Supplementary Material

1 Drivers description

Table S1. List of drivers currently available on *eDrivers* along with their respective accronym used in the figures in the supplementary material.

Groups	Drivers	Accronym	Source
Climate	Aragonite	ARAG	(Starr, 2019)
Climate	Hypoxia	HYP	(Starr, 2019)
Climate	Sea bottom temperature	SBT-	(Galbraith et al., 2018)
Climate	Sea bottom temperature	SBT+	(Galbraith et al., 2018)
Climate	Sea surface temperature	SST-	(Galbraith et al., 2018)
Climate	Sea surface temperature	SST+	(Galbraith et al., 2018)
Climate	Sea water level	SLR	(Halpern et al., 2015a)
Coastal	Aquaculture	AQUA	***TBD***
Coastal	Coastal development	CD	(Group, 2019)
Coastal	Direct human impact	DHI	(Halpern et al., 2015a)
Coastal	Inorganic pollution	IP	(Halpern et al., 2015a)
Coastal	Nutrient import	NI	(Halpern et al., 2015a)
Coastal	Organic pollution	OP	(Halpern et al., 2015a)
Coastal	Toxic algae	TA	(Bates, 2019)
Fisheries	Demersal, destructive	DD	(DFO, 2016)
Fisheries	Demersal, non-destructive, high-bycatch	DNH	(DFO, 2016)
Fisheries	Demersal, non-destructive, low-bycatch	DNL	(DFO, 2016)
Fisheries	Pelagic, high-bycatch	PHB	(DFO, 2016)
Fisheries	Pelagic, low-bycatch	PLB	(DFO, 2016)
Marine traffic	Invasive species	INV	(Halpern et al., 2015a)
Marine traffic	Marine pollution	MP	(Halpern et al., 2015a)
Marine traffic	Shipping	SHP	(Halpern et al., 2015a)

1.1 Climate

1.1.1 Aragonite

Describe data

We interpolated aragonite distribution using the exponential kriging model. We then built an index of aragonite level stress. When Ω aragonite levels decrease below 1, the water becomes corrosive to calcifying organisms such as echinoderms, corals and mollusks larvae. These organisms thus have to continually recalcify their shells to maintain their structure,

which becomes harder to do as Ω aragonite levels further decrease towards 0. We therefore considered stress associated to decreasing Ω aragonite levels to increase linearly between values of 1 and 0. The index of stress therefore has a value of 0 and 1 for values of Ω aragonite equal to 1 and 0, respectively.

1.1.2 Hypoxia

Describe data

We interpolated the dissolved oxygen using cokriging with depth as a covariable, as done in (Dutil et al., 2011). According to Diaz and Rosenberg (1995), severe hypoxia can be observed when oxygen saturation falls below 2 $ml\ L^{-1}$ or 62.5 $\mu mol\ L^{-1}$, which is considered the level necessary to maintain most animal life. This threshold is equal to 1.4 $ml\ L^{-1}$ (1 $mg\ L^{-1}$ = 0.7 $ml\ L^{-1}$). We used this threshold to create an index of hypoxia using an inverted logistic curve to transform the dissolved oxygen values as an index of hypoxic stress H_s ranging from 0 to 1, under the assumption that hypoxic stress will increase following a logistic curve as it approaches the hypoxic threshold:

$$H_s = \frac{-1}{200 * e^{-2.5 * DO_2}} + 1$$

1.1.3 Bottom temperature anomalies

The data used to characterize bottom temperature anomalies come from the Department of Fisheries and Oceans' (DFO) Atlantic Zone Monitoring Program (AZMP; Galbraith et al., 2018). We provide a brief summary of data and methods to characterize surface temperature climatology and anomalies in this document. For more details, refer to Galbraith et al. (2018).

Bottom temperatures are interpolated in the Gulf using conductivity-temperature-depth (CTD) sampling performed annually through DFO's multispecies surveys for the northern Gulf in August and for the Magdalen Shallows in September. Using this sampling survey, temperatures are interpolated at each 1 m depth layer on a 2 km resolution grid. Bottom temperature are then extracted from this grid by using a bathymetry layer from the Canadian Hydrographic Survey (Dutil et al., 2012).

We used temperature anomalies, *i.e.* deviations from long-term normal conditions, to measure an annual index of stress associated with extreme temperatures between 2013 and 2017. Temperature anomalies were calculated using the difference between grid cell monthly climatologies to the associated long-term averages generated from the reference period between 1981 and 2010. Grid cells whose monthly value exceeded ±0.5 standard deviation (SD) from the long-term average were considered as anomalous (Galbraith et al., 2018).

Outliers in the data were considered as those that fell beyond the interquartile range (IQR) * 3, identified as extreme ouliers by Tukey (1977). Outlier values were capped to correspond to the 5th and 95th percentiles. Anomalies were divided into positive and negative

anomalies and summed for the available months of September and August. By adding the monthly anomaly values, we essentially considered the summed deviation from the mean as an indicator of the annual intensity of surface temperature anomalies in the St. Lawrence.

1.1.4 Surface temperature anomalies

The data used to characterize surface temperature anomalies come from the Department of Fisheries and Oceans' (DFO) Atlantic Zone Monitoring Program (AZMP; Galbraith et al., 2018). We provide a brief summary of data and methods to characterize surface temperature climatology and anomalies in this document. For more details, refer to Galbraith et al. (2018).

The surface layer is characterized using a variety of methods:

1. Temperature-salinity sensors installed on various sihps and forming the shipboard thermosalinograph network,
2. Moored instruments recording water temperature every 30 minutes and forming the thermograph network,
3. Sea surface temperature (SST) monthly composite climatologies using Advanced Very High Resolution Radiometer (AVHRR) satellite images from the National Oceanic and Atmospheric Administration (NOAA) and European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) available from the Maurice Lamontagne Institute sea surface temperature processing facility at a 1km resolution from 1985-2013 and from the Bedford Institute of Oceanography (BIO) Operational Remote Sensing group at a 1.5km resolution since 2014.

Positive and negative surface temperature anomalies were identified following the same approach used for bottom temperature anomalies. Compared to the bottom temperature anomalies, however, data is available thoughout the year. Only the months of April to November were included to avoid biases associated with anomalies measurements due to the presence of ice cover.

1.1.5 Sea level rise

The data used to characterize sea level rise risk comes from the global cumulative impacts assessment on habitats (Halpern et al., 2015b, 2008) and available on the NCEAS online data repository (Halpern et al., 2015a). We provide a brief summary of data and methods in this document. For more details, refer to Halpern et al. (2015b).

Sea level rise was characterized by (**???**) using NASA's satellite altimetry data (Topex/Poseidon, Jason-1&2, GFO, ERS-1&2, and Envisat missions) and available at http://www.aviso.altimetry.fr/en/data/products/ocean-indicatorsproducts/mean-sea-level/products-images.html

The rate of sea level rise ($mm/year$) was measured between 1992 and 2012 and transformed as a net change value (mm) by multiplying by the number of years considered. Only positive

values were selected under the assumption that only positive sea level rise is likely to cause environmental stress.

