4 Andmed	17
4.1 Andmete päritolu	17
4.2 Andmete eeltöötlus	19
4.3 Andmete ülevaade	19
5 Meetod ja mudel	23
5.1 Mudelite kirjeldus	23
5.2 Mudeli treenimise parameetrid	25
5.2.1 Hüperparameetrid	25
5.2.2 Muud mudeli treenimise parameetrid	26
5.3 Treenimise kirjeldus	26
5.4 Valideerimise meetodite kirjeldus	27
5.5 Kuvamine	27
5.6 Ruumiline testimine	28
6 Tulemused	29
6.1 Treenimise tulemused	29
6.2 Valideerimine <i>in situ</i> mõõtmistega	30
6.3 Hinnangute kirjeldus	32
6.4 Järeldused	34

7 Kokkuvõte	35
Viited	37
Lisa 1 – Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks	40

Joonised

1	<i>Töös kasutatud võrreldavate punktide arv ja asukoht ning in situ punkt- mõõtmiste asukohad Läänemeres.</i>	18
2	<i>Tuule kiiruste ja suundade jaotuvus mõõtmistelt vahemikus 01/01/2016- 15/11/2021.</i>	20
3	<i>Mudel A kihtide arhitektuur.</i>	24
4	<i>Mudel C kihtide arhitektuur.</i>	25
5	<i>Mudel A valideerimise tulemused. Ennustuste tihedus on 0.25 m/s sammu- ga, prognoosi keskmine on arvatud 1 m/s sammuga.</i>	31
6	<i>Mudel B valideerimise tulemused. Ennustuste tihedus on 4° sammuga, prognoosi keskmine on arvatud 10° sammuga.</i>	32
7	<i>Mudel A ennustatud kiirused ning võrdluseks OWI komponendi tuule kiirused.</i>	33

Tabelid

1	<i>in situ</i> mõõtmiskeskused ja nende asukohad.	18
2	Töö käigus loodud mudelid.	23
3	Kihtide kirjeldused.	24
4	Mudelite täpsused kiiruse ennustamise korral.	29
5	Mudelite täpsused sisendiga vahemikus 0° - 180° suuna ennustamise korral.	30

1. Sissejuhatus

Ilma prognoosimine on pakkunud teadlastele väljakutseid tuhandeid aastaid. Ilmaennustus on mõjutatud paljudest faktoritest, mõned nendest on tuule kiirus ja suund. Olemasolevad tuule parameetrite hindamise meetodid tehisavaradari andmetest on keerulised ja hõlmavad mitut sammu, kus on vaja omavahel sobitada erineva päritoluga andmeid.

Üks võimalik viis protsessi lihtsustada, et saada lisateavet tuuleolude kohta, on rakendada masinõpet tehisavaradari piltidele. Kui õnnestub treenida efektiivne mudel, mis hindab täpselt tuule suunda ja kiirust, on rohkem informatsiooni olude kohta, mille sisendamine prognoosimudelitesse võib kaasa aidata ilma prognoosimisele. Tänapäeval on saadaval kõrglahutusega satelliitpildid, mida on võimalik kasutada ka tuule omaduste hindamiseks. Tehisavaradari andmete abil tuule parameetrite määramist on varasemalt katsetatud ning meetodil nähakse potentsiaali ilma ennustamise valdkonnas ka tulevikus.

Käesolev töö käsitleb masinõppe abil tuule parameetrite ennustamist Läänemere piirkonnas. Eesmärkideks on ennustada sügava närvivõrguga Sentinel-1 tehisavaradari andmetest tuule suunda ja kiirust. Masinõppe abil saadud tulemuste valideerimine ja võrdlemine toimub *in situ* mõõtmistega.

Antud töös käsitletakse järgnevaid etappe:

- Andmete eeltöötlus
- Kiiruse ja suuna mudeli treenimine
- Erinevate mudelite testimine
- Mudeli tulemuste valideerimine

2. Taust ja hüpotees

2.1 Varasemad katsed ja arengud

Tehisavaradari ehk SAR-i andmete abil ookeani tuule parameetrite ja lainetuse uurimine algas juba 1978. aastal [1]. Selleks kasutati SEASAT-A pardal asuvat tehisavaradarit. Tuule parameetrite leidmise jaoks on sellest ajast saati loodud erinevaid algoritme [1][2]. Mõningad näited meetoditest, millega umbkaudselt ilma välise informatsioonita tuletada tuule suunda, on Fourier' kiirteisendus ja pidev või diskreetne laineteisendus. Ükski neist ei ole puudusteta, näiteks võib mõni meetod töötada vaid kindlates oludes, kus vajalikud tingimused peavad olema täidetud. Et Fourier' kiirteisendust saaks kasutada, on vaja merepinnal "triipe" (tuule suunaga risti olevad nähtavad laineharjad), mida ei pruugi seal kogu aeg olla, sest nende tekkimiseks on tarvis kestvalt suurema kiirusega tuult [3]. Juhul, kui laineharjad on nähtavad ja tuule liikumise telg on leitav, puudub informatsioon selle kohta, kas tuul "tuleb" või "läheb" (180 kraadine määramatus).

Tehisavaradarilt saadud piltidelt tuulevälja määramise võimalus on tähtis rannikualadel ning väiksemates meredes. Tuule suuna ja kiiruse kättesaamine on oluline, sest see aitab luua seoseid ja mõista ranniku ning õhuvoolu koostoimimist ja vastastikmõju. Levinumad meetodid tuule suuna ja kiiruse hindamiseks tehisavaradaritelt on näiteks geofüüsikalise mudeli funktsioonid (GMF) CMOD4 ja CMOD5.N, mida rakendatakse samatasandiliste polarisatsioonide (HH või VV) jaoks [3]. GMF-i abil saab omavahel siduda radari ristlõike väärtustega järgmised kaks väärtust: anduri geomeetria ja tuuleolud [4].

SAR spetsiifiliste GMF-ide täpsus oleneb sellest, mis on mudelile antud tuule sisend-suund, mida ainuüksi tehisavaradarilt pole siiaaani suudetud kõikide olude jaoks täpselt hinnata. Sellepärast kasutab tuulevälja parameetrite kättesaamiseks Euroopa Kosmoseagentuur numbrilise ilmaennustuse (NWP) mudeli tuule suunda GMF-i sisendiks. Numbriline ilmaennustus jääb sageli hätta ruumilise variatsiooniga, millel on väike ulatus. Seega ei pruugi antud meetod hästi toimida rannikualadel ning võib teha valesid järeldusi [3].

Konvolutsiooniline närvivõrk on üks suurima edu saavutanud meetodeid piltidelt ruumilise teabe väljavõtmiseks. Seda kasutatakse mitmete erinevate probleemide lahendamiseks, mõned näited on piltide klassifitseerimine ja objektide tuvastamine. On alust arvata, et

tehisavaradari piltide muustritest, mis sisaldavad palju teavet tuulevälja struktuurist, on konvolutsioonilist närvivõrku kasutades võimalik ilma välise informatsioonita kätte saada tuule suund ja kiirus. Tehisnärvivõrkude abil on kalkuleeritud välise informatsiooni (atmosfäärilistest mudelitest saadud tuule suuna) abil tuule kiirust [3]. Kui GMF võtab lihtsustatult sisse ainult keskmise alampildi tagasihajumise (CMOD algoritm kasutab ainult HH- või VV-andmed), siis närvivõrgu jaoks on kõik pikslid võimalikud parameetrid. Täiendavalt on närvivõrgule võimalik lisainformatsiooni anda ka teiste polarimeetriliste kanalite näol (VH) või lisada muid tuule hindamise asukoha spetsiifilisi metaandmeid (merepunkti kaugus rannikust, aeg, koordinaat jms) [3].

Viimastel aastatel on tehtud mitmeid katseid tuule parameetrite kätte saamiseks masinõppe abil. Li 2020 kirjeldas protsessi, kus saadi Põhja-Jäämere piirkonnas Sentinel-1 HH-polarisatsiooniga SAR andmetelt masinõppe abil tuule kiirus. Antud uuringus kasutati polarisatsioone ja geofüüsikalise mudeli funktsioone [5]. Zanchetta ja Zecchetto 2021 tutvustasid võimalikku meetodit tuule suuna ennustamiseks SAR piltidelt ilma välise informatsioonita. Selleks kasutati sügava närvivõrgu tehnoloogiat ning vaadeldav ala oli Vahemeri ja Pärsia laht. Protsessi lihtsustamiseks rakendati tuule suunale valemit, mis ei võtnud arvesse tuule suuna 180 kraadist määramatust [3].

2.2 Läänemere tuuleolud

Läänemeri asub geograafiliselt 9 kuni 30° E laiuskraadi ja 53 kuni 66° N pikkuskraadi vahel ja on 392,978 km² pindalaga Atlandi ookeani mandrisisene meri [6]. Keskmise vee sügavus meres on 54m [7]. Läänemeri oma eripäradeka on keeruliste oludega piirkond. Madalad alad ja paljud väikesaared, kitsad lahed, lähedus rannikule ja talvine jää mõjutavad tuule kiiruse kasvu potentsiaali (*fetch*). Seega on tuulest põhjustatud lained, mis radarile tagasihajumist põhjustab, keerulised [8]. Nähtavat lainemustrit esineb harva ja see teeb lokaalse tuule suuna ja kiiruse hindamise väljakutsuvaks [1]. Masinõppe meetoditel on siinkohal eelis suure andmehulga pealt rahuldavate täpsusteni jõudmisel.

2.3 Eesmärk

Antud töö eesmärgiks on uurida, kas ja kui täpselt on võimalik tehisavaradari andmetest määrata tuule suunda ja kiirust kasutades masinõppe meetodeid.

Masinõppe meetodite abil tehisavaradari andmetest Läänemere tuule suuna ja -kiiruse määramisel on mitu eelist ja head omadust. Esiteks läheks praegune protsess tuule parameetrite kätte saamiseks mõnevõrra lihtsamaks. Olemasoleva protsessi (CMOD) jaoks on

tarvis välise mudeli andmete sissevõtmist, SAR-i andmete töötlemise üksust ning tuule kiiruse ja suuna määramise üksust [9]. Hüpotees on, et närvivõrgu põhisele mudelile ei ole vaja sobitada välist informatsiooni tuule kiiruse ja suuna kohta ning tuule kiiruse ja suuna saab kätte ainult mudelist. Seega tänu treenitud mudelile ei oleks tarvis toetuda välisele informatsioonile. Samuti on tehisavaradaril võime kirjeldada merepinda väikeses skaalas, kuna sellel on väga hea ruumiline katvus. See tuleb kasuks nendel aladel, kus *in situ* mõõtmiste hulk on väiksem ja neid mõõtmisi on hõredamalt.

