Next Generation Planning - Structuring and Sharing Environmental Drivers Data for the St. Lawrence

David Beauchesne[1,2,], Rémi M. Daigle[2], Steve Vissault[3], Dominique Gravel[3], Andréane Bastien[4], Simon Bélanger[5], Pascal Bernatchez[5], Marjolaine Blais [6], Hugo Bourdages[6], Clément Chion[7], Peter S. Galbraith[6], Benjamin Halpern[8,9], Camille Lavoie[2], Christopher W. McKindsey[6], Alfonso Mucci[10], Simon Pineault[11], Michel Starr[6], Anne-Sophie Ste-Marie[4], Philippe Archambault[2]*

2019-04-23

[1]Institut des sciences de la mer, Université du Québec à Rimouski, Rimouski, QC, Canada
[2]Québec Océan, Département de biologie, Université Laval, Québec, QC, Canada
[3]Département de biologie, Université de Sherbrooke, Sherbrooke, QC, Canada
[4]St. Lawrence Global Observatory, Rimouski, QC, Canada
[5]Département de Biologie, Chimie et Géographie, Université du Québec à Rimouski, Rimouski, QC, Canada
[6]Département des Sciences naturelles, Université du Québec en Outaouais, Gatineau, QC, Canada
[7]Fisheries and Oceans Canada, Maurice Lamontagne Institute, Mont-Joli, QC, Canada
[8]National Center for Ecological Analysis and Synthesis, University of California, Santa Barbara, CA, United States
[9]Bren School of Environmental Science and Management, University of California, Santa Barbara, CA, United States
[10]Department of Earth & Planetary Sciences, McGill University, Montréal, QC, Canada
[11]Ministère Environnement et Lutte contre les changements climatiques, Québec, QC, Canada

Correspondence:
David Beauchesne

david.beauchesne@uqar.com

Number of figures: 5

Number of tables: 1

1 Abstract

The St. Lawrence System is a vast and complex social-ecological system providing a wealth of ecosystem services sustaining numerous economic sectors. Related human pressures combine with climate driven environmental changes to produce intricate exposure regimes of overlapping and potentially interacting environmental drivers (*e.g.* hypoxia and fisheries) threatening ecosystems. There is a critical need to systematically characterize the distribution, intensity and overlap between drivers to support integrative initiatives such as ecosystem-based management. While portals providing knowledge on individual drivers exist, platforms collating comparable and interoperable knowledges on multiple drivers remain conspicuously missing. This paper presents two distinct, but related objectives. The first is to characterize the distribution and intensity of drivers in the St. Lawrence System. The second is to launch *eDrivers*, an open knowledge platform gathering experts committed to structuring, standardizing and sharing knowledges on drivers in support of science and management. We gathered data on 22 coastal, climate, fisheries and marine traffic drivers through collaborations, existing environmental initiatives and open data portals. We show that few areas of the St. Lawrence are free of cumulative exposure. The Estuary, the Anticosti Gyre and coastal areas are particularly exposed, especially in the vicinity of urban centers. We identified 6 areas of distinct cumulative exposure regime that show that certain drivers typically co-occur in different regions of the St. Lawrence and that coastal areas are exposed to all driver types. Of particular concern are two threat complexes capturing most exposure hotspots that show the convergence of contrasting exposure regimes at the head of the Laurentian Channel. These observations are destined to improve as *eDrivers* evolves through time to address knowledge gaps and refine current driver layers. In an effort to share the knowledge acquired and to ensure the lasting relevance of the description of drivers presented in this manuscript, *eDrivers* was built on a series of guiding principles upholding existing data management and open science standards. Ultimately, we believe that *eDrivers* represents a much needed solution that could radically influence broad scale research and management practices by increasing data accessibility and interoperability and by increasing research and decision-making efficiency.

2 Introduction

The St. Lawrence System, formed by one of the largest estuaries in the world and a vast interior sea, is a complex social-ecological system characterized by highly variable environmental conditions and oceanographic processes, both in space and time (Dufour and Ouellet, 2007; El-Sabh and Silverberg, 1990; White and Johns, 1997). It thus offers a unique and heterogeneous array of habitats suited for the establishment of diverse and productive ecological communities (Savenkoff et al., 2000). As a result, the St. Lawrence System provides a wealth of ecosystem services that have historically and contemporarily benefited the Canadian economy. It sustains a rich fisheries industry targeting more than 50 species, serves as the gateway to eastern North-America by granting access to more than 40 ports and the most densely populated Canadian region, hosts a booming tourism industry and an expanding aquaculture production, fosters emerging activities and boasts a yet untapped hydrocarbon potential (Beauchesne et al., 2016; Schloss et al., 2017 @archambault2017). With major investments recently made and more forthcoming in economic and infrastructure development and research (*e.g.* Québec, 2015; RQM, 2018), an intensification of the human footprint is expected in the St. Lawrence System.

As elsewhere in the world (see Halpern et al., 2015b), this intensifying human footprint will likely result in increasingly intricate environmental exposure regimes, *i.e.* suites of overlapping and potentially interacting environmental drivers threatening ecosystems, habitats or ecological communities. Drivers, often referred to as stressors or pressures, are any externalities that affect environmental processes and disturb natural systems. Drivers may originate from natural or human-induced biophysical processes (*e.g.* sea surface temperature anomalies and hypoxia) or directly from anthropogenic activities (*e.g.* fisheries and marine pollution). The potential for complex interactions between drivers is the largest uncertainty when studying or predicting environmental change (Côté et al., 2016; Darling and Côté, 2008). The effects of multiple drivers can combine non-linearly and result in effects that are greater (synergistic effect) or lower (antagonistic effect) than the sum of individual effects (Côté et al., 2016; Crain et al., 2008; Darling and Côté, 2008).

The uncertainty associated with complex driver interactions must therefore be taken into account when investigating environmental impacts (Côté et al., 2016), yet most research on driver effects in marine environments remains overwhelmingly focused on single driver assessments (O'Brien et al., 2019). Arising from the expected increase in environmental exposure and the experiences of past and current ecological tragedies such as the collapse of cod fisheries (Dempsey et al., 2018; Frank et al., 2005) and the decline of the beluga and right whale populations (**???**) is a need to characterize the distribution, intensity and overlap between drivers in the St. Lawrence System. This will provide critical information on areas most exposed to cumulative drivers and on the interaction potential of drivers in the St. Lawrence. It is also a necessary step towards the holistic and integrated management of the St. Lawrence System.

Gathering environmental data for large scale, systematic initiatives can, however, be a very challenging and time consuming - not to say painful - process. On one hand, there is an overwhelming and expanding wealth of data available. Such information overload may inhibit our ability to make decisions based on scientific information, promote massive effort duplication, disproportionately appropriate research funds to certain sectors, and obscure knowledge gaps amid a sea of information (Eppler and Mengis, 2004). On the other hand, crucial data are lacking and remain largely unavailable or inaccessible for a variety of reasons, including proprietary rights, lack of organizational time, capacity and training, and in some rare cases unwillingness to share, curtailing our ability for appropriate decision-making.

The current digital infrastructure is highly decentralized and the data management and sharing practices are highly heterogeneous, preventing us from maximizing benefits from research investments (Wilkinson et al., 2016). Yet there now exists multiple initiatives that address this issue by assembling, organizing and sharing vast arrays of environmental data. Biotic data can be accessed through web portals such as the Ocean Biogeographic Information System (OBIS; OBIS, 2018), the Global Biodiversity Information Facility (GBIF; GBIF, 2018), the Global Biotic Interactions platform (GloBI; Poelen et al., 2014), and the World Register of Marine Species (WoRMS; WoRMS Editorial Board, 2017). Abiotic data can also be accessed through WorldClim (Hijmans et al., 2005), Bio-ORACLE (Tyberghein et al., 2012), and MARSPEC (Sbrocco and Barber, 2013). Initiatives focused on sharing environmental data for specific areas also exist, such as the St. Lawrence Global Observatory (SLGO; https://ogsl.ca/en), the European Marine Observation and Data Network (EMODnet; http://www.emodnet.eu/) and the U.S. Integrated Ocean Observing System (IOOS; https://ioos.noaa.gov/). Essential environmental parameters are also organized, coordinated and acquired through global initiatives like the Group on Earth Observations Biodiversity Observation Network (GEO BON; Scholes et al., 2012), the Census of Marine Life (CoML; CoML, 2010) and the Global Ocean Observing System (GOOS; http://www.goosocean.org).