For the St. Lawrence, we overlaid the raw data layers (Halpern et al., 2015a) with our 1 km^2 grid cell using weighted area average.

1.2 Coastal

1.2.1 Aquaculture

Aquaculture data comes from a variety of sources in the St. Lawrence because aquaculture sites are mostly managed at the provincial level. We therefore had to gather the data on aquaculture sites from the 5 provinces dividing the St. Lawrence.

Invertebrates aquaculture is especially important in the southern and western Gulf. Fish and algae aquaculture, on the other hand, remains marginal. Considering this, we only considered invertebrates aquaculture for the aquaculture driver layer. However, if fish or algae farming were to become more important, these driver should be incorporated in future analyses as individual layers, as impacts vary between types of aquaculture.

Aquaculture activities are highly localized and potential effects do not or rarely extend beyond the location of the farms. We therefore only considered the actual location of sites to characterize the distribution of this driver. We were unable to characterize site production in terms of biomass farmed, which could provide an indication of the intensity of aquaculture activities. As such, we considered aquaculture as binary presence-absence data in our analyses.

1.2.2 Coastal development

We used lights at night as a proxy of coastal infrastructure development, as terrestrial stable lights at night represent light from human settlements and industrial sites with electricity.

The data comes from the Nighttime Lights Time Series. Nighttime light products are compiled by the Earth Observation Group at the National Oceanic and Atmospheric Administration's (NOAA) National Centers for Environmental Information (NCEI). They use glogally available nighttime data obtained from the Visible Infrared Imaging Radiometer Suite (VIIRS) Day/Night Band (DNB) of the Defense Meteorological Satellite Program (DMSP) to characterize global average radiance ($nanoWatts\ cm^{-2}\ sr^{-1}$) composite images at a 15 arcsecond (~200 m) resolution.

We used the annual Version 1 Nighttime VIIRS DNB composites between 2015 and 2016 (Group, 2019) to characterize coastal development in coastal areas of the St. Lawrence. As the effects of coastal development are likely acute in its direct vicinity, we extracted average radiance values using a 2 km buffer around grid cells within 2 km of the coast. We used a weighted area average to extract the radiance values.

1.2.3 Direct human impact

As in Halpern et al. (2008) and Halpern et al. (2015b), we used the sum of coastal populations as a proxy of direct human impact. We used Statistics Canada dissemination area population count from the 2016 census to obtain coastal population size around the St. Lawrence (**???**). Dissemination areas are the smallest standard geographic area in which census data are disseminated and they combine to cover all of Canada. The census provides population count within the boundary of each dissemination area, which we used to evaluate total coastal population.

As the effects of direct human impacts are likely acute mostly in coastal areas we calculated total population in grid cells within 2 km of the coast. Total population was measured in a 10 km buffer around each coastal cell. The total population in each buffer was the sum of intersecting dissemination areas divided by the intersection area between buffers and dissemination areas:

$$DHI_j = \sum_{k=1}^{n_j} P_k * \frac{A_{j,k}}{Atot, k}$$

where j is a buffered grid cell, k is a dissemination area intersecting j, P is the population in k, A is the area of the k overlapping with j and A_{tot} is the total area of k. This approach was favoured to reduce the effects of very large dissemination areas overlapping with buffers on a very small percentage of their total area.

1.2.4 Inorganic pollution

The data used to characterize inorganic pollution comes from the global cumulative impacts assessment on habitats (Halpern et al., 2015b, 2008) and available on the NCEAS online data repository (Halpern et al., 2015a). We provide a brief summary of data and methods in this document. For more details, refer to Halpern et al. (2015b).

Inorganic pollution was modelled using impervious surface area (*i.e.* artificial surfaces such as paved roads) under the assumption that most of this pollution source comes from urban runoff. Inorganic pollution originating from point-sources or in areas lacking paved roads is therefore not captured by this layer. The data obtained was aggregated at the watershed scale and spread into coastal and marine environments was modelled using a diffusive plume model from each watershed pourpoints (*e.g.* river mouths).

For the St. Lawrence, we overlaid the raw data layers (Halpern et al., 2015a) with our 1 km^2 grid cell using weighted area average.

1.2.5 Nutrient pollution

The data used to characterize nutrient pollution comes from the global cumulative impacts assessment on habitats (Halpern et al., 2015b, 2008) and available on the NCEAS online

data repository (Halpern et al., 2015a). We provide a brief summary of data and methods in this document. For more details, refer to Halpern et al. (2015b).

Annual fertilizer use in tonnes (t) was used as a proxy of nutrient pollution. The data used came from the Food and Agriculture Organization of the United Nations (FAO). Gaps in data were modelled using a linear regression between fertilizer and pesticides or agricultural gross domestic product (GDP). Dasymetric maps were then used to distribute fertilizer data over the landscape using 2009 data from the Moderate Resolution Imaging Spectroradiometer (MODIS) at ~500 m resolution and aggregated to watersheds. Diffusive plume models from each watershed pourpoint (*e.g.* river mouths) were then used to model the distribution and intensity of nutrient pollution in coastal and marine environments.

For the St. Lawrence, we overlaid the raw data layers (Halpern et al., 2015a) with our 1 km^2 grid cell using weighted area average.

1.2.6 Organic pollution

The data used to characterize organic pollution comes from the global cumulative impacts assessment on habitats (Halpern et al., 2015b, 2008) and available on the NCEAS online data repository (Halpern et al., 2015a). We provide a brief summary of data and methods in this document. For more details, refer to Halpern et al. (2015b).

Annual pesticide use in tonnes (t) was used as a proxy of organic pollution. The data used came from the FAO and gaps in data were modelled using a linear regression between pesticides and fertilizers or agricultural GDP. The same methodology as that used to characterized nutrient pollution was then applied to organic polllution.

For the St. Lawrence, we overlaid the raw data layers (Halpern et al., 2015a) with our 1 km^2 grid cell using weighted area average.

1.2.7 Toxic algae

The data we use to describe the risk of toxic algae comes from an expert based map delineating the areas where coastal areas are at risk from five different toxins (Bates, 2019). The map presents coastal areas at risk from 5 different toxins: 1) paralytic shellfish poisoning (PSP) toxins from the regular presence of the dinoflagellate *Alexandrium catenella* (previously *Alexandrium tamarense*) at high concentrations, 2) amnesic shellfish poisoning (ASP) toxins from domoic acid 3) diarrhetic shellfish poisoning (DSP) toxins, 4) spirolides and 5) pectenotoxins, two toxins produced by dinoflagellates occurring in the St. Lawrence.

