Horizon 2020 LC-SPACE-04-EO-2019-2020

Copernicus Evolution – Research for harmonised and Transitional-water Observation (CERTO)

Project Number: 870349

Deliverable No: D4.2		Work Package: 4	
Date:	13-JUL-2022	Contract delivery due date	31-AUG-2021
Title:	WP4 Optical water type set		
Lead Partner for Deliverable	PML		
Author(s):	Thomas Jackson, PML, Elizabeth Atwood, PML, Angus Laurenson, PML		
Dissemination level (PU=public, RE=restricted, CO=confidential)			PU
Report Status (DR = Draft, FI = FINAL)			FI

Acknowledgements

This project has received funding from the European Union's Horizon 2020 research and innovation programme grant agreement N° 870349

Table of Contents

Glossary ... 3
1 Executive Summary ... 4
2 Introduction ... 4
3 Regional OWT sets ... 6
3.1 S3 Input training data ... 6
3.2 S2 Input training data ... 6
3.3 Cluster set testing parameters ... 6
3.4 Curonian Lagoon ... 7
3.4.1 Sentinel 3 ... 8
3.4.2 Sentinel 2 ... 10
3.5 Danube Delta ... 11
3.5.1 Sentinel 3 ... 13
3.5.2 Sentinel 2 ... 15
3.6 Elbe Estuary ... 16
3.6.1 Sentinel 3 ... 18
3.6.2 Sentinel 2 ... 20
3.7 Tagus Estuary ... 21
3.7.1 Sentinel 3 ... 23
3.7.2 Sentinel 2 ... 25
3.8 Tamar River/Plymouth Sound ... 26
3.8.1 Sentinel 3 ... 28
3.8.2 Sentinel 2 ... 30
3.9 Venice ... 31
3.9.1 Sentinel 3 ... 33
3.9.2 Sentinel 2 ... 35
4 Pan-Regional Clusters ... 37
4.1 Sentinel 3 ... 37
4.1.1 Cluster reflectance spectra ... 37
4.1.2 Distribution of clusters in case study regions ... 39
4.2 Sentinel 2 ... 43
4.2.1 Cluster reflectance spectra ... 43
4.2.2 Distribution of clusters in case study regions ... 43
5 Cluster set files ... 48

Glossary

C3S	Copernicus Climate Change Service
CERTO	Copernicus Evolution – Research for Transitional-water Observation
CLMS	Copernicus Land Monitoring Service
CMEMS	Copernicus Marine Environment Monitoring Service
c-means	Soft (or fuzzy) clustering scheme in which a data point can belong to more than one cluster
EO	Earth observation
ESA	European Space Agency
k-means	Hard clustering scheme in which each data point belongs to a single cluster
OC-CCI	Ocean Colour Climate Change Initiative
OLCI	Ocean and Land Colour Instrument
OWT	Optical water type
PCA	Principal Component Analysis
R_{rs}	Remote sensing reflectance
S2	Copernicus Sentinel-2
S3	Copernicus Sentinel-3
std	Standard deviation
WP	Work Package

1 Executive Summary

- This short report provides a brief description of the OWT sets created for use within CERTO for algorithm analysis and blending purposes. It contains sections for both regional and pan-regional analysis, covering both the training sets used to create the optical water types and a description of the final class sets.

- This document is partnered by a set of .pkl (pickle) files containing the class sets in a format that can be read back into the classifier software for use with S3 and S2 data.

2 Introduction

This report presents a summary of the regional and pan-regional optical water class sets that have been created and used thus far within the CERTO project.

Within the context of this document optical water class sets use fuzzy membership (c-means) rather than hard membership (k-means) classification schemes, meaning that classified spectra are assigned a membership between 0 and 1 for each class in the set.

Each of the water class sets is summarised below and covered in greater detail in the regional reports that were provided to regional partners for CERTO field campaign planning.

For each OWT set we will cover:

- The input training data
- Cluster set testing parameters
- Cluster reflectance spectra
- Distribution of clusters in case study regions

Cluster set generation was performed using the CERTO developed OWT python package. Therefore, a number of elements, such as available testing metrics, were common across all clustering exercises. Figure 1 below shows a generic example output from the cluster generation process. The figure shows four panels with metric scores against possible number of clusters, from left to right corresponding to four different fuzziness factor 'm' values (as shown on the top bar from 1.5 to 5). The four coloured sets of points in each panel correspond to four different metrics or scores of cluster set performance (Xie-Beni index, Fuzzy partition coefficient, Silhouette score and the Davies Bouldin metric), with the error bars showing the variance generated by bootstrapping the analysis with 5 subsets of the total sample. For a set of K clusters and a dataset consisting of N data points, the Xie-Beni index is defined as:

$$C = \frac{1}{N} \frac{WGSS}{\min\limits_{k<k'} \delta_1(C_k\,,\ C_{k'})^2} \tag{1}$$

where $WGSS$ is the pooled within-cluster sum of squares (the sum of within-cluster dispersions for all clusters), C_k and $C_{k'}$ refer to separate clusters, and δ_1 is the single-linkage distance, defined as:

$$\delta_1 = \min_{\substack{i \in I_k \\ j \in I_{k'}}} d(M_i,\ M_j) \tag{2}$$

with M_i, M_j being pairs of observation points.

The Fuzzy Partition coefficient (F) was defined by Bezdek (1981):

$$F = \frac{1}{N}\sum_{j=1}^{K}\sum_{i=1}^{N}(\mu_{ij})^m \tag{3}$$

with μ_{ij} the membership value of point *i* to cluster *j*. The Silhouette score is calculated as:

$$\mathcal{C} = \frac{1}{K}\sum_{k=1}^{K} \mathfrak{s}_k \tag{4}$$

where $\mathfrak{s}_k$ is the per-cluster mean silhouette derived from all silhouette widths (s):

$$\mathfrak{s}_k = \frac{1}{n_k}\sum_{i \in I_k} s(i) \tag{5}$$

The Davies Bouldin index is calculated as:

$$DB = \frac{1}{K}\sum_{k=1}^{K} M_k = \frac{1}{K}\sum_{k=1}^{K}\max_{k' \neq k}\left(\frac{\delta_k + \delta_{k'}}{\Delta_{kk'}}\right) \tag{6}$$

where δ_k is the mean distance between the points belonging to cluster k and the cluster barycentre.

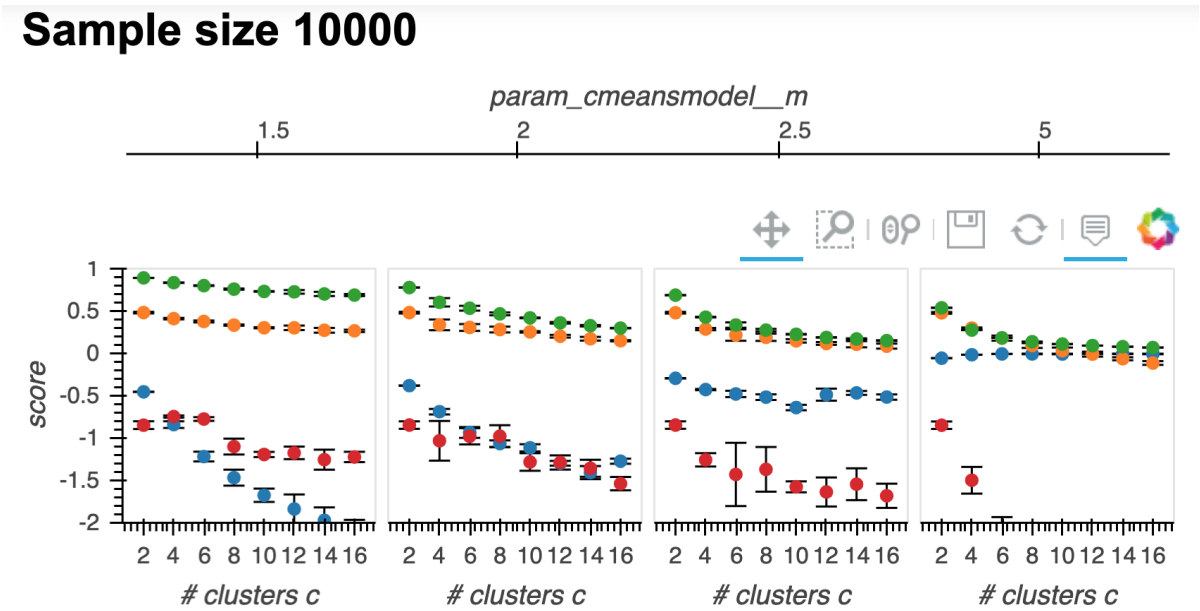

Figure 1: Cluster set scoring across a range of metrics, fuzziness values and number of clusters. Metrics, explained in the text, are shown by colour: Xie-Beni index (blue), Silhouette index (orange), Fuzzy partition coefficient (green) and Davies Bouldin index (red).

For all exercises listed below these four metrics were calculated, but the Xie-Beni index was used as the primary indicator of cluster set suitability.

3 Regional OWT sets

The regional OWT clustering exercises shared a common approach to selecting input training data and cluster searching. This is detailed below, with region-specific results following separately.

3.1 S3 Input training data

- OLCI full resolution data (300m, sensors A and B) from April 2016 to March 2021 sampled to build the training dataset.
- Data were atmospherically corrected with POLYMER, following the CALIMNOS lakes processing chain.
- R_w bands used for clustering: 400, 412, 443, 490, 510, 560, 620, 665, 674, 681, 709, 754, 779, 865, 885 nm.
- Region-specific winter month exclusion from the training dataset (based on incident light level <30° calculated using NOAA solcalc).
- Random sampling used inverse distance to coast weighting (more samples closer to the coast) to balance sampling in more complex waters with often larger area oceanic waters.
- Total training data size of 100,000 samples per region.

3.2 S2 Input training data

- MSI 60m resolution data (sensors A and B) from November 2016 to December 2020 sampled to build the training dataset.
- Data were atmospherically corrected with POLYMER, following the CALIMNOS lakes processing chain.
- R_w bands used for clustering: 443, 490, 560, 665, 705, 740, 783 and 865 nm.
- Region-specific winter month exclusion from training dataset (based on incident light level <30° calculated using NOAA solcalc).
- Random sampling used inverse distance to coast weighting (more samples closer to the coast) to balance sampling in more complex waters with often larger area oceanic waters.
- Total training data size of 100,000 samples per region.

3.3 Cluster set testing parameters

- Within the pipeline (prior to clustering), spectra were standardised using a spectrally-flat reflectance offset (0.015 for S3 data and 0.003 for S2 data) to remove negative reflectance values in NIR bands, followed by a log-transformation and principal component analysis (PCA).
- Cluster centre locations were optimised using Euclidian distance in principal-component-space, which is comparable to the Mahalanobis distance in untransformed space when data are normally distributed.

3.4 Curonian Lagoon

The regional extent and an example training data sample distribution are shown in Figure 2. Coverage of sampling points is roughly identical (allowing for resolution differences) for both S3 and S2 based clustering exercises.

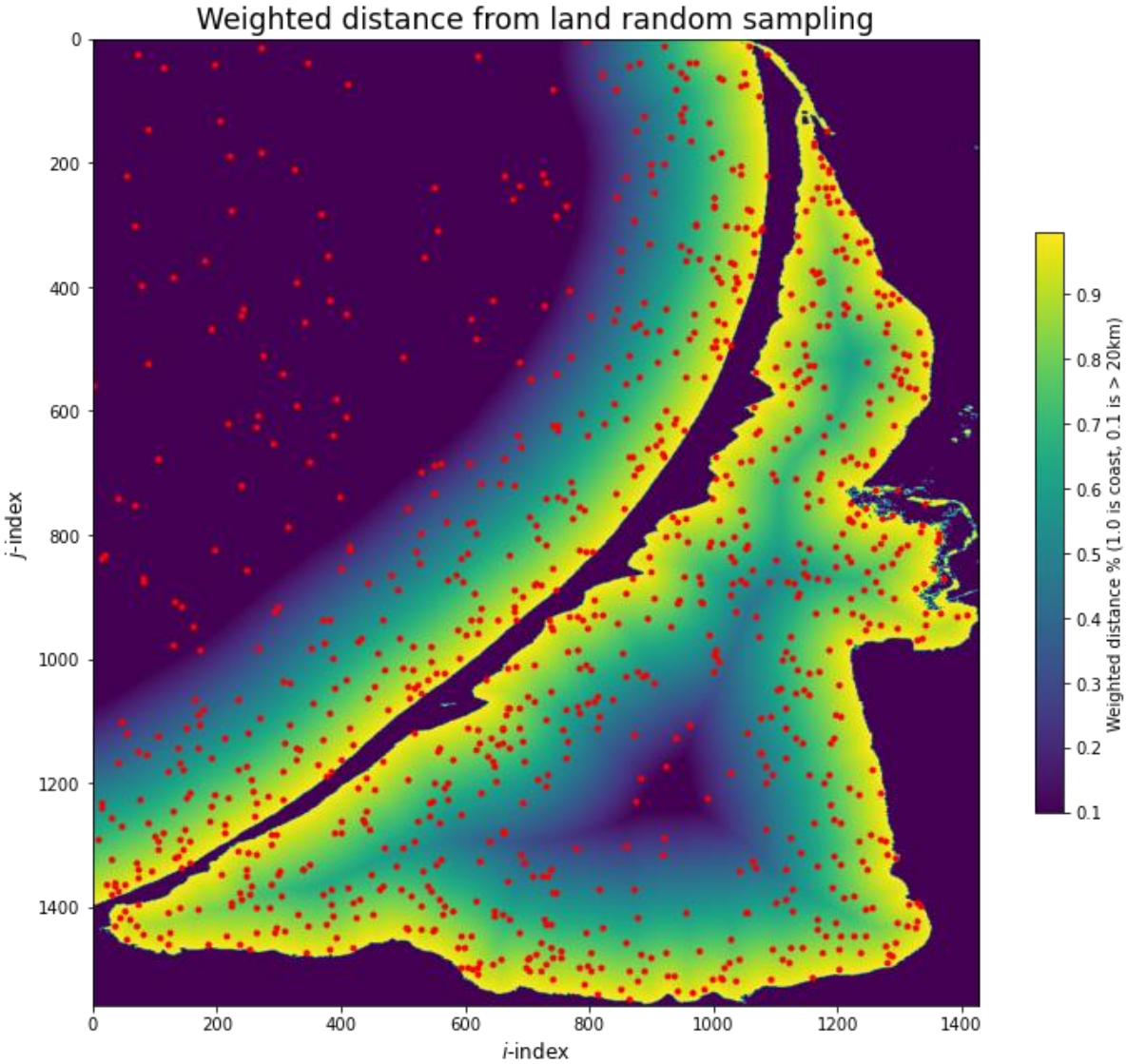

Figure 2: S2 Curonian Lagoon sample weighting map, background colour (dark blue to yellow) represents weighting using distance to coast. Example weighted random sample (red dots).

3.4.1 Sentinel 3

3.4.1.1 Cluster reflectance spectra

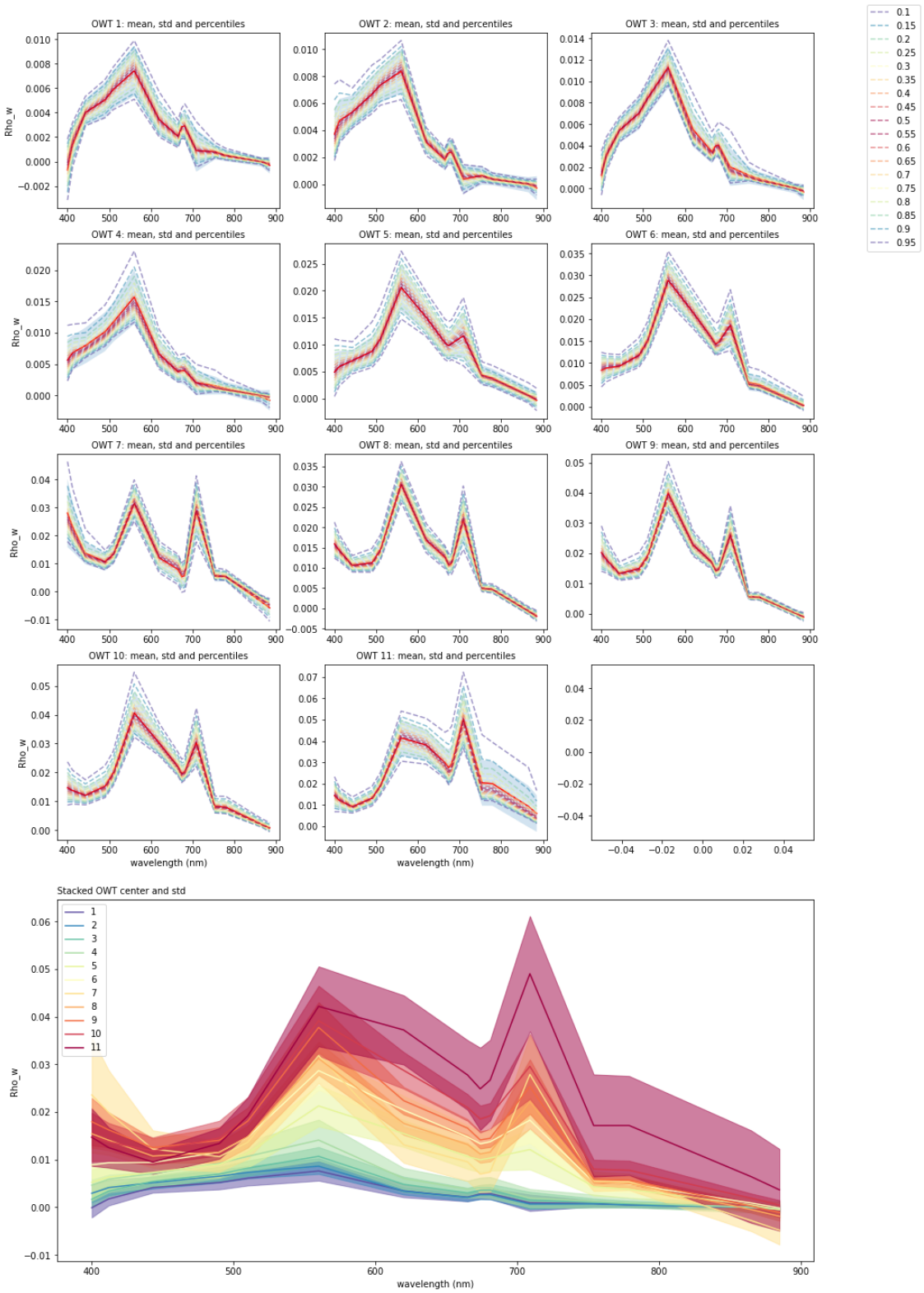

Figure 3: Sentinel-3 Optical Water Type (OWT) clusters created for the Curonian Lagoon. Top panels show cluster centre spectra (red solid line) for spectra distribution for training data dominated by membership to given cluster, together with ± 1 std (gray shading) and distribution percentiles (hatched lines, percentile indicated in legend). Lower panel shows centre spectrum ± 1 std for all classes overlain.

3.4.1.2 Distribution of clusters in case study region

Dominant OWT spatial distribution over the entire timeseries (Figure 4) shows that OWTs 1 to 4 primarily represent the Baltic Sea waters. OWTs 4 and 5 capture well the waters along the north-eastern coast of the lagoon, with OWT 5 in particular picking up the waters at the outflow of the Neman River. The plume of lagoon water entering the Baltic Sea to the north is not sufficiently present through the entire timeseries (it is likely variable in strength and location). OWT 6 primarily occurs in the northeast waters of the lagoon, and the remaining OWTs 7 to 10 represent most of inner lagoon waters. OWT 11 remains a rarely occurring and scattered water type. Monthly analysis shows that OWT 11 has significant coverage for some months (particularly February).

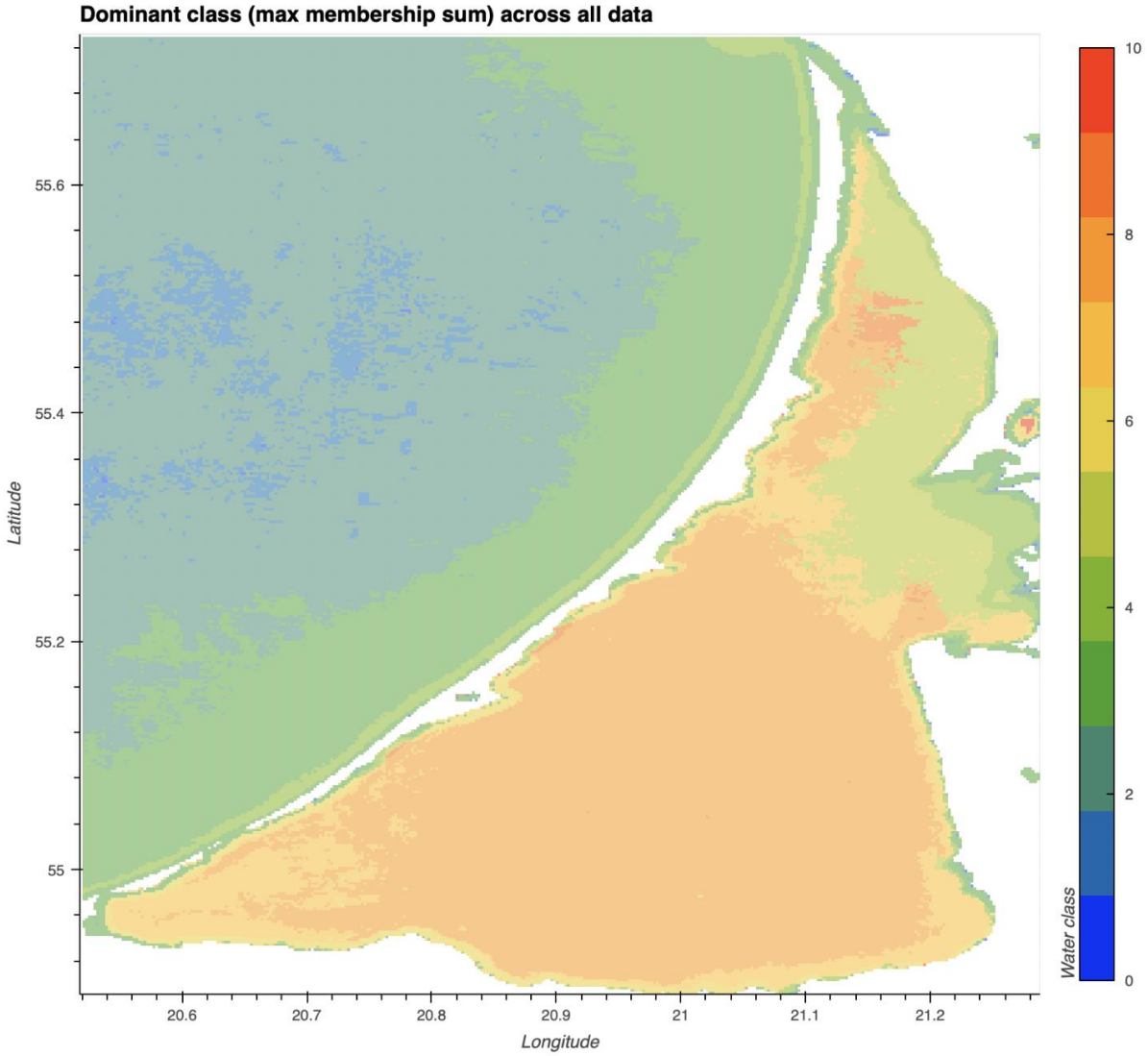

Figure 4: Dominant S3 OWT for each pixel within Curonian Lagoon region over the entire timeseries (2016 to 2021), note that colour bar is 0 indexed.

3.4.2 Sentinel 2

3.4.2.1 Cluster reflectance spectra

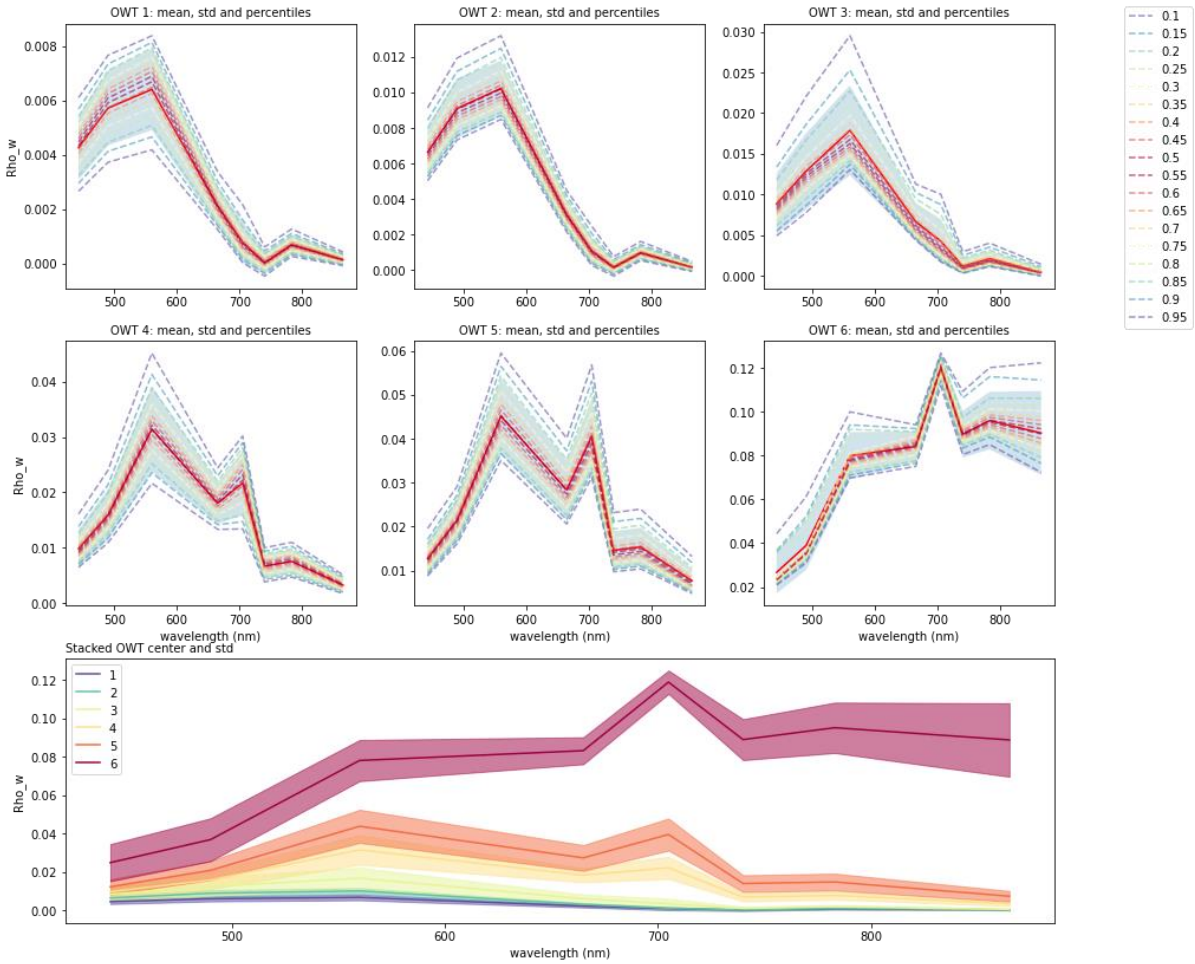

Figure 5: Sentinel-2 Optical Water Type (OWT) clusters created for the Curonian Lagoon. Top panels show cluster centre spectra (red solid line) for spectra distribution for training data dominated by membership to given cluster, together with ± 1 std (gray shading) and distribution percentiles (hatched lines, percentile indicated in legend). Lower panel shows centre spectrum ± 1 std for all classes overlain.