However, equivalent platforms for environmental drivers have largely focused on single drivers (e.g. Global Fishing Watch) and platforms collating data and knowledge on multiple drivers in a comparable and interoperable way remain conspicuously missing (but see Halpern et al., 2015b). This is in spite of integerated management and assessment approaches needing efficient data reporting, standardized data management practices and tools tailored to the study of the effects of multiple drivers (Dafforn et al., 2016; Stock et al., 2018). An additional objective thus emerged in the process of addressing our initial goals: sharing the knowledge gathered through the description of drivers in the St. Lawrence.

This manuscript thus has two distinct, but related, goals. The first is to characterize the distribution and intensity of drivers in the St. Lawrence System in order to: 1) identify areas of high cumulative exposure, 2) characterize areas with similar cumulative exposure regimes, dubbed threat complexes (Bowler et al., 2019) and 3) identify drivers that are likely to interact in the St. Lawrence. The second goal is to present how we are sharing the knowledge acquired by launching an open knowledge platform, *eDrivers*, and use this paper as a call for collaboration to any person, group or organization that might be interested in contributing to this initiative.

3 Materials and Methods

3.1 Drivers

The list of drivers for which we sought data was informed by a global cumulative impact assessment initiative (Halpern et al., 2015b, 2008), regional holistic evaluations of the state of the St. Lawrence (Benoît et al., 2012; Dufour and Ouellet, 2007), and communications with regional experts. Through the data gathering process, we developed and continue to develop collaborations with regional experts and data holders. We also use global data from the global cumulative impact assessment initiative (Halpern et al., 2015b, 2008) available from the National Center for Ecological Analysis and Synthesis (NCEAS) online data repository (Table 1; Halpern et al., 2015a). We selected global data that were unavailable at the regional scale and that were available at a resolution adequate for use at the scale of the St. Lawrence (*e.g.* marine pollution).

We were able to characterize the intensity and distribution of 22 drivers (Table 1;) and we are actively working on updating or developing additional driver layers. Drivers incorporated in the analyses are varied in origin, *i.e.* from terrestrial (*e.g.* nutrient input) to marine (*e.g.* shipping), and from large scale biophysical processes (*e.g.* temperature anomalies) to localized anthropogenic activities (*e.g.* fisheries). Drivers were divided into 4 groups: coastal, climate, fisheries and marine traffic (Table 1). All data layers and methodologies are described in the supplementary materials.

Drivers with non-normal frequency distributions were log-transformed (Figure S1) and all drivers were scaled between 0 and 1 to allow driver comparisons. The 99th quantile of individual driver distribution was used as the upper bound for scaling to control for extreme values and produce maps of individual drivers (Figure S2). All drivers were embedded in a regular grid composed of 245604 $1km^2$ exagonal cells to construct the integrated dataset used for the analyses.

3.2 Cumulative exposure

Areas with high cumulative exposure (objective 1) were identified by evaluating the cumulative footprint of combined drivers and by identifying hotspots of cumulative footprint. Cumulative footprint (F) was defined as the sum of the scaled intensity of all drivers in each grid cell:

$$F_x = \sum_{i=1}^{n} D_{i,x}$$

where x is a grid cell, i is a driver and D is the scaled intensity of driver i.

Cumulative hotspots (H) were defined as the number of drivers in each grid cell with scaled intensity contained over their respective 80th percentile:

$$H_x = \sum_{i=1}^{n} \mathbb{1}(D_{i,x} \, \epsilon \, P_{80,D_i})$$

where, x is a grid cell, i is a driver and D is the scaled intensity of driver i and P_{80,D_i} is the 80th percentile of driver i. Hotspots thus identify areas where drivers are co-occurring at high relative intensities.

3.3 Driver interactions

The distribution of driver interactions was investigated through the spatial overlap of combinations of drivers using the cumulative footprint (F) equation with pairs of drivers. The intensity at which pairs of drivers co-occur was evaluated using a two-dimensional kernel density estimate. As there are 231 pairwise combinations between 22 drivers, we focus on a single example using hypoxia and demersal destructive fisheries, two drivers known to occur mainly in deeper areas of the St. Lawrence and, hence, an interaction between the effects of the two drivers could be anticipated. Note, however, that *eDrivers* will provide the capacity to compare any combinations of drivers.

3.4 Threat complexes

Natural systems are likely to host multiple overlapping drivers. The dimensionality of the problem can quickly rise as additional drivers are considered. In order to decrease the dimensionality of the integrated dataset, we identify threat complexes, *i.e.* regions with similar cumulative exposure regimes, using a clustering approach (*e.g.* see Bowler et al., 2019).

3.4.1 Clustering

Threat complexes were identified using a partional *k-medoids* clustering algorithm, CLARA (CLustering for Large Applications; Kaufman and Rousseeuw, 1990), which was designed for 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 and that are central to the cluster they represent. The goal of the algorithm is to iteratively minimize intra-cluster dissimilarity. Iterations are compared on the basis of the average dissimilarity between cluster objects and representative medoid to select the optimal set of k medoids that minimizes average dissimilarity. We used 100 iterations using samples of 10000 observations (*i.e.* ~5% of observations) to identify clusters. Analyses were performed using the *cluster* R package (Maechler et al., 2018).

Partitional clustering algorithms require a user-defined number of clusters. Values of k ranging from 2 to 10 were tested 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 (Figure S3).

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 S3).

While *k-means* algorithms are more efficient since they do not compute pairwise dissimilarities, it is more sensitive to outliers through the use of the mean rather than a centroid. We therefore favored the use of the *k-medoids* algorithm, but used the *k-means* to validate clusters.

3.4.2 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 S4; 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).

3.4.3 Intra-cluster similarity

Intra-cluster similarity was evaluated using the Bray-Curtis similarity index (Figure S5). 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 with less than 10000 observations since computation time was manageable.

4 Results and discussion

4.1 Cumulative exposure

Apart from the northeastern Gulf, the cumulative footprint of drivers is ubiquitous in the St. Lawrence (Figure 1). Cumulative exposure is generally highest along the coast (Figure 1), with hotposts located in the vicinity of coastal cities (Figure **??**). In general, offshore areas are less exposed to cumulative drivers, with the Estuary and the Anticosti Gyre being notable exceptions (Figures 1 and **??**). This is not to say that offshore areas are free of exposure, as most of the St. Lawrence is exposed to multiple overlapping drivers (Figures 1 and **??**). For example, it is worthy to note high cumulative footprint observed at the heads of the Anticosti and Esquiman Channels (Figure 1).

These results are consistent with observations elsewhere in the world, where cumulative driver exposure conspicuously arises from and markedly intensifies close to coastal cities and at the mouth of rivers draining highly populated areas (e.g. Halpern et al., 2015b; Feist and Levin, 2016; Mach et al., 2017; Stock et al., 2018). These are areas where human activities (*e.g.* coastal development and shipping) and footprint (*e.g.* pollution runoff) are the most intense (Feist and Levin, 2016), and on which is overlaid a background of natural disturbances (Micheli et al., 2016). They are also the areas in which the most dramatic increases in exposure are expected, with populations increasing more rapidly along the coast than inland (Feist and Levin, 2016). In the St. Lawrence, large coastal cities are mostly located along the Estuary and the southwestern Gulf, while the northeastern Gulf is largely uninhabited or home to small coastal communities.

As for offshore exposure, the Estuary, along with the St. Lawrence River, provide access to and serve as the primary drainage outflow of the Great Lakes Basin, the most densely populated region in Canada (Canada, 2017). Most marine traffic thus converges to the Estuary.

While we cannot ascertain that high exposure areas are the most impacted, we can safely predict that these are the areas where studying ecosystem state will be the most complex due to the uncertainty associated with driver interactions, an uncertainty bound to increase rapidly with the number of interacting drivers (Côté et al., 2016).

4.2 Driver interactions

Hypoxia is mainly distributed in the Laurentian, Anticosti and Esquiman Channels, with the head of the Channels most exposed to hypoxia (Figure 2A). Demersal destructive fisheries are located along the Laurentian Channel, the heads of the Anticosti and Esquiman Channels and around the Magdalen Islands (Figure 2B). By combining both drivers, we can observe that hypoxia and demersal destructive fisheries overlap mostly at high relative intensity (Figure 2D) in the vicinity of the Anticosti Gyre and the heads of the Esquiman and Anticosti Channels (Figure 2C). The ease with which this figure can be created using *eDrivers* is demonstrated in box 1.

