`

Appendix 11: DATA-SHARING AND ACCESS Report

STATUS AND TREND AND RED LIST FOCUS GROUP

September 11th 2014

BACKGROUND

Hong Kong is probably one of the best studied/surveyed areas in Asia. We have an active community of researchers, government authorities, NGOs, consultants and individual experts and naturalists conducting studies, surveys, monitoring or making interesting anecdotal observations, both now and historically. There is a tremendous wealth of information gathered by this community, some dating back to early explorers to Hong Kong in 1880's. The knowledge is not only invaluable in helping us to understand the status and trends of biodiversity in Hong Kong and the changes through time, but also to guide us in developing conservation actions and making decisions. There is a strong need to share, catalogue and spread this knowledge so that it can be used in conservation, education, raising awareness, sustainable use of our natural resources, inspiring people, generating new knowledge, and determining knowledge gaps that need to be filled while avoiding repeating work already done. Hence, how to manage and share the existing data/information is of considerable interest and utmost importance, as is the establishment of a mechanism for long-term archiving and access.

This Focus Group was tasked to take stock of the existing information on, and examine the status and trends in, local biodiversity, discuss data-sharing and access, identify data gaps and develop

a framework for assessing Hong Kong species of conservation concern and for improved data-sharing and access into the future. In addition, local Rd List assessments were conducted, when possible, to initiate the process of understanding the status of Hong Kong's biodiversity. When funding is available, a list of species of conservation concern will also be compiled, using an already agreed-upon protocol based on alternative information, when full Red List assessments could not be completed. Species identified to be of conservation concern should be considered priorities for subsequent Red List assessments. Both terrestrial and marine biodiversity was considered. **Ten recommendations on data-sharing and access are provided based on the discussions of this Focus Group.**

INTRODUCTION TO THE ISSUES OF DATA-SHARING AND ACCESS

A fundamental part of understanding the status of biodiversity is to assess the conservation status of species (which can range from Least Concern [i.e. no problem] to Extinct using IUCN categories) by the compilation of relevant information on status and population trends. With sufficient information on populations and their trends over time, biology, habitat, exploitation, etc., and the application of a structured assessment framework, species status can be transparently assessed and data gaps on assessed species determined. The IUCN Red List assessment framework, a widely used and globally respected system of well-established and widely understood categories and criteria, was trialed and found to be valid for regional (Hong Kong) species assessments, with some adjustments. Using this system, a diverse range of species was assessed for their conservation status. Given limited manpower and resources, only a subset of all of Hong Kong's biodiversity could be assessed for this first phase of the BSAP process, largely through the donated time of experts in the Hong Kong community who assisted the HKSAR government in their BSAP process. Not all taxa could be assessed either due to lack of local expertise or because some local experts could or would not participate in the process or provide access to relevant information.

In compiling information from government sources to conduct assessments on Hong Kong's biodiversity, the FG members found several challenges to accessing reports and data of sufficient detail to be useful for assessments highlighting shortcomings in data access and archiving

practices. Examples include (1) reports from AFCD-sponsored and completed studies, many of which are likely to be relevant to the assessments, are not always available on the website due to issues including sensitivity of the information, inability to locate existing reports due to poor archiving,.or because hard copies of reports had not yet been made digitally available, (2) government surveys conducted often are not analysed or released in a timely manner such that the major findings may not be known for many years or, indeed, ever. Although survey findings are sometimes published in the AFCD newsletter 'Hong Kong Biodiversity', for some taxa very few studies are published (e.g. marine fishes and invertebrates) and methodologies may not be provided in sufficient detail to evaluate scientific rigour and hence validity of the study and (3) government study/survey/detailed fishery data were not available for FG members use in the BSAP work; it was not even possible, at least initially, to locate a full listing of studies that had been completed hence since members do not know what information is available to allow them to follow up and request access.

