

Information Management

Developing an advanced information system to support ballast water management

Mohammad Etemad¹, Amilcar Soares², Paul Mudroch³, Sarah A. Bailey^{4,*} and Stan Matwin^{5,6}

¹Department of Physical and Environmental Sciences, University of Toronto Scarborough, Toronto, ON, Canada

²Department of Computer Science, Memorial University of Newfoundland, St. John's, NL, Canada

³Transport Canada, Ottawa, ON, Canada

⁴Great Lakes Laboratory for Fisheries and Aquatic Sciences, Fisheries and Oceans Canada, Burlington, ON, Canada

⁵Institute for Big Data Analytics, Dalhousie University, Halifax, NS, Canada

⁶Institute of Computer Science, Polish Academy of Sciences, 00-901 Warsaw, Poland

*Corresponding author

E-mail: sarah.bailey@dfo-mpo.gc.ca

Citation: Etemad M, Soares A, Mudroch P, Bailey SA, Matwin S (2022) Developing an advanced information system to support ballast water management. *Management of Biological Invasions* 13(1): 68–80, <https://doi.org/10.3391/mbi.2022.13.1.04>

Received: 14 May 2021

Accepted: 6 August 2021

Published: 29 November 2021

Thematic editor: Joana Dias

Copyright: © Etemad et al.

This is an open access article distributed under terms of the Creative Commons Attribution License (Attribution 4.0 International - CC BY 4.0).

OPEN ACCESS

Abstract

Global ballast water management aims to reduce the transport and introduction of non-indigenous species through practices such as ballast water exchange and ballast water treatment. Comprehensive enforcement to ensure vessels are meeting ballast water management requirements are a key part of success, but such activities are limited by available resources. Targeted and/or stratified enforcement activities are one option to make best use of available resources. International vessels are required to submit ballast water reporting forms prior to arrival at many ports around the world, declaring quantities, geographic sources, management undertaken and expected discharge location. These data are essential for risk assessment and trends analysis, but the inflow of data can be overwhelming for daily operations, particularly for jurisdictions with many ports and/or high vessel traffic. Having near real-time access to ballast water data enhances opportunities for data validation and verification and facilitates customized reports such as mapping of exchange coordinates and ballast water discharge statistics. Customized software enables seamless application of best-available science through integration of decision-support tools. The Ballast Water Information System (BWIS) was developed to support daily ballast water enforcement activities and scientific research in Canada. The BWIS increases accessibility of ballast water report data and streamlines data processing to support decision-making using an on-line platform.

Key words: aquatic invasive species, data storage and standardization, data validation and verification, data visualization, decision support system, environmental management, nonindigenous species, risk assessment

Introduction

International shipping is recognized globally as the primary transport mechanism for the introduction of aquatic non-indigenous species (NIS) (Bailey et al. 2020). While the International Convention for the Control and Management of Ships' Ballast Water and Sediments recently entered into force (IMO 2004), countries such as Canada have had national requirements for more than a decade (e.g., Scriven et al. 2015). An important aspect of such regulations is the requirement to report on

completed ballast water management activities to port authorities to facilitate enforcement and compliance monitoring (e.g., Bailey et al. 2011).

A Ballast Water Reporting Form (BWRF) is required for every international vessel transit arriving to Canadian ports (7000+ forms per year). Initially, BWRFs were submitted by fax or email, ranging in quality from handwritten to typeset scanned images, MS Word, or Adobe printed (non-machine readable) files. BWRF templates were not standardized across all regions of Canada, and forms were periodically changed/updated without any version management. Coordinates of ballast water exchange might be mapped by inspectors to verify if depth and distance-from-shore requirements were met. However, in many cases, detailed verification was limited due to time constraints and absence/inaccessibility of secondary sources of information such as independent lists of vessels entering Canadian waters or Automated Identification Systems (AIS) data verifying vessel routes.

After initial assessment, BWRFs were no longer considered of business value for enforcement; however, recognizing the value of the data for scientific research and risk assessment, a basic relational database was set up to store the information in 2006. BWRFs forms were sent to a central office, printed out, and manually entered (initially into MS Access; an Oracle-based system was set up in 2007). With a single form containing entries for up to 50 ballast water tanks, it could take nearly an hour for an experienced data entry clerk to input each form.

Canada can receive up to 50 forms per day, and with only a few personnel conducting data entry (often as one component of daily duties), a backlog of BWRFs quickly developed. Data entry was often conducted weeks or months after the BWRF was submitted, making the correction of missing or incomplete data virtually impossible. Another challenge was finding and keeping employees willing to carry out such a large volume of repetitive and detail-oriented work.

In 2011, due to a shortage of resources, only BWRF for vessels heading into the Great Lakes were prioritized for data entry. BWRFs for Arctic, Atlantic, Quebec, and Pacific ports were stored in parent format on a central server, creating a backlog of approximately 60,000 forms between 2011 and 2018, from which data were not easily accessible for program reporting or research (such as has been conducted for the U.S. in Gerhard and Gunsch 2018). Further, the Oracle database was set up mainly as a data repository, with limited ability to retrieve and analyze stored data.