Fisheries in the St. Lawrence have historically affected biodiversity distribution and habitat quality (Moritz et al., 2015). Concurrently, hypoxia decreases overall habitat quality, but triggers species-dependent responses ranging from adaptation (*e.g.* northern shrimp *Pandalus borealis* and Greenland halibut *Reinhardtius hippoglossoides*; Pillet et al., 2016) to reduced growth rates (Dupont-Prinet et al., 2013) and avoidance of oxygen-depleted habitats (*e.g.* Atlantic cod *Gadus morhua*; Chabot and Claireaux, 2008) to increased mortality (*e.g.* sessile benthic invertebrates; Eby et al., 2005; Belley et al., 2010; Gilbert et al., 2007). Certain species may thus be adversely affected by fisheries and withstand hypoxia but still experience a decrease in prey availability, while others may be deleteriously affected by the compounded effect of both drivers (De Leo et al., 2017).

4.3 Threat complexes

While informative, the hypoxia-fisheries example focuses on a single pair of drivers and falls short of the number of drivers overlapping at high intensities throughout the St. Lawrence (Figure **??**). The number of drivers overlapping in the St. Lawrence increases with cumulative exposure (Figure S3). Areas with high exposure such as the Estuary, the Anticosti Gyre and the southwestern Gulf coastline (Figure 1 and **??**) are thus areas where driver interactions are most likely, and where they can arise between a host of different drivers.

The identification of threat complexes provides a crucial tool to simplify the multi-dimensional complexity of overlapping drivers to areas exposed to similar suites of drivers (Bowler et al., 2019). This may prove critical for a better understanding the state of species, habitats and ecosystems located within or moving through threat complexes and exposed to the combined effects of all drivers typical to those areas.

Six distinct threat complexes were identified in the St. Lawrence using the *k-medoids* and *k-means* algorithms (Figures S4, S5). Based on their distribution and representative drivers, threat complexes can be divided into 3 offshore and 3 coastal complexes (Figures 3, S6 and S7). Coastal threat complexes (1 to 3; Figure 3) include all types of drivers besides hypoxia and are the most exposed threat complexes, both in terms of driver overlap and intensity. Threat complex 2 is differentiated from other complexes by the presence of aquaculture sites. Threat complex 1 encompasses the coastline and is characterized by higher direct human impact (*i.e.* population density). Threat complex 3 is the most exposed complex and has a distribution similar to the most exposed coastal hotspots (Figure **??**). This complex is characterized by high intensities of land-based drivers (*e.g.* nutrient input), demersal non-destructive high-bycath fisheries (*e.g.* trap fishing), climate drivers and marine traffic drivers in the vicinity of ports.

Offshore threat complexes (4 to 6; Figure 3) are generally characterized by high intensity climate and marine traffic drivers. Threat complex 4 is differentiated by demersal non-destructive high-bycath fisherie, higher marine traffic drivers compared to complex 5 and generally corresponds to the whole Southern Gulf. Threat complex 5 is characterized by more fisheries types (*i.e.* demersal destructive and pelagic high-bycatch), generally lower intensity marine traffic drivers and is located almost exclusively in the Northern Gulf. Finally, threat complex 6 corresponds primarily to the Laurentian Channel, a deep (250-500 m) and long (1250 km) submerged valley connecting the Estuary to the Atlantic. It also incorporates parts of the Esquiman and Anticosti Channels, two deep channels that branch off from the the Laurential Channel to the north towards the Arctic and the north of Anticosti Island, respectively. This threat complex is the most exposed offshore threat complex and includes all offshore hotspots (Figure **??**). This complex is characterized by high intensity hypoxia, marine traffic and pollution, as well as demersal destructive and pelagic high-bycatch fisheries.

Of particular concern are threat complexes 3 and 6, which are the two most exposed complexes in the St. Lawrence and are characterized by distinct cumulative exposure regimes Between them, they capture most of the coastal and offshore hotspots identified in the St. Lawrence and discussed above.

They also offer some insight into the potential importance of considering spatial dynamics in areas intersecting multiple threat complexes. For example, threat complexes 3 and 6 meet at the mouth of the River Saguenay. This area is particularly dynamic, with deep Atlantic waters advected through estuarine circulation mixing with surface waters from the Great Lakes Basin and the Saguenay River (Dufour and Ouellet, 2007). This results in the convergence of climate drivers from the bottom of the Laurentian Channel and marine traffic drivers (threat complex 6) with terrestrial run-off from river outflows and direct human impacts (threat complex 3). This dynamic area is also highly productive and hosts large aggregations of krill exploited by numerous fish, marine mammal and marine bird species (**???**; **???**). It might therefore be reasonable to expect highly unpredictable environmental effects in this area.

5 Open Knowledge Platform: *eDrivers*

Sharing the knowledge acquired through the description of drivers in the St. Lawrence quickly emerged as a priority to curtail the need to reach dozens of experts across multiple organizations and over extensive periods of time to assemble the data needed to apply integrated research and management. It is also a requirement to ensure that this manuscript will not become a quickly outdated snapshot of drivers distribution and intensity in the St. Lawrence System, but rather serve as a stepping stone towards an adaptive and ever-improving collection of knowledges.

As such, we are launching *eDrivers*, an open knowledge platform focused on sharing knowledge on the distribution and intensity of drivers and on gathering a community of experts committed to structuring, standardizing and sharing knowledge on drivers in support of science and management. In launching this intiative, our objective is to uphold the highest existing standards of data management and open science. We identified four guiding principles to meet this objective and that guide the structure of the platform (Figure 4).

5.1 Unity and inclusiveness

Why: Operating over such large scales in time, space and subject matter requires a vast and diverse expertise that cannot possibly be possessed by any one individual or organization. Consequently, we envision an initiative that seeks to mobilize all individuals and entities with relevant expertise.

How: By promoting, consolidating and working with experts involved in existing and highly valuable environmental initiatives already in place in the St. Lawrence. Notable examples of environmental initiatives are the annual review of physical (Galbraith et al., 2018), chemical, and biological (Blais et al., 2018) oceanographic conditions in the St. Lawrence, the fisheries monitoring program (DFO, 2016), the annual groundfish and shrimp multidisciplinary survey (Bourdages et al., 2018), the characterisation of benthic (Dutil et al., 2011), epipelagic and coastal (Dutil et al., 2012) habitats of the St. Lawrence, and Canada's shoreline classification (ECCC, 2018). There are also nascent efforts to share information on several human activities in the St. Lawrence such as the Marine Spatial Data Infrastructure portal, which provides data on zoning, shipping, port activities, and other human activities in Canadian waters, including the St. Lawrence system (Canada, 2016).

By working with existing data portals whose objective is to share environmental data. We are thus collaborating actively with the St. Lawrence Global Observatory (SLGO) to develop the initiative and to host the platform on their web portal (https://ogsl.ca/en). The mission of SLGO is to promote and facilitate the accessibility, dissemination and exchange of official and quality data and information on the St. Lawrence ecosystem through the networking of organisations and data holders to meet their needs and those of users, to improve knowledge and to assist decision-making in areas such as public safety, climate change, transportation, resources and biodiversity conservation. The SLGO is also one of three regional association spearheading the Canadian Integrated Ocean Observing System (CIOOS; http://meopar.ca/research/cioos-call-for-proposals/), which will focus on integrating oceanographic data from multiple sources to make them accessible to end-users and to enable the national coordination of ocean observing efforts by integrating isolated or inaccessible data, and by identifying gaps or duplications in observations and research efforts. We are also developing collaborations with the Portal on water knowledge (http://www.environnement.gouv.qc.ca/

eau/portail/), an initiative from the Québec provincial government. This portal aims at collecting and sharing accurate, complete and updated ressources on water and aquatic ecosystems to support the mandate of relevant actors and stakeholders working in water and aquatic ecosystems management.

By actively inviting, seeking, and developing collaborations as well as encouraging constructive criticism from the inception and throughout the lifetime of the platform.

By inviting external community contributions (Figure 4). External researchers or entities wishing to submit marine data will be able to do so through the SLGO web portal (https://ogsl.ca/en). Submissions through other data portals will also be accepted either through the development of data sharing agreements or with the caveat that shared data are under an open source license and that they adhere to the platform data standards.

