CONTRIBUTION 3.13

The Evolution of Data Collection, Analysis and Reporting of Fisheries Data for the Commonwealth of Dominica

Derrick Theophille and Kurt Hilton

Ministry of Agriculture, Fisheries, Blue and Green Economy,

Dominica Fisheries Division, Commonwealth of Dominica
derrick.theophille@gmail.com/theophilled@dominica.gov.dm *and*
hiltonk@dominica.gov.dm/khiltonflo@gmail.com

Abstract

Fisheries data collection is essential for the long-term sustainability of the sector. Limited fisheries data presents uncertainty at varying levels leading to decreased productivity and economic loss. Utilising new and emerging technologies within the fisheries sector allows for more efficient decision-making since there is an ability to record and report more quickly, comprehensively and efficiently.

Spanning a 29-year period, data collection and management in Dominica have seen significant changes. This has included paper-based data collection at landing sites, a data input clerk inputting data, storage of data in Microsoft Access, and data analysis using Microsoft Excel. More recently, however, there has been a shift to utilising electronic data collection forms in Kobotoolbox and analysis and reporting of data using the R statistical software.

In this paper, we document the evolution of fisheries data collection and storage in the Commonwealth of Dominica, highlighting the challenges and future trajectory using low-cost and open-source methods and technologies for data collection, data manipulation, data analysis, and reporting. The use of new and inclusive methods of data collection presents a greater scope for the collection of information, leading to a reduced cost of data collection as well as time spent on the analysis and reporting of standardised or automated results.

KEYWORDS: Dominica, fisheries, data collection, kobotoolbox, fisheries management, R statistical software, R-programming

CITATION: Theophille, D. and K. Hilton, 2026. The Evolution of Data Collection, Analysis and Reporting of Fisheries Data for the Commonwealth of Dominica. Contribution 3.13, p238-248 *IN* CRFM, 2026. Volume I: Technical Papers. Presented at the CRFM 20th Anniversary Scientific Conference, 28-31 August 2023. CRFM Special Publication No. 37, CRFM Secretariat https://www.crfm.int/index.php?option=com_k2&view=item&id=918

1. INTRODUCTION

The extraction of marine resources is being closely monitored worldwide owing to the growing concerns about the decline of fish stocks and the negative impacts of fishing on their habitats, as well as the resulting effects on broader ecological processes. Despite this, the demand for fish-based protein has not diminished and there are further concerns about the quality of reporting coming from the sector (Regueiro et al., 2022). Consequently, the collection of fisheries data is essential for the sustainability of the sector, allowing for decisions to be made that reduce conflicts, monitor fish stocks and fishing pressure, and guide marine spatial planning. Furthermore, data collection aids in developing policy and strengthens management authorities. According to Gebremedhin et al. (2021), two key pillars should guide decisions, data analysis and data availability. To combat deficiencies in these areas, there needs to be an improvement in the data collection process to better understand constraints within the sector and identify areas for development.

In many countries, fisheries management authorities lack focus on data and information collection, leading to subpar fisheries management practices. While in some departments there is an effort to gather the required information, the processing, analysis and storage may be deficient. The purpose of data collection should not be to simply collect the data if it is not properly assessed. For effective fisheries management it is essential to collect the necessary data and use it to make informed decisions (*A Fishery Manager's Guidebook - Second Edition*, n.d.). However, as it relates to the fisheries sector there are many countries which are still considered data poor and lack information that can be considered reliable (Salas et al., 2007).

In the scientific community, it is recognized that recent advancements in data collection have led to systems that are accessible to fishers and other stakeholders in the sector. These systems are designed to address technological, social, and economic barriers to data collection and play a critical role in verifying the sustainability of the sector (Bradley et al., 2019). Modernizing fisheries data systems is crucial for the adaptive management of fisheries in the face of climate change and complex dynamics. While progress may be slow in some areas due to challenges such as funding, capacity constraints, or political issues, it is important to assess paper-based methods and address the global problem of data collection. The emergence of mobile devices, cloud-based systems, and mapping solutions offers a smooth transition away from paper-based systems for data collection(Merrifield et al., 2019).