After raising these issues with the government, a list of titles of some AFCD-sponsored studies was produced and some reports relevant to the work of this FG were made available. This is significant progress but many of the key findings and/or data accumulated by government's many-years of surveys remain unavailable, including many older reports (decades old). Moreover, some reports located contained too little information (such as simple species lists without quantitative information of any kind or survey methods used in the case of some Marine Protected Area surveys) to be of use in species assessments. Hence the stock-taking work and use of government information was limited by the access to government data and was not as comprehensive we would have liked for some of the red list assessments. It is not known whether older/other reports of work known to have been or being conducted, are available, but all available information was compiled and considerable progress was also made in locating old reports through compiling reports filed within the personal collections of local academics. Combined with expert knowledge, scientific publications, available reports and unpublished data, many species assessments were satisfactorily completed.

This Focus Group has identified three issues in relation to data-sharing and access which need attention under the ongoing BSAP process that will also greatly help biodiversity knowledge

management in Hong Kong in general:(1) sensitive data, (2) transparency and access to information (e.g. the form in which government or other information is made publicly available and level of detail released--in particular when data have been collected using public funds by government, academic institutions or others, and (3) long-term data management/archiving. These issues were discussed and recommendations for follow up action made.

Current background on data access:

It was not possible to comprehensively address the issue of data access across all sectors in Hong Kong due to time constraints during the voluntary focal group process. Inevitably, discussion initially focused on data access from government because BSAP is a government process convened by government which also holds much environmental and species data of direct relevance to species and environmental assessments. However, it was recognised that data on biodiversity are being collected by several different parties, including the government, academics, NGOs, consultants as well as amateurs and that the discussion undertaken, especially in the case of public fund use, applies to all sectors. Should funding be available to continue BSAP work, then a more comprehensive discussion on data access and sharing can be conducted.

(1) *Government*: although the HKSAR is bound by its Code on Access to Information (Code) this is not clear in a number of areas which need to be addressed. In particular, the Code does not (1) provide a clear/standard definition of "sensitive data" nor is it specified (2) how much data should be released and how it should be released. Currently, clauses in the Code most relevant to release of environmental data include, in the case of sensitive data: 'information the disclosure of which would increase likelihood of damage to the environment or to a rare or endangered species or its habitat'. In the case of general information release, the Code limits public access to government "information relating to incomplete analysis, research or statistics, where disclosure could be misleading or deprive the department or others of priority of publication or commercial value". There is also a clause on restricting access when insufficient funds are available to be able to compile requested data. There does not appear to be any clear requirement for government departments to regularly archive or maintain completed reports or data, nor any

centralized or systematic access to these that facilitates access to information by the professionals and/or the general public.

There is much scope and considerable importance for government to clarify the following: (1) what constitutes sensitive data; sensitive to whom, who decides, and with what interest; public/private? (2) How much data should be released, within what time frame, to whom, how and in what form? A simple summary of a species list in a country park, for example, without quantitative details, such as numbers or study methodologies, may be good for general public information but is useless for meaningful species assessments or analyses, or determining trends over time. In such cases, how much of the original study should be released, how is this decided, and by whom?

In general there should be good reasons not to release publicly funded data back to the public unless it is clearly identified to be sensitive. This should apply to data collected not only by government but also by academic institutions, NGOs and other organizations for the purpose of sharing and analysis of data and information access.

Once studies are conducted and completed, reports should be made available, according to the Code (unless sensitive). Yet if they have not been archived, or made available on the internet (even just in the form of report titles), or otherwise accessible, their existence cannot be known. Long-term monitoring and surveys conducted directly by AFCD need particular attention because these represent valuable insights into trends and history that are extremely important for evaluating changes (positive and negative) due to management, development, mitigation efforts, etc. Other than AFCD's studies and surveys, there are studies on biodiversity conducted by academic institutions and other parties that should also be covered under the broader discussion on the access to information if public funding is involved. Further discussions and actions on this issue are needed subject to recommendations and funding made available for the next BSAP phase. There are precedents of possible approaches to the issue of data release such as the requirement for reporting/publication by the research grants council or a time limit introduced for OPCF-HK funded work which allows several years for researchers to publish their findings

after which projects reports can be posted on-line. Such measures ensure that data are not locked up forever.