The information provided on this expert map on the 5 toxins (Bates, 2019). was georeferenced and transformed as vecctorized objects. We calculated a toxic algae risk (T) index for each cell (x) in the 1 km^2 study grid. For each toxin (t), a value of 1 was attributed to all grid cells overlapping with areas identified at risk on the expert map and a value of 0.5 for grid cells overlapping with areas where ASP and DSP toxins were observed without exceeding legal thresholds. The value for all 5 toxins was them summed for all grid cells:

$$TA_{i,x} = \sum_{i=1}^{5} i_x$$

1.3 Fisheries

The impacts of fisheries activities in the St. Lawrence are evaluated using DFO's fisheries logbooks program (DFO, 2016). While logbooks are not mandatory for all fisheries in the St. Lawrence, they still provide a very thorough overview of the spatial distribution and intensity of fishing activities in the St. Lawrence. The data we used spans 6 years from 2010 to 2015 and details 218323 fishing events (36387 $\pm$ 3147 fishing events per year). There were 31 targetted species and a total of 53 caught species in the dataset.

Fishing activities are performed using a variety of gear types: trap, trawl, dredge, driftnet, hand line, longline, scuba diving, purse seine, seine, beach seine and jig fishing. Intensity of fishing activities was divided among gear types and based on their respective types of environmental impacts (Table S2). For example, traps and trawls have very different effects on a system. Gear classification was done using the classification presented in Halpern et al. (2008) and Halpern et al. (2015a) and is broken down into 5 distinct classes: demersal destructive (DD), demersal, non-destructive, low-bycatch (DNL), demersal, non-destructive, high-bycatch (DNH), pelagic, low-bycatch (PLB) and pelagic, high-bycatch (PHB). This categorization therefore divides the fisheries data into 5 distinct driver layers characterizing fishing activities.

Gear types can also be further classified into fixed or mobile engines based on their mobility. We used these two mobility classes to generate a buffer of impact around each fishing activity coordinates to consider potential spatial uncertainty associated with locations and the fact that mobile engines can be tracted over several kilometers during fishing activities and that we do not have the beginning and end points of mobile fishing events. Buffer sizes for fixed and mobile engine was of 200 and 2000 meters, respectively.

Table S2. Classification of gear types in the fisheries dataset based on their environmental impact and mobility

Gear type (EN)	Classification	Mobility
Trap	DNH	Fixed
Trawl	DD	Mobile
Dredge	DD	Mobile
Driftnet	PHB	Fixed
Hand lines	PLB	Fixed
Longline	PHB	Fixed
Scuba diving	DNL	Fixed
Purse seine	PLB	Fixed
Seine	DNH	Fixed
Beach seine	DNH	Fixed
Trap	DNH	Fixed

Gear type (EN)	Classification	Mobility
Jig fishing	PLB	Fixed

In order to characterize the intensity of fishing activities (FI), we used a biomass yield density index. We multipled the total annual biomass captured in each grid cell j, regardless of species, by the proportion of fishing area in each grid cell:

$$FI_j = \sum_{k=1}^{n_j} B_{tot,k} * \frac{A_{j,k}}{A_{tot,k}}$$

where j is a grid cell, k is a fishing event, B_{tot} is the total biomass of a fishing event k, A is the area of a fishing event k overlapping a cell j and A_{tot} is the total area of the fishing event k. This formula gives an intensity measurement in biomass units, which is kg in our case. Since we measure the intensity within a 1 km^2 grid cell, the intensity evaluation is in $kg * km^{-2}$. This metric distributes the biomass captured within each grid cell as a function of overlapping fishing area and provides an overview of how impacted each grid cell is in terms of extracted biomass.

1.4 Marine traffic

1.4.1 Shipping

The data used to characterize shipping comes from the global cumulative impacts assessment on habitats (Halpern et al., 2015b, 2008) and available on the NCEAS online data repository (Halpern et al., 2015a). We provide a brief summary of data and methods in this document. For more details, refer to Halpern et al. (2015b).

Two data sources were used to characterize shipping. The first set of data is gathered as part of the World Meteorological Organization Voluntary Observing Ships' (VOS) scheme (**???**). Ships participating in the program gather meteorological data along with observation location as part of an open-ocean climate dataset. The data spans 20 years and annually covers 10-20% of ships worlwide. Data used spanned 2003 to 2011.

The second set of data comes from the Automatic Identification System (AIS), an initiative launched in 2002 that sought to improve marine safety by providing mariners with real-time vessel traffic (**???**). Through the International Maritime Organization SOLAS agreement, all vessels of over 300 gross tonnage on international voyages and those carrying passengers are now required to be equipped with AIS transceivers. These transceivers use Global Positioning System technology to locate vessels every 10 minutes. The data used was from November 2010 to December 2011.

Data used come mostly from vessels that move globally (*i.e.* cargo, tanker and passenger), as they are required to carry AIS transceivers, but also include data from fishing, high-speed, pleasure and support classes. Shipping intensity was evaluated as the number of fishing

tracks at a 0.1 decimal degrees resolution. For more details on data and methods used, consult (**???**).

For the St. Lawrence, we overlaid the raw data layers (Halpern et al., 2015a) with our 1 km^2 grid cell using weighted area average.

1.4.2 Invasive species

The data used to characterize invasive species risk comes from the global cumulative impacts assessment on habitats (Halpern et al., 2015b, 2008) and available on the NCEAS online data repository (Halpern et al., 2015a). We provide a brief summary of data and methods in this document. For more details, refer to supplementary materials provided in Halpern et al. (2008) and Halpern et al. (2015b).

Cargo volume was used as a proxy of invasion risk under the assumption that risk of invasion is proportional to tonnes of goods transferred through ports. Cargo throughput in metric tonnes for the year 2011 was accessed through a variety of sources (see supplementary material in Halpern et al., 2015b for more details) and cross-matched with entries in the World Port Index database (WPI; available from the National Geospatial-Intelligence Agency). A gap-filling procedure using linear regression and sets of predictors related to port volume and available in the WPI dataset was then applied to the WPI dataset to predict missing cargo volume entries. Finally, volume data was distributed in marine environments adjacent to ports using a diffusive plume model with an exponential decay function that set the maximum spread distance to approximately 1000 km. The plume model was then clipped to areas less than 60 m deep, as invasives species are more likely to invade shallow areas.

For the St. Lawrence, we overlaid the raw data layers (Halpern et al., 2015a) with our 1 km^2 grid cell using weighted area average.

1.4.3 Marine pollution

The data used to characterize marine pollution risk comes from the global cumulative impacts assessment on habitats (Halpern et al., 2015b, 2008) and available on the NCEAS online data repository (Halpern et al., 2015a). Marine pollution was considered to be mainly driver by the shipping industry. As such, the driver layer was constructed by combining the shipping (*i.e.* shipping lanes) and invasives species (*i.e.* cargo volume) layers. invasive. For more details, refer to supplementary materials provided in Halpern et al. (2008) and Halpern et al. (2015b).