In the Commonwealth of Dominica, the need for data to inform, educate, and make scientific based decisions is prevalent. The Fisheries Division, being managers of the marine resources, has realized the need for data collection and the importance of furthering the need for incorporating new and emerging technology to aid in the collection of data. Here, we provide an overview of the evolution of fisheries data collection and storage in the Commonwealth of Dominica, highlighting the challenges and future trajectory using low-cost and open-source methods and technologies for data collection, data manipulation, analysis, and reporting. We show how the use of technology can assist the managing authority in making decisions at a faster rate due to decreased delays in decision making due to the faster reporting and analysis of data.

2. METHODS

Study Area

Dominica is an island located in the eastern Caribbean Sea, nestled between the French islands of Guadeloupe to the north and Martinique to the south, in an archipelago known as the Lesser Antilles. The country spans an area of 750 km2 and is home to approximately 72,000 people, who reside primarily along the shore or in low-lying areas (Andereck, 2007). Dominica's Exclusive Economic Zone (EEZ) covers an area of 28,653 km2, while its Territorial Sea extends twelve nautical miles from land (IMIS, 2023). The continental shelf surrounding Dominica is narrow, with the 50m depth contour reaching a maximum distance of 2.8 km from the coast.

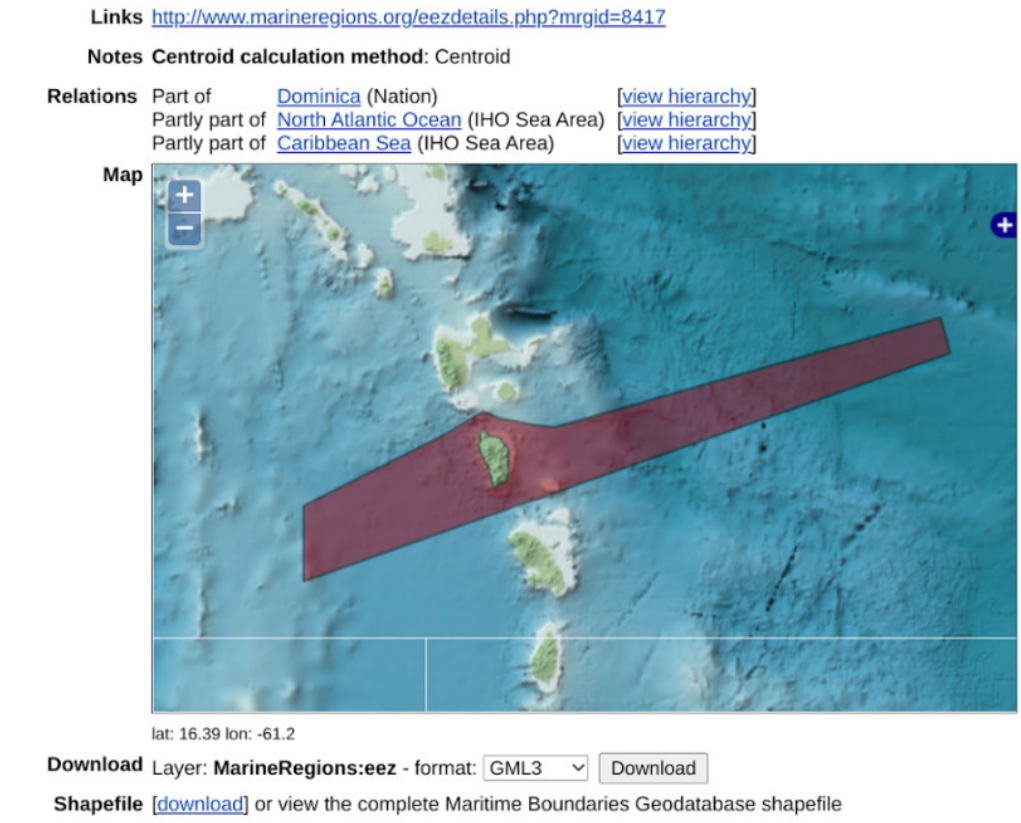


Figure 1. Map of the Commonwealth of Dominica showing range of Exclusive Economic Zone (EEZ). Source: IMIS, 2023

Fisheries in Dominica

The Fisheries sector in Dominica plays a crucial role in food security, despite its relatively low economic output when being compared to other economic activities in the country (e.g., tourism). Additionally, it offers a source of employment, support, and income to vulnerable families.

