

Mašīnmācīšanās izmantošana tālzondēšanas datu apstrādes automatizācijā

AGRIS TRAŠKOVS

DARBA VADĪTĀJS: ASOCIĒTAIS PROFESORS JĀNIS ZUTERS, DR. DAT.

Pamatproblēma (s)

- Kā izveidot automātisku vides monitorēšanas sistēmu, kas spējīga noteikt gan naftas plankumu esamību jūrā, gan dažādus sauszemes noklājuma klases (meži, pilsētas, lauki u.t.t.)?
- Kādā vidē veidot, lai būtu pieejama no dažādām vietām?
- Kādus algoritmus izmantot, lai izpildītu uzdevumus?

Problēmvide

Jūra: Lielas platības, maz novērotāju, strauji mainīgi apstākļi, piesārņotāji un piesārņojums pārvietojas. Apgrūtināta novērošana – ar lidmašīnām prasa lielus resursus (Latvijai piemēram tādu nav, 2010-2014 lidojumu stundas kopā: 7)

Sauszeme: Liela platība, bet liela daļa pārvaldīta, daudz dažādu klašu, gan lēna gan ātra klašu izmaiņa iespējama.

Satelītattēli iespējami dažādās joslās, kas dod dažādu informāciju par redzamo: radara, infrasarkanie, redzamās gaismas diapozoni arī «uzvedās» dažādi interpretācijā.

Nafta

Apmēram 62% no visiem noplūdes gadījumiem ir industrijas darbības rezultāts nevis negadījumi.

Nafta slāpē kapilāros viļņus, radar-attēlos radot tumšus «mierīgus» plankumus, kad virs jūras ir neliels vējš.

Izmanto tieši radarus lai veiktu novērojumus mākoņu noseguma gadījumā un/vai nakts novērojumiem.

Zemes virsmas pārklājuma klases

Var izmantot optiskos datus, iespējamās lēnās pārmaiņu norises dēļ, neuztraucoties par mākoņiem un citiem traucēkļiem.

Infrasarkanās frekvences un optiskās dod iespēju veidot indeksus, piemēram NDVI (**normalized difference vegetation index**), kas dod iespēju kartēt atvasinātās vērtības

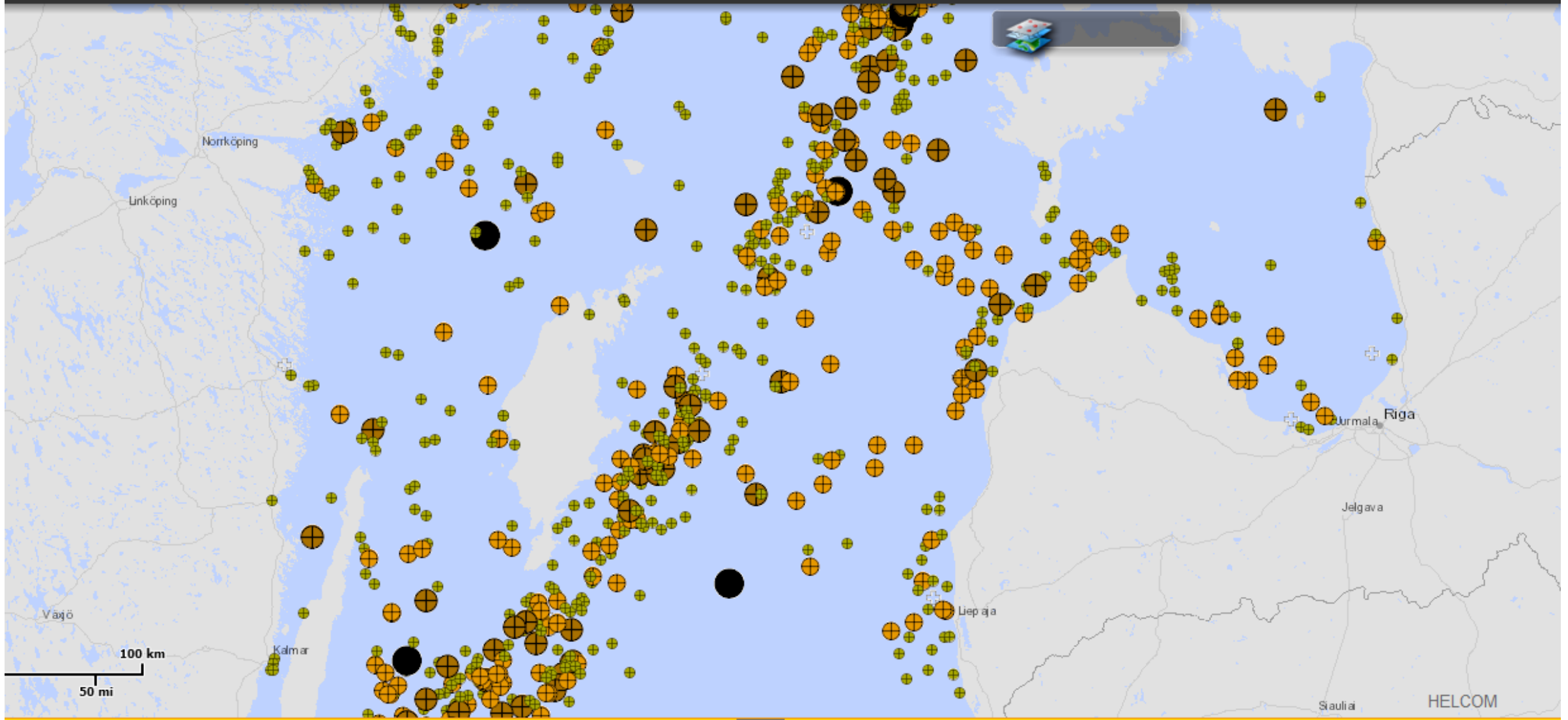
Iespējams veidot dažādas kombinācijas no papild-datiem un frekvenču datiem, pateicoties vieglai piekļuvei.