For the St. Lawrence, we overlaid the raw data layers (Halpern et al., 2015a) with our 1 km^2 grid cell using weighted area average.

2 Driver intensity and distribution

We evaluated the frequency distribution of each drivers to verify whether data should be transformed (Figure S1). In light of this, we log-transformed the following driver layers

- Coastal development
- Direct human impact
- All fisheries data
- Hypoxia
- Inorganic pollution
- Invasive species
- Nutrient pollution
- Organic pollution
- Sea bottom temperature anomalies
- Shipping

To allow for relative intensity comparison, all driver layers were subsequently normalized between 0 and 1 using the 99th quantile to further control for extreme values (Figure S2).

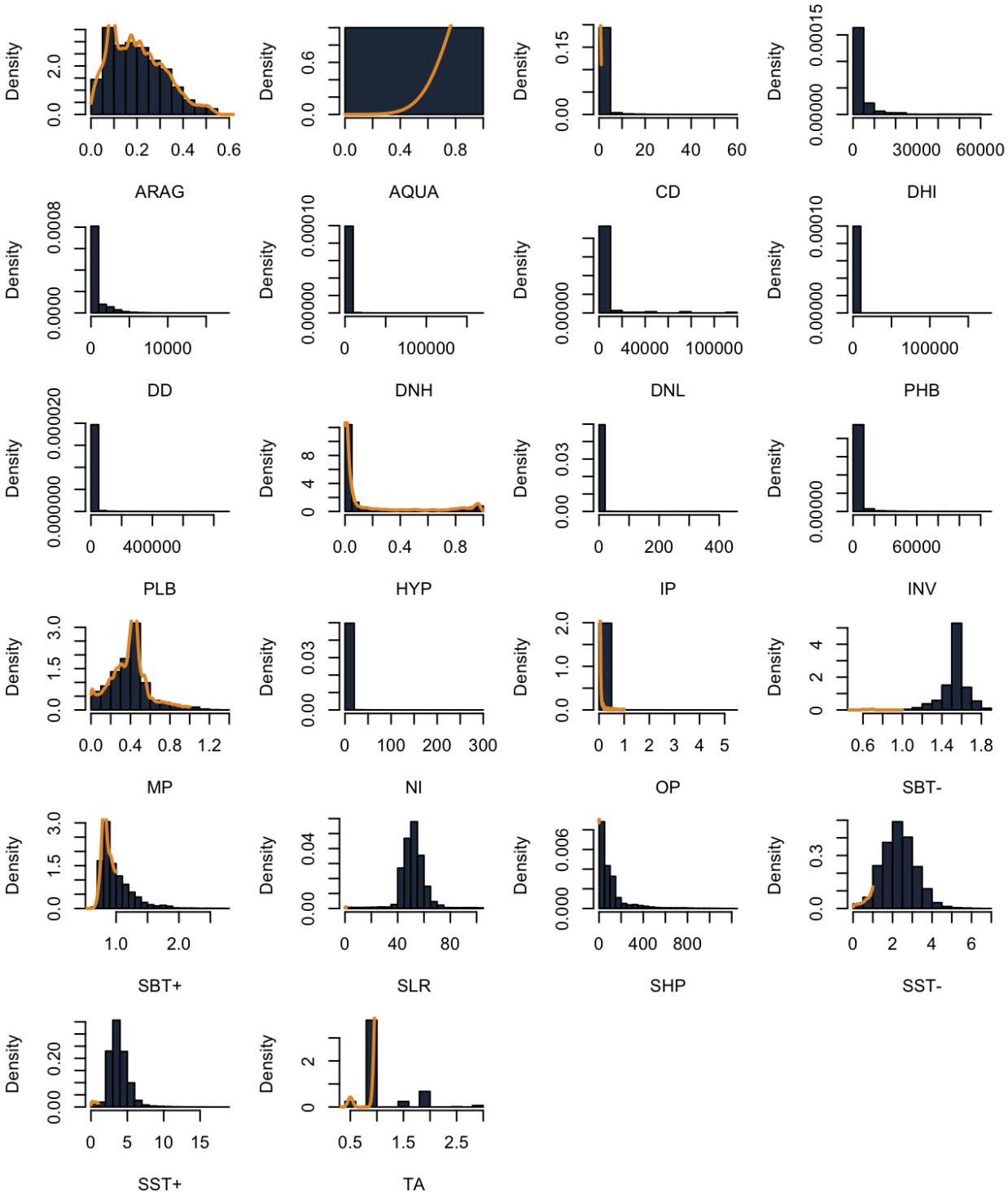

Figure 1: Frequency distribution of the untransformed data for all driver layers.