The Fisheries Sector in Dominica is small-scale and artisanal, consisting of individual fishers or fisher groups using small, open fishing vessels for short trips lasting only a few hours each day. There is no industrial fleet segment, and fish exports are minimal, with all landings being consumed locally (Theophile, 2016). According to the 2011 Fisheries Industry Census, there were 734 fishermen, most of whom were part-time and used small, open vessels such as

wooden keel boats, traditional canoes, and 434 fibreglass-reinforced plastic (FRP) pirogues under 25ft in length (Dominica Fisheries Division, 2011). Demersal fish species are targeted across the island, with pelagic species such as Yellowfin tuna and Common dolphin fish contributing approximately 65% to overall yearly landings. These migratory species are targeted from January to June.

The Fisheries Division (FD) of Dominica is responsible for developing and managing living marine and fisheries resources to ensure their optimal utilization. For almost three decades, the FD has maintained a fishery-dependent field data collection program and database, recording the minimum data required for fishery resource management at many fish landing sites around the island. The data collected include catch and fishing effort, as well as associated variables about the sampled fishing trip (Theophile, 2016).

However, the optimal collection of field data is often challenged by budgetary and human resource constraints. Despite these challenges, quality control and consistency obligations are maintained, encouraging revisions and updates that reflect changes in the fishery. Small scale artisanal fisheries, which are critical to many communities, require data collection programs that are tailored to the specific national scenario. However, monitoring these fisheries can be difficult because of their inherent complexity. These factors may contribute to the general paucity in fisheries data, which affects management in developing countries such as Dominica (Theophile, 2016).

Catch and effort data collection in Dominica

The Fisheries Division is mandated under the Fisheries Act No. 11 of 1987 of the Laws of Dominica to manage the affairs of the marine environment. Its functions include research and development, compilation, collection and reporting of fisheries information (fish catch and effort, registration, research), exploratory fishing for new and underexploited species, introduction of new and enhanced fishing gear, marine environmental impact assessments e.g., reef monitoring, development of fishing communities and maintaining fisheries exploitation at sustainable levels (Hilton, 2023).

In Dominica, the unit of effort for fishing is defined as a fishing trip, sometimes referred to as a day-trip or simply, a trip. Data collection is conducted at the time of landing, where it is possible to accurately identify, weigh, and count the species landed, including catch per gear specification. The Fisheries Division (FD) routinely records field catch and effort data from 14 out of the 31 (Fig. 2) designated fish landing sites around the island (Theophile, 2016). The field data collection process in Dominica involves assigning trained data collectors to one or two landing sites, which they visit at least three days within the work week, to document fishing activities for about one third of all boats landed at each site per day (Theophile, 2016). The data collectors interview the crew about the trip duration, locations fished, costs of operation, gear, and bait used. The vessel details are also recorded, including vessel registration number and whether it is motorized. The weight and species of all fish caught are recorded (Theophile, 2016.

Before January 2021, all field data were recorded using printed forms (Jullan Defoe, pers. comm.). These forms were compiled in a data book and submitted for data entry at the main office at the Fisheries Division. However, from February 2021, the FD upgraded its process and now uses electronic forms on mobile devices to record and submit information from the field in near real-time. These forms were developed by Derrick Theophile and after testing and training with staff, was deployed for use of data collection.