In 2017, a project was initiated to design and implement a modern ballast water information system (BWIS) through a partnership between the Government of Canada and the Dalhousie University Big Data Institute. Inspired by the National Ballast Information Clearinghouse (NBIC) which was initiated in 1999 to collect, analyze, and interpret data on the ballast water management practices of commercial ships operating in the waters of the United States (National Ballast Information Clearinghouse 2021),

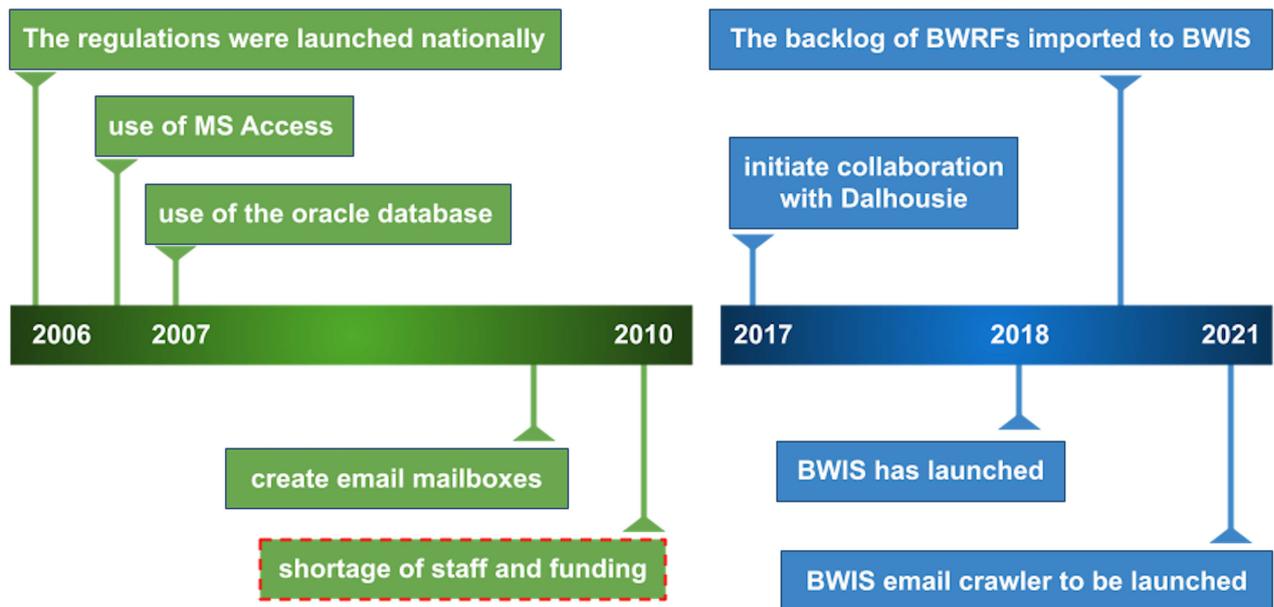


Figure 1. A timeline of important milestones in the development of the Canadian Ballast Water Information System (BWIS). BWRF = Ballast Water Reporting Forms.

the Canadian BWIS was set up to facilitate automated data collection and verification, integrate science-based decision-support tool(s), and enable data analysis, potentially with machine learning techniques.

The main objective was to design a centralized and uniform system where BWRF data could be properly organized, assessed, stored, and queried by the scientific and regulatory communities. A secondary objective was to address the backlog of BWRF, fusing all different formats of BWRFs used during the past 15 years, making data available for trends analysis and risk assessment. The vision was to decrease the daily workload related to ballast water enforcement while also improving data quality, supporting science-based policy development, ultimately decreasing the risk of introduction of aquatic NIS. Figure 1 shows a timeline of important events resulting in the development of the BWIS. This paper describes the BWIS in detail, including BWIS functionalities, users, and architecture (Section 2), next steps and future functionalities to be added (Section 3), and final remarks regarding the experience of assembling the BWIS (Section 4).

Ballast Water Information System (BWIS)

The BWIS was designed to accomplish the following four major functions:

1. Store and retrieve ballast water reports;
2. Clean and integrate ballast water data (i.e., to combine data from multiple sources into a unified dataset);
3. Generate reports related to ballast water data;
4. Implement available science-based tools to support decision-making procedures.

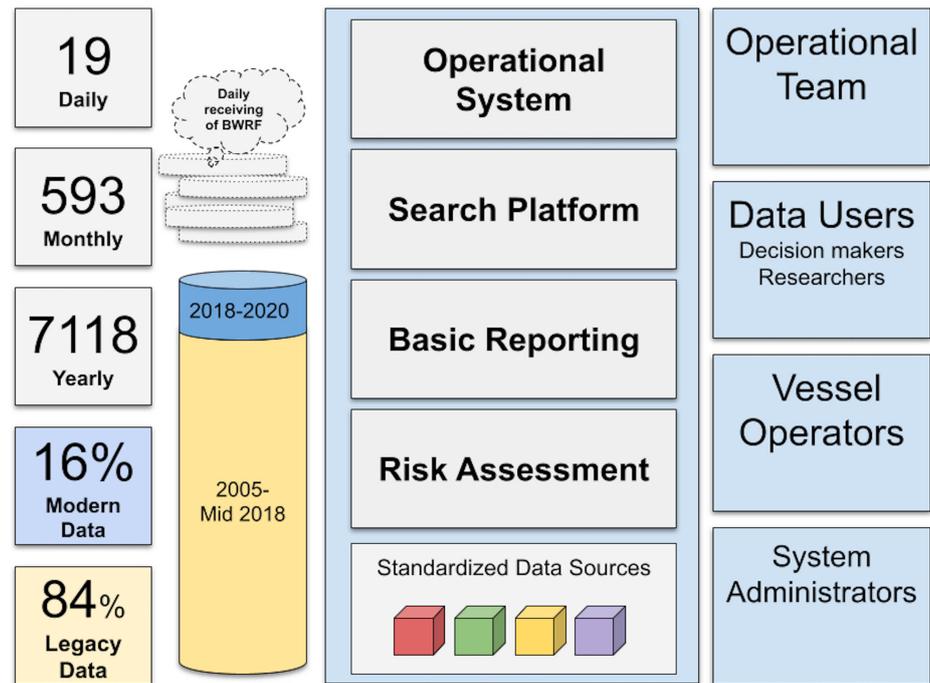


Figure 2. The Ballast Water Information System provides services to Operational Team, Data Users, Vessel Operators, and System Administrators such as: 1) facilitating data entry, validation and verification for Ballast Water Reporting Forms (BWRFs); 2) providing a search functionality for BWRFs; 3) providing basic reports on BWRf data; and 4) implementing research on risk assessment to build decision-support reports.