5.2 Findability, accessibility, interoperability and reusability

Why: Open data has been propelled to the forefront of scientific research in an era of open, collaborative and reproducible science. By moving towards large scale, cross-disciplinary research and management projects, there is a growing need to increase the efficiency of data discovery, access, interoperability and analysis (Reichman et al., 2011; Wilkinson et al., 2016). Our goal is to foster efficient and functional open science by creating a fully open, transparent and replicable open knowledge platform.

How: By building an infrastructure adhering to the FAIR Data Principles, which states that data and metadata must be Findable, Accessible, Interoperable and Reusable. These principles focus on the ability of humans and machines to automatically find and (re)use data and knowledge (Wilkinson et al., 2016).

By making data and associated tools accessible through a variety of ways: the SLGO web portal (https://ogsl.ca/en), two R packages called *eDrivers* (https://github.com/orgs/eDrivers/eDrivers) and *eDriversEx* (https://github.com/orgs/eDrivers/eDriversEx) to access the data through SLGO's API and to provide analytical tools to explore data, respectively, and a Shiny application (https://david-beauchesne.shinyapps.io/eDriversApp/) to explore drivers data interactively (Figure 4). Note that the data are currently contained within and accessible through the *eDrivers* R package only, as we are actively working to allow users to download selected layers from SLGO's web portal and geoserver. The functions available in *eDrivers* to access the data have however been developed to ensure forward compatibility once the data migrate to SLGO's geoserver.

By defining clear data and metadata standards and specifications to support the regional standardization of current and future protocols and practices and to favour interoperability with national and international initiatives like the Essential Ocean Variables (EOV) identified by the Global Ocean Observing System [GOOS; http://www.goosocean.org]. As such, we will adopt the metadata standard currently targetted for the upcoming CIOOS, *i.e.* the North American Profile of ISO 19115:2014 - Geographic information - Metadata, a schema favoured for geospatial data in Canada and the United-States.

By providing version control and code access to the workflows set up to generate driver layers from raw data, the R packages and the Shiny application through a GitHub organization called *eDrivers* (https://github.com/orgs/eDrivers/).

5.3 Adaptiveness

Why: In the face of uncertainty and in an effort to address impending environmental changes, adaptive management has been identified as the chief strategy to guide efficient decision-making (*e.g.* Costanza et al., 1998; Jones, 2016; Keith et al., 2011; Margules and Pressey, 2000) and has already been discussed in the context of multi-drivers and cumulative impact assessments (Beauchesne et al., 2016; Côté et al., 2016; Halpern et al., 2015b; Schloss et al., 2017). Adaptive management can only be truly achieved through a commitment to adaptive monitoring and data reporting (Halpern et al., 2012; Lubchenco and Grorud-Colvert, 2015; Margules and Pressey, 2000). We further contend that adaptive management requires the development of adaptive monitoring tools and infrastructures, which we seek to address through a continuously-evolving platform.

How: By setting up mechanisms structuring cyclic reviews of platform content, for the integration of new material (*e.g.* data and methods) as it becomes available or accessible, and by striving to provide time-series data that are crucial to assess temporal trends and potentially early-warning signals of ecosystem change (Figure 4).

5.4 Recognition

Why: Like peer-reviewed publications, data must also be given its due importance in scientific endeavors and thus be considered as legitimate citable products contributing to the overall scientific output of data providers (Data Citation Standards and PractOut of Mind: The Current Sices, 2013; FORCE11, 2014). Appropriate citations should therefore be provided for all data layers used and shared by the platform.

How: By adhering to the Data Citation Principles (FORCE11, 2014), which focus on citation practices that provide appropriate credit to data products.

6 Perspectives

Understanding how ecosystem state will be affected by global change requires a comprehensive understanding of how threats are distributed and interact in space and time, which in turn hinges on appropriate data tailored to multi-driver studies (Bowler et al., 2019; Dafforn et al., 2016; Stock et al., 2018). In the St. Lawrence, we found that few areas are free of cumulative exposure and that the whole Estuary, the Anticosti Gyre and coastal southwestern Gulf are particularly exposed to cumulative drivers, especially close to urban areas. We also identified six geographically distinct threat complexes that display similar cumulative exposure regimes. These complexes reveal that coastal areas are particularly exposed to all types of drivers and that multiple drivers typically co-occur in space. These resuts allow us to efficiently identify areas in need of heightened scrutiny from a science and management perspective.

Through *eDrivers*, these observations will be iteratively improved towards an increasingly robust assessment of cumulative exposure and threat complexes as gaps in knowledge are addressed or approaches to describe drivers are refined. Arguably, the most meaningful benefit anticipated from *eDrivers* will be the gain in efficient access to comparable data-based knowledge on the exposure of ecosystems to multiple threats. This could pay quick scientific and management dividends by drawing on the knowledge and efforts of a wide range of contributors, by expanding avenues of scientific inquiry, by decreasing overall effort duplication and research costs, and by increasing research efficiency (Franzoni and Sauermann, 2014).

Critically - and we emphasize this point - *eDrivers* will allow the scientific and governmental communities to identify key knowledge gaps that will assist in prioritizing and optimizing research efforts. Ultimately, we believe that *eDrivers* will operationalize evidence-based decision-making by streamlining data management and research, allowing science output to be available and interpretable on a time scale relevant to management (see Reichman et al., 2011). The platform will thus greatly facilitate the application of broad scale, holistic research and management approaches such as ecosystem-based management, strategic environmental assessments and social-ecological metanetworks (*e.g.* Halpern et al., 2015b; Dee et al., 2017; Jones, 2016).

Significant effort is still needed to bring our vision to fruition. Foremost is to maintain our efforts to foster collaborations, develop platform content and identify key knowledge gaps. A fair and efficient organizational structure will be developed in order to manage *eDrivers* as a community and appropriate funding must be secured to continue building this community and ensure the long-term viability of the initiative, although the partnership with SLGO partly addresses this issue.

Finally, terrestrial and coastal environments must be incorporated, as sources of stress within those habitats extend to the marine environments. Moreover, despite coastal areas being recognized as the most exposed to environmental threats, we continue to delineate terrestrial and marine realms, considering coastlines as an impermeable barrier. While there is a sensible rationale for this division, we must strive to eliminate it if we are to appropriately study and predict the impacts of global change (*e.g.* see Bowler et al., 2019).

Despite the challenges and work ahead, we are hopeful that this initiative will be very successful. Ultimately, *eDrivers* represents a much needed solution to address important issues in data management that could radically shift broad scale research and management practices towards efficient, adaptive and holistic ecosystem-based management in the St. Lawrence and elsewhere in the world. All it requires to be successful is for the scientific and political communities to fully commit to open knowledge, adaptive monitoring and, most of all, an integrated vision of ecosystem management.

7 Acknowledgements

We thank the Fond de Recherche Québécois Nature et Technologie (FRQNT) and the Natural Science and Engineering Council of Canada (CRSNG) for financial support. This project is also supported by Québec Océan, the Quebec Centre for Biodiversity Science (QCBS), Takuvik, and the Notre Golfe networks. This research is also sponsored by the NSERC Canadian Healthy Oceans Network and its Partners: Department of Fisheries and Oceans Canada and INREST (representing the Port of Sept-Îles and City of Sept-Îles).

8 Author contributions statement

TO WRITE

9 Conflict of interest statement

The authors declare that the submitted work was carried out in the absence of any personal, professional or financial relationships that could potentially be construed as a conflict of interest.