3.4.2.2 Distribution of clusters in case study region

Dominant OWT spatial distribution over the entire timeseries (Figure 6) shows that OWT 1 and 2 primarily represent the Baltic Sea waters. OWT 3 captures well the waters along the north-eastern coast of the lagoon but the plume of lagoon water entering the Baltic Sea to the north is not sufficiently present through the entire timeseries (it is likely variable in strength and location). OWT 4 represents primarily inner lagoon waters, with OWT 5 representing the lagoon waters in the South. OWT 6 remains a rarely occurring and scattered water type. Monthly analysis shows that that OWT 6 does occur in higher frequencies for some months (like January and December), in a similar manner to OWT 11 of the S3 cluster set. Also, the band of OWT 3 along the north-eastern coast of the lagoon appears to break up for the winter months.

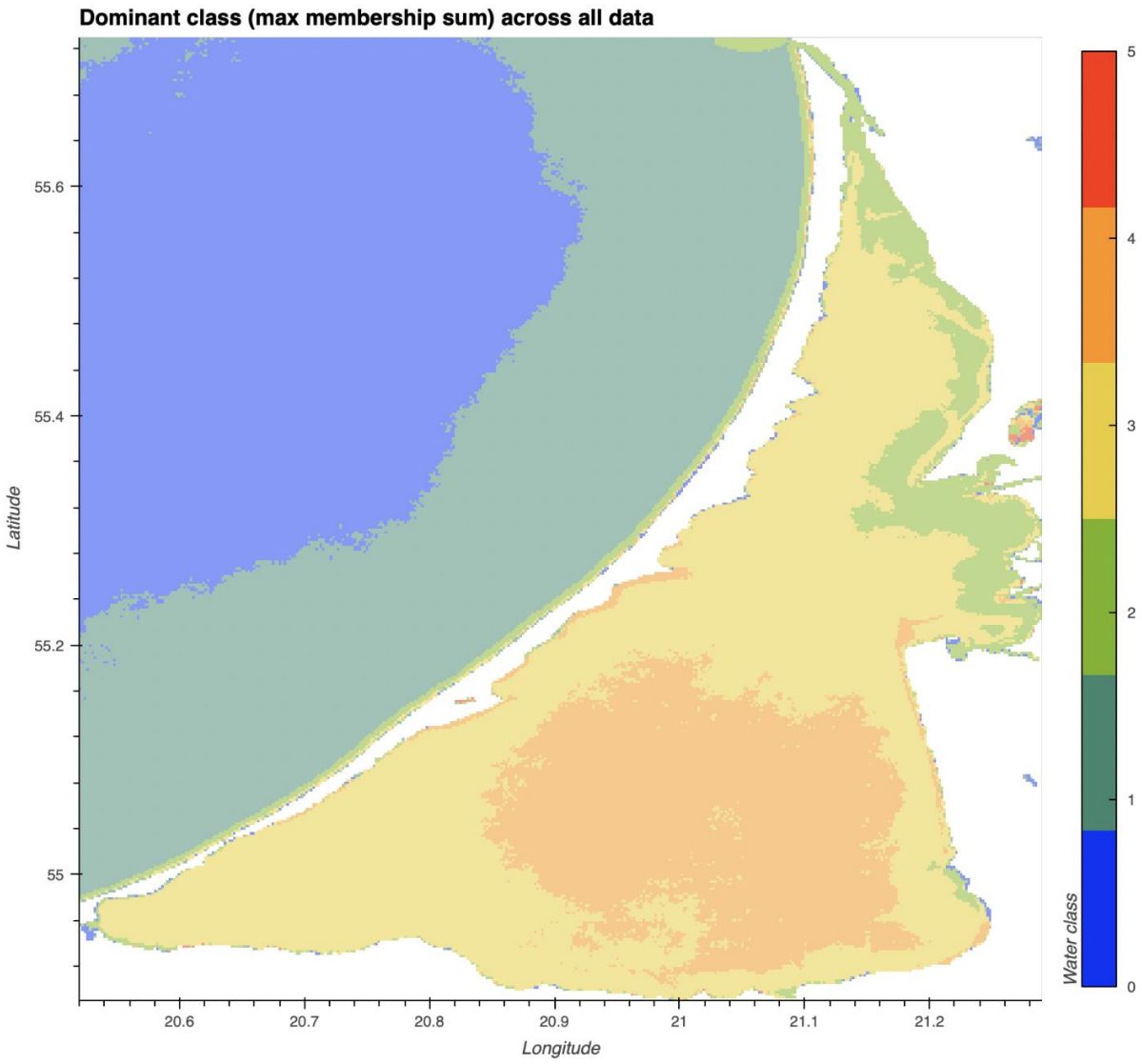

Figure 6: Dominant S2 OWT for each pixel within Curonian Lagoon region over the entire timeseries (2016 to 2020), note that colour bar is 0 indexed.

3.5 Danube Delta

The regional extent and an example training data sample distribution are shown in Figure 7. Coverage of sampling points is roughly identical (allowing for resolution differences) for both S3 and S2 based clustering exercises.

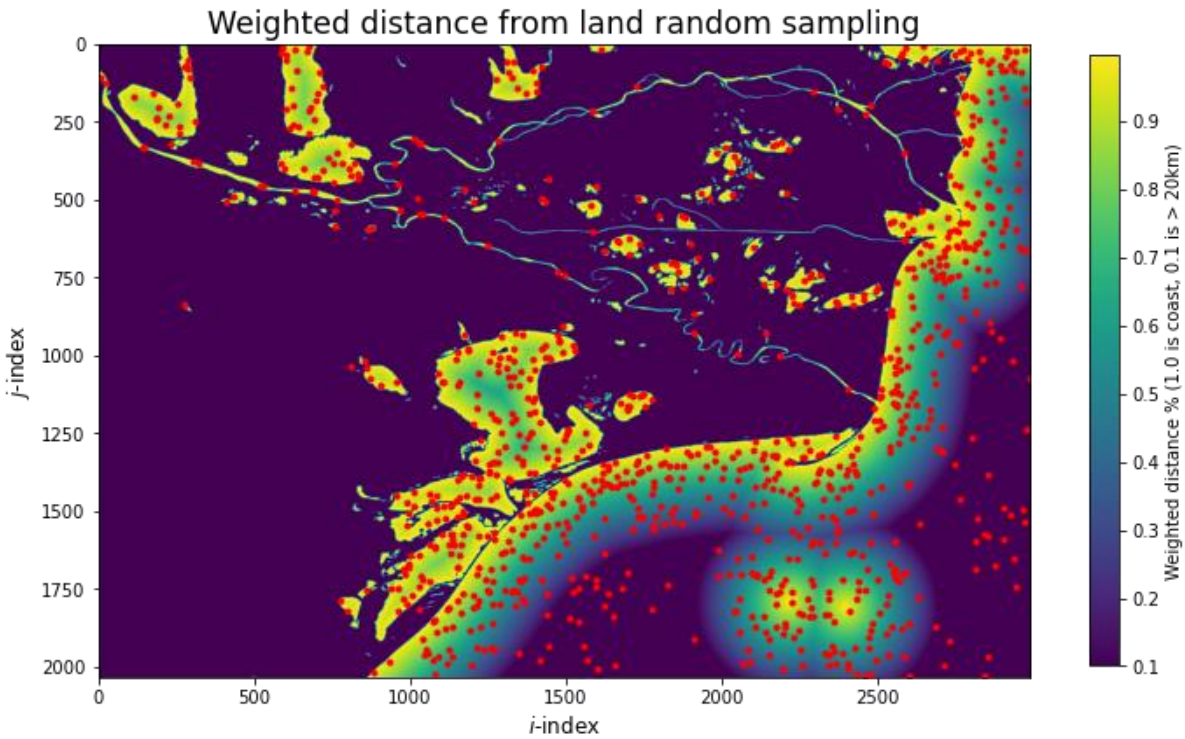

Figure 7: S2 Danube Delta sample weighting map, background colour (dark blue to yellow) represents weighting using distance to coast. Example weighted random sample (red dots).

3.5.1 Sentinel 3

3.5.1.1 Cluster reflectance spectra

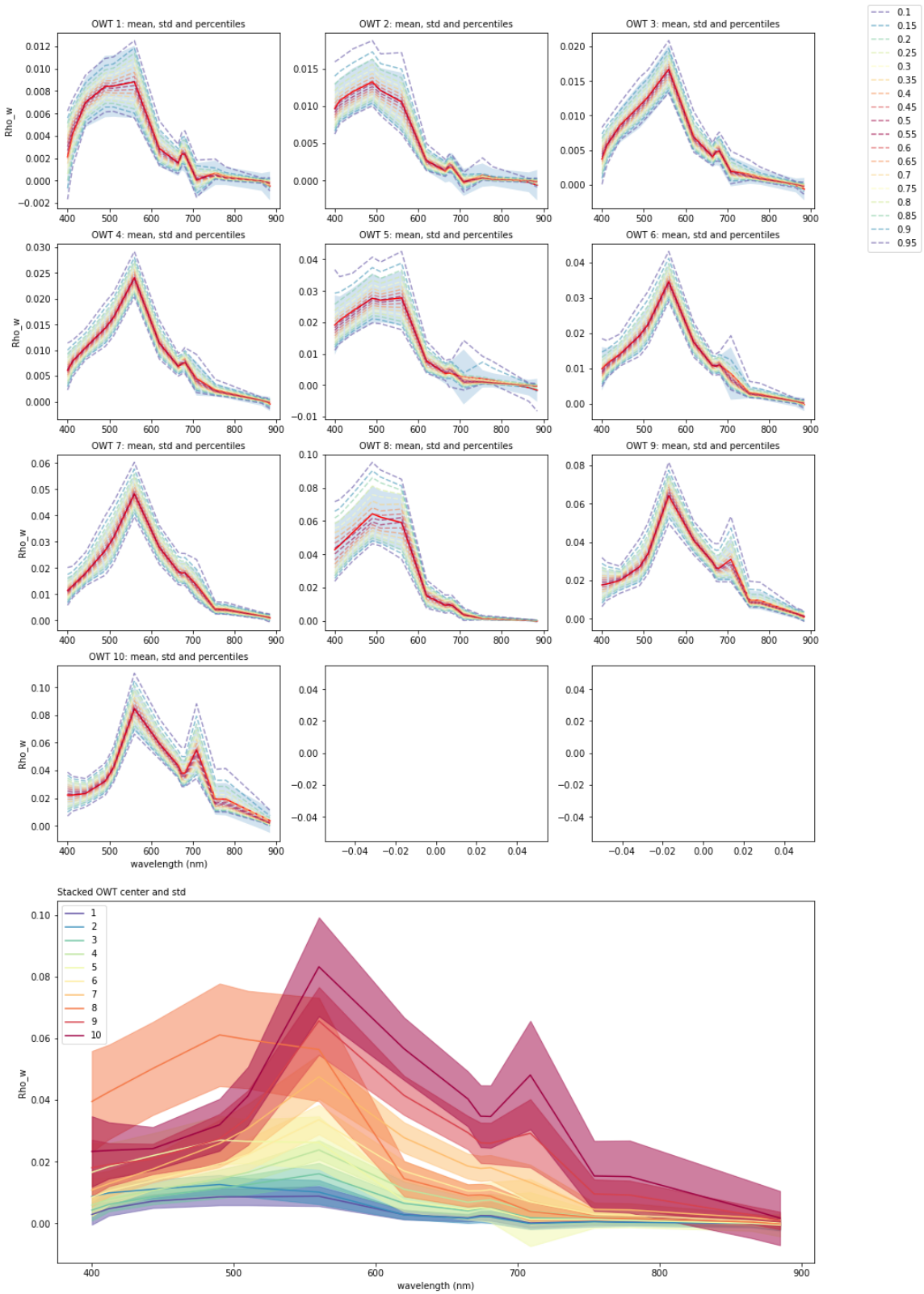

Figure 8: Sentinel-3 Optical Water Type (OWT) clusters created for the Danube Delta. Top panels show cluster centre spectra (red solid line) for spectra distribution for training data dominated by membership to given cluster, together with ± 1 std (gray shading) and distribution percentiles (hatched

lines, percentile indicated in legend). Lower panel shows centre spectrum ± 1 std for all classes overlain.

3.5.1.2 Distribution of clusters in case study region

Dominant OWT spatial distribution over the entire timeseries (Figure 9) shows that OWT 1 is primarily represented over the entire timeseries by the smaller ponds within the delta, while OWTs 2 through 5 primarily represent the Baltic Sea waters, as well as portions of these ponds. OWT 6 and 7 capture well the river plume waters. Razelm and Golovita waters are best captured by OWT 9 and Sinoe with OWT 10. Monthly analysis shows seasonal shifting between OWT 9 and 10 within Razelm, Golovita and Sinoe. Sometimes all three lakes vary coherently and sometimes they differ.

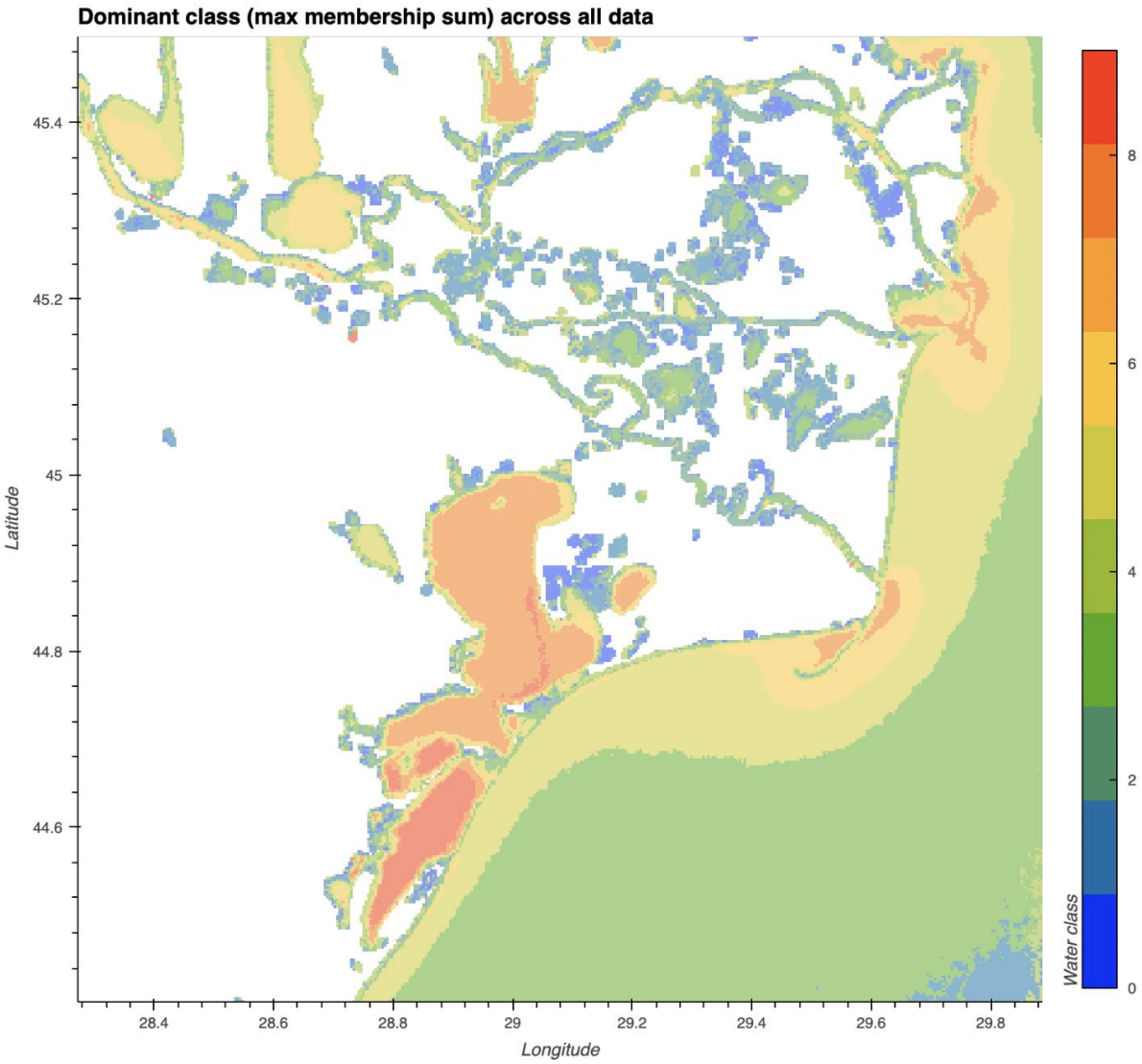

Figure 9: Dominant S3 OWT for each pixel within Danube Delta region over the entire timeseries (2016 to 2021), note that colour bar is 0 indexed.

3.5.2 Sentinel 2

3.5.2.1 Cluster reflectance spectra

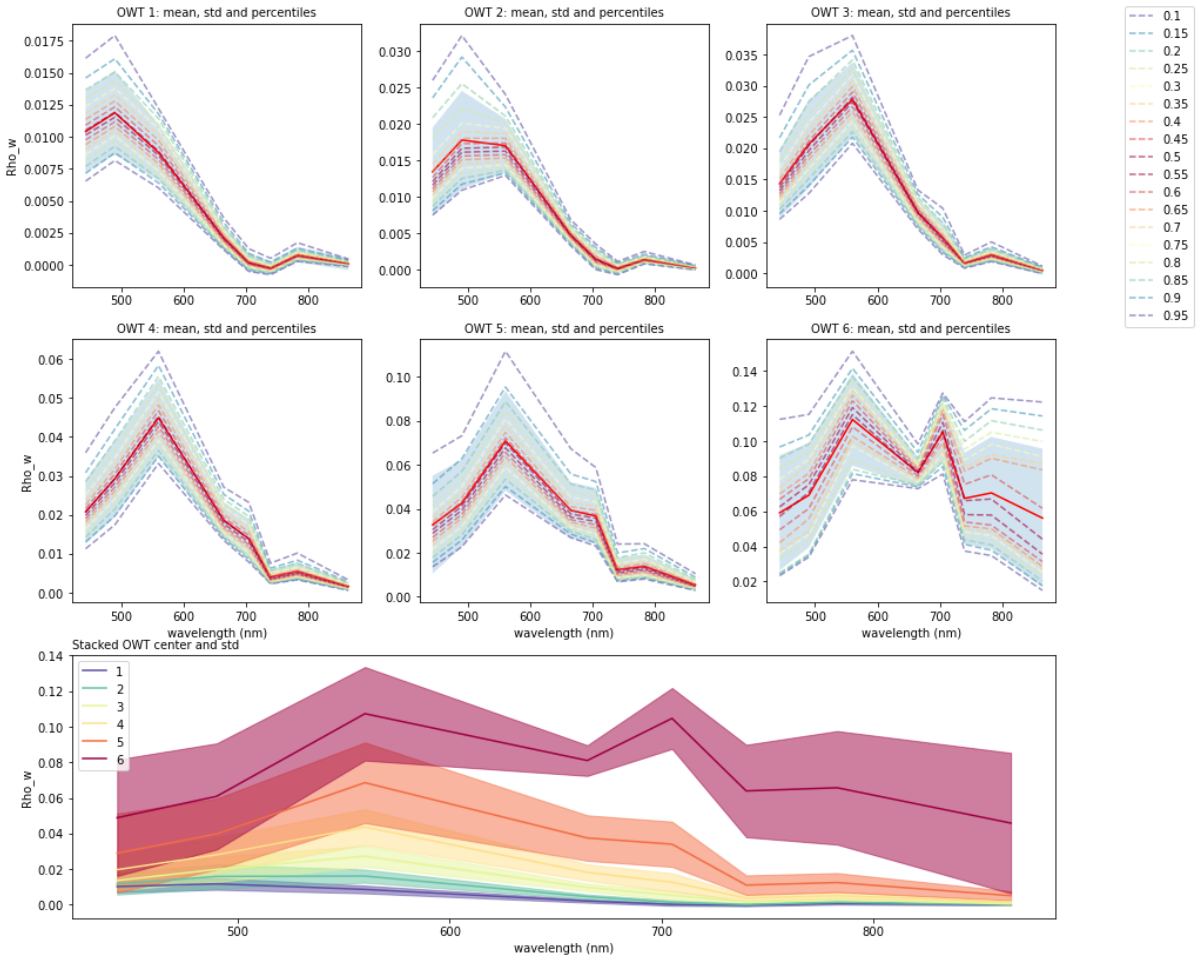

Figure 10: Sentinel-2 Optical water type (OWT) clusters created for the Danube Delta. Top panels show cluster centre spectra (red solid line) for spectra distribution for training data dominated by membership to given cluster, together with ± 1 std (gray shading) and distribution percentiles (hatched lines, percentile indicated in legend). Lower panel shows centre spectrum ± 1 std for all classes overlain.

3.5.2.2 Distribution of clusters in case study region

Dominant OWT spatial distribution over the entire timeseries (Figure 11) supports that OWT 1 and 2 primarily represent the Black Sea waters. OWT 3 captures well the waters along the coast of the lagoon complex, perhaps due to mixing with the river plume and Black Sea waters, as well as many of the smaller ponds within the delta. OWT 4 represents much of the river water and parts of the large freshwater lake on the Ukrainian side of the border, while OWT 5 captures well the southern portion of that same lake as well as much of the Razelm and Golovita lakes. OWT 6 is best represented within the Sinoe lake and southern portions of both Razelm and Golovita.

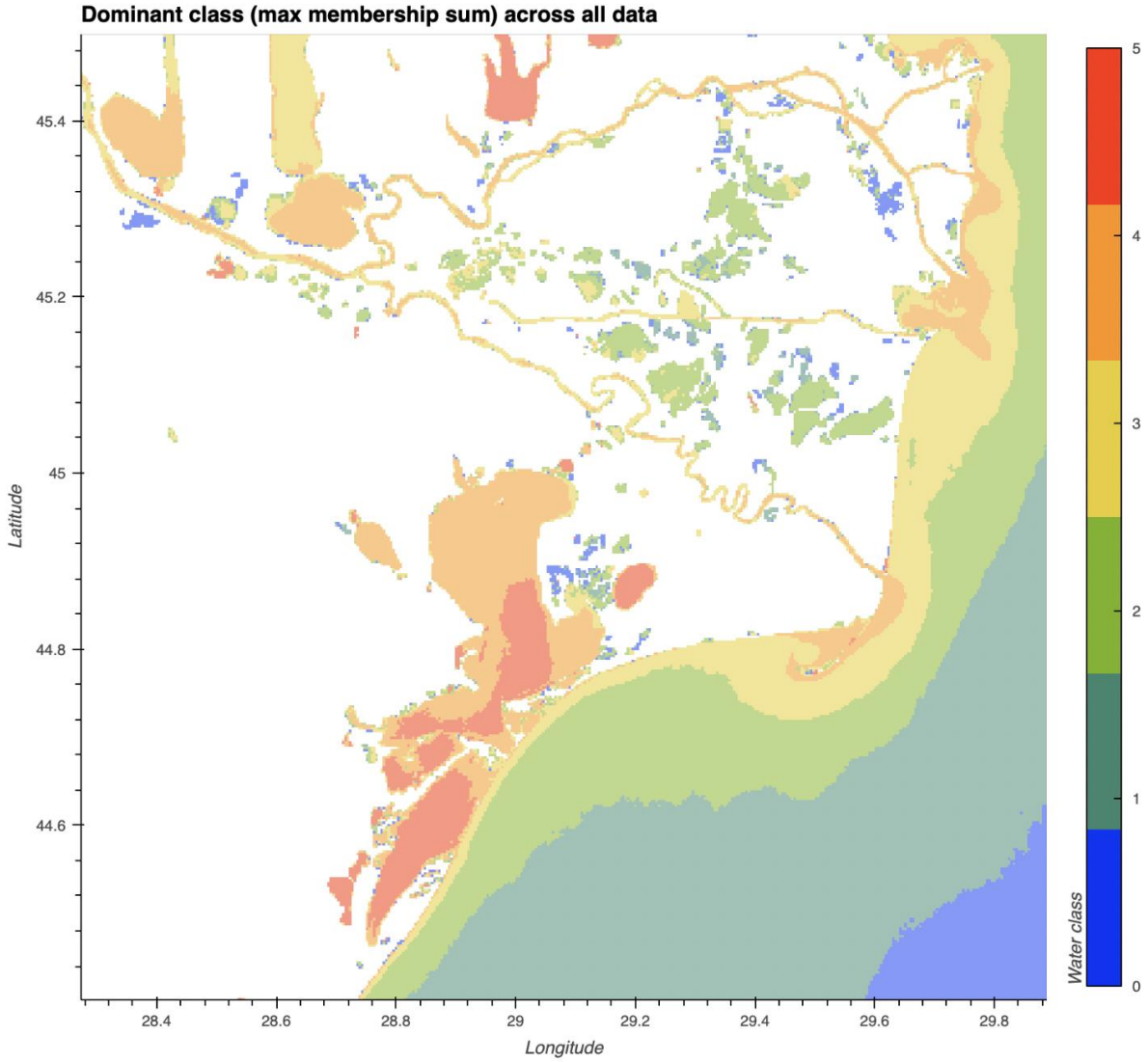

Figure 11: Dominant S2 OWT for each pixel within Danube Delta region over the entire timeseries (2016 to 2020), note that colour bar is 0 indexed.

3.6 Elbe Estuary

The regional extent and an example training data sample distribution are shown in Figure 12. Coverage of sampling points is roughly identical (allowing for resolution differences) for both S3 and S2 based clustering exercises.