Rannikualad on tehisavaradarsatelliidiga paremini kättesaadavad, sest SAR mõõtmised pole rannikust mõjutatud - võrreldes skatteromeetri või radiomeetriga saab informatsiooni rannikule lähemalt [10]. Mudel, mis suudab satelliidipiltidelt määrata tuule suunda ja kiirust, võib mõjutada positiivselt ka ilma prognoosimist. Mida rohkem on sisendit ja informatsiooni olude kohta, seda paremini ja täpsemini saab ka ilmaennustusi teha. Näiteks on tuule parameetrid peamised sisendid lainetuse prognoosimisel. Kui mudel hindab tuule andmeid hästi, on võimalik ka teisi parameetreid paremini määrata [4].

Käesoleva töö eesmärkideks on:

- hinnata tuule kiiruse määramise täpsust tehisnärvivõrgu abil ja rakendada erinevaid meetodeid parima tulemuse leidmiseks
 - kui piltide kolmandaks kanaliks on kiire langemisnurk ehk IA
 - kui piltide kolmandaks kanaliks on polarisatsioonide ristkorrelatsiooni valemi arvutuse väärtus
 - kui pilte keeratakse, et luua suurem andmehulk
- hinnata tuule suuna määramise täpsust tehisnärvivõrgu abil
 - samadel tingimustel kui kiirust
 - võttes arvesse 180 kraadist määramatust
- kirjeldada mudelite efektiivsust erinevate mõõdikutega

3. Radarseire instrumendid

3.1 Pinnalt radarisignaali tagasihajumise mõjurid

Pinnalt signaali tagasihajumist mõjutavad mitmed radarispetsiifilised tegurid, näiteks polarisatsioon ja langemisnurk [11].

Elektromagnetlainete võnketasandit kirjeldav omadus on polarisatsioon. Laineid, mis on fikseeritud võnketasandiga, kutsutakse polariseeritud laineteks. Enim kasutatud polarimeetrilised kanalid radarseires on HH, HV, VH ning VV. Välja saadetud signaali polarisatsiooni saab välja lugeda polarimeetrilise kanali nime esimesest tähest ning vastu võetud signaali polarisatsiooni saab teisest tähest:

- HH - radar saadab välja ja võtab vastu horisontaalselt polariseeritud kiirgust
- HV - radar saadab välja horisontaalselt polariseeritud kiirgust ja võtab vastu vertikaalselt polariseeritud kiirgust
- VH - radar saadab välja vertikaalselt polariseeritud kiirgust ja võtab vastu horisontaalselt polariseeritud kiirgust
- VV - radar saadab välja ja võtab vastu vertikaalselt polariseeritud kiirgust [11]

Erinevate kanalite piltidel on ka erinevad tagasihajumise mehhanismid: peamiselt pinnapeegeldus HH- ja VV- kanali puhul ning ruumiline tagasihajumine VH- ja HV-kanali puhul [11].

Kiire langemisnurk, mis samuti mõjutab radarisignaali tagasihajumist, on maapinna normaali ja radarikiire vaheline nurk. Tagasihajumismehhanismide esinemisele avaldab mõju langemisnurk. Üldiselt mida väiksem on langemisnurk, seda tugevam on pinnapeegeldus [11].

3.2 SAR - tehisavaradar

Tehisavaradar ehk SAR kasutab andmete kogumiseks aktiivset meetodit - andur edastab mikrolainesignaale ja salvestab selle energia hulga, mis allolevalt pinnalt tagasi saabub või hajub [12]. Tehisavaradar simuleerib palju suuremat antenni kombineerides lühema

antenni võtete jada ja pakkudes seega suurema eraldusvõimega andmeid [13]. Vastupidiselt optilisele tehnoloogiale suudab tehisavaradar teostada mõõtmisi olenemata valguse hulgast, pilvede olemasolust või vihmast [12].

Tehisavaradaril on mitmeid erinevaid lainepikkuseid, mida nimetatakse ribadeks ja tähistatakse tähtedega (C-riba, X-riba jms). Elektromagnetlaineline lainepikkus on SAR-il oluline. Riba lainepikkus määrab keskkonda tungimise võime ehk kui lähedale maismaani signaal levib. X-riba (mille signaali lainepikkus on umbes 3cm) on näiteks vähevõimeline tungima laialehistesse metsa, seevastu L-riba (mille signaali lainepikkus on umbes 23cm) võimaldab rohkemat vastastikmõju radari signaali ning suurte okste ja puutüvede vahel [13].

Hetkel on levinuim C-riba, mille kasutusala on ülemaailmne kaardistamine, muutuste tuvastamine, madala kuni keskmise läbitungimisega alade seire, jää ja ookeani merenavigatsioon. Levinud on ka X-riba, mida kasutatakse muuhulgas näiteks linnade seireks [13].

SAR kogub aluspinnalt tagasi peegeldunud signaali tugevust kontrollides analüüsitud polarisatsiooni nii edastamise- kui ka vastuvõtuteel. Tasandi orientatsioon, milles edastatav elektromagnetlaineline võngub, viitab polarisatsioonile. SAR-andurid edastavad orientatsiooni tavaliselt lineaarselt polariseeritult, kuigi orientatsioon võib toimuda ükskõik millise nurga all. Vertikaalset polarisatsiooni tähistab V ja horisontaalset H. Erinevatest polarisatsioonidest tuleneva signaali tugevuse analüüs annab informatsiooni kujutise pinna struktuuri kohta, mis põhineb järgmistel hajumise tüüpidel: maht, kare pind ja topeltpõrge. Kõige tundlikum pinnapealse krobeline hajumise suhtes on VV. HV või VH ehk ristpolariseeritud andmete suhtes on kõige tundlikum mahuline hajumine (põhjustavad näiteks suured lained). Topeltpõrke hajumisele on tundlik HH polariseeritud signaal, seda hajumist põhjustavad näiteks hooned ja taimestik [13].

Tehisavaradarilt saab kätte mitu produkti, mis on erinevate tasemetega: tase 0, tase 1 ja tase 2 [14]. Tasemest 1 on võimalik kätte saada taseme 2 produktid, mis on geograafilise asukohaga ja geofüüsikalised. Lainete, tuule ja hoovuste jaoks sisaldab tase 2 ookeani (inglise keeles *Ocean* ehk OCN) produkt järgmisi komponente: ookeani tuuleväli (inglise keeles *Ocean Wind field* ehk OWI), ookeani lainespekter (inglise keeles *Ocean Swell spectra* ehk OSW) ja pinna radiaalkiirus (inglise keeles *Surface Radial Velocity* ehk RVL). Kuna Sentinel-1 tehisavaradar opereerib erinevates režiimides, siis neist mõne puhul ei saa kätte kõiki eelnimetatud OCN komponente. Näiteks IW (inglise keeles *Interferometric Wide Swath*) ja EW (inglise keeles *Extra-Wide Swath*) režiimides tehtud andmetele ei saa rakendada OSW komponenti, kuna puudub piisav eraldusvõime [15]. Ookeani pinna kahedimensionaalse lainete spektri komponent (OSW) kirjeldab laine suuna ja kõrguse

väärtusi vastavate sagedusvahemike jaoks. Rannikualade jaoks eespool mainitud lainetuse spektrit sisaldavat OSW produkti olemas ei ole [16]. Küll aga on OWI komponendis tuule parameetrite informatsioon olemas ka Läänemere piirkonna jaoks. CMOD funktsiooni kasutades saadakse NWP mudeli abil need parameetrid kätte.

3.3 SCAT - skatteromeeter

Skatteromeeter ehk SCAT on võimeline lisaks tuule kiirusele määrama ka tuule suunda vee kohal. Suutelisus määrata tuule suunda on skatteromeetri eelis teiste kaugseire instrumentide ees [17]. On võimalik eristada kahte tüüpi skatteromeetreid - C-band ja Ku-band. Nendel kahel tüübil saab vahet teha kiirguse lainepikkuse põhjal, samuti on neil näiteks ka erinevad polarisatsioonid, C-bandil VV polarisatsioon ja Ku-bandil duaalne polarisatsioon [18].

Skatteromeetri tööpõhimõte põhineb aktiivsel meetodil - skatteromeeter saab informatsiooni kätte nii, et peale elektromagnetiliste impulsside väljasaatmist tehakse kindlaks, kui palju signaalist tagasi jõuab [17]. Geofüüsikalise mudeli funktsioon (GMF) on skatteromeetri mõõtmiste juures oluline osa. Selle abil on võimalik tagasi peegeldatud signaali polarisatsioonist, nurkadest (näiteks kiire langemisnurgast), pinnalähedasest tuule vektorist ja radari sagedusest koosnevast funktsioonist kätte saada ookeani pinna radari ristlõige [18]. Tagasi peegeldatud signaali tugevus on mõjutatud tuule suuna ja tuule kiiruse poolt ning seetõttu on võimalik neid kätte saada tagasi hajutatud signaalist. Mikrolainete hajumist, mis skatteromeetrile signaalina tagasi peegelduvad, kirjeldab Braggi seadus [17]. Tagasi hajutatud signaal hajub otse peegeldumise asemel mitmes erinevas suunas, kuna merepind ei ole tasane pind. Braggi seadusest tulenevalt muutub tagastatava energia hulk proportsionaalselt tuule kiirusest ja suunast merepinnal. Skatteromeetrid ja radarid üldisemalt on tundlikud vaid suuremate lainete peal asetsevatele väiksemate kapillaarlainete lainetustele, mis on tuulest mõjutatud. Tõusu-mõõna lainete suhtes need radarid tundlikud pole. [18].