10 Listings

Box 1. Code snippet demonstrating how to use the *eDrivers* to reproduce figure 2 in R.

```
# Install and load eDrivers package
devtools::install_github('eDrivers/eDrivers')
library(eDrivers)

# Load data
drivers <- fetchDrivers(drivers = c('hypoxia','fishDD'))

# Get data from `eDrivers` class object
driverData <- getData(drivers)

# Normalize data
driverData <- driverData / cellStats(driverData, 'max')

# Visualize data and combination
plot(driverData$fishDD)   # Demersal destructive fisheries
plot(driverData$hypoxia)  # Hypoxia
plot(sum(driverData))     # Combination

# Identify values > 0 and not NAs
driverData$fishDD[driverData$fishDD < 0] <- NA
driverData$fishDNH[driverData$hypoxia < 0] <- NA
id0 <- !is.na(values(driverData$fishDD)) &
       !is.na(values(driverData$hypoxia))

# 2D kernel for driver co-intensity
library(MASS)
coInt  <- kde2d(x = values(driverData$fishDD)[id0],
                y = values(driverData$hypoxia)[id0],
                n = 500, lims = c(0, 1, 0, 1))
image(coInt, zlim = c(0,max(coInt$z)))

# Driver density distribution
plot(density(driverData$fishDD[id0]))   # Demersal destructive
plot(density(driverData$hypoxia[id0]))  # Hypoxia
```

547 # 11 Figures

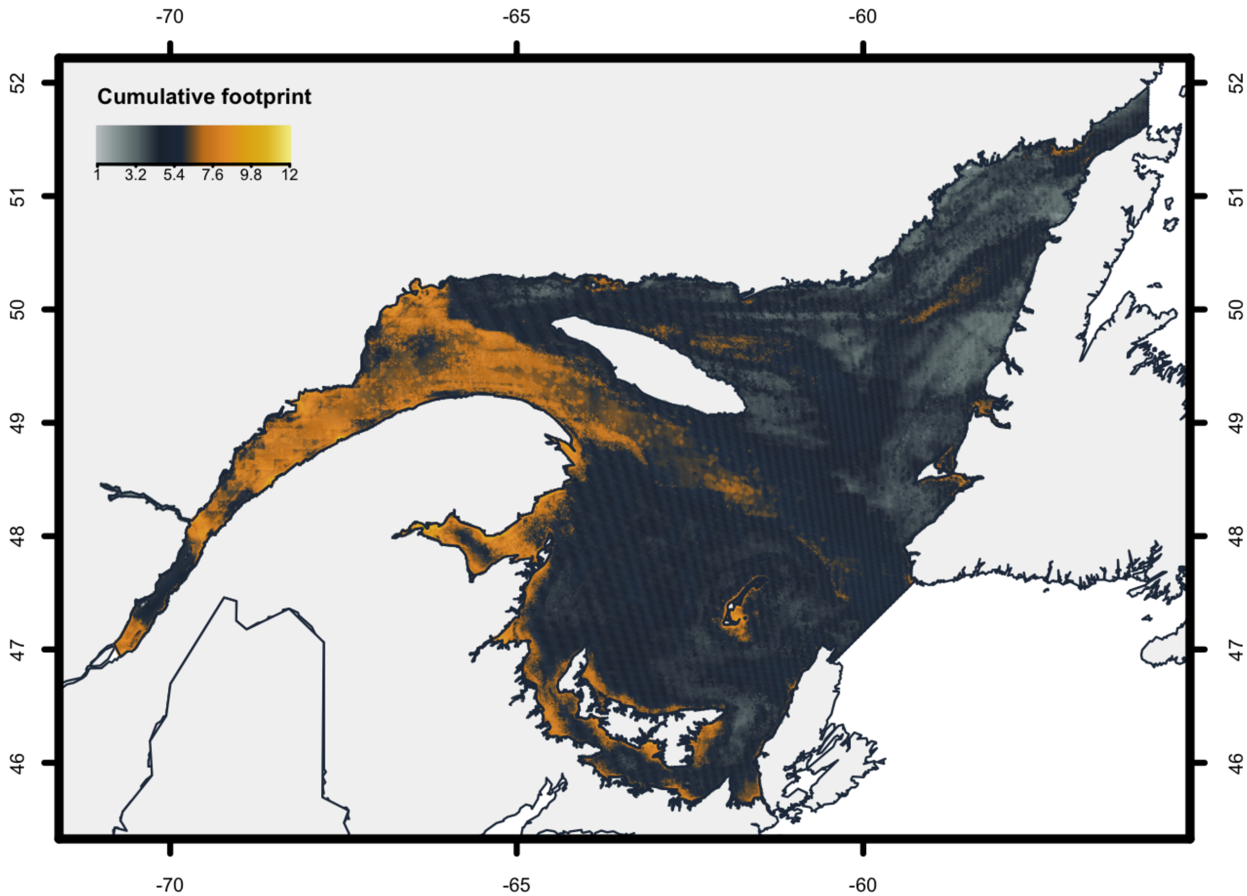

Figure 1: Distribution of driver footprint in the Estuary and Gulf of St. Lawrence. The footprint is measured by summing the relative intensity of all drivers for each grid cell: $F_d = \sum_{i=1}^{n} D_i$. Driver layers were also log-transformed when their frequency distribution was non-normal. All driver layers were also normalized between 0 and 1 using the 99th quantile to allow for direct comparison of relative intensities between drivers and control for extreme values.

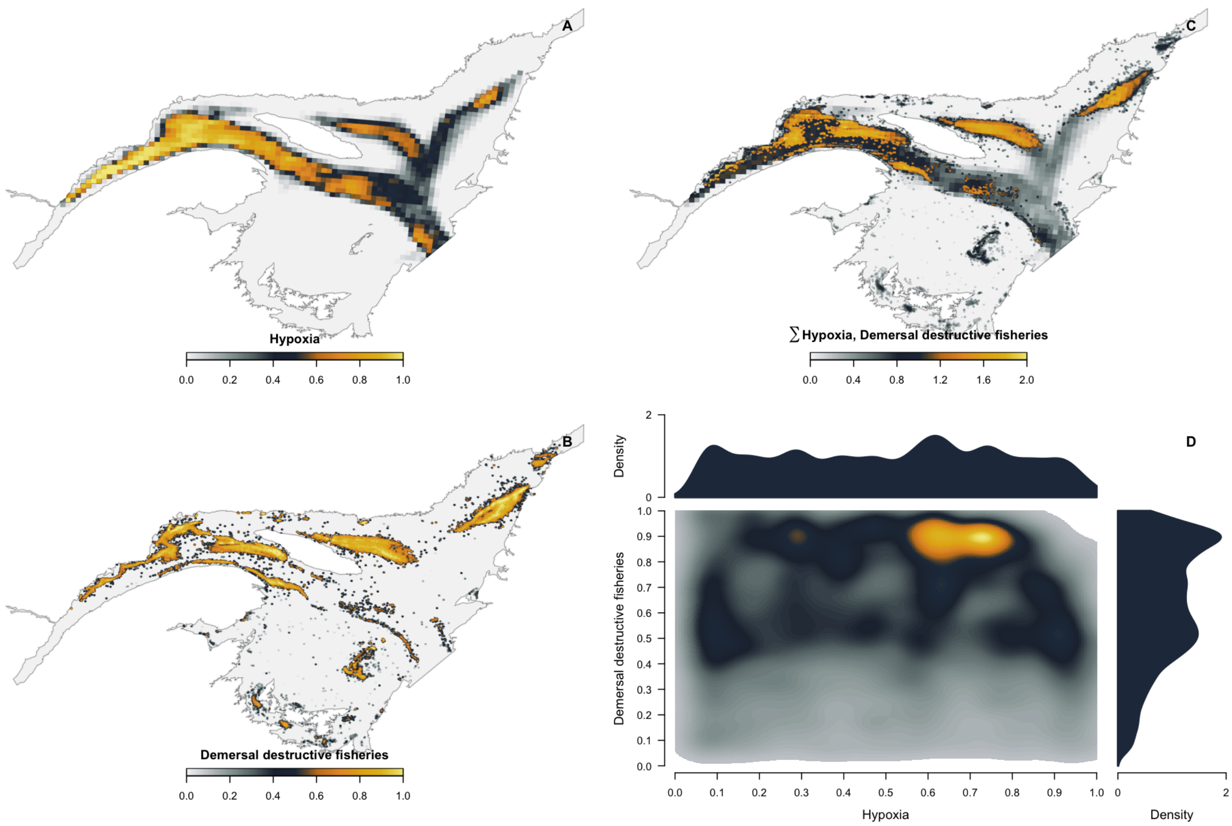

Figure 2: Example of platform content using the spatial distribution and the intensity of hypoxia and demersal destructive fisheries in the St. Lawrence. An index of hypoxia (**A**) was created using bottom-water dissolved oxygen between 2013 and 2017 [@blais2018]. Demersal destructive fisheries (*i.e.* trawl and dredges) (**B**) intensity was evaluated from fisheries catch data collected between 2010 and 2015 used to measure annual area weighted total biomass (kg) in i km^2 grid cells [@dfo2016]. See supplementary materials for more information on specific methodologies. Relative hypoxic stress and demersal destructive fisheries intensity was summed (**C**) to visualize their combined spatial distribution and intensity. Finally, individual density and the co-intensity of hypoxia and demersal destructive fisheries was investigated with a two-dimensional kernel analysis (**D**).

