

NetCDF-muotoisen epäsäännöllisen liidinaineiston visualisointimenetelmät merentutkimuksessa

Ada Hytönen

Maantiede

LuK-tutkielma

24.4.2026
Turku

Kandidaatintutkielma

Tutkinto-ohjelma, oppiaine: Luonnontieteiden kandidaatti, maantiede

Tekijä: Ada Hytönen

Otsikko: NetCDF-muotoisen epäsäännöllisen liidinaineiston visualisointimenetelmät merentutkimuksessa

Ohjaaja: Harri Tolvanen

Sivumäärä: 72 sivua

Päivämäärä: 24.4.2026

Tässä kandidaatintutkielmassa tarkastellaan moniulotteisen liidinaineiston visualisointia ArcGIS Pro -ohjelmistolla siten, että veden lämpötilan ja tiheyden horisontaalinen ja pystysuuntainen vaihtelu hahmottuvat selkeästi. Aineistona käytettiin Ilmatieteen laitoksen keräämää NetCDF-formaattiin tallennettua epäsäännöllistä aineistoa Tvärminnen merialueelta. Tutkimuksen tavoitteena oli vertailla visualisointimenetelmiä ja arvioida niiden tulkittavuutta ja käytettävyyttä moniulotteisen meridatan esittämisessä.

Menetelmäkokonaisuus sisälsi pisteaineiston esittämisen, IDW-, Kriging- ja natural neighbor-interpolointikartat, hajontakuviot ja laatikkojanakuviot. Interpolointikartat tuottivat nopeasti tulkittavan yleiskuvan syvyyskerroksittain horisontaalisesta rakenteesta. Hajontakuviot paljastivat selkeästi pystysuuntaisen kerrostuneisuuden ilman interpolointivirheitä. Laatikkojanakuviot täydensivät tarkastelua tuottamalla kvantitatiivisen näkymän syvyysluokkien sisäisestä vaihtelusta. ArcGIS Pro osoittautui toimivaksi työkaluksi sen jälkeen, kun aineisto oli ensin jaettu pienempiin osiin suorituskyvyn takaamiseksi.

Lisäksi tulokset osoittivat, että lämpötilan ja tiheyden vaihtelun rakenteet erosivat horisontaalisesti ja pystysuunnassa. Horisontaalinen vaihtelu näkyi parhaiten interpolointikartoissa, jotka näyttivät pintakerroksen tasaisen lämpimän vesimassan ja syvemmän välikerroksen viileät alueet. Pystysuuntainen kerrostuneisuus hahmottui parhaiten hajontakuvioista, joissa erottuivat termokliini ja tiheyden kasvu syvyyden myötä. Visualisointimenetelmät täydensivät toisiaan ja tarjosivat kokonaiskuvan Tvärminnen merialueen fysikaalisesta rakenteesta.

Tulokset viittaavat siihen, että synoptisen ja elementaarisen tarkastelun yhdistäminen on keskeistä liidinaineiston tulkinnassa. Interpolointikartat, poikkileikkaushajontakuviot ja laatikkojanakuviot yhdessä muodostavat toistettavan ja tulkittavan visualisointiprosessin, joka soveltuu veden kerrostuneisuuden rakenteen esittämiseen. Lisäksi työ tunnistaa menetelmien rajoitteita, ja nostaa jatkokehitysmahdollisuuksia. Tutkielma tuottaa käytännönläheistä tietoa liidinaineiston visualisoinnin käytännöistä Suomen rannikkovesien tutkimuksessa.

Avainsanat: NetCDF, liidin (glider), ArcGIS Pro, geovisualisointi, veden kerrostuneisuus, ympäristödata, Itämeri

Sisällysluettelo

1	Johdanto	5
2	Teoreettinen viitekehys ja tutkimuksen tausta	8
2.1	Merentutkimus ja liitimet	8
2.1.1	Liitimet	8
2.1.2	Mittauslaitteet ja sensorit	8
2.1.3	Aineiston tallennus ja paikkatietoympäristö	9
2.1.4	Tvärminnen mittausjaksot	10
2.2	Itämeren fysikaalinen rakenne ja kausivaihtelu	10
2.3	Aineisto ja ilmiö visualisoinnin kohteena	11
2.3.1	Moniulotteisen meridatan luonne ja 4D-rakenne	11
2.3.2	Aineiston epäsäännöllisyys	11
2.3.3	Kerrostuneisuuden visualisointi	12
2.4	Moniulotteisen aineiston visualisoinnin teoreettinen perusta	13
2.4.1	Geovisualisointi ja visuaalinen analytiikka	13
2.4.2	Vuorovaikutus	13
2.4.3	Big data -ympäristö	14
2.4.4	Värikartat	14
2.5	NetCDF-aineiston visualisointi	15
2.5.1	Esitystavat	15
2.5.2	2D-tasot yleiskuvana	16
2.5.3	Kohde- ja taulukkopohjainen tarkastelu	17
2.5.4	Voxel-pohjainen tilavuusesitys	17
2.6	Tulkittavuus ja käytettävyys	18
2.6.1	Tulkittavuus	18
2.6.2	Käytettävyys	18
3	Aineistot ja menetelmät	20
3.1	Aineisto ja metatiedot	20
3.2	Aineiston esikäsittely	21
3.3	Päivän valinnan periaate	22
3.4	Visualisointimenetelmät	23
3.4.1	Pisteaineisto	23
3.4.2	IDW-interpolointi	24
3.4.3	Kriging- ja natural neighbor -interpolointi	25
3.4.4	Poikkileikkaushajontakuviot	25
3.4.5	Muut visualisointimenetelmät	26
3.5	Menetelmävertailu	27
3.6	Meriveden ominaisuuksien kvantifiointi	27
4	Tulokset	28
4.1	Valittu päivä	28
4.2	Havaintopisteiden sijainti	28
4.3	Lämpötilan horisontaalinen vaihtelu syvyyskerroksittain	30
4.3.1	0–10 m	30
4.3.2	10–20 m	31
4.3.3	20–30 m ja 30–40 m	32

4.3.4	40–50 m	32
4.3.5	0–50 m	32
4.4	Interpolointimenetelmien vertailu	32
4.5	Lämpötilan vertikaalinen poikkileikkaus	34
4.6	Tiheyden vertikaalinen poikkileikkaus	35
4.7	Lämpötilan ja tiheyden vaihtelu syvyyskerroksittain	36
5	Keskustelu	39
5.1	Menetelmävalinnat	39
5.2	Menetelmien rajoitukset	40
5.2.1	Interpoloinnit	40
5.2.2	Hajontakuviot	41
5.2.3	Laatikkojanakuviot	41
5.2.4	Liidin ja aineisto	42
5.3	Toistettavuus ja käytettävyys	43
5.4	Jatkokehitys	44
6	Johtopäätökset	47
	Kiitokset	49
	Lähteet	50
	Liitteet	60
	Liite 1. Ohjelmisto- ja laskentaympäristö	60
	Liite 2. Aineiston esikäsittelyn Python-koodi (split_data.py) sekä tuloste	62
	Liite 3. ArcGIS työnkulku	67

1 Johdanto

Merivesi on ominaisuuksiltaan monimutkainen ja jatkuvasti muuttuva. Veden ominaisuudet kuten lämpötila, tiheys ja suolapitoisuus vaihtelevat. Vaihtelua aiheuttavat muun muassa vuodenaajat, auringon säteily, tuuli, aallot ja virtaukset. Vesipatsaan rakenne voi muuttua nopeasti, minkä vuoksi meriveden tutkimus edellyttää tiheää ja moniulotteista mittausaineistoa. Näistä syistä autonomiset mittausalustat ovat nousseet keskeiseen rooliin.

Autonomiset vedenalaiset liitimet eli gliderit tarjoavat mahdollisuuden kerätä fysikaalisia, kemiallisia ja optisia havaintoja jopa kuukausien ajan, eivätkä ne vaadi jatkuvaa tarkkailua (Testor ym. 2019). Ne ovat vakiinnuttaneet asemansa osana merien tarkkailujärjestelmää. Liitimet profiloivat vesipatsaan syöksy- ja nousuliikkeiden avulla, ja samalla ne tuottavat havaintojoukkoja, jotka ovat tiheämpiä kuin perinteiset luotaukset (Ilmatieteen laitos 2022). Suomessa liitimellä toteutetaan mittauksia osana FINMARI-infrastruktuuria (FINMARI s.a.). Ilmatieteen laitos operoi liitimiä esimerkiksi Itämerellä Tvärminnen ja Utön merialueilla (Closset ym. 2015).

Liitimen keräämä aineisto on luonteeltaan epäsäännöllistä, moniulotteista ja hyvin tiheää. Havaintopisteillä on aina sijainti, syvyys ja aika. Näiden lisäksi ne voivat sisältää muita fysikaalisia, kemiallisia ja optisia muuttujia. Aineiston käsittely vaatii tiedostomuodon, joka on suunniteltu moniulotteisen havaintoaineiston tallentamiseen. NetCDF (Network Common Data Form) on alun perin kehitetty ilmasto- ja meritieteilijöiden tarpeisiin (Weather Stream 2026). Liidinaineisto tallennetaan useimmiten NetCDF-formaattiin ja CF-metadatakonventioiden mukaiseen DSG-rakenteeseen (Eaton ym. 2025; Esri s.a. a).

NetCDF tukee moniulotteista aineistoa, mutta sen visualisointi paikkatietoympäristössä on haastavaa. Kaksiulotteiset karttaesitykset eivät itsessään pysty esittämään syvyys- ja aikaulottuvuuksia, ja rasteripohjaiset 3D-kartat vaativat tasavälisen hilarakenteen (Esri s.a. m). Moniulotteisen aineiston käsittelyssä valinnoilla on merkittävästi vaikutusta siihen, miten rakenteet hahmottuvat lukijalle (Kelleher & Braswell 2021). Liidinaineisto antaa paljon tietoa, mutta sisältää myös visualisointi- ja käsittelyhaasteita, kuten horisontaalisen ja pystysuuntaisen vaihtelun samanaikainen

tarkastelu sekä ohjelmiston suorituskyvyn rajoitteet. ArcGIS Pro tukee suoraan NetCDF-formaattia, ja tarjoaa erilaisia visualisointityökaluja (Esri s.a. a). Visualisointitapojen vertailu on tärkeää, koska tulkinta perustuu usein visuaaliseen havaintoon. Lisäksi visualisointi ei ole ainoastaan karttojen tuottamista vaan prosessi, johon liittyy esimerkiksi värikartan valinta (Crameri ym. 2020).

Ilmastonmuutoksen kiihtyessä meriympäristön seurannan tarpeet kasvavat (Itämeri.fi s.a. b). Lisäksi seuranta on tärkeää myös käytännön toimijoille kuten meriliikenteelle ja merivoimille. Autonomisten liidinten tuottama aineisto mahdollistaa seurannan myös alueilla, joissa perinteisten mittausmenetelmien käyttö on logistisesti hankalaa ja kallista. Tämä korostaa visualisointimenetelmien merkitystä, sillä moniulotteisen aineiston tulkinta edellyttää teknisesti toimivia ja selkeitä esitystapoja.

Liidinaineistoja on visualisoitu ja käsitelty aiemmin esimerkiksi Python-pohjaisilla työkaluilla (Hanson ym. 2016; Gregor ym. 2019), mutta ArcGIS Pro -ohjelmiston käyttö moniulotteisen NetCDF-muotoisen liidinaineiston visualisointiin on jäänyt vähälle huomiolle. NetCDF on vakiintunut de facto -standardiksi ilmakehän, ilmaston ja merimallien hilamuotoiselle aineistolle (Hankin ym. 2009). Lisäksi se on keskeinen IHO- (The International Hydrographic Organization) ja NATO-yhteensopiva formaatti hilalliselle meri- ja ympäristödatalle (NATO Geospatial Maritime Working Group 2022). Tämä korostaa tarvetta arvioida, miten hilaton eli epäsäännöllinen liidinaineisto voidaan visualisoida tulkittavasti ja toistettavasti.

Tämän tutkielman tavoitteena on selvittää, miten moniulotteista ja epäsäännöllistä liidinaineistoa pystytään visualisoimaan ArcGIS Pro -paikkatieto-ohjelmistolla niin, että veden fysikaaliset rakenteet ovat tulkittavia ja toistettavia. Tutkimuksessa hyödynnetään Ilmatieteen laitoksen Koskelo-liitimen keräämää aineistoa Tvärminnen merialueelta, ja työ toteutetaan yhteistyössä Puolustusvoimien, erityisesti Pansion rannikkolaivaston kanssa. Tarkastelun kohteena ovat erityisesti visualisointimenetelmien käytettävyys ja tulkittavuus.

Tämän tutkielman tutkimuskysymykset ovat:

1. Miten NetCDF-muotoinen liidinaineisto voidaan käsitellä ja visualisoida ArcGIS Prossa toistettavasti ja käytännöllisesti?
2. Miten eri visualisointimenetelmät tuottavat selkeän esityksen veden fysikaalisesta vaihtelusta?
3. Miten lämpötilan ja tiheyden horisontaalinen ja pystysuuntainen vaihtelu ilmenevät aineiston Tvärminnen alueen visualisoinneissa?

2 Teoreettinen viitekehys ja tutkimuksen tausta

2.1 Merentutkimus ja liitimet

2.1.1 Liitimet

Tämän tutkielman kontekstina on suomalainen merentutkimusinfrastruktuuri. Ilmatieteen laitos operoi FINMARI-verkoston puitteissa Itämerellä Slocum-liitimiä (FINMARI s.a.). Liitimet ovat autonomisia vedenalaisia ajoneuvoja (AUV), ja ne ovat yhä tärkeämpi osa maailmanlaajuisista merien tarkkailujärjestelmää (Global Ocean Observing System, GOOS) sekä vuonna 2016 perustettua OceanGliders-verkostoa (Cauchy 2021; GOOS 2026b). Suomessa (Euro)GOOS:n kansallinen yhteyspiste on Ilmatieteen laitos (WMO 2025; GOOS 2026a).

Liitimiä käytetään merien tutkimuksessa rannikolta avomerelle, ja toiminta vedessä perustuu sukelluksiin ja nousuihin sekä sukelluksen aikana kerättyjen havaintojen lähettämiseen satelliittiyhteydellä. Liitimen keräämä aineisto rakentuu näytteenoton pohjalta, mikä vaikuttaa aineiston rakenteen ja tulkinnan ymmärtämiseen. NGDAC:n (National Glider Data Assembly Center) määritelmän mukaan yksittäiset nousut ja laskut muodostavat profiilin (profile), kaksi peräkkäistä profiilia muodostavat yhden sukelluksen (dive), kahden peräkkäisen pintaan nousun välinen kokonaisuus on segmentti (segment), ja sarja segmenttejä muodostaa liitimen lentoradan (trajectory) (IOOS 2025a). Liitimen reitti ohjelmoidaan etukäteen, mutta etäohjaus on myös mahdollista. Liitimiä on käytetty esimerkiksi mesoskaalan ilmiöiden, valtamerten rajavirtojen ja veden muutosten tutkimiseen (Rudnick ym. 2004; Testor ym. 2019).

2.1.2 Mittauslaitteet ja sensorit

Merentutkimuksessa yksi keskeisimmistä mittausvälineistä on CTD-luotain (Conductivity, Temperature, Depth), joka mittaa veden lämpötilaa, sähkönjohtavuutta ja painetta. Näistä voidaan johtaa esimerkiksi suolapitoisuus, syvyys ja tiheys (Ilmatieteen laitos 2022). Usein liitimissä on kiinni pieni CTD-luotain. Tutkimuksen aineiston keräyksessä käytetyssä liitimessä oli varusteena SBE GLDCTD, joka on liitimiä varten tehty kevyt luotain (Closset ym. 2025). Liitimet tuottavat CTD- ja bio-optisia mittauksia,

ja mahdollistavat hyvin kustannustehokkaan ja autonomisen seurannan sääolosuhteista huolimatta jopa useiden kuukausien ajalta, muodostaen jatkuvasti profiloivan CTD-järjestelmän (Rudnick & Cole 2011).

Lisäksi merentutkimuksessa käytetään liitimen lisäensoreina erilaisia optisia mittalaitteita, kuten läpäisymittareita (transmissometer) ja hiukkaskoon analysointilaitteita (particle size analyzer) sekä muita LISST-laitteita (Sequoia 2021; Sequoia 2022). Optinen läpäisykyky mittaa, kuinka suuri osa valosta läpäisee samean väliaineen (Sequoia 2012). Valo voi vaimentua väliaineen imeytymisen vuoksi tai sirota hiukkasista. LISST-laitteissa keskitytään usein sirontaan eli optiseen läpäisyyn. Hiukkanalysointilaitteet kuten LISST-200X määrittävät hiukkasten koon laserdiffraktiolla. Tämä on myös tällä hetkellä yksi yleisimmistä menetelmistä tutkimuksessa ja teollisuudessa.

2.1.3 Aineiston tallennus ja paikkatietoympäristö

Aineisto tallennetaan NetCDF-muotoon, joka on itsekuvaileva (self-describing), koneesta riippumaton array-pohjainen formaatti (UCAR s.a.). Se soveltuu hyvin moniulotteisen aineiston jakamiseen ja osajoukkojen tehokkaaseen hakuun. Yhdysvaltain IOOS National Glider Data Assembly Center asettaa NetCDF-tiedostoille standardivaatimuksia (IOOS 2025b). NetCDF tallentaa lisäksi metatiedot yhteen tiedostoon. CF-metadatakonventiot (Climate and Forecast) takaavat, että muuttujien merkitys, yksiköt, koordinaatit ja prosessointitapa ovat yksiselitteisesti määriteltyjä ohjelmistolle ja ihmiselle (Eaton ym. 2025). Tämä tukee yhteentoimivuutta ja toistettavuutta menetelmävertailussa. ArcGIS Pro käsittelee NetCDF-aineistoja suoraan, ja tukee sekä 2D-näkymiä että 3D/4D-näkymiä, mikä tekee siitä sopivan ympäristön liidinaineiston visualisointitapojen vertailemiseen (Esri s.a. a). Tässä tutkielmassa ohjelmistoa tarkastellaan käytännöllisenä paikkatietoympäristönä, jossa voidaan tarkastella eri esitystapojen vahvuuksia ja rajoituksia moniulotteisen meriympäristödatan tulkinnessa.

2.1.4 Tvärminnen mittausjaksot

Liidinaineiston havaintojen sijainti keskittyy Suomen rannikkovesiin, sillä Tvärminnen edustan merialue on yksi Suomen merentutkimuksen havaintopaikoista. Syksyllä 2023 Tvärminnessä suoritettiin kaksi mittausjaksoa (Closset ym. 2025). Alustavat lyhyet testit (Tvärminne I) osoittivat ensin laitteiston toimivuuden. Seuraavalla mittausjaksolla (Tvärminne II) keskityttiin tasapainotuksen parantamiseen ja erilaisten sukellusmoodien testaamiseen syvemmissä vesissä mahdollisimman hiljaisen toiminnan löytämiseksi. Seuraavan vuoden toukokuussa suoritettiin kolmas mittausjakso (Tvärminne III), josta kerätty data toimii myös tämän tutkielman aineistona (TVAR20242). Mittauksen tarkoituksena oli tutkia matalalla rannikkoalueella muun muassa liidinten soveltuvuutta passiiviseen akustiseen monitorointiin (PAM), jossa tallennetaan vedenalaisia ääniä ja päätellään tietoa niiden lähteistä. Kiinnostuksen kohteena oli erityisesti ihmisen toiminnan tuottamat vedenalaiset äänet, jotka voisivat aiheuttaa merkittäviä uhkia meriekosysteemeille. Passiivisen akustisen monitoroinnin ja fysikaalisten mittausten yhdistäminen korostaa liidinten monipuolisuutta ja vahvistaa niiden asemaa osana Suomen merentutkimuksen kehitystä ja FINMARI-infrastruktuuria.

2.2 Itämeren fysikaalinen rakenne ja kausivaihtelu

Itämeri on matala, murtovedestä koostuva ja geologisesti nuori meri, jonka voimakas kerrostuneisuus johtuu suolapitoisuuden ja lämpötilan pienistä vaihteluista (Itämeri.fi s.a. a). Veden rakenteeseen vaikuttavat vuodenaikojen vaihtelu, sää ja valuma-alueiden olosuhteet (Ilmatieteen laitos s.a.). Suomenlahti muodostaa Itämeren koillisen jatkeen, joka on suhteellisen matala allas. Sen keskimääräinen syvyys on 37 m ja maksimisyvyys 123 m (Aavaste ym. 2021).