4. Andmed

4.1 Andmete päritolu

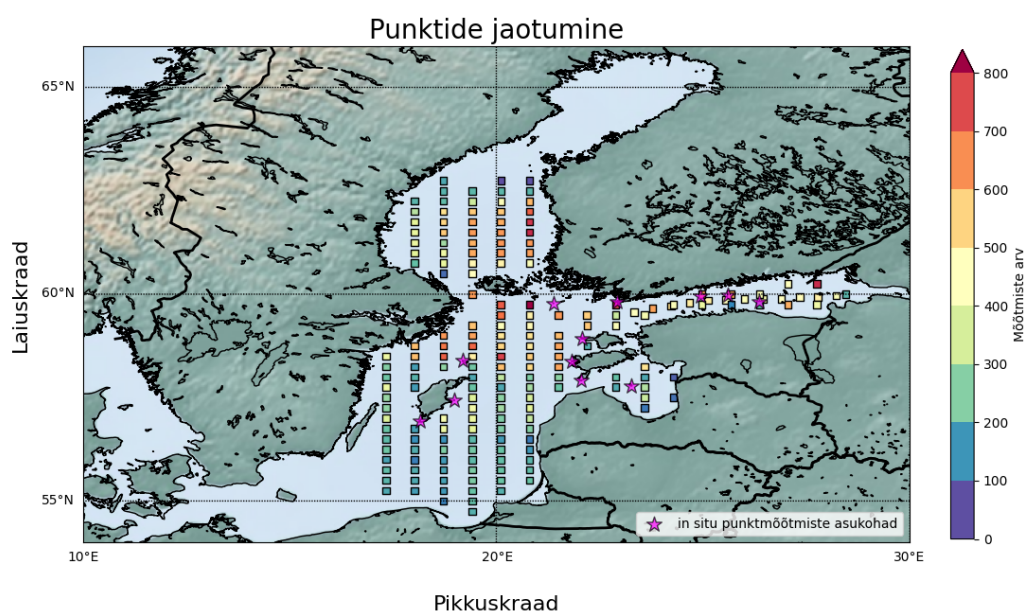
Mudeli treenimiseks kasutatavad Sentinel-1 kahedimensioonilised satelliidipildid tehti IW töörežiimis ja C-riba lainepikkusel. Pildid pärinevad Eesti riiklikult satelliidiandmete keskusest ESTHubilt. Seal hangitakse Eestile huvipakkuvate piirkondade kohta Sentinel satelliitide andmeid. ESTHub laeb alla ja arhiveerib Eesti riigipiirist ligikaudu 200 km ulatuses satelliidi Sentinel-1 andmed. Üldiselt on sarnaste andmete jaoks Open Access Hub portaal levinum ja mõeldud tavakasutajatele. Tänu Ettevõtluse Arendamise Sihtasutuse sõlmitud kokkuleppele Euroopa Kosmoseagentuuriga on kasutajatel võimalik soovitud andmeid tunduvalt kiiremini alla laadida põhiliselt riikide jaoks mõeldud portaali ColHub kaudu [19].

Treenimiseks kasutatavad lähima aja ja asukohaga skatteromeetri mõõtmised pärinevad Copernicuse merekeskkonna seireteenuselt (CMEMS). CMEMS teeb kasutajale kättesaadavaks okeanograafilised teenused ja produktid, mille abil on võimalik näiteks ennustada ilma ja kliimat, saada infot mere- ja rannikukeskkonna kohta. Rohkem kui 150 produkti, mida CMEMS kasutajatele pakub, tulenevad kaugseirepiltidelt satelliidilt, ookeani prognoosimudelilt ning merel tehtud mõõtmistelt [20]. Tuule suund ja -kiirus on üldjuhul kirjeldatud kahe komponendiga - u ja v , kus u on ida-lääne ja v on põhja-lõuna suunaline komponent.

Antud töös kasutatavad Läänemere piirkonna *in situ* punktmõõtmised pärinevad Eesti (KAUR), Soome (FMI) ja Rootsi (SMHI) vastavatest meteoroloogiateenistustest, kokku 12 mõõtmispunktist. Mõõtmispunktid ja nende koordinaadid on näha Tabelis 1. Lisaks on *in situ* punktmõõtmiste asukohad Läänemeres ning töös kasutatud võrreldavate punktide arv ja asukoht nähtavad Joonisel 1.

Tabel 1. *in situ* mõõtmiskeskused ja nende asukohad.

Mõõtmiskeskuse asukoht	Laiuskraad	Pikkuskraad	Päritolu
Parainen Utö	59.77909	21.37479	FMI
Porvoo Kalbådagrund	59.98518	25.59855	FMI
Hanko Tulliniemi	59.80861	22.91248	FMI
Helsinki Helsingin majakka	59.94898	24.92631	FMI
Gotska Sandön A	58.3943	19.1975	SMHI
Hoburg A	56.9213	18.1541	SMHI
Östergarnsholm A	57.4413	18.9871	SMHI
Ristna	58.92083	22.06639	KAUR
Ruhnu	57.78333	23.25889	KAUR
Sõrve	57.91361	22.05806	KAUR
Vaindloo	59.81694	26.36111	KAUR
Vilsandi	58.38278	21.81417	KAUR



Joonis 1. Töös kasutatud võrreldavate punktide arv ja asukoht ning *in situ* punktmõõtmiste asukohad Läänemeres.

4.2 Andmete eeltöötlus

SAR andmete eeltöötlus tehti ESTHub töötlusplatvormi peal. See hõlmas endas radiomeetrilist kalibreerimist, müra eemaldamist ja projektsiooni keeramist. Suurest SAR GeoTIFF failist lõigati väiksemad pildid suurusega 1024x1024 pikslit. Lisaks arvutati alampildi statistika, mis salvestati CSV formaati.

Piltidel oli kolm kanalit, kolmanda kanali jaoks oli mitu varianti. Esimene variant oli kolmanda kanalina kasutada kiire langemisnurka. Teine variant kolmanda kanali jaoks oli kohandatud kanalite ristkorrelatsioon Valemiga 4.1. Ristkorrelatsiooni arvutamisel kasutati valemis Sentinel-1 IW režiimis tehtud piltide kahte polarisatsiooni (VV ja VH) [21].

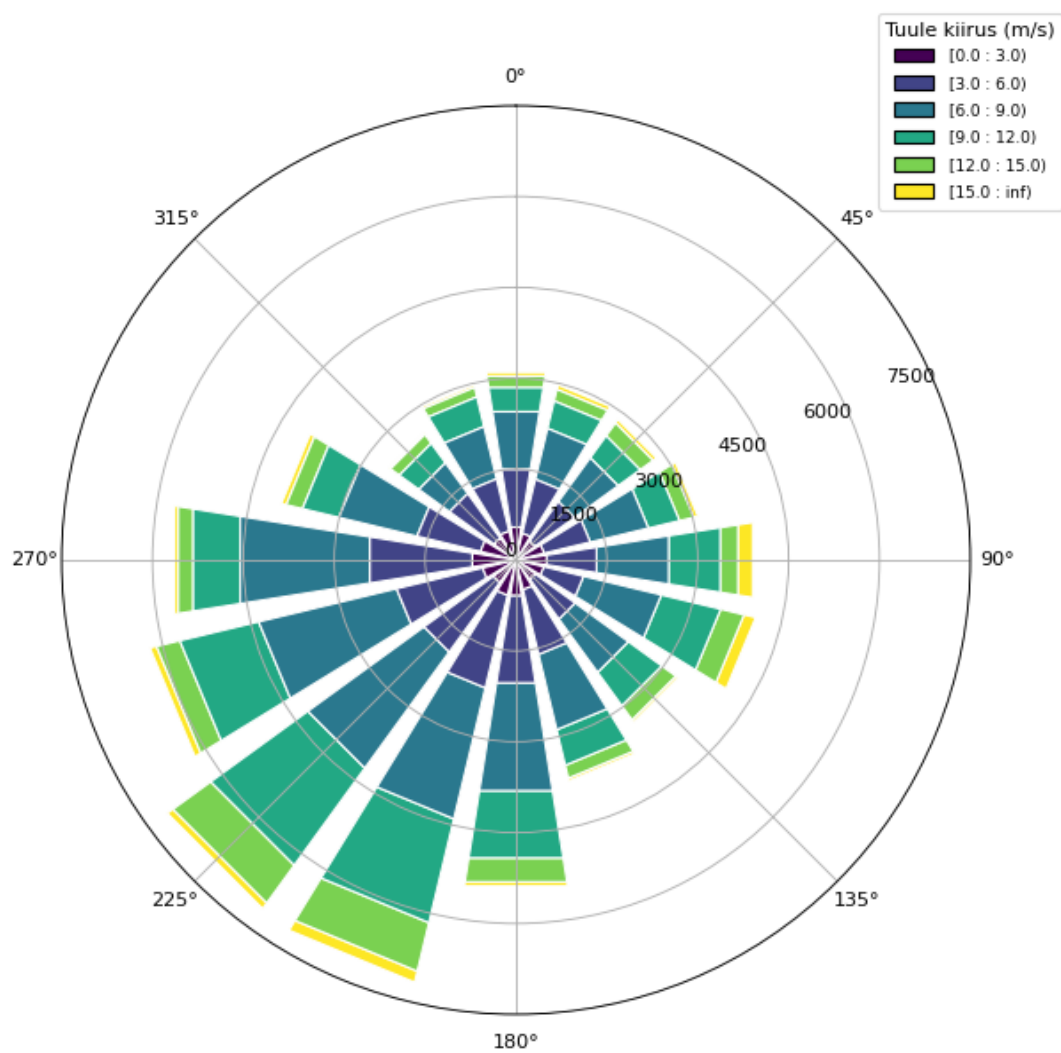
Kanalite ristkorrelatsiooni arvutamise valem

$$C_c(k, l) = \frac{1}{N\sigma_{VV}\sigma_{VH}} \times \sum_{i,j \in W} (I_{VV}(k+i, l+j) - \mu_{VV}) (I_{VH}(k+i, l+j) - \mu_{VH}) \quad (4.1)$$

4.3 Andmete ülevaade

Tuule suunda ja kiirust ennustava mudeli sisendiks olid tehisavaradari pildid ja skatteromeetrilt pärinevad mõõtmised CSV failis. Mõõtmiste hulgas on kiirus ja statistika tehisavaradari pildi kohta. CSV failis on mõõtmisi mitme aasta jagu - alates 2016. aasta algusest kuni 2021. aasta lõpuni. Keskmise ajavahe satelliitpiltide ja skatteromeetri mõõtmiste vahel on ligikaudu 2.9 tundi. Keskmise distants satelliitpiltide ja skatteromeetri mõõtmiste vahel on 4.2 kilomeetrit.

