REVIEW OF THE DATA COLLECTION AND MANAGEMENT SYSTEMS OF THE MARINE FISHERIES IN THE DOMINICAN REPUBLICAN

FINAL REPORT

Paul A. Medley
Sunny View
Jack Hole
Alne, YO61 1RT
UK
paul.medley@virgin.net

Recommendations

1. Review Objectives of the Data Collection

It is necessary to consider information required to implement the current policy and how that information should be presented so that it can be used in decision-making. Although the decision on how to approach this remains with the Dominican Republic Government, the following are recommended:

- A simple catch-per-unit-effort index should be chosen as the basic performance index of all fisheries.
- For all fisheries, a recovery programme needs to be defined which will come into force when the fishery passes into an overfished state. The overfished state is defined by a precautionary reference point.
- A precautionary reference point can be developed for all types of fisheries from Barahona historical data, interviews with fishers and the scientific literature. The precautionary reference point will have to be agreed with decision-makers to ensure it is consistent with policy.

Based on the review of objectives, other additional performance indices such as fishers net earnings and employment may also be considered useful. For some species, like shrimp, more rigorous monitoring system may be used. In all cases, reference points need to be explicitly linked to management actions whenever possible.

Check the legal basis for the industry's provision of information. Current legislation is vague on this point, and some thought may need to be given on what data must be given for licensing.

2. Develop the Barahona Data Collection System

The Barahona data collection system needs to be adapted to meet the objectives. In general, more detailed information may be required and enumerators require training to ensure they are recording information correctly. The following is recommended:

2.1 Growth, Fishing Mortality and Discards

Mortality and growth cannot be estimated with current data. Data collection needs to include a biological data sampling programme, which takes measurements on sex, length, weight and maturity of sub-samples of the catches. Whereas growth and natural mortality could use information published elsewhere in the region, fishing mortality estimation requires length frequencies. However, these data would only be important for a minority of important species in the catches and are lower priority than collecting the basic catch and effort data.

Discards should be estimated in a combined approach. 1) Interviews should be carried out to get fishers to subjectively estimate their discards on that day. 2) Full catch and discards should be obtained by gear and area by either using a research vessel simulating commercial fishing, or requiring a small sample of fishers to land everything they catch. For example, the Fisheries Department might buy everything they land, and then the fisher indicates what would have been thrown back. 3) An estimate of total discards extrapolating the sample to all landings using both 1) and 2) above.

2.2 Economic Data

Economic variables are well catered for in the Barahona system. Revenue and crew shares are available and appear accurate. Costs are recorded, but are not broken down into their sources. This suggests only average costs are given, rather than specific observations on the cost for that trip. More specific cost data collection would improve data quality.

2.3 Develop Sampling Programme

The current sampling procedure does not follow a statistically rigorous method. In this report, an efficient, simple, robust sampling method has been proposed. Staff will need some training in learning to use this method.

2.4 Develop Database

Although Barahona possesses a large amount of data, it is under-utilised, as much of it is inaccessible. This can only be addressed by developing an appropriate database. The database reporting should be in line with objectives defined above. There are two options:

- Develop a database locally. The database can be adapted from one used in another country. However, the database must be supported locally and will require expertise currently lacking in the fisheries department.
- Use a standard regional database, such as the CFRAMP TIP/LRS system. This has the advantage of external support and will require a lower level of local expertise. However, the system may not fit exactly what is required locally and may need to be adapted.

The preferred option is to use the CFRAMP TIP/LRS system if it is appropriate, as it is more likely to be sustainable.

The database should maintain older statistics as well as new data being collected.

The programme should begin to compile literature on fisheries, fishing communities and the living resources in Dominican Republic as part of an information resource. Some literature is already available in the Department.

2.5 Training

Training is central to developing a data collection programme. While enumerators and other staff have learnt how to use the current system in practice, training will improve data quality and allow use of a database.

Training should include:

- Workshops on the objectives of data collection. All staff should be clear on why data is collected and how it will be used.
- Data collection methods. This will mainly be interview techniques, species identification, vessel registration and collecting biological data (e.g. length, sex, maturity etc.)
- Database use. For most staff, this will be limited to techniques for data entry, although at least two staff members should be able to carry out basic data management tasks.
- Reporting standard information indicating the state of the fishery. This will include statistics, which are already provided from Barahona, and additional information on

the performance indices decided from the review of objectives. Some training in simple statistics and estimation will be necessary.