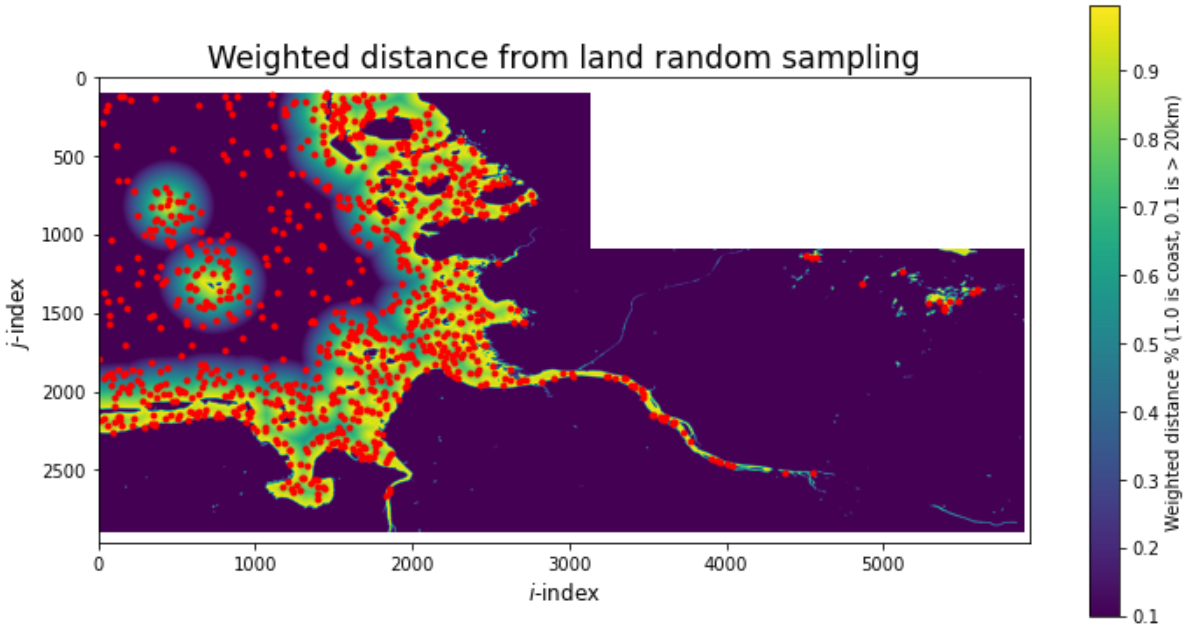

Figure 12: S2 Elbe region sample weighting map, background colour (dark blue to yellow) represents weighting using distance to coast. Example weighted random sample (red dots).

3.6.1 Sentinel 3

3.6.1.1 Cluster reflectance spectra

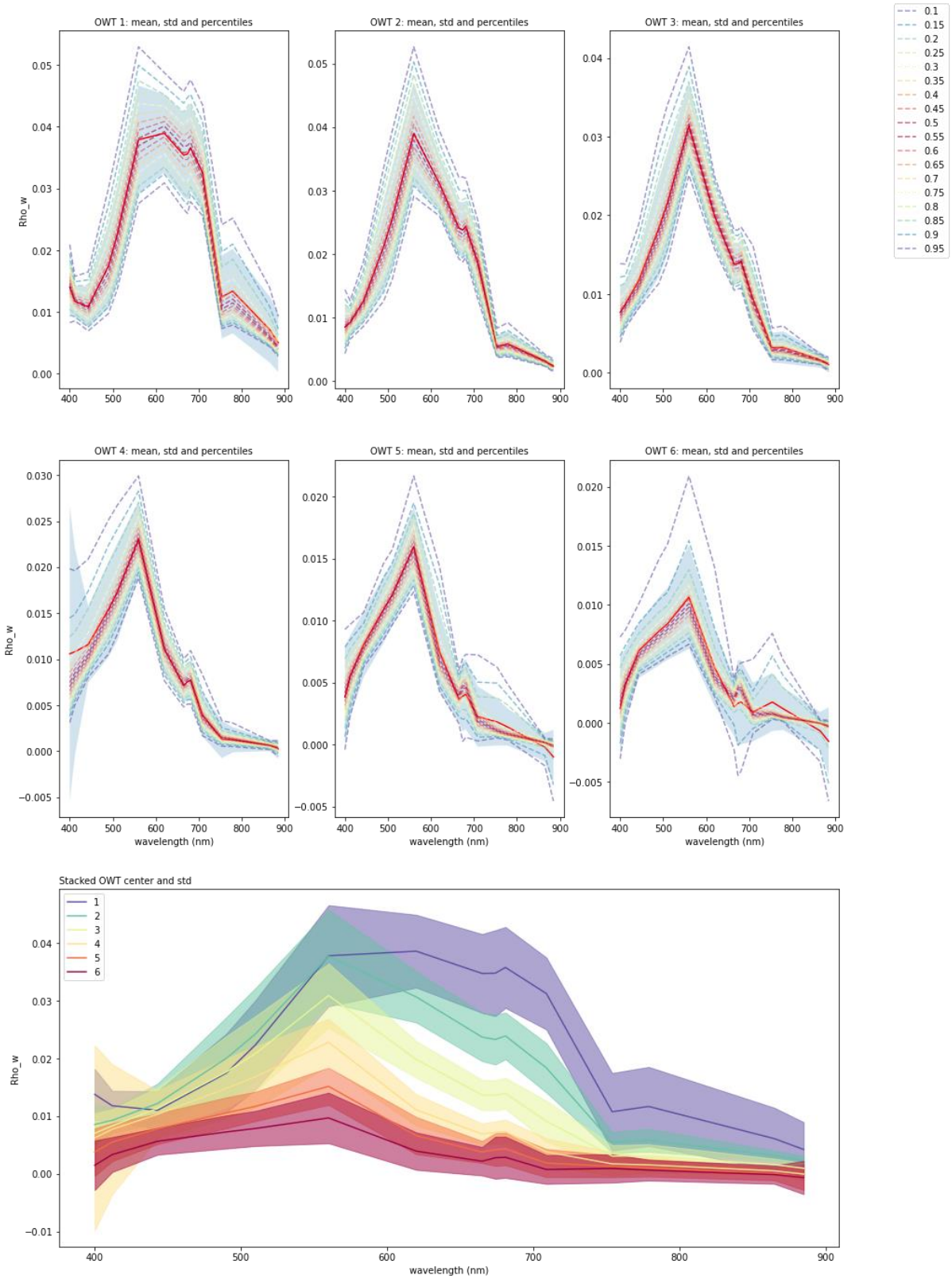

Figure 13: Sentinel-3 Optical water type (OWT) clusters created for the Elbe region. Top panels show cluster centre spectra (red solid line) for spectra distribution for training data dominated by membership to given cluster, together with ± 1 std (gray shading) and distribution percentiles (hatched

lines, percentile indicated in legend). Lower panel shows centre spectrum ± 1 std for all classes overlain.

3.6.1.2 Distribution of clusters in case study region

Note that OWT numbering is reversed for this region/sensor combination relative to other regions. Dominant OWT spatial distribution over the entire timeseries (Figure 14) shows that OWT 6 primarily represents clearer offshore waters, most consistently represented in the Baltic Sea waters, while OWT 4 and 5 capture better the North Sea waters. OWT 3 best represents the river plume mixing area and OWT 2 the intertidal areas. OWT 1 picks up the clear signal of the Turbidity Maximum Zone. The clearer OWTs, i.e., OWT 5 and 6, are also observed within the river around Hamburg and further upstream. Monthly analysis shows that OWT 1, representing in this case the more highly reflecting turbid waters, extends further outside of the Elbe during the months of December to March. The clearer water OWTs (5 and 6) are consistently present in the Baltic Sea waters and only appear in the North Sea waters for the spring and summer months (April to September).

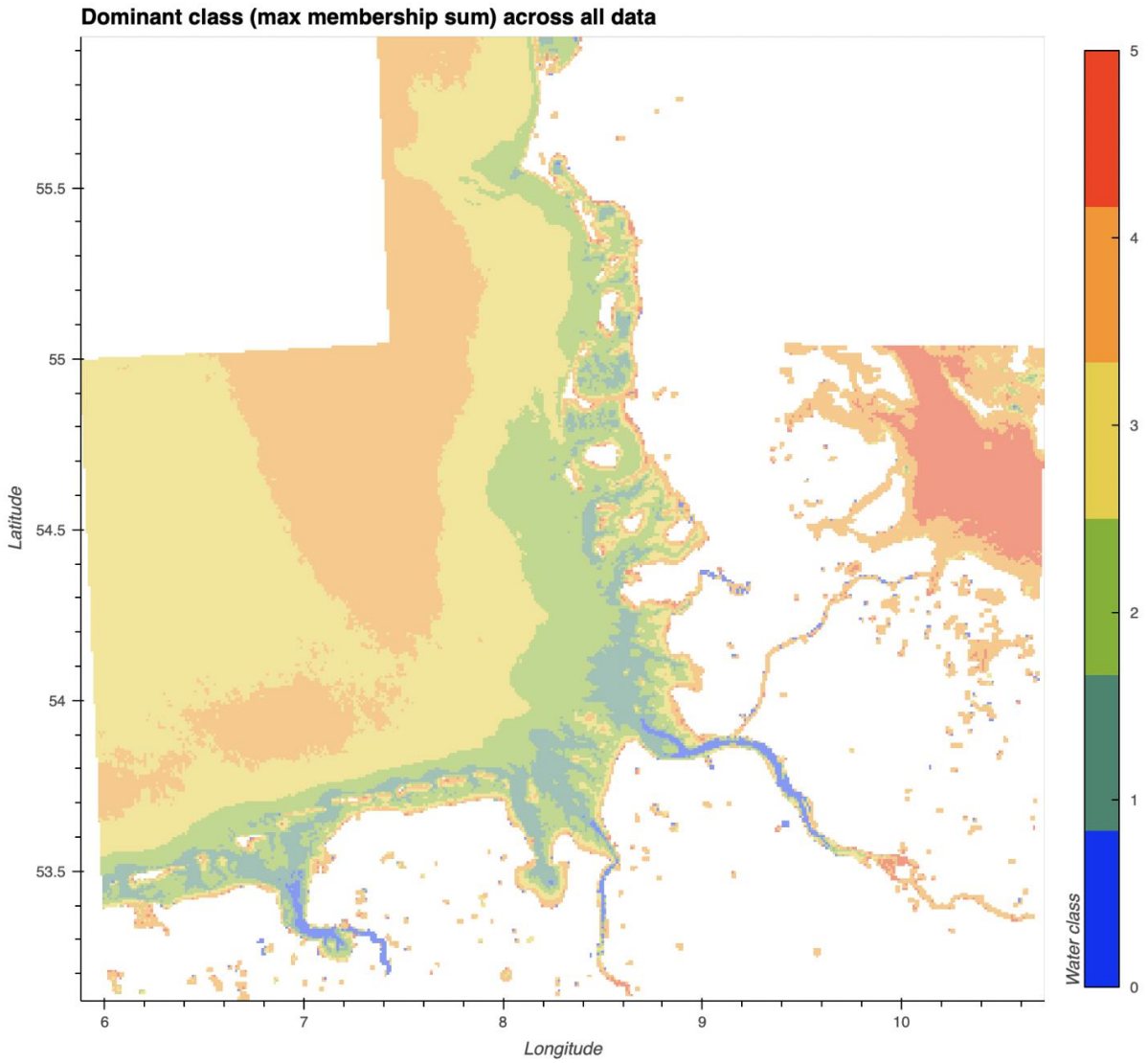

Figure 14: Dominant S3 OWT for each pixel within Elbe region over the entire timeseries (2016 to 2021), note that colour bar is 0 indexed.

3.6.2 Sentinel 2

3.6.2.1 Cluster reflectance spectra

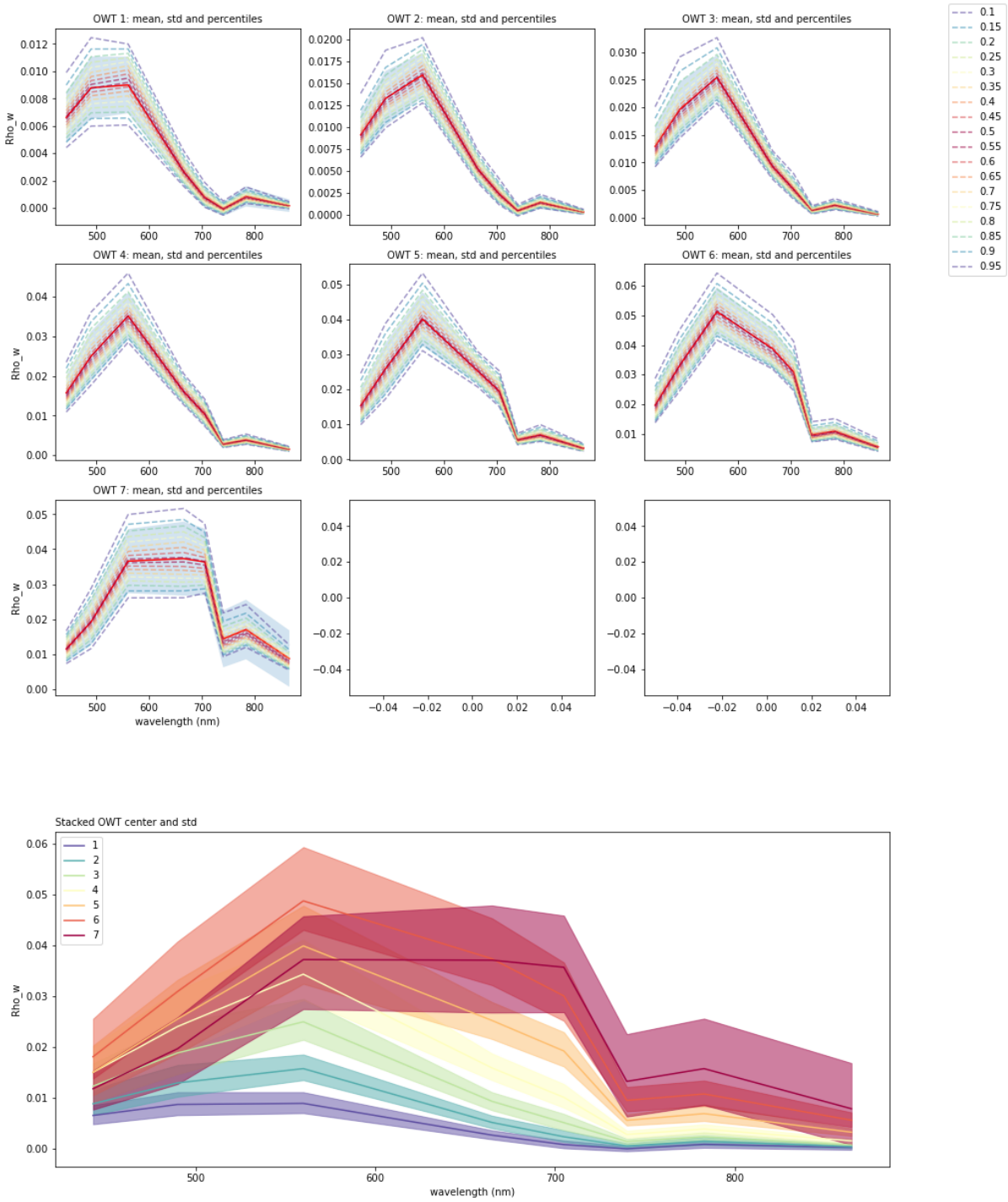

Figure 15: Sentinel-2 Optical water type (OWT) clusters created for the Elbe region. Top panels show cluster spectra (red line) for spectra distribution for training data dominated by membership to given cluster. Lower panel shows mean spectrum ± 1 std for all classes.

3.6.2.2 Distribution of clusters in case study region

Dominant OWT spatial distribution over the entire timeseries (Figure 16) supports that OWT 1 and 2 primarily represent the North Sea waters. OWT 3 and 4 capture well the mixed waters along the coastline but the Elbe river plume entering the North Sea is not sufficiently

present through the entire timeseries (it is likely variable in strength and location). OWT 5 & 6 represents various regions of the Wadden Sea and intertidal zones of the region. OWT 7 picks up the location of the Elbe Turbidity Maximum Zone, as well as other outflowing river mouth within the region. Monthly analysis shows that the extension of OWT 7 appears seasonal, reaching much further outside of the Elbe from December through May. This OWT is also present upstream of the tidal influenced region of the Elbe (taking this to be the Geesthacht weir) for the months April and May.

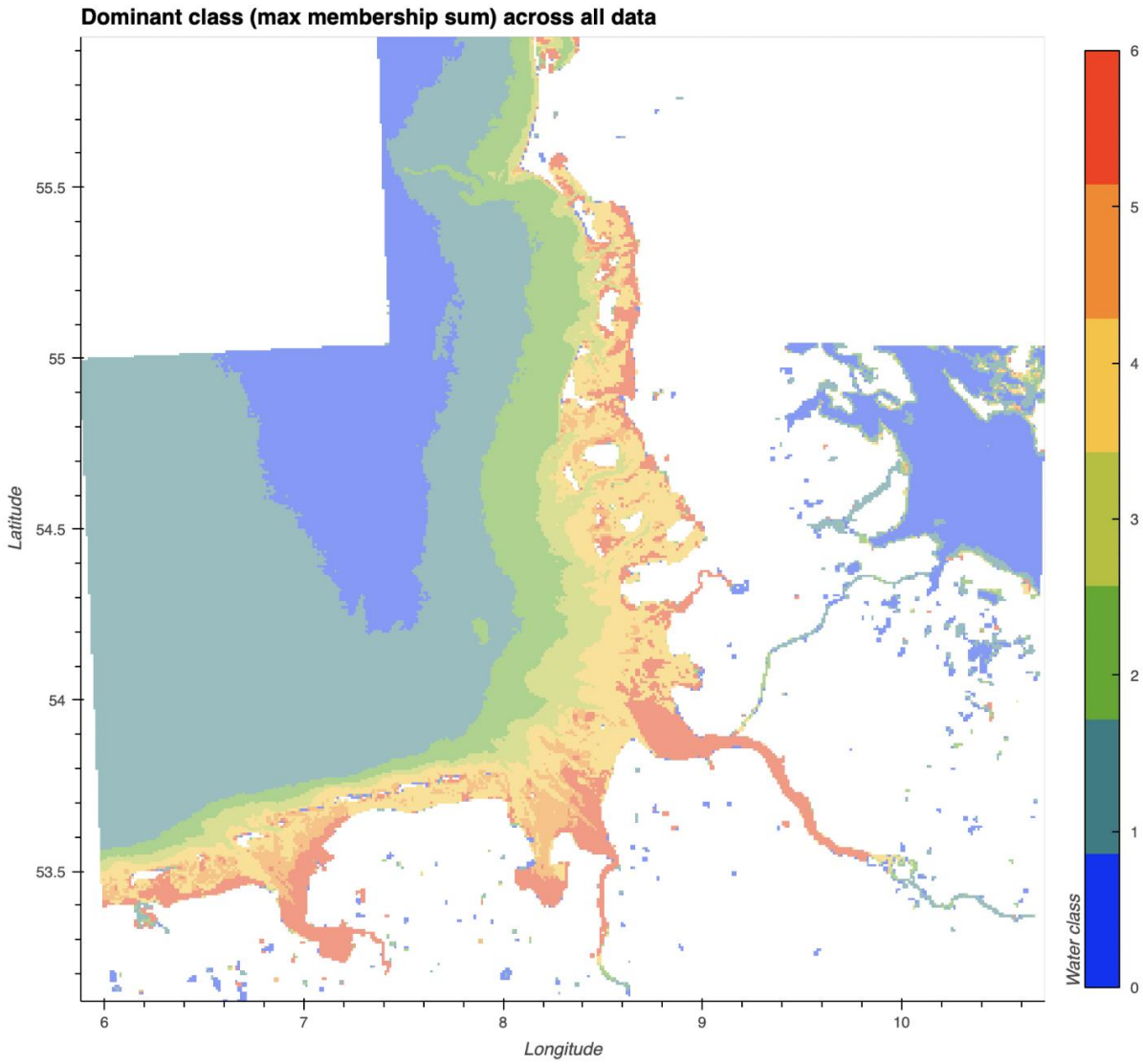

Figure 16: Dominant S2 OWT for each pixel within Elbe region over the entire timeseries (2016 to 2020), note that colour bar is 0 indexed.

3.7 Tagus Estuary

The regional extent and an example training data sample distribution are shown in Figure 17. Coverage of sampling points is roughly identical (allowing for resolution differences) for both S3 and S2 based clustering exercises.

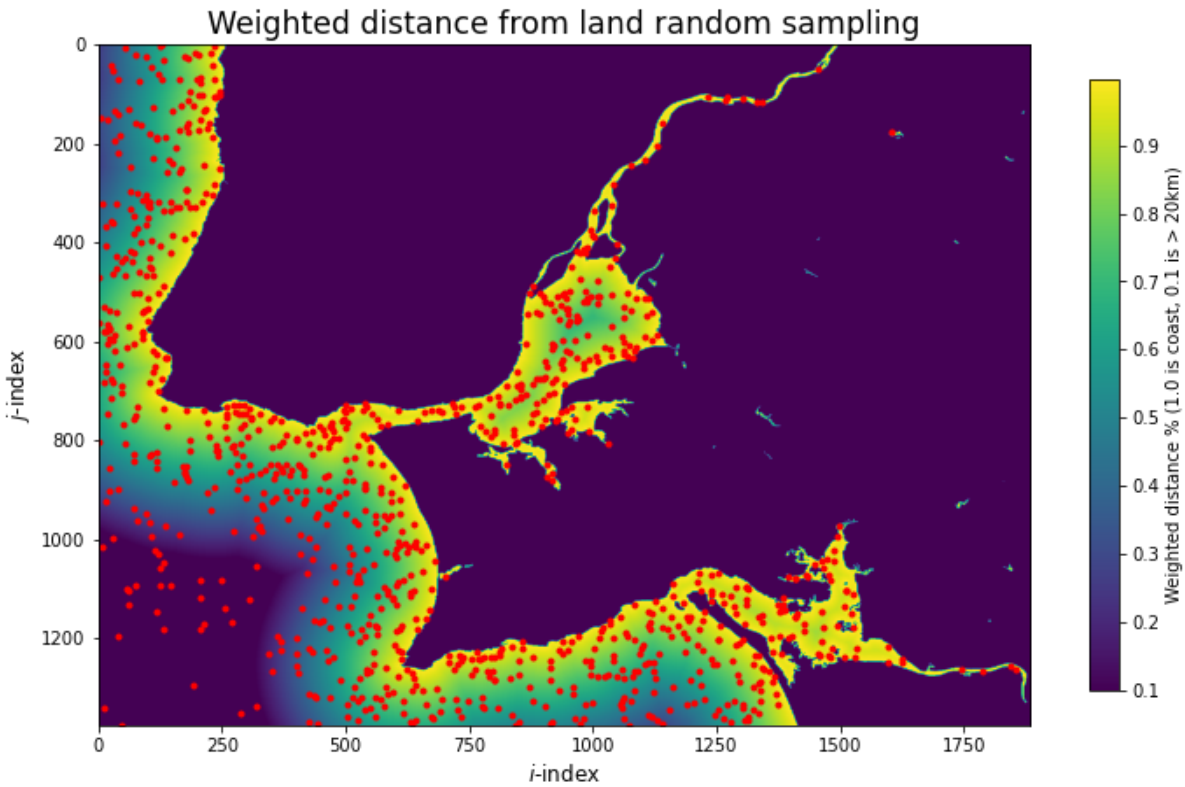

Figure 17: S2 Tagus region sample weighting map, background colour (dark blue to yellow) represents weighting using distance to coast. Example weighted random sample (red dots).

3.7.1 Sentinel 3

3.7.1.1 Cluster reflectance spectra

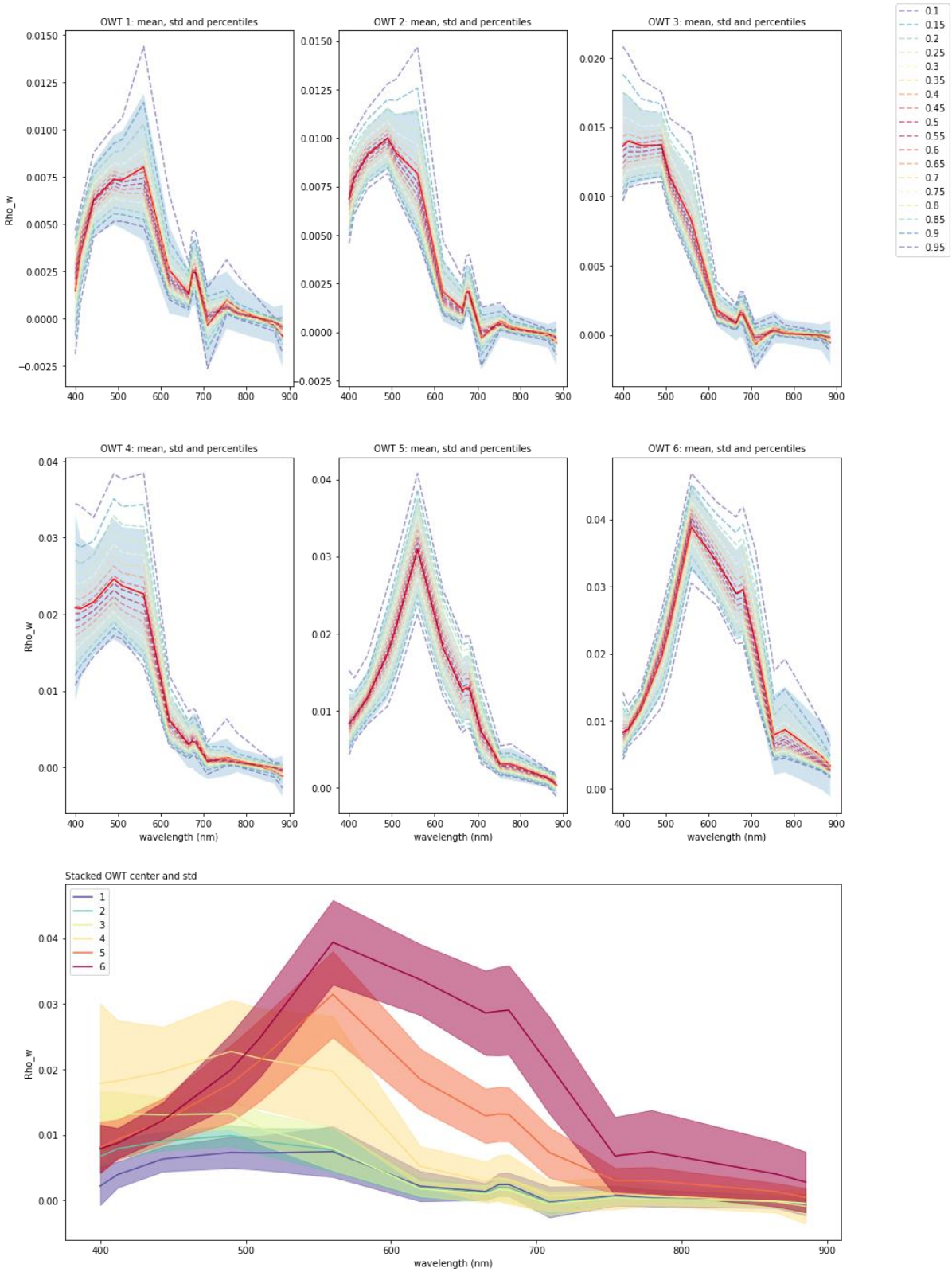

Figure 18: Sentinel-3 Optical water type (OWT) clusters created for the Tagus region. Top panels show cluster centre spectra (red solid line) for spectra distribution for training data dominated by membership to given cluster, together with ± 1 std (gray shading) and distribution percentiles (hatched

lines, percentile indicated in legend). Lower panel shows centre spectrum ± 1 std for all classes overlain.