Tuule kiiruste ja suundade jaotuvus mõõtmistelt vahemikus 01/01/2016-15/11/2021 on näha Joonisel 2.



Joonis 2. Tuule kiiruste ja suundade jaotuvus mõõtmistelt vahemikus 01/01/2016-15/11/2021.

Mõõtmiste hulgas on eraldi põhja- ja ida-suunalised kiirused ehk tuule u ja v komponendid, viimast kahte kasutatakse ka tuule suuna arvutamiseks. Tuule suund arvutatakse kasutades numpy *atan2* funktsiooni, mille tulemus konverteeritakse radiaanidest kraadidesse. Kuna antud arvutus annab tulemuseks matemaatilise suuna ja vajaminev suund on geograafiline, siis lahutatakse tulemus 270st. Geograafiline suund kirjeldab seda, et tuul puhub sellest suunast ning põhi on tähistatud 0 kraadiga. Kuna antud valemi tulemused võisid olla olenevalt tingimustest ka negatiivsed, siis võeti moodul 360-ga. Geograafilise suuna arvutamine on kirjeldatud Valemiga 4.2 [22].

Geograafilise suuna arvutamise valem

$$\begin{aligned} \text{Dir}_{\text{geo}} &= \text{atan2}(-U_{\text{geo}}, -V_{\text{geo}}) * \text{DperR} \\ &= 270 - (\text{atan2}(V_{\text{geo}}, U_{\text{geo}}) * \text{DperR}) \end{aligned} \quad (4.2)$$

Andmeid on võimalik soovi korral filtreerida, näiteks on olemas võimalus filtreerida välja potentsiaalseid jääalaseid või liiga suuri ajalisi erinevusi mõõtmistes. Alad, kus suure tõenäosusega võib jaanuaris, veebruaris ja märtsis esineda arvestatav kogus jääd, mis segab tuule suuna ja kiiruse ennustamist, on Soome laht ja Botnia meri. Mõõtmised, mis on tehtud eelnimetatud piirkondades eelmainitud kuudel, kus tuule kiirus on võrdne või suurem kui 15.0 m/s (ehk juba arvestatava suurusega kiirus, mida võiks ka mudeli treenimisel arvesse võtta) tuleb käsitsi üle kontrollida. Seda selle jaoks, et tuvastada jää olemasolu ning selle osakaal vastaval SAR pildil ning otsustada, kas anda pilt mudelile ette või mitte. Potentsiaalsete jääalade kohta tehti analüüs. Sellest selgus, et sellist filtrit rakendades leitakse algsete andmete hulgast üle 250 rea, mis tuleks jää olemasolu ja selle hulga tuvastamiseks eraldi läbi vaadata.

Teine filtreerimine, mida on võimalik andmetele rakendada, on ajalise parameetri järgi andmetest mõõtmiste filtreerimine, kus esimese tingimusena SAR-i ja skatteromeetri mõõtmiste vaheline aeg on suurem kui neli tundi ning teiseks tuule kiirus on väiksem kui 10.0 m/s. Sellised mõõtmised jäetakse mudelile ette antavatest andmetest välja.

On loodud võimalus piltide mõõtkava muuta. Seda on võimalik rakendada piltidele kahes erinevas variandis - kas suurendada pilti või vähendada (*upscale* ja *downscale*). Potentsiaalsete jääalade filtreerimise ning erinevatel viisidel piltide mõõtude muutmise mõju treenimisele vajab edasist uurimist ja testimist.

Andmetehulka suurendati ka pilte töödeldes - kasutusel on pildi pööramine 90 kraadi kaupa ja pildi peegelpildis kasutamine. Nii saadi ühe pildi asemel, millele vastas üks kiirus, neli

või rohkem sama väljundiga pilti. Suuna puhul tuli suunda vastavalt pildi keeramisele ka keerata. Samuti on võimalik filtreerida pilte olenevalt kuupäevast ja keskmisest kiirusest.

5. Meetod ja mudel

Konvolutsioonilise närvivõrgu abil on võimalik kätte saada piltidelt ruumiline teave. Kuna Sentinel-1 tehisavaradari piltidel olevates mustrites sisaldub informatsiooni tuulevälja struktuuri kohta, siis võiks sügav närvivõrk sobida hästi tuule suuna ja kiiruse hindamiseks.

Projekti tegemisel kasutati erinevaid teke ja tehnoloogiaid. Projektis kasutatav programmeerimiskeel oli Python. Erineva struktuuriga närvivõrgud implementeeriti Kerase ja Tensorflow abil. Eelmainitud tehnoloogiad said valitud sellepärast, et nende kasutamine oli lihtsasti arusaadav varem valdkonnaga mitte kokkupuutunud inimeste jaoks. Lisaks leidis antud tehnoloogiatega õpetusi piltidelt regressiooni mudeli treenimiseks. Erinevate andmemassiividel operatsioonide läbiviimiseks ning nende haldamiseks ja töötlemiseks olid kasutusel laialdaselt levinud Pythoni teegid Numpy ja Pandas. Treenimise keskkonnana kasutati AI Lab ehk Tehismõistuse laborit.

5.1 Mudelite kirjeldus

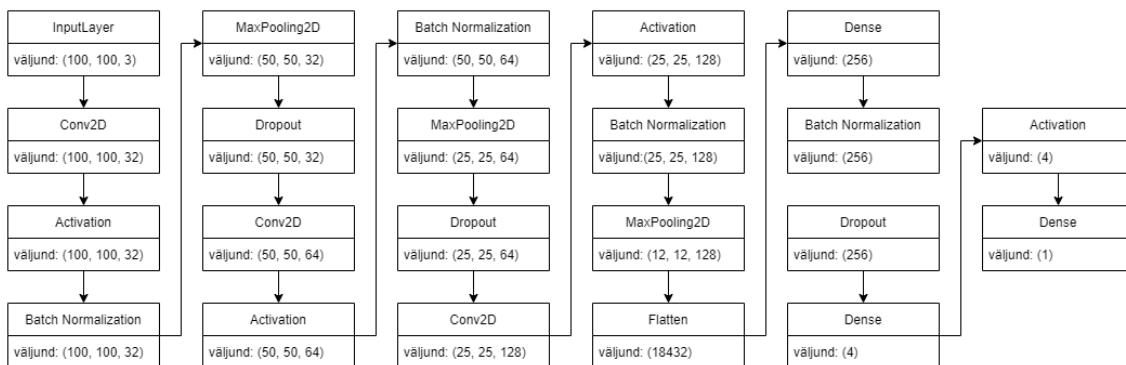
Töö käigus loodi 3 erinevat mudelit tuuleparameetrite kättesaamiseks. Kaks mudelit võtsid sisendiks SAR pildid, kolmas mudel loodi selleks, et treenida piltide pealt arvutatud statistika väärtustel.

Tabel 2. Töö käigus loodud mudelid.

	Mudel	Kihtide arv	Kirjeldus
Mudel A	Sügav närvivõrk	20	Sisendiks pildid. Treenitavaid parameetreid 8 945 833 ja mitte treenitavaid 1408
Mudel B	ResNet50	50	Sisendiks pildid. Treenitavaid parameetreid 23 536 641 ja mitte treenitavaid 53 120
Mudel C	Sügav närvivõrk	11	Sisendiks CSV statistika. Treenitavaid parameetreid 18 603 ja mitte treenitavaid 512

Esimene mudel, mida katsetati oli mudel A. Sellel mudelil oli 20 kihti. Mudeli arhitektuur

on pärit veebilehe PyImageSearch õpetusest ning vastavalt vajadusele kohandatud [23]. Mudel A arhitektuur on näha Joonisel 3. Selle mudeli sisendiks olid pildid.



Joonis 3. Mudel A kihtide arhitektuur.

Erinevate kihtide kirjeldused on nähtavad Tabelis 3.

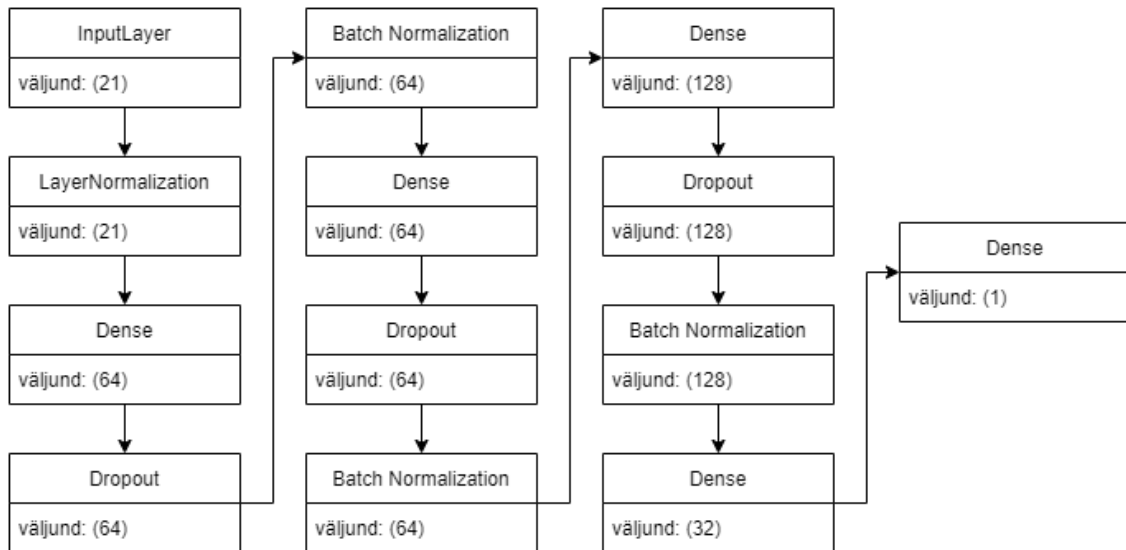
Tabel 3. Kihtide kirjeldused.