Ir liels pieprasījums vides ietekmes modeļu izveidei un novērtējumiem.

Kompleksa datu plūsma, daudzo dimensiju iespaidā.

Baltic Sea data and map service

Provided by HELCOM



Risinājumi

Ir mēģinājumi dažādas apakšproblēmas risināt: modelēt piesārņojuma pārvietošanos jūrā (CleanSeaNet)

Ir programmas, kas fokusējas uz novērojumu organizēšanu un vainīgo pušu sodīšanu (HELCOM)

Šobrīd jau arī ir automātiskās sistēmas satelītdatu apstrādei un naftas plankumu meklēšanai – Spānijas piekraste (Sentinazos) – izmanto jaunākās web tehnoloģijas, gan sadalīto datu bāzi, gan webGIS. (vienas valsts ietvaros var izmantot, sākas juridiskas problēmas multinacionālā gadījumā)

Jau ilgāku laiku ir dažādi globālie «produkti», piemēram MODIS (Moderate Resolution Imaging Spectroradiometer)

Kas nepieciešams produkta lietotājiem?

Viens standarts, ar dažādām informācijas detalizācijām.

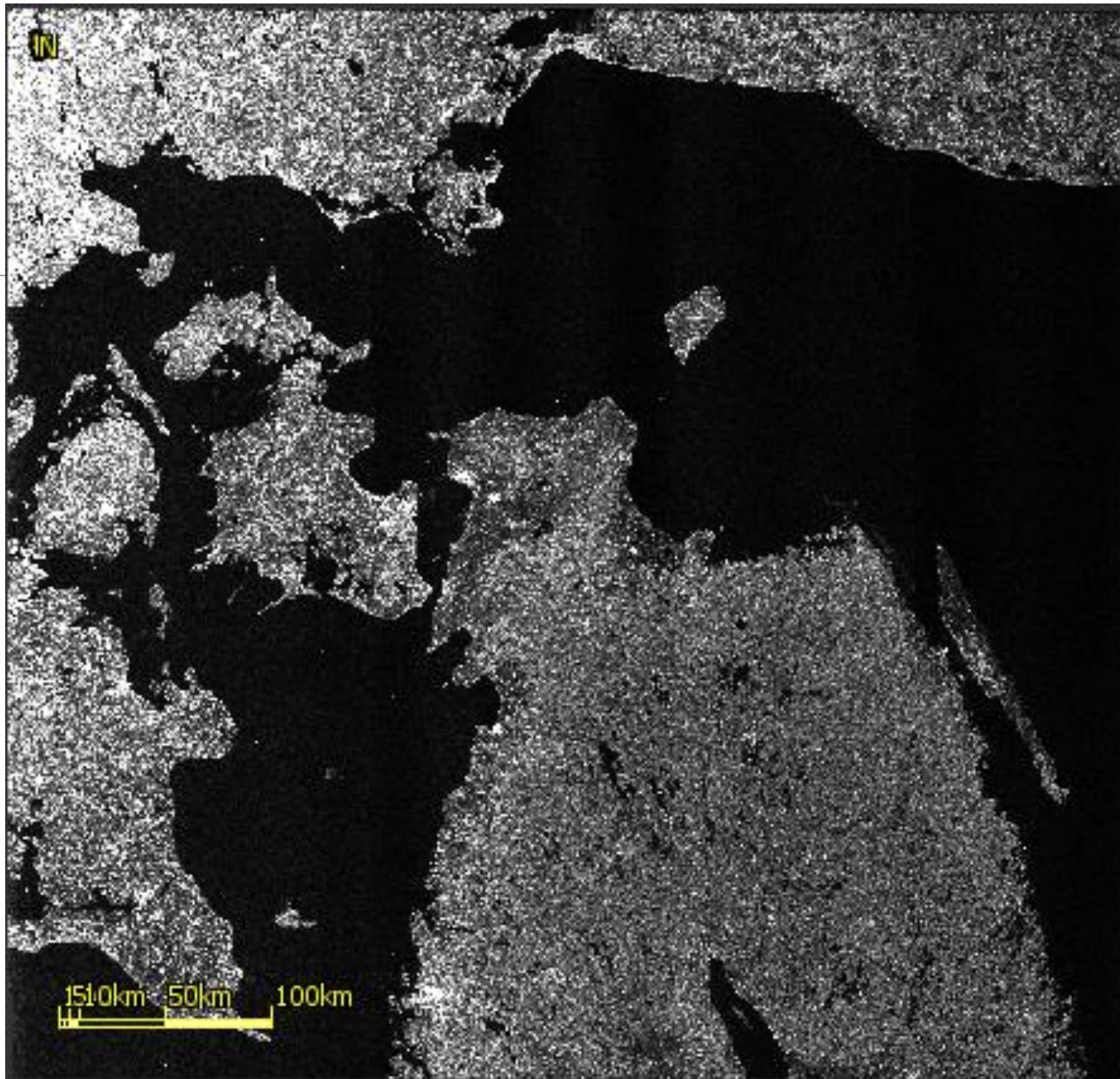
Apvienojums ar jau esošajām sistēmām un to datiem (bieži vien katram dažādas sistēmas pieejamas)