Vesipatsas on kerrostunut, ja kerrostuneisuus seuraa kausivaihtelua. Kausittainen termokliini muodostuu Itämerellä yleensä toukokuun alussa ja alkaa heiketä elokuun lopulla (Aavaste ym. 2021). Termokliini sijaitsee yleensä 10–20 m:n syvyydessä, ja vesipatsaassa voidaan erottaa kolmitasoinen rakenne. Siihen kuuluu lämmin ja hyvin sekoittunut pintakerros, sen alla oleva kylmä välikerros sekä lähellä pohjaa oleva tasalämpöinen kerros (Alenius ym. 1998; Liblik & Lips 2011).

Fysikaalinen ja optinen monimuotoisuus tekee Itämerestä tutkimuksellisesti haastavan, minkä vuoksi tiheä ja jatkuva havaintoaineisto on erityisen arvokasta. Perinteiset luotaukset eivät tavoita pienen mittakaavan muutoksia, minkä vuoksi liitimet tarjoavat ainutlaatuisen mahdollisuuden seurata tarkasti rakenteellisia muutoksia.

2.3 Aineisto ja ilmiö visualisoinnin kohteena

2.3.1 Moniulotteisen meridatan luonne ja 4D-rakenne

Liidinaineistot ovat moniulotteista ympäristödataa, jossa havaintojen ulottuvuudet eivät rajoitu pelkästään sijaintiin. Syvyys ja aika muodostavat aineistoon kaksi keskeistä lisäulottuvuutta, jolloin ilmiöt rakentuvat käytännössä neljän ulottuvuuden varaan: x , y , z , t . Tällaisessa aineistossa kiinnostavat piirteet eivät aina erotu pelkästään yksittäisessä karttanäkymässä, vaan esimerkiksi pystysuuntaisessa profiilissa, ajallisena muutoksena tai niiden yhteisvaikutuksena. Tämän vuoksi tavallinen karttaesitys ei aina riitä rakenteiden havaitsemiseen, vaan tarvitaan visualisointitekniikoita, jotka tukevat moniulotteista tarkastelua ja tulkintaa (Kilsedar & Brovelli 2020; Kelleher & Braswell 2021). Moniulotteisen liidinaineiston visuaalinen analyysi on keskeinen osa meriympäristön havainnointia, sillä liitimet tuottavat suuria määriä syvyys- ja aikasidonnaisia profiileja (Hanson ym. 2017). Aineisto kattaa usein sekä fysikaalisia että biokemiallisia muuttujia. Moniulotteisuus näkyy myös sen käyttötilanteissa, sillä liidinaineistoja hyödynnetään yleensä prosessien ja ilmiöiden seurantaan, tilannekuvan muodostamiseen ja rakenteellisten piirteiden tunnistamiseen. Visualisoinnin tehtävänä ei vain ole pelkästään arvojen esittäminen, vaan tukea rakenteiden havaitsemista ja tulkintaa.

2.3.2 Aineiston epäsäännöllisyys

Liidinaineistoilla on erityispiirre, joka erottaa ne valmiiksi hilamuotoisista ympäristödataista. Havaintojen sijainti ja jakautuminen voivat olla epäsäännöllisiä. Koska liidin liikkuu ja kerää havaintoja jatkuvasti, data on yleensä piste- tai profiilipohjaista eikä tasavälinen rasteri (GRIIDC 2023). Tämä on merkittävää visualisoinnin kannalta, sillä epäsäännöllinen aineisto ei aina sovi suoraan kaikkiin esitystapoihin ilman ylimääräistä tulkintaa tai muokkaamista. Keskeistä tulkinnan

kannalta on myös se, että moniulotteisen ympäristödatan visualisoinnissa esitystavan valinta voi vaikuttaa siihen, millaiset rakenteet korostuvat ja mitkä jäävät piiloon. Esimerkiksi eri ulottuvuuksien esittämisen valinnat sekä visuaaliset ratkaisut voivat vahvistaa tai heikentää lukijan kykyä hahmottaa muun muassa kerrostuneisuutta tai ajallista vaihtelua. Liidinten liikevirheet, virtausolosuhteet ja reittipoikkeamat lisäävät epäsäännöllisyyttä sekä vaikeuttavat havaintojen tulkintaa (von Oppeln-Bronikowski ym. 2021). Tämän vuoksi tutkimuksen viitekehukseen sisällytetään sekä geovisualisoinnin periaatteet että arviointikriteerit, jotka auttavat tarkastelemaan esitystapoja systemaattisesti.

2.3.3 Kerrostuneisuuden visualisointi

Tässä tutkielmassa aineiston keskeinen rakenteellinen piirre on vesimassan kerrostuneisuus, joka ilmenee havaintoaineistossa tyypillisesti pystysuuntaisina gradientteina ja kerrospintoina. Kerrostuneisuus on seurausta lämpötilan ja suolaisuuden aiheuttamasta tiheyserosta, jossa lämpimämpi ja kevyempi vesi asettuu kylmemmän ja tiheämmän veden yläpuolelle (Li ym. 2020; Cheng ym. 2025). Kerrospintojen jyrkkyys ja ajallinen vaihtelu vaikuttavat suoraan veden sekoittumiseen ja lämpörakenteeseen, mikä heijastuu myös meriekosysteemeihin (Capotondi ym. 2012; Roch ym. 2023).

Kerrostuneisuuden tulkinta vaatii rakenteen hahmottamista, eikä se rajaudu yksittäisen arvon lukemiseen. Tätä moniulotteisuutta korostetaan myös kerrostumisen tutkimuksessa, jossa tarkastellaan esimerkiksi sekakerroksen syvyyttä ja termokliinin sekä pyknokliinin (tiheyden harppauskerros) muodostamia gradientteja (Cheng ym. 2025). Useissa tutkimuksissa on myös osoitettu, että kerrostuneisuuden vahvistuminen tai heikkeneminen voivat tapahtua nopeina muutoksina pienelläkin alueella, mikä korostaa tarkemman ajallisen ja paikallisen analyysin tarvetta (Roch ym. 2023; Cheng ym. 2025).

Nämä tutkimushaasteet kytkeytyvät myöhemmin käsiteltävään tehtäväluokitteluun (elementaariset ja synoptiset tehtävät) ja siihen, miten eri esitystavat tukevat rakenteen havaitsemista ja kuvaamista. Kerrostumisen havainnointi on synoptinen tehtävä, sillä se edellyttää usean ulottuvuuden tarkastelua yhtä aikaa (Andrienko & Andrienko 2006).

Moniulotteista aineistoa voidaan tarkastella myös Hovmöller-diagrammien avulla, jotka ovat graafin ja kartan yhdistelmä (Di Liberto 2021). Diagrammi mahdollistaa esimerkiksi vuoden karttamateriaalin tiivistämisen yhdeksi kuvaksi. Yhdellä akselilla on aika, ja toisella esimerkiksi syvyys, pituus- tai leveysaste (Heuzé ym. 2015). Liitimen keräämä aineisto sopisi hyvin myös tällaiseen visualisointiin, erityisesti jos tavoitteena on analysoida ajallista kehitystä.

2.4 Moniulotteisen aineiston visualisoinnin teoreettinen perusta

2.4.1 Geovisualisointi ja visuaalinen analytiikka

Geovisualisointi hyödyntää tehtäväkeskeisiä ja vuorovaikutteisia esityksiä aika-paikkasidonnaisten ilmiöiden tutkimukseen ja hypoteesien muodostamiseen, kun taas visuaalinen analytiikka yhdistää visualisointiin algoritmisia analyysejä ja ihmisen päättelyä (Keim ym. 2008). Lähtökohtana on suurten, heterogeenisten ja dynaamisten aineistojen tulkinta. Tämä näkökulma sopii myös liidinaineistoihin, joissa havainnot sijaitsevat epäsäännöllisesti, ja syvyys ja aika muodostavat 4D-rakenteen, jolloin normaali kartta ei riitä rakenteiden havaitsemiseen (Kilsedar & Brovelli 2020; Kelleher & Braswell 2021).

Geovisualisoinnin näkökulmasta keskeistä on vuorovaikutuksen rooli, sillä käyttäjän on pystyttävä tarkastelemaan ilmiötä useista näkökulmista, tekemään rajauksia ja tarkentamaan yksityiskohtiin. Visuaalisen analytiikan näkökulma täydentää tätä tuomalla esiin, ettei tulkinta synny pelkästään visualisoinnista, vaan ihmisen päättelyn ja mahdollisten analyysivaiheiden yhdistelmästä. Tässä tutkielmassa pääpaino on kuitenkin visualisointimenetelmien vertailussa sekä niiden tulkittavuuden ja käytettävyyden arvioinnissa paikkatietoympäristössä.

2.4.2 Vuorovaikutus

Ben Shneiderman (1996) jäsentää tavoitteellisen analyysin klassiseen periaatteeseen eli ensin yleiskuva (overview), sen jälkeen rajaukset (zoom and filter) ja lopuksi tarkennus (details on demand). Tara Munznerin (2014) What-Why-How ja nelikerroksinen ”Nested model” täydentävät tätä periaatetta ohjaamalla, millainen näkymä tukee parhaiten mitäkin tehtävää. Natalia ja Gennady Andrienko (2006)

ehdottavat erottelua elementaarisiiin (elementary tasks) ja synoptisiin tehtäviin (synoptic tasks). Elementaariset tehtävät viittaavat yksittäisiä elementtejä käsitteleville alkeistehtäville, kun taas synoptiset tehtävät edellyttävät rakenteen hahmottamista kokonaisuutena. Tässä tapauksessa kerrostuneisuuden tulkinta on synoptinen tehtävä, koska se vaatii useiden ulottuvuuksien yhtäaikaista tarkastelua. Vuorovaikutuksen näkökulmasta olennaista on myös moniulotteisen aineiston tulkinnan eteneminen iteratiivisesti, sillä yleiskuvan ja yksityiskohtien välillä liikutaan useita kertoja. Esitystavan ja käyttöliittymän on siksi tuettava sujuvaa siirtymistä näkymien välillä, sekä mahdollisuutta tarkastella yksityiskohtia ilman, että laajempi rakenne katoaa käyttäjän mielestä.

2.4.3 Big data -ympäristö

Suurten tietokokonaisuuksien visualisointia kuvataan usein niin sanotuilla neljällä V:llä, jotka ovat volyymi (volume), vaihtelu (variety), luotettavuus (veracity) ja prosessointinopeus (velocity). Nämä ominaisuudet edellyttävät tarkoituksenmukaisia päätöksiä esimerkiksi sen suhteen, mitä yksityiskohtia säilytetään ja mitä yksinkertaistetaan, jotta monimutkaisen tiedon esittäminen pysyy tulkittavana (Kelleher & Braswell 2021). Liidinaineistossa volyymi liittyy pitkään seurantajaksoon ja tiheään mittaukseen, vaihtelu eri mittausmuuttujiin ja niiden yhdistelmiin, luotettavuus mittausten laatuun ja mahdollisiin epävarmuuksiin, ja prosessointinopeus käytännössä käsittely- ja visualisointinopeuteen. Vaikka kaikki neljä V:tä eivät korostu yhtä paljon kaikissa tapauksissa, malli ohjaa tarkastelemaan aineiston ominaisuuksia sekä esitystavan vaikutusta tulkinnan kuormittavuuteen. Suurissa ja moniulotteisessa aineistossa visualisointi on myös valintaa siitä, mitä korostetaan ja mitä yksinkertaistetaan.

2.4.4 Värikartat

Tutkimus on osoittanut, että määrällisen tiedon tarkin visuaalisen esittäminen (visual encoding) perustuu sijaintiin yhteisellä asteikolla. Seuraavaksi tarkin vaihtoehto on pituus, jota seuraavat kulma, kaltevuus ja pinta-ala, jotka ovat yleensä epätarkempia tapoja esittää sama tieto (Cleveland & McGill 1984). Väriin osalta on tärkeää erottaa kirkkaus ja sävy. Kirkkaus välittää esimerkiksi järjestystä, mutta se ei yleensä yllä

sijainnin tai pituuden tarkkuuteen määrän esittämisessä. Siksi pelkkä sävyn vaihtelu ei ole täysin luotettava keino arvojen keskinäisen järjestyksen esittämiseen (Ware 2020).

Sateenkaari-värikartta on ongelmallinen, sillä se sisältää epätasaisia kirkkaushyppyjä ja värihyppyjä, jotka voivat peittää todellisia gradientteja. Lisäksi se on värinäkörajoitteisille huonosti luettavissa. Tässä yhteydessä perseptuaalisuudella tarkoitetaan sitä, että gradientit ja rakenteet hahmottuvat oikein ilman harhaanjohtavia korostuksia. Tämän vuoksi suositellaan perseptuaalisesti yhtenäisiä ja värinäkörajoitteisille (color-vision-deficiency, CVD) soveltuvia värikarttoja, joissa kirkkaus etenee tasaisesti (Crameri ym. 2020). Käytännössä toimivia vaihtoehtoja ovat Scientific Colour Maps -värikartat, jotka ovat perseptuaalisesti yhtenäisiä sekä dokumentoituja ja versioituja työn toistettavuuden tukemiseksi (Crameri ym. 2020; Crameri 2023). Sen lisäksi, että ne ovat luettavissa eri käyttäjille, ne säilyvät tulkittavina myös mustavalkotulosteissa.

Selkeän ja perseptuaalisesti yhtenäisen sekä johdonmukaisen värikartan valinta on erityisen tärkeää ei-asiantuntijalukijoille. Se on tärkeää myös silloin, kun visualisoidaan suuria ympäristödattoja, sillä tietoiset värikartta- ja esitystapavalinnat vähentävät kuormitusta ja parantavat tiedon tulkittavuutta (Kelleher & Braswell 2021).

Perseptuaalisuuden merkitys korostuu moniulotteisessa datassa, jossa muuttujien, kuten lämpötilan ja tiheyden, pienet vaihtelut voivat olla rakenteen kannalta merkityksellisiä. Värien merkitys ei ole ainoastaan esteettinen, sillä vääränlainen värikartta voi peittää esimerkiksi termokliinin tai muita ilmiöitä.

2.5 NetCDF-aineiston visualisointi

2.5.1 Esitystavat

Tässä tutkielmassa ArcGIS Pro -paikkatieto-ohjelmiston mahdollistamia esitystapoja tarkastellaan ensisijaisesti tapoina esittää moniulotteista aineistoa siten, että tietyt tehtävät ja rakenteet tulevat havaittaviksi. Olennaista ei ole siis yksittäisen työkalun tekninen toteutus, vaan se, miten esitystavat tukevat Shneidermanin (1996) analyysiprosessia, Munznerin (2014) What-Why-How-ajattelua ja Andrienkojen (2006)

tehtäväluokitusta. ArcGIS Pro valikoitui tutkimuksessa käytettäväksi ohjelmistoksi, koska sillä on mahdollista käsitellä NetCDF-aineistoa suoraan (Esri s.a. a).

Esitystavat on mahdollista jäsentää kolmeen pääluokkaan:

1. 2D-tasot, joissa tarkastellaan karttanäkymässä valittua aika- ja syvyystasoa.
2. Kohde- ja taulukkopohjaiset esitykset, joissa tarkastellaan aineistoa havaintoina tai profiileina.
3. Voxel-pohjaiset tilavuusesitykset (3D/4D), joissa tarkastellaan aineistoa tilavuutena.

2.5.2 2D-tasot yleiskuvana

Moniulotteisen datan 2D-tasot ovat usein luonteva lähtökohta, koska ne hyödyntävät kartografisesti tuttua esitystapaa sekä Shneidermanin ”overview”-periaatetta (1996). Käyttäjä voi muodostaa ensin yleiskuvan tilanteesta, ja tarkentaa tämän jälkeen kiinnostaviin aikoihin, alueisiin ja syvyystasoihin. 2D-tasot soveltuvat erityisesti elementaarisiiin tehtäviin, kuten yksittäisen ajankohdan tai lämpötilajakauman tarkasteluun tietyllä syvyydellä sekä yksinkertaiseen aikasarjavertailuun, jos viipaleita tarkastellaan ajallisesti peräkkäin (Andrienko & Andrienko 2006). Tasojen etu on myös se, että määrällisten erojen havaitseminen perustuu kartalla sijaintiin ja yhteiseen asteikkoon, mikä tukee tarkkaa lukemista verrattuna muuhun visuaaliseen esittämiseen (Cleveland & McGill 1984). Rajoitteena on kuitenkin, että kerrostuneisuus voi jäädä piiloon, jos tarkasteltavana on vain yksi syvyys- tai aikataso ilman kokonaiskuvaa.

2D-tasojä tuottavia interpolointimenetelmiä käytetään vaihtelevasti eri tutkimuksissa. Tässä tutkielmassa hyödynnettiin IDW-interpolointia (Inverse Distance Weighted), Kriging-menetelmää sekä natural neighbor -interpolointia, jotka soveltuvat epäsäännöllisen aineiston interpolointiin. Kriging on tilastollinen menetelmä, sillä se huomioi alueellisen autokorrelaation ja sitä käytetään laajasti ympäristötieteissä (Cressie 1993; Munyati & Sinthumule 2021). Natural Neighbor on deterministinen eli sääntöpohjainen menetelmä, joka perustuu pisteiden lähinaapurisuhteisiin ja tuottaa interpolointipinnan mittauspisteiden muodon mukaan (Esri s.a. f; Lucas s.a.).

Kirjallisuudessa hyödynnetään lisäksi regressiopohjaisia menetelmiä, kuten LOESS-menetelmää, jota on sovellettu epäsäännölliseen meridataan (Ridgway ym. 2002).

2.5.3 Kohde- ja taulukkopohjainen tarkastelu

Kohde- ja taulukkopohjaiset esitykset korostavat havaintojen ja profiilien tarkastelua kohteina, jolloin yksittäiseen havaintoon liittyvät ominaisuudet ja metatiedot ovat helposti saatavilla. Tämä tukee erityisesti ”details on demand”-periaatetta, jolloin käyttäjä voi tarkentaa kiinnostavaan kohteeseen tai tarkastella yksityiskohtia ilman, että laajempi rakenne hävietään. Tällainen esitys on hyödyllinen epäsäännöllisesti näytteenotetuille aineistoille, koska se ei oletta tasavälistä hilarakennetta (Esri s.a. k). Tätä visualisointitapaa käytetään esimerkiksi pystysuuntaisen rakenteen tarkasteluun (Gregor ym. 2019), ja se on paras keino havainta esimerkiksi termokliini.

2.5.4 Voxel-pohjainen tilavuusesitys

Voxel-pohjainen tilavuusesitys pyrkii tekemään moniulotteisen aineiston rakenteen näkyväksi tilavuudellisen kokonaisuutena (Esri s.a. i; Esri s.a. m). Tällöin kerrostuneisuus voidaan hahmottaa jatkuvana rakenteena. Tämä liittyy erityisesti synoptisiin tehtäviin, joissa tavoitteena on hahmottaa kokonaisrakennetta ja sen muutosta, ei yksittäisiä arvoja (Andrienko & Andrienko 2006). Samalla tilavuusesitykseen liittyy kuitenkin tulkinnallisia haasteita. Kolmiulotteisessa esityksessä rakenteet voivat peittyä toistensa taakse, ja jotkin osat voivat korostua suhteettomasti toisiinsa nähden. Näin ollen tilavuusesitystapaa arvioidaan erityisesti sen mukaan, miten hyvin se tukee synoptista rakenteen havaitsemista.

Tilavuusesityksen tuottaminen edellyttää kuitenkin tasavälisen hilarakenteen eli datan tulee olla järjestetty säännölliseen ruudukkoon. CF-konventiot painottavat, että moniulotteisen aineiston tulee muodostaa eksplisiittisesti määritelty hila (Esri s.a. m). Tämän vuoksi liidinaineisto ei ole suoraan yhteensopiva esitystavan kanssa ilman lisäkäsittelyä.

2.6 Tulkittavuus ja käytettävyys

2.6.1 Tulkittavuus

Tässä tutkielmassa tulkittavuudella tarkoitetaan sitä, kuinka hyvin visualisointi tukee käyttäjää aineiston rakenteellisten piirteiden, erityisesti vesimassan kerrostuneisuuden havaitsemisessa ja kuvaamisessa. Tulkittavuus liittyy myös aiemmin mainittuun Shneidermanin (1996) vuorovaikutteisen analyysin periaatteeseen, jossa käyttäjä muodostaa ensin yleiskuvan, rajaa kiinnostavat osat ja lopuksi tarkentaa yksityiskohtiin. Keskeistä tulkittavuuden kannalta on se, millaisia tehtäviä visualisoinnilla tuetaan. Andrienko ja Andrienko (2006) erottavat elementaariset tehtävät ja synoptiset tehtävät. Kerrostuneisuuden tulkinta on luonteeltaan synoptinen tehtävä, koska se edellyttää useiden ulottuvuuksien, kuten sijainti-syvyys-aika, yhtäaikaista tarkastelua. Tulkittavuuteen vaikuttaa myös visuaalisen esittämisen tarkkuus, sillä määrällisen tiedon esittäminen on tarkinta silloin, kun havainnot ovat sijoitettuina yhteiselle asteikolle (Cleveland & McGill 1984). Tämä korostaa tarvetta arvioida, milloin tulkinta perustuu yksityiskohtaiseen arvojen lukemiseen ja milloin rakenteiden hahmottamiseen.