2.6 Consultation and Education

Organise consultations with the users of the data collection programme, primarily the central and local fisheries administration and the bodies dealing with management and development of fisheries.

Create contacts with local authorities, associations, co-operatives, and fishers representing their communities and organise meetings to discuss data collection and use. Posters and leaflets may be useful in explaining why the data is needed.

3. Extend the Data Collection Programme in Barahona to other Districts

The data collection programme developed in Barahona should be extended to other districts. This should not threaten the Barahona programme, so it is recommended that where possible staff from other district visit Barahona to observe the programme and for training.

Each district should agree a plan to adapt and apply the Barahona experience. Each district should develop a brief project to implement a programme in their area addressing the costs of sampling and resources required, noting in particular the need for manpower and transport.

A first priority would be to extend the experience in Barahona to Samaná. A significant fishery is present in the area around Samaná, accounting for much of the fishing activity in the country.

4. Central Fisheries Administration

The central Fisheries Department should be responsible for national issues and providing technical support to the administrative stations.

A central database needs to be developed which holds, at least, summary information from the databases distributed among the different stations. If raw data are held centrally, the database would become very big and expensive software might be required. Data from the administrative stations can uploaded through the Internet or other means. However, because of the potential complexity of the system, it is not recommended that direct access the database is allowed through the World Wide Web at this stage.

The Central Fisheries Administration should provide analysis and scientific advice. It will not be possible for individual districts to develop the necessary skills for this.

Contents

1. REVIEW OF CURRENT DATA COLLECTION

1.1 Overview

The current review and recommendations are derived from standard approach recommended in the FAO Guidelines for the Routine Collection of Capture Fishery Data (FAO 1999). It is strongly recommended these guidelines are used in developing the data collection programme in the Dominican Republic.

The current system primarily was set up to report total catch by particular species groups. These catches are reported to FAO and used for general management advice.

The most important issue for the Dominican Republic is the cost of data collection and added value to management through obtaining information. This requires a careful consideration of how much information is necessary to manage the fishery and the least expensive way that information can be obtained.

The scope of this report covers marine fisheries only. However, inland fisheries were included in the review, as the data collection system should cover both. In general, the same recommendations apply, although inland fisheries will have their own specific issues that will need to be addressed. These are dealt with briefly as part of this review.

There are two special cases where much more data is collected. These cases, Samaná and Barahona, have collected information for their own purposes as part of previous or current development projects. Both projects demonstrate a capability to collect detailed information given adequate resources.

The Dominican Republic is currently developing a new data collection system. Many of the recommendations in this report reinforce what is already planned. However, it will be still worthwhile for the Fisheries Department to review these plans once the data collection objectives are agreed.

1.2 Fishery Structure

A fishery census conducted in 1990 contains the most up-to-date detailed information on the fishery. There are 160 landing sites with 8597 fishermen and 3700 boats of which 3252 were active. The census identified 19 gear types and in many cases vessels used more than one gear. Gear usage and target fish varied with location. The coast was divided into sixteen districts by the 1990 fishery census. Broadly, the main types of fishery are:

(i) Inshore multispecies fisheries using a variety of gears, notably hook and line, traps, nets and diving. Hand lines are the predominant gear. Fishing is also carried out at night, which changes gear selectivity.

(ii) Offshore fisheries, including catches from areas not under Dominican Republic management

(iii) Shrimp fisheries (*Penaeus schmitti*), mainly from the North-East coast.

(iv) Bottom set longline: a new fishery being developed off the North cost.

(v) FADs are used to catch tuna and kingfish off the southern coast.

Fishing is of low social status and fishers generally have poor education. Earnings are usually between US$200 - 400 per month. Fishing is centred around communities, which are dependent on fishing and therefore sensitive to problems in this industry.

Very few women are involved in the capture fisheries. Women used to sell fish, but the fish buyers have recently tended to consolidate purchases and export bulk to towns for distribution through standard retail outlets, such as supermarkets and restaurants.

Credit is provided by fish buyers who market the fish. Often fishers process (scale and gut) the fish at the landing site before sale. Fish are sold in commercial groups based on species and size.