While many AFCD studies have reports variously available on the AFCD website, with varying levels of detail, or the summary findings might be published in Hong Kong Biodiversity Newsletter - (2002-2012), the quality and detail of the reporting is highly variable and at least one major component of Hong Kong biodiversity, its marine biodiversity of fishes and invertebrates (thousands of species), is rarely reported on. Given the considerable public expense and effort expended to collect such data, it is extremely important that it is more readily access for public use and good, and as part of the historic record of the city.

Overall, archiving, in general, and from the experiences of members of this Focus Group. as well as of some other taxa focus groups who attempted to access relevant government data in relation to biodiversity, is not organized and much is not readily available in meaningful format. There is clearly a need to improve data management, access and transparency. It is also important to ensure that existing websites or databases are updated and maintained. If studies have been publicly funded, completed reports should be made available unless considered, with justification, to be sensitive.

(2) Wider society: considerable information is also held in the wider Hong Kong society ranging from consultancies (many such studies may be government-funded), NGOs, academia, institutions and individual experts. While much may be published in the scientific literature or available in reports (details of data are often nowadays published as appendices to journal publications, in theses, etc.) there may be issues of 'publication rights' or 'intellectual property' that may need to be considered, with various licensing options now available to restrict commercial usage (e.g. https://creativecommons.org/choose/). Where it is not published, there may be a time cost in the informed update, compilation and interpretation of this data that needs to be considered.

In order that data are not tied up indefinitely, funders could strongly encourage publication in their research agreements or stipulate a reasonable timeframe for release of data that ensures the

investigator a reasonable time for publication while ensuring that data or analyses are not locked up indefinitely, especially in the case of publicly funded studies, and there need to be clearer guidelines on access to reports and format for data release. Note that EIA reports submitted under EIAO are available on the EPD website. Ecological information in the EIAs including results of baseline surveys and habitat maps are available online. Ecological studies under ECF could be made available electronically rather than just in the form of a soft copy format which can be extremely inconvenient.

DISCUSSION

Data sensitivity: A clear understanding on what constitutes 'sensitivity' in relation to data release by the government, and other publicly-funded bodies/sources is needed. Valid (yet to be clarified) reasons for refusing a general request for data should be stated in writing. Confidential data should still be made available to the relevant bodies to enable appropriate conservation action and decision-making regarding the sensitive area or issue to be taken, if necessary, and there could be some binding terms restricting the use and circulation of such data. The default position should be to share/disclose information as this will further conservation, with data only being withheld if certain criteria are met. The criteria should be transparent, and relate to the protection of the species in question if threatened or otherwise of concern. Examples of sensitive data are most likely to be those that involve information that might facilitate access for collection or exploitation of valuable species, or encourage, facilitate damage to be done to species or habitat.

Hong Kong is a small area with intense development pressure. Species distribution data at fine scale are normally considered, by AFCD, sensitive and hence not released to the public, as these may be used in malicious ways resulting in damage to species or habitats [Note that this is generally agreed upon within government, not limited to the AFCD]. The FG has discussed the need to establish a protocol for allowing differential data access, with more access rights granted to genuine researchers. While AFCD agrees with this approach in principle, the protocol is yet to be discussed and agreed upon. Currently, release of information to researchers is decided case-

by-case considering their need and intended use, following the principles of the Code. Data on population status (numbers, densities) and trends over time may typically be less sensitive and hopefully generally considered for release.

Data release - what format, how much detail? Release of information from government and other public-funded studies requires discussion regarding the level of detail, and the quality and form of information released. To be useful for biodiversity assessment and conservation planning, data need to be openly, readily and freely accessible and accepted by experts to be reliable (for example species identifications, methodology of studies conducted, etc.). The BSAP process has been an excellent opportunity for our community to learn more about what data on Hong Kong species are held, readily available or otherwise, within the wider community and government, compile information on work being done/completed, foster working relations among experts inside and outside of government and identify means to improve access to information in general. However, there remains much that can be improved to ensure better access to information and to improve the sharing of information.