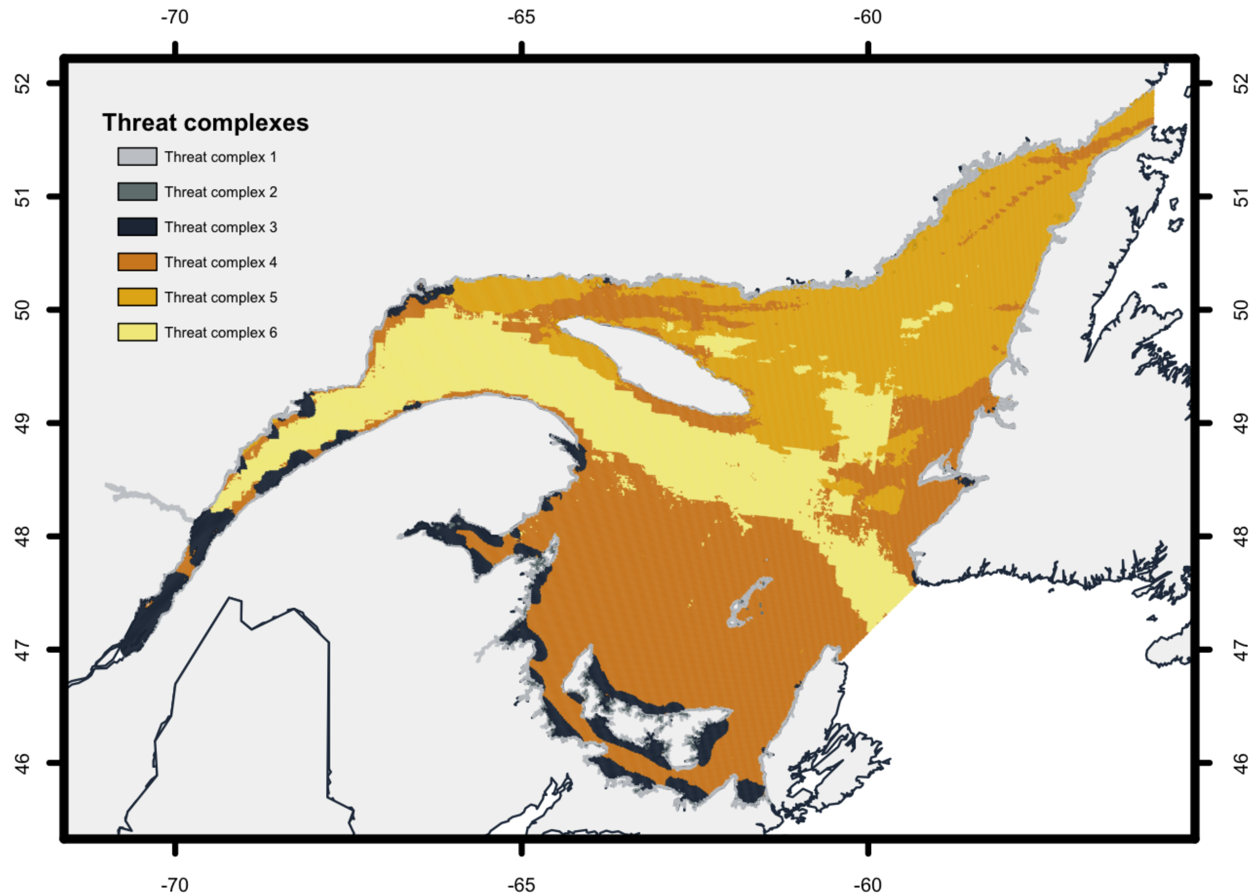

Figure 3: Distribution of threat complexes in the Estuary and Gulf of St. Lawrence. Threat complexes [a term coined by @bowler2019] are areas with similar cumulative driver exposure regimes. Threat complexes were identified using a partitional clustering algorithm [@kaufman1990]. The appropriate number of clusters (k) was tested using a range of values and validated by selecting the number of clusters that maximized the average silhouette width [@kaufman1990] and minimized the total within-cluster sum of squares (WSS). Six distinct threat complexes were identified in the St. Lawrence. The partitional clustering algorithm analysis were performed using the *cluster* R packages [@maechler2018]. Refer to the Supplementary Materials for more details.

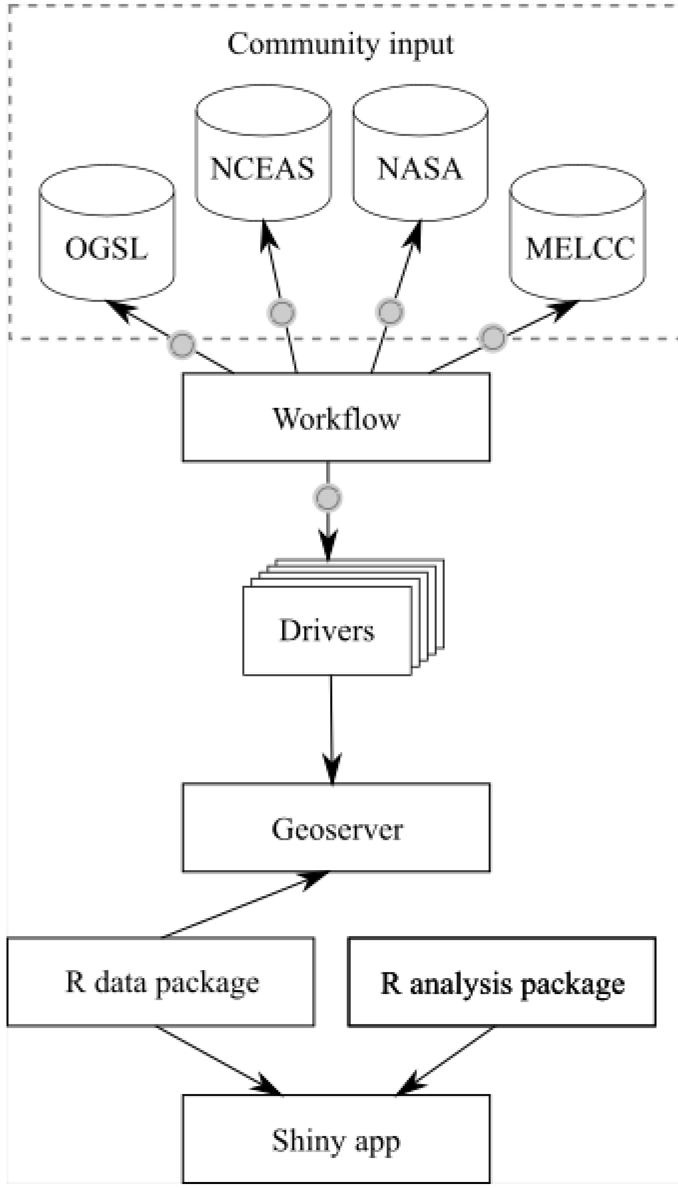

Figure 4: Diagram of the platform structure. Community input in the form of raw data is accessed through the St. Lawrence Global Observatory (SLGO); https://ogsl.ca/en) repository - the platform host - or through open access repositories (*e.g.* NASA data). The raw data are then processed through a workflow hosted on the *eDrivers* GitHub organization (https://github.com/orgs/eDrivers/). Data processing may be as simple as data rescaling (*e.g.* night lights) or make use of more complex methodologies (*e.g.* acidification). All data is then hosted on SLGO's geoserver and accessible through their API. We developed a R package called *eDrivers* to access the driver layers through R and we are actively developing a second R package called *eDriversEx* that includes analytical tools to explore drivers data. Finally, we have developed a Shiny application that allows users to explore drivers data interactively (https://david-beauchesne.shinyapps.io/eDriversApp/). All R components of the project are hosted and available on the *eDrivers* GitHub organization.

12 Tables

Table 1. List of drivers currently available on *eDrivers* and used for the analyses presented in this paper.