3.7.1.2 Distribution of clusters in case study region

Dominant OWT spatial distribution over the entire timeseries (Figure 19) supports that OWT 2 primarily represents the Atlantic waters, but OWT 1 is only consistently observed in some of the estuary edges over the entire time series (which could likely be an artefact from the atmospheric correction or adjacency effects). OWT 3 and 4 capture well the coastal waters outside the estuary. OWT 5 best represents the lower estuary and the primary river plume emerging from the estuary. OWT 6 primarily occurs in the upper estuary. Monthly analysis shows that OWT 6 occurs over a larger region of the upper estuary for the months November through February. This also coincides with the OWT 5 river plume extending further away from shore. The coastal mixed water OWTs 3 & 4 are much less present for the summer months June to September.

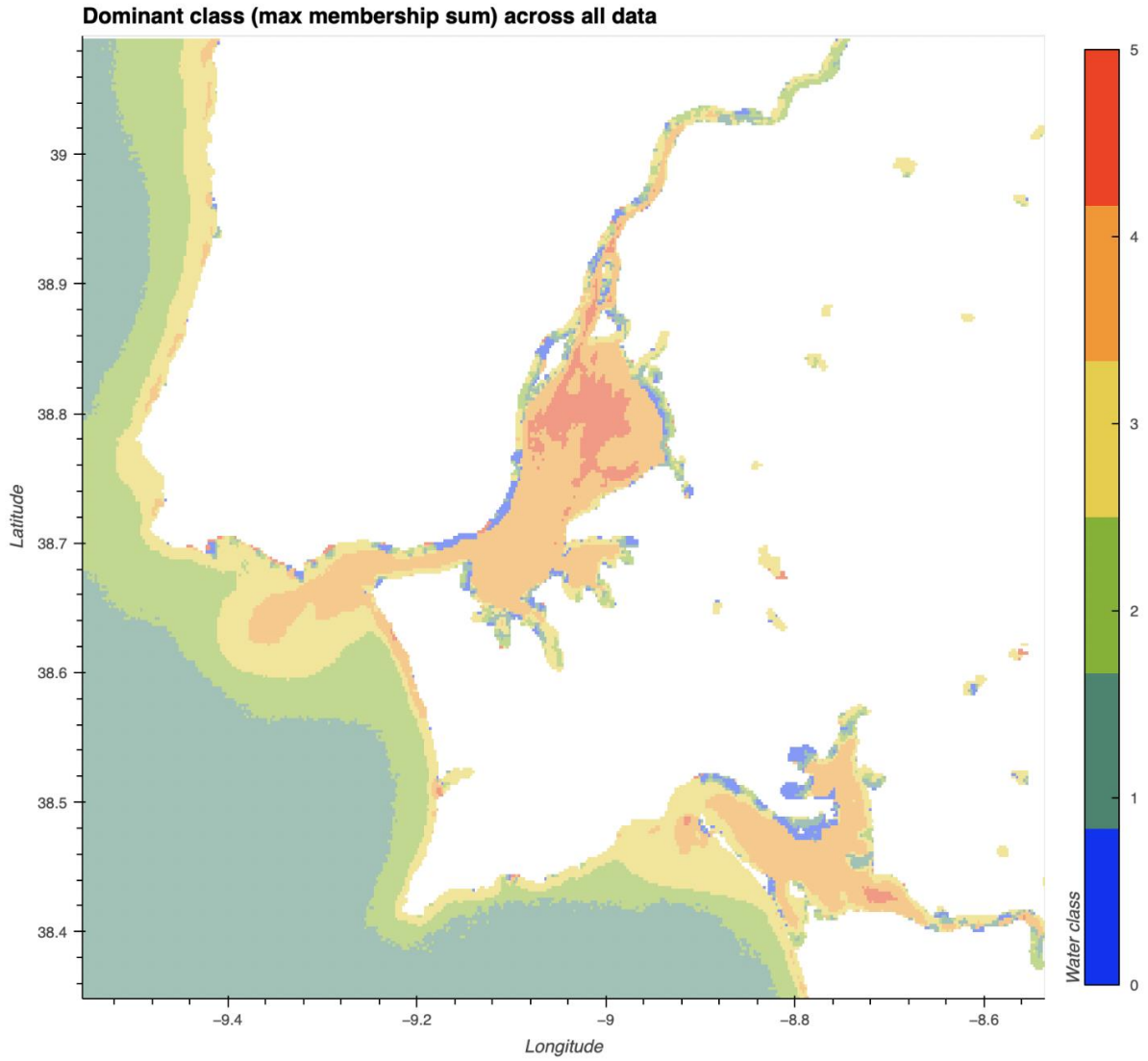

Figure 19: Dominant S3 OWT for each pixel within Tagus region over the entire timeseries (2016 to 2021), note that colour bar is 0 indexed.

3.7.2 Sentinel 2

3.7.2.1 Cluster reflectance spectra

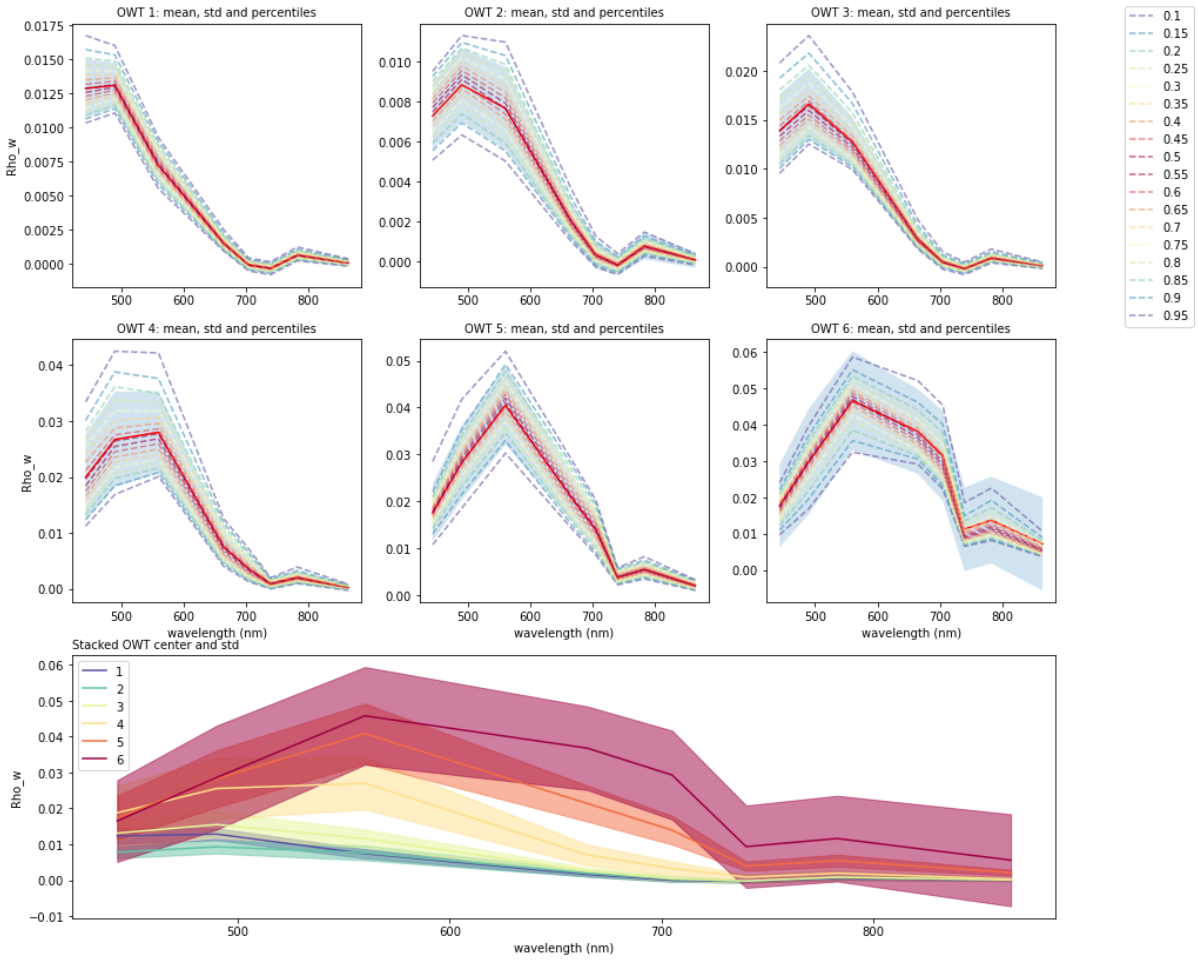

Figure 20: Sentinel-2 Optical water type (OWT) clusters created for the Tagus region. Top panels show cluster centre spectra (red solid line) for spectra distribution for training data dominated by membership to given cluster, together with ± 1 std (gray shading) and distribution percentiles (hatched lines, percentile indicated in legend). Lower panel shows centre spectrum ± 1 std for all classes overlain.

3.7.2.2 Distribution of clusters in case study region

Dominant OWT spatial distribution over the entire timeseries (Figure 21) supports that OWT 1 and 2 primarily represent the Atlantic waters, while OWT 3 captures well the waters along the coastline. OWT 4 represents the mid-estuary waters and OWT 5 those of the upper estuary.

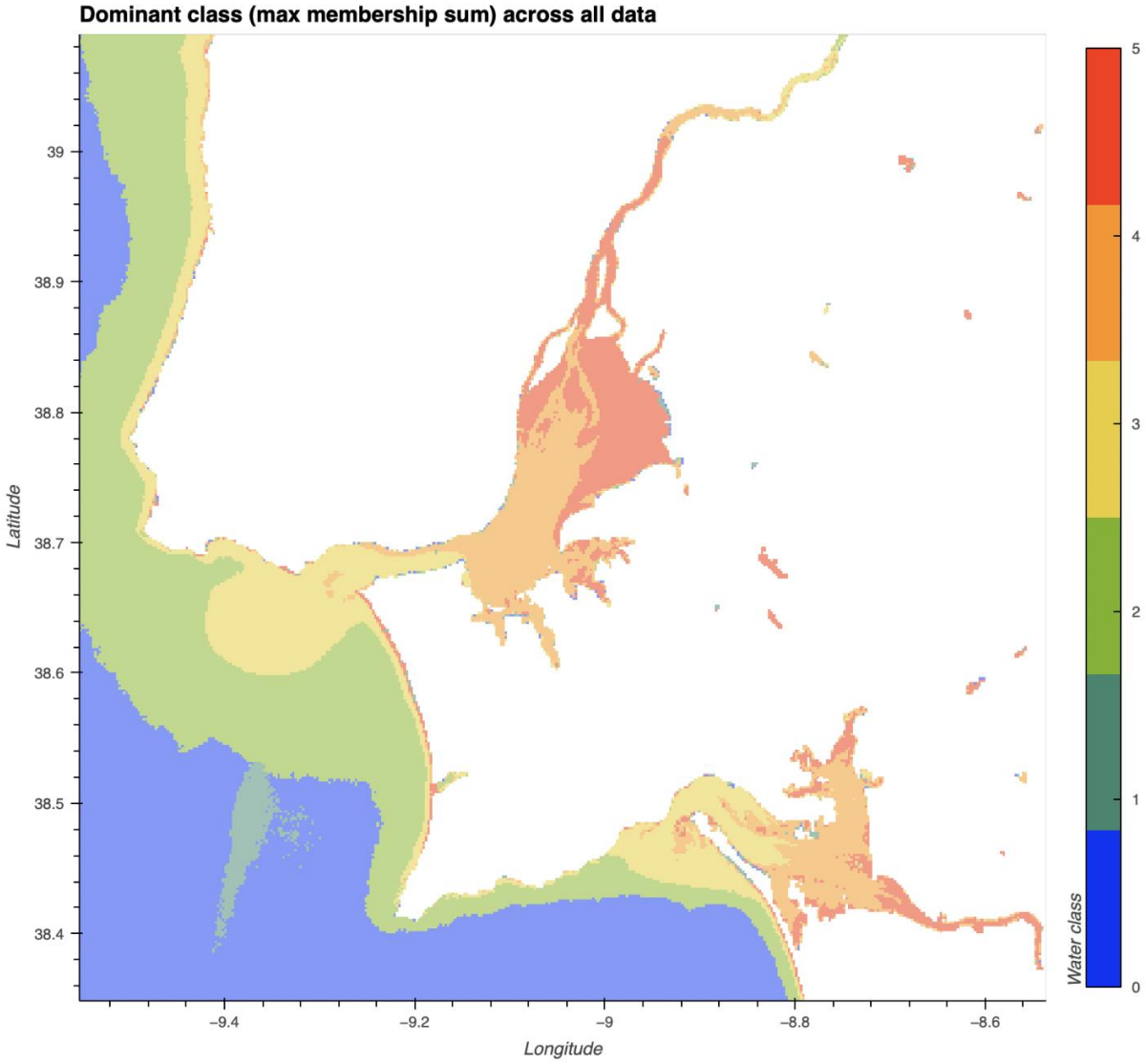

Figure 21: Dominant S2 OWT for each pixel within Tagus region over the entire timeseries (2016 to 2020), note that colour bar is 0 indexed.

3.8 Tamar River/Plymouth Sound

The regional extent and an example training data sample distribution are shown in Figure 1722. Coverage of sampling points is roughly identical (allowing for resolution differences) for both S3 and S2 based clustering exercises, with the exception of the Eddystone lighthouse being recognised as land in the S2 data but not in the S3 data.

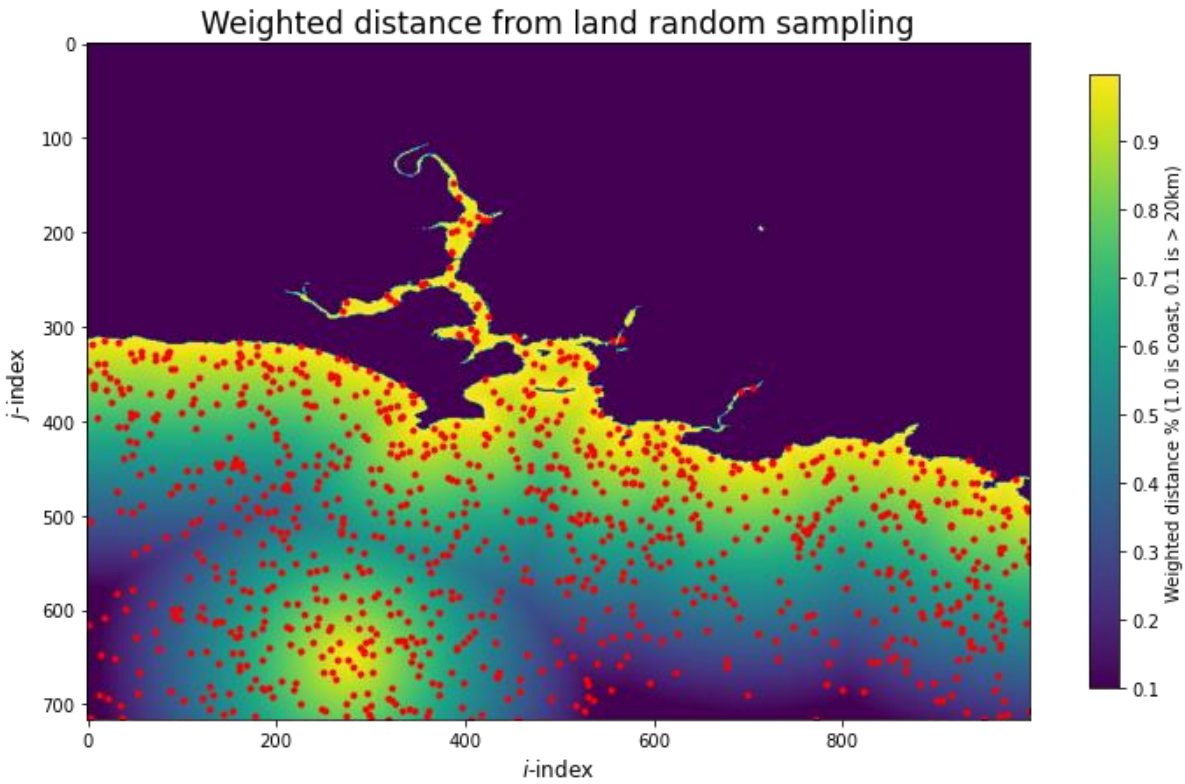

Figure 22: S2 Tamar region sample weighting map, background colour (dark blue to yellow) represents weighting using distance to coast. Example weighted random sample (red dots).

3.8.1 Sentinel 3

3.8.1.1 Cluster reflectance spectra

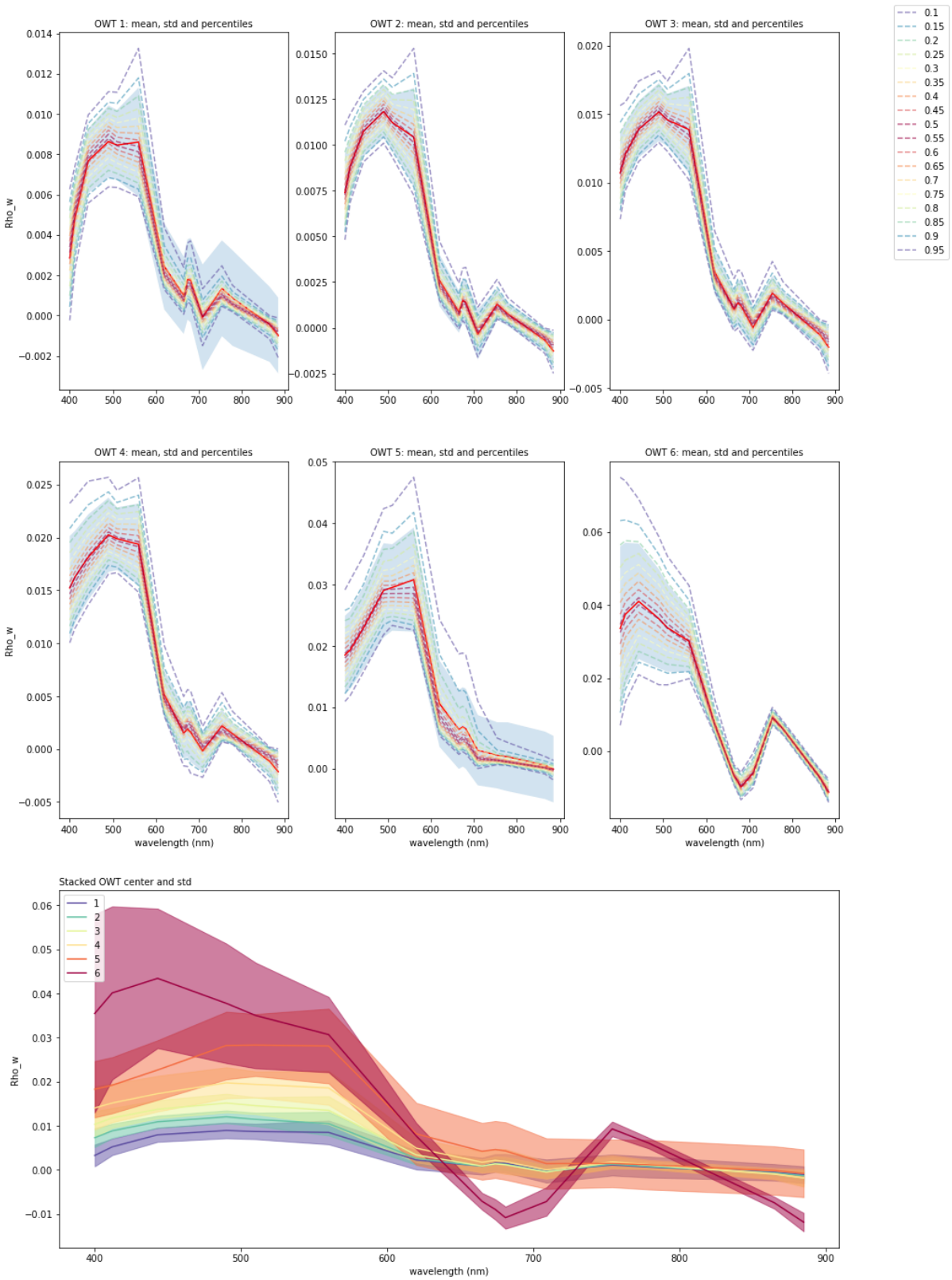

Figure 23: Sentinel-3 Optical water type (OWT) clusters created for the Tamar region. Top panels show cluster centre spectra (red solid line) for spectra distribution for training data dominated by membership to given cluster, together with ± 1 std (gray shading) and distribution percentiles (hatched

lines, percentile indicated in legend). Lower panel shows centre spectrum ± 1 std for all classes overlain.

3.8.1.2 Distribution of clusters in case study region

Dominant OWT spatial distribution over the entire timeseries (Figure 24) reveal spatial patterns that are better interpretable than for the single day observations, similar to that observed for the MSI analysis. OWT 1 and 2 are not well represented in the full timeseries data, with the exception of the occasional estuary edge pixels which are more likely artefacts from factors such as the atmospheric correction or adjacency effects. OWT 3 primarily represents the Western Channel waters. OWT 4 captures well the coastal waters and up into the estuary, while OWT 5 the near coast and the central shipping channel portion of the Tamar estuary. OWT 6 best represents the closest to shore regions and only the very upper portions of some of the estuaries.

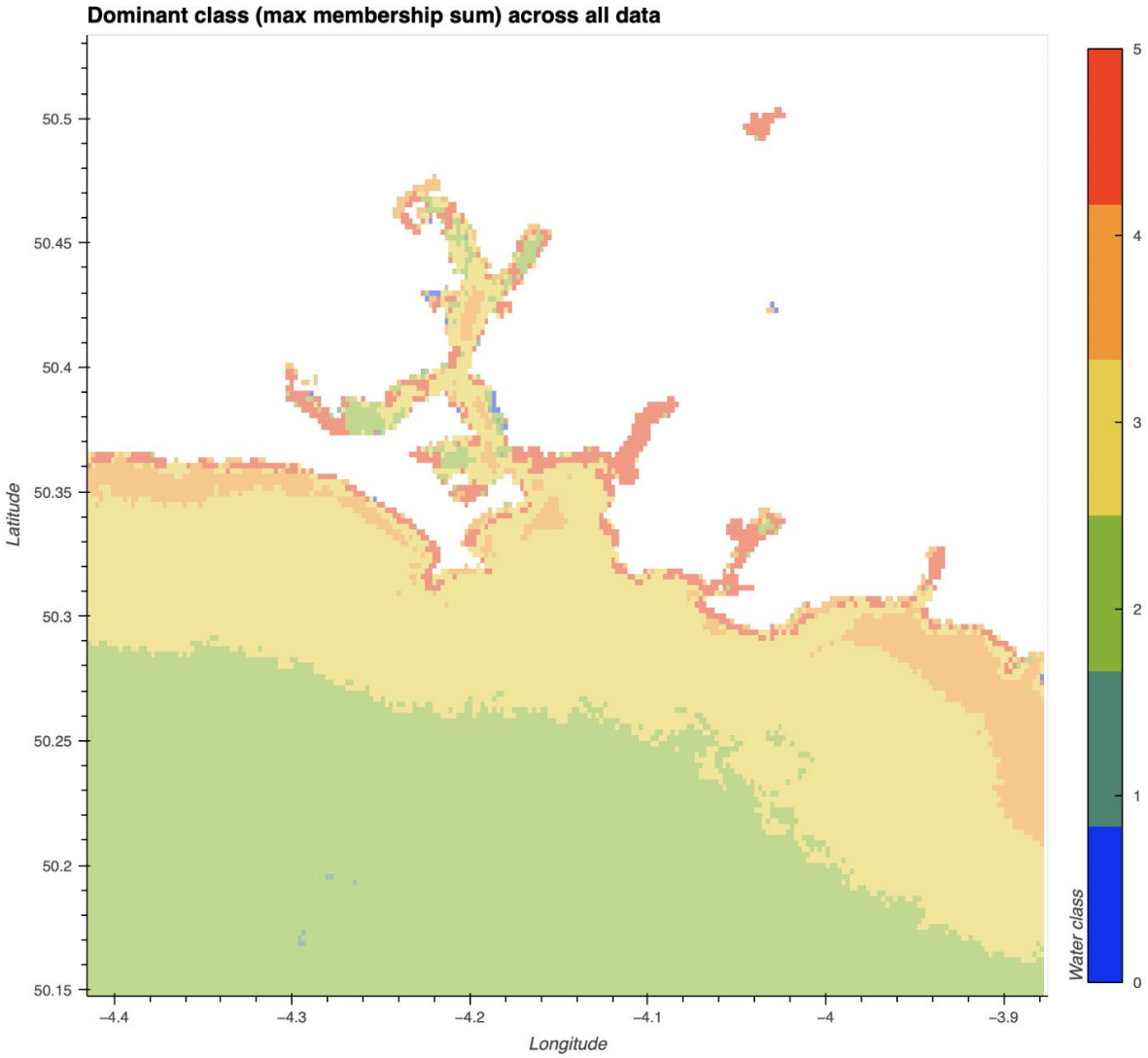

Figure 24: Dominant S3 OWT for each pixel within Tagus region over the entire timeseries (2016 to 2021), note that colour bar is 0 indexed.