2.6.2 Käytettävyys

Tässä tutkielmassa käytettävyyttä tarkastellaan ISO-9241-11 -standardin perustaisesti käytön tuloksena tietyssä käyttötilanteessa eikä järjestelmän sisäisenä ominaisuutena (ISO 2018). Käytettävyyttä tarkastellaan tuloksellisuuden (effectiveness), tehokkuuden (efficiency) ja tyytyväisyyden (satisfaction) näkökulmista, ja se riippuu aina määritellyistä käyttäjistä, tavoitteista ja käyttöympäristöstä (specified context of use) (Temoshok ym. 2025; NIST s.a.). Tämä kontekstisidonnaisuus on olennaista, koska sama järjestelmä voi olla eri tavoin käytettävä riippuen tehtävistä ja olosuhteista (Bevan 1995). Bevan korostaa, että käytettävyys voidaan ymmärtää tavoitteiden saavuttamisen onnistumisena (tuloksellisuus), siihen kuluvana työmääränä (tehokkuus) ja käytön hyväksyttävyytenä (tyytyväisyys).

Tässä tutkimuksessa käytettävyys tarkoittaa erityisesti sitä, kuinka sujuvasti NetCDF-pohjaisesta liidinaineistosta saadaan tehtyä ja muokattua visualisointeja ArcGIS Prossa

niin, että työvaiheet, ajan kulutus ja prosessin hallittavuus ovat perustellulla tasolla. Käytettävyyden tarkastelu liittyy myös suurten datamäärien käsittelyn reunaehtoihin. Moniulotteisissa aineistoissa käytettävyys kytkeytyy erityisesti siihen, kuinka nopeasti käyttäjä saa muodostettua yleiskuvan, kuinka helposti esitystapaa voidaan muokata, ja kuinka hyvin ohjelmisto tukee työskentelyä. Tämä liittyy big data -ympäristöön, jossa volyymi, vaihtelu, luotettavuus ja prosessointinopeus vaikuttavat esitystapojen valintaan (Kelleher & Braswell 2021). Jotkin menetelmät voivat olla tulkinnan kannalta hyviä, mutta eivät käytännöllisiä, jos niiden tuottaminen on liian raskasta.

Käytettävyyden näkökulmasta liidinaineiston esikäsittelyyn liittyy usein myös datan leikkaamista ja uudelleenmallintamista. GliderTools-julkaisussa (Sutton ym. 2019) korostetaan toisen vaiheen laadunvalvonnan tärkeyttä, kuten aineiston kalibroimista ja korjausta. Tätä varten on tarjolla ohjelmistopaketteja, jotka tarjoavat menetelmiä tähän. Vaikka tässä tutkielmassa esikäsittely tehtiin Pythonilla, periaate on sama sekä välttämätön käytettävyyden ja tulkittavuuden kannalta.

3 Aineistot ja menetelmät

3.1 Aineisto ja metatiedot

Tässä tutkielmassa käytettävä aineisto on Ilmatieteen laitoksen liitimen keräämä havaintoaineisto, joka on tallennettu NetCDF-muotoon ja varustettu standardoiduilla metatiedoilla. Metatiedot ovat keskeisiä, sillä ne määrittävät muuttujien merkityksen, kuten yksiköt ja standardoidut nimet, sekä miten havainnot sijoittuvat ajassa ja paikassa. Tämä parantaa aineiston yhteentoimivuutta eri analyysi- ja visualisointityökaluissa (Eaton ym. 2025). Jokaiselle havainnolle on liitetty aika, sijainti- ja syvyystieto, mikä on liidinaineiston uudelleenikäytön sekä tulkinnan kannalta olennaista (GRIIDC 2023).

Aineisto on kerätty 6.5.–15.5.2024, ja sen maantieteellinen ulottuvuus on 59.57–59.87°N ja 23.19–23.38°E, mikä sijoittuu lähelle Hankoa Tvärminnen edustalle. Tutkielman keskeisimmät fysikaaliset muuttujat ovat lämpötila (temperature) ja tiheys (density), joiden avulla tarkastellaan veden kerrostuneisuutta sekä horisontaalista ja pystysuuntaista vaihtelua. Osa muuttujista on suoraan mitattuja, kun taas osa johdettu toisista havainnoista. Aineiston metatiedot kuvaavat tämän yksityiskohtaisesti.

Esimerkiksi tiheys on johdettu muuttuja, joka lasketaan mittausjärjestelmän tuottamista lämpötila-, suolapitoisuus- ja painehavainnoista. Tämä käy ilmi metatietojen kohdista `observation_type = "calculated"` sekä `sources = "salinity temperature pressure"`. Suoraan mitattuja muuttujia edustavat esimerkiksi lämpötila, jonka mittausinstrumentiksi on määritelty metatiedoissa olevan CTD-luotain. Leveys- ja pituusarvot on merkitty mitatuiksi arvoiksi, mutta metatiedoissa huomautetaan, että sijainti ei perustu jatkuvaan GPS-mittaukseen, vaan tarkka sijainti saadaan liitimen noustessa pintaan (estimated between surface fixes). Pintanousujen välissä sijainti arvioidaan liikkeen, suunnan ja nopeuden perusteella. Lisäksi muuttujat `profile_index` ja `profile_direction` tukevat aineiston profiilirakennetta, sillä niiden avulla havainnot voidaan jäsentää nousu- ja laskuprofiileiksi sekä muodostaa poikkileikkausesityksiä.

Metatiedoista selviää myös aineiston olevan käsittelytasoa Level 0 (L0), eli siinä ei ole tehty korjauksia tai laaduntarkastusta. Tämän vuoksi työn tulokset painottuvat

rakenteiden ja vaihtelun selkeään esittämiseen ja niiden ymmärrettävyyteen (EUMETSAT s.a.). Aineistokuvauksen tarkoituksena on varmistaa, että saadut tulokset ovat tulkittavissa myös lukijalle, joka ei tunne liidinmittausten yksityiskohtia.

Aineisto sisältää yli 600 000 havaintopistettä, mikä johtuu tiheästä mittaustaajuudesta ja liitimen jatkuvasta liikkeestä. Suuri havaintomäärä sekä epäsäännöllinen näytteenotto vaikuttivat siihen, että visualisointeja ja tulkintaa varten aineisto jaettiin päiväkohtaisesti osiin, joista valittiin yksi edustava päivä tarkempaa tarkastelua varten. Esikäsittely tehtiin Python-ympäristössä (liite 2).

3.2 Aineiston esikäsittely

Ennen visualisointeja NetCDF-aineisto jaettiin Pythonilla Visual Studio Code -ympäristössä päiväkohtaisesti osiin tulkittavuuden ja suorituskyvyn vuoksi (liite 1 ja liite 2). Jokaiselle päivälle laskettiin kolme eri tunnuslukua: horisontaalinen liike, syvyysvaihtelu ja havaintojen lukumäärä. Näiden perusteella saatiin muodostettua yksinkertainen aktiivisuusindeksi, jonka avulla valittiin kaikkein edustavin päivä.

Aineisto luettiin ensin xarray-kirjastolla, ja päivän valintaa varten se muutettiin pandas-taulukoksi (The pandas development team 2026; xarray Developers 2026). Aikaleimat muutettiin pandas datetime-muotoon UTC-ajassa, ja havainnot järjestettiin loogisesti aikajärjestykseen. Puuttuvat havainnot poistettiin jättämällä pois rivit, joilta puuttui jokin keskeinen arvo muuttujista latitude, longitude tai depth. Lisäksi puuttuvien havaintojen määrä tarkistettiin vertaamalla alkuperäisen NetCDF-datan aikadimension pituutta pandas-dataframeen. Päivittäisten havaintojen summa laskettiin, ja varmistettiin, että se vastaa suodatetun taulukon rivien määrää. Tämä varmistaa analyysin läpinäkyvyyden ja datan eheyden. Päiväkohtainen ryhmittely tehtiin UTC-ajasta johdetun päivämäärän perusteella.

Horisontaalinen liike laskettiin peräkkäisten koordinaattipisteiden välisinä isoympyräetäisyyksinä käyttämällä haversine-kaavaa (Veness s.a.). Laskentaan käytettiin NumPy-kirjaston trigonometrisia funktioita. Etäisyys pisteiden välillä määriteltiin seuraavasti:

$$\Delta\varphi = \varphi_2 - \varphi_1$$

$$\Delta\lambda = \lambda_2 - \lambda_1$$

$$a = \sin^2\left(\frac{\Delta\varphi}{2}\right) + \cos(\varphi_1) \cos(\varphi_2) \sin^2\left(\frac{\Delta\lambda}{2}\right)$$

$$c = 2\arctan2(\sqrt{a}, \sqrt{1-a})$$

$$d = R * c$$

missä λ ja φ ovat pituus- ja leveyspiiri radiaaneina ja R on maapallon säde (6371 km).

Syvyysvaihtelu määriteltiin syvyysmuuttujan (depth) maksimin ja minimin erotuksena jokaisen päivän havaintojoukolta. Vertailun tueksi syvyysvaihtelu ja horisontaalinen matka normalisoitiin välille 0–1 jakamalla kunkin päivän arvo jakson maksimiarvolla:

$$\Delta z_{norm} = \frac{\Delta z}{\max(\Delta z)} \text{ ja } \Delta d = \frac{d}{\max(d)}$$

Näistä muodostettiin aktiivisuusindeksi, jossa molemmat muuttujat ovat samaa painoarvoa (0.5):

$$\text{Aktiivisuus} = 0,5 * \Delta z_{norm} + 0,5 * d_{norm}$$

Painoarvot valittiin siksi, että tavoitteena oli tunnistaa mahdollisimman monipuolinen mittauspäivä. Aktiivisuusindeksi tulkitaan päivien välisenä suhteellisenä mittana, ja tulokset tallennettiin CSV-tiedostoon (tulokset.csv) (Bandaru 2022).

3.3 Päivän valinnan periaate

CSV-tiedosto (tulokset.csv) muunnettiin taulukkomuotoon Excelissä, josta se tuotiin Wordiin. Taulukossa syvyysvaihtelu ja horisontaalinen matka pyöristettiin yhteen desimaaliin luettavuuden parantamiseksi. Aktiivisuusindeksi päätettiin kuitenkin pyöristää kahteen desimaaliin, koska esiintyvät erot ovat pieniä, eikä kolmas desimaali tuottaisi lisäinformaatiota rakenteen tulkinnan kannalta.

Aineiston koko vaikutti huomattavasti esikäsittelyvaiheisiin, sillä alkuperäinen NetCDF-tiedosto sisälsi yli 600 000 havaintopistettä. Koko datasetin käsittely kerralla ArcGIS Prossa ei ollut mahdollista ilman erittäin merkittävää hidastumista tai koko ohjelmiston

kaatumista. Tämän vuoksi aineisto jaettiin päiviin, jolloin karttaprojektit pysyivät tarpeeksi kevyinä ja työskentely oli sujuvaa.

3.4 Visualisointimenetelmät

3.4.1 Pisteaineisto

ArcGIS Prossa toteutetut visualisointimenetelmät rakennettiin tukemaan tutkielman viitekehyksessä kuvattuja periaatteita, erityisesti Shneidermanin (1996) analyysimallia ”overview first, zoom and filter, then details on demand”, sekä menetelmien tulkittavuutta ja käytettävyyttä. Kaikissa visualisoinneissa käytettiin samaa perseptuaalisesti yhtenäistä batlow-värikarttaa, ja interpolointikartoissa käytettiin samaa manuaalista skaala 2–6 °C (Crameri ym. 2020). Poikkileikkausten hajontakuvioissa (scatter plot) x-akseli esittää liitimen profiilien järjestystä (profile_index) ja y-akseli syvyyttä (depth). Väri kuvaa tarkastellun muuttujan arvoja. ArcGIS Pron käytännön työskentelyssä havaittiin myös, että suurten NetCDF-aineistojen käsittely lisää projektin kaatumisriskiä, minkä vuoksi karttaprojekti tallennettiin jokaisen työvaiheen jälkeen.

Pythonin avulla valitun päivän aineisto tuotiin pisteaineistona käyttäen Make NetCDF Feature Layer (Multidimension) -työkalua (Esri 2025b; Esri s.a. j). Jokaisella pisteellä on sijainti (longitude, latitude), syvyys (depth) ja lämpötila (temperature). Tämä mahdollistaa pisteiden karttaesityksen ja myöhemmän interpoloinnin.

Datasta poistettiin attribuuttitaulukossa pisteet ajalta 6:15:00–8:20:22, sillä ne eivät tuoneet tulkinnallisesti lisäarvoa tuloksiin heikon syvyysvaihtelun vuoksi. Lisäksi osa havainnoista sijoittui epärealistisesti virheen vuoksi maalle tai sisälsi vain yhden syvyyden havainnot. Suodatuksena tarkoituksena oli varmistaa visualisointien tulosten kuvaavan meriveden rakennetta eikä mittausvirheitä. Liitimen perhosmaista pääreittiä ei muutettu, sillä poistetut pisteet sijaitsivat sen ulkopuolella. Pisteiden poiston seurauksena hajontakuvioihin muodostuu pieni katkos reitille.

Koska lämpötila vaihtelee syvyyden mukaan, pisteaineiston attribuuttitauluun lisättiin uusi kenttä depth_classes, johon syvyys jaettiin viiteen 10 m:n luokkaan: 0–10 m, 10–20 m, 20–30 m, 30–40 m, 40–50 m. Valitun päivän syvyysvaihtelu oli noin 47 m, mikä johti

käytettyyn luokitteluun. Luokittelu toteutettiin puoliavoimina väleinä, tällöin rajat eivät mene päällekkäin toistensa kanssa. Esimerkiksi 10–20 tarkoittaa aineistossa (10, 20]. Tämän jälkeen jokaiselle luokalle muodostettiin oma kerros Definition Query -toiminnon avulla, jotta pisteet voidaan tarkastella kerroksittain. Lisäksi säilytettiin kerros ilman suodatusta, jolloin koko aineisto näkyy yhdellä silmäyksellä. Pisteaineisto projisoitiin myös ETRS89 / TM35FIN -koordinaattijärjestelmään, joka mahdollistaa metripohjaisen analyysin sekä vähentää projektiovääristymiä (GISBox 2025).

Pisteaineiston selkeä visualisointi oli tulkittavuuden kannalta hyvin tärkeää.

Värikarttana käytettiin teoreettisessa viitekehyksessä mainittuja perseptuaalisesti yhtenäisiä ja CVD-ystävällisiä väripaletteja, tarkemmin sanottuna batlow-värikarttaa (Crameri ym. 2020). Tämä tukee myös ei-asiantuntijalukijoiden tulkintaa ja toimii mustavalkotulosteessa. Värikartan lisäys ArcGIS Prohon tehtiin lataamalla tiedosto netistä ja lisäämällä sen jälkeen projektiin (Esri 2025a, Esri s.a. b). Pistekooksi asetettiin 4 pt, ja reunaviivat poistettiin selkeyden vuoksi. Pisteet jaettiin meriveden lämpötilan mukaan kahdeksaan luokkaan, ja lämpötila esitettiin kahden desimaalin tarkkuudella. Tarkat ääriämpötilat merivedellä olivat valittuna päivänä noin 2,13 °C ja noin 5,62 °C. Selkeyden ja vertailtavuuden vuoksi päätettiin kuitenkin manuaalisesti asettaa jokaiseen kerrokseen minimi- ja maksimilämpötiloiksi 2 °C ja 6 °C.

Lisäksi pistekartasta kopioitiin toinen versio, jossa pisteet esitettiin pienemmällä symbolikoolla (2 pt) ja läpinäkyvän harmaina. Tätä versiota hyödynnettiin erityisesti interpolointikerrosten tulkinnan tukena, jotta pistehavainnot on helppo paikantaa interpolointipinnan suhteen. Tällöin taustaväri ei vaikuta tulkintaan.

3.4.2 IDW-interpolointi

IDW (Inverse Distance Weighted) arvioi interpolointimenetelmänä solujen arvot niiden ympäröivien mittauspisteiden perusteella, painottaen lähellä olevia pisteitä (Kusuma ym. 2018). ArcGIS Prossa on valmiiksi IDW-työkalu (Esri s.a. g; Esri s.a. h). Ennen interpolointia pistekerrokselle muodostettiin puskurivyöhyke (Buffer), jotta interpolointi rajautuisi vain halutulle alueelle (Esri s.a. d). Jokaiselle syvyyskerrokselle suoritettiin interpolointi samoilla asetuksilla puskurivyöhykekerroksen koon mukaan. Myös koko mitatulle syvyysvälille suoritettiin IDW-interpolointi.

Interpolointikarttoja visualisoinnissa käytettiin batlow-värikarttaa, jotta tulkittavuus ja selkeys paranisivat. Lisäksi minimi- ja maksimilämpötilat asetettiin kaikissa rastereissa samalla tavalla kuin pisteaineistossa. Yhtenäiset parametrit mahdollistavat suoran visuaalisen vertailun syvyyskerrosten välillä.

3.4.3 Kriging- ja natural neighbor -interpolointi

Lämpötilarakennetta tarkasteltiin IDW-interpolointien lisäksi myös Kriging- ja natural neighbor -menetelmillä. Näin voidaan arvioida eri menetelmien vaikutusta horisontaalisen vaihtelun tulkintaan. Molemmat interpoloinnit suoritettiin käyttämällä samoja syvyysluokkia ja puskurisaluetta kuin IDW-interpoloinnissa. Tämä mahdollisti interpolointimenetelmien tulosten suoran vertailun. Sekä Kriging- että natural neighbor-rasterit muodostettiin ArcGIS Prossa niille tarkoitetuilla työkaluilla. Kriging tuottaa rasterin, joka huomioi myös alueellisen autokorrelaation, ja natural neighbor tuottaa arvot ainoastaan pistejoukon koon mukaan. Kaikkiin kerroksiin lisättiin myös sama visualisointi kuin IDW-interpoloinneissa, mikä mahdollisti menetelmien rakenteellisen vertailun.

Kriging- ja natural neighbor -interpolointikartat eivät olleet varsinaisen interpolointianalyysin päämenetelmiä, vaan ne otettiin mukaan ensisijaisesti vertailun vuoksi. Tarkoituksena oli arvioida vaikuttaako interpolointimenetelmän valinta liidinaineiston horisontaalisen vaihtelun tulkintaan. Koska analyysi tehtiin alusta alkaen IDW-interpoloinnilla ja vertailumenetelmät eivät tuottaneet merkittäviä eroja rakenteeseen, IDW toimi ensisijaisena ja tulkintaa ohjaavana interpolointimenetelmänä. Kriging ja natural neighbor toimivat siis täydentävinä tarkistusmenetelminä, joilla varmistettiin, ettei tulkittava rakenne ollut riippuvainen interpolointimenetelmien valinnasta.

3.4.4 Poikkileikkaushajontakuviot

Uuteen karttapohjaan lisättiin pistekerros, mutta tällä kertaa mukaan otettiin leikatusta datasta useampi muuttuja, kuten tiheys (density) ja profiilin järjestys sekä suunta (profile_index/profile_direction). Liidin liikkuu vuorotellen ylös-alas ja muodostaa näin nousu- ja laskuprofiilin suunnat. Tämä kerros suodatettiin samalla työkalulla kuin

syvyyskerrokset niin, että siinä näkyvät vain laskuprofiilit (`profile_direction = 1`). Merentutkimuksessa laskuprofiili on yleisin valinta, sillä laskuvaiheen profiilin mittaukset ovat yleensä laadukkaampia ja sisältävät vähemmän häiriöitä (Ohnemus ym. 2018).