Groups	Drivers	Spatial resolution	Temporal resolution	Years	Units	Source
Climate	Aragonite	Lat/long	August-September	2018	Ω *Aragonite*	(Starr, 2019)
Climate	Hypoxia	Lat/long	August-September	2018	$ml\ L^{-1}$	(Starr, 2019)
Climate	Sea bottom temperature	~2 km^2	Monthly	1981-2010 vs. 2013-2017	n negative anomalies	(Galbraith et al., 2018)
Climate	Sea bottom temperature	~2 km^2	Monthly	1981-2010 vs. 2013-2017	n positie anomalies	(Galbraith et al., 2018)
Climate	Sea surface temperature	~2 km^2	Monthly	1981-2010 vs. 2013-2017	n negative anomalies	(Galbraith et al., 2018)
Climate	Sea surface temperature	~2 km^2	Monthly	1981-2010 vs. 2013-2017	n positive anomalies	(Galbraith et al., 2018)
Climate	Sea water level	Modeled 0.25 degree	10 days	1992-2012	mm	(Halpern et al., 2015a)

Groups	Drivers	Spatial resolution	Temporal resolution	Years	Units	Source
Coastal	Aquaculture	Lat/long	-	Variable, between 1990-2016	$presence - absence$	***TBD***
Coastal	Coastal development	15 arc-second	Annual	2015-2016	$nanoWatts\ cm^{-2}\ sr^{-1}$	(Group, 2019)
Coastal	Direct human impact	Modeled 1 km^2	Annual	2011	$population\ 10km^{-2}$	(Halpern et al., 2015a)
Coastal	Inorganic pollution	Modeled 1 km^2	Annual	2000-2001	***TBD***	(Halpern et al., 2015a)
Coastal	Nutrient import	Modeled 1 km^2	Annual	2007-2010	t fertilizer	(Halpern et al., 2015a)
Coastal	Organic pollution	Modeled 1 km^2	Annual	2007-2010	t pesticide	(Halpern et al., 2015a)
Coastal	Toxic algae	-	-	-	Expert based	(Bates, 2019)
Fisheries	Demersal, destructive	Lat/long	Event based	2010-2015	kg	(DFO, 2016)
Fisheries	Demersal, non-destructive, high-bycatch	Lat/long	Event based	2010-2015	kg	(DFO, 2016)
Fisheries	Demersal, non-destructive, low-bycatch	Lat/long	Event based	2010-2015	kg	(DFO, 2016)

Groups	Drivers	Spatial resolution	Temporal resolution	Years	Units	Source
Fisheries	Pelagic, high-bycatch	Lat/long	Event based	2010-2015	kg	(DFO, 2016)
Fisheries	Pelagic, low-bycatch	Lat/long	Event based	2010-2015	kg	(DFO, 2016)
Marine traffic	Invasive species	Modeled 1 km^2	Annual	2011	t port volume	(Halpern et al., 2015a)
Marine traffic	Marine pollution	Modeled 1 km^2	Event based & annual	2003-2011 & 2011	n lanes + t port volume	(Halpern et al., 2015a)
Marine traffic	Shipping	0.1 degree	Event based	2003-2011	n lanes	(Halpern et al., 2015a)

13 References

Archambault, P., Schloss, I. R., Grant, C., and Plante, S. eds. (2017). *Les hydrocarbures dans le golfe du Saint-Laurent - Enjeux sociaux, économiques et environnementaux.* Notre Golfe, Rimouski, Qc, Canada.

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

Beauchesne, D., Grant, C., Gravel, D., and Archambault, P. (2016). L'évaluation des impacts cumulés dans l'estuaire et le golfe du Saint-Laurent : Vers une planification systémique de l'exploitation des ressources. *Le Naturaliste canadien* 140, 45–55. doi:10.7202/1036503ar.

Belley, R., Archambault, P., Sundby, B., Gilbert, F., and Gagnon, J.-M. (2010). Effects of hypoxia on benthic macrofauna and bioturbation in the Estuary and Gulf of St. Lawrence, Canada. *Continental Shelf Research* 30, 1302–1313. doi:10.1016/j.csr.2010.04.010.

Benoît, H. P., Gagné, J. A., Savenkoff, C., Ouellet, P., and Bourassa, M.-N. (2012). State of the Ocean Report for the Gulf of St. Lawrence Integrated Management (GOSLIM). Department of Fisheries; Oceans Available at: http://publications.gc.ca/site/eng/9.575021/publication.html.

Blais, M., Devine, L., Lehoux, C., Galbraith, P. S., Michaud, S., Plourde, S., et al. (2018). Chemical and Biological Oceanographic Conditions in the Estuary and Gulf of St. Lawrence during 2016. 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].

Bourdages, H., Marquis, M.-C., Nozères, C., and Ouellette-Plante, J. (2018). Assessment of northern shrimp stocks in the Estuary and Gulf of St. Lawrence in 2017: Data from the research survey. Department of Fisheries; Oceans.

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.

Canada, S. (2017). Population and Dwelling Count Highlight Tables. 2016 Census. Statistics Canada Catalogue no. 98-402-X2016001. Ottawa. Released February 8, 2017. Available at: https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/hlt-fst/pd-pl/comprehensive.cfm.

Canada, G. of (2016). Tri-Agency Open Access Policy on Publications. Available from http://www.science.gc.ca/eic/site/063.nsf/eng/h_F6765465.html?OpenDocument.

Accessed 2019-09-19. Available at: http://www.science.gc.ca/eic/site/063.nsf/eng/h_F6765465.html?OpenDocument.

Chabot, D., and Claireaux, G. (2008). Environmental hypoxia as a metabolic constraint on fish: The case of Atlantic cod, Gadus morhua. *Marine Pollution Bulletin* 57, 287–294. doi:10.1016/j.marpolbul.2008.04.001.

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.

CoML (2010). Census of marine life, a decade of discovery. Available from http://www.coml.org. Accessed 2019-09-19.

Costanza, R., Andrade, F., Antunes, P., Belt, M. van den, Boersma, D., Boesch, D. F., et al. (1998). Principles for Sustainable Governance of the Oceans. *Science* 281, 198–199. doi:10.1126/science.281.5374.198.

Côté, I. M., Darling, E. S., and Brown, C. J. (2016). Interactions among ecosystem stressors and their importance in conservation. *Proc. R. Soc. B* 283, 20152592. doi:10.1098/rspb.2015.2592.

Crain, C. M., Kroeker, K., and Halpern, B. S. (2008). Interactive and cumulative effects of multiple human stressors in marine systems. *Ecology Letters* 11, 1304–1315. doi:10.1111/j.1461-0248.2008.01253.x.

Dafforn, K. A., Johnston, E. L., Ferguson, A., Humphrey, C. L., Monk, W., Nichols, S. J., et al. (2016). Big data opportunities and challenges for assessing multiple stressors across scales in aquatic ecosystems. *Marine and Freshwater Research* 67, 393–413. doi:10.1071/MF15108.

Darling, E. S., and Côté, I. M. (2008). Quantifying the evidence for ecological synergies. *Ecology Letters* 11, 1278–1286. doi:10.1111/j.1461-0248.2008.01243.x.

Data Citation Standards, C.-I. T. G. on, and PractOut of Mind: The Current Sices, C.-I. (2013). Out of Cite, Out of Mind: The Current State of Practice, Policy, and Technology for the Citation of Data. *Data Science Journal* 12, CIDCR1–CIDCR7. doi:10.2481/dsj.OSOM13-043.

Dee, L. E., Allesina, S., Bonn, A., Eklöf, A., Gaines, S. D., Hines, J., et al. (2017). Operationalizing Network Theory for Ecosystem Service Assessments. *Trends in Ecology & Evolution* 32, 118–130. doi:10.1016/j.tree.2016.10.011.

De Leo, F. C., Gauthier, M., Nephin, J., Mihály, S., and Juniper, S. K. (2017). Bottom trawling and oxygen minimum zone influences on continental slope benthic community structure off Vancouver Island (NE Pacific). *Deep Sea Research Part II: Topical Studies in Oceanography* 137, 404–419. doi:10.1016/j.dsr2.2016.11.014.

Dempsey, D. P., Gentleman, W. C., Pepin, P., and Koen-Alonso, M. (2018). Explanatory Power of Human and Environmental Pressures on the Fish Community of the Grand Bank before and after the Biomass Collapse. *Frontiers in Marine Science* 5. doi:10.3389/fmars.2018.00037.

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.

Dufour, R., and Ouellet, P. (2007). Estuary and Gulf of St. Lawrence marine ecosystem overview and assessment report. Department of Fisheries; Oceans Available at: http://publications.gc.ca/site/eng/9.574302/publication.html.

Dupont-Prinet, A., Vagner, M., Chabot, D., and Audet, C. (2013). Impact of hypoxia on the metabolism of Greenland halibut (Reinhardtius hippoglossoides). *Canadian Journal of Fisheries and Aquatic Sciences* 70, 461–469. doi:10.1139/cjfas-2012-0327.

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.