3.8.2 Sentinel 2

3.8.2.1 Cluster reflectance spectra

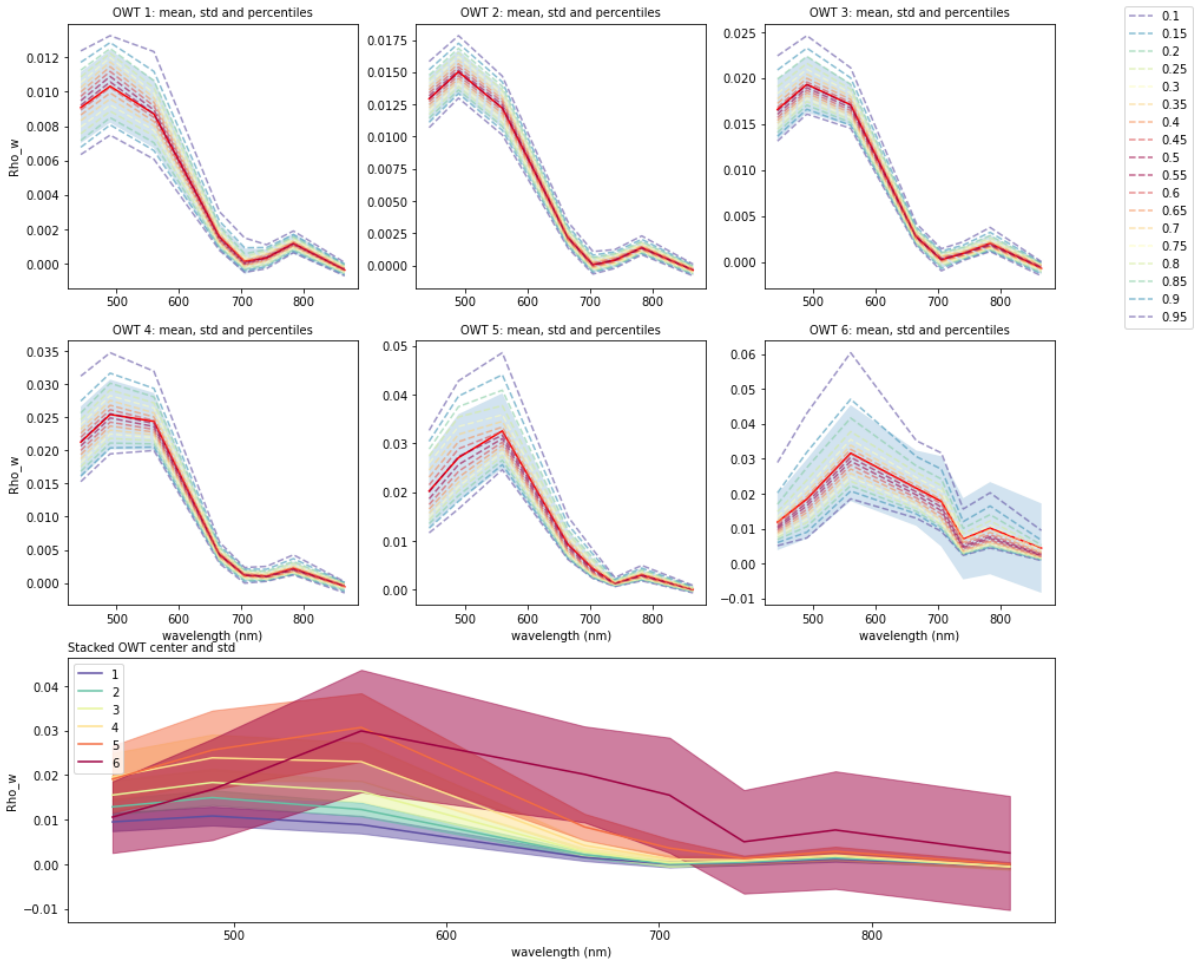

Figure 25: Sentinel-2 Optical water type (OWT) clusters created for the Tamar region. Top panels show cluster centre spectra (red solid line) for spectra distribution for training data dominated by membership to given cluster, together with ± 1 std (gray shading) and distribution percentiles (hatched lines, percentile indicated in legend). Lower panel shows centre spectrum ± 1 std for all classes overlain.

3.8.2.2 Distribution of clusters in case study region

Dominant OWT spatial distribution over the entire timeseries (Figure 26) reveals spatial patterns that are better interpretable than for the single day observations. OWT 1 and 2 primarily represent the Western Channel waters. OWT 3 and 4 capture well the coastal waters, while OWT 5 the near coast (potentially with bottom reflectance) and the lower portion of the Tamar estuary. OWT 6 best represents the upper portions of the estuaries. OWT 1 does appear in some edge areas of the estuary, which may represent errors from the atmospheric correction or adjacency effects.

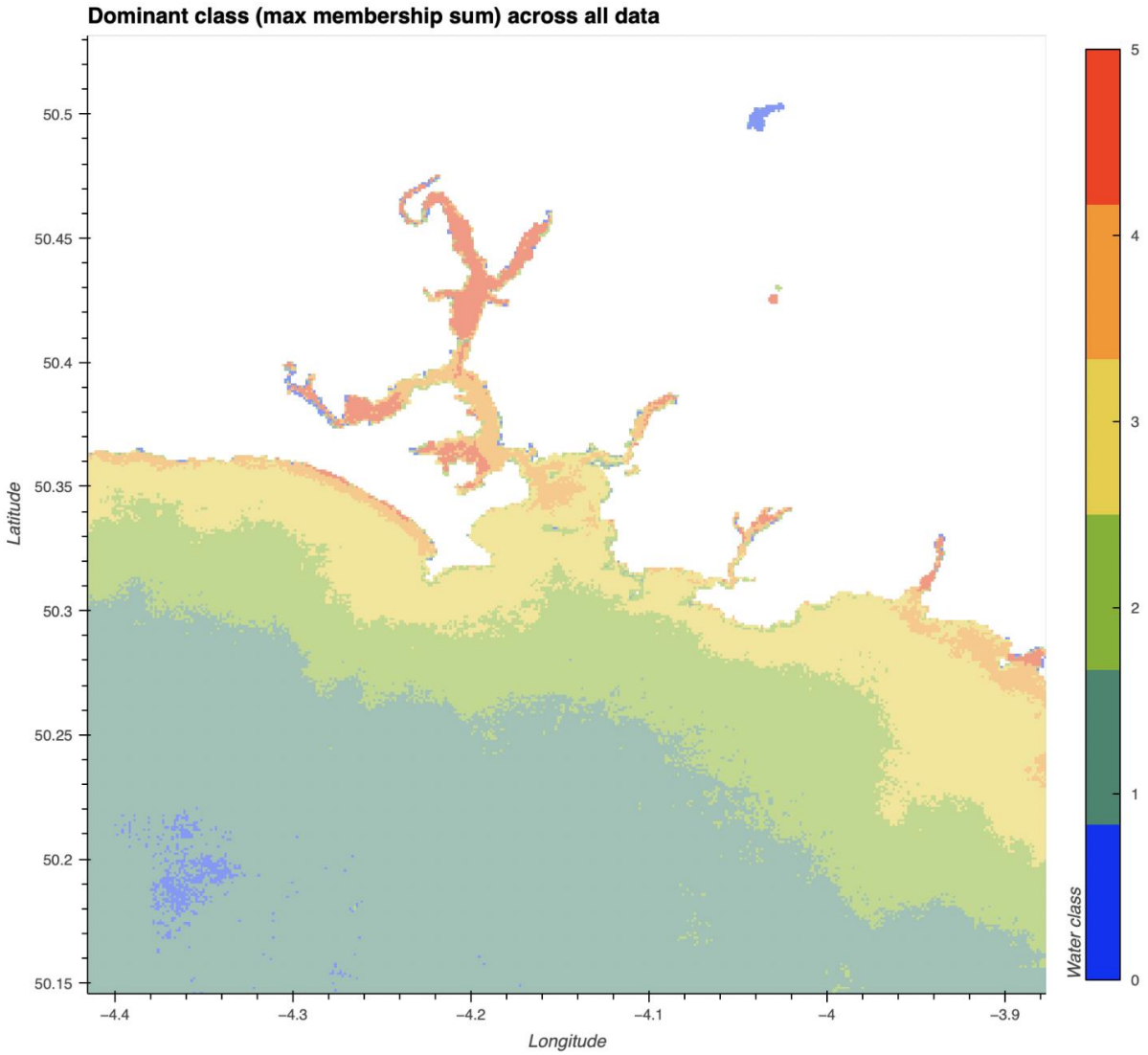

Figure 26: Dominant S2 OWT for each pixel within Tagus region over the entire timeseries (2016 to 2020), note that colour bar is 0 indexed.

3.9 Venice

The regional extent and an example training data sample distribution are shown in Figure 27. Coverage of sampling points is roughly identical (allowing for resolution differences) for both S3 and S2 based clustering exercises, with the exception of the Acqua Alta observation platform being recognised as land in the S2 data but not in the S3 data.

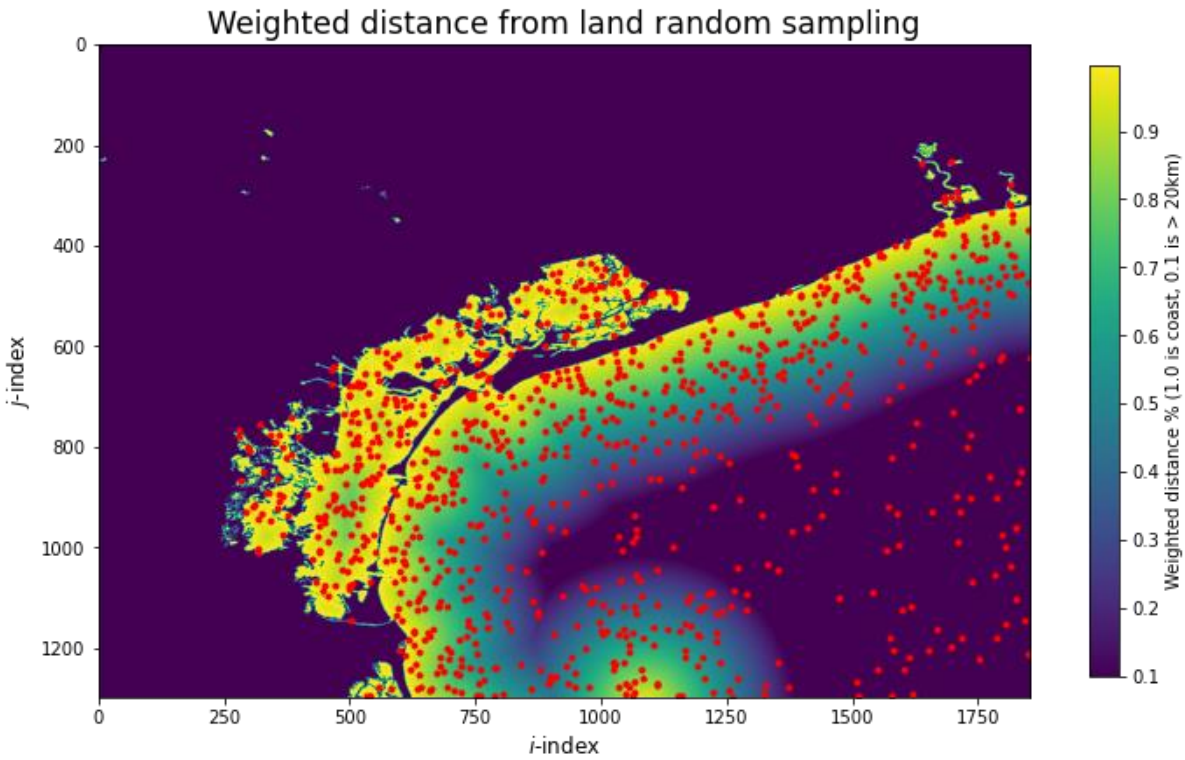

Figure 27: S2 Venice region sample weighting map, background colour (dark blue to yellow) represents weighting using distance to coast. Example weighted random sample (red dots).

3.9.1 Sentinel 3

3.9.1.1 Cluster reflectance spectra

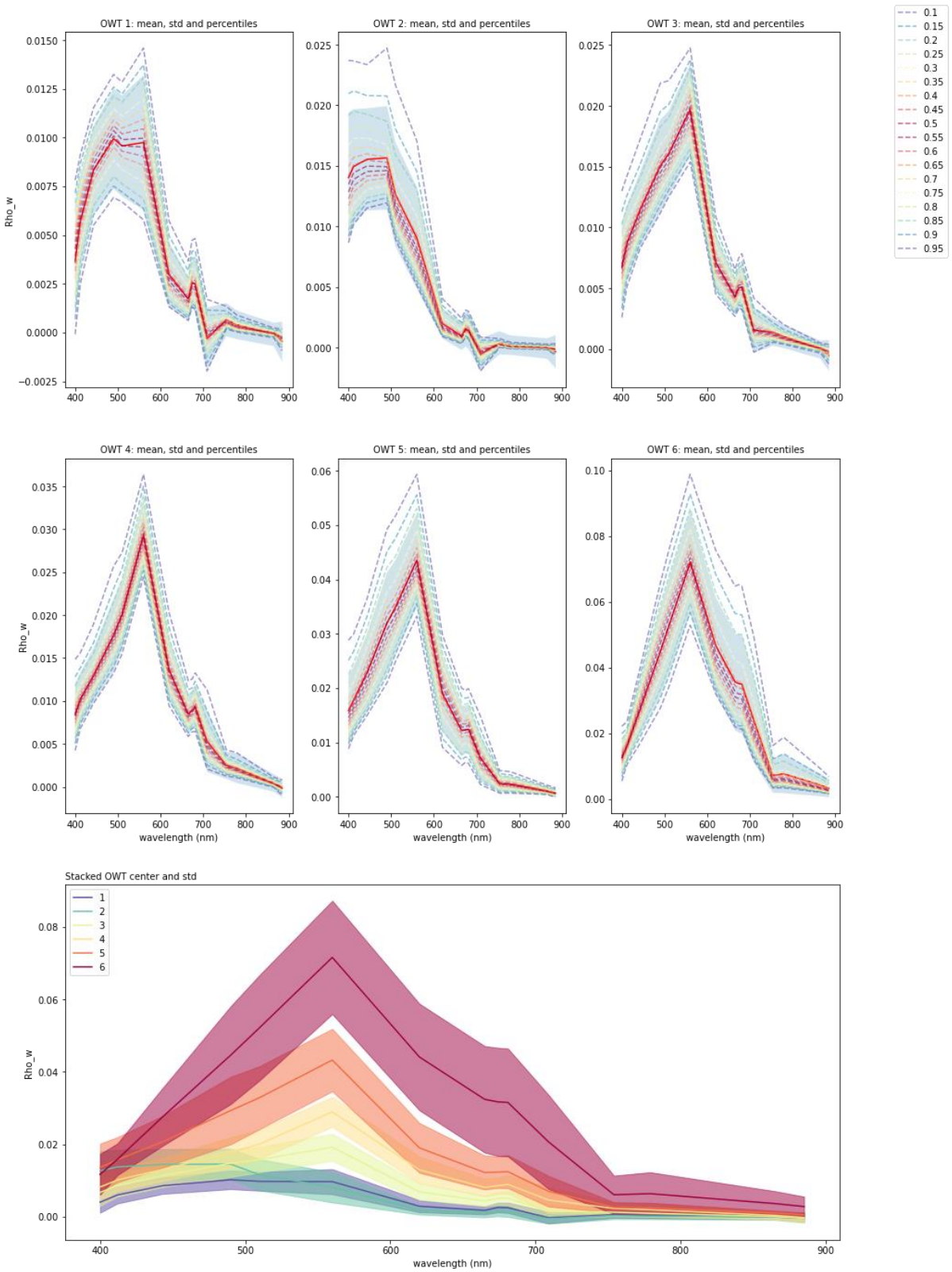

Figure 28: Sentinel-3 Optical water type (OWT) clusters created for the Venice region. Top panels show cluster centre spectra (red solid line) for spectra distribution for training data dominated by membership to given cluster, together with ± 1 std (gray shading) and distribution percentiles (hatched

lines, percentile indicated in legend). Lower panel shows centre spectrum ± 1 std for all classes overlain.

3.9.1.2 Distribution of clusters in case study region

Dominant OWT spatial distribution over the entire timeseries (Figure 29) supports that OWT 2 and 3 primarily represent the Northern Adriatic Sea waters. OWT 1 does not show up as a sufficiently consistent representative of the offshore waters to appear in the full timeseries analysis. This OWT is present in some edge coastal areas within lagoons and river areas, which are most likely due to artefacts from the atmospheric correction or adjacency effects. OWT 4 and 5 capture well the combined river plume waters flowing along the coastline past the Venice lagoon, as well as covering much of the southern portion of the lagoon. OWT 6 is picked up within the northern portion of the Venice lagoon, some areas of the Po Delta, and most strongly in the Valli di Comacchio.

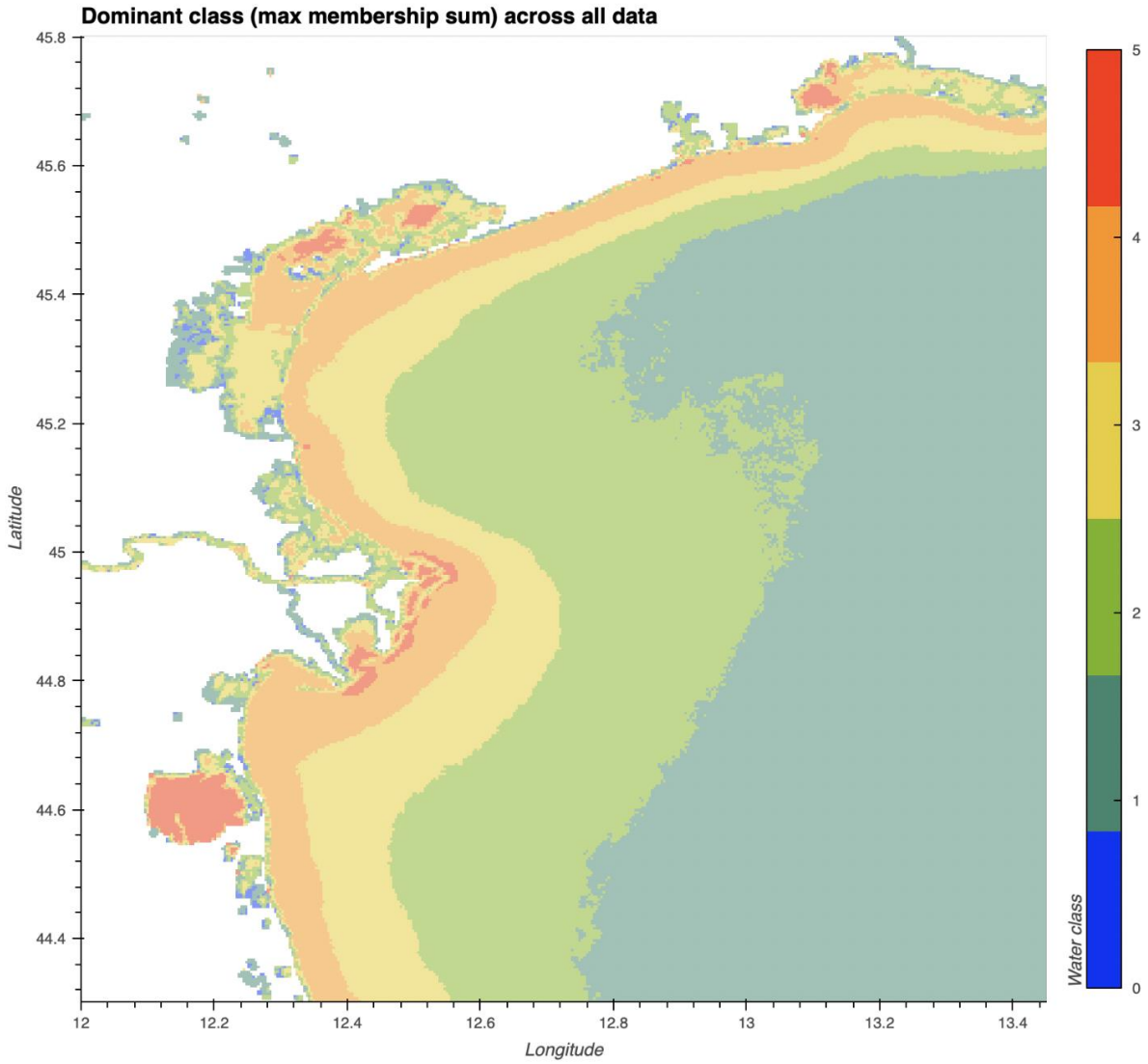

Figure 29: Dominant S3 OWT for each pixel within Tagus region over the entire timeseries (2016 to 2021), note that colour bar is 0 indexed.

3.9.2 Sentinel 2

3.9.2.1 Cluster reflectance spectra

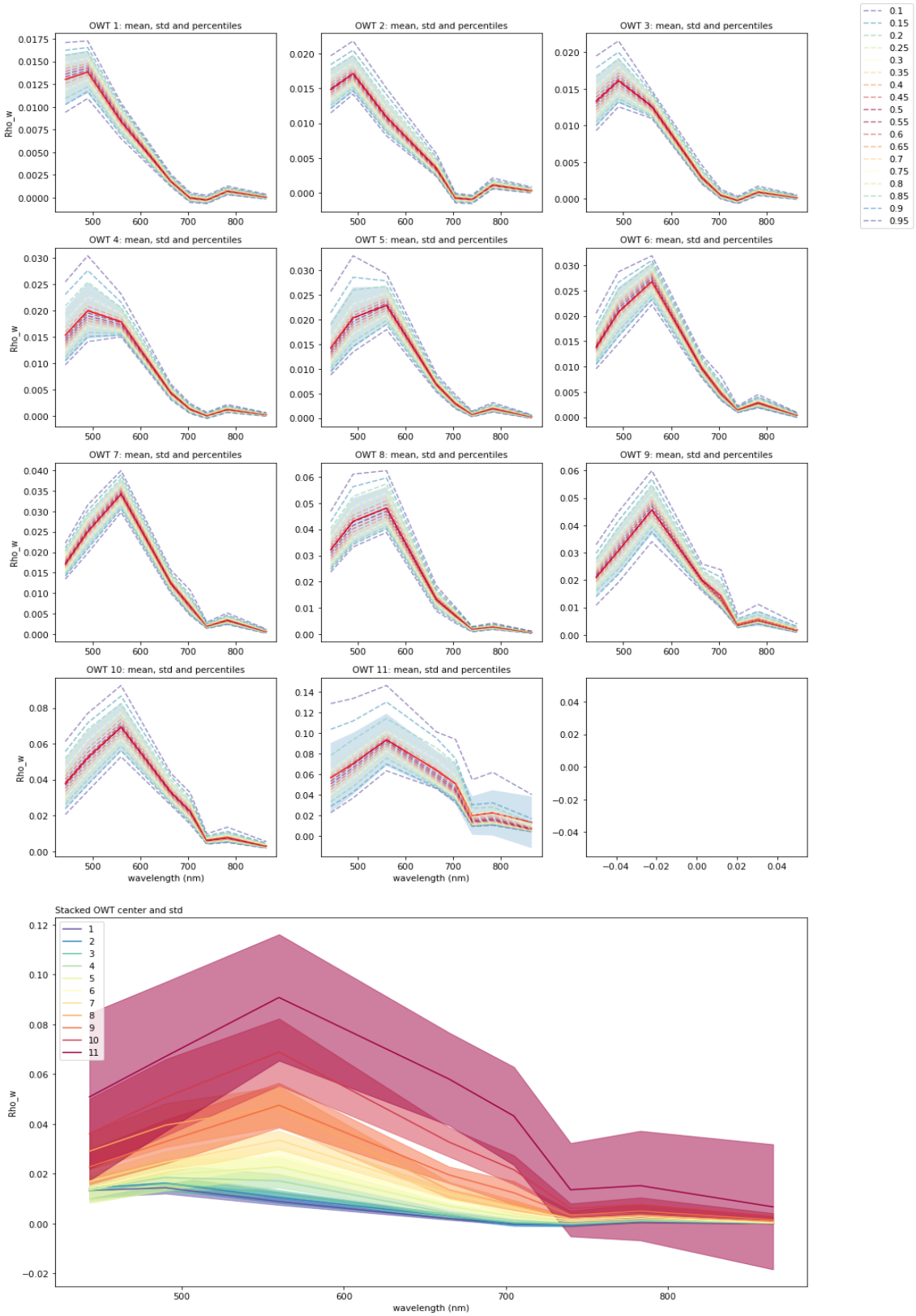

Figure 30: Sentinel-2 Optical water type (OWT) clusters created for the Venice region. Top panels show cluster centre spectra (red solid line) for spectra distribution for training data dominated by membership to given cluster, together with ± 1 std (gray shading) and distribution percentiles (hatched

lines, percentile indicated in legend). Lower panel shows centre spectrum ± 1 std for all classes overlain.

3.9.2.2 Distribution of clusters in case study region

Dominant OWT spatial distribution over the entire timeseries (Figure 31) provides some more interesting patterns than those observed for the single day analysis. OWTs 1 to 4 primarily represent the Northern Adriatic Sea waters. OWTs 5 to 8 capture well the combined river plume flowing along the coastline past the Venice lagoon, joining with the Po River plume to the South, with OWT 7 best representing the lower middle of the Venice lagoon over the entire timeseries. OWT 9 represents well the waters flowing out of the Po Delta and combined with OWT 10 captures the northerly section of the Venice lagoon. OWT 11 remains a rarely occurring and scattered water type within the Venice lagoon but seems to capture well some portions of the Po Delta lagoon system and the Valli di Comacchio further to the South.

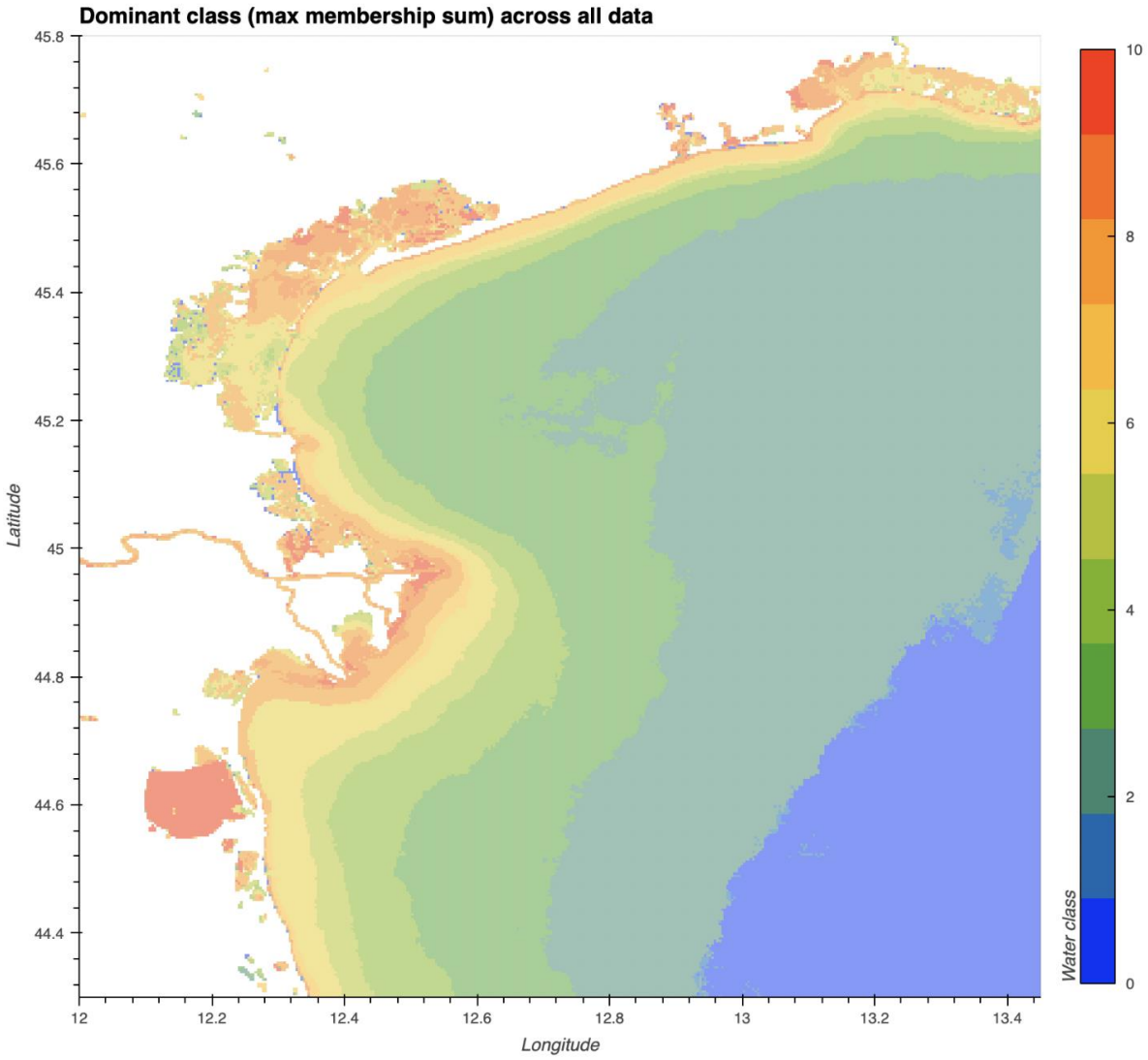

Figure 31: Dominant S2 OWT for each pixel within Tagus region over the entire timeseries (2016 to 2021), note that colour bar is 0 indexed.