1.3 Policy and Legislation

The government of the Dominican Republic is currently reviewing proposed legislation. The legislation was drafted with expert help from FAO, and reflects the FAO Code of Conduct. The legislation states, in general terms, the policy aims in managing the fishery. I assume that, while enacted legislation may differ in detail, this policy will remain in general form. In no particular order, the policy can be divided into the broad aims:

a) Resource conservation through controlled access. However, conservation should not prevent exploitation and development
b) Conservation measures should be adopted based on accumulated knowledge of the resource base. This implies it is recognized that data collection is necessary to manage resources.
c) Maximize economic benefits
d) Ensure food security
e) Develop fisheries and aquaculture
f) Protect non-commercial species

However, policy details are not provided, and to a large extent policy with respect to balancing development and management needs to be developed. In particular, there is no guidance given on how to manage overexploited stocks. It is likely that at least some inshore stocks are overexploited and would benefit from a rebuilding programme.

Using the general aims, specific policy might be developed for the different parts of the fishing industry. These policies need to take into account possible states of the resource, such as overfishing. A primary aim will be to define when a resource is overexploited and to develop a plan to rebuild should this occur. It is possible that once this is done, legislation may need to be reviewed to ensure it provides adequate tools to implement the management scheme.

1.4 National Data Collection System

There are 282 registered inspectors, who are available as enumerators in the country covering data collection in both inland and marine fisheries. There are 7 trained biologists and 22 technicians for collecting scientific data. These staff cover both inland and marine fisheries. Inspectors are allocated to administrative districts, which are currently being re-arranged. Significant problems for inspectors include transport to landings areas and a lack of training.

Fish processors, other non-government organisations, such as the fisheries co-operatives, and other government departments are currently not used for data collection.

There was a census of the marine fisheries in 1990 (PROPESCAR-SUR, 1994). This describes the distribution of landing sites, vessels, gears and fisheries around the coast. It forms the basis for the data collection programme. No other surveys, such as a consumption survey, have been conducted.

Fishers do not migrate in the marine fishery, but tend to be stationed at particular landing sites. This makes the fixed vessel distribution obtained from the census valid for structuring the data collection system. There is, however, seasonal migration in the inland fisheries. The inland fisheries were not covered by the census.

Most fishers go on one day (or night) trips. There are some potential problems on whether fishers are full or part time.

Data are collected as the best estimates that inspectors can obtain for the catches in their area. There is no systematic or standard practice implemented for the collection of data, and practices vary among areas. Information from vessels are obtained through interview and inspection of vessel log-books. Inspectors attempt complete enumeration of all landings and no overall sampling programme is implemented. Statistical errors are not accounted for.

There are no observations of vessels activities at sea (through, for example observer programmes). This means there are no estimates of discards or knowledge of discarding practice. In some cases, it may not be clear where catches come from. For example, local fishers on both sides of the Haiti - Dominican Republic border cross the border to fish.

Currently data is limited to catches by commercial category. There is no biological sampling or estimates of effort. No information on discarding exists. However, there are a number of data collection initiatives currently being implemented:

- Vessel and gear registration and licensing represents the most important initiative. Currently the amount of information being requested is limited, and some thought should be given to considering introducing further information gathering using this method.
- FADs, used to catch tuna and kingfish off the south coast, are registered.
- Database development. This is at very early stages. There was an expression of interested in developing a system with Internet access. Microsoft software products, including MS ACCESS, are probably the most widely used in government.

1.5 Data from Other Sources

Processors and fish buyers are a very useful source of information. One processor was interviewed who seemed willing to provide information, although he claimed not to keep written records. The processor did not indicate that any particular data would be confidential, and he would be willing to complete data forms as supplied by the fisheries department.

Information on weather (rainfall patterns and river out flow) may be available from the National Hydrographic Institute. Other data, such as exports, which may provide important auxiliary information, may be available from other government departments. These potential sources of data have not yet been explored.

1.6 Barahona Data Collection

The PROPESCAR-SUR project based in Barahona set up a data collection system which has been in place since 1988. Landings are sampled on a per-site basis and information is recorded from trip interviews. All landings at each site are recorded, but only a proportion of sites are covered. Visits to landing sites are allocated to enumerators at the beginning of each month. Currently there are 4 data collectors.

Since the aid agency withdrew support, no biological data collection has been undertaken. Original biological data forms have become separated from the catch and effort forms, however the links could probably be re-established.

The data collection system currently gathers relatively accurate catch and effort information based on trip interviews. The effort variables, in particular, seem well thought out and adequate for monitoring CPUE and estimating fishing mortality.

