11TH MEETING OF THE FINANCE AND ADMINISTRATION COMMITTEE (FAC)

Manta, Ecuador, 25, 26, 27 & 30 January 2024

FAC 11 – Doc 07.2
Data Management Update and Development
Secretariat

Executive Summary

This paper provides a comprehensive overview of the status of the database system development for the South Pacific Regional Fisheries Management Organisation (SPRFMO). The database, originally developed and hosted by FINNZ, faced various limitations in data receipt and processing, storage and management, standardisation, and data integrity. This paper outlines the milestones reached in addressing these limitations. It also discusses the future direction of the system and highlights the benefits enjoyed by continuously improving the system largely driven through member consultations on necessary functionalities to improve operational and reporting efficiencies.

1. Background

The SPRFMO database, initiated in 2011, encountered several limitations which included diverse data formats, iterative data processing, reliance on external data services, evolving data structures, and issues of standardisation affecting meaningful analyses. This limitations were highlighted and a proposal FAC10 – Doc06.2 (Data Management Update and Development) submitted to the Commission for approval. This proposal aimed to develop and improve a set of processes and tools used to automate the collection, processing, transformation, and storage of data in a systematic and efficient manner. Once in place this will greatly improve the management and moving of data from various sources to a central repository where it can be analysed, visualised, or otherwise utilised for decision-making. This paper presents the progress made in re-engineering and highlights future works that will greatly improve data related operational efficiencies in the Secretariat.

2. Achieved Milestones

The short-term project objectives set by SPRFMO have been successfully achieved by the service provider[1] in the following key areas among others, the migration of record of vessels management system and data, the development of the SPRFMO Reporting System[2] , delivery of web services, integration of the VMS data link, documentation enhancements, and the implementation of a geospatial tool.

2.1. Database Migration and Hosting

The 2023 project accomplished the successful migration of the existing database to a new hosting platform, with full responsibility for hosting, security, backup measures, and technical support assumed. The transition was executed seamlessly to maintain the full functionality of existing database services. Additionally, the extant fisheries database was smoothly transitioned to the improved database, with continuous hosting and technical support provided.

2.2. Database Schema and Design

The development of an effective, comprehensive, and normalised database schema was achieved. A robust and flexible framework was engineered to ensure adaptability for future needs. The schema is well documented, with a comprehensive data dictionary to enhance understanding and accessibility.

---

[1] Dragonfly co.nz
[2] The reporting system is currently hosted at https://gateaux.sprfmo.org/reports

2.3. Data Transition and Functionality

All existing data holdings were successfully moved to the improved database, with full responsibility maintained for hosting, security, and backup measures. The web-based connections to the database, such as the Record of Vessels and data uploads complete with change history or audit trail tracked data modifications, ensuring transparency and accountability.

2.4. Data Transparency, Reproducibility, and Quality

The project successfully developed code-based data processing for raw data uploads, along with comprehensive data validation checks based on existing functionality. These measures ensured the smooth flow of raw data into production/operational tables. Code-based reproducible approaches were also implemented to address data issues, enhancing data transparency and quality.

2.5. Web-based Solutions

Modules enabling user uploads of standardised data submissions were developed, complete with error checking and correction features. Users were empowered to edit the Record of Vessels, and an easy-to-use web-based tool for standardised queries and data extracts was created, providing enhanced accessibility and usability.

2.6. Security and Confidentiality

Comprehensive and robust processes based on the latest international standard for information security management (ISO/IEC 27002:2022) were implemented. Recognising the critical importance of data security, the project included a comprehensive backup and disaster recovery plan to safeguard information.

2.7. Data Connections

A secure data connection was successfully established between the SPRFMO Record of Vessels table and the SPRFMO VMS data provider. This connection facilitated near real-time monitoring of vessel activities, ensuring the continuity of data connections with other SPRFMO contractors.

2.8. Geospatial Support

Geospatial support has been incorporated into the database, including the storage of geographical data describing standard areas of interest. Furthermore, infrastructure support for generating standard geospatial plots and integrating with desktop tools such as QGIS has been provided. The Secretariat intends to explore this new functionality as capacity and organisational priorities allow.

3. Future Directions:

Continued efforts are required to maintain and enhance the SPRFMO database system. Focus areas include among others, ongoing standardisation initiatives, upload of historical data, proactive data validation, and continued engagement with Members/CNCPs to ensure a complete vessel record[3]. Regular updates and refinements will ensure the database remains a robust and reliable resource for decision-making. Building upon the successful achievements in database integration, migration, and system design, we propose a comprehensive plan to further enhance the SPRFMO data management system. The following initiatives aim to improve data ingestion, reporting mechanisms, and utilise the newly established Data Working Group (DWG) to ensure continuous enhancement and collaboration with Members.