4 Pan-Regional Clusters

For both S2 and S3 pan-regional clustering exercises, we merged the six sample sets of 100,000 samples into a single dataset of 600,000 samples (giving each region equal weighting).

Membership distribution resulting from an initial cluster optimisation exercise on this dataset produced broad classes that were well distributed across all six study sites, but some important areas (such as the Razelm, Golovita and Zmeica lakes by the Danube Delta) were not well represented. Furthermore, based on feedback from study site leads on the site-specific clustering results, one sees that some site-specific variability has been lost in this initial global clustering. The cluster grid search optimisation resulted with only 6 representative clusters, which seem few when considering some of the variability better captured with site-specific results. This spurred further analysis that aimed to better retain site-specific variability in the global cluster set, achieved using a cluster grouping and semi-supervised clustering approach.

We performed a cluster set comparison across all the site-specific cluster sets to identify groups of similar spectral signature clusters or clusters which are unique across all study sites. Both parametric (Welch's t-test, since reflectance data within a cluster is log normal distributed but still indicates variance heterogeneity) and non-parametric methods were explored: only the non-parametric results are presented here for brevity. Cluster sets were compared using the Adjusted Rand Index (ARI) which compares cluster membership similarity and dissimilarity between different cluster set results. ARI values less than or equal to 0 indicate that the two data cluster results do not agree on any pair of points, while an ARI value of 1 indicates that cluster sets are exactly the same. The clusters that were deemed to be similar were averaged to give a mean spectra per grouping and used as an initialisation for a fresh clustering exercise.

4.1 Sentinel 3

4.1.1 Cluster reflectance spectra

Pairwise comparisons between particular clusters for site-specific cluster sets were performed following sample separation by dominant OWT (samples were either class under examination or other). The resulting pairwise comparisons are presented in Table 1, which has been coloured using a threshold of ARI = 0.35 into "similar" (green) and "dissimilar" (red) clusters between cluster sets.

From this comparison, we see that many of the cluster spectra from the Danube Delta/Razelm-Sinoe Lake Complex overlap with cluster spectra from other study sites, while those from the Curonian Lagoon and Tamar Estuary are generally more unique.

The cluster grouping resulted in 18 clusters that were then used in a semi-supervised clustering exercise (as a prior to the cluster optimisation step) and resulted in the clusters shown in Figure 32.

Table 1: Adjusted Rand Index (ARI) for each S3 cluster pair. Comparison between cluster sets from all six CERTO focus regions (note clusters zero indexed). Threshold value of ARI=0.35 indicates clusters pairs that are similar (green) or dissimilar (red). Cluster pair comparisons were not made within a study site cluster set (grey). Cluster sets have been reordered by ARI > 0.35 threshold, so highlight similarity and difference along diagonal.

	CS_curonian_1	CS_danube_0	CS_elbe_5	CS_tagus_0	CS_tamar_0	CS_venice_0	CS_tagus_1	CS_tamar_1	CS_curonian_3	CS_danube_1	CS_elbe_4	CS_tagus_2	CS_tamar_2	CS_venice_1	CS_danube_3	CS_danube_4	CS_elbe_3	CS_tagus_3	CS_tamar_3	CS_danube_6	CS_elbe_1	CS_tamar_4	CS_venice_4	CS_curonian_9	CS_curonian_10	CS_danube_8	CS_danube_9	CS_elbe_0	CS_tagus_5	CS_venice_5	CS_curonian_4	CS_curonian_5	CS_danube_5	CS_elbe_2	CS_tagus_4	CS_venice_3	CS_curonian_8	CS_danube_7	CS_danube_2	CS_venice_2	CS_curonian_0	CS_curonian_2	CS_curonian_6	CS_curonian_7	CS_tamar_5
CS_danube_0	0.35		0.71	0.60	0.70	0.59	0.01	-0.09	0.01		-0.05	-0.09	-0.08	-0.09			-0.09	-0.09	-0.08		-0.07	-0.04	-0.09	-0.06	-0.05			-0.06	-0.09	-0.07	-0.07	-0.07		-0.09	-0.09	-0.10	-0.04			-0.09	0.25	0.13	-0.03	-0.03	-0.01
CS_curonian_1		0.35	0.55	0.13	0.23	0.29	0.23	0.09		0.07	-0.07	-0.02	-0.08	-0.01	-0.08	-0.08	-0.11	-0.10	-0.09	-0.08	-0.08	-0.06	-0.09			-0.06	-0.02	-0.06	-0.10	-0.08			-0.09	-0.10	-0.10	-0.10		-0.01	-0.09	-0.10					0.00
CS_elbe_5	0.55	0.71		0.46	0.55	0.61	0.09	-0.01	0.09	-0.02		-0.05	-0.07	-0.05	-0.06	-0.06		-0.07	-0.07	-0.06		0.01	-0.07	-0.05	-0.04	-0.04	-0.02		-0.07	-0.06	-0.05	-0.06	-0.07		-0.06	-0.07	-0.03	-0.01	-0.06	-0.07	0.17	0.10	-0.02	-0.02	0.00
CS_tagus_0	0.13	0.60	0.46		0.69	0.45		-0.09	-0.02	-0.08	-0.03		-0.08	-0.10	-0.08	-0.08	-0.10		-0.08	-0.08	-0.07	-0.04	-0.09	-0.07	-0.05	-0.05	-0.02	-0.06		-0.07	-0.06	-0.07	-0.08	-0.09		-0.10	-0.04	-0.01	0.10	-0.04	0.26	0.22	-0.03	-0.03	-0.01
CS_tamar_0	0.23	0.70	0.55	0.69		0.61	0.00		0.01	-0.04	-0.02	-0.07		-0.06	-0.04	-0.06	-0.06	-0.07		-0.06	-0.05		-0.06	-0.05	-0.04	-0.04	-0.02	-0.04	-0.06	-0.05	-0.03	-0.05	-0.06	-0.06	-0.05	-0.06	-0.03	-0.01	0.06	-0.04	0.15	0.13	-0.02	-0.02	
CS_venice_0	0.29	0.59	0.61	0.45	0.61		0.16	0.01	0.01	-0.01	-0.01	-0.05	-0.05		-0.04	-0.04	-0.03	-0.04	-0.05	-0.04	-0.04	0.08		-0.04	-0.03	-0.03	-0.01	-0.03	-0.05		-0.04	-0.04	-0.05	-0.05	-0.02		-0.02	-0.01	0.01		0.11	0.09	-0.02	-0.02	-0.01
CS_tagus_1	0.23	0.01	0.09		0.00	0.16		0.45	0.06	0.17	0.13		-0.04	-0.06	-0.07	-0.08	-0.09		-0.09	-0.08	-0.08	-0.06	-0.09	-0.07	-0.05	-0.06	-0.02	-0.06		-0.08	-0.07	-0.07	-0.09	-0.10		-0.10	-0.04	-0.01	0.16	0.08	-0.07	-0.06	-0.03	-0.03	-0.01
CS_tamar_1	0.09	-0.09	-0.01	-0.09		0.01	0.45		0.10	0.39	0.22	0.11		0.13	0.00	-0.08	-0.06	-0.10		-0.08	-0.08		-0.09	-0.07	-0.05	-0.06	-0.02	-0.06	-0.10	-0.08	-0.02	-0.07	-0.09	-0.09	-0.07	-0.08	-0.04	-0.01	0.10	0.09	-0.07	-0.06	-0.03	-0.03	
CS_curonian_3		0.01	0.09	-0.02	0.01	0.01	0.06	0.10		0.19	0.53	0.16	0.17	0.19	0.02	0.07	0.09	0.12	0.08	-0.04	-0.04	0.03	-0.04			-0.03	-0.01	-0.03	-0.04	-0.04			-0.04	-0.04	-0.03	-0.03		-0.01	0.14	0.23					0.00
CS_danube_1	0.07		-0.02	-0.08	-0.04	-0.01	0.17	0.39	0.19		0.39	0.53	0.29	0.52			-0.09	-0.05	-0.02		-0.07	-0.05	-0.09	-0.07	-0.05			-0.06	-0.09	-0.07	-0.07	-0.07		-0.09	-0.10	-0.10	-0.04			-0.05	-0.07	-0.06	-0.03	-0.03	0.00
CS_elbe_4	-0.07	-0.05		-0.03	-0.02	-0.01	0.13	0.22	0.53	0.39		0.33	0.24	0.29	-0.07	-0.05		-0.04	-0.03	-0.07		-0.01	-0.08	-0.06	-0.04	-0.04	-0.05	-0.02		-0.08	-0.06	-0.05	-0.06	-0.07		-0.08	-0.04	-0.01	0.27	0.11	-0.06	-0.02	-0.02	-0.03	0.01
CS_tagus_2	-0.02	-0.09	-0.05		-0.07	-0.05		0.11	0.16	0.53	0.33		0.57	0.66	-0.08	-0.08	-0.10		-0.06	-0.08	-0.08	-0.08	-0.09	-0.07	-0.05	-0.06	-0.02	-0.06		-0.08	-0.07	-0.07	-0.09	-0.09		-0.10	-0.05	-0.02	-0.08	-0.04	-0.07	-0.06	-0.03	-0.03	0.01
CS_tamar_2	-0.08	-0.08	-0.07	-0.08		-0.05	-0.04		0.17	0.29	0.24	0.57		0.34	0.02	-0.08	-0.02	-0.04		-0.08	-0.07		-0.09	-0.07	-0.05	-0.06	-0.02	-0.06	-0.09	-0.07	-0.05	-0.07	-0.08	-0.09	-0.06	-0.08	-0.05	-0.02	0.01	0.13	-0.07	-0.06	-0.03	-0.03	
CS_venice_1	-0.01	-0.09	-0.05	-0.10	-0.06		-0.06	0.13	0.19	0.52	0.29	0.66	0.34		-0.09	0.11	-0.02	0.11	0.08	-0.08	-0.08	-0.07		-0.07	-0.05	-0.06	-0.02	-0.06	-0.10		-0.07	-0.07	-0.09	-0.10	-0.11		-0.04	-0.01	-0.10		-0.07	-0.06	0.02	-0.03	0.05
CS_danube_3	-0.08		-0.06	-0.08	-0.04	-0.04	-0.07	0.00	0.02		-0.07	-0.08	0.02	-0.09			0.47	0.05	0.36		-0.07	-0.05	-0.08	-0.07	-0.05			-0.06	-0.08	-0.07	0.48	-0.07		-0.03	0.32	0.28	-0.04			0.21	-0.06	-0.06	-0.01	-0.03	-0.01
CS_danube_4	-0.08		-0.06	-0.08	-0.06	-0.04	-0.08	-0.08	0.07		-0.05	-0.08	-0.08	0.11			0.42	0.70	0.25		-0.07	0.08	0.08	-0.07	-0.05			-0.06	-0.08	-0.07	-0.06	-0.07		-0.05	-0.07	0.03	-0.01			0.06	-0.06	-0.06	0.15	0.17	0.08
CS_elbe_3	-0.11	-0.09		-0.10	-0.06	-0.03	-0.09	-0.06	0.09	-0.09		-0.10	-0.02	-0.02	0.47	0.42		0.49	0.40	-0.08		0.03	-0.02	-0.07	-0.05	-0.06	-0.02		-0.10	-0.08	0.25	-0.07	-0.02		0.17	0.25	-0.04	0.01	-0.06	0.23	-0.07	-0.06	0.08	0.06	0.00
CS_tagus_3	-0.10	-0.09	-0.07		-0.07	-0.04		-0.10	0.12	-0.05	-0.04		-0.04	0.11	0.05	0.70	0.49		0.42	-0.08	-0.07	0.05	0.06	-0.07	-0.05	-0.06	-0.02	-0.06		-0.08	-0.07	-0.07	-0.08	-0.05		0.01	0.02	0.08	-0.07	0.16	-0.07	-0.06	0.15	0.08	0.07
CS_tamar_3	-0.09	-0.08	-0.07	-0.08		-0.05	-0.09		0.08	-0.02	-0.03	-0.06		0.08	0.36	0.25	0.40	0.42		-0.08	-0.07		-0.08	-0.07	-0.05	-0.06	-0.02	-0.06	-0.09	-0.07	0.15	-0.06	-0.04	-0.05	0.07	0.12	-0.04	-0.02	-0.08	0.16	-0.07	-0.06	0.04	-0.02	
CS_danube_6	-0.08		-0.06	-0.08	-0.06	-0.04	-0.08	-0.08	-0.04		-0.07	-0.08	-0.08	-0.08			-0.08	-0.08	-0.08		0.71	0.24	0.54	0.51	-0.03			-0.02	0.40	0.10	-0.07	0.15		0.09	0.10	-0.03	0.28		-0.08	-0.06	-0.06	-0.02	-0.03	-0.01
CS_elbe_1	-0.08	-0.07		-0.07	-0.05	-0.04	-0.08	-0.08	-0.04	-0.07		-0.08	-0.07	-0.08	-0.07	-0.07		-0.07	-0.07	0.71		0.22	0.38	0.70	-0.05	0.14	-0.02		0.55	0.25	-0.06	0.07	-0.06		0.00	-0.03	0.16	0.02	-0.08	-0.08	-0.06	-0.06	-0.03	-0.03	-0.01
CS_tamar_4	-0.06	-0.04	0.01	-0.04		0.08	-0.06		0.03	-0.05	-0.01	-0.08		-0.07	-0.05	0.08	0.03	0.05		0.24	0.22		0.31	0.19	0.12	0.14	0.04	0.15	0.33	0.21	0.02	0.18	0.24	0.28	0.24	0.20	0.10	0.02	-0.07	-0.07	-0.05	-0.05	0.03	0.06	
CS_venice_4	-0.09	-0.09	-0.07	-0.09	-0.06		-0.09	-0.09	-0.04	-0.09	-0.08	-0.09	-0.09		-0.08	0.08	-0.02	0.06	-0.08	0.54	0.38	0.31		0.24	-0.04	-0.01	-0.02	-0.03	0.17		-0.07	0.17	0.06	0.28	0.17		0.40	0.12	-0.09		-0.07	-0.06	0.07	0.13	-0.01
CS_curonian_9		-0.06	-0.05	-0.07	-0.05	-0.04	-0.07	-0.07		-0.07	-0.06	-0.07	-0.07	-0.07	-0.07	-0.07	-0.07	-0.07	-0.07	0.51	0.70	0.19	0.24			0.33	-0.01	0.14	0.64	0.42			-0.07	-0.07	-0.06	-0.05		-0.02	-0.07	-0.07					-0.01
CS_curonian_10		-0.05	-0.04	-0.05	-0.04	-0.03	-0.05	-0.05		-0.05	-0.04	-0.05	-0.05	-0.05	-0.05	-0.05	-0.05	-0.05	-0.05	-0.03	-0.05	0.12	-0.04			0.51	0.50	0.85	0.46	0.63			-0.05	-0.05	-0.05	-0.04		-0.02	-0.05	-0.05					-0.01
CS_danube_8	-0.06		-0.04	-0.05	-0.04	-0.03	-0.06	-0.06	-0.03		-0.05	-0.06	-0.06	-0.06			-0.06	-0.06	-0.06		0.14	0.14	-0.01	0.33	0.51			0.65	0.51	0.67	-0.05	-0.05		-0.06	-0.06	-0.06	-0.02			-0.06	-0.05	-0.05	-0.03	-0.03	-0.01
CS_danube_9	-0.02		-0.02	-0.02	-0.02	-0.01	-0.02	-0.02	-0.01		-0.02	-0.02	-0.02	-0.02			-0.02	-0.02	-0.02		-0.02	0.04	-0.02	-0.01	0.50			0.42	0.19	0.31	-0.02	-0.02		-0.02	-0.02	-0.02	-0.02			-0.02	-0.02	-0.02	-0.02	-0.02	-0.01
CS_elbe_0	-0.06	-0.06		-0.06	-0.04	-0.03	-0.06	-0.06	-0.03	-0.06		-0.06	-0.06	-0.06	-0.06	-0.06		-0.06	-0.06	-0.02		0.15	-0.03	0.14	0.85	0.65	0.42		0.57	0.73	-0.05	-0.05	-0.06		-0.06	-0.05	-0.04	-0.02	-0.06	-0.06	-0.05	-0.05	-0.03	-0.03	-0.01
CS_tagus_5	-0.10	-0.09	-0.07		-0.06	-0.05		-0.10	-0.04	-0.09	-0.08		-0.09	-0.10	-0.08	-0.08	-0.10		-0.09	0.40	0.55	0.33	0.17	0.64	0.46	0.51	0.19	0.57		0.72	-0.07	-0.01	-0.07	-0.09		-0.05	-0.01	-0.01	-0.09	-0.10	-0.07	-0.06	-0.03	-0.03	-0.01
CS_venice_5	-0.08	-0.07	-0.06	-0.07	-0.05		-0.08	-0.08	-0.04	-0.07	-0.06	-0.08	-0.07		-0.07	-0.07	-0.08	-0.08	-0.07	0.10	0.25	0.21		0.42	0.63	0.67	0.31	0.73	0.72		-0.06	-0.06	-0.07	-0.08	-0.08		-0.03	-0.01	-0.07		-0.06	-0.06	-0.03	-0.03	-0.01
CS_curonian_4		-0.07	-0.05	-0.06	-0.03	-0.04	-0.07	-0.02		-0.07	-0.05	-0.07	-0.05	-0.07	0.48	-0.06	0.25	-0.07	0.15	-0.07	-0.06	0.02	-0.07			-0.05	-0.02	-0.05	-0.07	-0.06			0.17	0.10	0.36	0.44		-0.02	-0.05	-0.02					-0.01
CS_curonian_5		-0.07	-0.06	-0.07	-0.05	-0.04	-0.07	-0.07		-0.07	-0.06	-0.07	-0.07	-0.07	-0.07	-0.07	-0.07	-0.07	-0.06	0.15	0.07	0.18	0.17			-0.05	-0.02	-0.05	-0.01	-0.06			0.54	0.53	0.33	0.29		-0.02	-0.07	-0.07					-0.01
CS_danube_5	-0.09		-0.07	-0.08	-0.06	-0.05	-0.09	-0.09	-0.04		-0.07	-0.09	-0.08	-0.09			-0.02	-0.08	-0.04		-0.06	0.24	0.06	-0.07	-0.05			-0.06	-0.07	-0.07	0.17	0.54		0.69	0.50	0.51	0.01			-0.09	-0.07	-0.06	0.08	0.15	-0.01
CS_elbe_2	-0.10	-0.09		-0.09	-0.06	-0.05	-0.10	-0.09	-0.04	-0.09		-0.09	-0.09	-0.10	-0.03	-0.05		-0.05	-0.05	0.09		0.28	0.28	-0.07	-0.05	-0.06	-0.02		-0.09	-0.08	0.10	0.53	0.69		0.56	0.38	0.23	0.06	-0.09	-0.10	-0.07	-0.06	0.04	0.13	-0.01
CS_tagus_4	-0.10	-0.09	-0.06		-0.05	-0.02		-0.07	-0.03	-0.10	-0.08		-0.06	-0.11	0.32	-0.07	0.17		0.07	0.10	0.00	0.24	0.17	-0.06	-0.05	-0.06	-0.02	-0.06		-0.08	0.36	0.33	0.50	0.56		0.56	0.13	0.00	-0.09	-0.02	-0.07	-0.06	0.02	0.06	-0.01
CS_venice_3	-0.10	-0.10	-0.07	-0.10	-0.06		-0.10	-0.08	-0.03	-0.10	-0.08	-0.10	-0.08		0.28	0.03	0.25	0.01	0.12	-0.03	-0.03	0.20		-0.05	-0.04	-0.06	-0.02	-0.05	-0.05		0.44	0.29	0.51	0.38	0.56		-0.04	-0.02	-0.10		-0.07	-0.06	0.05	0.08	-0.01
CS_curonian_8		-0.04	-0.03	-0.04	-0.03	-0.02	-0.04	-0.04		-0.04	-0.04	-0.05	-0.05	-0.04	-0.04	-0.01	-0.04	0.02	-0.04	0.28	0.16	0.10	0.40			-0.02	-0.02	-0.04	-0.01	-0.03			0.01	0.23	0.13	-0.04		0.31	-0.05	-0.05					-0.01
CS_danube_7	-0.01		-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01		-0.01	-0.02	-0.02	-0.01			0.01	0.08	-0.02		0.02	0.02	0.12	-0.02	-0.02			-0.02	-0.01	-0.01	-0.02	-0.02		0.06	0.00	-0.02	0.31			-0.02	-0.02	-0.02	0.15	-0.01	0.09
CS_danube_2	-0.09		-0.06	0.10	0.06	0.01	0.16	0.10	0.14		0.27	-0.08	0.01	-0.10			-0.06	-0.07	-0.08		-0.08	-0.07	-0.09	-0.07	-0.05			-0.06	-0.09	-0.07	-0.05	-0.07		-0.09	-0.09	-0.10	-0.05			0.39	-0.06	0.10	-0.03	-0.03	-0.01
CS_venice_2	-0.10	-0.09	-0.07	-0.04	-0.04		0.08	0.09	0.23	-0.05	0.11	-0.04	0.13		0.21	0.06	0.23	0.16	0.16	-0.08	-0.08	-0.07		-0.07	-0.05	-0.06	-0.02	-0.06	-0.10		-0.02	-0.07	-0.09	-0.10	-0.02		-0.05	-0.02	0.39		-0.07	-0.06	-0.02	-0.03	-0.01
CS_curonian_0		0.25	0.17	0.26	0.15	0.11	-0.07	-0.07		-0.07	-0.06	-0.07	-0.07	-0.07	-0.06	-0.06	-0.07	-0.07	-0.07	-0.06	-0.06	-0.05	-0.07			-0.05	-0.02	-0.05	-0.07	-0.06			-0.07	-0.07	-0.07	-0.07		-0.02	-0.06	-0.07					-0.01
CS_curonian_2		0.13	0.10	0.22	0.13	0.09	-0.06	-0.06		-0.06	-0.02	-0.06	-0.06	-0.06	-0.06	-0.06	-0.06	-0.06	-0.06	-0.06	-0.06	-0.05	-0.06			-0.05	-0.02	-0.05	-0.06	-0.06			-0.06	-0.06	-0.06	-0.06		-0.02	0.10	-0.06					-0.01
CS_curonian_6		-0.03	-0.02	-0.03	-0.02	-0.02	-0.03	-0.03		-0.03	-0.02	-0.03	-0.03	0.02	-0.01	0.15	0.08	0.15	0.04	-0.02	-0.03	0.03	0.07			-0.03	-0.02	-0.03	-0.03	-0.03			0.08	0.04	0.02	0.05		0.15	-0.03	-0.02					0.22
CS_curonian_7		-0.03	-0.02	-0.03	-0.02	-0.02	-0.03	-0.03		-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	0.17	0.06	0.08	-0.02	-0.03	-0.03	0.06	0.13			-0.03	-0.02	-0.03	-0.03	-0.03			0.15	0.13	0.06	0.08		-0.01	-0.03	-0.03					-0.01
CS_tamar_5	0.00	-0.01	0.00	-0.01		-0.01	-0.01		0.00	0.00	0.01	0.01		0.05	-0.01	0.08	0.00	0.07		-0.01	-0.01		-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	0.09	-0.01	-0.01	-0.01	-0.01	0.22	-0.01	

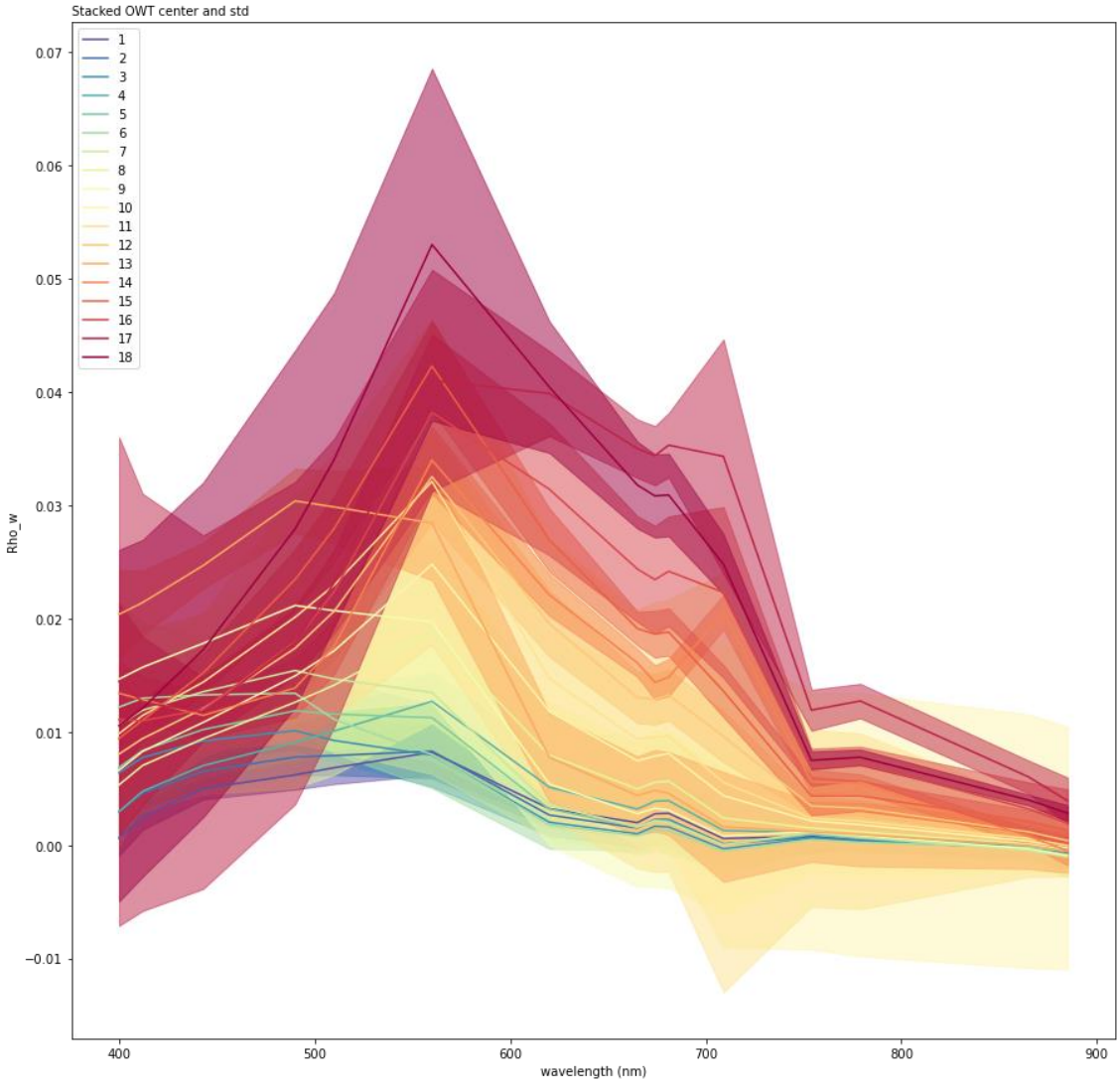

Figure 32: Prior-supported pan-regional S3 c-means cluster set of Optical Water Types (OWT) representing all six CERTO study sites. Cluster centre and ± 1 std is shown for each OWT, ordering the OWT from lowest overall reflectance (purple) to highest (red).