Not all data recorded are direct observations. Weights are not measured, but often guessed due to the current lack of weighing scales.

The costs that are recorded are approximate or average costs, rather than what was paid on that day. Prices and therefore revenue, are easily obtained at time of landing. Information is also obtained on the share system (how the revenue or profits are distributed to the owner of the boat, captain and crew).

The enumerator who was interviewed made several suggestions for improving the way data are collected. There appears to be limited discussion amongst staff on ways to improve the system.

There was no database, although all paper records are stored in the office. Monthly summary information is assembled from these data for monitoring purposes.

Nevertheless, the Barahona system represents a good model to apply to other administrative districts. The system has been sustained despite a withdrawal of external funding. This is a very important criterion for data collection systems, that they are robust and can be sustained despite the many problems they are bound to face.

The main problems that need to be addressed are:

Observations: Catch weight is estimated rather than measured directly. While this may give a reasonable indication of this variable, it introduces an unknown error. Directly measured weights, using a spring balance for example, is a much better approach as the error can be quantified.

Fishing costs appear to be given as an "average" in many cases. That is, the value entered on data sheets gives a rounded estimate on the general operational cost for a days fishing rather than specific information on what was paid in fuel costs that day. The problem is that this is not real raw data, but a subjective summary of costs incurred. More specific questions detailing what was actually bought to allow that days fishing (including sporadic purchases, such as replacement gear) would be much more valuable. Such detailed interviews need not occur for every vessel.

Sampling: The sampling methodology could be improved to avoid bias and, probably more importantly in this case, improve sampling efficiency. Setting up a good sampling procedure will be particularly important for other districts besides Barahona.

Data storage and retrieval: Data is stored on paper and monthly summaries are calculated by hand. This makes limited use of an extensive data set, which could provide a great deal more information for management advice. The only way to obtain this information is to hold the raw data on a database, which must be developed. This would make all paper summary forms obsolete.

Analysis and use of information in decision-making: The data are adequate for at least one type of stock assessment. No stock assessment has been undertaken yet, so the state of the resource is unknown.

1.7 Samaná Fisheries Development Project (CEDEP)

CEDEP is a Japanese financed project developing demersal resources using bottom set longlines. There are currently three vessels operating, each of which maintains a logbook detailing set by set operations and catch data. The project is not involved with other fisheries in the area and the primary aim is not management of the resources. Data collection is rigorous, but strongly orientated towards assessing the profitability of the operation only.

The logbooks are not formally structured. The captains of the vessels are trained to record relevant information: date, latitude, longitude, depth, set and haul times, number of ganglions (hooks) and number of fish caught. By-catch is not recorded. At landing, the catch weight and price are recorded and can be related to the trip. Detail cost records are also kept, so the profitability of the vessels can be calculated. However, these data are not kept for monitoring purposes and not maintained on a database. As they represent the early days in the fishery, such data are critical and should be archived for future use in assessments.

1.7.1 Inland Fisheries

The inland fisheries are outside the scope of the current report. However, they will form a significant part of the data collection programme, and are subject to much the same problems. In particular, the sampling system needs review and should be brought into line with the marine fishery. This requires that vessels and fishers be licensed and that landings sites (which are seasonal in this case) be recorded. Random sampling of landing sites and locations where fishers sell product would form the foundation for monitoring this fishery.

1.8 Summary

Table 1 A summary of the strengths and weaknesses of the current data collection programmes. Barahona currently has the best data collection programme. Stock assessments and various analyses can be conducted on these data. With some improvements, the Barahona system should be extended to other parts of the country.

Strength	Weakness
Fishery is well covered by enumerators	In many cases, landings data only collected
System is sustainable	No statistical sampling system has been implemented
	Enumerators have not been formally trained
	No database for raw data storage and retrieval
	Enumerators are not necessarily distributed proportionally to fishing effort
	Data not used explicitly in decision making
Barahona	
Catch, effort and economic data have been collected since 1988	Biological data collection has been discontinued
Individual observations made and recorded	
Samaná	
Accurate logbook system initiated for small number of vessels	Logbook system needs to be formalised with printed forms and database for monitoring purposes

2. DATA AND DECISION-MAKING

2.1 Overview

The standard approach to managing fish stocks has been to carry out periodic assessments and then give general advice to management on their options. This approach has generally not been successful, for the following reasons:

- Disagreement among stakeholders with the assessment
- It is often unclear among different groups on the decision to be made
- Risk assessment is poor, so often the most optimistic forecasts are chosen even though this is bound to lead to overfishing in the longer term.
- There is improper interpretation of scientific information, as reports are too technical and can be misunderstood.
- The older approach has not been pro-active. The assessment focuses on the current state of the fishery rather than considering what might happen and what would be an acceptable response.