Although for this first BSAP phase, the resources (time and manpower) were insufficient to make a comprehensive compilation and assessment of available information to assess all species in Hong Kong, considerable progress was nonetheless made in compiling existing reports from the scientific community, government studies, consultancies/NGOs and from the knowledge held individually in the community by the various experts. It is clear that conservation planning and biodiversity assessment should embrace knowledge from the wider community (i.e. which has considerable expertise, and a great willingness to share it). Indeed, it is largely due to the wider, non-government, community that the completion of the Hong Kong initial Red List assessments for the BSAP process has been possible. This voluntary effort supported AFCD-led efforts to conduct this work under the BSAP process. Overall, the engagement of the wider community was a very positive initiative by the government that should be lauded and further fostered.

Hong Kong Government should take the lead in setting up an effective management system for compiling and storing existing data/information held by government and within the wider community, especially in the case of public-funded projects. There was not sufficient time to

discuss how this, and what exactly, could be done but having some kind of data depository in a city the size of Hong Kong would be a very worthy goal and valuable resource. Perhaps it could be done together with the development of a Natural History museum or facility. This topic needs further discussion and consideration to make better and easier use of data available in Hong Kong.

For the wider community to put their data/information into the government system there is a need for clear guidelines on intellectual property and data-sharing to ensure that it can be appropriately accessed. There could be guiding principles on how the process of data-sharing and access could be managed, who will be in charge, and how the data will be released. Otherwise, some individuals/bodies might be reticent or mistrustful of sharing their data if they worry these would be used without due credit or they will be denied access to it at a later date.

Archiving and long-term data access and management: Once study reports are completed, publications issued, etc., their existence should be made publicly known. Data analysis and report compilation should be built into long-term monitoring/surveys and reports shared as outlined above. Unpublished work, especially if publicly funded, should be properly archived and catalogued. One option could be to create a centralized database or institution in which government, and other available (but non-sensitive), biodiversity studies/reports/papers could be digitised and to keep a live database to make information more widely available. At the very least, titles and abstracts of existing reports should be made available.

To aid discussion of the above issues, there are protocols already in existence in other countries that can be used for guidance to develop a process for Hong Kong to move towards better practices for data management, access, and use of sensitive data. Examples from other countries/organisations of how biodiversity data have been collated, standardised and compiled, can be examined, as well as examples of decision-making processes for what was made publicly available or available to the conservation/research community. For example, the USGS protocol (**www.usgs.gov/gov/datamanagement/shar/sensitivedata.php**), IUCN guidelines, and the Global Biodiversity Information Facility (GBIF), National Biodiversity Network (NBN) in UK, Global Biodiversity Information Facility (GBIF) and biodiversity database in Taiwan. The UK's

National Biodiversity Network (NBN) was highlighted as an example of how data are collated from multiple sources, filtered, integrated and shared. In the GBIF context, the most relevant model to look at is the 'Participant node: http://www.gbif.org/participation/list . The Chinese Academy of Sciences recently joined as an associate participant organization and has published its first dataset through GBIF (see http://www.gbif.org/node/e760dc6f-dd68-474d-ab41-bd3588571793.

It is clear that actions are needed to address (1) the current situation regarding management and access to existing data (sensitive and otherwise) in government and other publicly-funded studies, and secondly (2) to plan for a system of improved data-sharing and access across government and non-government sectors in Hong Kong in the future that pulls in expertize from the wider community. Both actions are contingent upon receiving additional funding and hence two scenarios should be considered: (1) options, opportunities and priorities available if additional funds are received, and (2) actions needed under the *Status Quo*, i.e. no additional funding.