4.1.2 Distribution of clusters in case study regions

Checking geographic distribution coverage of all 18 OWTs across the six separate CERTO study sites, good coverage is observed for each site. Looking at the dominant OWT class distribution (Figure 33 to Figure 38), we observe good coverage of clear coastal sea or ocean by the lower OWTs, while the higher OWTs better represent the lagoon or estuary waters. In some cases, the lowest OWT class is seen in very shallow waters. This is likely a result of poor atmospheric correction and investigation into this phenomenon is ongoing within CERTO.

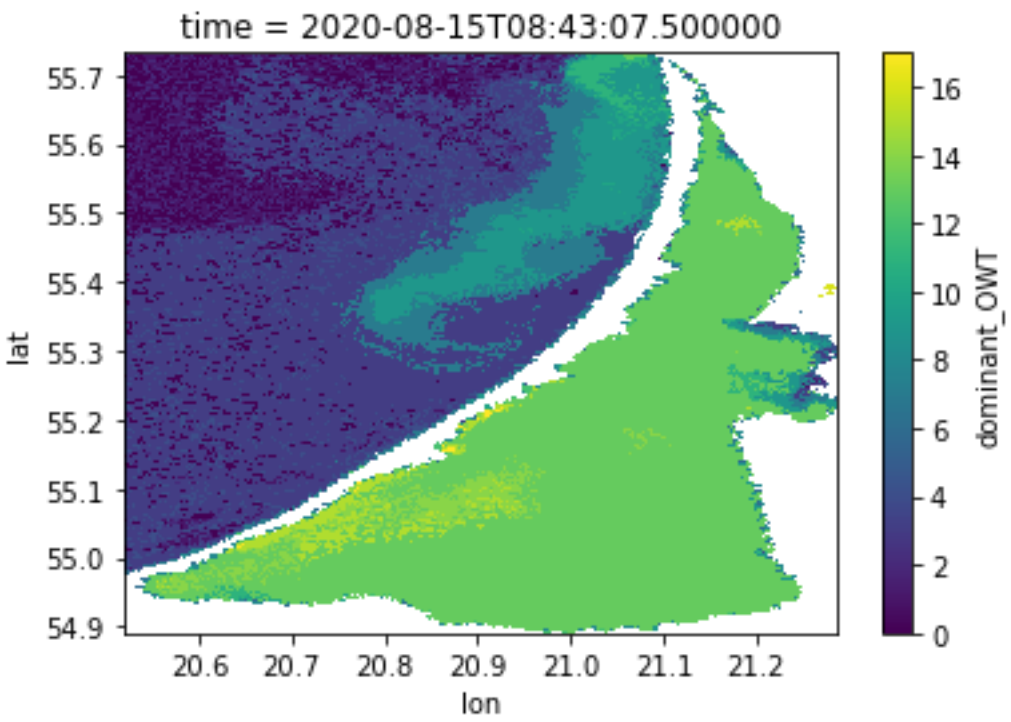

Figure 33: Pan-regional S3 dominant cluster for Curonian Lagoon for daily image.

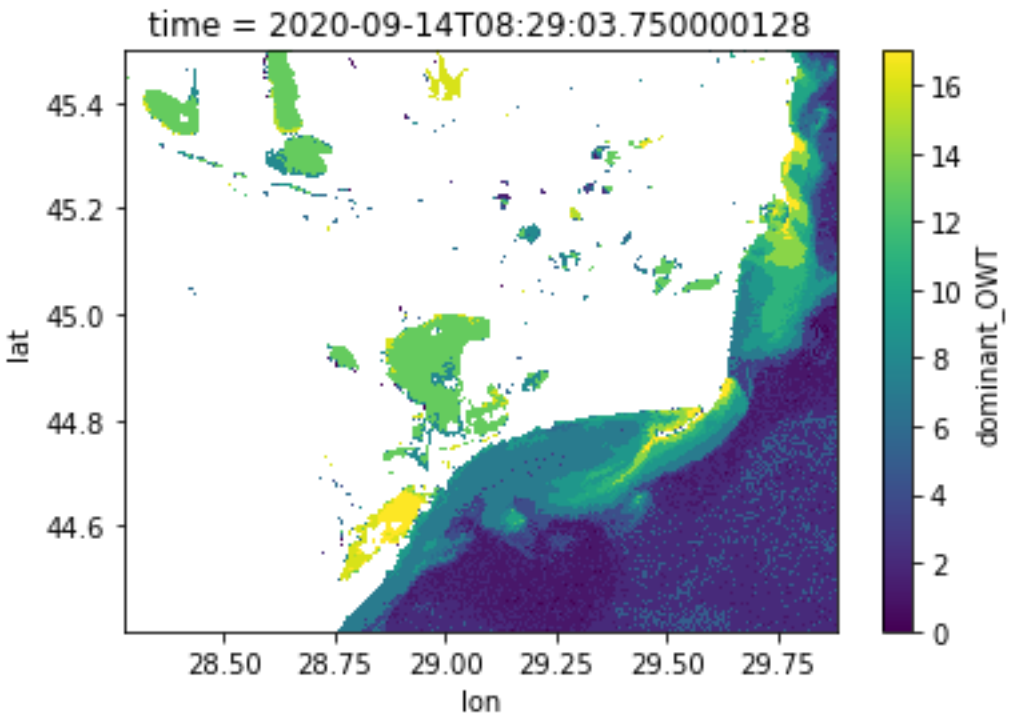

Figure 34: Pan-regional S3 dominant cluster for Danube Delta region for daily image.

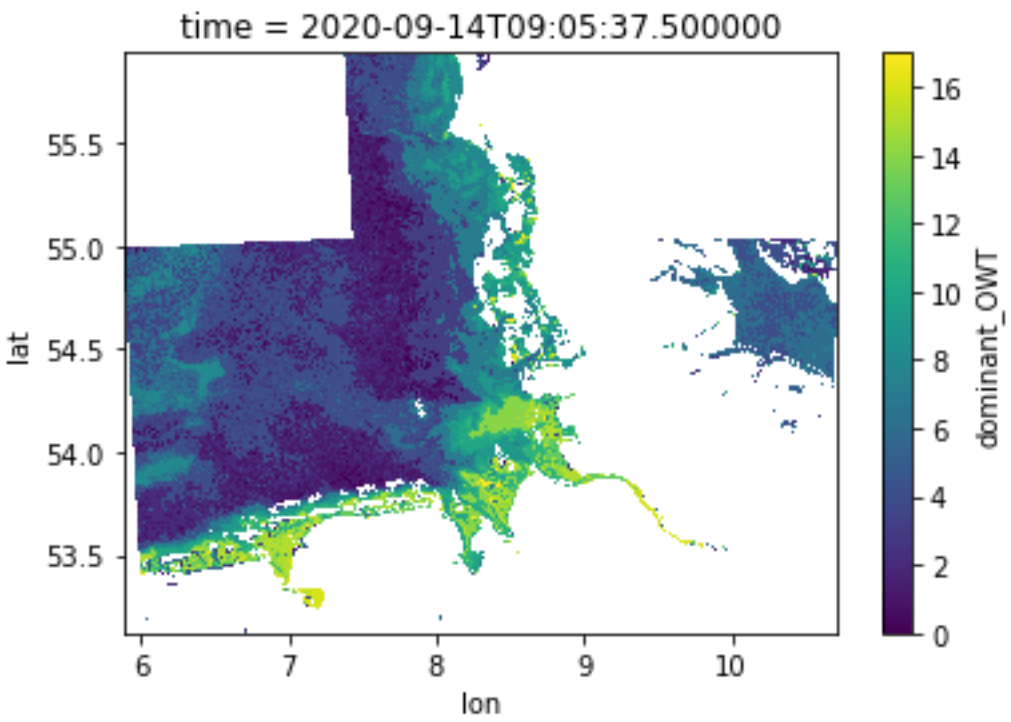

Figure 35: Pan-regional S3 dominant cluster for Elbe region for daily image.

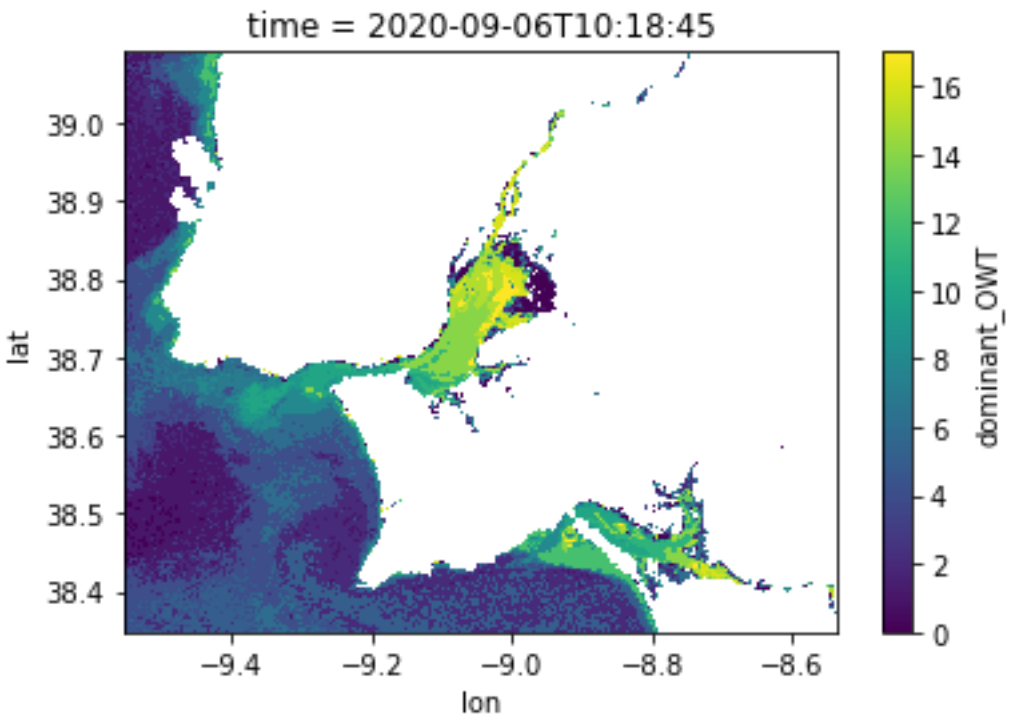

Figure 36: Pan-regional S3 dominant cluster for Tagus region for daily image.

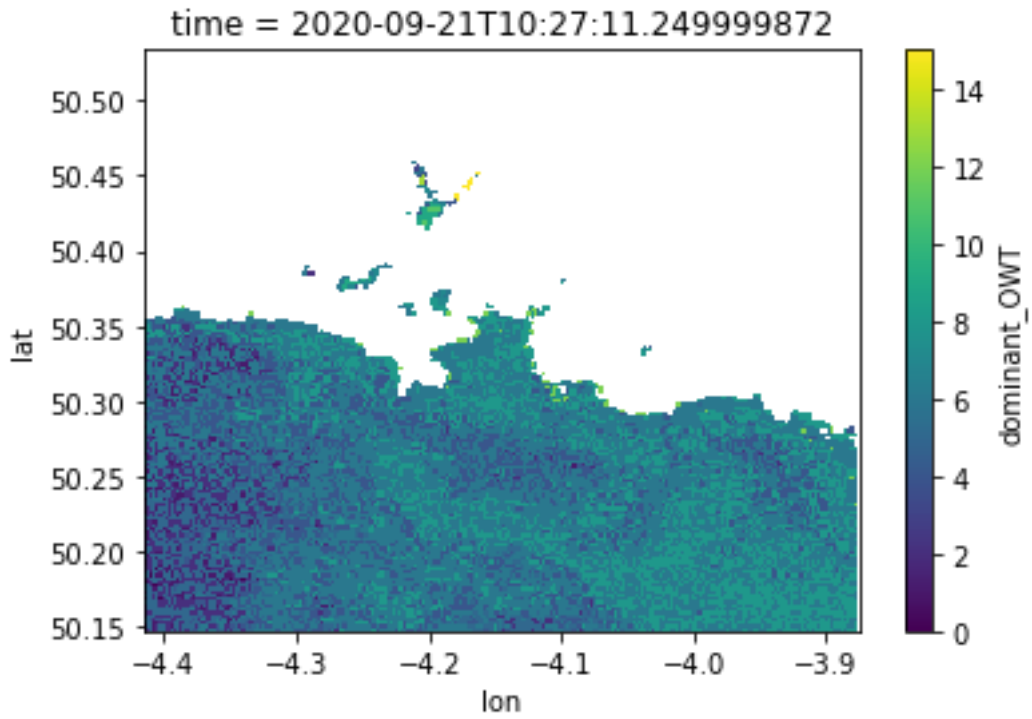

Figure 37: Pan-regional S3 dominant cluster for Tamar region for daily image.

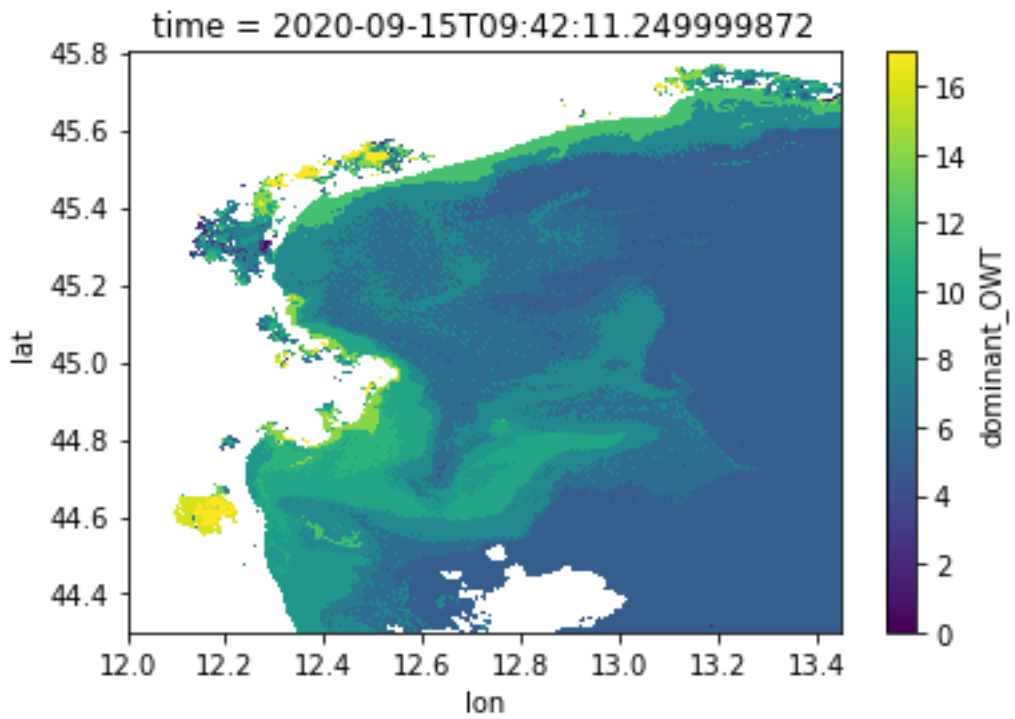

Figure 38: Pan-regional S3 dominant cluster for Venice region for daily image.

4.2 Sentinel 2

4.2.1 Cluster reflectance spectra

As for S3, pairwise comparisons between particular clusters for site-specific cluster sets were performed following sample separation by dominant OWT (samples were either class under examination or other). The resulting pairwise comparisons are presented in Table 2, which has been coloured using a threshold of ARI = 0.35 into “similar” (green) and “dissimilar” (red) clusters between cluster sets.

From this comparison, we see that all the cluster spectra from the Elbe Estuary/German Bight overlap with cluster spectra from other study sites, while those from the Venice Lagoon/North Adriatic Sea are relatively more unique.

The cluster grouping resulted in 17 clusters that were then used in a semi-supervised clustering exercise (as a prior to the cluster optimisation step) and resulted in the clusters shown in Figure 39.

4.2.2 Distribution of clusters in case study regions

Checking geographic distribution coverage of all 17 OWTs across the six separate CERTO study sites, good coverage is observed for each site. Looking at the dominant OWT class distribution (Figure 40 to Figure 45) we observe good coverage of clear coastal sea or ocean by the lower OWTs, while the higher OWTs better represent the lagoon or estuary waters.

In some cases, the lowest OWT class is seen in very shallow waters. This is likely a result of poor atmospheric correction and investigation into this phenomenon is ongoing within CERTO.

Table 2: Adjusted Rand Index (ARI) for each S2 cluster pair. Comparison between cluster sets from all six CERTO focus regions (note clusters zero indexed). Threshold value of ARI=0.35 indicates clusters pairs that are similar (green) or dissimilar (red). Cluster pair comparisons were not made within a study site cluster set (grey). Cluster sets have been reordered by ARI > 0.35 threshold, so highlight similarity and difference along diagonal.

	CS_curonian_1	CS_danube_0	CS_elbe_0	CS_tagus_1	CS_tamar_0	CS_venice_0	CS_tagus_2	CS_tamar_1	CS_danube_1	CS_elbe_1	CS_tamar_2	CS_tamar_3	CS_venice_3	CS_curonian_2	CS_danube_2	CS_elbe_2	CS_tagus_3	CS_tamar_4	CS_venice_5	CS_curonian_3	CS_danube_3	CS_elbe_4	CS_tagus_4	CS_tamar_5	CS_venice_8	CS_curonian_4	CS_danube_4	CS_elbe_5	CS_elbe_6	CS_tagus_5	CS_venice_9	CS_venice_10	CS_curonian_5	CS_danube_5	CS_elbe_3	CS_venice_6	CS_curonian_0	CS_tagus_0	CS_venice_1	CS_venice_2	CS_venice_4	CS_venice_7
CS_curonian_1		0.50	0.48	0.10	0.27	0.25	0.15	0.23	-0.01	0.00	-0.01	-0.04	-0.04		-0.03	-0.04	0.00	-0.02	-0.04		-0.04	-0.04	-0.04	0.02	-0.04		-0.04	-0.03	-0.03	-0.04	-0.04	-0.02		0.00	-0.04	-0.04		0.18	0.06	0.23	-0.04	-0.02
CS_danube_0	0.50		0.94	0.32	0.57	0.63	0.00	0.08		-0.01	-0.04	-0.04	-0.04	0.11		-0.04	0.00	-0.02	-0.04	-0.03		-0.04	-0.04	0.02	-0.04	-0.04		-0.03	-0.03	-0.04	-0.04	-0.02	0.00		-0.04	-0.04	0.13	0.19	0.04	0.05	-0.04	-0.02
CS_elbe_0	0.48	0.94		0.36	0.62	0.65	-0.01	0.06	0.00		-0.04	-0.04	-0.05	0.10	-0.03		0.00	-0.03	-0.04	-0.04	-0.04		-0.04	0.01	-0.05	-0.05	-0.05			-0.05	-0.04	-0.02	0.00	0.00		-0.04	0.13	0.20	0.04	0.04	-0.04	-0.02
CS_tagus_1	0.10	0.32	0.36		0.63	0.38		-0.11	-0.07	-0.07	-0.07	-0.07	-0.07	-0.06	-0.10	-0.11		-0.11	-0.07	-0.11	-0.10	-0.08		-0.09	-0.10	-0.09	-0.09	-0.06	-0.06		-0.08	-0.04	0.00	0.00	-0.09	-0.07	0.40		-0.04	0.08	-0.07	-0.03
CS_tamar_0	0.27	0.57	0.62	0.63		0.61	-0.06		-0.05	-0.06			-0.06	0.00	-0.08	-0.08	-0.06		-0.06	-0.08	-0.08	-0.06	-0.08		-0.08	-0.07	-0.08	-0.05	-0.05	-0.08	-0.06	-0.03	0.00	0.00	-0.07	-0.06	0.22	0.20	-0.02	0.03	-0.06	-0.03
CS_venice_0	0.25	0.63	0.65	0.38	0.61		-0.05	-0.01	-0.07	-0.08	-0.06	-0.06		-0.01	-0.09	-0.09	-0.08	-0.09		-0.09	-0.09	-0.07	-0.09	-0.07		-0.08	-0.08	-0.05	-0.05	-0.09			0.00	0.00	-0.08		0.25	0.32				
CS_tagus_2	0.15	0.00	-0.01		-0.06	-0.05		0.54	0.38	0.47	0.29	-0.03	0.27	0.02	-0.10	-0.10		-0.10	-0.07	-0.10	-0.10	-0.08		-0.09	-0.09	-0.09	-0.09	-0.06	-0.06		-0.07	-0.04	0.00	0.00	-0.09	-0.07	-0.07		0.11	0.29	-0.05	-0.03
CS_tamar_1	0.23	0.08	0.06	-0.11		-0.01	0.54		0.12	0.17			0.03	-0.05	-0.11	-0.11	-0.10		-0.08	-0.11	-0.10	-0.08	-0.10		-0.10	-0.09	-0.10	-0.06	-0.06	-0.10	-0.08	-0.04	0.00	0.00	-0.09	-0.07	-0.08	0.04	0.21	0.28	-0.07	-0.03
CS_danube_1	-0.01		0.00	-0.07	-0.05	-0.07	0.38	0.12		0.82	0.32	0.27	0.53	0.25		-0.05	0.05	-0.08	-0.07	-0.10		-0.08	-0.10	-0.09	-0.09	-0.09		-0.06	-0.06	-0.10	-0.07	-0.04	0.00		-0.09	-0.07	-0.07	-0.09	-0.01	0.13	0.13	-0.03
CS_elbe_1	0.00	-0.01		-0.07	-0.06	-0.08	0.47	0.17	0.82		0.35	0.17	0.50	0.20	-0.09		0.01	-0.08	-0.07	-0.10	-0.10		-0.10	-0.09	-0.09	-0.09	-0.09			-0.10	-0.07	-0.04	0.00	0.00		-0.07	-0.07	-0.09	0.00	0.16	0.09	-0.03
CS_tamar_2	-0.01	-0.04	-0.04	-0.07		-0.06	0.29		0.32	0.35			0.38	0.10	-0.08	-0.08	-0.05		-0.06	-0.08	-0.08	-0.07	-0.08		-0.07	-0.07	-0.07	-0.05	-0.06	-0.07	-0.06	-0.04	0.00	0.00	-0.07	-0.06	-0.06	-0.07	0.01	0.05	0.02	-0.03
CS_tamar_3	-0.04	-0.04	-0.04	-0.07		-0.06	-0.03		0.27	0.17			0.33	0.17	-0.04	0.11	0.26		-0.06	-0.07	-0.07	-0.06	-0.07		-0.07	-0.07	-0.07	-0.05	-0.05	-0.07	-0.06	-0.04	0.00	0.00	-0.07	-0.06	-0.06	-0.07	-0.04	-0.07	0.26	-0.03
CS_venice_3	-0.04	-0.04	-0.05	-0.07	-0.06		0.27	0.03	0.53	0.50	0.38	0.33		0.26	-0.10	-0.06	0.03	-0.10		-0.10	-0.10	-0.08	-0.10	-0.09		-0.09	-0.09	-0.06	-0.06	-0.09			0.00	0.00	-0.09		-0.08	-0.09				
CS_curonian_2		0.11	0.10	-0.06	0.00	-0.01	0.02	-0.05	0.25	0.20	0.10	0.17	0.26		0.41	0.42	0.64	0.38	0.17		-0.07	-0.06	-0.06	-0.04	-0.07		-0.07	-0.04	-0.05	-0.07	-0.06	-0.03		0.00	-0.01	0.06		-0.07	-0.03	-0.07	0.23	0.02
CS_danube_2	-0.03		-0.03	-0.10	-0.08	-0.09	-0.10	-0.11		-0.09	-0.08	-0.04	-0.10	0.41		0.72	0.61	0.72	0.47	-0.06		-0.08	-0.01	-0.06	-0.10	-0.09		-0.06	-0.06	-0.10	-0.08	-0.04	0.00		0.11	0.28	-0.08	-0.09	-0.05	-0.10	0.19	0.05
CS_elbe_2	-0.04	-0.04		-0.11	-0.08	-0.09	-0.10	-0.11	-0.05		-0.08	0.11	-0.06	0.42	0.72		0.69	0.59	0.50	-0.10	-0.10		-0.09	-0.08	-0.10	-0.09	-0.09			-0.10	-0.08	-0.04	0.00	0.00		0.08	-0.08	-0.09	-0.05	-0.10	0.30	0.02
CS_tagus_3	0.00	0.00	0.00		-0.06	-0.08		-0.10	0.05	0.01	-0.05	0.26	0.03	0.64	0.61	0.69		0.60	0.30	-0.10	-0.09	-0.08		-0.09	-0.09	-0.09	-0.09	-0.06	-0.06		-0.07	-0.04	0.00	0.00	-0.01	0.09	-0.07		-0.05	-0.10	0.34	0.05
CS_tamar_4	-0.02	-0.02	-0.03	-0.11		-0.09	-0.10		-0.08	-0.08			-0.10	0.38	0.72	0.59	0.60		0.35	-0.01	0.01	-0.08	0.03		-0.07	-0.09	-0.09	-0.06	-0.06	-0.10	-0.07	-0.04	0.00	0.00	0.21	0.24	-0.08	-0.09	-0.05	-0.10	0.19	0.16
CS_venice_5	-0.04	-0.04	-0.04	-0.07	-0.06		-0.07	-0.08	-0.07	-0.07	-0.06	-0.06		0.17	0.47	0.50	0.30	0.35		-0.06	-0.08	-0.06	-0.04	-0.03		-0.07	-0.07	-0.05	-0.06	-0.07			0.00	0.00	-0.05		-0.06	-0.07				
CS_curonian_3		-0.03	-0.04	-0.11	-0.08	-0.09	-0.10	-0.11	-0.10	-0.10	-0.08	-0.07	-0.10		-0.06	-0.10	-0.10	-0.01	-0.06		0.85	0.55	0.79	0.49	0.66		-0.03	-0.04	0.00	0.07	0.02	-0.04		0.00	0.41	0.14		-0.09	-0.05	-0.10	-0.09	0.06
CS_danube_3	-0.04		-0.04	-0.10	-0.08	-0.09	-0.10	-0.10		-0.10	-0.08	-0.07	-0.10	-0.07		-0.10	-0.09	0.01	-0.08	0.85		0.55	0.74	0.40	0.69	-0.09		-0.06	-0.01	0.03	-0.02	-0.04	0.00		0.41	0.09	-0.08	-0.09	-0.05	-0.10	-0.09	0.12
CS_elbe_4	-0.04	-0.04		-0.08	-0.06	-0.07	-0.08	-0.08	-0.08		-0.07	-0.06	-0.08	-0.06	-0.08		-0.08	-0.08	-0.06	0.55	0.55		0.41	0.33	0.60	-0.07	-0.02			0.10	0.07	-0.04	0.00	0.00		-0.05	-0.07	-0.08	-0.05	-0.08	-0.07	-0.03
CS_tagus_4	-0.04	-0.04	-0.04		-0.08	-0.09		-0.10	-0.10	-0.10	-0.08	-0.07	-0.10	-0.06	-0.01	-0.09		0.03	-0.04	0.79	0.74	0.41		0.36	0.49	-0.09	-0.08	-0.05	-0.06		-0.01	-0.04	0.00	0.00	0.55	0.24	-0.08		-0.05	-0.10	-0.09	0.09
CS_tamar_5	0.02	0.02	0.01	-0.09		-0.07	-0.09		-0.09	-0.09			-0.09	-0.04	-0.06	-0.08	-0.09		-0.03	0.49	0.40	0.33	0.36		0.44	0.41	0.45	0.23	0.24	0.54	0.30	0.13	0.00	0.00	0.10	0.03	-0.07	-0.09	-0.04	-0.09	-0.08	-0.03
CS_venice_8	-0.04	-0.04	-0.05	-0.10	-0.08		-0.09	-0.10	-0.09	-0.09	-0.07	-0.07		-0.07	-0.10	-0.10	-0.09	-0.07		0.66	0.69	0.60	0.49	0.44		-0.01	0.02	-0.05	0.12	0.16			0.00	0.00	0.16		-0.07	-0.09				
CS_curonian_4		-0.04	-0.05	-0.09	-0.07	-0.08	-0.09	-0.09	-0.09	-0.09	-0.07	-0.07	-0.09		-0.09	-0.09	-0.09	-0.09	-0.07		-0.09	-0.07	-0.09	0.41	-0.01		0.93	0.61	0.55	0.80	0.62	0.41		0.00	-0.08	-0.07		-0.08	-0.05	-0.09	-0.08	-0.04
CS_danube_4	-0.04		-0.05	-0.09	-0.08	-0.08	-0.09	-0.10		-0.09	-0.07	-0.07	-0.09	-0.07		-0.09	-0.09	-0.09	-0.07	-0.03		-0.02	-0.08	0.45	0.02	0.93		0.59	0.52	0.85	0.65	0.36	0.00		-0.08	-0.07	-0.07	-0.09	-0.05	-0.09	-0.08	-0.04
CS_elbe_5	-0.03	-0.03		-0.06	-0.05	-0.05	-0.06	-0.06	-0.06		-0.05	-0.05	-0.06	-0.04	-0.06		-0.06	-0.06	-0.05	-0.04	-0.06		-0.05	0.23	-0.05	0.61	0.59			0.49	0.58	0.27	0.00	0.00		-0.05	-0.05	-0.06	-0.04	-0.06	-0.06	-0.03
CS_elbe_6	-0.03	-0.03		-0.06	-0.05	-0.05	-0.06	-0.06	-0.06		-0.06	-0.05	-0.06	-0.05	-0.06		-0.06	-0.06	-0.06	0.00	-0.01		-0.06	0.24	0.12	0.55	0.52			0.52	0.26	0.35	0.00	0.00		-0.05	-0.06	-0.06	-0.04	-0.06	-0.06	-0.03
CS_tagus_5	-0.04	-0.04	-0.05		-0.08	-0.09		-0.10	-0.10	-0.10	-0.07	-0.07	-0.09	-0.07	-0.10	-0.10		-0.10	-0.07	0.07	0.03	0.10		0.54	0.16	0.80	0.85	0.49	0.52		0.54	0.30	0.00	0.00	-0.09	-0.07	-0.08		-0.05	-0.10	-0.08	-0.04
CS_venice_9	-0.04	-0.04	-0.04	-0.08	-0.06		-0.07	-0.08	-0.07	-0.07	-0.06	-0.06		-0.06	-0.08	-0.08	-0.07	-0.07		0.02	-0.02	0.07	-0.01	0.30		0.62	0.65	0.58	0.26	0.54			0.00	0.00	-0.07		-0.06	-0.07				
CS_venice_10	-0.02	-0.02	-0.02	-0.04	-0.03		-0.04	-0.04	-0.04	-0.04	-0.04	-0.04		-0.03	-0.04	-0.04	-0.04	-0.04		-0.04	-0.04	-0.04	-0.04	0.13		0.41	0.36	0.27	0.35	0.30			0.00	0.01	-0.04		-0.04	-0.04				
CS_curonian_5		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.00	0.00		0.39	0.00	0.00		0.00	0.00	0.00	0.00	0.00
CS_danube_5	0.00		0.00	0.00	0.00	0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.01	0.39		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CS_elbe_3	-0.04	-0.04		-0.09	-0.07	-0.08	-0.09	-0.09	-0.09		-0.07	-0.07	-0.09	-0.01	0.11		-0.01	0.21	-0.05	0.41	0.41		0.55	0.10	0.16	-0.08	-0.08			-0.09	-0.07	-0.04	0.00	0.00		0.48	-0.07	-0.08	-0.05	-0.09	-0.08	0.26
CS_venice_6	-0.04	-0.04	-0.04	-0.07	-0.06		-0.07	-0.07	-0.07	-0.07	-0.06	-0.06		0.06	0.28	0.08	0.09	0.24		0.14	0.09	-0.05	0.24	0.03		-0.07	-0.07	-0.05	-0.05	-0.07			0.00	0.00	0.48		-0.06	-0.07				
CS_curonian_0		0.13	0.13	0.40	0.22	0.25	-0.07	-0.08	-0.07	-0.07	-0.06	-0.06	-0.08		-0.08	-0.08	-0.07	-0.08	-0.06		-0.08	-0.07	-0.08	-0.07	-0.07		-0.07	-0.05	-0.06	-0.08	-0.06	-0.04		0.00	-0.07	-0.06		-0.06	-0.04	-0.07	-0.07	-0.03
CS_tagus_0	0.18	0.19	0.20		0.20	0.32		0.04	-0.09	-0.09	-0.07	-0.07	-0.09	-0.07	-0.09	-0.09		-0.09	-0.07	-0.09	-0.09	-0.08		-0.09	-0.09	-0.08	-0.09	-0.06	-0.06		-0.07	-0.04	0.00	0.00	-0.08	-0.07	-0.06		0.07	-0.09	-0.08	-0.04
CS_venice_1	0.06	0.04	0.04	-0.04	-0.02		0.11	0.21	-0.01	0.00	0.01	-0.04		-0.03	-0.05	-0.05	-0.05	-0.05		-0.05	-0.05	-0.05	-0.05	-0.04		-0.05	-0.05	-0.04	-0.04	-0.05			0.00	0.00	-0.05		-0.04	0.07				
CS_venice_2	0.23	0.05	0.04	0.08	0.03		0.29	0.28	0.13	0.16	0.05	-0.07		-0.07	-0.10	-0.10	-0.10	-0.10		-0.10	-0.10	-0.08	-0.10	-0.09		-0.09	-0.09	-0.06	-0.06	-0.10			0.00	0.00	-0.09		-0.07	-0.09				
CS_venice_4	-0.04	-0.04	-0.04	-0.07	-0.06		-0.05	-0.07	0.13	0.09	0.02	0.26		0.23	0.19	0.30	0.34	0.19		-0.09	-0.09	-0.07	-0.09	-0.08		-0.08	-0.08	-0.06	-0.06	-0.08			0.00	0.00	-0.08		-0.07	-0.08				
CS_venice_7	-0.02	-0.02	-0.02	-0.03	-0.03		-0.03	-0.03	-0.03	-0.03	-0.03	-0.03		0.02	0.05	0.02	0.05	0.16		0.06	0.12	-0.03	0.09	-0.03		-0.04	-0.04	-0.03	-0.03	-0.04			0.00	0.00	0.26		-0.03	-0.04				