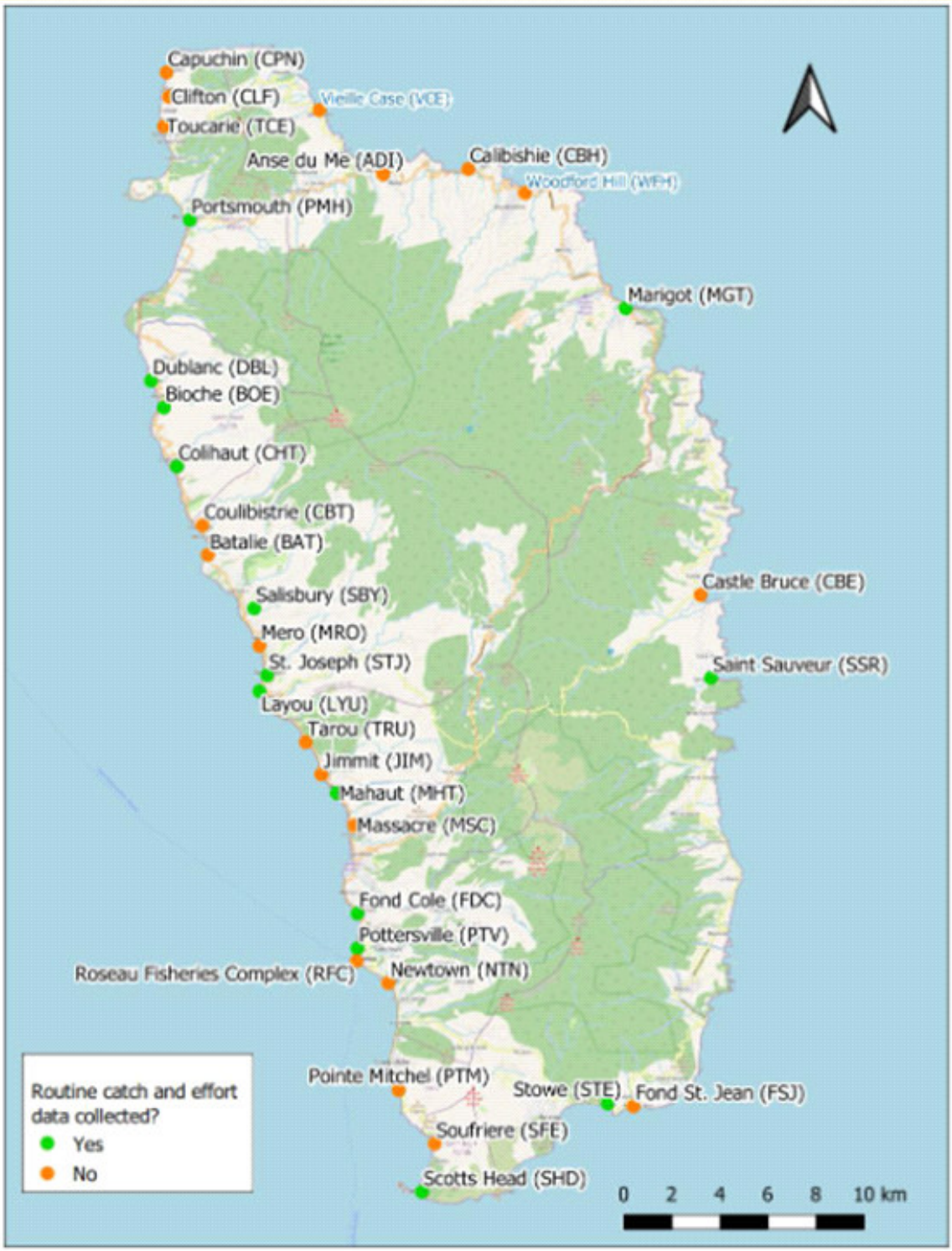


Figure 2. Map of Dominica indicating fish landing sites with active data collection.

Over the past 27 years, the data collection program has undergone minor changes, such as the addition or removal of certain data collection sites owing to changes in fishing activity by location or the availability of temporary funding to expand data collection. Additionally, new, or revised variables were incorporated into routine surveys, namely variables that would increase the resolution of the fishery location , the identity of anchored FADs visited on a fishing trip, the separation of primary and secondary gear, and vessel activity at the data

collection site for the day of sampling (recording the number of boats by type and fishery). New variables include bait used; time of day the fish was caught; bycatch data; number of fish caught; discard data; weighing method (whether a visual estimate or scale was used); biological data (length, weight, sex, and maturity); type of ice used; and amount of fuel used. Changes to variables were largely guided by recommendations in the Anchored Fisheries Model Logbook FAD, developed by CRFM in collaboration with the JICA-funded Caribbean Fisheries Co-management (CARIFICO) project (Theophile, 2016).