Kihi nimi	Selgitus
InputLayer	Närvivõrgu sisend punkt.
Conv2	Mitu filtrit, mille parameetrid on treenitavad.
Activation	Kasutab väljundi peal aktiveerimis-funktsiooni.
Batch Normalization	Muudab andmeid nii, et keskvärtus oleks 0 ja standardhälve 1.
MaxPooling2D	Skaleerib andmeid väiksemaks valides alamhulkadest suurimad väärtuse.
Dropout	Eemaldab suvalised neuronid, et ära hoida üle sobitamist.
Dense	Neuron on sügavalt seotud neuronitega eelmisest kihist.
Flatten	Teeb andmed ühemõõtmeliseks.

Teise mudelina katsetati mudel B-d, mis oli 50 kihti sügav konvolutsiooniline närvivõrk. Sellele lisati viimasena üks tihe närvikiht, et saavutada regressioon. Mudeli sisendiks olid pildid. Antud juhul on kasutusel Kerase loodud ResNet50 arhitektuuri implementeeriv mudel.

Kolmas mudel, mida prooviti, oli mudel C. Selle mudeli sisendiks olid CSV-st loetud

andmed ja sellel oli 11 kihti, kihtide arhitektuur nähtav Joonisel 4. Erinevalt mudel A-st ja mudel B-st ennustas mudel C tuule kiiruseid ja suundi CSV-st loetud parameetrite ja andmete põhjal. Kiiruse ennustamisel olid sisenditeks kiire langevusnurk, laiuskraad, pikkuskraad, mõlema polarisatsiooni keskmine, standardhälbed ja ristkorrelatsiooni aritmeetiline keskmine. Suuna ennustamisel olid sisenditeks lisaks mõlema polarisatsiooni X ja Y suunalise spektri piigi väärtus ja selle asukoht.



Joonis 4. Mudel C kihtide arhitektuur.

5.2 Mudeli treenimise parameetrid

5.2.1 Hüperparameetrid

Mudeli loomise funktsiooni parameetrik on mudeli hüperparameetrid, millega on võimalik määrata mudelile erinevaid näitajaid. Antud juhul määratakse mudelis esimese tiheda kihi neuronite arv ja mudeli õppimiskiirus, mida kasutatakse mudelit kompileerides.

Võimalikult efektiivsed parameetrid võimaldavad vähendada mudeli kadu - kadu on üldine nimetus, millega võrreldakse saadud tulemusi tegelike tulemustega. Valideerimise andmete kadu saadakse testandmetelt. Hüperparameetrite leidmiseks kasutati Keras Tunerit, mis leiab parimad võimalikud parameetrid. Keras Tuner laseb mudelil kõigepealt natuke treenida, seejärel tehakse varajane peatamine ja Tuner leiab kõige paremini toimivad parameetrid, et vähendada valideerimisandmete kadu. Sarnaselt leitakse ka prima valideerimise kaoga epohhi. Pärast parameetrite ülevaatamist luuakse mudel uuesti parandatud parameetritega ja jätkatakse treenimisega [24].

Mudeli optimeerijaks sai valitud Adam. Kuna sügava mudeli treenimine on iteratiivne protsess, siis hästi häälestatud parameetrid on olulised, et treenimine kiiremini ja efektiivsemalt toimiks. Adami nimi tähendab adaptiivset hetkehinnangut (adaptive moment estimation), see optimeerija ühendab impulsi ja RMSpropi algoritmid. Adam seab mudelile mitmed parameetrid, olulisim neist on õppimiskiirus [25].

Kuna tuule suuna ja kiiruse ennustamine masinõppega toimib regressiooni, mitte klassifitseerimisega, siis kadu ja ennustustäpsust hinnati keskmise absoluutse protsentuaalse vea meetodil. Antud meetod võimaldab hinnata seda, kui lähedale ennustuse väärtus on tegelikule väärtusele.

Lisaks loodi eraldi tuulenurga kao arvutamise valem. Masinõppe jaoks on keeruline aru saada kahe nurga erinevusest. Näiteks nurga 350° ja 10° vahe on 340° , kuid teisalt on antud kahe nurga erinevus ainult 20° . Valem 5.1 on tuulenurga kao arvutamise valem (n - andmete koguarv, y - mudeli ennustus, \hat{y} - tegelik väärtus). Selle valemi abil suudab masinõppe 180 kraadise määramatusega mõista nurkade sarnasust isegi juhtudel, kui nurgad on erinevates veerandites.

Kao arvutamise valem

$$\frac{\sum_{i=1}^n |1 - \cos^2(y - \hat{y})|}{n} \quad (5.1)$$

5.2.2 Muud mudeli treenimise parameetrid

Mudelit treenides on võimalik aktiveerida erinevaid parameetreid. Nendega saab näiteks määrata seda, kas mudel võtab sisse pildid, millel on lisaks VV- ja VH-polarisatsioonile kolmas kanal IA või polarisatsioonide ristkorrelatsiooni valemiga arvutatud väärtused. Lisaks saab aktiveerida piltide keeramist 45 või 90 kraadi kaupa.

5.3 Treenimise kirjeldus

Treenimisprotsessi käivitas treeningfail, mis sisaldab endas andmete sisse lugemist, treenimist ja mudeli testimist. Treenimise käigus käivitati tsükkel, kus igas iteratsioonis loeti 1000 rida andmeid. Andmeteks olid satelliidipildid ja igale pildile vastav tuule suund või kiirus. Väärtused viidi vastavusse piltidega failinimede järgi. Pildid skaleeriti väiksemaks, et vähendada andmemahutu. Andmed jagati suvaliselt treening- ja testandmeteks nii, et 80% pildi-väärtuse paaridest kasutati treenimiseks ja 20% kasutati testimiseks. Sügava närvivõrgu treenimine vajab tavaliselt väga suuri andmehulki, seega loodi andmehulga suurendamiseks ka erinevaid operatsioone. Pildihulka saab soovi korral suurendada pilte

keerates 90, 180, 270 kraadi ja peegeldades. Pärast andmete asukohtade lugemist treeniti mudelit selle iteratsiooni jooksul laetud andmetel. Testandmetel testiti mudeli täpsust ja loodi hajuvusgraafikud.

Treenimise failis on võimalik täpsustada, et kas laetakse sisse pildid, mis kasutavad kiire langemisnurka või polarisatsioonide ristkorrelatsiooni valemit. Samuti on võimalik kasutada varianti, kus kolmandaks kanaliks on korrutustehte tulemus kasutades kahte esimest kanalit (VV-polarisatsioon ja VH-polarisatsioon).

Treenimiseks kasutati Tallinna Tehnikaülikooli AI-Lab keskkonda. AI-Lab keskkonnast küsiti sbatch käsuga treeningu jaoks 8 tuuma protsessorist, 30 gigabaiti muutmälu ja ühte videokaarti. Mudeli treeninguid jooksutanud arvuti oli AMD Threadripper 3960X, NVidia 2080TI. Treenimise esimeseks sammuks oli YML faili baasil uue Anaconda keskkonna loomine. Treenimisprotsessi käivitati skriptiga, mis aktiveeris Anaconda keskkonna ja käivitas selles varasemalt OOP kujul loodud treeningfaili.

5.4 Valideerimise meetodite kirjeldus

Mudeli tulemusi hinnates võrreldakse ennustatud väärtusi *in situ* mõõdetud väärtustega. Mudelite tulemuste täpsuste hindamiseks on kasutusel kolm erinevat mõõdikut: Pearsoni korrelatsioon (r), R^2 ja ruutkeskmine viga ($RMSE$). Pearsoni korrelatsioon ehk r arvutatakse kasutades Numpy teegi funktsiooni *corrcoef*. Mida suurem on korrelatsioonikordaja absoluutväärtus, seda tugevam on korrelatiivne seos tunnuste vahel - antud juhul tähendab see seda, et ennustused järgivad mõõdetud väärtusi. See meetod iseloomustab hästi kahe andmestiku suhet. Samas on meetod tundlik erindite suhtes, näiteks kui mõni väärtus on teistest väga erinev. R^2 meetod iseloomustab mudeli kirjeldusvõimet. Erinevalt korrelatsioonist mõõdab R^2 meetod seda, kui suure osa sõltuva tunnuse koguhajuvusest moodustab regressioonhajuvus. R^2 puhul tähendab skoor 1.00 seda, et iga hinnatud suurus on täpne mõõdetud väärtusele. Seevastu skoor 0.00 on mudelil, mis ennustab alati ennustatava muutuja keskmist [26] [27]. Ruutkeskmise vea puhul arvutatakse iga ennustuse vea ruutudest ruutjuur. See on üks meetoditest, mille abil võrrelda erinevate mudelite ennustuste vigasid [28].

5.5 Kuvamine

Pärast treenimist luuakse kao ja valideerimisandmete kao graafik. Kadu saadakse treeningandmetest ja valideerimisandmete kadu testandmetest leides mudeli ennustuse ja tegeliku väärtuse vahe. Kuna eesmärgiks on vähendada kao ja valideerimisandmete kao väärtuste

erinevust, siis graafik aitab näha, millal oli kahe väärtuse vahe vähim.

Mudeli hindamise käigus luuakse ka hajuvusgraafik, mis kujutab tegelike kiiruste või suundade ja ennustatud väärtuste suhteid. Lisaks väljastatakse kaks CSV faili, millest ühes on tegelikud väärtused ja teises ennustatud väärtused.

5.6 Ruumiline testimine

Terve SAR pildi töötlemiseks on loodud eraldi meetod, kuhu on võimalik sisse anda SAR-i GeoTIFF fail. Funktsiooni sees loetakse esialgu TIFF fail GDAL-i (inglise keeles *The Geospatial Data Abstraction Library*) abil rasterfailina sisse ning seejärel hakatakse seda faili töötleva. Rasterfailist filtreeritakse ebasobivad väärtused ja andmed viiakse soovitud skaalale, kas logaritmilisele või lineaarsele skaalale. Samuti rakendatakse andmete funktsiooni, mis loob sisendfailiga samade dimensioonide ja projektsioonidega vea- ja maamaski - maatriksi väärtustega 0 ja 1. Vigased väärtused või maa tähistatakse selles maatriksis 0-ga.