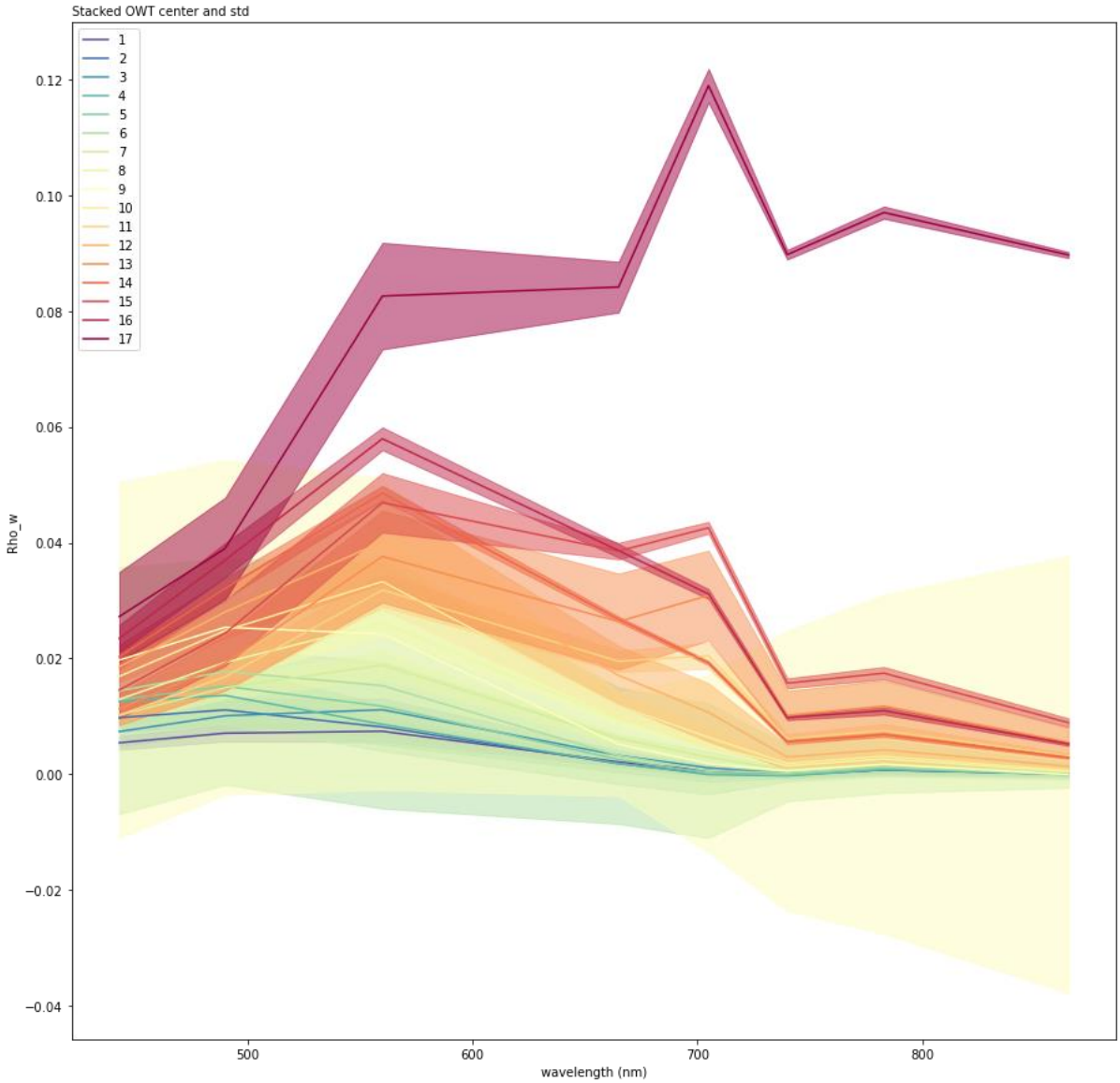

Figure 39: Prior-supported pan-regional S2 c-means cluster set of Optical Water Types (OWT) representing all six CERTO study sites. Cluster center and ± 1 std is shown for each OWT, ordering the OWT from lowest overall reflectance (purple) to highest (red).

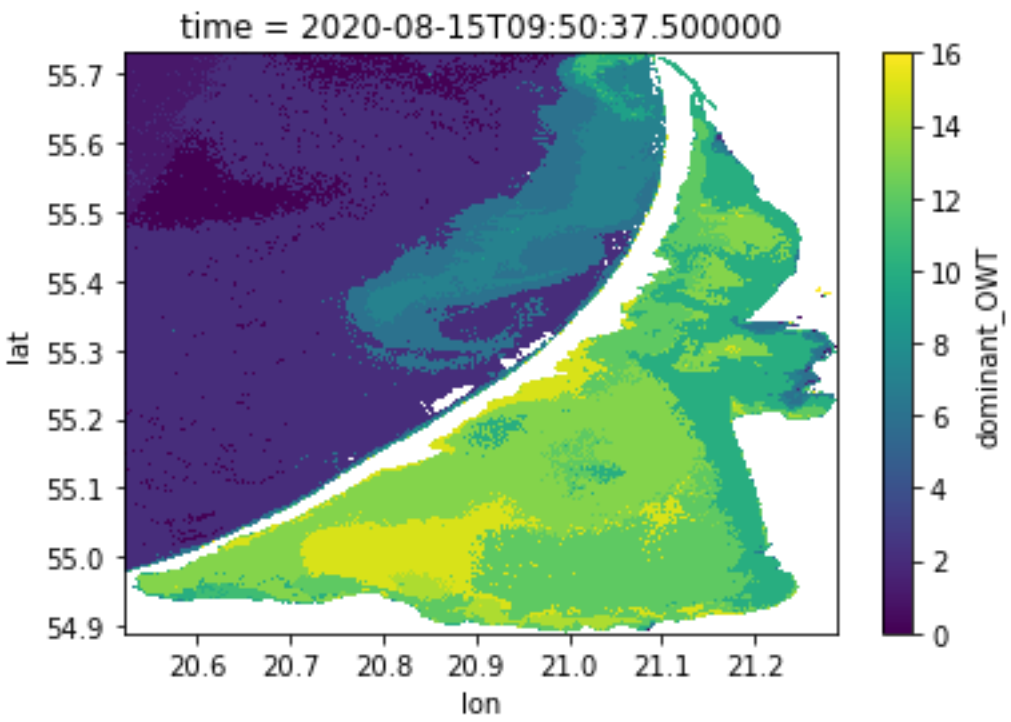

Figure 40: Pan-regional S2 dominant cluster for Curonian Lagoon region for daily image.

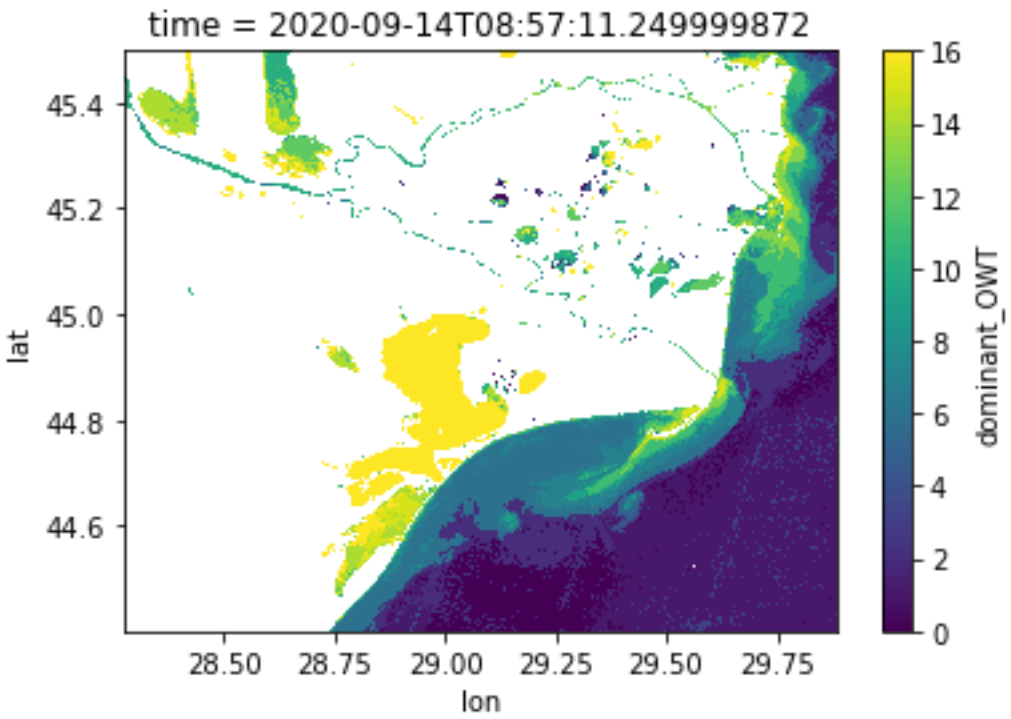

Figure 41: Pan-regional S2 dominant cluster for Danube Delta region for daily image.

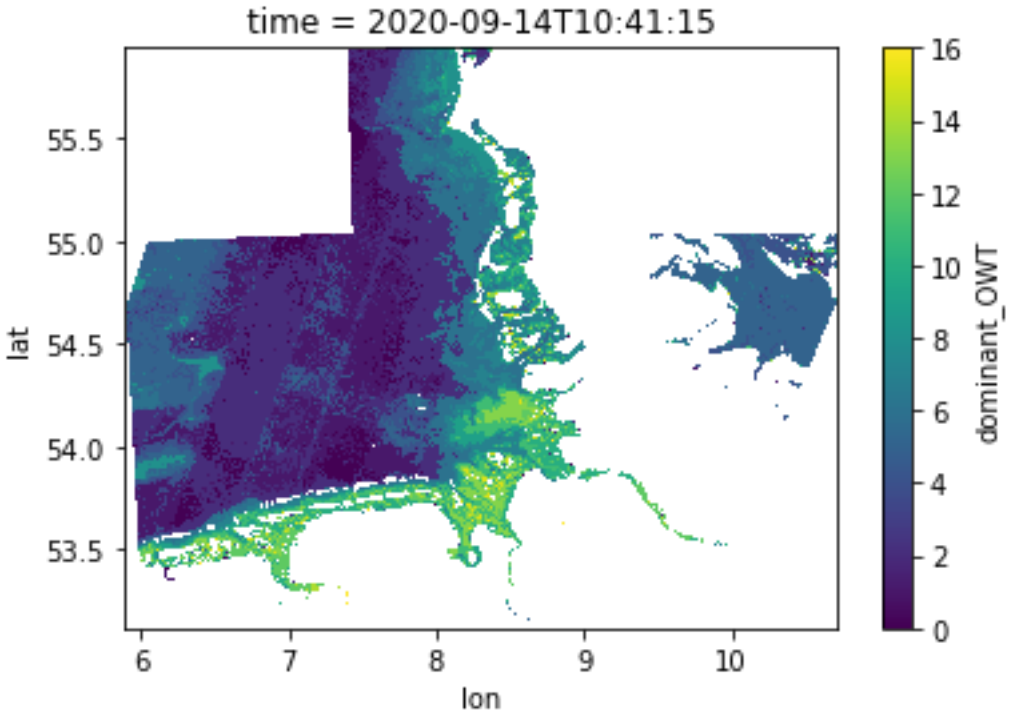

Figure 42: Pan-regional S2 dominant cluster for Elbe region for daily image.

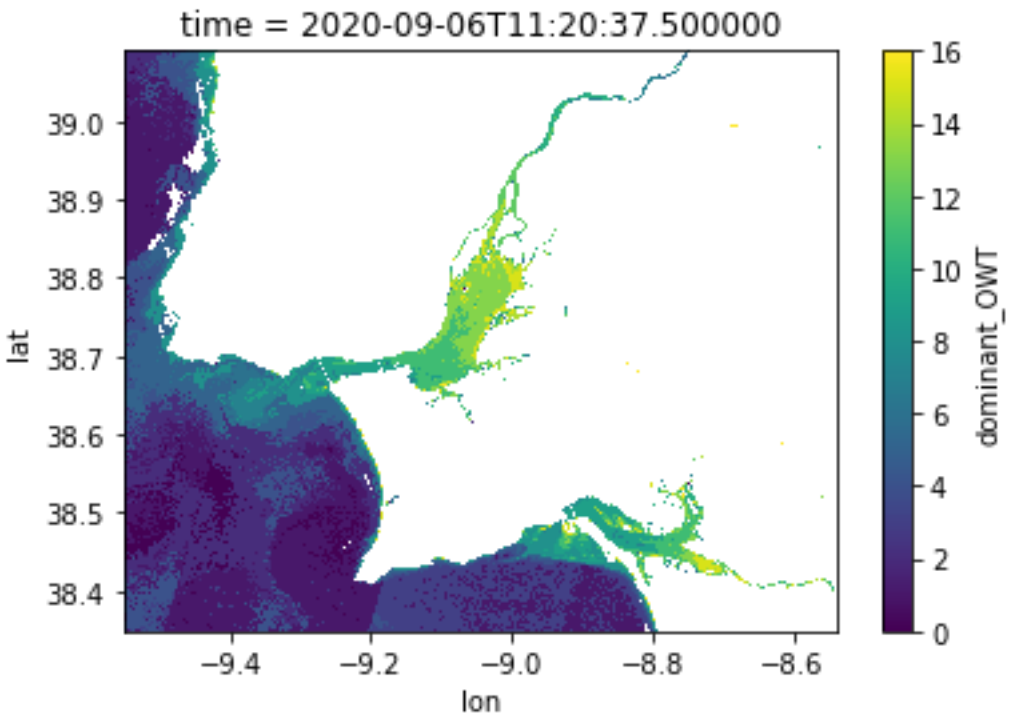

Figure 43: Pan-regional S2 dominant cluster for Tagus region for daily image.

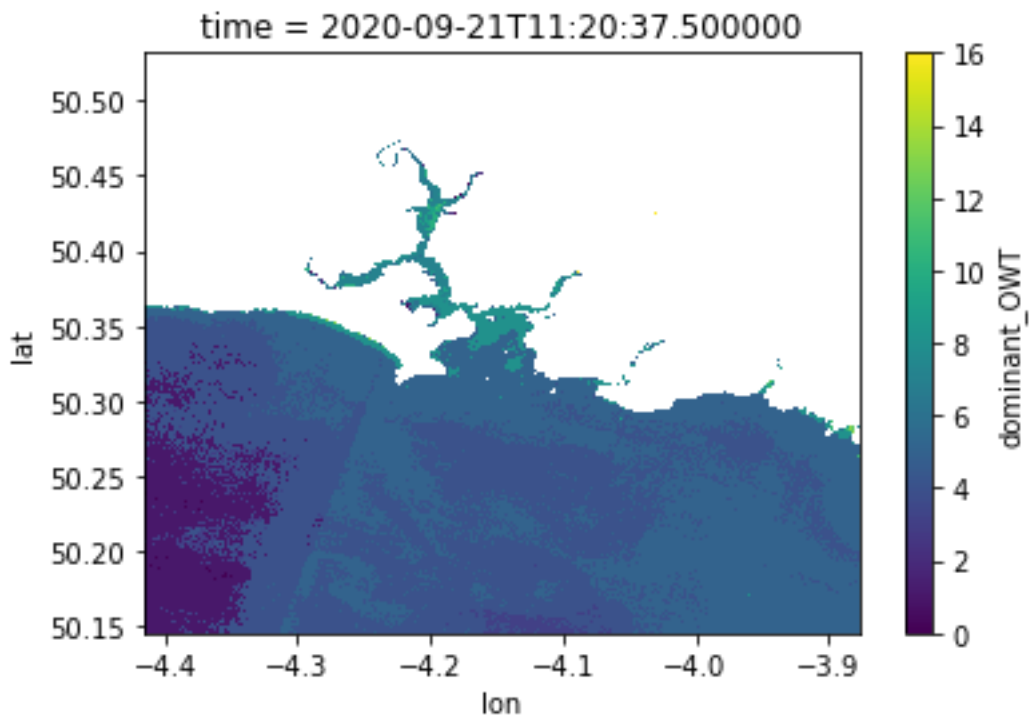

Figure 44: Pan-regional S2 dominant cluster for Tamar region for daily image.

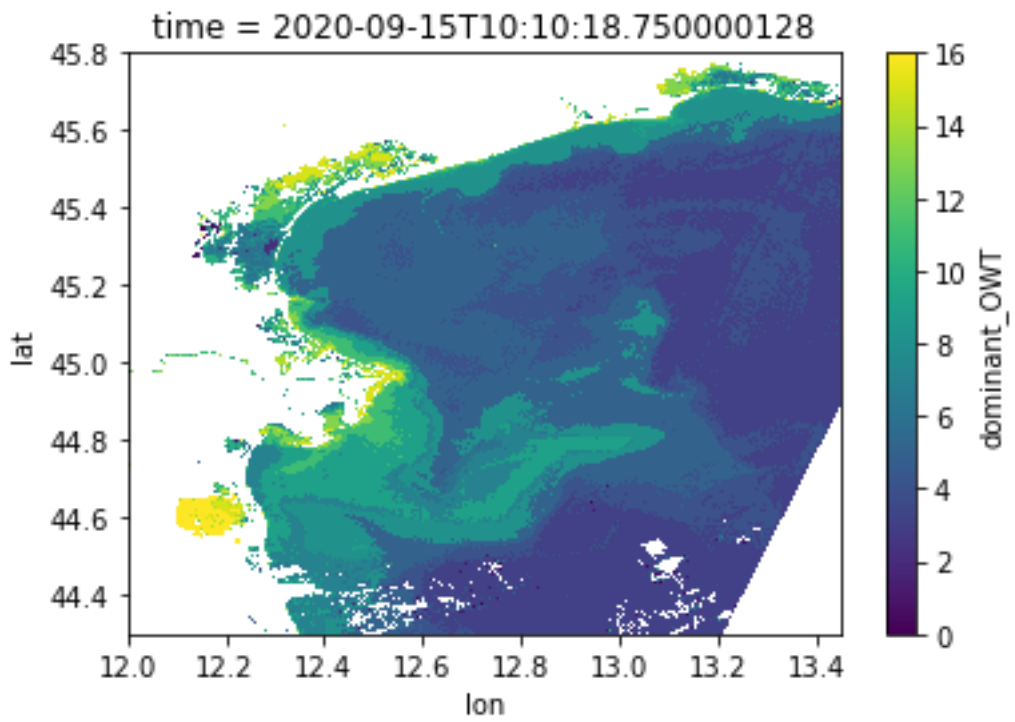

Figure 45: Pan-regional S2 dominant cluster for Venice region for daily image.

5 Cluster set files

All the cluster sets described above are available within the zipped file available at: https://doi.org/10.5281/zenodo.6724663