While these problems have been addressed in a number of ways, an approach using reference points is rapidly becoming the standard, being applied among fisheries throughout the world. It addresses these issues by focusing decisions on indices, which summarise information on the state of the fishery. The indices are based on the objectives of management, and hence measure something about the fishery that is of interest to decision-makers. Indices are estimated based on a data collection programme, which should produce data that objectively represents the state of the fishery.

2.2 Management Units

Before considering indicators, we must consider what makes up a complete fishery, the unit to be managed. Assessments and management deal with single stocks as the appropriate unit of control. The concept of "stock" is complicated and difficult to define in practice. A living stock could be defined as a group of animals from one species, which share a common gene pool. For the management of fisheries, however, this definition is academic rather than practical, and it is the concept of “management unit” rather than stock that is useful. A management unit is a resource, which we choose to management as a single unit. For this to be meaningful, the unit must possess a way to measure its status and predict how that state might change under different levels of exploitation. In practice for the Dominican Republic, management units will be defined by species or commercial group and location of fishing.

Until more scientific information comes available, management units based on administrative areas might be declared and assessed. Management units should be clearly defined in much the same way that administrative centres are, so that the sea area is divided into statistical areas. Management for the Dominican Republic will primarily depend upon the way the fishery is administered, at least until more scientific research can lead to improved definitions. In terms of data collection, this means that statistical areas will need to be defined and catch and effort data allocated to these areas.

As in the case of the resource, the vessels can be allocated and managed in fleets. A fleet is a group of similar vessels, which are approximately the same size and have the same construction. The vessels of a fleet use the same type of gear and fishing techniques and most often, they share fishing grounds. Identifying fleets is also important in simplifying data collection and analysis. For the Dominican Republic, the registration of vessels must be complete before fleets could be defined.

2.3 Performance Indices and Reference Points

'Performance indicator' is a general term for any quantifiable value that gives some idea of how well a fishery is doing in some respect. For example, the total annual landings gives some indication of how well a fishery is doing. Assuming they are sustainable, high landings may be better than low landings as they imply greater economic output. However, interpretation of indicators is not necessarily simple. If total landings are low, we do not know whether this is because the resource is small, the fishery is in a chronically overfished state or the fishery is undeveloped. To help interpret an indicator we use reference points.

Reference points indicate a target we might want to achieve or a limit we might not wish to cross. For example, the total annual landings may be compared against a reference point, the Maximum Sustainable Yield (MSY), which marks the point beyond which landings are thought to be unsustainable. Although MSY indicates the potential size of the resource, it still does not automatically determine the current state of the resource. Often several indicators are required to manage a fishery effectively.

The main attributes of performance indicators and reference points is they reduce complex information into forms which are relatively easy to interpret. Therefore, they are an important aid to decision-makers in simplifying scientific information and focusing decisions on the important aspects of the fishery.

Decision rules may also be considered. Once we define MSY, we may require a management action to maintain or reduce catches below this figure. A decision rule can be proposed and negotiated that achieves this aim well before the problem arises. This ensures the fishers and the fishing industry are aware of what might happen, making implementation of fishing controls much easier.

Table 2 Some examples of the types of indicators and reference points used in fisheries. Empirical indicators do not have a strong dependence on underlying models, although implicit relationships are assumed. Non-empirical indicators depend directly on the choice of analytical model, which must fit the data, and generally require a scientific research programme.

Performance Indicators		Minimum Data Variables Required
Empirical	CPUE	Catch and effort
	Vessel Profit	Catch, prices and costs
	Average Fish Size	Catch composition; Age-at-first-maturity
Non-Empirical	Yield-per-recruit	A long time series of catch composition and effort
	Maximum Sustainable Yield	A long time series of catch and effort
	Net Present Value	A long time series of costs, prices, catch and effort

2.4 Examples

To illustrate the approach, two examples of performance indices are given. In both cases, performance indices are calculated or estimated from the available data. These indices are compared with reference points, which represent a decision rule. The approach has several important attributes.