These functionalities were designed for four user groups: Operational Team, Data Users, Vessel Operators, and System Administrators (Figure 2). The Operational Team utilizes the BWIS to facilitate their daily activities (e.g., Ballast Water Inspectors). Regular operational users can upload BWRFs, run reports and use decision-support tools. Authorized operational users can also correct or delete data. Data Users are consumers of BWIS data (e.g., Researchers, Decision Makers) with read-only access. They can generate reports and use science-based tools to become informed about ballast water operations within Canada. Vessel Operators provide data into the BWIS by submitting BWRFs by email but do not have access to data within the BWIS. System Administrators have highest access to the BWIS; they can define parameters of the BWIS such as standardizing port names and can make modifications to the operational aspects of the system (i.e., Computer Programmers).

The Django (<https://www.djangoproject.com/>) platform and MongoDB (<https://www.mongodb.com/>) were selected for implementation of this project, as both applications are open source. Django is a high-level Python Web framework with Object Relational Mapping. This feature helps us to define an object layer that the front end interacts with. This object layer can be linked to most available data storage systems such as PSQL, SQL Server, and MongoDB. Django also has the capability of defining Application Programming Interfaces (API), allowing implementation of API code in the same repository as the front-end code. MongoDB was selected as the

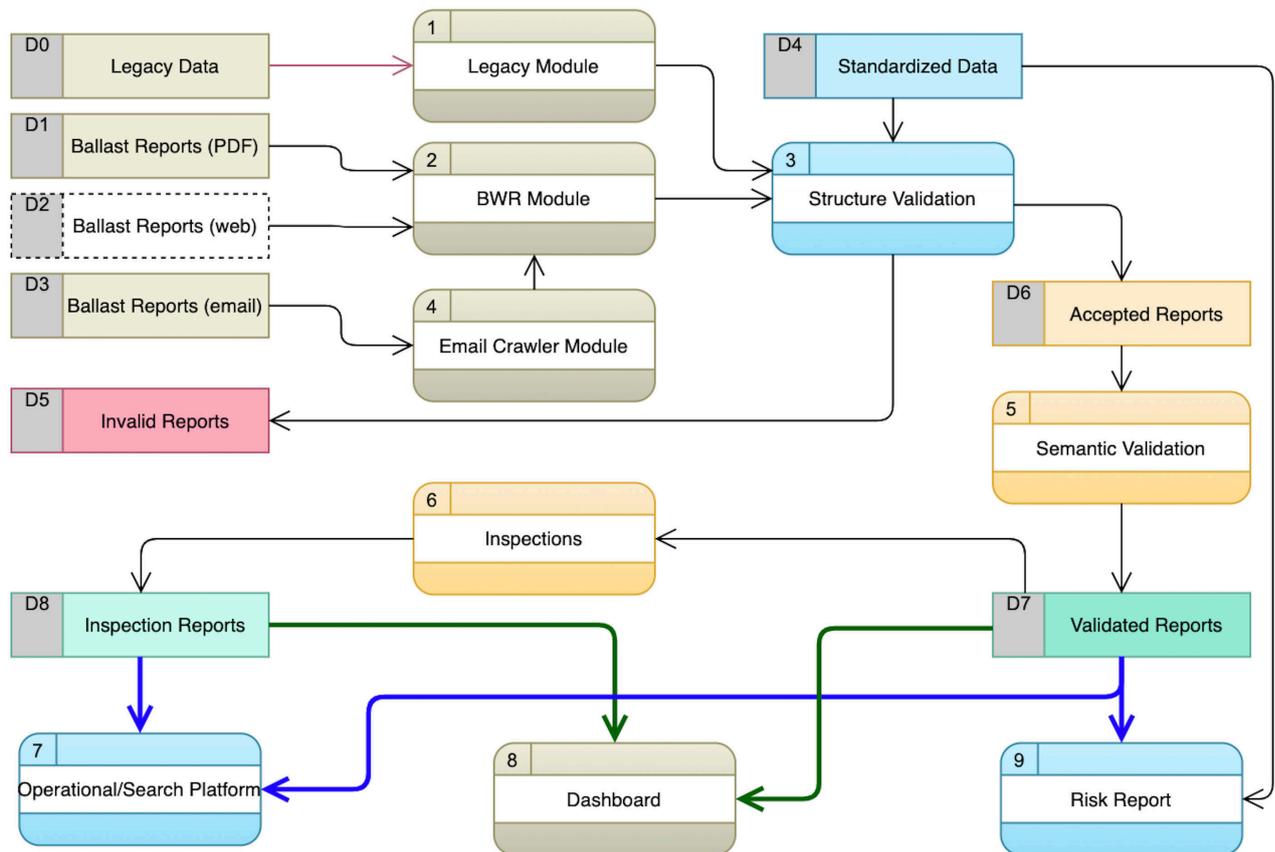


Figure 3. An overall view of all modules of the Ballast Water Information System.

main data storage software since it allows for flexible and dynamic data structure, ideal for multiple data sources with different data structures while also allowing for changes in the future.