Poikkileikkaus tehtiin ArcGIS Pron hajontakuviolla. X-akseli kuvaa profiilien järjestystä reitillä (`profile_index`), y-akseli syvyyttä (`depth`) ja batlow-värikartta tarkasteltavaa muuttujaa. Syvyyskaalan suunta on käännetty tulkittavuuden helpottamiseksi. Lämpötilalle ja tiheydelle muodostettiin seitsemän luokkaa. Nämä perustuivat valitun päivän minimi- ja maksimiarvoihin. Hajontakuviot vietiin kartta-asetteluun, jossa lisättiin karttaselitteenä muuttujan arvot esittävät väriluokat. Symboliksi valittiin ääriiviivaton kolmio, jotta vaihtelu ja kerrostuneisuus erottuisivat selkeästi.

3.4.5 Muut visualisointimenetelmät

Interpolointikarttojen ja vertikaalisten poikkileikkausten lisäksi tuotettiin myös muita täydentäviä visualisointeja. Ne tukevat liitimen keräämän aineiston pistemäistä rakennetta ja auttavat tulkitsemaan horisontaalista ja pystysuuntaista vaihtelua eri näkökulmista.

Jotta liitimen liike ylös ja alas havainnollistuisi, tuotettiin syvyys-aika-viivadiagrammi Line Chart -työkalulla. X-akselin muuttujana on aika ja y-akselilla syvyys käännettynä. Kaavio esittää jatkuvan nousu- ja laskuliikkeen synnyttämän epäsäännöllisen profiiliarjan, josta myös muut visualisoinnit muodostuvat. Tämä toimi tukena erityisesti hajontakuvion poikkileikkausten rakenteen tulkinnassa.

Syvyysluokkien välistä jakauman vaihtelua tutkittiin laatikkojanakuvioden avulla, mitkä toteutettiin Box Plot -työkalulla. Laatikkojanakuviot muodostettiin sekä lämpötilasta että tiheydestä. Luokkamuuttujana käytettiin syvyyskerroksia (`depth classes`). Laatikkojanakuviot tuottivat myös automaattisesti tilastolliset tunnusluvut (minimi, maksimi, mediaani, keskiarvo, kvartiilit sekä kvartiiliväli).

3.5 Menetelmävertailu

2D-interpolointirasterikerroksia käytettiin muodostamaan yleiskuva syvyyskerrosten horisontaalisesta vaihtelusta. Yhdenmukaiset ja manuaalisesti asetetut minimi- ja maksimiarvot (2–6 °C) tekevät syvyyskerrosten välisestä visuaalisesta vertailusta yksinkertaista ja tulkittavaa. Interpolointi tuottaa kuitenkin menetelmänä arvion, minkä vuoksi tehtiin pisteaineiston poikkileikkaus hajontakuvion muodossa. Näin pystysuuntainen gradientti eli muutoksen nopeus ja kerrostuminen ovat helposti nähtävissä havaintopisteistä. Hajontakuvioita täydennettiin laatikkojanakuvioilla, jotka kuvaavat jokaisen syvyysluokan tilastollista vaihtelua. Pisteaineisto, IDW-interpolointi, vertikaalisen poikkileikkauksen hajontakuvio ja laatikkojanakuvio muodostavat menetelmäkokonaisuuden, jossa jokainen menetelmä täydentää toisiaan.

3.6 Meriveden ominaisuuksien kvantifiointi

Lämpötilasta ja tiheydestä muodostettiin syvyysjakaumat numeerisen vaihtelun tarkastelua varten. Luokkamuuttujana käytettiin syvyyskerroksia ja arvoina lämpötilaa ja tiheyttä. Syvyysjakaumien tarkastelussa hyödynnettiin laatikkojanakuvioita, jotka muodostivat myös jakaumaa kuvaavat tunnusluvut. Ne vietiin Exceeliin ja koottiin taulukoiksi.

4 Tulokset

4.1 Valittu päivä

Aktiivisuusindeksin perusteella edustavimmaksi päiväksi valikoitui 11.5.2024 (taulukko 1). Päivällä on poikkeuksellisen suuri horisontaalinen liike (noin 88 km) sekä riittävä syvyysvaihtelu. Mittauspisteitä päivällä on hieman yli 77 700. Valittu päivä eroteltiin alkuperäisestä NetCDF-aineistosta ja kirjoitettiin erilliseksi tiedostoksi (valittu_paiva.nc), jota käytettiin myöhemmissä visualisoinneissa.

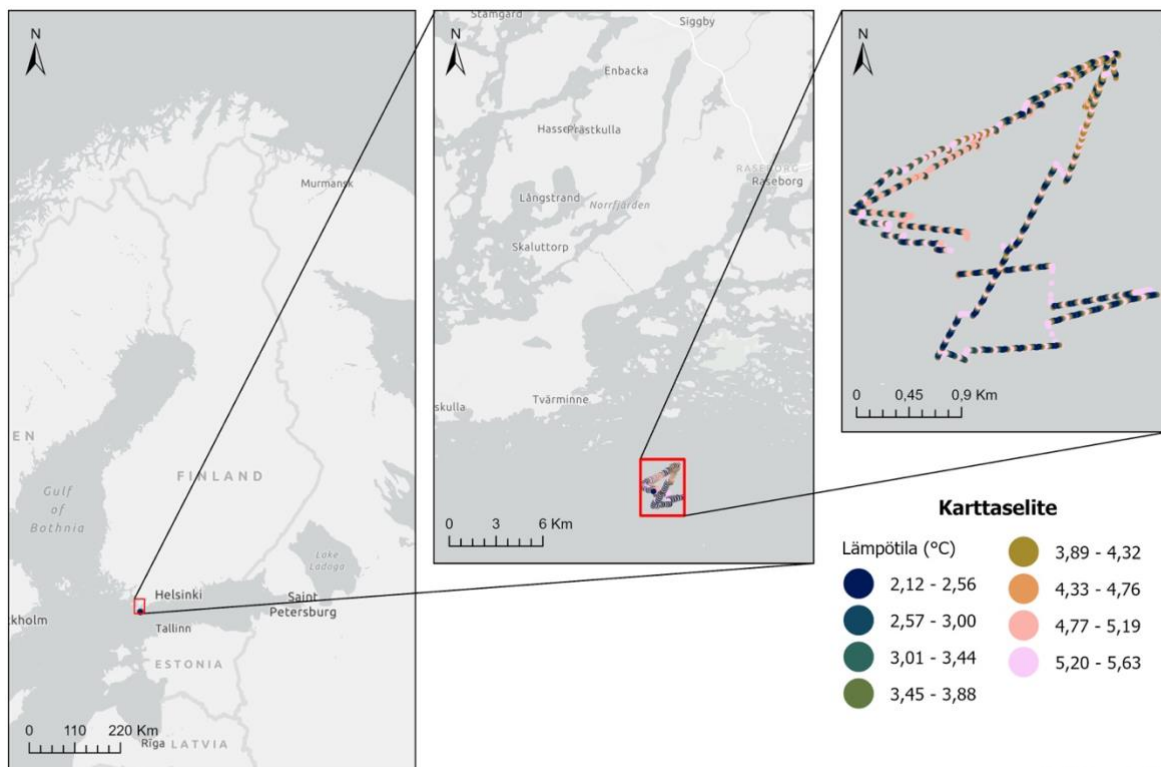
Taulukko 1. Liitimen NetCDF-aineiston päiväkohtaiset tunnusluvut ja laskettu aktiivisuusindeksi 6.-15.5.2024. Havaintojen määrä kuvaa suodatettujen pisteiden lukumäärää. Syvyysvaihtelu on päivän syvimmän ja matalimman havainnon erotus metreinä. Horisontaalinen matka on kuljettu matka kilometreinä. Tämä laskettiin koko vuorokauden havaintopisteistä ennen visuaalista suodatusta, minkä vuoksi karttakuvioissa näkyvä reitin laajuus ei vastaa kaikkia horisontaalisen matkan muodostavia pisteitä. Normalisoidut arvot on skaalattu välille 0–1. Aktiivisuusindeksi kuvaa päivien keskinäistä aktiivisuutta suhteellisenä mittana.

Päivä	Havaintojen määrä	Syvyysvaihtelu (m)	Horisontaalinen matka (km)	Syvyysnorm. (0–1)	Matkanorm. (0–1)	Aktiivisuus (0–1)
6.5.2024	34 013	38,5	28,7	0,67	0,33	0,50
7.5.2024	66 374	46,4	19,9	0,80	0,23	0,52
8.5.2024	76 891	55,1	20,0	0,95	0,23	0,59
9.5.2024	79 290	53,7	20,1	0,93	0,23	0,58
10.5.2024	78 051	57,8	20,8	1	0,24	0,62
11.5.2024	77 715	46,9	88,0	0,81	1	0,91
12.5.2024	78 471	55,8	20,0	0,97	0,23	0,60
13.5.2024	87 386	49,0	20,2	0,85	0,23	0,54
14.5.2024	36 277	33,4	10,2	0,58	0,12	0,35
15.5.2024	1	0	0	0	0	0

4.2 Havaintopisteiden sijainti

Sijaintikartassa (kuva 1) esitetään liitimen reitti sekä havaintopisteet mittauspäivältä 11.5.2024. Kuvassa on kolme eri mittakaavan karttaa. Vasemmanpuoleisessa kartassa näkyy liitimen sijoittuminen Suomenlahden alueella. Tämä antaa kontekstin ympäröivään merialueeseen. Keskimmaisessä kartassa näkyy etäisyys rannikkoon. Mittaukset on otettu Hankoniemen ja Tammisaaren ulkosaariston eteläpuolelta

Tvärminnen merialueelta. Oikeanpuoleisimmasta kartasta on mahdollista nähdä mitatut lämpötilat ja pisteaineiston muodostama perhosmainen geometria ja alue, johon myöhemmät visualisoinnit suhteutetaan.



Ada Hytönen, 3.3.2026
Ilmatieteenlaitos, "Tvärminne deployment May 2024"

Kuva 1. Havaintopisteiden sijainti ja lämpötilat 11.5.2024. Kolmen eri mittakaavan kartta näyttää liitimen reitin Suomenlahden alueella sekä mitatut lämpötilat.

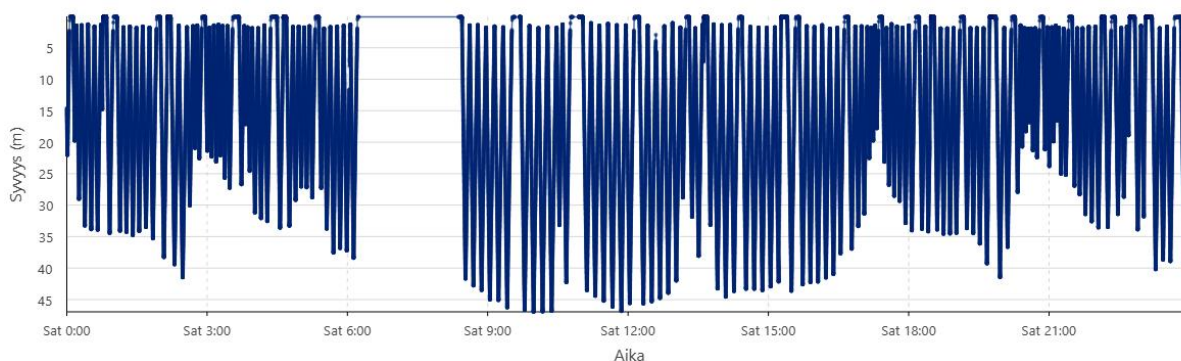
Liitimen reitti muodostuu selkeästä pisteprofiilien jonosta. Kartasta on nähtävissä yhdellä silmäyksellä liitimen liikkeen aiheuttama lämpötilavaihtelu. Havaintopisteet muodostavat tiheän, mutta epäsäännöllinen ajan ja syvyyden mukaan rakentuvan sarjan. Liitimen jatkuva liike ylös ja alas estää tasavälisen hilarakenteen muodostumisen. Tämä on tärkeää huomioida myöhemmissä tulkinnoissa, koska se selittää, miksi esimerkiksi interpolointikerrokset voivat olla paikoin epätasaisia. Poikkileikkauskuviot perustuvat myös tämän vuoksi `profile_index`-muuttujaan eikä todellisiin etäisyyksiin.

Sijaintikartta antaa lähtökohdan tulosten tulkinneille, koska se osoittaa, missä horisontaalinen vaihtelu tapahtuu, mitä pisteitä käytetään, ja kuinka poikkileikkaukset

sijoittuvat reitin kannalta. Tämä helpottaa myöhemmin tulkittavien visuaalisten rakenteiden paikantamista suhteessa reittiin.

Liitimen vuorokauden aikana kulkema kokonaismatka oli noin 88 km, mutta kartalla näkyvä reitti vaikuttaa huomattavasti lyhyemmältä. Tämä johtuu siitä, että kartta näyttää liikkeen kaksiulotteisena reittinä. Todellisuudessa matka muodostuu kymmenistä tuhansista pienistä siirtymistä liitimen jatkuvan liikkeen seurauksena. Käytetty Haversine-kaava huomioi nämä, mikä kasvattaa kokonaismatkaa. Lisäksi kartasta poistettiin poikkeamia ylhäältä ja alhaalta, mikä lyhensi reitin näkyvää osuutta.

Liitimen syvyyskuvaava viivadiagrammi (kuva 2) havainnollistaa selkeästi, miten havaintopisteet ovat muodostuneet liitimen jatkuvan liikkeen seurauksena, ja miksi tasavälistä hilarakennetta ei syntynyt. Liidin tuotti yhden vuorokauden aikana mittavan määrän nousu- ja laskuprofiileja, jotka ulottuivat maksimissaan noin 47 m:n syvyyteen. Profiilit muodostavat epäsäännöllisen mutta tiheän pistesarjan. Tämä sarja muodostaa poikkileikkausdatan rakenteen, jota tarkastellaan myöhemmin.

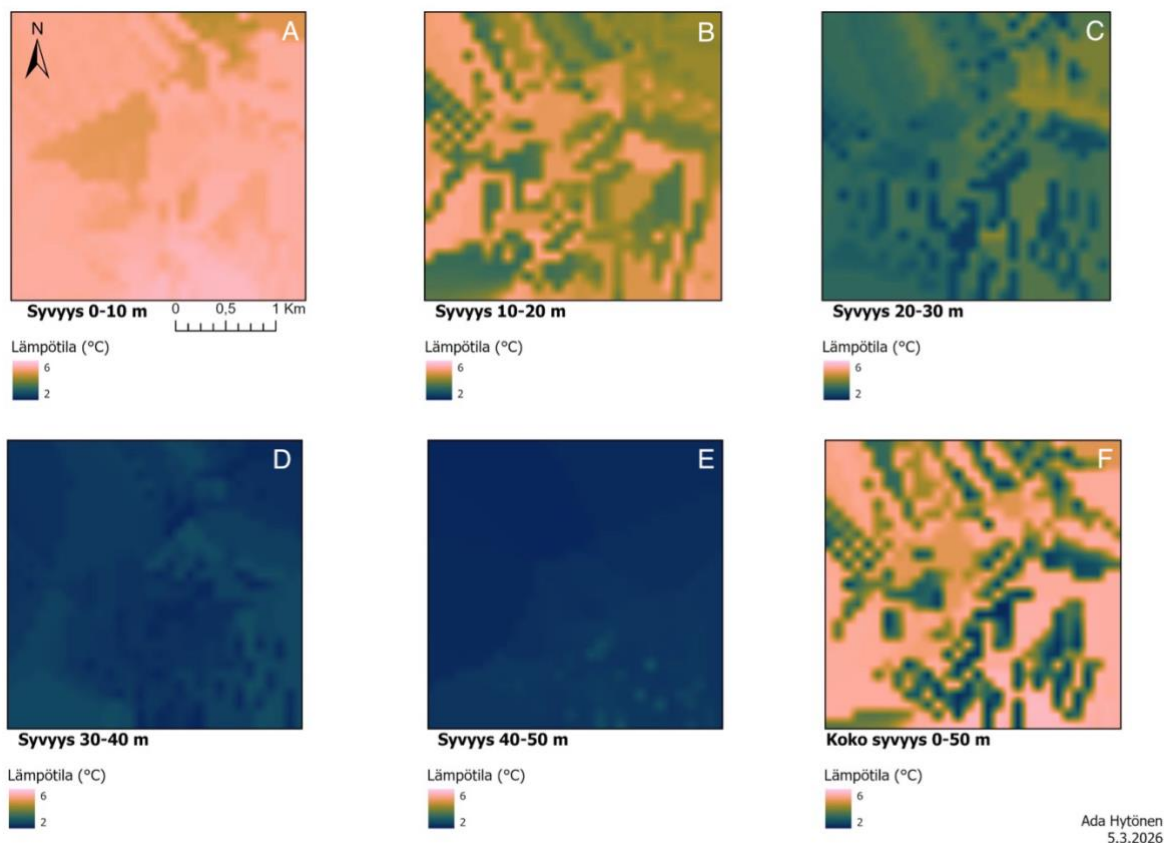


Kuva 2. Liitimen syvyysprofiili yhden vuorokauden aikana (11.5.2024). Liitimen jatkuva liike ylös ja alas muodostaa profiileita, jotka ulottuvat noin 47 m:n syvyyteen. Tämä tuottaa epäsäännöllisen kuvion, joka auttaa ymmärtämään interpolointipintojen muodon sekä poikkileikkausten profiilipohjaisen x-akselin.

4.3 Lämpötilan horisontaalinen vaihtelu syvyyskerroksittain

4.3.1 0–10 m

Syvyyskerrosten IDW-interpoloinnit näyttävät lämpötilan rakenteen 0–50 m:n syvyydessä havaintopäivänä (kuva 3). Jokaisessa kartassa käytettiin manuaalisesti asetettua 2–6 °C:n skaalaa ja samaa värikarttaa, jotta visualisointien välinen vertailu olisi helposti tulkittavaa. Lisäksi jokainen karttakuva on rajattu täsmälleen samaan alueeseen, mikä varmistaa yhdenmukaisen tarkastelun.



Kuva 3. Interpolointikartat syvyyskerroksittain 2–6 °C:n väriskaalalla. A) 0–10 m, B) 10–20 m, C) 20–30 m, D) 30–40 m, E) 40–50 m ja F) 0–50 m. Yhtenäinen väriskaala mahdollistaa helpon visuaalisen vertailun.

Pintakerroksen (kuva 3A) interpolointikartta näyttää suurimman osan alueen olevan 5–6 °C. Tämä on yhtenäinen Ilmatieteen laitoksen (2021) pintalämpötilan 30 vuoden (1991–2020) tilaston kanssa, jossa Suomenlahdella sijaitsevassa Harmajassa meriveden pintalämpötila on ollut toukokuussa keskimäärin 6,8 °C. Tämä syvyyskerros on horisontaalisesti melko yhtenäinen.

4.3.2 10–20 m

Toinen syvyyskerros (kuva 3B) sisältää selkeää horisontaalista vaihtelua. Osa alueista on yhtä lämpimiä kuin pintakerros, mutta joissain kohdissa lämpötila laskee kohti 3–4 °C:ta. Tämä vastaa kausittaisen termokliinin muodostumisen vaihetta, joka kehittyy Itämerellä tyypillisesti toukokuun aikana ja sijoittuu 10–20 m:n syvyyteen (Aavaste ym. 2021). Lämpötilakerrostuneisuus ei siis ole keväällä vielä kokonaan vakiintunut, mutta lämpimän pintavesikerroksen alle alkaa muodostua lämpötilagradientti.

4.3.3 20–30 m ja 30–40 m

Syvemmillä kerroksilla (kuva 3C ja kuva 3D) interpolointikartat tummenevat selkeästi, mikä viittaa kylmään välikerrokseen, jonka lämpötila oli mittauspäivänä noin 2,5–4 °C. Tämä vastaa Itämeren tyypillistä kausirakennetta, jossa havaitaan kolmitasoinen lämpötilaprofiili. Ensimmäinen kerros on lämmin pintakerros, jonka jälkeen on termokliini ja kylmä välikerros (Aavaste ym. 2021). 30–40 m:n interpolointikerros on väriltään kokonaan sinisävyinen eli horisontaalista lämpenemistä ei ole vielä tapahtunut.

4.3.4 40–50 m

Syvin interpolointikerros (kuva 3E) on lähes täysin tummansininen eli meriveden lämpötila on noin 2–3 °C. Syvin vesikerros on horisontaalisesti heikosti vaihteleva, koska tuulen ja aaltojen vaikutus ei yllä enää niin syvälle.