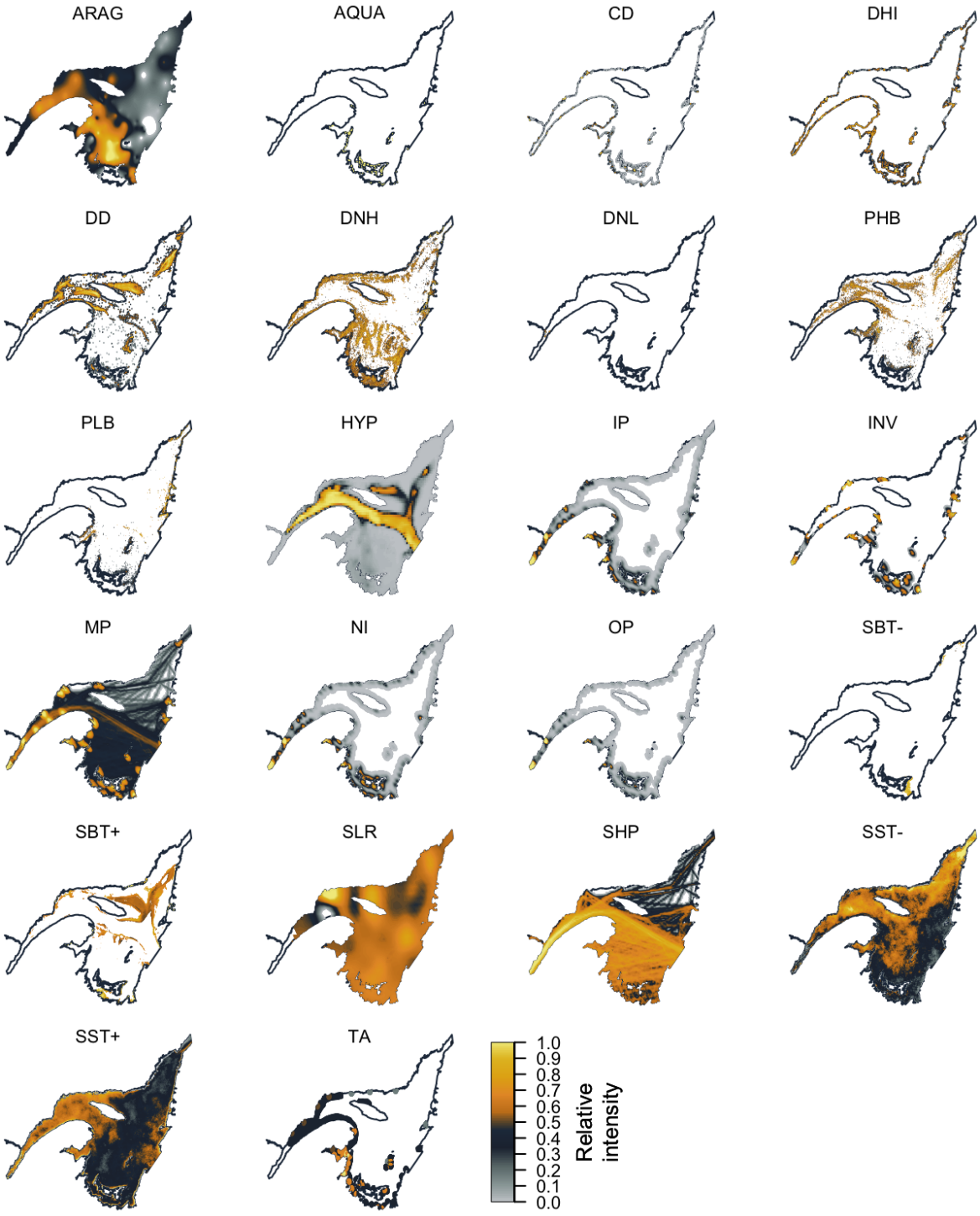

Figure 2: Distribution and intensity of transformed and normalized drivers in the Estuary and Gulf of St. Lawrence available on *eDrivers*.

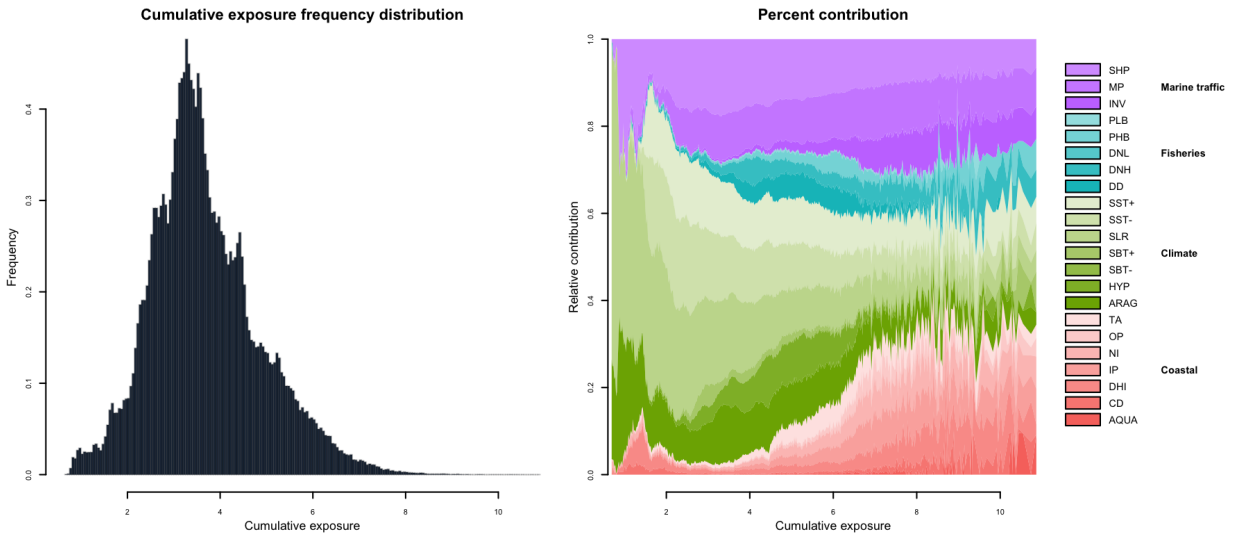

Figure 3: Frequency distribution of cumulative exposure (*i.e.* sum of normalized driver intensity in each grid cell) and percent contribution of each driver to the frequency distribution of cumulative exposure in the Estuary and Gulf of St. Lawrence

3 Cumulative exposure

4 Threat complexes

4.1 Clustering

We identified regions with similar cumulative exposure regimes, referred to as threat complexes after (Bowler et al., 2019), using a partional *k-medoids* clustering algorithm, CLARA (CLustering for Large Applications; Kaufman and Rousseeuw, 1990), which was designed to use with large datasets. The CLARA algorithm uses the PAM (Partition Around Medoids) algorithm on a sample from the original dataset to identify a set of k objects that are representative of all other objects, *i.e.* medoids (also referred to as examplars) and that are central to the cluster they represent. The goal of the algorithm is to iteratively minimize dissimilarity between objects in the cluster and the representative medoid. Sets of k medoids are thus selected and replaced in order to iteratively minimize the intra-cluster dissimilarity. With the CLARA algorithm, this process is repeated and iterations are compared on the basis of their average dissimilarity between objects and cluster medoids to select the optimal set of k medoids that minimizes average dissimilarity. We performed the analyses for a total of 100 iterations using samples of 10000 observations, representing ~5% of the drivers dataset. Analyses were performed using the *cluster* R package (Maechler et al., 2018).

Because the algorithm identifies clusters by minimizing dissimilarity around a set of k medoids, it requires a pre-defined number of clusters. The appropriate number of clusters k was thus tested using values ranging from 2 to 10 and validated by selecting the number of

clusters that maximized the average silhouette width (Kaufman and Rousseeuw, 1990) and minimized the total within-cluster sum of squares (WSS; Figure S4).

We also validated the clustering by comparing *k-medoids* clustering with *k-means* clustering with the *Lloyd* algorithm (Lloyd, 1982). The *k-means* approach is similar to the *k-medoids*, but identifies observations belonging to a cluster iteratively by minimizing the mean intra-cluster squared distance until it converges to an optimal solution. We used 25 random sets and set a maximum of 1000 iterations for the analysis. Analyses were performed using the *stats* R package (R Core Team, 2018). We used the same validation procedure to select the optimal number of clusters k than with the *k-medoids* clustering (Figure S4).

Both the *k-medoids* and *k-means* clustering validation procedure suggest that there are 7 distinc threat complexes in the St. Lawrence (Figure S4). While the comparison of of the spatial distribution of clusters results in only a 66% cell correspondance between the *k-medoids* and *k-means* approaches, the bulk of the difference is attributable to a single undivided cluster that combines most of the southern and norther Gulf in a single cluster in favor of a more localised cluster by the *k-means* algorithm (Figure S5). While *k-means* algorithms are much faster than *k-medoids* since it does not need a pairwise dissimilarity measurement of all observations, it is also more sensitive to outliers since it focuses on the mean rather than a centroid. We therefore opted the use of *k-medoids* clustering.