Peale eelnimetatud sammude läbimist hakatakse tsükli abil tükeldama TIFF faili üksikhaaval väiksemateks alampiltideks. Alampildile rakendatakse varasemalt leitud vea- ja maamaski. Alampilt antakse ennustamiseks mudelile ette ainult siis, kui maad või vigaseid väärtuseid on pildil vähem kui 30%. Koordinaatide sobivale kujule saamiseks alampildi projektsioon transformeeritakse. Erinevate kanalite (VV-polarisatsioon, VH-polarisatsioon ja kiire langemisnurk ehk IA) alampilte loetakse eraldi ning alampildid skaleeritakse ümber 100x100 piksli suuruseks. Ümberskaleeritud alampiltidest pannakse kokku 100x100 piksli suurune pilt, mis antakse treenitud mudelile ette. Pildi keskpunkti pikkus- ja laiuskraad, rea ja tulba alguspunkt, mudeli väljastatud tuule kiirus ja tuule suund kraadides lisatakse peale iga tsükli iteratsiooni lõppu järjendisse. Järjendi sisu kirjutatakse tsükli lõppedes CSV formaadis väljundfaili. Samuti genereeritakse PNG pilt, millel on kujutatud Läänemere piirkonna kaart ning koordinaatide kaupa ennustatud tuule kiirused. Funktsiooni sees tehakse automaatselt kõike seda, mis enne mudeli treenimist tehti iga pildiga.

6. Tulemused

6.1 Treenimise tulemused

Töö käigus treeniti kolme erinevat mudelit nii kiiruse kui ka suuna ennustamiseks Sentinel-1 satelliidipiltidelt: mudel A, mudel B, mudel C. Mudelite tulemuste täpsuste hindamiseks on kasutusel Pearsoni korrelatsioon (r), R^2 ja ruutkeskmine viga. Esimene mõõdik näitab seost andmete vahel, R^2 kirjeldab mudeli täpsust ja ruutkeskmine viga näitab hindamise viga. Mudeli tulemusi hinnates võrreldakse ennustatud väärtusi *in situ* mõõdetud väärtustega. Erinevate mudelite tulemused on kirjeldatud kiiruse ennustamise puhul Tabelis 4 ja suuna ennustamise puhul Tabelis 5.

Tabel 4. Mudelite täpsused kiiruse ennustamise korral.

Mudel		Kiiruse täpsus		
		r	R^2	RMSE
Mudel A (Sügav närvivõrk)	ristkorrelatsioon	0.69	0.31	3.01
	ristkorrelatsioon, 90 kraadi kaupa keeramisega	0.65	0.25	4.02
	IA	0.71	0.33	2.95
	IA, 90 kraadi kaupa keeramisega	0.70	0.32	2.99
Mudel B (ResNet50)	ristkorrelatsioon	0.67	0.30	3.02
	ristkorrelatsioon, 90 kraadi kaupa keeramisega	0.66	0.27	3.81
	IA	0.69	0.31	3.00
	IA, 90 kraadi kaupa keeramisega	0.70	0.33	2.97
Mudel C (CSV mudel)	CSV	0.68	0.32	2.99

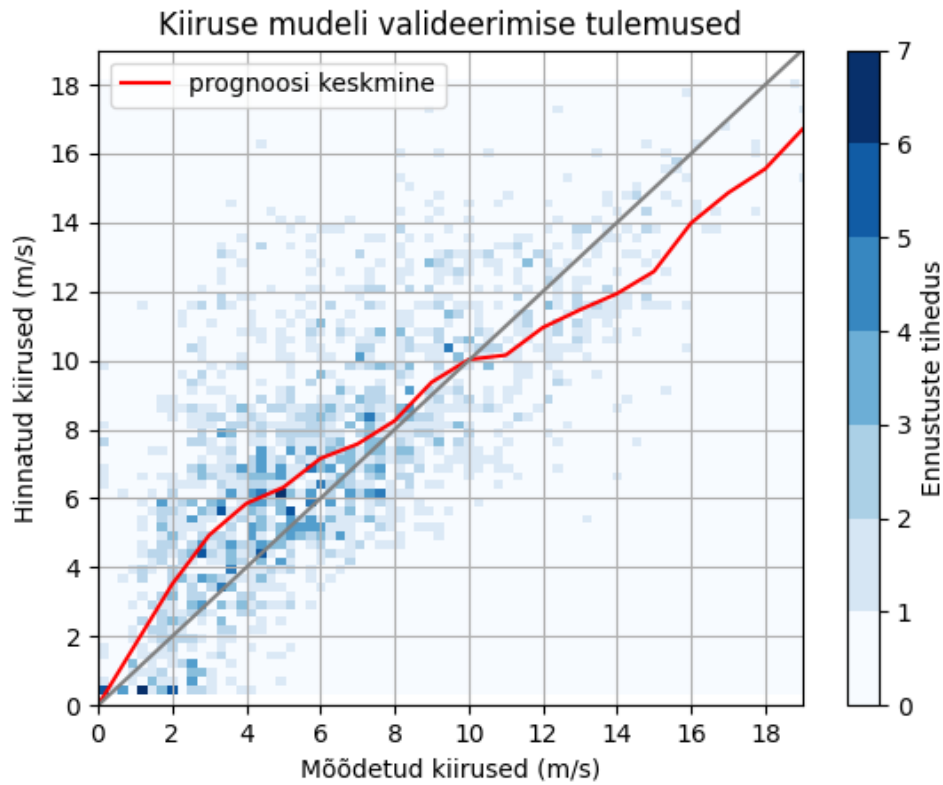
Tabel 5. Mudelite täpsused sisendiga vahemikus 0°-180° suuna ennustamise korral.

Mudel		Suuna täpsus		
		r	R^2	RMSE
Mudel A (Sügav närvivõrk)	ristkorrelatsioon	0.16	-0.78	66.22
	ristkorrelatsioon, 90 kraadi kaupa keeramisega	0.00	-1.03	70.05
	IA	0.22	-0.56	61.17
	IA, 90 kraadi kaupa keeramisega	-0.01	-1.23	72.53
Mudel B (ResNet50)	ristkorrelatsioon	0.17	-0.75	65.48
	ristkorrelatsioon, 90 kraadi kaupa keeramisega	0.00	-1.05	70.11
	IA	0.23	-0.52	60.28
	IA, 90 kraadi kaupa keeramisega	0.01	-1.15	71.87
Mudel C (CSV mudel)	CSV	0.01	-0.06	64.51

6.2 Valideerimine *in situ* mõõtmistega

Kiiruse ennustamise parima tulemuse andis mudel A, kui sisendiks olid pildid, mille kolmas kanal oli kiire langemisnurk. Selle tulemuseks oli Pearsoni korrelatsioon ehk r 0.71, R^2 skoor 0.33 ja ruutkeskmine viga 2.95 m/s. Mudeli tulemuse valideerimiseks ning hajuvusgraafiku loomiseks kasutati *in situ* mõõtmisi, mis pärinevad rannikujaamade mõõtmistest. Mudeli täpsuse võrdlemine *in situ* mõõtmistega annab teavet selle kohta, kui täpselt on mudel suutnud hinnata soovitud parameetrit - antud juhul tuule kiirust.

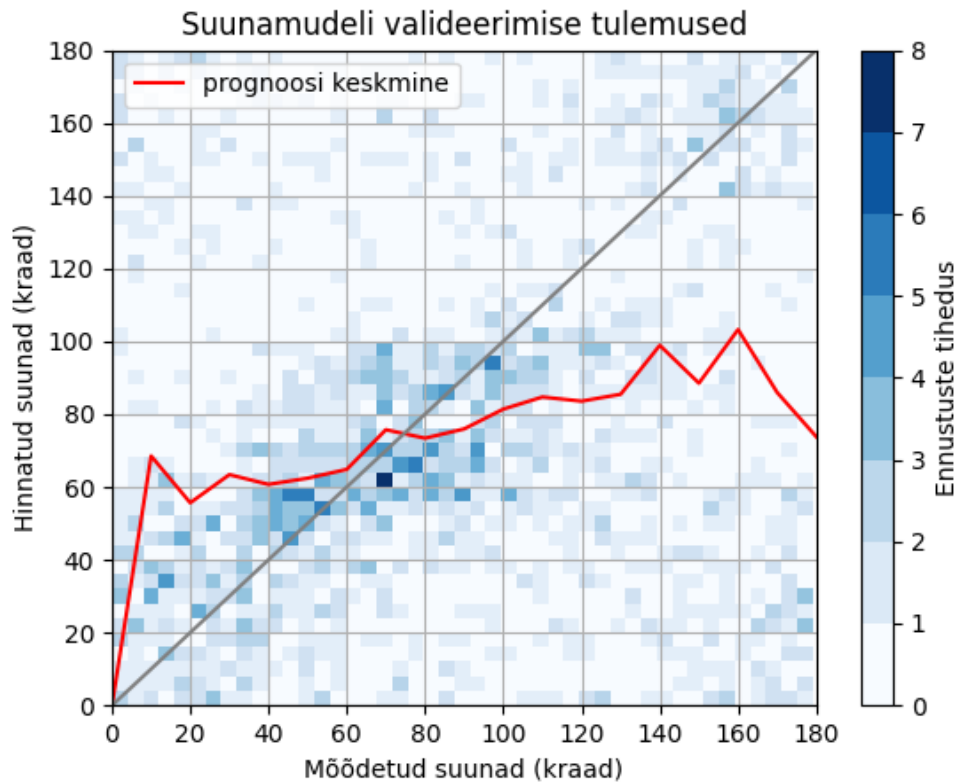
Joonisel 5 kujutatud hajuvusgraafikul on kirjeldatud hinnatud kiiruste ja mõõdetud kiiruste suhe.



Joonis 5. Mudel A valideerimise tulemused. Ennustuste tihedus on 0.25 m/s sammuga, prognoosi keskmine on arvatatud 1 m/s sammuga.

Suuna ennustamise parima tulemuse andis mudel B, kui sisendiks olid pildid, mille kolmas kanal oli kiire langemisnurk. Selle tulemuseks oli Pearsoni korrelatsioon ehk r 0.23, R^2 skoor -0.52 ja ruutkeskmine viga 60.28°.

Joonisel 6 kujutatud hajuvusgraafikul on kirjeldatud hinnatud nurkade ja mõõdetud nurkade suhe.