Store and retrieve ballast water reports

The BWIS is designed to receive BWRFs by multiple methods, as shown in Figure 3:

- Import of legacy data by System Admin (D0)
- Upload BWRF by operational user (D1)
- Submit BWRF through web form (D2)
- Receive BWRF by email from vessel operator (D3)

The first step towards a uniform, centralized BWIS was the alignment of regional BWRFs into a single standard national reporting form in ADOBE PDF, using a fillable, machine-readable format, and creating a matching “PDF-reader”. The form was also developed in consultation with U.S. agencies to ensure compatibility in neighboring jurisdictions, particularly the North American Great Lakes. This allows the Operational User to directly import BWRFs (D1) through “drag and drop” of forms received by email into the BWIS. The time to enter data is thus reduced from up to one hour to less than one minute per BWRF. The Operational Team began

using the standardized national BWRf and automated data entry system in July 2018, with a complete transition by January 2019.

Since July 2018, the use of the BWIS for automated data entry has enabled the Operational Team to verify data (assuring data entered correctly, accurately, and is consistent with standard data dictionaries in the BWIS) on each BWRf within 24 h of receipt. The BWIS receives the PDF file and makes a copy of it on the server immediately. After that, the PDF file is processed to read the data fields submitted by vessels and generate a BWRf record in the database with the extracted information.

In the next step, the BWRf is validated with some pre-defined business logic rules to check that data quality is at the expected level (see next section). If there is any conflict with the business rules, a report of all errors is generated and displayed to the user. The operational user can then determine if the issues must be resolved at the time of validation or they can be skipped (are non-critical). The BWIS is a growing dataset, receiving, on average, 19 BWRfs daily (593 monthly, and 7118 per year).

The second gateway of the system is receiving BWRfs by email (D3) from Vessel Operators. An email crawler was set up on an operational email server to find PDF email attachment/s, and extract BWRf data. This function is under a testing phase and is expected to become the primary method of data entry later in 2021. This subsystem will be responsible for some data cleaning and validation as the email crawler can communicate with ballast water inspectors and provide guidance to fix any errors and warnings related to the received BWRf.

A web form (D2) is also possible for submitting BWRf directly to the BWIS. The online system can execute additional data standardization procedures to produce the highest quality data through interaction with the data submitter. However, this is currently not in use due to vessel bandwidth limitations. Finally, the import of legacy data (D0) is another gateway for the BWIS.

BWRf received prior to July 2018 are considered legacy data. Optical Character Recognition (OCR) was used to retrieve data from backlogged BWRfs, and both OCR and manual entry legacy data have been imported into the BWIS by the System Administrators (currently about 84% of total data).

Clean and integrate ballast water data

As BWRfs contained in the BWIS originated from differing sources (i.e., OCR legacy reports, Oracle database, direct import), cleaning and integration procedures are key for improving the quality and integrity of BWIS data. Both operational and legacy data can have data quality issues. First, BWRfs may contain missing values. Second, the data is not missing, but it is wrong – for example, erroneous entries or misspellings. Third, the data is neither missing nor wrong but is unusable. Ambiguous data such as abbreviations and different representations of compound data such as date fields are

examples of this category. A more detailed taxonomy of dirty data is proposed by Kim and colleagues, including possible resolutions (Kim et al. 2003).

One of the objectives of implementing the BWIS is to access high-quality data. Data quality is a gray subject and domain dependent; however, six major factors contribute to high-quality data: Validity, Accuracy, Completeness, Consistency, Uniformity, and Timeliness (Askham et al. 2013). Utilizing the BWIS, we can enhance the quality of data in some of these aspects. Data profiling (i.e., reviewing the BWRFs available in BWIS and collecting statistics or summary reports) with intervention by a domain expert, who has considerable experience in processing and reviewing BWRFs, is one of the key steps for clean data recommended by Kim and colleagues, along with automatic tools such as regular expressions, format checkers and type checkers (Kim et al. 2003). Therefore, the BWIS is interactive – enabling Operational Team subject matter experts to review data using tools to filter and update data.

The BWIS also contains dictionaries and lookup tables for categorical data and compound fields (e.g., vessel type, port name) in a data store called standardized data (step D4, Figure 3). The rapid upload of data using the BWIS also enables the Operational Team to identify issues while there is time to ask Vessel Operators for corrections. As a result, legacy data contain more “dirty” data (i.e., data that contains erroneous information) while direct import by Operational Users is more complete and more reliable.

For data fields determined to be “critical”, the BWIS was designed to generate errors to prevent a user from storing data unless the required quality level is satisfied. Similarly, warnings are generated for non-critical data. The BWIS applies business rules to assist the Operational Team in the identification of data errors. For example, the arrival date can have a defined range between the years 2005–2021; any date outside of this interval can be flagged for review by the operational user. Another business rule can be the chronological order of date fields related to ballast water operations (intake date, exchange date, discharge date). The rules can be executed against the direct import and legacy data to flag poor quality records for review and possible correction by operational users.

Generate reports related to ballast water data

One of the most important functionalities of the BWIS is to generate basic reports of ballast data – components seven and eight in Figure 3. Basic reports are flat queries that can be executed by selecting fields within the BWRF (Codd 1990). These reports are designed so that the operational team can access any BWRF based on specific criteria in a search platform. The basic report does not support join, and aggregation operations (Codd 1990); however, a few frequently used aggregation reports have been customized for the operational team to support their daily work. These reports are accessible from the main dashboard of the operating system, although they can be slow depending on the nature of aggregation.