4.2 Inter-cluster dissimilarity

The difference between clusters was explored by measuring the total inter-cluster dissimilarity and the contribution of each driver to the total inter-cluster dissimilarity using a similarity percentage analysis (SIMPER) with Bray-Curtis dissimilarity (Figure S6; Clarke, 1993). As the drivers dataset is too large, we used a bootstrap procedure for the SIMPER analysis, randomly selecting 5% of each cluster to run the analysis and repeating the process over 300 iterations. We also compared the mean intensity of each driver within each cluster to better capture the inter-cluster dissimilarity. Analyses were performed using the *vegan* R package (Oksanen et al., 2018).

4.3 Intra-cluster similarity

Intra-cluster similarity was evaluated using the Bray-Curtis similarity index (Figure S7). As with the inter-cluster dissimilarity, we used a bootstrap procedure for the intra-cluster similarity, randomly selecting 5% of each cluster observation to run the analysis and repeating the process over 300 iterations. We however did not use the bootstraping procedure for clusters 2 and 6, as they both contain less than 10000 observations, making the calculation time more manageable.

References

Bates (2019). Reference to come, toxic algae.

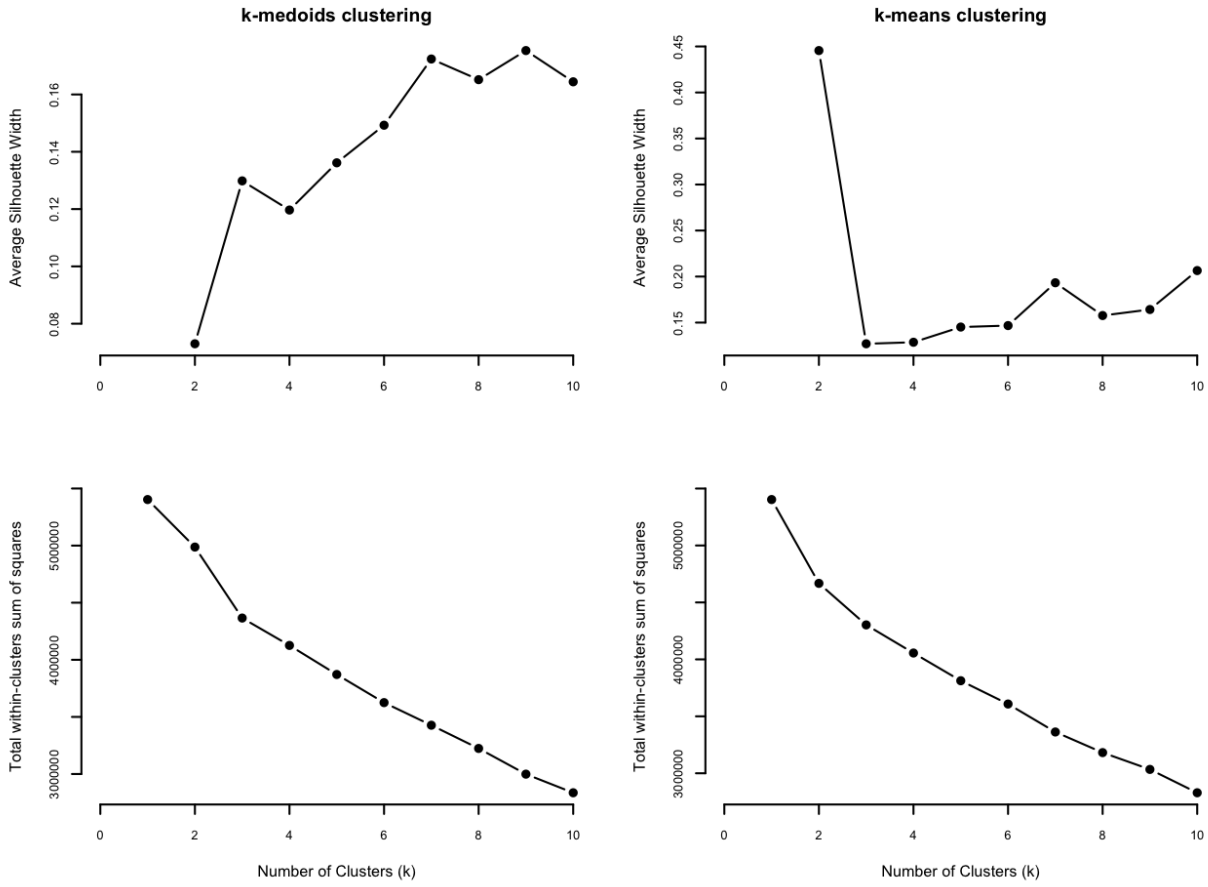

Figure 4: Validation procedure for the *k-medoids* and *k-means* clustering algorithms based on the number of cluster that maximizes average silhouette width (upper panels; Kaufman and Rousseeuw, 1990) and minimizes the total within-cluster sum of squares (WSS; lower panels).