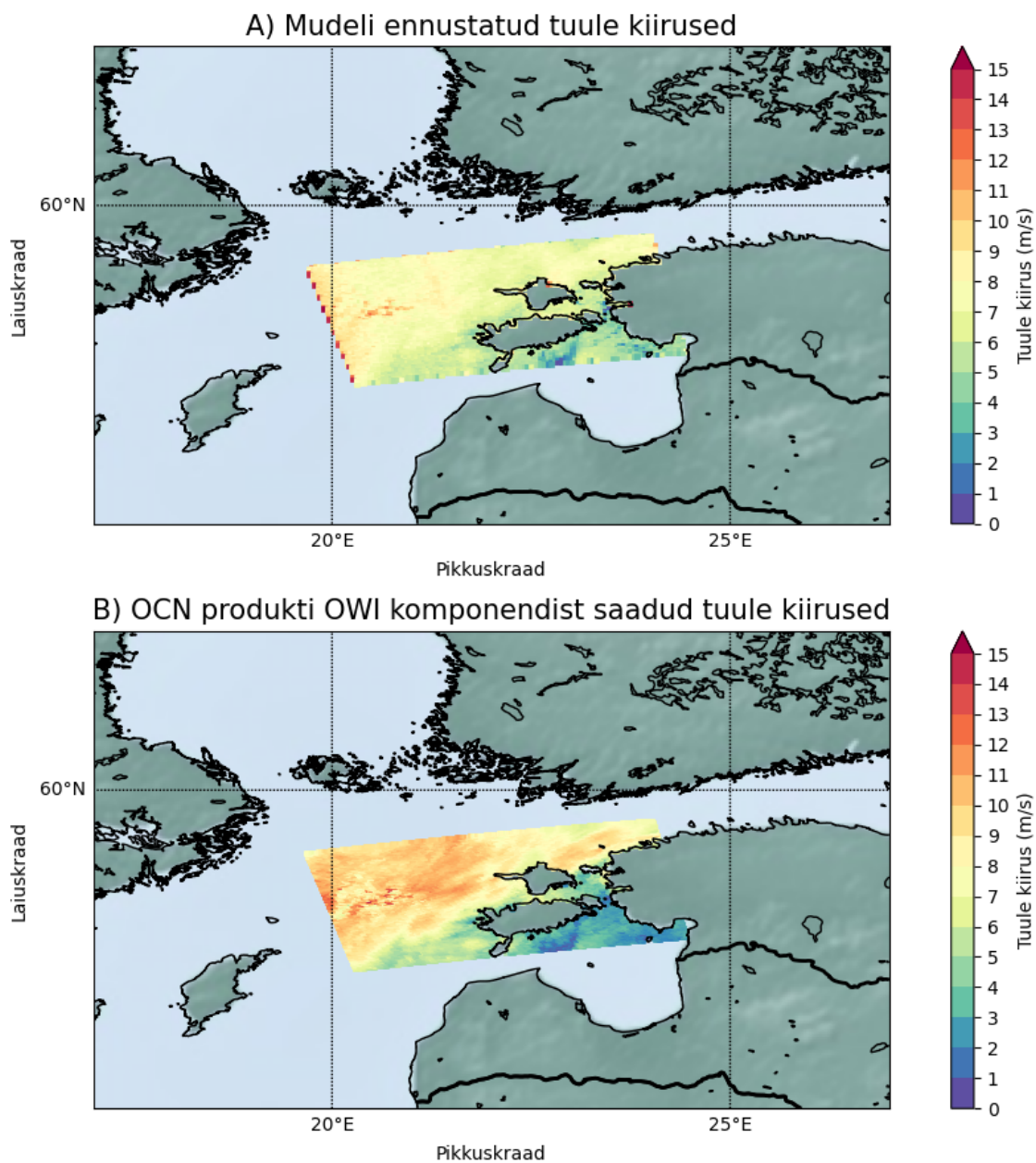
Joonis 6. Mudel B valideerimise tulemused. Ennustuste tihedus on 4° sammuga, prognoosi keskmine on arvatatud 10° sammuga.

6.3 Hinnangute kirjeldus

Tuule kiiruse ennustamisega saavutas mudel häid tulemusi. Kuna suurte tuule kiirustega pilte ja möötmisi oli vähem, siis neid ennustas mudel võrreldes väiksemate tuule kiirustega halvemini. Mudelil on mõningane kalduvus madalamaid tuule kiiruseid (ehk madalamaid kui 10 m/s) ennustada tegelikust väärtusest suuremaks. Ning vastupidi suuremaid tuule kiiruseid (ehk kõrgemaid kui 10 m/s) ennustab mudel pigem väiksemalt kui tegelikke väärtuseid (Joonis 5).

Kiiruste puhul parima tulemuse andnud mudeli (Mudel A, kui piltide kolmandaks kanaliks on kiire langemisnurk ehk IA) ennustatud kiirused kujutati Läänemere kaardil. Lisaks parima mudeli kiirustele on võrdluseks lisatud OCN produktist saadud tuule kiirused. Jooniselt 7, kus on kujutatud mudel A ennustatud kiirused ning OWI komponendist pärit tuule kiirused, saab teha mitu järeldust. Näiteks suuremate tuule kiiruste puhul on mudelil kalduvus ennustada kiirust tegelikust madalamaks. Positiivse tulemusena saab välja tuua, et Saaremaast lõunas on tuule kiiruste ühtivus erinevate produktide vahel suur. Märkatavaid sarnasusi OWI komponendist pärit tuule kiirustega on näha Läänemere keskosas, kus

kõrgemad tuule kiirused on hästi presenteeritud.



Joonis 7. Mudel A ennustatud kiirused ning võrdluseks OWI komponendi tuule kiirused.

Andes mudelile suunad vahemikus 0° - 360° ei suutnud mudel tuule suunda ennustada. Selle tulemuseks oli olukord, kus mudel hakkas ennustama alati ligikaudu suundade keskmist väärtust. Seevastu andes mudelile suunad vahemikus 0° - 180° koos 180 kraadise määramatuse kao funktsiooniga suutis mudel suundasid mõningal määral ennustada. Siiski jäi suundade ennustamine alla keskmise. Mudeli puudulikul suundade ennustamisel võis olla mitmeid põhjuseid. Näiteks võib tulemusi mõjutada piltide skaala vähendamine - algseid 1024×1024 pilte vähendati umbes 10 kordselt treenimisel mälu kokku hoidmiseks. Vähendatud skaalaga piltidelt ei tule tuulelainete puhul piisavalt selgelt lainemuster välja, mille järgi saaks suunda määrata. Kui oleks treeninud algsete piltidega, siis oleks tõenäoliselt

suuna määramine täpsem olnud.

6.4 Järeldused

Töö käigus testiti kolme mudelit. Kiiruse ennustamise puhul andis mudel A parima Pearsoni korrelatsiooni mõõtmiste ja ennustuste vahel ning R^2 skoori. Kõikide mudelite täpsused olid samas suurusjärgus ja vahed mudelite vahel olid väikesed. Tulemustest võib täheldada, et kiiruse ennustamine on mudelile keeruline juhul, kui kiirused on suuremad kui 12.5 m/s. Selle põhjuseks võib olla väiksem andmete hulk antud vahemikus. Siiski on valideerimise andmetel saavutatud 0.7 korrelatsioon hea algus täpsemate mudelite arendamiseks. Hüpooteesi osa, mis väitis, et Sentinel-1 satelliidipiltidelt on võimalik ennustada tuule kiirust, osutus tõseks.

Tulemustest järeldub samuti asjaolu, et loodud mudelid ei suuda hästi ennustada tuule suunda. Sellel võib mitu põhjust olla, kuid üheks olulisemaks võib olla piltide skaala vähendamine, mille tõttu ei pruugi lainemuster olla piltilt hästi eristuv. Seega hüpoteesi osa, kus väideti, et Sentinel-1 satelliidipiltidelt saab välise informatsiooni sobitamisetu kätte tuule suuna, osutus vaid osaliselt tõseks, kuna vaatamata keskpärastele ennustustele on nende ja *in situ* mõõtmiste vahel siiski seos.

7. Kokkuvõte

Läänemeri on piirkond, kus on palju madalaid alasid, saari ja kitsaid lahtesid. Need eripärad tingivad paratamatult ka keerulised laineolud, mille pärast peab tuule parameetreid hindama suuresti müra andmete põhjal, lainemustri tuvastamine tehisavaradari pildilt on keeruline.

Mudel, mis suudab Läänemere piirkonnas hästi ennustada tuule kiirust ja suunda, lihtsustaks praegust protsessi tuule parameetrite kättesaamiseks, teeks suure tõenäosusega paremaks lainetuse prognoosimist ning võimaldaks ka ilma täpsemini ja paremini ennustada.

Antud töö hüpotees oli, et närvivõrgu põhisele mudelile ei ole vaja sobitada välist informatsiooni tuule kiiruse ja suuna kohta ning tuule kiiruse ja suuna saab kätte ainult mudelist. Konvolutsioonilise närvivõrgu abil saab väga hästi piltidelt kätte ruumilise teabe. Võttes arvesse Sentinel-1 tehisavaradari piltide mustrites sisalduva informatsioonihulka tuulevälja struktuuri kohta, on alust arvata, et närvivõrku kasutades on võimalik piltidelt ennustada tuule suunda ja kiirust.

Töö tulemusena eeltöödeldi andmed, loodi erinevad filtreerimise funktsionaalsused, väiksemad pildi töötlemise funktsionaalsused ning testiti mitut mudelit. Näiteks loodi töö käigus filtreerimise funktsionaalsused potentsiaalsete jääalade filtreerimiseks ning ajalise parameetri järgi filtreerimiseks. Samuti on olemas võimalused pilte erinevat moodi suurendada ja vähendada. Funktsionaalsusena loodi ka ruumiline testimine, mis ennustab olemasoleva mudeli põhjal alampiltidelt tuule kiiruseid ja suundi ning kirjutab tulemused hiljem faili ning genereerib ennustatud tuule kiirustest pildi.

Treenimise käigus katsetati kolme erinevat mudelit: mudel A, mudel B, mudel C. Mudeli A ja mudeli B treenimiseks oli mitmeid variante: piltide kolmanda kanalina oli võimalik kasutada näiteks kiire langemisnurka või polarisatsioonide ristkorrelatsiooni valemit. Mõlemal juhul sai lisaks andmehulga suurendamiseks pilte keerata ja peegeldada. Mudel C-d treeniti CSV-st loetud SAR piltidelt arvutatud statistikute põhjal.