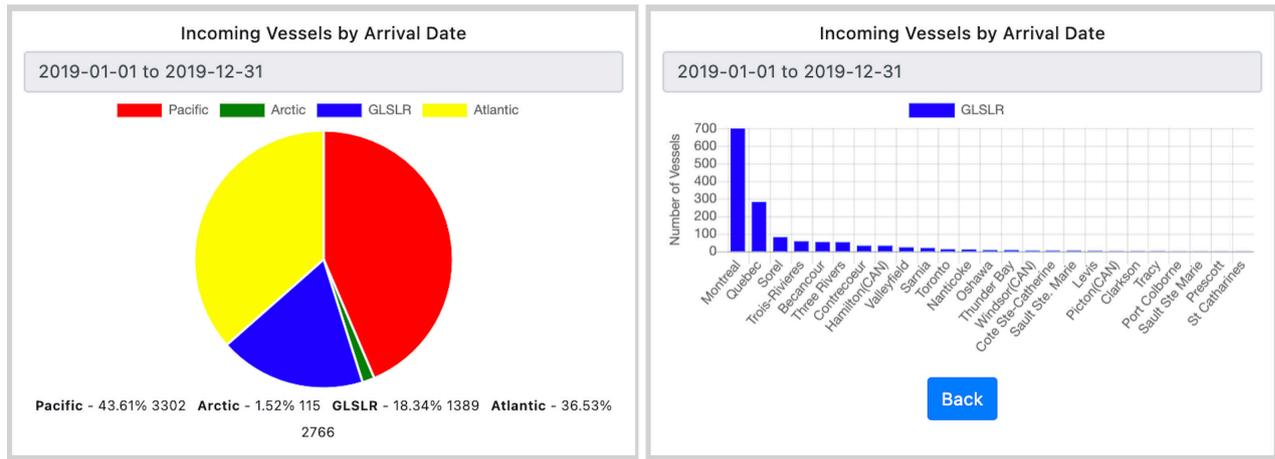


Figure 4. An example two-level report for Incoming Vessels by Arrival Date. Level one is displayed on the left and level two is displayed on the right.

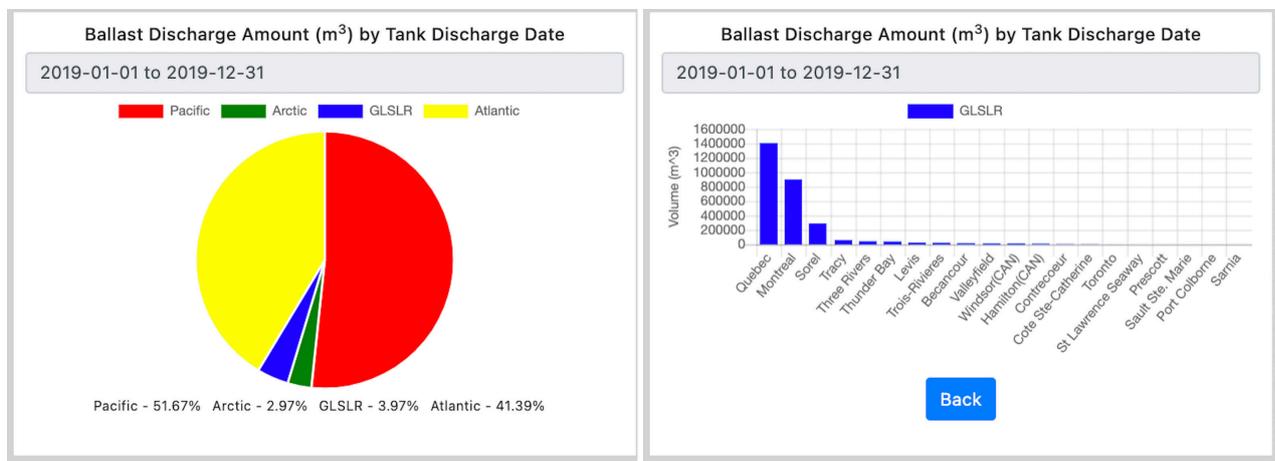


Figure 5. An example two-level report for Ballast Discharge Amount (m³) by Tank Discharge Date. Level one is displayed on the left and level two is displayed on the right.

Incoming Vessels by Arrival Date

The incoming vessels by arrival date report is a two-level aggregation query to provide a summary of the number of vessel arrivals during a defined period of time in a specific region. An example of this report is shown in Figure 4 for the year 2019. In the first level of the report, shown on the left side of Figure 4, a pie diagram is presented where each slice represents the number and percentage of vessel arrivals by region (Pacific, Arctic, Great Lakes-St. Lawrence River (GLSLR), and Atlantic). By selecting an individual slice, the user can filter data on the selected area to see the number of arrivals at specific ports within the selected region, as in the right side of Figure 4.

Ballast Discharge Amount by Tank Discharge Date

This report is another two-level aggregation query that provides a summary of ballast water discharge volume during a defined period of time in a specific region. The summary reported in level one is the sum of the volume of ballast water discharged in each region, shown on the left side of Figure 5. By selecting an individual slice, the sum of the volume of ballast

Table 1. An example of a Risk Assessment Report generated for each ballast tank on the ship “Penn 92”, with estimated time of arrival (ETA), showing ballast source and discharge ports, discharge amount, environmental distance and risk level.