Databases

The FD has utilized three databases for data entry and management since 1994. These include the Trip Interview Programme (TIP), the Unified Fisheries Database (UFD), and KoBoToolbox. TIP was introduced by CFRAMP in the early 1990s to digitize data management systems in member states (CARICOM Fisheries Unit 2001). It is an MS DOS-based desktop application designed to store and manage data related to the data collector, fishing activity, fishing effort, catch, gear used, fish prices, and fish biology (length, weight, otolith samples) (CARICOM Fisheries Unit 2001). TIP also can export data in various formats such as database files (DBF), text, and spreadsheets (CARICOM Fisheries Unit 2001).

The Licensing and Registration System (LRS) was a partner MS-DOS database application to TIP, used for managing fisher and fishing fleet details. TIP was supported for many years by CFRAMP and then by the CRFM and remains in use by some member states in the region despite attempts by CRFM to upgrade to the Caribbean Fisheries Information System (CARIFIS) in the 2000s.

The UFD is an MS Access database developed in 2007 as a response to the challenges arising from the attempts to upgrade from TIP to CARIFIS. CRFM member states did not use CARIFIS due to technical issues with the software and insufficient training of staff. The technical issues were of most concern to the Dominica FD. Revisions and upgrades were made to the UFD at intervals to add necessary functionality and capture new data as the fishery and monitoring programme evolved.

The UFD comprised modules that were analogues for the original TIP and LRS but joined to create a single database, making use of the relational database structure. UFD included functionality to generate reports, print licensing and registration forms, and print fisher identification cards. Data could also be exported in spreadsheet formats.

KoBoToolbox is a free, open-source, cloud-based data platform that enables the capture, management, and visualization of data. An electronic form is created using the webform builder or the XLSFORM standard and deployed to the KoBoToolbox server. In the field, the electronic form is accessed on a mobile device via a dedicated mobile application or via a modern web browser. In both cases, the form can be used offline to capture records repeatedly while in the field.

Data is automatically submitted to the KoBoToolbox server when the mobile device regains internet access. With this system, field data collection benefits include speedier and more efficient data collection and entry, reduced cost of data collection, higher quality data capture through built-in form logic and data validation, and improved fieldwork control and flexibility through easy questionnaire revision and deployment. KoBoToolbox brings the benefits of a database to field data collection.

Data handling

The R statistical software

The R statistical software (R Core Team 2023), simply called R, is used for data preparation, statistical analyses and generation of plots and tables and data visualization. R is a free, opensource software environment used for reproducible, computational statistical research, which runs on desktop and server operating systems, including Linux, Mac, Unix, and Windows, and hosts a plethora of tools for performing a large diversity of operations on data, including acquisition, manipulation and restructuring, cleaning, statistical analysis, and data visualization.

The R workflow is structured in a way that allows for easy interpretation of the code, as it is executed in the order it is written, from top to bottom. R script also allows for the inclusion of notes and commentary, providing as much detail as needed. This makes it easier to document the workflow, comment on the process, and discuss results immediately after the code has been executed.

In addition, the structured workflow in R promotes the adoption of best practices for working with data and data files. These include standardized naming conventions for files and datasets, standardized formats for date-time variables using the ISO 8601 standard, and the adoption of tidy data principles. These principles state that each variable should have its own column, each observation its own row, and each observational unit its own table. Datasets that do not adhere to these principles are considered messy. As part of the data preparation process, imported datasets are tidied and a final tidy dataset is produced for analysis (Theophile, 2016).