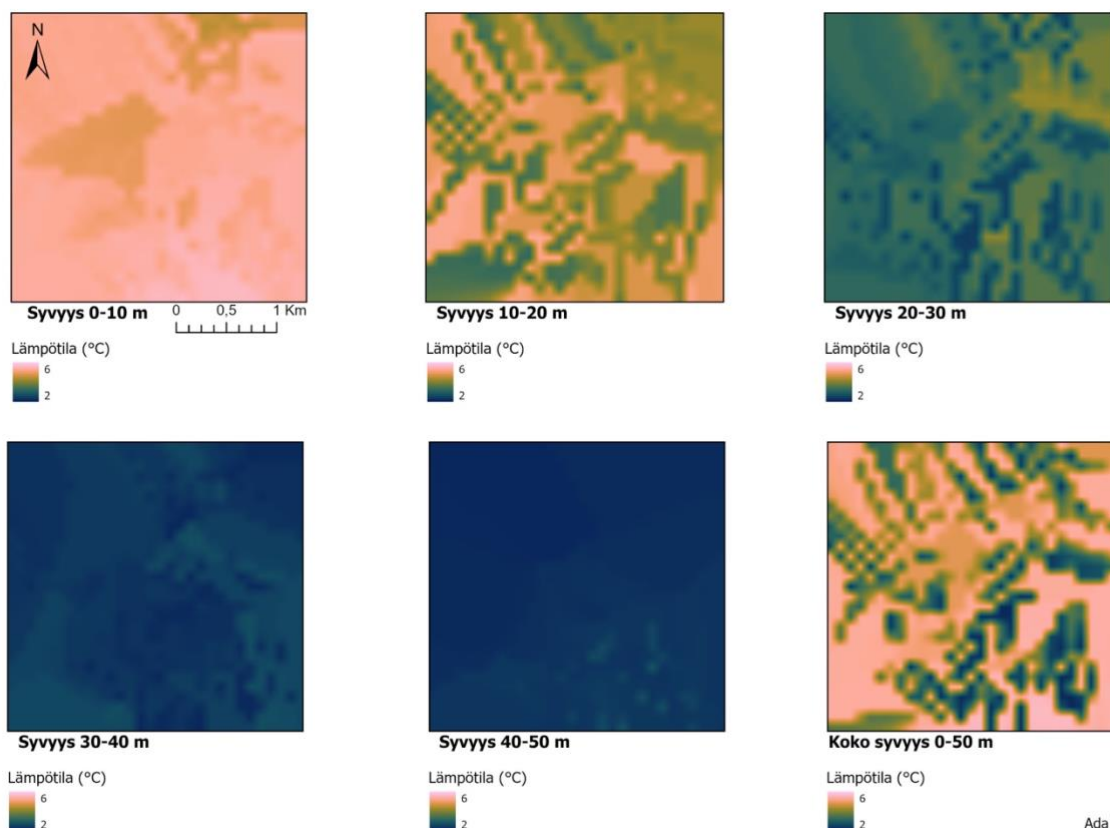
4.3.5 0–50 m

Koko 0–50 m:n kerroksen interpolointi (kuva 3F) yhdistää syvyysrakenteen yhdeksi pinnaksi ja tarjoaa yleiskuvan koko mitatun vesipatsaan tilasta mittauspäivänä. Tämä interpolointirasteri kuvaa pintakerroksen ja syvän veden keskimääräistä rakennetta. Sen tulkintaa voi hyödyntää esimerkiksi silloin, kun pyritään hahmottamaan, kuinka paljon pystysuuntainen kerrostuneisuus heijastuu horisontaaliseen lämpötilarakenteeseen. Suurin osa alueista on 4–5 °C, mikä on odotettua, sillä interpolointi ottaa huomioon voimakkaan lämpötilagradientin, joka oli hyvin havaittavissa syvyyskerroksissa sekä myöhemmässä lämpötilan poikkileikkauksessa. Koko syvyysvälin interpolointikartta toimi erityisesti laajempaan yleiskuvan visualisointina, mikä täydentää rakenteellista tulkintaa.

4.4 Interpolointimenetelmien vertailu

Syvyyskerrosten interpoloinnit tehtiin IDW:n lisäksi myös Krigingillä (kuva 4) ja natural neighbor -menetelmällä (kuva 5). Tarkoituksena oli arvioida, kuinka interpoloinnin

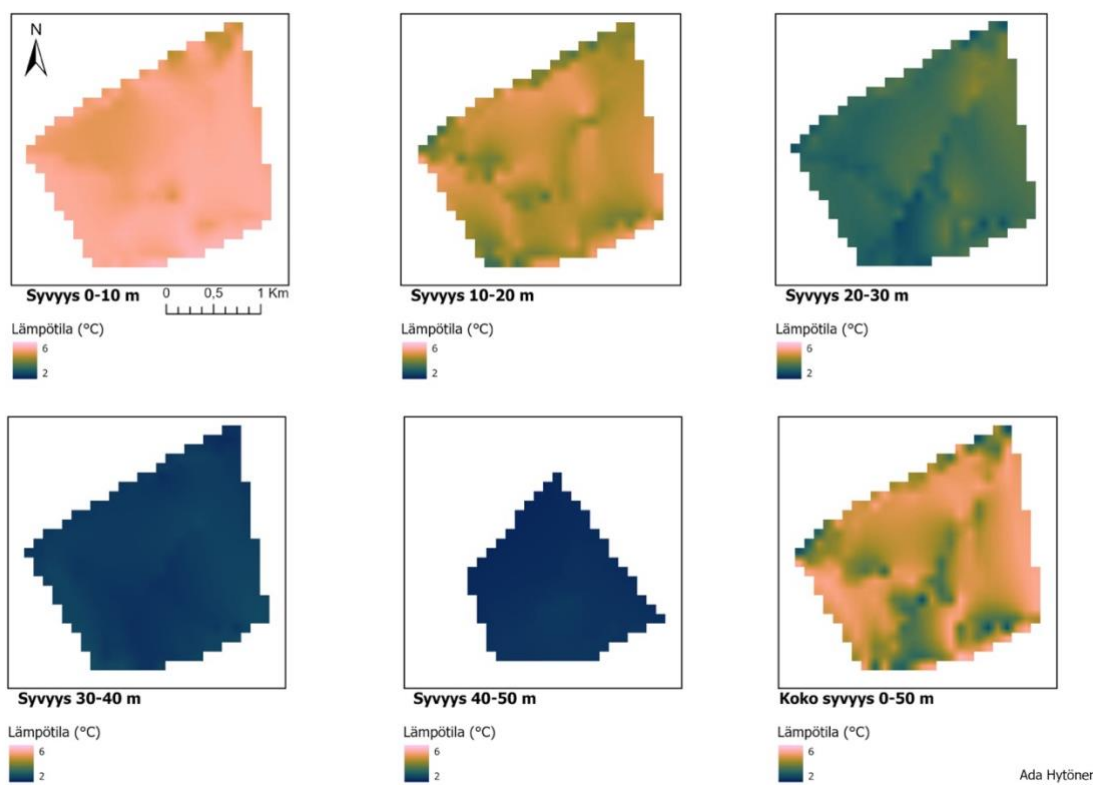
menetelmävalinta vaikuttaa horisontaalisen lämpötilarakenteen tulkintaan sekä sitä, onko käytetyllä menetelmällä merkittävästi vaikutusta.



Ada Hytönen
26.3.2026

Kuva 4. Kriging-interpolointimenetelmän kartat yhtenäisellä 2–6 °C:n väriskaalalla syvyyskerroksittain (0–10 m, 10–20 m, 20–30 m, 30–40 m, 40–50 m) sekä koko mitatulle vesipatsaalle (0–50 m).

Kaikki kolme menetelmää tuottivat rakenteeltaan hyvin samanlaisia interpolointikarttoja. Erityisesti IDW ja Kriging näyttävät lähes identtisiltä. Tämä johtuu aineiston suhteellisen vähäisestä vaihtelusta ja siitä, että lämpötilan gradientti määräytyy pääasiassa syvyyden mukaan. Kriging painottaa tällöin läheisiä havaintoja suhteellisen samankaltaisesti kuin IDW, mikä johtaa visuaalisesti hyvin samanlaiseen lopputulokseen. Natural neighbor -interpolointi tuotti myös rakenteeltaan samansuuntaiset lämpötilakerrokset, mutta niiden pinta oli tasaisempi eikä yhtä pikselimäinen. Menetelmä tuottaa arvot ainoastaan alueelle, jonka sisään mittauspisteet sijoittuivat, minkä vuoksi rasterin ulkoreunat seuraavat pistejoukon muotoa.



Kuva 5. Natural neighbor-interpolointimenetelmän kartat yhtenäisellä 2–6 °C:n väriskaalalla syvyyskerroksittain (0–10 m, 10–20 m, 20–30 m, 30–40 m, 40–50 m) sekä koko mitatulle vesipatsaalle (0–50 m).

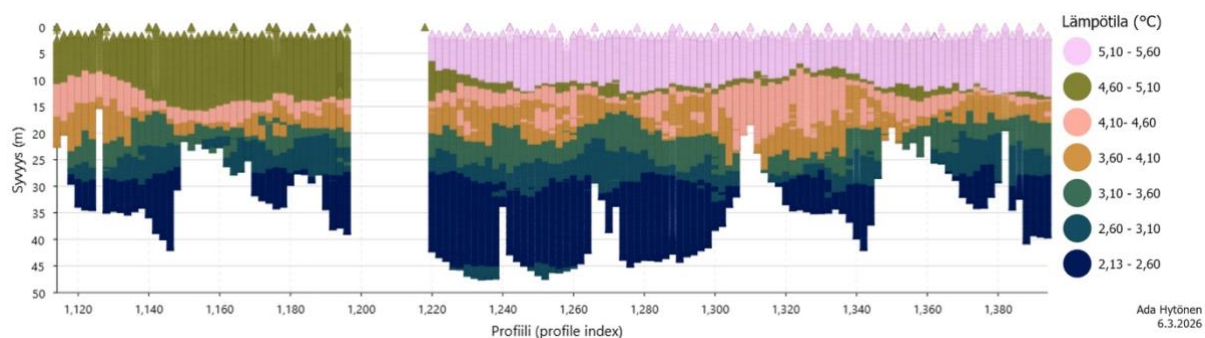
Vertailu tehtiin samoilla syvyysluokilla, interpolointiasetuksilla ja yhtenäisellä väriskaalalla, jotta menetelmien väliset erot perustuivat ainoastaan laskentaperiaatteisiin eikä visualisoinnin asetuksiin. Vertailu osoittaa, että voimakkaimmat horisontaaliset gradientit näkyvät kaikissa interpolointikerroksissa. Liidinaineiston lämpötilarakenteen pääpiirteet eivät siis ole interpolointimenetelmästä riippuvaisia. IDW valittiin interpolointien päämenetelmäksi, koska se tuotti selkeän lämpötilarakenteen, ja menetelmä sopii myös tilanteisiin, joissa alueellinen korrelaatio on heikkoa. Kriging ja natural neighbor toimivat kuitenkin tärkeinä tukimenetelminä, jotka vahvistavat IDW-interpoloinnin luotettavuutta, ja osoittivat, että lämpötilarakenteen keskeiset piirteet ovat samat menetelmävalinnasta riippumatta.

4.5 Lämpötilan vertikaalinen poikkileikkaus

Lämpötilan vertikaalinen poikkileikkaus liitimen reitiltä valittuna päivänä (kuva 6) täydentää interpolointikarttojen tulkintaa, sillä se paljastaa lämpötilan rakenteen pystysuunnassa. X-akseli kuvaa profiilien järjestystä liitimen kulkemalla linjalla. Y-akseli puolestaan kuvaa syvyyttä, ja lisäksi akselin suunta on käännetty tulkittavuuden vuoksi.

Kuvasta erottuu keväällä muodostuva kolmitasoinen lämpötilaprofiili, joka tulkittiin myös interpolointikerroksista. Lämmin pintakerros on noin 0–10 m ja se on lähes koko reitin ajan yhtenäisesti noin 4,5–5,5 °C. Pintakerros on siis hyvin homogeeninen horisontaalisesti. Termokliini sijoittuu 10–20 m:n syvyyteen lämpötilan laskiessa nopeasti. Tämä näkyy jyrkkänä värigradienttina ja kertoo kausitermokliinin muodostumisesta. Näiden alapuolella on kylmä välikerros, jossa vesi on tasaisemmin noin 2–3 °C. Vesimassa on tässä kerroksessa myös suhteellisen homogeeninen koko reitin ajan.

Vertikaalisen poikkileikkauksen ja interpolointikerrosten yhdistäminen osoittaa, että horisontaalinen vaihtelu on vähäisempää kuin pystysuuntainen. Hajontakuvioissa oli myös havaittavissa vuorokausivaihtelua. Aamuyöllä pintavesi oli viileämpää, ja päivän aikana auringon säteily lämmitti pintavettä. Tämä voimisti lämpötilagradienttia pintakerroksen ja välikerroksen välillä. Tämä vastaa tyypillistä Itämeren keväistä termodynamiikkaa, sillä lyhytaikainenkin säteilyvaihtelu voi vaikuttaa nopeasti pintakerroksen lämpötilaan (Aavaste ym. 2021; Cheng ym. 2025). Hajontakuvio on erityisen hyödyllinen pintakerroksen muutosten havaitsemiseen, sillä interpolointi pehmentää nopeita vaihteluita.



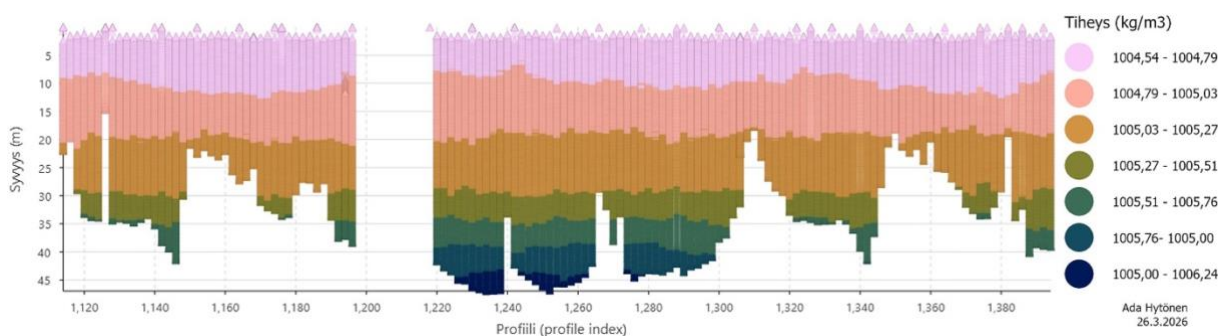
Kuva 6. Lämpötilan vertikaalinen poikkileikkaus liitimen reitiltä 11.5.2024. Poikkileikkauksesta erottuu kolmitasoinen pystyrakenne, jossa on lämpimin pintakerros, termokliini sekä tasalämpöinen syvempi vesi.

4.6 Tiheyden vertikaalinen poikkileikkaus

Tiheyden vertikaalinen poikkileikkaus liitimen reitiltä valittuna päivänä (kuva 7) paljastaa fysikaalisen rakenteen, joka puolestaan täydentää lämpötilatarkastelua. Tiheyden vaihteluväli $1004\text{--}1006\text{ kg/m}^3$ on tyypillinen Itämerelle, jossa meriveden tiheys on keskiarvolta noin 1005 kg/m^3 (Luoma & Okkonen 2014). Suomenlahdella tiheys

vaihtelee tyypillisesti 1001–1006 kg/m³ välillä. Tiheysvaihtelut ovat pieniä mutta kerrostuneisuuden kannalta merkittäviä.

Kuva osoittaa pinnalla (0–10 m) olevan kevyen kerroksen, jossa tiheys pysyy koko matkan ajan hyvin tasaisena. Lämmin pintavesi on siis kevyempää. Tämän jälkeen tiheys kasvaa 10–20 m:n syvyydessä, mikä vastaa myös lämpötilan poikkileikkauksessa havaittua termokliinia. Syvemmällä vesi on tiheämpää ja tiheys kasvaa tasaisesti alaspäin mentäessä.

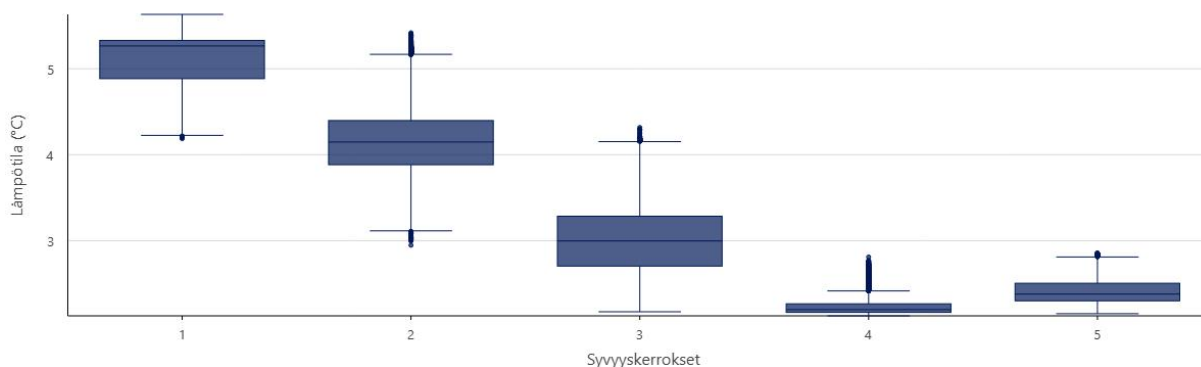


Kuva 7. Tiheyden vertikaalinen poikkileikkaus liitimen reitiltä 11.5.2024. Kevyt pintavesi erottuu vaaleimpana, minkä jälkeen tiheys kasvaa kohti syvempiä kerroksia.

4.7 Lämpötilan ja tiheyden vaihtelu syvyyskerroksittain

Laatikkojanakuviot kvantifioivat lämpötilan ja tiheyden vaihtelua viidessä syvyysluokassa: 0–10 m, 10–20 m, 20–30 m, 30–40 m, 40–50 m. Nämä kuvat täydentävät interpolointikarttojen sekä poikkileikkausten visuaalisia havaintoja tarjoamalla tilastollisen näkökulman vesipatsaan rakenteellisiin eroihin.

Lämpötilan laatikkojanakuva (kuva 8) osoittaa myös vesipatsaan kolmitasoisen rakenteen, joka havaittiin jo aikaisemmin. Pintakerros on lämpimin ja tasalaatuinen, sillä mediaani on 5,27 °C ja IQR (kvartiiliväli) 0,44 °C (taulukko 2). Havaintojen vaihteluväli on toisessa kerroksessa (10–20 m) 0,51 °C, mikä osoittaa termokliinin paikkaa. IQR on kuitenkin hieman suurempi (0,58 °C) kolmannessa syvyyskerroksessa, mutta tämä johtuu termokliinin alla olevan veden ja tasaisen viileän vesimassan sekoittumisesta. Syvyyskerroksessa on siis kaksi erilaista vesimassaa, mikä laajentaa kvartiiliväliä. Mediaani laskee tässä kerroksessa jo selkeästi (2,99 °C). Kahdessa mittauksen syvimässä kerroksessa mediaanit ovat 2,20–2,38 °C eli vesi on jo kylmää ja lähes tasalämpöistä. Vaihtelu on siis hyvin vähäistä.



Kuva 8. Lämpötilan jakauma syvyyskerroksittain 11.5.2024. Laatikkojanakuva näyttää lämpötilan jakaumat 10 m:n syvyyskerroksissa 0–50 m välillä. Pintakerros on lämmin ja toinen syvyyskerros havainnollistaa termokliinin gradientin. Syvin vesi on kylmää mutta stabiilia, mikä ilmenee pienenä vaihteluna ja kapeina laatikkoina.

Taulukko 2. Lämpötilan tunnusluvut syvyyskerroksittain 11.5.2024 (°C).

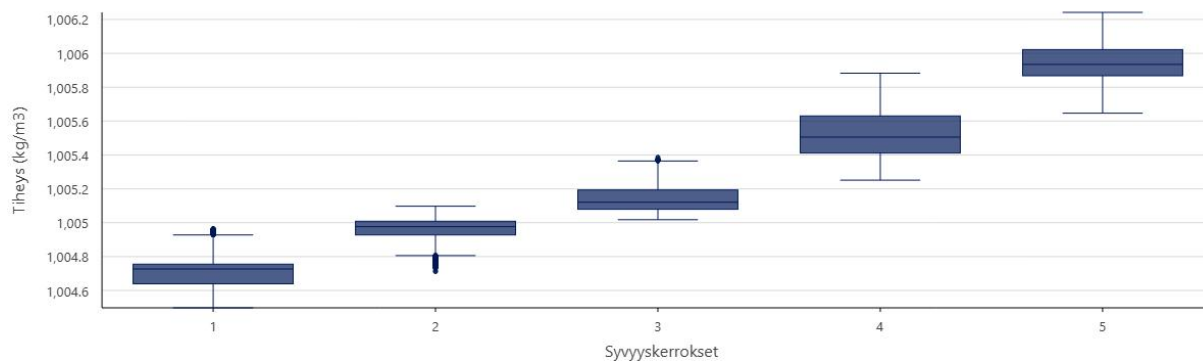
Lämpötila esitetään celsiusasteina. 1. kvartiili ja 3. kvartiili kuvaavat havaintojen jakauman 25 % ja 75 % pisteitä. IQR (Interquartile Range) on kvartiiliväli, joka kuvaa syvyyskerrosten sisäistä vaihtelua.

