

Working Paper 14

Original: English

**Towards a data sharing policy for survey and monitoring data
collected and/or stored by the SPC Coastal Fisheries Programme
on behalf of member countries**

*Prepared by Franck Magron, Reef Fisheries Information Manager
SPC Coastal Fisheries Science and Management Section*



TOWARDS A DATA SHARING POLICY FOR SURVEY AND MONITORING DATA COLLECTED AND/OR STORED BY THE SPC COASTAL FISHERIES PROGRAMME ON BEHALF OF MEMBER COUNTRIES

1. The SPC Coastal Fisheries programme provides assistance to member countries in a number of domains and in various ways such as conducting surveys, advising on management issues, providing and maintaining databases or simply backing-up data on behalf of countries and territories. All these activities imply that data is collected in collaboration with SPC staff or shared with SPC staff by fisheries departments, NGOs, regional and international organisations, private companies etc.
2. As a regional organisation, SPC is regularly approached by the scientific community, other organisations or even the private sector, for fisheries, trade, socio-economic and resource data. As we are the custodians but not the data owners, we won't release it on countries' behalf without their express consent, regardless of whether it is data provided to SPC, collected in-country by SPC staff or stored in a system we established or maintain.
3. At a time where more and more governments adopt open data policies to increase transparency and accountability, the systematic refusal of data disclosure by SPC is often perceived as data protectionism while we merely protect country and private sector information.
4. Data sharing can continue to be conducted on a case by case basis by countries, but it could also be facilitated by SPC for non-sensitive data. For the latter we need guidance from countries to establish a policy that clarifies which datasets we can release (under conditions set by data owners) and what needs to remain confidential and protected (the default policy).

Benefits of data sharing

5. While not all data is shareable, there are some benefits of sharing non-sensitive datasets. For example showing that surveys are regularly conducted and providing data about the resource status, catch levels and the importance of the fisheries to communities raises the profile of coastal/inshore fisheries in the region and can play important roles in global analyses, feeding directly back to countries as advice for best-practice local-level management practices.
6. Aggregated outputs and indicators are often required by donors to assess project outcomes and make decisions on further support and funding. In the absence of information, the case for continuing or promoting specific activities in the region is weaker even if the project successfully attained its broad objectives such as creating employment, increasing food security or providing income.
7. Releasing data also allows independent reanalysis by third parties (reinforcing the credibility of conclusions) or comparison at later point in time with new survey data to highlight changes. In the absence of real data some scientists and NGOs use information available from other regions of the world such as South-East Asia or the Caribbean and apply general assumptions from those studies to the Pacific, even if the situation is not comparable, and often to the detriment of the Pacific countries.

8. Finally sharing data is a first step towards greater collaboration between government offices and local and regional partners to standardise survey methods, split the load of data collection and maximise possible outputs and outcomes. At the regional level, sharing aggregated data would not only allow countries to position themselves within the region, but also provide them with information to strengthen their case and attract donors.

Sensitivity of information

9. The sensitivity of specific datasets highly not only depends on the nature of the data and when it was collected but also on the level of aggregation, whether it is already publicly available from other sources and what the benefits / consequences may be if the information was to be released.
10. For example, 5-year old underwater survey data might be shared with the scientific community and serve as a baseline for temporal and spatial comparison. Sharing that dataset whereby countries would also obtain in return the newly collected data and results would be beneficial to data owners without impacting the resource. On the other hand, detailed trade information such as importers' contacts, volume and value traded at the level of a single and specific exporter need to remain confidential to protect the business. The case would be different when aggregated at country level for example, also noting that depending on the commodity some trade information is available via other sources (country declarations to FAO and CITES, statistics office, customs etc.).
11. At a recent sub-regional workshop on aquarium trade & CITES databases, we asked participants to classify each type of data as low, medium or highly sensitive and for a combination of a given type of data and a variety of potential users, whether the data should be considered open access, of restricted access (e.g., permission needs to be sought) or confidential. In instances where prior authorisation would be needed we also asked participants to specify who should give it. Finally we asked participants to differentiate between raw and aggregated data at the country level.
12. Most participants determined that data can be shared with other government services and with SPC staff either automatically or once authorised, but that data should not be disclosed to other countries, CROP organisations, donors, NGOs or the public and that scientists and researchers need to obtain permission from data owners (fisheries department, environment department or company depending on the type of data).
13. The fact that the majority of participants felt that researchers should seek data use permission directly with data owners reflects the majority of experiences to data whereby countries feel that not enough information has been provided on how the data will be used, and to what end (e.g., design of MPAs etc.). Moreover, data providers feel they are not sufficiently involved in the analysis and that outputs are rarely communicated in a timely manner if at all (something that is exacerbated by the publication delays of scientific papers and theses). To improve upon these experiences and perceptions, data provision by data owners should be conditioned by the acceptance of specific terms of use and (possibly) require in return that the researcher provides any new data related to the country to ensure a two way data sharing.