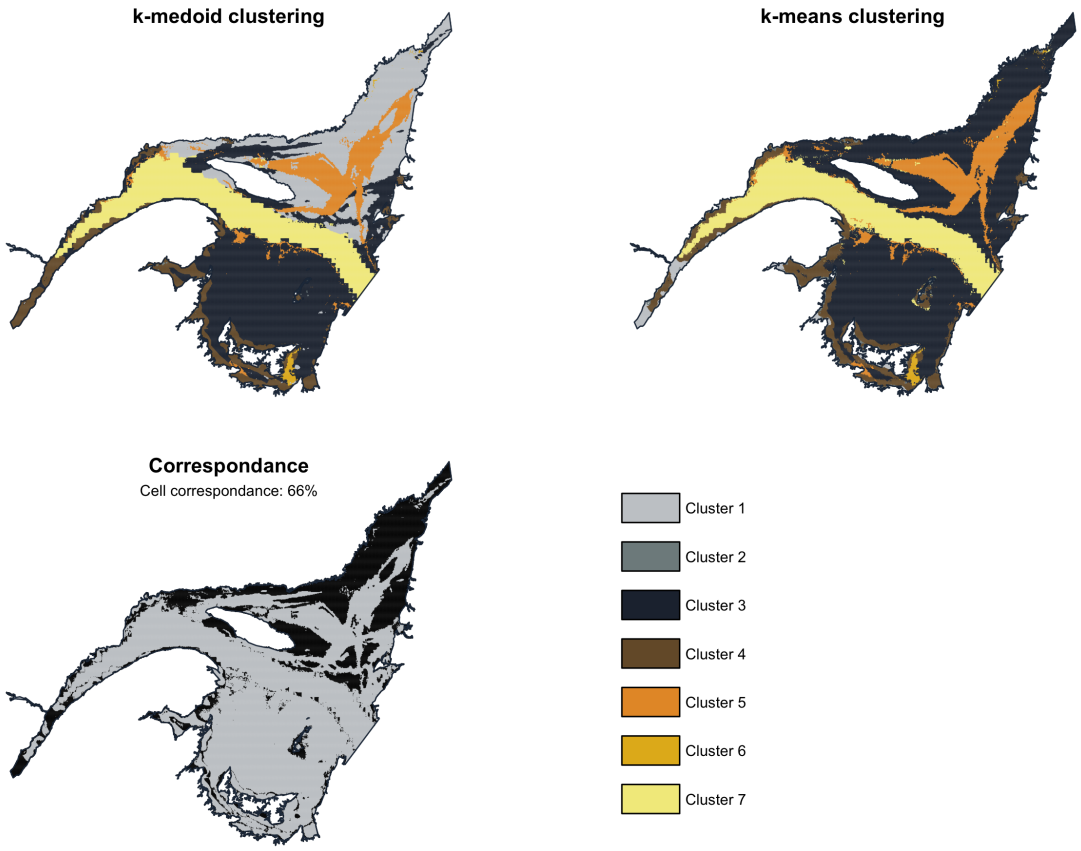

Figure 5: Comparison of the distribution of the 7 clusters identified using the *k-medoids* and *k-means* clustering algoriths and the percent grid cell correspondance between the two approaches

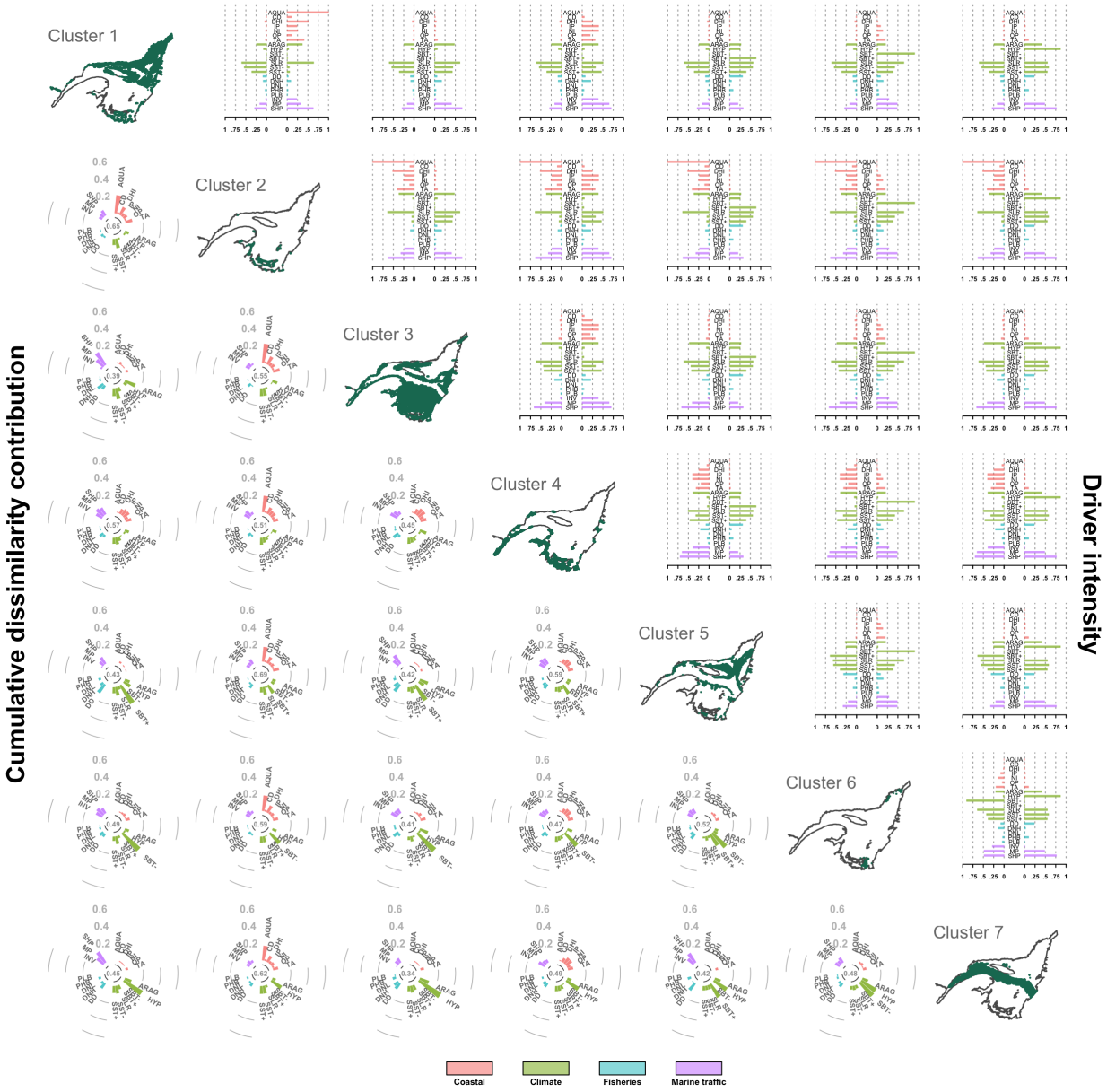

Figure 6: Evaluation of inter-cluster dissimilarity using a similarity percentage analysis (SIMPER) with Bray-Curtis dissimilarity (Clarke, 1993). The figure diagonal presents the distribution of the 7 clusters identified using the *k-medoids* clustering algorithm. The lower triangle shows all combinations of inter-cluster dissimilarity with circular barplots showing the percent contribution to total dissimilarity of each driver and with the total inter-cluster dissimilarity in the center of the barplots. The upper triangle shows the average relative intensity of each driver for all driver combinations, with barplots to the left and the right representing the row and columns clusters, respectively.