Syvyyskerrokset	Minimi (°C)	1. kvartiili (°C)	Mediaani (°C)	3. kvartiili (°C)	Maksimi (°C)	IQR (°C)	Keskiarvo (°C)
1 (0–10 m)	4,19	4,89	5,27	5,33	5,63	0,44	5,15
2 (10–20 m)	2,95	3,88	4,15	4,40	5,42	0,51	4,18
3 (20–30 m)	2,17	2,70	3,00	3,28	4,32	0,58	3,02
4 (30–40 m)	2,12	2,17	2,20	2,27	2,81	0,10	2,24
5 (40–50 m)	2,15	2,30	2,38	2,50	2,85	0,20	2,42

Suurin fysikaalinen lämpötilagradientti on siis syvyyskerroksessa 2 (10–20 m), joka edustaa termokliinia. Syvyyskerroksen 3 (20–30 m) suurempi kvartiiliväli johtuu kahden eri vesimassan sekoittumisesta termokliinin alapuolella.

Tiheyden laatikkojanakuva (kuva 9) näyttää rakenteen olevan samankaltainen lämpötilan kanssa, mikä vahvistaa kerrostuneisuuden fysikaalista tulkintaa.

Syvyyskerros 1 sisältää kevyimmän vesimassan (mediaani $1004,73 \text{ kg/m}^3$) (taulukko 3). Toisessa kerroksessa mediaani on korkeampi ($1004,98 \text{ kg/m}^3$), eli tiheys nousee, mutta vaihtelu pysyy kohtuullisena (IQR $0,081 \text{ kg/m}^3$). Seuraavissa kerroksissa (20–30 m ja 30–40 m) tiheys kasvaa koko ajan, ja syvimässä kerroksessa mediaani on $1005,93 \text{ kg/m}^3$ eli vesi on selkeästi tiheintä. Veden tiheys nousee siis koko ajan tasaisesti ja syvemmillä muodostuu homogeeninen ja tiheä kerros.



Kuva 9. Tiheyden jakauma syvyyskerroksittain 11.5.2024. Laatikkojanakuvio näyttää tiheyden kasvun syvyyttä kohti. Pintavesi on kevyttä, ja tiheämpi syvä vesi muodostavat kerrostuneisuuden.

Taulukko 3. Tiheyden tunnusluvut syvyyskerroksittain 11.5.2024 (kg/m^3).

Tiheyden yksikkö on kilogrammaa kuutiometriä kohden. 1. kvartiili ja 3. kvartiili kuvaavat havaintojen jakauman 25 % ja 75 % pisteitä. IQR (Interquartile Range) on kvartiiliväli, joka kuvaa syvyyskerrosten sisäistä vaihtelua.

Syvyyskerrokset	Minimi (kg/m^3)	1. kvartiili (kg/m^3)	Mediaani (kg/m^3)	3. kvartiili (kg/m^3)	Maksimi (kg/m^3)	IQR (kg/m^3)	Keskiarvo (kg/m^3)
1 (0–10 m)	1004,50	1004,64	1004,73	1004,76	1004,97	0,12	1004,70
2 (10–20 m)	1004,71	1004,93	1004,98	1005,01	1005,10	0,08	1004,95
3 (20–30 m)	1005,02	1005,08	1005,12	1005,19	1005,39	0,11	1005,14
4 (30–40 m)	1005,25	1005,41	1005,51	1005,63	1005,88	0,22	1005,53
5 (40–50 m)	1005,64	1005,87	1005,93	1006,02	1006,24	0,15	1005,95

5 Keskustelu

5.1 Menetelmävalinnat

Käytetty menetelmäkokonaisuus koostui pisteaineistosta, interpoloinneista, hajontakuviosta ja laatikkojanakuviosta. Se muodostettiin tukemaan geovisualisoinnin ja visuaalisen analytiikan (visual analytics) periaatteita. Yleiskuva tuotettiin interpolointikartoilla, jotka esittävät syvyyskerrosten horisontaalisen vaihtelun yhtenäisellä ja selkeällä väriskaalalla (Shneiderman 1996). Interpolointikarttoja täydennettiin ja tarkennettiin käyttäen hajontakuviota pystysuuntaisen rakenteen esittämiseen. Yksityiskohtaisempi ja tarkoituksenmukainen visualisointirakenne vastasivat Munznerin (2014) What-Why-How mallia. Käytetty kaksitasoinen lähestymistapa vastaa Andrienkon ja Andrienkon (2006) kuvailemia synoptisia ja elementaarisia tehtäviä. Menetelmävalinnat sovitettiin datan rakenteeseen, minkä vuoksi pistepohjaiset hajontakuviot valittiin ensisijaiseksi tavaksi esittää pystysuuntainen rakenne. Laatikkojanakuviot täydentävät jo käytettyjä menetelmiä, sillä ne antavat tilastollisen näkymän kerrosten sisäiseen vaihteluun. Yhdistettynä interpolointikartat, hajontakuviot ja laatikkojanakuviot luovat menetelmäkokonaisuuden, joka on toistettavissa sekä näyttää horisontaalisen ja pystysuuntaisen vaihtelun moniulotteisesta liidinaineistosta.

Moniulotteisen liidinaineiston tulkinta edellyttää useamman ulottuvuuden, kuten sijainnin, syvyyden ja ajan tarkastelua samanaikaisesti. Kelleher ja Braswell (2021) korostavat, että moniulotteinen ympäristödata edellyttää menetelmiä, joiden avulla voidaan tunnistaa rakenteita, jotka eivät ilmene yhdestä näkymästä. Synoptinen ja elementaarinen tarkastelu osoittautuivat välttämättömiksi, mikä vahvistaa Andrienkojen (2006) tehtäväluokituksen soveltuvuuden moniulotteisen datan tulkintaan.

Visualisointiprosessien eteneminen noudatti selvästi Shneidermanin (1996) esittämää ”overview first, zoom and filter, details on demand” -mallia. Interpolointikartat toimivat yleiskuvana, syvyyskerrokset vastasivat rajauksia ja hajontakuviot mahdollistivat yksityiskohtiin keskittymisen. Visualisointimenetelmien kokonaisuus vastaa myös

Munznerin (2014) What-Why-How-mallin mukaista analyysiprosessia, jossa näkymän valinta perustuu tehtävän luonteeseen.

Värikartan vaikutus tulkintaan oli huomattavissa. Perseptuaalisesti yhtenäinen batlow-värikartta paransi rakenteiden luettavuutta ja vähensi virhetulkinnan riskiä, mikä on myös linjassa Cramerin ym. (2020) tekemien havaintojen kanssa. Sateenkaari-värikartan kaltaiset epälineaariset paletit saattavat peittää todellisia gradientteja tai korostaa vähemmän oleellisia kontrasteja. Valittu väripaletti osoittautui erittäin toimivaksi, koska horisontaaliset ja pystysuuntaiset lämpötilan muutokset olivat hienovaraisia. Nyt rakenteiden tulkittavuus on myös nähtävissä mustavalkotulosteissa.

Havaittu kolmitasoista lämpötilarakennetta muistuttava kerrostuneisuus vastasi kerrostuneisuuden fysikaalista mallia, jossa on lämmin pintakerros, termokliini ja kylmä välikerros (Aavaste 2021). Nämä heijastivat myös keväistä tilannetta, jossa pintavesi lämpenee nopeasti mutta syvämpi vesi pysyy lähes saman lämpöisenä. Visualisoinnit eivät ainoastaan kuvanneet aineistoa, vaan tuottivat myös fysikaalisesti realistisen näkymän meriveden rakenteesta.

5.2 Menetelmien rajoitukset

5.2.1 Interpoloinnit

IDW-interpolointi on havainnollistava ja helppokäyttöinen, mutta se tuottaa kuitenkin aina arvion eikä suoraa mittaustulosta jokaiselle rasterisolulle (Esri s.a. c). Luodut interpolointikartat ovat siis parhaimmillaan hyvä informatiivinen arvio, ei jokaisen pisteen fysikaalinen mittaustulos. Interpoloidut arvot pysyvät aina havaittujen minimi- ja maksimiarvojen välissä, minkä vuoksi todelliset gradientit voivat pehmentyä. IDW on myös altis niin sanotulle bull's-eye -ilmiölle, jossa yksittäisen pisteen lähiympäristöön piirtyy rengasmaisen korostus. Tätä tapahtuu etenkin silloin kun havaintotiheys on epätasainen tai piste poikkeaa selvästi ympäröivistään pisteistä (GITTA 2016).

IDW valittiin interpolointien päämenetelmäksi sen selkeyden ja suorituskyvyn vuoksi, mutta tuloksia täydennettiin kuitenkin vertaamalla niitä myös Kriging- ja natural neighbor -interpolointeihin. Lisäksi yhtenäinen sekä vertailtava visualisointitapa tuotti nopeasti tulkittavan ja toistettavan yleiskuvan merivedenlämpötilan vaihtelun

pääpiirteistä. Menetelmä on myös ympäristötieteissä laajasti käytetty (Li & Heap 2011). Lisäksi yhtenäisen väriskaalan (2–6 °C) käyttö paransi kerrosten välistä vertailtavuutta, mikä on suositeltavaa suurten ympäristödatojen visualisoinnissa (Crameri ym. 2020; Kelleher & Braswell 2021).

5.2.2 Hajontakuviot

Hajontakuviot näyttävät pystysuuntaisen rakenteen suoraan pisteaineistosta, jolloin interpolointia ei tarvita. Tämä vähentää mallinnusvirheitä. Pisteiden epäsäännöllisyys ja jakautuminen voivat kuitenkin aiheuttaa joidenkin syvyyksien yliedustusta tai aliedustusta. On myös hyvä huomioida, ettei pisteiden x-akselin järjestys (`profile_index`) vastaa tarkkaa horisontaalista sijaintia.

Liidindatalle tyypillinen epäsäännöllinen näytteenotto näkyi selvästi visualisointien välillä. Kuten von Oppeln-Bronikowski ym. (2021) ja GRIIDC (2023) kuvaavat, liitimen jatkuva liike ylös ja alas tuottaa tiheän mutta epäsäännöllisen profiilisarjan, joka soveltuu huonosti tasavälisiin rasteriesityksiin. Tämä havaittiin myös tässä tutkielmassa, sillä hajontakuviot paljasti pystysuuntaisen rakenteen luotettavammin kuin interpolointikartat, jotka tasoittavat epäsäännöllistä rakennetta. Kuten oletettiin, aineiston luonne ohjasi merkittävästi menetelmävalintoja.

ArcGIS Pron Scatter Plot -työkalu ei tue väripalkin ja selitteen näyttämistä, joten sekä lämpötilan että tiheyden väripalkit jouduttiin lisäämään kartta-asetelmassa manuaalisesti. Skaala olisi mahdollista kirjata esimerkiksi legendan otsikkoon tai kuvaukseen, mutta tämä heikentäisi huomattavasti etenkin välitöntä tulkittavuutta (Munzner 2014; Esri s.a. a). Tulkittavuuteen vaikuttaa lisäksi kuvion muut visuaaliset valinnat kuten symbolin muoto ja koko. Muun muassa ääriviivaton symboli vähentää merkittävästi päällekkäisyyksien aiheuttamaa tummumista.

5.2.3 Laatikkojanakuviot

Vaikka laatikkojanakuviot tuottavat arvokasta kvantitatiivista tietoa syvyyskerrosten jakaumista, niillä on kuitenkin rajoitteita. Kuviot visualisoivat ainoastaan jakauman perustunnuslukuja, eli mediaania, kvartiileja, minimiä ja maksimia, eikä niistä näy profiilien sisäinen vaihtelu tai horisontaalinen sijainti. Näiden lisäksi kvartiileihin

perustuvat tunnusluvut voivat liioitella tai heikentää muutoksia kahden vesimassan yhdistyttyä yhden syvyysluokan sisällä, kuten syvyyskerroksessa 3. Tämä saattaa kasvattaa kvartiiliväliä eikä se kuvaa todellista fysikaalista vaihtelua. Laatikkojanakuviot ovat siis hyödyllisiä kvantitatiiviseen arviointiin, mikä vastaa Clevelandin ja McGillin (1984) näkemystä siitä, että sijainti on tarkoin määrällisen tiedon visuaalisen esittämisen muoto. Mutta ne eivät kuitenkaan korvaa poikkileikkausta kuvaavia hajontakuviota pystysuuntaisen rakenteen tulkinnessa.

5.2.4 Liidin ja aineisto

Liidin toimii CTD-luotaimen mittausalustana, ja mittausten perustan muodostavat lämpötilan, johtavuuden sekä paineen mittaus (Ilmatieteen laitos 2022). Mittalaitteisiin liittyy kuitenkin rajoitteita, kuten ylläpito ja mahdollinen rikkoutuminen. Liitimessä on usein mukana optisia LISST-sensoreita kuten läpäisymittari ja hiukkaskoon analysaattori (Sequoia 2021; Sequoia 2022). Laitteet ovat itsessään herkkiä, ja iskut tai huono käsittely voivat herkästi vaikuttaa laitteen mittaukseen negatiivisesti. Ympäristötekijöillä kuten saastuneella vedellä voi myös olla vaikutusta.

Käytetty aineisto on CF-konventioihin sopiva (CF 1.6), mikä mahdollistaa standardoidun tallennuksen, mutta silläkin on rajoitteita. Liitimen keräämä aineisto kuuluu CF-konventioiden mukaiseen DSG (Discrete Sampling Geometry) -malliin (Esri s.a. k), eikä se muodosta säännöllistä hilarakennetta. Tämän vuoksi aineistoa ei voi visualisoida suoraan ArcGIS Prossa voxel-muotoisena tilavuusesityksenä (Esri s.a. m).

Vaikka ArcGIS Pro tarjoaa työkalut voxel-visualisoinnin, tilavuusesitys vaatii säännöllistä 3D-hilarakennetta (Esri s.a. e; Esri s.a. l). Tässä tutkielmassa käytetty liidinaineisto muodostaa kuitenkin epäsäännöllisen liikeradan (trajectory), jossa havainnot tallennetaan aikasarjana. Aineiston muuttaminen voxel-esitykseen sopivaksi vaatisi merkittävästi työtä tutkielman laajuuteen nähden. Voxel-pohjaisen tilavuusesityksen visualisoinnista luovuttiin siten kokonaan. Tämä korostaa sitä, että kaikki ohjelmiston tarjoamat visualisointimenetelmät eivät ole suoraan sovellettavissa epäsäännölliseen aineistoon ilman merkittävää esikäsittelyä. Menetelmätoteutuksessa täytyi siis huomioida epäsäännöllisen datarakenteen aiheuttamat rajoitteet.

5.3 Toistettavuus ja käytettävyys

Erityisesti visualisointimenetelmien toistettavuus ja käytettävyys ovat tärkeitä arviointikriteereitä, koska tavoitteena oli rakentaa selkeä menetelmäkokonaisuus. Tulkittavuus ja käytettävyys kytkeytyvät siihen, miten hyvin käyttäjä pystyy hahmottamaan aineiston rakenteen, suorittamaan tehtävät ja saavuttamaan tavoitteet tehokkaasti ja hallittavasti (Bevan 1995; Shneiderman 1996; Andrienko & Andrienko 2006; ISO 2018).

Tässä tutkielmassa toistettavuus varmistettiin dokumentoimalla jokainen visualisointivaihe yksityiskohtaisesti (liite 2). ArcGIS Pro:n käyttö tukee toistettavuutta, sillä se pystyy lukemaan suoraan CF-konventioiden mukaista NetCDF-formaattia (Esri s.a. k). Aineisto noudatti CF 1.6-standardia eli aineiston muuttujat, dimensiot ja koordinaatit tulkittiin automaattisesti, mikä vähensi virheiden riskiä (Eaton ym. 2025). Tämä tukee myös tehokkuutta, sillä käyttäjä voi keskittyä ensisijaisesti tulkintaan eikä rakenteen selvittämiseen.

Käytettävyyden kannalta ArcGIS Pro toimi tarkoituksenmukaisesti, kun tavoitteena oli yleiskuvan muodostus, helppo raja-alue syvyyskerroksiin sekä yksityiskohtainen pystysuuntaisen rakenteen tarkastelu. Käytettävyys toteutui tässä mielessä ISO 9241-11-standardin mukaisesti, sillä työkalun avulla saatiin toteutettua määritellyt tehtävät niin, että tuloksellisuus oli riittävä ja työmäärä kohtuullinen (ISO 2018). Lisäksi tämä tukee ISO-standardin ajatusta, jonka mukaan käytettävyys syntyy järjestelmän, tehtävän ja käyttäjän yhteensopivuudesta.

ArcGIS Pro -työskentelyssä havaittiin merkittäviä suorituskykyyn liittyviä rajoitteita, jotka olivat kuitenkin ratkaistavissa. Suuri NetCDF-aineisto kasvatti latausaikoja ja hidasti karttatasojen piirtymistä. Kelleher ja Braswell (2021) korostavat, että suurien ympäristödatojen visualisoinnissa laskennalliset rajoitteet ja järjestelmän suorituskyky vaikuttavat keskeisesti analyysin sujuvuuteen. Raskaat prosessit voivat selkeästi heikentää mahdollisuuksia tarkastella kokonaisuutta. Tämä korostaa tarvetta esimerkiksi datan pilkkomiselle, jota tässäkin tutkielmassa tehtiin. Ilman aineiston ajallista pilkkomista käsittely olisi ollut käytännössä mahdotonta, sillä ArcGIS Pro ei pystynyt lataamaan koko aineistoa kerralla ilman kaatumista. Jo alle kymmenen

muuttujan lataaminen aiheutti viivettä ja kaatumisia, mikä osoittaa, että NetCDF-aineiston koko jo moniulotteisuus ylittävät ohjelman käytettävissä olevan muistinkäsittelyn. Tämän vuoksi liidinaineisto jaettiin yksittäisiin päiviin, mikä vähensi kuormitusta merkittävästi ja mahdollisti sujuvan analyysin.

Perseptuaalisesti yhtenäinen batlow-värikartta tuotti selkeän ja yhdenmukaisen väriskaalan, mikä vähensi tulkintavirhettä sekä tuki rakenteen synoptista hahmottamista (Andrienko & Andrienko 2006; Crameri ym. 2020). Elementaarisissa tehtävissä, kuten yksittäisten pisteiden tarkastelussa, visuaalisen erottelun tarkkuus perustui datapisteiden sijaintiin akselilla yhteisellä asteikolla (Cleveland & McGill 1984). Tämä toteutui hajontakuvioissa hyvin, sillä tulokset esitettiin suorakulmaisella akselilla. Symbolin koon muuttaminen ja ääriviivojen poistaminen paransi myös tehokkuutta vähentämällä visuaalista ylikuormitusta.

Big data -näkökulma esiintyi tutkielmassa erityisesti prosessointinopeuden ja datan volyymin kohdalla. NetCDF-aineiston koko vaikutti latausaikoihin ja viiveisiin. Tämä tukee Kelleherin ja Braswellin (2021) tutkimusta siitä, että järjestelmän kuormitus ympäristödatassa on keskeinen osa käytettävyyttä. Lisäksi datan vaihtelu ja luotettavuus korostuivat, sillä epäsäännöllinen pistejakauma ja laatuvaihtelut näkyivät hajontakuvioissa. Tämä osoittaa 4V-mallin luovan konkreettiset reunaehdot.

Myös menetelmien laskennallinen raskaus vaikuttaa käytettävyyteen. Tutkielmassa käytetyt menetelmät olivat laskennallisesti suhteellisen kevyitä datan esikäsittelyn jälkeen. Kuitenkin esimerkiksi tilavuusesitykset olisivat lisänneet työvaiheita ja vaatineet aikaa.

5.4 Jatkokehitys

Menetelmäkokonaisuus soveltui hyvin kandidaatintutkielman laajuuteen, mutta jatkokehitys olisi silti mahdollista. Koska aineisto sisältää jatkuvia profiileja kymmeneltä päivältä, laajempi ajallinen tarkastelu voisi näyttää eri säätilanteiden tuomat vaikutukset termokliiniin. Tämän voisi toteuttaa esimerkiksi animaatiolla. Tällaisten pitkittäisanalyysien mahdollisuus mainitaan aiemmissa AUV-tutkimuksessa (Testor ym.