Preparation of data

The process of data preparation involves: (i) importing datasets into the statistical software working environment; (ii) documenting datasets and contained variables; (iii) selecting datasets and variables for analysis; (iv) cleaning datasets; (v) standardising datasets; (vi) merging datasets; and (vii) checking sample size for species analysis. Data were first imported into the R working environment, either in the format of an MS Excel file for the TIP and KoBoToolbox datasets, or an R data file (RDA) for the UFD datasets.

The datasets would include the raw data and accompanying code references for decoding raw data responses. For example, fish and gear were encoded in the raw data as an alpha or numeric code, so the reference allowed for decoding to a human readable label such as a common English name. Each database would be comprised of multiple tables or datasets, separated by themes, denoting the type of data contained.

Cleaning which is a required step before analysis is done solely within R. The process involves identifying inconsistencies in the data and correcting these inconsistencies in some fashion. The variables are cleaned by reviewing individual responses, identifying and addressing typing and spelling errors, improbable values, or outliers (for numeric variables), improbable responses (for text or categorical variables), and duplicate records.

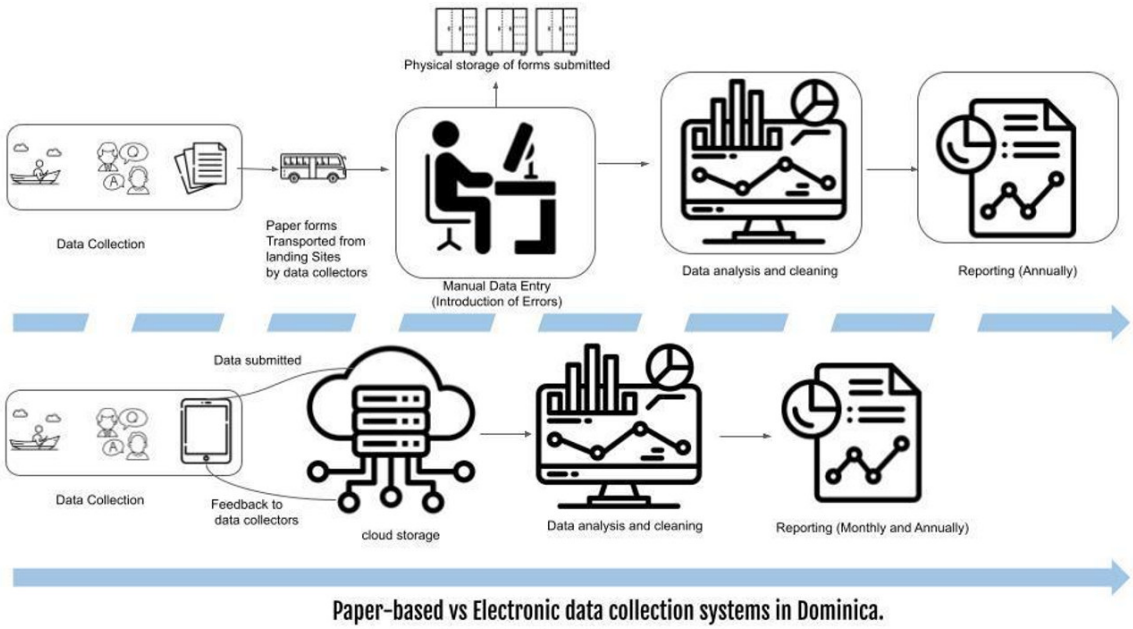
Labels within variables are then standardized to account for changes in labels or codes used over time and across databases. The data type and formats are also standardized to improve consistency throughout the datasets. This includes setting certain numerical values to decimal, or using a standard date format, or correctly ordering categorical variables (for example, months were organised in chronological order).

Datasets within each database are merged using a key variable, common for all datasets, and specific to that database. Key variables contain a unique identifier, numeric or alpha-numeric, for each record, joining the main table to the associated sub tables for the database. In the case of TIP, the key variable was "sequence". In the UFD, it was "site_day_id" for the "FCEsite" and "trip_id" for the "FCEtrip" main tables and "tripIDLink" for the sub table. In KoBoToolbox the key was "uuid" for the main table and "submission_uuid" for the sub tables. Each dataset was structured in a similar fashion in preparation for merging.