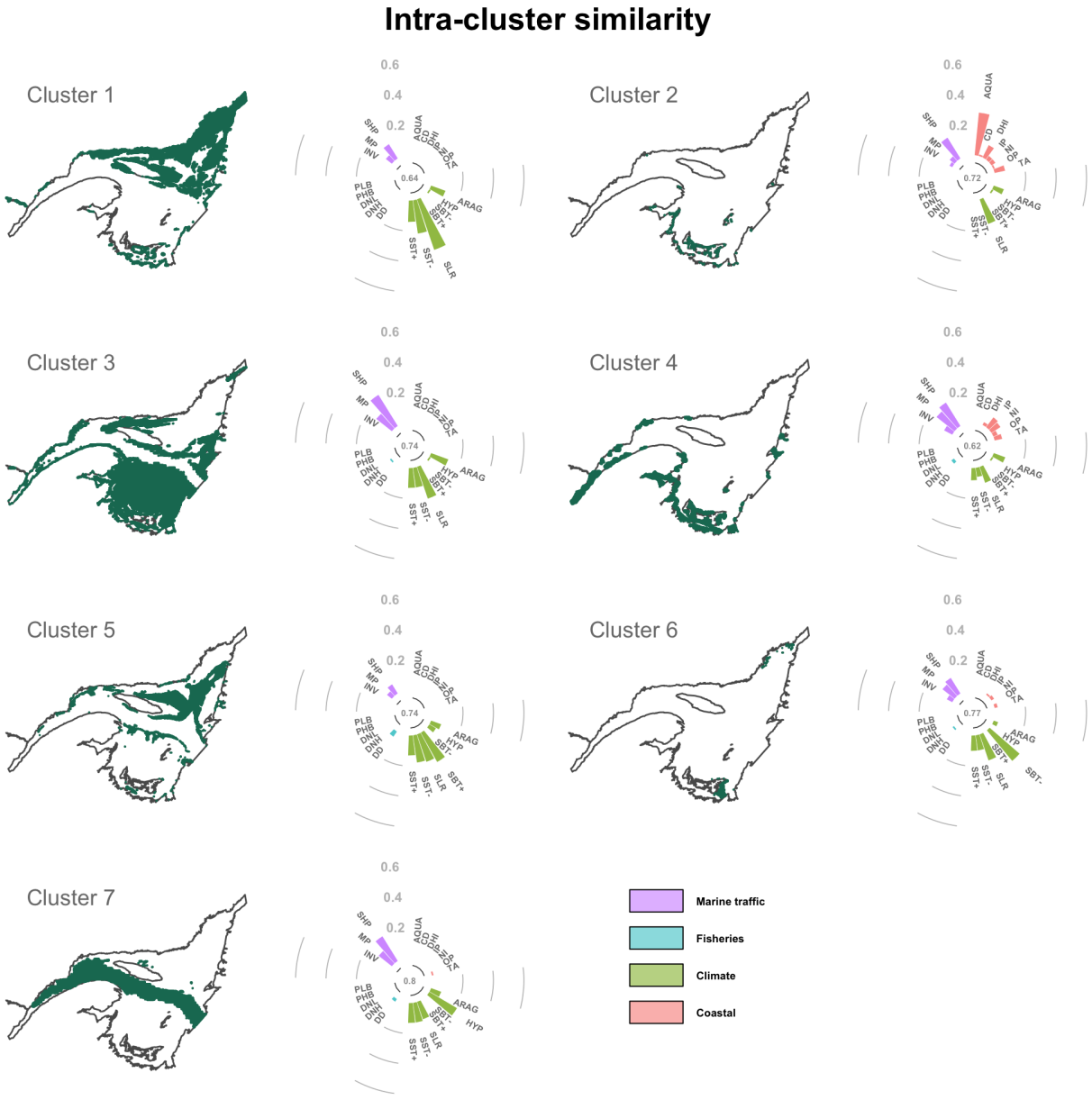

Figure 7: Evaluation of intra-cluster similarity using the Bray-Curtis similarity index. The distribution of all 7 clusters is presented along with circular barplots showing the percent contribution to total similarity of each driver and with the total intra-cluster similarity in the center of the barplots.

Bowler, D., Bjorkmann, A., Dornelas, M., Myers-Smith, I., Navarro, L., Niamir, A., et al. (2019). The geography of the Anthropocene differs between the land and the sea. *bioRxiv*. doi:10.1101/432880.

Clarke, K. R. (1993). Non-parametric multivariate analyses of changes in community structure. *Australian Journal of Ecology* 18, 117–143. doi:10.1111/j.1442-9993.1993.tb00438.x.

DFO (2016). Departement of Fisheries and Oceans Canada's Fisheries and Oceans Canada Zonal Interchange File Format (ZIFF) data. A compilation of landing data from logbook data between 2010 and 2015.

Diaz, R. J., and Rosenberg, R. (1995). Marine benthic hypoxia: A review of its ecological effects and the behavioural responses of benthic macrofauna. *Oceanography and marine biology. An annual review* 33, 245–303.

Dutil, J.-D., Proulx, S., Chouinard, P.-M., and Borcard, D. (2011). A Hierarchical Classification of the Seabed Based on Physiographic and Oceanographic Features in the St. Lawrence. Department of Fisheries; Oceans.

Dutil, J.-D., Proulx, S., Galbraith, P. S., Chassé, J., Lambert, N., and Laurian, C. (2012). Coastal and epipelagic habitats of the estuary and Gulf of St. Lawrence. Department of Fisheries; Oceans.

Galbraith, P. S., Chassé, J., Caverhill, C., Nicot, P., Gilbert, D., Lefaivre, D., et al. (2018). Physical Oceanographic Conditions in the Gulf of St. Lawrence during 2017. Department of Fisheries; Oceans Available at: http://www.dfo-mpo.gc.ca/csas-sccs/Publications/ResDocs-DocRech/2018/2018_050-fra.html [Accessed November 26, 2018].

Group, E. O. (2019). Version 1 VIIRS Day/Night Band Nighttime Lights. NOAA National Centers for Environmental Information (NCEI).

Halpern, B. S., Frazier, M., Potapenko, J., Casey, K. S., Koenig, K., Longo, C., et al. (2015a). *Cumulative human impacts: Raw stressor data (2008 and 2013)*. KNB Data Repository doi:10.5063/f1s180fs.

Halpern, B. S., Frazier, M., Potapenko, J., Casey, K. S., Koenig, K., Longo, C., et al. (2015b). Spatial and temporal changes in cumulative human impacts on the world's ocean. *Nature Communications* 6. doi:10.1038/ncomms8615.

Halpern, B. S., Walbridge, S., Selkoe, K. A., Kappel, C. V., Micheli, F., D\textquotesingleAgrosa, C., et al. (2008). A Global Map of Human Impact on Marine Ecosystems. *Science* 319, 948–952. doi:10.1126/science.1149345.

Kaufman, L., and Rousseeuw, P. (1990). Finding groups in data: An introduction to cluster analysis. in (Wiley, New York), 342.

Lloyd, S. (1982). Least squares quantization in PCM. *IEEE Transactions on Information Theory* 28, 129–137. doi:10.1109/TIT.1982.1056489.

Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., and Hornik, K. (2018). *Cluster: Cluster Analysis Basics and Extensions*.

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., et al. (2018). *Vegan: Community Ecology Package*. Available at: https://CRAN.R-project.org/package=vegan.

R Core Team (2018). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing Available at: https://www.R-project.org/.

Starr, M. (2019). Reference to come, aragonite and hypoxia.

Tukey, J. W. (1977). *Exploratory Data Analysis*. Reading, Massachusetts: Addison-Wesley.