The complex information is reduced to a relatively simple form that summarise the state of the fishery and the associated decision-making process. This clarity helps management and industry understand what is going on with the fishery.

The decision rule ensures management and industry have considered their options should the fishery be threatened with overfishing. This focuses the fishery on future options than simply react to current assessments.

2.4.1 USA

In the North East USA, decisions are standardised and encapsulated into a decision rule based on fishing mortality and the spawning stock biomass. Both these performance indicators are estimated using stock assessment. Fishing mortality is closely related to the proportion of the stock that is removed each year by the activities of the fishing fleets. The spawning stock biomass is closely related to the numbers of mature females that provide eggs and therefore recruits in future years.

The decision rule encapsulates two principles (Fig. 1). Firstly, the rate of exploitation should never go above some maximum level. If the fishing mortality is found to be above this level, then the TAC or fishing effort should be reduced. Secondly, if the spawning stock falls below some precautionary level, then fishing mortality should be reduced to rebuild the stock. If the spawning stock biomass falls to a level that endangers the stock, the fishery will have to be closed for the stock to be rebuilt.

The reference points are set according to what is known about the biology of the stock and its past behaviour. In general however, this approach is sophisticated and requires an in depth knowledge of the stock and has limited application in areas which do not have a long tradition of scientific research. However, this or similar approaches may form the ultimate aim of a data collection programme which aims to generate information indefinitely into the future.

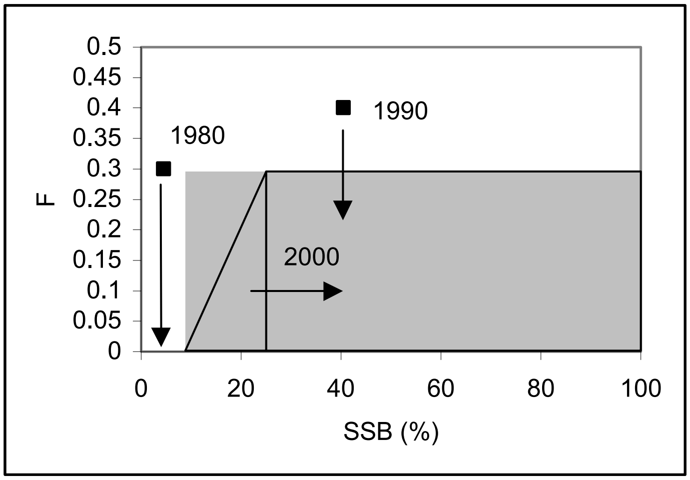

Figure 1 The decision rule indicates the decision to be made based on two performance indices, fishing mortality (F) and spawning stock biomass (SSB) as a percentage of the unexploited spawning stock. The shaded area represents the desirable states in which the stock should reside. In this example, three scenarios are displayed. In 1980, the spawning stock is so low that any reduction in fishing mortality will not automatically drive the fishery into the desirable region. Under these conditions, the rule states the fishery should be closed. The arrow represents the decision, and indicates the fishing mortality should be set to zero (or as low as possible if the species is taken as bycatch). For 1990, the spawning stock biomass is high, but the fishing mortality is also higher than the maximum. If left alone, the spawning stock would be driven to low levels threatening the stock. Again, the decision is to reduce effort or catch quota in line with reducing the fishing mortality below the maximum level. Finally 2000 is in the desirable region. In fact, it lies in the rebuilding area where both fishing mortality and spawning stock are at low levels. The projection, if the decision rule is set up correctly, suggests that the spawning stock will be rebuilt to safe levels as long as fishing mortality is not allowed to increase beyond the limits.

2.4.2 Jamaica

Even where scientific information is lacking, it is quite possible to use the same approach as the USA. With inadequate scientific information, a reference point may be more arbitrary, but can nevertheless guide management in planning actions to maintain the fishery in a desirable state. Reference points in this instance are usually empirical - that is they are based on past observations rather than on any theory.

To illustrate the approach, the shrimp fishery in Discovery Bay, Jamaica is used (from FAO 2001). The objective of the assessment was to decide on a precautionary reference point for an empirical CPUE index. The CPUE index is obtained from a small scale sampling programme. The estimate is subject to a great deal of statistical error, but nevertheless indicates the state of the fishery and resource. The hypothetical decision which management has to make was whether to implement a recovery programme or not. A recovery programme would be appropriate if the stock was declining, otherwise no action is necessary for conservation purposes.