Reporting

One of R's core strengths is the possibility to produce reproducible reports which can be used on an annual basis. This is accomplished by preparing a script using R markdown and integrating cleaned datasets into R. This allows for reports to be prepared at a faster rate once a text has been established for the report. This process significantly reduces the time which one would have required to produce a 30–50 page report using excel or any other statistical software.

The areas for reporting include catch and effort, cost of operation, estimated fish landings specific to landings sites, percentage for use of FADs and breakdowns of totals by fish classifications. Inside of this script all analysis and development of graphs can be prepared including texts. Therefore, there is no need to move between an analysis program and a word processing program. The introduction of these steps has altered the work process allowing for



safer storage of information as well as analysis and reporting. A comparison between the two processes is highlighted in Figure 3.

Figure 3. Comparison between manual data collection and analysis vs Electronic data collection and storage in Dominica.

Future plans

The Fisheries Division, while making significant headway with the collection, analysis and reporting of data, has seen the need for further integration of external resources to help with the collection and management of data. Three such interventions include the integration of FAO's

Calipseo, the Government of Dominica´s National Information management system, and the digitization of registration forms for various stakeholders within the fisheries sector as well as fishing vessels.

Calipseo can be considered FAO's response for support to fisheries management authorities to assist in the production of statistics at various stages. This could be either data collection or reporting for policy guidance and stock assessments. This web-based application was developed with open-source frameworks which include Java spring boot, Bootstrap, MySQL for the management and storage of data. Further, R is used for two reasons, the computation of data and R shiny for processing and reporting.

The National Agriculture Information Management Systems (NAIMS) will serve as the Governments of Dominica data repository for reporting of agricultural information. This system once launched will provide persons with access to total amount of farmer and fishers as well as key details needed for reporting within the agricultural sector. This project is still in development.

3. DISCUSSION

Data collection within the fisheries sector is essential to ensuring that marine resources are harvested in a sustainable manner, allowing management authorities to make decisions based on facts. This paper explored the evolution of data collection systems used in Dominica starting from the collection of data to reporting of data highlighting the key achievements and future plans.

While data collection happens in many countries, there is still the need for processing and analysis of the data to give out meaningful information. One setback to this process, however, is ultimately the format in which the data is stored, which is on paper. With the passing of Hurricane Maria in Dominica, this prompted a need to rethink this paper-based collection and storage of information, which has resulted in the integration of free and open source applications for data analysis and storage.

The use of electronic data collection systems allows for even fishers to actively participate in the collection of data. With the growing adoption of technology by the fishing community, electronical data collection allows fishers to feel involved at decision making levels (Granek, et al., 2008). Moreover, this provides the Fisheries Division with further insight and capacity to monitor areas where data collection may not be active, as well as gather additional information which can be used for analysis. In turn, these data collected can be used to assist fishers in the acquisition of financial aid from lending institutions.

Despite the potential to increase reporting and assist in the decision-making process, the shift from paper requires technical staff to ensure that the forms and reporting of information is correct and that submissions are being sent in via each site. Some of these factors however can be reduced by repurposing already present data entry staff or liaison officers.

Of course, with every new technology there are challenges to adoption, such as procuring suitable mobile devices, which may be expensive, and the difficulties of using the technology by people who are not familiar with it. Data in KoBoToolbox may be exported in several formats, including spreadsheets, or accessed directly via third-party applications such as the R statistical software. Currently KoBoToolbox is mostly used for fish catch and effort data collection by the FD.

The digitization of fisheries data collection in Dominica has standardized the typology and frequency of data collection. However, there is still potential for biased data input. Theophille

(2016) noted that data collectors may selectively choose which fishers to interview based on personal preferences and cooperation. While there has been no exhaustive investigation into this issue, supervisory visits with data collectors over several years have not shown these incidents to be widespread or continuous.