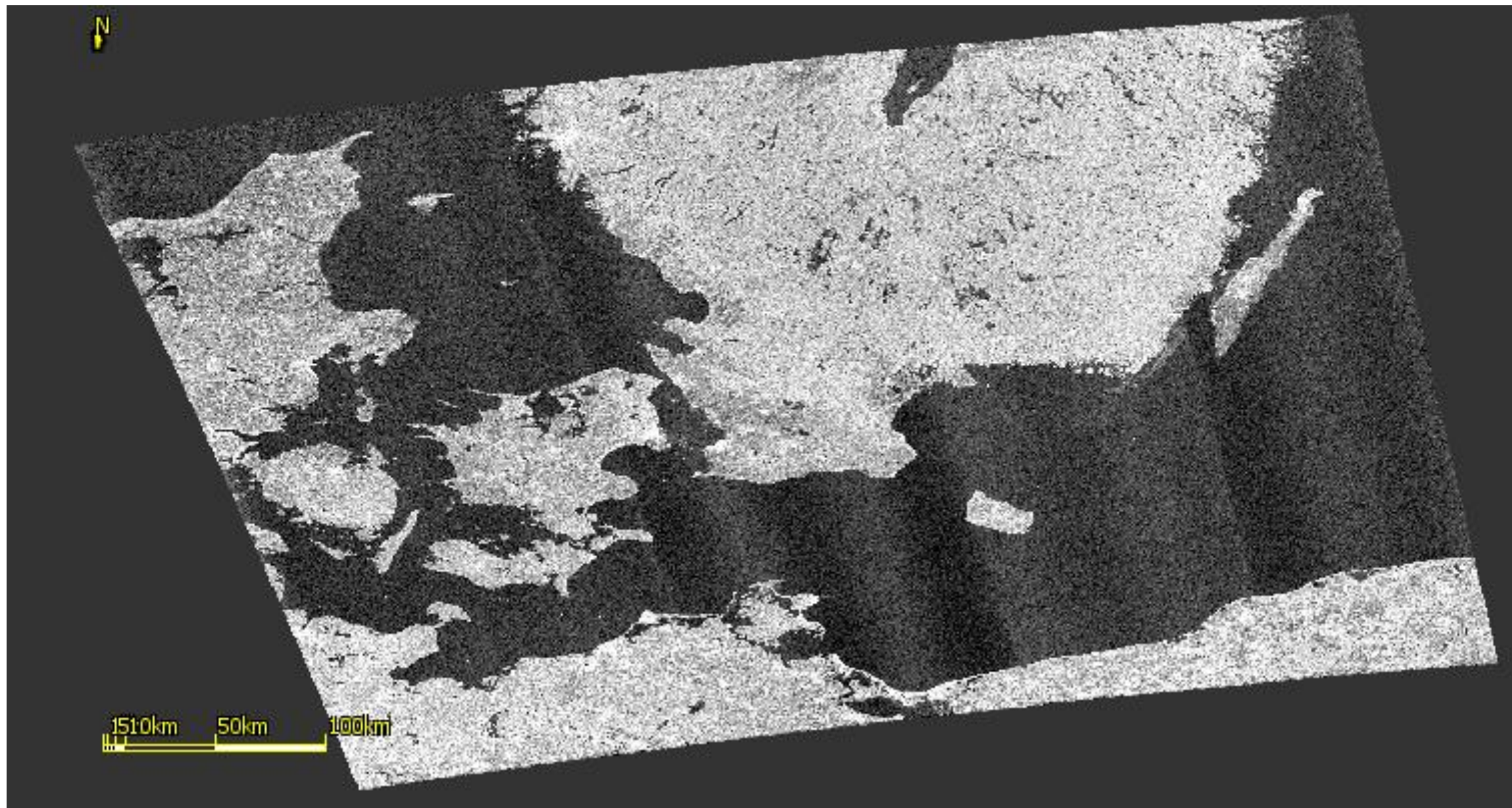
Gandrīz reālā laika dati – lai ātrāk un precīzāk noteiktu izmaiņu iemeslus.

Pieejams ar «zemu» interneta ātrumu.

Nepieciešama liela precizitāte, lai būtu pielietojumam derīgi dati.

Datiem jābūt viegli interpretējamā formātā.





Jau izmantotie risinājumi

Sentenazos: Izmanto daudzpakāpju risinājumu, kur iegūst daudz papild datus (vēja informācija, nokrišņu, temperatūras u.c. dati), gala rezultātus iegūst izmantojot mākslīgos neironu tīklus un lēmumu kokus.

Atsevišķos risinājumos tiek izmantoti ģenētiskie algoritmi un SVM.

Ir nopietnas diskusijas par meklējamo klašu skaitu nafta plankumu gadījumā.

Visi principā vienisprātis – *nav viena ideāla piegājiena/risinājuma* – nepieciešama algoritmu kombinācija.

Izdarītais

- ✓ Datu lejuplādēšanas aplikācija pārveidota lai darbojas ar jauno Sentinel Data Hub API.
- ✓ Automātiskā apstrāde Sentinel datiem () , RADARSAT-2 datiem.
- ✓ Pētījumi (ģenētiskie algoritmi + neironu tīkli) kā izmantot nevadītos (unsupervised) algoritmus, lai apmācītu un automātiski lietotu vadītos algoritmus (supervised).
- ✓ Literatūras apskats par jūras novērošanas sistēmām.
- ✓ Iesniegts raksts «Assessment of different estimation algorithms and remote sensing data sources for regional level wood volume mapping in hemiboreal mixed forests» žurnālā «*Baltic Forestry*» (šoreiz recenzents un redaktors nepazuda)

Plānotais

- ❖ Sagatavot datu bāzes vidi (Hadoop), lai veiksmīgi varētu darboties ar attēliem pilnā apmērā (attēli mēdz būt izmēros 1-7 GB).
- ❖ Pabeigt rakstus par algoritmu un sistēmu automatizāciju.
- ❖ Izveidot Web servisu, kas pārvaldītu datu izdali/apstrādi.
- ❖ Izveidot literatūras apskatu visām pieejām un aktīvajām sistēmām vides novērošanai.
- ❖ Sagatavot algoritmus ALOS PALSAR-2 un TerraSAR-X datu apstrādei.

Laiks jautājumiem!