Vessel Name	Tank	Ballast Source	Ballast Discharge	Disch. Amount (m ³)	Environmental Distance	Risk Level	Differential Risk
Penn 92	1P	Boston(USA)	Saint John(CAN)	580.00	2.26	Moderate Risk	12
Penn 92	1S	Boston(USA)	Saint John(CAN)	580.00	2.26	Moderate Risk	12
Penn 92	2P	Providence(USA)	Saint John(CAN)	606.00	3.20	Low Risk	9
Penn 92	2S	Providence(USA)	Saint John(CAN)	606.00	3.20	Low Risk	9
Penn 92	3P	Boston(USA)	Saint John(CAN)	606.00	2.26	Moderate Risk	12
Penn 92	3S	Boston(USA)	Saint John(CAN)	606.00	2.26	Moderate Risk	12
Penn 92	4P	Providence(USA)	Saint John(CAN)	606.00	3.20	Low Risk	9
Penn 92	4S	Providence(USA)	Saint John(CAN)	606.00	3.20	Low Risk	9
Penn 92	5P	Boston(USA)	Saint John(CAN)	626.00	2.26	Moderate Risk	12
Penn 92	5S	Boston(USA)	Saint John(CAN)	626.00	2.26	Moderate Risk	12

water discharged in each port within the selected region is displayed, as shown on the right side of Figure 5. This report and the previous report can be used together for better interpretation. For example, the number of arrivals in the Pacific region in 2019 was 43.61% while the volume of ballast water discharged was 51.67% of ballast discharged in 2019 in Canada.

Implement available science-based tools to support decision-making

Insight on the level of risk each vessel arrival poses to a port can be used to inform prioritization of compliance activities (David et al. 2012). The BWIS includes a risk assessment module based on a previously developed decision-support tool (Bradie and Bailey 2021). Having BWRFs readily available in the BWIS presents opportunity for automatizing the risk calculation. The relative risk of different arrivals is calculated based on environmental distance between source and recipient ports, as well as an estimate of propagule pressure (the number of organisms expected to be released in a given discharge) based on ballast volume and ballast age; all of which can be deduced from BWRF data for individual ballast tanks (Bradie and Bailey 2021). Environmental distance is calculated based on a comparison of salinity and mean, minimum and maximum temperature at source and recipient ports (following methodology originally developed by Keller et al. 2010), while propagule pressure is predicted based on a model examining the joint effects of ballast water age and ballast water volume on organism abundances from previous biological studies (see Bradie and Bailey 2021, for full details on the decision-support tool). The information is summarized for each ballast tank discharge as a risk ranking (being the lowest of either environmental or propagule pressure risks) and a numeric risk score (called differential risk) allowing more refined ranking capabilities (Table 1). For each ballast water tank, the BWIS also produces a map showing the source port and the destination port with the environmental distance calculation (Figure 6). Points within 50 km of the destination port are also shown to better visualize potential risk to the local area; polygons are used to categorize the risk levels.

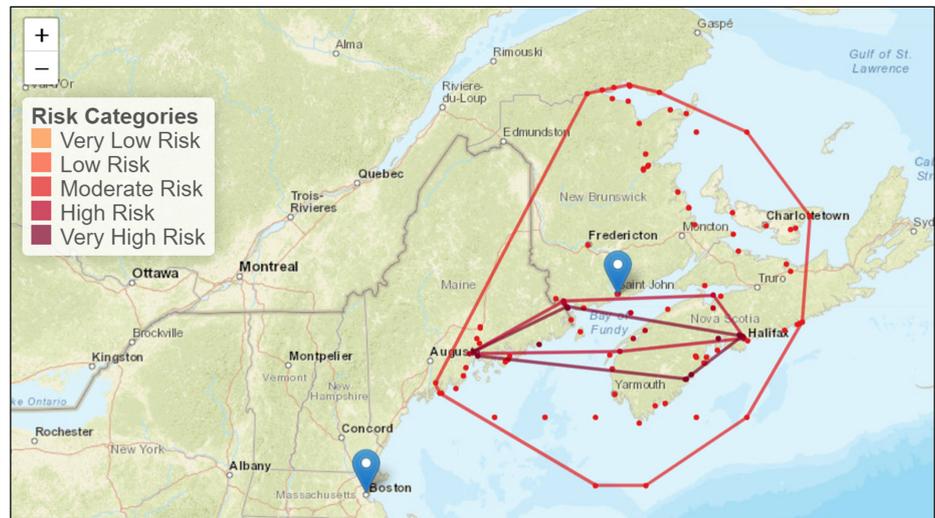


Figure 6. An example Risk Map showing environmental risk between source port Boston, USA and destination port Saint John, Canada.