2019). Merentutkimuksessa on myös yleistynyt Hovmöller-diagrammit, jotka ovat kartan ja graafin yhdistelmä (Heuzé ym. 2015; Di Liberto 2021).

Jatkokehityksenä voisi tarkastella myös vielä useamman interpolointimenetelmän hyödyntämistä. Merellisen datan interpolointi AUV-havainnoista tarjoaa keinon hahmottaa merien tilaa, vaikka pisteet ovat epäsäännöllisesti jakautuneita. Tässä tutkielmassa käytetty Kriging-interpolointi on tilastollisesti edistyneempi interpolointimenetelmä, sillä se huomioi pisteiden välisen alueellisen autokorrelaation (Cressie 1993; Munyati & Sinthumule 2021). Mahdollinen vaihtoehto jatkotarkastelulle on myös LOESS-menetelmä (locally weighted least squares), sillä se pohjautuu regressiomalleihin, joissa arviot perustuvat läheisiin havaintoihin. Merkittävä esimerkki tästä on CSIRO (Commonwealth Scientific and Industrial Research Organisation) Atlas of Regional Seas (CARS), jossa epäsäännöllisesti havaittu merellinen data interpoloitiin tasaväliseen hilarakenteeseen LOESS-suodatuksen avulla (Ridgway ym. 2002). Menetelmässä painotukset skaalautuvat automaattisesti datapisteiden tiheyden mukaan. Liidinaineistoon sovellettuna LOESS voisi toimia ratkaisuna, kun halutaan tuottaa tasavälisiä 3D-pintoja. Tämä voisi mahdollistaa myös tasavälisten tilavuusesitysten tuottamisen.

Visualisointeja voisi laajentaa jatkokehityksessä fysikaalisten muuttujien ohella myös kemiallisiin ja optisiin suureisiin. Käytetyssä NetCDF-aineistossa on mitattu muun muassa klorofyllipitoisuus, fluoresoiva liukoinen orgaaninen aine, valon takaisinsironta 700 nm aallonpituudella, näkyvyys ja liuenneen hapen pitoisuus. Näiden suureiden avulla olisi mahdollista saada lisätietoa veden ominaisuuksista. Nämä suureet kuvaavat muun muassa hiukkasten määrää vedessä, planktonin runsautta sekä liuenneiden aineiden vaikutusta valon kulkuun, minkä vuoksi ne ovat laajasti käytettyjä meriekosysteemien tarkastelussa (Dickey & Chang 2002). Esimerkiksi valon takaisinsironta on yhteydessä vedessä olevien hiukkasten määrään ja kokojakaumaan (Reynolds ym. 2016). LISST-laitteisiin perustuvat mittaukset voivat paljastaa sameuskerroksia ja hiukkaspitoisuuden ajallista vaihtelua (Sequoia 2021).

Mahdollisena jatkokehityksenä voisi olla datassa esiintyvän mittausvarmuuden ja laadunvalvonnan(QC) tarkastelu. Sekä CTD-luotain että optiset sensorit sisältävät

mahdollisia virhelähteitä, kuten optisen signaalin kohina tai GPS-paikannuksen virheet, joita esiintyi myös tässä tutkielmassa käytetyssä aineistossa (GRIIDC 2023).

Visualisointeihin voisi liittää esimerkiksi poikkeavien havaintojen korostuksen laatikkojanakuvoissa tai interpolointimenetelmien tuottamia virhearvioita.

Epävarmuuden visualisointia on käsitelty ympäristödatan tutkimuksessa, ja muun muassa Kelleher ja Braswell (2021) korostavat, että virhe- ja epävarmuustietojen havainnollistaminen parantaa tulkittavuutta. Epävarmuuksien havainnollistaminen olisi luonteva jatkokehityksen kannalta.

Viides jatkokehitysmahdollisuus liittyy dynaamisiin analyysijärjestelmiin. Tuore tutkimus esittelee uuden datapohjaisen dynaamisen sovellusjärjestelmän (Dynamic Data Driven Application Systems, DDDAS), joka on tarkoitettu alueellisten kolmiulotteisten meriympäristöjen reaaliaikaiseen ennustamiseen (Sang ym. 2026). Se perustuu autonomisten vedenalaisten liidinten havaintoihin. Viitekehys mahdollistaa mukautuvan ja tarkan ympäristöennusteen. Mallinnusmenetelmää kehitetään yhdistämällä dynaaminen moodihajotelma (Dynamic Mode Decomposition, DMD) sekä Gaussin prosessiregressio (Gaussian Process Regression, GRP), jolloin liidinaineiston epäsäännöllinen eli hilton näytteenotto voidaan sovittaa dynaamiseen ennusteeseen.

6 Johtopäätökset

Tämän tutkielman tavoitteena oli vertailla ArcGIS Pron tarjoamia visualisointitapoja moniulotteisen liidinaineiston esittämiseen sekä arvioida menetelmien tulkittavuutta ja käytettävyyttä. Aineistona käytettiin Ilmatieteen laitoksen keräämää, muun muassa veden lämpötilaa ja tiheyttä kuvaavaa NetCDF-aineistoa. Visualisointimenetelmät rakennettiin tukemaan geovisualisoinnin periaatteita, erityisesti Shneidermanin (1996) ”overview first, zoom and filter, then details on demand” -periaatetta sekä Andrienkon & Andrienkon (2006) tehtäväluokitusta elementaarisiiin ja synoptisiin tehtäviin.

Tulokset osoittivat interpolointikarttojen toimivan tehokkaana yleiskuvan muodostamisen menetelmänä, mikä tekee syvyyskerrosten horisontaalisen rakenteen helposti ja nopeasti hahmotettavaksi. Menetelmän rajoitteet painottavat kuitenkin sen käyttöä ensisijaisesti yleiskuvan muodostamiseen eikä yksittäisten arvojen tulkintaan. Interpolointikarttojen vahvuutena oli hyvä visuaalinen vertailtavuus, mihin vaikutti yhtenäisen väriskaalan käyttö.

Hajontakuviot osoittautuivat välttämättömiksi syvyysrakenteen tarkastelussa. Aineisto ei muodosta säännöllistä hilarakennetta, jolloin pistepohjainen esitystapa mahdollistaa pystysuuntaisen rakenteen havaitsemisen ilman interpoloinnin tuomia vääristymiä. Rajoitteena on kuitenkin x-akselin epäsuora suhde tarkkaan horisontaaliseen sijaintiin sekä lisäksi ArcGIS Pron väriskaalarajoitukset. Hajontakuviot tarjosivat silti selkeimmän näkymän kerrostuneisuudesta.

Lämpötilan ja tiheyden horisontaalinen ja pystysuuntainen vaihtelu ilmeni visualisoinneissa hyvin. Interpolointi toi esiin pintakerroksen suhteellisen tasaisen lämpötilarakenteen sekä kylmän välikerroksen. Pystysuuntainen vaihtelu ja kerrostuneisuus näkyivät parhaiten hajontakuvioissa, joissa termokliini ja tiheyden tasainen nousu erottuivat, kuitenkin ilman interpoloinnin vääristymiä.

Laatikkojanakuviot vahvistivat havainnot tilastollisten lukujen avulla. Näiden yhdistelmä osoitti, että pystysuuntainen rakenne on Tvärminnen merialueella merkittävämpi kuin horisontaalinen vaihtelu.

Laatikkojanakuviot täydensivät aikaisempia menetelmiä tarjoamalla tilastollisen näkymän syvyyskerrosten välisiin eroihin. Ne eivät siis yksinään sovellu rakenteen tulkintaan. Yhdessä käytettynä menetelmät muodostivat kuitenkin kokonaisuuden, joka mahdollistaa sekä synoptisen että elementaarisen tarkastelun.

Käytettävyyden näkökulmasta ArcGIS Pro osoittautui toimivaksi työkaluksi aineiston kevennyksen ja päiväkohtaisen jaon jälkeen. Koska ohjelmisto tukee NetCDF-formaattia suoraan, toteutus oli melko suoraviivaista. Toistettavuus varmistettiin dokumentoimalla työvaiheet ja käyttämällä standardoituja värikarttoja, mikä mahdollistaa menetelmien soveltamisen muihin liidinaineistoihin.

Tutkielma osoitti, että ArcGIS Pro soveltuu hyvin liidinaineiston visualisointiin silloin, kun visualisointimenetelmät valitaan datan rakenteen mukaan ja huomioidaan epäsäännöllisen näytteenoton vaikutukset. IDW-kartat, hajontakuviot ja laatikkojanakuviot muodostavat yhdessä toimivan menetelmäkokonaisuuden. Rajoitteet osoittavat, että menetelmien jatkokehys on perusteltua. Mahdollisia kehityssuuntia ovat esimerkiksi optisten sekä kemiallisten muuttujien laajempi hyödyntäminen.

Kiitokset

Suuret kiitokset Sari Penttiselle Puolustusvoimista, jolta sain tämän kandidaatintutkielman aiheen sekä tukea erityisesti työn alkuvaiheessa.

Kiitokset kuuluvat myös Ilmatieteen laitokselle ja erityisesti Lauri Laaksolle, jonka tutkimusryhmän keräämä liidinaineisto mahdollisti tämän työn toteuttamisen.

Lämmin kiitos ohjaajalleni Harri Tolvaselle ohjauksesta ja rakentavista kommentteista.

Lisäksi haluan kiittää Aleksanteri Nissistä avusta Python-työskentelyssä ja ohjelmiston käytössä sekä kommentteista ja tuesta.

Lähteet

- Aavaste, A., Sipelgas, L., Uiboupin, R. & Uudeberg, K. (2021) Impact of thermohaline conditions on vertical variability of optical properties in the Gulf of Finland (Baltic Sea): Implications for water quality remote sensing. *Frontiers in Marine Science* 8 674065. <https://doi.org/10.3389/fmars.2021.674065>
- Alenius, P., Myrberg, K. & Nekrasov, A. (1998) The physical oceanography of the Gulf of Finland: a review. *Boreal Environment Research* 3 97–125. <https://helda.helsinki.fi/server/api/core/bitstreams/3132c831-966b-4e4b-875c-eae19145d630/content> 30.3.2026.
- Andrienko, N. & Andrienko, G. (2006) *Exploratory analysis of spatial and temporal data: A systematic approach*. Springer, Berlin & New York. <https://doi.org/10.1007/3-540-31190-4>
- Bandaru, A. (2022) Min-max normalization. Medium. <https://medium.com/@abhi1achiever/min-max-normalization-db1f515b08b4> 15.2.2026.
- Bevan, N. (1995) Measuring usability as quality of use. *Software Quality Journal* 4 (2) 115–130. <https://doi.org/10.1007/BF00402715>
- Capotondi, A., Alexander, M. A., Bond, N. A., Curchitser, E. N. & Scott, J. D. (2012) Enhanced upper ocean stratification with climate change in the CMIP3 models. *Journal of Geophysical Research: Oceans* 117 C04031 <https://doi.org/10.1029/2011JC007409>
- Cauchy, P. (2021) *Ocean of sound: Underwater gliders observing the oceanic environment*. Väitöskirja. University of East Anglia, Norwich. <https://ueaeprints.uea.ac.uk/id/eprint/79694/1/2020CauchyPPhD.pdf> 11.3.2026.
- Cheng, L., Li, G., Long, S.-M., Li, Y., von Schuckmann, K., Trenberth, K. E., Mann, M. E., Abraham, J., Du, Y., Cheng, X., Liu, H., Xu, Z., Liu, M., Peng, Q., Gong, X., Ma, Z. &

- Yuan, H. (2025) Ocean stratification in a warming climate. *Nature Reviews Earth & Environment* 6 637–655. <https://doi.org/10.1038/s43017-025-00715-5>
- Cleveland, W. S. & McGill, R. (1984) Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American Statistical Association* 79 (387) 531-554.
<https://doi.org/10.1080/01621459.1984.10478080>
- Cressie, N. A. C. (1993) *Statistics for spatial data*.
<https://doi.org/10.1002/9781119115151>
- Closset, I., Tikka, K., Roine, T., Seppänen, J., Siiriä, S.-M. & Laakso, L. (2025) Use of underwater gliders on marine surveillance and observations. *Maanpuolustuksen tieteellisen neuvottelukunnan julkaisuja* 2025:5. Puolustusministeriö.
<https://urn.fi/URN:ISBN:978-951-663-245-5>
- Crameri, F. (2023) Scientific colour maps (version 8.0.1). Zenodo.
<https://doi.org/10.5281/zenodo.1243862>
- Crameri, F., Shephard, G. E. & Heron, P. J. (2020) The misuse of colour in science communication. *Nature Communications* 11 5444.
<https://doi.org/10.1038/s41467-020-19160-7>
- Di Liberto, T. (2021) Hovmöller diagram: A climate scientist's best friend.
<https://www.climate.gov/news-features/understanding-climate/hovmoller-diagram-climate-scientists-best-friend> 12.3.2026.
- Dickey, T. & Chang, G. (2002) Recent advances and future visions: Temporal variability of optical and bio-optical properties of the ocean. *Oceanography* 14 (3) 15–29.
<https://doi.org/10.5670/oceanog.2001.21>
- Eaton, B., Gregory J., Drach, B., Taylor, K., Hankin, S. ym. (2025) *NetCDF Climate and Forecast (CF) Metadata Conventions*. *CF Community*.
<https://doi.org/10.5281/zenodo.17801666>

Esri (2025a) Scientific Colour Maps 8.0.1.

<https://www.arcgis.com/home/item.html?id=db36aff062ec4218a9bd384b74dc6e9d> 28.2.2026.

Esri (2025b) Work with NetCDF files in ArcGIS Pro. <https://support.esri.com/en-us/knowledge-base/working-with-netcdfs-in-arcgis-pro-basics-000037232> 24.1.2026.

Esri (s.a. a) A quick tour of NetCDF data. <https://pro.arcgis.com/en/pro-app/latest/help/data/multidimensional/a-quick-tour-of-netcdf-data.htm> 24.1.2026.

Esri (s.a. b) Add styles to a project. <https://pro.arcgis.com/en/pro-app/latest/help/projects/connect-to-a-style.htm> 28.2.2026.

Esri (s.a. c) An overview of the Interpolation toolset. <https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/an-overview-of-the-interpolation-tools.htm> 26.3.2026.

Esri (s.a. d) Buffer (analysis). <https://pro.arcgis.com/en/pro-app/latest/tool-reference/analysis/buffer.htm> 1.3.2026.

Esri (s.a. e) Create a voxel layer. <https://pro.arcgis.com/en/pro-app/latest/help/mapping/layer-properties/create-a-voxel-layer.htm> 6.3.2026.

Esri (s.a. f) How natural neighbor works. <https://pro.arcgis.com/en/pro-app/3.4/tool-reference/spatial-analyst/how-natural-neighbor-works.htm> 28.2.2026.

Esri (s.a. g) IDW (geostatistical analyst). <https://pro.arcgis.com/en/pro-app/latest/tool-reference/geostatistical-analyst/idw.htm> 28.2.2026.

Esri (s.a. h) IDW (spatial analyst). <https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/idw.htm> 28.2.2026.

Esri (s.a. i) Make multidimensional voxel layer (Multidimension). <https://pro.arcgis.com/en/pro-app/latest/tool-reference/multidimension/make-multidimensional-voxel-layer.htm> 23.1.2026.

Esri (s.a. j) Make NetCDF feature layer (multidimension). <https://pro.arcgis.com/en/pro-app/latest/tool-reference/multidimension/make-netcdf-feature-layer.htm>

28.2.2026.

Esri (s.a. k) NetCDF points to feature class (discrete sampling geometry) (multidimension). <https://pro.arcgis.com/en/pro-app/latest/tool-reference/multidimension/netcdf-points-to-feature-class-discrete-sampling-geometry-.htm> 11.3.2026.

Esri (s.a. l) Supported voxel formats. <https://pro.arcgis.com/en/pro-app/latest/help/mapping/layer-properties/supported-voxel-formats.htm>

24.1.2026.

Esri (s.a. m) What is a voxel layer? <https://pro.arcgis.com/en/pro-app/latest/help/mapping/layer-properties/what-is-a-voxel-layer-.htm> 6.3.2026.

EUMETSAT (European Organisation for the Exploitation of Meteorological Satellites) (s.a.) Data processing and file formats. Teoksessa: *Short_course_07: The color of the oceans*.

<https://classroom.eumetsat.int/mod/book/tool/print/index.php?id=13919>

15.2.2026

Finnish Marine Research Infrastructure (FINMARI) (s.a.) Autonomous platforms.

<https://www.finmari-infrastructure.fi/facilities/autonomous-platforms/> 23.1.2026.

Geographic Information Technology Training Alliance (GITTA) (2016) Distance-based interpolation.

http://www.gitta.info/ContiSpatVar/en/html/Interpolatio_learningObject2.xhtml

12.3.2026.

GISBox (2025) TM35FIN (Finnish Transverse Mercator coordinate system – EPSG:3067).

<https://www.gisbox.com/en/articles/v1/82an8s6iaudk/> 28.2.2026.

The Global Ocean Observing System (GOOS) (2026a) GOOS national focal points.

<https://goosocean.org/who-we-are/goos-national-focal-points/group/> 18.3.2026.

The Global Ocean Observing System (GOOS) (2026b) Ocean Gliders.

<https://goosocean.org/who-we-are/observations-coordination-group/global-ocean-observing-networks/ocean-gliders/> 17.3.2026.

Gregor, L., Ryan-Keogh, T. J., Nicholson, S-A., du Pleiss, M., Giddy, I. & Swart, S. (2019)

GliderTools: A Python toolbox for processing underwater glider data. *Frontiers Marine Science* 6. <https://doi.org/10.3389/fmars.2019.00738>

Gulf of Mexico Research Initiative Information & Data Cooperative (GRIIDC) (2023)

Glider Data Guidance. https://www.griidc.org/media/glider_data_guidance.pdf
27.1.2026.

Hankin, S., Blower, J. D., Carval, T., Casey, K. S. Donlon, D. ym. (2009) NetCDF-CF-

OPeNDAP: Standards for ocean data interoperability and object lessons for community data standards processes. *Proceeding of OceanObs'09*.

<https://doi.org/10.5270/OceanObs09.cwp.41>

Hanson, C. E., Woo, L. M., Thomson, P. G. & Pattiaratchi, C. B. (2016) Observing the

ocean with gliders: Techniques for data visualization and analysis. *Oceanography* 30 (2) 222–227. <https://doi.org/10.5670/oceanog.2017.210>.

Heuzé, C., Ridley, J., Calvert, D. Stevens, D. & Heywood, K. (2015). Increasing vertical

mixing to reduce Southern Ocean deep convection in NEMO3.4. *Geoscientific Model Development* 8 3119–3130. <https://doi.org/10.5194/gmd-8-3119-2015>

Ilmatieteen laitos (2021) Meriveden lämpötilatilastot.

<https://www.ilmatieteenlaitos.fi/meriveden-lampotilatilastot> 9.3.2026.

Ilmatieteen laitos (2022) CTD measures temperature and salinity.

<https://en.ilmatieteenlaitos.fi/ctd> 11.3.2026.

Ilmatieteenlaitos (s.a.) Itämeren kerrostuneisuus.

<https://www.ilmatieteenlaitos.fi/itameren-kerrostuneisuus> 30.3.2026.

IOOS (2025a) Glider background and sampling terminology. Integrated Ocean

Observing System. <https://ioos.github.io/glider-dac/glider-background-and-sampling-terminology.html> 17.3.2026.

IOOS (2025b) NGDAC NetCDF file format version 2. Integrated Ocean Observing System. <https://ioos.github.io/gliders-dac/ngdac-netcdf-file-format-version-2>
19.3.2026.

International Organization for Standardization (ISO) (2018) *ISO 9241-11:2018 Ergonomics of human-system interaction – Part 11: Usability: definitions and concepts*. <https://www.iso.org/obp/ui/en/#iso:std:iso:9241:-11:ed-2:v1:en>
26.1.2026.

Itämeri.fi (s.a. a) Luonto ja sen muutos. <https://itameri.fi/luonto-ja-sen-muutos/>
30.3.2026.

Itämeri.fi (s.a. b) Ilmastonmuutos. <https://itameri.fi/itameren-tila/ilmastonmuutos/>
30.3.2026.