3.1. Improve Data Ingestion:

Efforts to enhance data ingestion aim to alleviate reporting burdens on Members/CNCPs, increase Secretariat efficiency, and reduce reporting errors. Increased standardisation and improved data templates have been pursued. However, the short-term goal is to transition towards Electronic Reporting (ER) automation especially for updating the SPRFMO Record of Vessels. ER development necessitates substantial Member/CNCP input to ensure fitness for purpose and a framework adaptable to evolving reporting requirements.

Members and CNCPs will need support and capacity building to successfully transition to new tools including portal based self-management of their Vessel data within appropriate constraints.

---

[3] A directive in (second footnote on page 2 of CMM 05-2022) requiring Members/CNCPs to update vessel records.

Transhipment Notifications are a major area of concern and another high priority candidate for automation. The handling of transhipment notifications presents significant challenges due to the high volume of data involved *(with over 2,000 notification data files submitted in 2021 and approximately 4,000 transhipment-related emails received in 2022)*. Usability issues arise as notifications and details are received from both fishing and carrier vessels, with the potential for modifications or cancellations. Discrepancies between the reports from these vessels after a transhipment further complicate matching and compliance assessment.

Through collaborative efforts, a solution can be arrived at, similar to the one utilised by the WCPFC. This application would enable the addition, modification, or cancellation of transhipment notifications and the creation of a unique identifier for efficient data matching. Although such a shift represents a major change in reporting, it is anticipated to offer numerous benefits. Leveraging the experience of other RFMOs in the development of similar platforms could expedite SPRFMO's implementation, with potential collaboration opportunities.

3.2. *Enhancing data reporting and dissemination*

The Secretariat currently disseminates regular reports on various data aspects, such as monthly catches, transhipments, vessel activity, Records of Vessels, SC, and Commission papers. These data summaries are either presented on the SPRFMO website, circulated via email correspondence, or detailed in meeting papers. There may be, however, added benefits of making data/summary reports more available to Members directly (acknowledging appropriate data sharing and confidentiality agreements outlined in CMM 02) as well as to the broader SPRFMO community.

One envisioned approach was the implementation of dynamic reports accessible through the website. Members could query and generate real-time reports, promoting timely access to critical data. Additionally, a focus on data accessibility involves exploring models allowing Members/CNCPs to directly query and download standardised data. This shift in data dissemination would empower Members to independently extract information for monitoring, research, and management purposes.

SPRFMO, working with Dragonfly, has implemented a reporting framework which validates, extracts and updates a reporting table whose data is considered valid for analytics and reporting. The system has been built on top of GitHub and taking advantage of Continuous Integration and Continuous Deployment (CI/CD). This solution offers a surface area where more innovative approaches to reporting can be explored and extended.

3.3. *Integration of Machine Learning (ML) and Artificial Intelligence (AI)*

Long term, the Secretariat could also aim to revolutionise its data reporting and dissemination by leveraging Machine Learning (ML) and Artificial Intelligence (AI). The organisation can employ predictive analytics to forecast monthly catches, utilising historical data and environmental factors for more precise resource management. Anomaly detection algorithms can enhance monitoring by identifying irregularities in transshipment activities, bolstering enforcement capabilities. ML models can analyse vessel activities, distinguishing between routine fishing and suspicious behaviour for improved compliance efforts. Natural Language Processing (NLP) can streamline the extraction of insights from commission papers, while AI-driven dashboards and NLG tools facilitate interactive data exploration and automated report generation.

Collaboration with data scientists (SC), domain experts and Members/CNCPs is crucial to tailor these solutions to fisheries management needs, while addressing ethical considerations and ensuring transparency builds trust with Members. The integration of these AI and ML applications enhances SPRFMO's ability to manage and disseminate data effectively, promoting informed decision-making in the realm of regional fisheries management.

These longer-term plans reflect SPRFMO's commitment to evolving data management practices, embracing technology for streamlined reporting, and fostering a culture of enhanced data accessibility and utilisation. Continuous collaboration with Members and stakeholders remains paramount in realising these ambitious goals.

3.4. Data Working group

The SC established a Data Working group in 2023. The DWG will support the continuous evolution of data management practices. The DWG will focus on discussions related to data collection, management, and dissemination. The group's responsibilities include advising on data needs aligned with the SPRFMO Convention, leading tasks to assess data appropriateness, and supporting the Scientific Committee by providing data summaries and analyses. The DWG will initially report through the SC but could be used directed by other subsidiary bodies, and/or the Commission regarding its activities, data-related challenges, and opportunities.

4. Conclusion

The SPRFMO database system has undergone significant improvements, overcoming initial limitations. The achieved milestones demonstrate a commitment to enhancing data management and accessibility. By transforming the SPRFMO database into a data engineering pipeline, the organisation sets the stage for continuous improvement, innovation, and leveraging data-driven insights to fulfil its objectives. With a strengthened foundation, the SPRFMO database is poised to support informed decision-making and research within the organisation.