Kõik mudelid said tuule kiiruse ennustamisega satelliidipiltidelt hästi hakkama, erinevused meetodite tulemuste vahel olid minimaalsed. Kõige paremaks osutus mudel A variandiga,

kus pildi kolmanda kanalina kasutati kiire langemisnurka - Pearsoni korrelatsiooniga mõõdetud mudeli täpsus *in situ* valideerimisandmetel oli 0.7, R^2 skoor 0.3 ja ruutkeskmine viga 2.95 m/s. Seega hüpoteesi osa, kus väidetakse, et närvivõrgu põhisele mudelile ei ole vaja sobitada välist informatsiooni tuule kiiruse kohta ning tuule kiiruse saab kätte vaid mudelist, osutus tõseks.

Tuule suuna ennustamisega satelliidipiltidelt mudelid hästi hakkama ei saanud, kuid tulemuste ja mõõtmiste vahel oli siiski seos. Kõige paremad tulemused saadi mudel B variandiga, kus pildi kolmanda kanalina kasutati kiire langemisnurka - Pearsoni korrelatsiooniga mõõdetud mudeli täpsus *in situ* valideerimisandmetel oli 0.23, R^2 skoor -0.52 ja ruutkeskmine viga 60.28°. Hüpoteesi osa, et närvivõrgu põhisele mudelile ei ole vaja sobitada välist informatsiooni tuule kiiruse kohta ning tuule kiiruse saab kätte vaid mudelist, osutus osaliselt tõseks, kuna ennustuste ja *in situ* mõõtmiste vahel oli seos.

Loodud mudel on hea alus paremate ja efektiivsemate mudelite arendamiseks. Edasist uurimist vajab veel erinevate filtrite rakendamise mõju andmetel mudeli efektiivsusele.

Viited

- [1] Sander Rikka *et al.* „Radarkaugseire kasutamine Läänemere lainetuse ja tuuleväljade määramiseks“. Teoses: *Kaugseire Eestis 2018*. Observatooriumi 1, Tõravere: Tartu Ülikooli Tartu observatoorium, 2018, lk. 32–37.
- [2] Rolando L. Jordan. „The Seasat-A synthetic aperture radar system“. *IEEE Journal of Oceanic Engineering* 5.2 (1980), lk. 154–164. DOI: 10.1109/JOE.1980.1145451.
- [3] Andrea Zanchetta ja Stefano Zecchetto. „Wind direction retrieval from Sentinel-1 SAR images using ResNet“. *Remote Sensing of Environment* 253 (2021), lk. 112178. ISSN: 0034-4257. DOI: <https://doi.org/10.1016/j.rse.2020.112178>. URL: <https://www.sciencedirect.com/science/article/pii/S0034425720305514>.
- [4] Sander Rikka *et al.* „Meteo-Marine Parameters from Sentinel-1 SAR Imagery: Towards Near Real-Time Services for the Baltic Sea“. *Remote Sensing* 10.5 (2018). ISSN: 2072-4292. DOI: 10.3390/rs10050757. URL: <https://www.mdpi.com/2072-4292/10/5/757>.
- [5] Xiao-Ming Li. „Retrieval of Sea Surface Wind Speed by Spaceborne SAR Based on Machine Learning“. Teoses: *IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium*. 2020, lk. 4011–4014. DOI: 10.1109/IGARSS39084.2020.9323135.
- [6] Victor Alari *et al.* „Surface wave effects on water temperature in the Baltic Sea: simulations with the coupled NEMO-WAM model“. *Ocean Dynamics* 66.8 (august 2016), lk. 917–930. ISSN: 1616-7228. DOI: 10.1007/s10236-016-0963-x. URL: <https://doi.org/10.1007/s10236-016-0963-x>.
- [7] Matti Leppäranta ja Kai Myrberg. *Physical oceanography of the Baltic Sea*. Berlin, Heidelberg: Springer, 2009.
- [8] Sander Rikka. „Estimation of wave field parameters from SAR imagery in the Baltic sea“. Magistritöö. Tallinn, Eesti: Tallinna Tehnikaülikool, juuni 2014.
- [9] C. Hasager *et al.* „SAR-Based Wind Resource Statistics in the Baltic Sea“. *Remote Sensing* 3 (detsember 2011), lk. 119–121. DOI: 10.3390/rs3010117.

- [10] Jekaterina Služenikina. „Applications of Marine Scatterometer Winds and Quality Aspects of their Assimilation into Numerical Weather Prediction Model HIRLAM“. Doktoritöö. Tallinn, Eesti: Tallinna Tehnikaülikool, juuni 2016.
- [11] Anni Sisas. „Ehitiste tuvastamine radarsatelliidi Sentinel-1A mõõtmiste põhjal“. Magistritöö. Tartu, Eesti: Tartu Ülikool, juuni 2015.
- [12] The Alaska Satellite Facility. *What is SAR?* [Accessed: 27-03-2022]. URL: <https://asf.alaska.edu/information/sar-information/what-is-sar/>.
- [13] Kelsey Herndon *et al.* *Synthetic Aperture Radar*. [Accessed: 27-03-2022]. URL: <https://earthdata.nasa.gov/learn/backgrounders/what-is-sar>.
- [14] The European Space Agency. *Product Slice Handling*. [Accessed: 09-04-2022]. URL: <https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-1-sar/products-algorithms/product-slice-handling>.
- [15] The European Space Agency. *Level-2 Algorithms and Products*. [Accessed: 09-04-2022]. URL: <https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-1-sar/products-algorithms/level-2-algorithms>.
- [16] The European Space Agency. *Level 2 OCN Ocean Swell Spectra (OSW) Component*. [Accessed: 09-04-2022]. URL: <https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-1-sar/products-algorithms/level-2/products/ocean-swell-spectra-component>.
- [17] Florida State University Center for Ocean-Atmospheric Prediction Studies. *Scatterometry & ocean vector winds*. [Accessed: 24-03-2022]. URL: <https://www.coaps.fsu.edu/scatterometry/about/overview.php>.
- [18] EUMeTrain. *Scatterometers*. [Accessed: 24-03-2022]. URL: <http://www.eumetrain.org/data/4/438/navmenu.php?tab=2&page=3.0.0>.
- [19] Maa-amet. *Riiklik satelliidiandmete keskus ESTHub*. [Accessed: 29-04-2022]. URL: <https://geoportaal.maaamet.ee/est/Ruumiandmed/Riiklik-satelliidiandmete-keskus-ESTHub-p443.html>.
- [20] Copernicus. *Copernicus marine environment monitoring service*. [Accessed: 24-03-2022]. URL: <https://insitu.copernicus.eu/FactSheets/CMEMS/>.

- [21] Juha Karvonen. „Baltic Sea Ice Concentration Estimation From C-Band Dual-Polarized SAR Imagery by Image Segmentation and Convolutional Neural Networks“. *IEEE Transactions on Geoscience and Remote Sensing* 60 (2022), lk. 1–5. DOI: 10.1109/TGRS.2021.3097885.
- [22] Earth Observing Laboratory. *Wind Direction Quick Reference*. [Accessed: 24-03-2022]. URL: <https://www.eol.ucar.edu/content/wind-direction-quick-reference>.
- [23] Adrian Rosebrock. *Keras, Regression, and CNNs PyImageSearch tutorial*. [Accessed: 25-05-2022]. 2019. URL: <https://pyimagesearch.com/2019/01/28/keras-regression-and-cnns/>.
- [24] Tom O’Malley *et al.* *KerasTuner*. [Accessed: 25-05-2022]. 2019. URL: https://keras.io/guides/keras_tuner/getting_started/.
- [25] Marco Peixeiro. *The 3 Best Optimization Methods in Neural Networks*. [Accessed: 20-03-2022]. URL: <https://towardsdatascience.com/the-3-best-optimization-methods-in-neural-networks-40879c887873>.
- [26] Krishna Rao. *r² or R² — When to Use What*. [Accessed: 25-05-2022]. 2020. URL: <https://towardsdatascience.com/rC2B2-or-rC2B2-when-to-use-what-4968eee68ed3>.
- [27] Ako Sauga. *Determinatsioonikordaja*. [Accessed: 26-05-2022]. URL: <https://www.sauga.pri.ee/gretl/popup/determinatsioonikordaja.html>.
- [28] James Moody. *What does RMSE really mean?* [Accessed: 25-05-2022]. 2019. URL: <https://towardsdatascience.com/what-does-rmse-really-mean-806b65f2e48e>.

Lisa 1 – Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks ¹

Mina Tiia Sildam, Mario Veelaid, Andra Velve

1. Annan Tallinna Tehnikaülikoolile tasuta loa (lihtlitsentsi) enda loodud teose “Me-retuule kiiruse ja suuna leidmine Sentinel-1 satelliidipiltidelt”, mille juhendaja on Sander Rikka and Martin Simon
 - 1.1. reprodutseerimiseks lõputöö säilitamise ja elektroonse avaldamise eesmärgil, sh Tallinna Tehnikaülikooli raamatukogu digikogusse lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;
 - 1.2. üldsusele kättesaadavaks tegemiseks Tallinna Tehnikaülikooli veebikeskkonna kaudu, sealhulgas Tallinna Tehnikaülikooli raamatukogu digikogu kaudu kuni autoriõiguse kehtivuse tähtaja lõppemiseni.
2. Olen teadlik, et käesoleva lihtlitsentsi punktis 1 nimetatud õigused jäävad alles ka autorile.
3. Kinnitan, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest ning muudest õigusaktidest tulenevaid õigusi.

30.05.2022

¹Lihtlitsents ei kehti juurdepääsupiirangu kehtivuse ajal vastavalt üliõpilase taotlusele lõputööle juurdepääsupiirangu kehtestamiseks, mis on allkirjastatud teaduskonna dekaani poolt, välja arvatud ülikooli õigus lõputööd reprodutseerida üksnes säilitamise eesmärgil. Kui lõputöö on loonud kaks või enam isikut oma ühise loomingu tegevusega ning lõputöö kaas- või ühisautor(id) ei ole andnud lõputööd kaitsvale üliõpilasele kindlaksmääratud tähtjaks nõusolekut lõputöö reprodutseerimiseks ja avalikustamiseks vastavalt lihtlitsentsi punktidele 1.1. ja 1.2, siis lihtlitsents nimetatud tähtaja jooksul ei kehti.