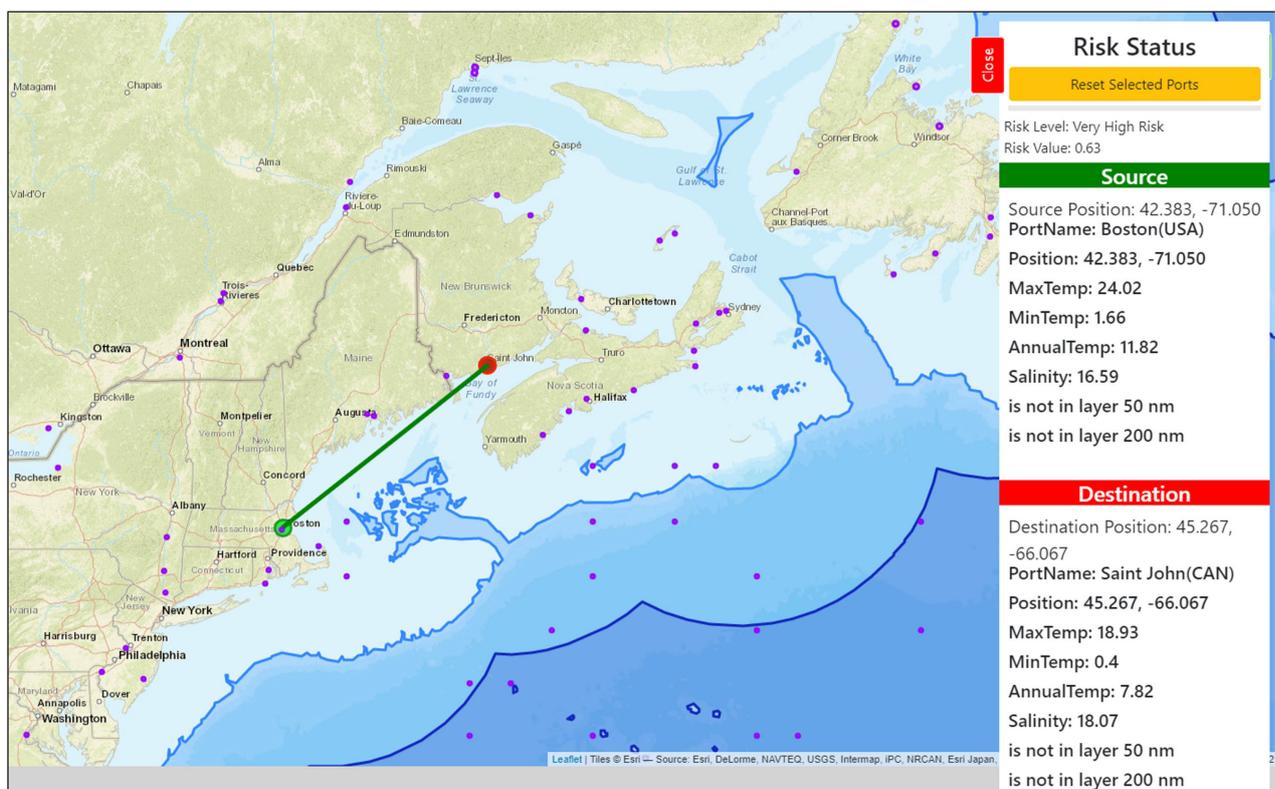


Figure 7. The application of the on-demand visualization tool for assessing the risk of ballast water sourced from Providence, USA and discharged in Louisbourg, Canada.

The operational team can also utilize the risk assessment tool as an on-demand risk calculator. An underlying standardized environmental factors dataset has a mesh of global geo-locations with environmental factors. When the user selects a geo-location on the map, the system finds the nearest point/port with available environmental data (salinity and temperature). A visualization tool draws a green line from the nearest port to the source and destination ports (Figure 7). In the right panel, each port's environmental factors are displayed, and the level of risk is calculated.

This visualization tool also contains layers to facilitate the assessment of ballast water exchange positional requirements. The first layer shows where the ocean's depth is more than 2000 meters and 200 nautical miles distance from shore (as required by Canadian regulations, dark blue). The second layer shows where the ocean's depth is 200 meters and is 50 nautical miles distance from shore (matching the depth and distance specified in the Convention, a secondary option in Canada if the first cannot be met, medium blue).

Future work

A key next step for the BWIS is to build on the centralized online system to derive annual statistics about ballast water data management in relation to Canadian regulatory requirements (trends analysis). It will be particularly interesting to cross-check the accuracy of self-reported data within BWRf to independent data sources (e.g. compare ballast water exchange positional coordinates with satellite-based information systems). There is also increasing interest to understand the interaction of shipping activities with current and proposed marine protected areas (Kenchington 2010), and a new visual analytic tool is being considered to assist with such assessments. Further, the risk assessment tool can be refined/developed (using Machine Learning Methods) to incorporate advances in risk understanding, particularly related to the use of new ballast water treatment methods (Bradie et al. 2021; Casas-Monroy et al. 2015).

Work will also continue on the improvement of data quality, potentially introducing additional business rules for use during direct or web-based data import. Double-checking the submitted information against a secondary data source is an important step after the implementation of the BWIS because the BWRf data is self-reported. As future work, we plan to cross check data in the BWIS against secondary sources such as Automatic Identification System (AIS) positional information (e.g. Kim et al. 2014).

The BWIS is a prototype that is running on a managed platform with limited scalability and basic security. As maintenance of a managed server can be costly and requires many specialties, transfer to a cloud service with enhanced security and scalability can be considered. The BWIS could also be replicated for use in other jurisdictions using the code repository for the software.

Conclusions

The Ballast Water Information System (BWIS) has enabled regulators and scientists to efficiently capture and understand data related to ballast water management for the last 15 years while decreasing daily operational workloads and improving data quality for scientific use. Having real-time, online access to BWRfs increases data availability, while customized data

reporting features streamline operational tasks and the decision-making process. The BWIS expands operational users' ability to receive forms, enabling reporting by all vessels (domestic and international) for all transits in the future. The BWIS offers a more collaborative environment that utilizes online procedures and routines, enabling integration of best-available science into the operational decision-making procedure, such as the risk assessment tool embedded in the BWIS.

In summary, the BWIS is an advanced platform to store, clean, and integrate ballast water reports from multiple data sources. This platform increases the accuracy and availability of ballast water reports, supports decision-making procedures, promotes research findings in the operational environment, ultimately improving ballast water management practices.