Eby, L. A., Crowder, L. B., McClellan, C. M., Peterson, C. H., and Powers, M. J. (2005). Habitat degradation from intermittent hypoxia: Impacts on demersal fishes. *Marine Ecology Progress Series* 291, 249–262. doi:10.3354/meps291249.

ECCC (2018). Environment and Climate Change Canada's (ECCC) Atlantic Shoreline Classification Available from https://open.canada.ca/data/en/dataset/30449352-2556-42df-9ffe-47ea8e696f91 Accessed 2019-09-19. Available at: https://open.canada.ca/data/en/dataset/30449352-2556-42df-9ffe-47ea8e696f91.

El-Sabh, M. I., and Silverberg, N. (1990). *Oceanography of a Large-Scale Estuarine System.*, eds. M. I. El-Sabh and N. Silverberg Springer New York doi:10.1007/978-1-4615-7534-4.

Eppler, M. J., and Mengis, J. (2004). The Concept of Information Overload: A Review of Literature from Organization Science, Accounting, Marketing, MIS, and Related Disciplines. *The Information Society* 20, 325–344. doi:10.1080/01972240490507974.

Feist, B. E., and Levin, P. S. (2016). Novel Indicators of Anthropogenic Influence on Marine and Coastal Ecosystems. *Frontiers in Marine Science* 3. doi:10.3389/fmars.2016.00113.

FORCE11 (2014). Data Citation Synthesis Group: Joint Declaration of Data Citation Principles. Martone M. (Ed.) San Diego CA. Available at: https://doi.org/10.25490/a97f-egyk.

Frank, K. T., Petrie, B., Choi, J. S., and Leggett, W. C. (2005). Trophic Cascades in a Formerly Cod-Dominated Ecosystem. *Science* 308, 1621–1623. doi:10.1126/science.1113075.

Franzoni, C., and Sauermann, H. (2014). Crowd science: The organization of scientific research in open collaborative projects. *Research Policy* 43, 1–20. doi:10.1016/j.respol.2013.07.005.

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].

GBIF (2018). GBIF: The Global Biodiversity Information Facility (year) What is GBIF? Available from https://www.gbif.org/what-is-gbif Accessed 2018-09-19.

Gilbert, D., Chabot, D., Archambault, P., Rondeau, B., and Hébert, S. (2007). Appauvrissement en oxygène dans les eaux profondes du Saint-Laurent marin: Causes possibles et impacts écologiques. *Naturaliste Canadien* 131, 67–75.

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., Longo, C., Hardy, D., McLeod, K. L., Samhouri, J. F., Katona, S. K., et al. (2012). An index to assess the health and benefits of the global ocean. *Nature* 488, 615–620. doi:10.1038/nature11397.

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.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., and Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25, 1965–1978. doi:10.1002/joc.1276.

Jones, F. C. (2016). Cumulative effects assessment: Theoretical underpinnings and big problems. *Environmental Reviews* 24, 187–204. doi:10.1139/er-2015-0073.

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

Keith, D. A., Martin, T. G., McDonald-Madden, E., and Walters, C. (2011). Uncertainty and adaptive management for biodiversity conservation. *Biological Conservation* 144, 1175–1178. doi:10.1016/j.biocon.2010.11.022.

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

Lubchenco, J., and Grorud-Colvert, K. (2015). Making waves: The science and politics of ocean protection. *Science* 350, 382–383. doi:10.1126/science.aad5443.

Mach, M. E., Wedding, L. M., Reiter, S. M., Micheli, F., Fujita, R. M., and Martone, R. G. (2017). Assessment and management of cumulative impacts in California's network of marine protected areas. *Ocean & Coastal Management* 137, 1–11. doi:10.1016/j.ocecoaman.2016.11.028.

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

Margules, C. R., and Pressey, R. L. (2000). Systematic conservation planning. *Nature.* doi:10.1038/35012251.

Micheli, F., Heiman, K. W., Kappel, C. V., Martone, R. G., Sethi, S. A., Osio, G. C., et al. (2016). Combined impacts of natural and human disturbances on rocky shore communities. *Ocean & Coastal Management* 126, 42–50. doi:10.1016/j.ocecoaman.2016.03.014.

Moritz, C., Gravel, D., Savard, L., McKindsey, C. W., Brêthes, J.-C., and Archambault, P. (2015). No more detectable fishing effect on Northern Gulf of St Lawrence benthic invertebrates. *ICES Journal of Marine Science* 72, 2457–2466. doi:10.1093/icesjms/fsv124.

OBIS (2018). Ocean Biogeographic Information System. Intergovernmental Oceanographic Commission of UNESCO. Www.iobis.org. Accessed 2018-09-19.

O'Brien, A. L., Dafforn, K. A., Chariton, A. A., Johnston, E. L., and Mayer-Pinto, M. (2019). After decades of stressor research in urban estuarine ecosystems the focus is still on single stressors: A systematic literature review and meta-analysis. *Science of The Total Environment.* doi:10.1016/j.scitotenv.2019.02.131.

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.

Pillet, M., Dupont-Prinet, A., Chabot, D., Tremblay, R., and Audet, C. (2016). Effects of exposure to hypoxia on metabolic pathways in northern shrimp (Pandalus borealis) and Greenland halibut (Reinhardtius hippoglossoides). *Journal of Experimental Marine Biology and Ecology* 483, 88–96. doi:10.1016/j.jembe.2016.07.002.

Poelen, J. H., Simons, J. D., and Mungall, C. J. (2014). Global biotic interactions: An open infrastructure to share and analyze species-interaction datasets. *Ecological Informatics* 24, 148–159. doi:10.1016/j.ecoinf.2014.08.005.

Québec, G. of (2015). Stratégie maritime, The maritime strategy by the year 2030. 2015-2020 action Plan.

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/.

Reichman, O. J., Jones, M. B., and Schildhauer, M. P. (2011). Challenges and Opportunities of Open Data in Ecology. *Science* 331, 703–705. doi:10.1126/science.1197962.

RQM (2018). Réseau Québec Maritime (RQM). Available from http://rqm.quebec/en/home/. Accessed 2019-09-19.

Savenkoff, C., Vézina, A. F., Roy, S., Klein, B., Lovejoy, C., Therriault, J. C., et al. (2000). Export of biogenic carbon and structure and dynamics of the pelagic food web in the Gulf of St. Lawrence Part 1. Seasonal variations. *Deep Sea Research Part II: Topical Studies in Oceanography* 47, 585–607. doi:10.1016/S0967-0645(99)00119-8.

Sbrocco, E. J., and Barber, P. H. (2013). MARSPEC: Ocean climate layers for marine spatial ecology. *Ecology* 94, 979–979. doi:10.1890/12-1358.1.

Schloss, I. R., Archambault, P., Beauchesne, D., Cusson, M., Ferreyra, G., Levasseur, M., et al. (2017). "Cumulative potential impacts of the stress factors associated with human activities on the St. Lawrence marine ecosystem," in *Hydrocarbon in the Gulf of St. Lawrence - Social, economic and environmental issues*, eds. P. Archambault, I. R. Schloss, C. Grant, and S. Plante (Notre Golfe, Rimouski, Qc, Canada), 133–165.

Scholes, R. J., Walters, M., Turak, E., Saarenmaa, H., Heip, C. H., Tuama, É. Ó., et al. (2012). Building a global observing system for biodiversity. *Current Opinion in Environmental Sustainability* 4, 139–146. doi:10.1016/j.cosust.2011.12.005.

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

Stock, A., Haupt, A. J., Mach, M. E., and Micheli, F. (2018). Mapping ecological indicators of human impact with statistical and machine learning methods: Tests on the California coast. *Ecological Informatics* 48, 37–47. doi:10.1016/j.ecoinf.2018.07.007.

Tyberghein, L., Verbruggen, H., Pauly, K., Troupin, C., Mineur, F., and Clerck, O. D. (2012). Bio-ORACLE: A global environmental dataset for marine species distribution modelling. *Global Ecology and Biogeography* 21, 272–281. doi:10.1111/j.1466-8238.2011.00656.x.

White, L., and Johns, F. (1997). *Marine Environmental Assessment of the Estuary and Gulf of St. Lawrence*. Department of Fisheries; Oceans.

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* 3, 160018. doi:10.1038/sdata.2016.18.

WoRMS Editorial Board (2017). World Register of Marine Species. Available from http://www.marinespecies.org at VLIZ. Accessed 2018-09-19. Https://doi.org/10.14284/170. doi:10.14284/170.