Keim, D., Andrienko, G., Fekete, J.-D., Görg, C., Kohlhammer, J. & Melançon, G. (2008) Visual analytics: Definition, process, and challenges. Teoksessa Kerren, A., Stasko, J., Fekete, J.-D. & North, C. (toim.) *Information visualization* (LNCS 4950, s. 154–175). Springer. https://doi.org/10.1007/978-3-540-70956-5_7

Keim, D.A., Mansmann, F., Schneidewind, J., Thomas, J. & Ziegler, H. (2008) Visual analytics: Scope and challenges. Teoksessa Simoff, S. J., Böhlen, M. H. & Mazeika, A. (toim.) *Visual data mining* (LNCS 4404, s. 76–90). Springer.
https://doi.org/10.1007/978-3-540-71080-6_6

Kelleher, C. & Braswell, A. (2021) Introductory overview: Recommendations for approaching scientific visualization with large environmental datasets. *Environmental Modelling & Software* 143 105113.
<https://doi.org/10.1016/j.envsoft.2021.105113>

Kilsedar, C. E. & Brovelli, M. A. (2020) Multidimensional visualization and processing of big open urban geospatial data on the web. *ISPRS International Journal of Geo-Information* 9 (7) 434. <https://doi.org/10.3390/ijgi9070434>

- Kusuma, D. W., Murdimanto, A., Sukresno, B. & Jatisworo, D. (2018) Comparison of interpolation methods for sea surface temperature data. *Journal of Fisheries and Marine Research* 2 (2) 103–115. <https://doi.org/10.21776/ub.jfmr.2018.002.02.7>
- Li, G., Cheng, L., Zhu, J., Trenberth, K. E., Mann, M. E. & Abraham, J. P. (2020) Increasing ocean stratification over the past half-century. *Nature Climate Change* 10 1116–1123. <https://doi.org/10.1038/s41558-020-00918-2>
- Li, J. & Heap, A. D. (2011) A review of comparative studies of spatial interpolation methods in environmental sciences: Performance and impact factors. *Ecological Informatics* 6 (3–4) 228–241. <https://doi.org/10.1016/j.ecoinf.2010.12.003>
- Liblik, T. & Lips, U. (2011) Characteristics and variability of vertical thermohaline structure in the Gulf of Finland in summer. *Boreal Environment Research* 16 73–83. <https://www.borenv.net/BER/archive/pdfs/ber16/ber16A-073.pdf> 30.3.2026.
- Lucas, G. W. (s.a.) An introduction to natural neighbor interpolation. <https://gwlucastrig.github.io/TinfourDocs/NaturalNeighborIntro/index.html> 26.3.2026.
- Luoma, S. & Okkonen, J (2014) Impacts of future climate change and Baltic Sea level rise on groundwater recharge, groundwater levels, and surface leakage in the Hanko aquifer in southern Finland. *Water* 6 (12) 3671–3700. <https://doi.org/10.3390/w6123671>
- Munzner, T. (2014) *Visualization Analysis and Design*. A K Peters/CRC Press, New York. <https://doi.org/10.1201/b17511>
- Munyati, C. & Sinthumule, N. I. (2021) Comparative suitability of ordinary kriging and Inverse Distance Weighted interpolation for indicating intactness gradients on threatened savannol woodland and forest stands. *Environmental and Suitability Indicators* 12 <https://doi.org/10.1016/j.indic.2021.100151>
- National Institute for Standards and Technology (NIST) (s.a.) Usability. *NIST Computer Security Resource Center (CSRC) Glossary*. <https://csrc.nist.gov/glossary/term/Usability> 26.1.2026.

- NATO Geospatial Maritime Working Group (2022) AML Handbook: Additional military layers – information superiority in the maritime environment.
https://assets.admiralty.co.uk/public/2022-06/20220322_NATO%20AML%20Handbook%20Edition%205.pdf?VersionId=in9okGHLWSu0hqMdO0ihTlNgWggo86hq 5.4.2026.
- NumPy Developers (2025) *NumPy User Guide*.
<https://numpy.org/doc/stable/user/index.html#user> 15.2.2026.
- Ohnemus, D. C., Lam, P. J. Twining, B. S. (2018) Optical observation of particles and responses to particle composition in the GEOTRACES GP16 section. *Marine Chemistry* 201 124–136. <https://doi.org/10.1016/j.marchem.2017.09.004>
- Reynolds, R. A., Stramski, D. & Neukermans, G. (2016) Optical backscattering by particles in Arctic seawater and relationships to particle mass concentration, size distribution and bulk composition. *Limnology and Oceanography* 61 1869–1890.
<https://doi.org/10.1002/lno.10341>
- Ridgway, K., Dunn, J. R. & Wilkin, J. L. (2002) Ocean interpolation by four-dimensional weighted least squares – application to the waters around Australasia. *Journal of Atmospheric and Oceanic Technology* 19 (9) 1357–1375.
[https://doi.org/10.1175/1520-0426\(2002\)019<1357:OIBFDW>2.0.CO;2](https://doi.org/10.1175/1520-0426(2002)019<1357:OIBFDW>2.0.CO;2)
- Roch, M., Brandt, P. & Schmidtko, S. (2023) Recent large-scale mixed layer and vertical stratification maxima changes. *Frontiers in Marine Science* 10 1277316.
<https://doi.org/10.3389/fmars.2023.1277316>
- Rudnick, D. L. & Cole, S. T. (2011) On sampling the ocean using underwater gliders. *Journal of Geophysical Research* 116. <https://doi.org/10.1029/2010JC006849>
- Rudnick, D. L., Davis, R., Eriksen, C. C. & Fratantoni, D. M. (2004) Underwater gliders for ocean research. *Marine Technology Society Journal* 38 (2) 73–84.
<https://doi.org/10.4031/002533204787522703>
- Sang, Q., Tian, Y., Yu, J. & Zhang, F. (2026) A dynamic data driven application system for ocean environment prediction using reduced-order modeling and adaptive glider

sensing networks. *Ocean Engineering*.

<https://doi.org/10.1016/j.oceaneng.2026.124812>

Sequoia (2012) Optical transmission – what is it?

<https://www.sequoiasci.com/article/optical-transmission-what-is-it/> 11.3.2026.

Sequoia (2021) LISST-Tau high precision transmissometer.

https://www.sequoiasci.com/wp-content/uploads/2021/03/LISST-Tau_Users_Manual_v1_35.pdf 11.3.2026.

Sequoia (2022) LISST-200X particle size analyzer. https://www.sequoiasci.com/wp-content/uploads/2016/02/LISST-200X_Users_Manual_v2_35.pdf 11.3.2026.

Shneiderman, B. (1996) The eyes have it: A task by data type taxonomy for information visualizations. *Proceedings of the IEEE Symposium on Visual Languages* 336–343. <https://doi.org/10.1109/VL.1996.545307>

Temoshok, D., Proud-Madruga, D., Choong, Y.-Y., Galluzzo, R., Gupta, S., LaSalle, C., Lefkovitz, N. & Regenscheid, A. (2025) *Digital Identity Guidelines (NIST SP 800-63-4)*. National Institute of Standards and Technology. <https://doi.org/10.6028/NIST.SP.800-63-4>

Testor, P., de Young, B., Rudnick, D. L., Glenn, S., Hayes, D. ym. D (2019) OceanGliders: A component of the integrated GOOS. *Frontiers in Marine Science* 6 422. <https://doi.org/10.3389/fmars.2019.00422>

Tian, B. (2016) *GIS technology applications in environmental and earth sciences*. CRC Press, Boca Raton. <https://doi.org/10.1201/9781315366975>

The pandas development team (2026) *User Guide*. Pandas Documentation.

https://pandas.pydata.org/docs/user_guide/index.html#user-guide 15.2.2026

Unidata Program Center (University of Corporation for Atmospheric Research, UCAR) (s.a.) *The NetCDF user's guide*.

https://docs.unidata.ucar.edu/nug/current/?rsource=https%3A%2F%2Flinks.esri.com%2Fnetcdf_UserGuide 23.1.2026.

- Veness, C. (s.a.) Calculate distance, bearing and more between latitude/longitude points. Movable Type Scripts. <https://www.movable-type.co.uk/scripts/latlong.html> 13.2.2026.
- Veness, C. (s.a.). Vincenty's formula for distance between points of the ellipsoid. Movable Type Scripts. <https://www.movable-type.co.uk/scripts/latlong-vincenty.html> 14.2.2026.
- Von Oppeln-Bronikowski, N., de Young, B., Atamanchuk, D. & Wallace, D. (2021) Glider-based observations of CO₂ in the Labrador Sea. *Ocean Science* 17 (1) 1–16. <https://doi.org/10.5194/os-17-1-2021>
- Ware, C. (2020) *Information visualization: Perception for design* (4. painos). Morgan Kaufmann. Cambridge, MA. <https://doi.org/10.1016/C2016-0-02395-1>
- Weather Stream (2026) What is a NetCDF? <https://weatherstream.com/what-is-a-netcdf/> 19.3.2026.
- World Meteorological Organization (WMO) (2025) The global ocean observing system 2025 status report released. World Meteorological organization. <https://wmo.int/media/news/global-ocean-observing-system-2025-status-report-released> 18.3.2026.
- Xarray Developers (2026) *User Guide*. <https://docs.xarray.dev/en/stable/user-guide/index.html> 15.2.2026.

Liitteet

Liite 1. Ohjelmisto- ja laskentaympäristö

Laitteisto:

- CPU: AMD Ryzen 7 5800X3D
- Motherboard: ASRock B450 Pro4
- GPU: NVIDIA GeForce RTX 4070 Ti
- RAM: 32Gb DDR4 (3200 MHz)
- OS: Windows 11 Pro

Sovellukset:

- Python 3.13.12
- Pip 26.0.1
- ArcGIS Pro 3.6.1
- Visual Studio Code 1.109.0

Python-kirjastot:

- pandas 3.0.0
- xarray 2026.1.0
- numpy 2.4.2

Tiedostot:

- ArcGIS Prohon ladattava väripaletti: `crameri_colourmaps.stylx`

NetCDF-data: TVAR20242.nc

Koodin tuottama data:

- tulokset.csv, jossa päiväkohtaiset tunnusluvut sekä aktiivisuusindeksi
- valittu_paiva.nc, jossa valitun päivän kaikki havainnot sekä dimensiot NetCDF-muodossa

Liite 2. Aineiston esikäsittelyn Python-koodi (split_data.py) sekä tuloste

```
import xarray as xr
import pandas as pd
import numpy as np

# Nimetään NetCDF-tiedosto dataksi
data = xr.open_dataset("data/TVAR20242.nc")

# Muutetaan data taulukoksi ja otetaan päivänvalintaan tarvittavat muuttujat
dataframe = data[["latitude", "longitude", "depth"]].to_dataframe().reset_index()

# Aikaleimat
dataframe["time"] = pd.to_datetime(dataframe["time"], utc=True)
dataframe = dataframe.sort_values("time")

# Puuttuvien arvojen poisto
dataframe = dataframe.dropna(subset=["latitude", "longitude", "depth"])

# Tarkistetaan puuttuvien havaintojen määrä
puuttuvat = data.sizes["time"] - len(dataframe)
print("Puuttuvien havaintojen määrä: ")
print(puuttuvat)

# Lasketaan päivittäisten havaintojen summa
paivittaiset = dataframe.groupby(dataframe["time"].dt.date).size().sum()
print("Päivittäisten havaintojen summa: ")
print(paivittaiset)

# Varmistetaan, että päivittäisten havaintojen summa täsmää dataframeen
assert paivittaiset == len(dataframe),
"Päivittäisten havaintojen summa ei täsmää dataframeen"

# Sarake päivämäärille
dataframe["date"] = dataframe["time"].dt.date

# Horisontaalisen liikkeen mukaan aktiivisuus
# Lasketaan havaintopisteiden välinen etäisyys vaakasuunnassa
# Haversine-kaavan käyttö
def matkan_lasku(lat, lon):
```

```

# Radiaanit
lat_rad = np.radians(lat)
lon_rad = np.radians(lon)

# Pisteiden väli
ero_lat = np.diff(lat_rad)
ero_lon = np.diff(lon_rad)

# Kaava
a = (np.sin(ero_lat/2)**2 + np.cos(lat_rad[:-1]) * np.cos(lat_rad[1:]) *
      np.sin(ero_lon/2)**2)
c = 2 * np.arctan2(np.sqrt(a), np.sqrt(1-a))

# Maan säde
R = 6371

# Etäisyydet pisteiden välillä
return R * c

# Päivän gliderin aktiivisuus
tiedot_paivista = []

for paiva, joukko in dataframe.groupby("date"):
    # Järjestetään päivän tulokset aikajärjestykseen
    joukko = joukko.sort_values("time")

    # Syvyysvaihtelu
    syvyysvaihtelu = joukko["depth"].max() - joukko["depth"].min()

    # Kuljettu matka horisontaalisesti (km)
    matka = matkan_lasku(joukko["latitude"].values, joukko["longitude"].values).sum()

    # Havaintopisteiden määrä
    havainto_maara = len(joukko)

    # Tulosten talletus listaan
    tiedot_paivista.append({"Päivä": paiva, "Havaintojen määrä": havainto_maara,
                           "Syvyysvaihtelu (m)": syvyysvaihtelu, "Matka (km)": matka})

# Muutetaan tiedot pandas.DataFrame tyyppiin ja indeksoidaan 1 alkaen

```

```

tulokset = pd.DataFrame(tiedot_paivista)
tulokset.index += 1

# Normalisoidaan aktiivisuusindeksin muuttujat
tulokset["Syvyys_norm"] = (tulokset["Syvyysvaihtelu (m)"] / tulokset["Syvyysvaihtelu (m)"].max())
tulokset["Matka_norm"] = (tulokset["Matka (km)"] / tulokset["Matka (km)"].max())

# Painotetaan molemmat 0.5
tulokset["Aktiivisuus"] = (0.5 * tulokset["Syvyys_norm"] + 0.5 * tulokset["Matka_norm"])

# Tallennetaan tulokset csv-taulukoksi
tulokset.to_csv("tulokset.csv", encoding="utf-8-sig")

print("Kaikki päivät:")
print(tulokset)

paras_aktiivisuus = tulokset.sort_values("Aktiivisuus", ascending = False).iloc[0]
print("Valitaan testipäiväksi: ")
print(paras_aktiivisuus)

# Valitun päivän datan erottaminen alkuperäisestä NetCDF-tiedostosta
alku = pd.to_datetime(paras_aktiivisuus["Päivä"])
loppu = alku + pd.Timedelta(days=1)

# Varmistetaan, että alkuperäinen NetCDF-tiedosto käyttää aikaa indeksinä
data = data.set_index(time="time")

paivan_data = data.where((data["time"] >= alku) & (data["time"] < loppu), drop=True)
paivan_data.to_netcdf("valittu_paiva.nc")

```

Puuttuvien havaintojen määrä:

1

Päivittäisten havaintojen summa:

614469

Kaikki päivät:

	Päivä	Havaintojen määrä	Syvyysvaihtelu (m)	Matka (km)	Syvyys_norm	Matka_norm	Aktiivisuus
1	2024-05-06	34013	38.490686	28.678919	0.666355	0.325797	0.496076
2	2024-05-07	66374	46.443451	19.897913	0.804034	0.226043	0.515039
3	2024-05-08	76891	55.128740	19.998911	0.954395	0.227191	0.590793
4	2024-05-09	79290	53.742307	20.072815	0.930393	0.228030	0.579212
5	2024-05-10	78051	57.763009	20.813863	1.000000	0.236449	0.618224
6	2024-05-11	77715	46.938645	88.027044	0.812607	1.000000	0.906304
7	2024-05-12	78471	55.812068	20.041362	0.966225	0.227673	0.596949
8	2024-05-13	87386	48.968848	20.235886	0.847754	0.229883	0.538819
9	2024-05-14	36277	33.350441	10.229188	0.577367	0.116205	0.346786
10	2024-05-15	1	0.000000	0.000000	0.000000	0.000000	0.000000

Valitaan testipäiväksi:

Päivä 2024-05-11

Havaintojen määrä 77715

Syvyysvaihtelu (m) 46.938645

Matka (km) 88.027044

Syvyys_norm 0.812607

Matka_norm 1.0

Aktiivisuus 0.906304

Name: 6, dtype: object

Liite 3. ArcGIS työnkulku

Taustakartta

Light Gray Base

Pisteaineiston luonti NetCDF-tiedostosta:

Make NetCDF Feature Layer

Input netCDF File: valittu_paiva.nc

Variables: longitude, latitude, depth, temperature, profile_index,
profile_direction, density

X Variable: longitude

Y Variable: latitude

Output Feature Layer: Pisteet

Row Dimensions: time

Z Variable: depth

Syvyysluokkien lisäys:

Attribute Table → Add Field →

Field Name: depth_classes

Data Type: short

Calculate Field

Field = depth_classes(!depth_classes!)

Code Block:

```
def depth_classes (d):
```

```
    if 0 <= d < 10:
```

```
    return 1

elif 10 <= d < 20:

    return 2

elif 20 <= d < 30:

    return 3

elif 30 <= d < 40:

    return 4

elif 40 <= d <= 50:

    return 5
```

Syvyyserrokset:

Data → Definition Query → New definition query

```
0–10 m: depth_classes = 1

10–20 m: depth_classes = 2

20–30 m: depth_classes = 3

30–40 m: depth_classes = 4

40–50 m: depth_classes = 5
```

Koordinaattijärjestelmän muutos:

Project (Data Management Tools)

Input Dataset or Feature Class: Pisteet

Output Dataset or Feature Class: pisteet

Output Coordinate System: EUREF-FIN TM35FIN

Virheellisen ajanjakson poisto:

Poistetaan pisteet ajalta 6:15:00–8:22:20 attribuuttitaulukossa

Pisteaineiston visualisointi:

Symbology → Graduated Colors

Field: temperature

Method: Equal Interval

Classes: 8

Color Scheme: batlow

Template: 4pt, no outline

Advanced symbology

Maximum Sample Size: 80 000

Bufferointi:

Buffer

Input Features: pisteet

Output Feature Class: pisteet_Buffer_100m

Distance: 100 m

Method: Planar

Dissolve Type: No Dissolve

IDW-interpolointi:

IDW (Statistical Analysis Tools)

Input Features: pisteet

Use the filtered records (0-10 m, 10-20 m, 20-30 m, 30-40 m, 40-50 m,
pois päältä)

Z-value field: temperature

Output cell size: 100

Search neighbourhood:

Max neighbors: 20

Min neighbors: 10

Sector Type: 1

Output raster layers:

IDW_0_10m

IDW_10_20m

IDW_20_30m

IDW_30_40m

IDW_40_50m

IDW_0_50m

Processing Extent: pisteet_Buffer_100m

Mask: pisteet_Buffer_100m

Kriging-interpolointi:

Kriging (Statistical Analysis Tools)

Input Features: pisteet

Use the filtered records (0-10 m, 10-20 m, 20-30 m, 30-40 m, 40-50 m,
pois päältä)

Z-value field: temperature

Output cell size: 100

Semivariogram properties:

Ordinary

Spherical

Output raster layers:

Kriging_0_10m

Kriging_10_20m

Kriging_20_30m

Kriging_30_40m

Kriging_40_50m

Kriging_0_50m

Processing Extent: pisteet_Buffer_100m

Mask: pisteet_Buffer_100m

Natural Neighbor-interpolointi:

Natural Neighbor (Statistical Analysis Tools)

Input point features: pisteet

Use the filtered records (0-10 m, 10-20 m, 20-30 m, 30-40 m, 40-50 m,
pois päältä)

Z-value field: temperature

Output cell size: 100

Output raster layers:

Natural_0_10m

Natural_10_20m

Natural_20_30m

Natural_30_40m

Natural_40_50m

Natural_0_50m

Interpoloitujen karttatasojen visualisointi:

Symbology → Stretch

Color Scheme: batlow

Min: 2

Max: 6

Stretch Type: Minimum Maximum

Gamma: 1.0

Statistic: Custom

Apply Symbology from Layer:

Input Layer: IDW_10_20m, IDW_20_30m, IDW_30_40m, IDW_40_50m,
IDW_0_50m

Symbology Layer: IDW_0_10m

Viivadiagrammi:

Charts → Create Chart → Line plot

Date or Number: time

Aggregation: none

Numeric field: depth

Poikkileikkauskuvat:

Charts → Create Chart → Scatter plot

X-axis Number: profile_index

Y-axis Number: depth (inverted)

Laatikkojanakuvat :

Charts → Create Chart → Box plot

Numeric field: temperature / density

Category: depth_class

Viivadiagrammi:

Charts → Create Chart → Line plot

Date or Number: time

Numeric field: depth