TEN GENERAL RECOMMENDATIONS FOR DATA-SHARING AND ACCESS

It should again be highlighted that while the focus is largely on the government because this is a government process and BSAP is government-led, with government possessing much data needed for assessments, the issue of public data-sharing and access, at least when public funds are involved, also applies more widely to the community as a whole, including academics, NGOs, etc. Time constraints for this Focus Group precluded the much-needed broader discussion on this issue to address all sectors of society on the issue of data-sharing and access.

1. AFCD to develop/establish some form of database/centralized archive on Hong Kong biodiversity to pull together the wide range of materials, reports, data, studies, etc. that have been completed, already published, etc. To be considered first are those held by AFCD and other publicly-funded studies. PDF copies of releasable reports should ultimately be made available to

ensure access to completed studies, as well as reports of historical interest, and compilations of titles of available reports be posted on-line to increase awareness of government work. There are a number of databases held in other countries that could be useful as reference examples.

2. There is considerable expertise on biodiversity in Hong Kong which is little-tapped and could greatly assist the government, limited as it is in resources in this area, on species information, interpretation of past taxonomy, phenology, threats, assessments, survey designs, data analysis, etc. AFCD should proactively seek advice from them. The current BSAP process has done much to bring these experts together and it would be a positive step for government to continue to engage experts more broadly in their (government) work and consultations. Hong Kong is a small place and it cannot be assumed it can maintain a full expertize among its full-times staff across all taxa and environmental issues.

3. Considerable data are collected by AFCD. While the findings of some are published through a number of channels such as newsletters, papers, reports etc much of it is not, evidently, readily available to the public conservation community, or has not been updated, or data have not been organized/analysed/archived to be readily available to the public. In some cases, data release, is restricted under contract with researchers, is inconsistent/incomplete or insufficiently detailed to ensure transparency and to be useful for assessments. Guidelines are needed to improve data management, release and access which deal with these issues.

4. There is a need to clarify the issue of data 'sensitivity'. Generally speaking, highly sensitive, mainly spatial, data should not be released to the public and should only be circulated among relevant officers in the authorities and relevant experts/ scholars. A mechanism should be established that relevant expert(s) should be consulted before releasing the information especially in cases that the information is indeed provided by the expert(s). This issue needs further discussion, as noted above under 'Data sensitivity'.

5. For raw data considered non-releasable, summaries or publications covering methodology and data analysis (such as a more detailed version of the AFCD newsletter Hong Kong Biodiversity) could be produced to make information available on all taxa studied in a

form useful enough to advance understanding of the species, or habitat. While this occurs for some taxa, it does not occur for all (such as marine fishes or fisheries). Information from MPAs and AR studies from the last decade, for example, could be made available in more detail than is currently the case.

6. There is a shortage of available information on marine species, and their fisheries, that needs to be addressed taking into account the large number of marine species that make up a considerable part of Hong Kong's biodiversity. AFCD could conduct more studies on species of commercial importance in relation to sustainable management and encourage/fund academics to conduct more research on local species.

7. It is necessary to complete data compilation in Hong Kong and conduct further local Red List assessments, with initial priority on identified unassessed species of concern.

8. A list of the ongoing significant monitoring and/or surveys of natural resources by AFCD including objectives, methodology and geographic coverage should be compiled and released publicly across all taxa studied. While some studies are available, others are not yet available in sufficient detail, for example, for red listing (e.g. (dragonfly and butterfly). Data analysis and compilation of full report should be built into such studies which will be released when completed.

9. Ecological data received by AFCD using ECF and other relevant studies should be assessed (for their accuracy) and centralised. Sometimes these studies/ reports provide updated and valuable data. As one example, the HKU Biodiversity Survey data has been incorporated into the AFCD departmental GIS. However, other studies might also be highly relevant to the current BSAP process, for example, surveys on horseshoe crabs and amphioxus, from which data could be extracted and made available, along with other studies (e.g., ecological monitoring).

10. To action many of the identified items, there is a need for resources; however, consideration is also needed regarding what can and should be accomplished without additional funding.