4. CONCLUSION

The integration of technology presents a unique opportunity for the monitoring and management of the fisheries sector in Dominica. This new technology allows for the collection of data, analysis, and reporting to be done in a more efficient manner. The potential use of these can be seen through their use in Dominica, which has caused an evolution in the way tasks are done from the field to reporting. This has resulted in reduced loss of data, reduced storage space when compared to space requirements for paper-based system, repurposed personnel to fit in with new systems and faster reporting.

While there is an associated cost attached to moving away from paper-based systems for data collection and steep learning curves associated with the adoption of new technology, there is a growing demand for systems that allow for monitoring of resources and faster decision-making. This improved data collection will not be the solution to all problems but is the start to making more informed decisions and at a faster rate.

Management authorities should therefore explore options that support the integration of technology into the fishing sector. Further, they should support and build the capacity of staff to better understand and maintain the systems needed for data collection and analysis. And lastly, they should work collaboratively with external partners and fishers to ensure that the key source of data collection understands the need and benefits from the implementation of these systems.

LITERATURE CITED

A fishery manager's guidebook—Second edition. (n.d.).

Andereck, Z. (2007). *Mapping vulnerability of infrastructure to destruction by slope failures on the island of Dominica WI: A Case Study of Grand Fond, , Petite Soufrière, and Mourne Jaune.* Miami: Miami University.

Bradley, D., Merrifield, M., Miller, K. M., Lomonico, S., Wilson, J. R., & Gleason, M. G. (2019). Opportunities to improve fisheries management through innovative technology and advanced data systems. *Fish and Fisheries*, *20*(3), 564–583. https://doi.org/10.1111/faf.12361

CARICOM Fisheries Unit. 2001. 'Fisheries Management Data System Terminal Worksop Proceedings: The Way Forward - a Review and Planning Session. 25-28 November, 2000. Castries St. Lucia.' CARICOM Fisheries Unit. https://www.crfm.int/images/documents/Fishery%20Research%20Documents/Fisheries%20Management%20Data%20System%20Terminal%20Worksop%20Proceedings.pdf.

Dominica Fisheries Division. (2011). *Fisheries Industry Census of Dominica.* Roseau.
Eastern Caribbean Central Bank. (2022, 11 3). Retrieved from Eastern Caribbean Central Bank: https://www.eccb-centralbank.org/statistics/dashboard-datas/

Granek, E. F., Madin, E. M., Brown, M. A., Figuera, W., Cameron, D. S., & Hogan, Z. (2008). Engaging recreational fishers in management and conservation: global case studies. *Conserv. Biol*, 1125-1134.

Hilton, K.A. 2023. Assessment of demersal reef fisheries in the Commonwealth of Dominica, GRO-FTP, Reykjavík

IMIS. (2023, January 6). *Maritime Boundaries Geodatabase: Maritime Boundaries and Exclusive Economic Zones (200NM), version 11.* Retrieved from Flanders Marine Institute (2019): https://www.vliz.be/en/imis?module=dataset&dasid=6316

Merrifield, M., Gleason, M., Bellquist, L., Kauer, K., Oberhoff, D., Burt, C., Reinecke, S., & Bell, M. (2019). eCatch: Enabling collaborative fisheries management with technology. *Ecological Informatics*, *52*, 82–93. https://doi.org/10.1016/j.ecoinf.2019.05.010

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Regueiro, L., Newton, R., Soula, M., Méndez, D., Kok, B., Little, D. C., Pastres, R., Johansen, J., & Ferreira, M. (2022). Opportunities and limitations for the introduction of circular economy principles in EU aquaculture based on the regulatory framework. *Journal of Industrial Ecology*, *26*(6), 2033–2044. https://doi.org/10.1111/jiec.13188

Salas, S., Chuenpagdee, R., Seijo, J. C., & Charles, A. (2007). Challenges in the assessment and management of small-scale fisheries in Latin America and the Caribbean. *Fisheries Research*, 5-16.

Steiner, S. (2015). Coral Reefs of Dominica (Lesser Antilles). *Annalen des Naturhistorischen Museums in Wien*, 47-119.

Theophile, D. (2016). *Collection, Management and Primary Analysis of Fisheries Data in the Commonwealth of Dominica.* Reykjavic: United Nations University- Fisheries Training .