14. The classification exercise otherwise showed little differences in responses between raw and aggregated data, and we spotted some inconsistencies such as more restrictions for aggregated data than for raw data, and discrepancies between sensitivity and sharing policy. In hindsight the exercise would have benefitted from clarifications and examples regarding aggregated versus raw data as well as some more time for decision makers to weigh benefits and drawbacks of sharing data for each type of data and potential user.

Establishing a regional data sharing policy

15. The exercise conducted for aquarium trade data showed that it will probably take time to reach a regional consensus on data sharing. Moreover there are many other datasets to consider, which could be grouped in broad categories, for example:
 - snapshot surveys conducted by SPC staff and country counterparts as part of a regional project;
 - snapshot surveys conducted by countries for which SPC expertise is sought for analysis;
 - long-term fisheries dependent surveys through logbooks, market / creel surveys, biological sampling etc.;
 - trade and fishery data through monitoring of domestic shipments and exports; and
 - inventory of aquaculture installations and annual production.
16. Currently by default, all raw data collected by or shared with SPC staff is confidential and will be treated as such. Yet some data could potentially be shared with a larger audience for the benefit of countries.
17. Ideally a data sharing policy and associated general principles could be decided at the regional level for the various categories of surveys, level of data aggregation, types of users and year of data collection. The policy would include a list of public domain and non-public domain datasets with a risk classification (sensitivity) and associated dissemination and protection rules. Similarly to what has been established by WCPFC for tuna data, the policy would also include procedures to follow for obtaining data.
18. Countries, in collaboration with SPC and relevant partners if and where appropriate and relevant, would need to determine how this policy would be established and how often they would seek to revise it. The policy would need to incorporate exceptions, in other words, special cases or specific adjustments that might be required by a country or a project; something likely to happen because of the large scope of activities that the coastal fisheries programme conduct for member countries.
19. At this point in time, whether a regional data sharing policy agreement is necessary and achievable (or not) for coastal fisheries data remains an open question.

Questions to the heads of fisheries

20. We propose a set of questions to the Heads of Fisheries to guide us on how the SPC coastal fisheries programme should handle requests for information and data.
- i. Is there a need for some regionally agreed rules and procedures that SPC could apply for coastal fisheries and aquaculture data collected in country by SPC staff — or shared with SPC staff by governmental services? [YES or NO]
 - ii. In the absence of a data policy, should SPC continue to assume that all data requests require country authorisation and redirect requestors to country government services for approval and terms of use? [YES or NO]
 - iii. Should SPC publish the list and description of existing datasets for surveys conducted by SPC staff and country counterparts? [YES or NO]
 - iv. Should SPC publish the list and description of existing datasets for other types of surveys and long term monitoring conducted independently by countries? [YES or NO]
 - v. Should SPC provide a regional central data repository so that countries can publish their own lists and descriptions of datasets? [YES or NO]
 - vi. Should SPC provide an archival service for country datasets? [YES or NO]
 - vii. Should SPC provide an online system where that government services can decide by themselves the level of “sharing” they feel comfortable with for each data type and aggregation level? [YES or NO]
 - viii. For resource surveys (e.g., species, size, number, of fish, sea cucumber, other invertebrates counted on a given transect at a given site) conducted by SPC staff with country counterparts, can validated raw data become open access 5 years after the survey? [YES or NO]
 - ix. For socio-economic surveys (e.g. household and fisher interviews) conducted by SPC staff with country counterparts, can validated raw data become open access 5 years after the survey as long as it is not possible to identify specific households and individuals from the data? [YES or NO]