Acknowledgements

Thanks to Salil Kapur, Jordan Rose, and Nader Zare for their contributions to the development of the BWIS, and to the many personnel conducting data entry since 2005. Also, thanks to Jennifer Strang, GIS Analyst at the Dalhousie University Libraries, for help preparing the geographic data for analysis. Feedback from Transport Canada colleagues and insight provided by Greg Ruiz and A. Whitman Miller from the U.S. National Ballast Information Clearinghouse are gratefully acknowledged. Suggestions from two reviewers to improve the article are also appreciated.

Funding declaration

This project was funded by Transport Canada, Fisheries and Oceans Canada, NSERC, Dalhousie University and Memorial University of Newfoundland. The funding institutions had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript (outside the contributions of the authors).

Authors' contributions

SAB and PM conceived the idea for the Ballast Water Information System (BWIS); ME, AS and SM designed the BWIS; ME and AS wrote some parts of BWIS code; PM conducted beta testing; ME and SB led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

Data availability statement

The BWIS is available at <https://ballast.research.cs.dal.ca/> and access can be granted upon reasonable request to the corresponding author.

References

- Askham N, Cook D, Doyle M, Fereday H, Gibson M, Landbeck U, Lee R, Maynard C, Palmer G, Schwarzenbach J (2013) The six primary dimensions for data quality assessment: defining data quality dimensions. DAMA UK Working Group on "Data Quality Dimensions", Bristol. <https://sil0.tips/download/the-six-primary-dimensions-for-data-quality-assessment> (accessed 11 October 2021)
- Bailey SA, Deneau MG, Jean L, Wiley CJ, Leung B, MacIsaac HJ (2011) Evaluating efficacy of an environmental policy to prevent biological invasions. *Environmental Science and Technology* 45: 2554–2561, <https://doi.org/10.1021/es102655j>
- Bailey SA, Brown L, Campbell ML, Canning-Clode J, Carlton JT, Castro N, Chainho P, Chan FT, Creed JC, Curd A, Darling J, Fofonoff P, Galil BS, Hewitt CL, Inglis GJ, Keith I, Mandrak NE, Marchini A, McKenzie CH, Occhipinti-Ambrogi A, Ojaveer H, Pires-Teixeira LM, Robinson TB, Ruiz GM, Seaward K, Schwindt E, Son MO, Therriault TW, Zhan A (2020) Trends in the detection of aquatic non-indigenous species across global marine, estuarine and freshwater ecosystems: A 50-year perspective. *Diversity and Distributions* 26: 1780–1797, <https://doi.org/10.1111/ddi.13167>

- Bradie JN, Bailey SA (2021) A decision support tool to prioritize ballast water compliance monitoring by ranking risk of non-indigenous species establishment. *Journal of Applied Ecology* 58: 587–595, <https://doi.org/10.1111/1365-2664.13822>
- Bradie JN, Drake DAR, Ogilvie D, Casas-Monroy O, Bailey SA (2021) Ballast Water Exchange plus Treatment Lowers Species Invasion Rate in Freshwater Ecosystems. *Environmental Science and Technology* 55: 82–89, <https://doi.org/10.1021/acs.est.0c05238>
- Casas-Monroy O, Linley RD, Adams JK, Chan FT, Drake DAR, Bailey SA (2015) Relative invasion risk for plankton across marine and freshwater systems: Examining efficacy of proposed international ballast water discharge standards. *PLoS ONE* 10: 1–16, <https://doi.org/10.1371/journal.pone.0118267>
- Codd EF (1990) *The Relational Model for Database Management: Version 2*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, United States, 538 pp
- David M, Perkovic M, Suban V, Gollasch S (2012) A generic ballast water discharge assessment model as a decision supporting tool in ballast water management. *Decision Support Systems* 53: 175–185, <https://doi.org/10.1016/j.dss.2012.01.002>
- Gerhard WA, Gunsch CK (2018) Analyzing trends in ballasting behavior of vessels arriving to the United States from 2004 to 2017. *Marine Pollution Bulletin* 135: 525–533, <https://doi.org/10.1016/j.marpolbul.2018.07.001>
- IMO (2004) International Maritime Organization. International convention for the control and management of ships' ballast water and sediments. In: BWM/ CONF/36, London, UK
- Keller RP, Drake JM, Drew MB, Lodge DM (2010) Linking environmental conditions and ship movements to estimate invasive species transport across the global shipping network. *Diversity and Distributions* 17: 93–102, <https://doi.org/10.1111/j.1472-4642.2010.00696.x>
- Kenchington TJ. (2010) Environmental Monitoring of the Gully Marine Protected Area: A Recommendation. DFO Canadian Science Advisory Secretariat Research Document 2010/075, vi + 59 pp, <http://waves-vagues.dfo-mpo.gc.ca/Library/342194.pdf>
- Kim W, Choi BJ, Hong EK, Kim SK, Lee D (2003) A Taxonomy of Dirty Data. *Data Mining and Knowledge Discovery* 7: 81–99, <https://doi.org/10.1023/A:1021564703268>
- Kim E-C, Oh J-H, Lee S-G (2014) Port risk assessment for ballast water managements in the Republic of Korea. *Indian Journal of Geo-Marine Sciences* 43(11): 2013–2023
- National Ballast Information Clearinghouse (2021) NBIC Online Database of the Smithsonian Environmental Research Center and United States Coast Guard. <https://nbic.si.edu> (accessed 11 October 2021)
- Scriven DR, Dibacco C, Locke A, Therriault TW (2015) Ballast water management in Canada: A historical perspective and implications for the future. *Marine Policy* 59: 121–133, <https://doi.org/10.1016/j.marpol.2015.05.014>