The suggested rule was to test the average annual CPUE against the initial three year average (the data collection programme was three years old). This was considered reasonable as the empirical base line representing a state where the stock is not overexploited. The specific decision rule is: if the average annual CPUE of both gears has fallen below x_{RP}% of the base line (i.e. the reference point), then a recovery programme would be instituted, otherwise it would not. The question is what is an appropriate reference point in this sort of system.

There are two kinds of error possible. We may implement the recovery programme when there has been no decline in stock size (Type I error), or alternatively we do not implement the programme when there has been a decline (Type II error). These incorrect decisions may be made because the CPUE is not an exact measure of stock size and includes significant observation error. Both incorrect decisions introduce a cost. When the correct decision is made we assume that there is no cost as this is the best action we could take (i.e. we have no regrets). In the absence of proper costs (for example, no recovery programme has been proposed, so its potential cost is unknown) a relative cost was assumed between the two types of error. If this method was to be applied, these costs would have to be known, and the recovery programme agreed with the fishers. A recovery programme would include reducing catches of shrimp, but might include some form of compensation, for example.

To find the optimal x_{RP}%, the computer was used to simulate CPUE data with statistical errors consistent with those observed in the real data. Different recruitment trends were introduced to represent possible scenarios.

For each random simulation, the management rule was applied by calculating the average CPUE and comparing it to the base line value multiplied by the reference point percentage. If the average annual CPUE was below the reference point, the recovery programme would be implemented. This simulation was done a very large number of times and the costs recorded.

The optimal reference point was found to be approximately 80% of the average annual CPUE assuming a precautionary cost structure. This cost structure in this case tended to favour implementing a recovery programme, with a relative cost ratio (Type II : Type I) of 1.7 (Fig. 2), although results were robust to this assumption. With real cost information, a true optimal point could be found.

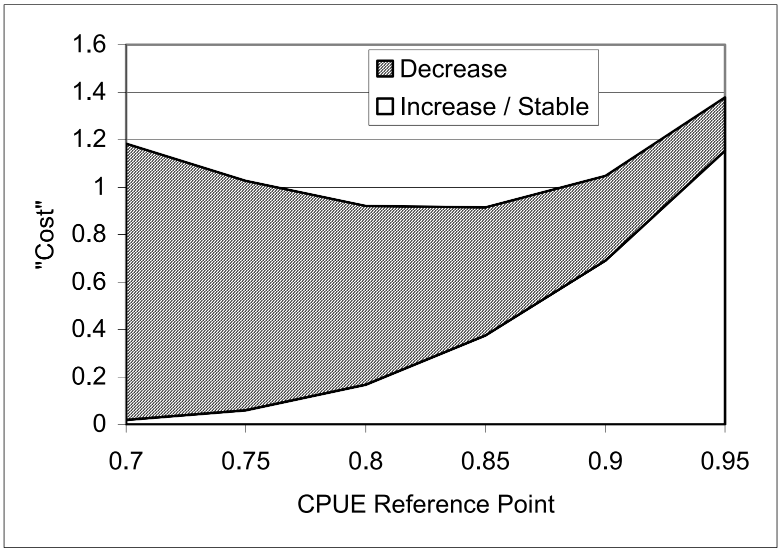

Figure 2 Cumulative two sources of costs from the regret function. The "Decrease" cost is incurred when the population is decreasing but the decrease is not detected by the reference point, so no recovery is implemented. The "Increase / stable" cost is incurred when the reference point method indicates a declining stock when in fact the stock is stable or actually increasing, and an unnecessary recovery programme is implemented. The minimum cost occurs between 80 and 85%.

The method illustrates an explicit risk assessment based on the data collection programme. It recognises that the data collection system is not perfect (they never are), and that decisions should be based on the relative costs of making the wrong decision. For small scale fisheries, this is an appropriate approach. It has a low cost of implementation, allows management to plan ahead and uses information that has a clear meaning to fishers. Although the analysis above would take a little research, the application of the method only requires using averaged CPUE, estimated directly from trip interview data. Therefore, once the reference point is known, the decision rule is very simple to understand and apply as part of the decision-making process.

This system is very dependent on the monitoring of catch per unit effort carried out as part of the continuous data collection programme. It should be noted that CPUEs are more than indicators of just population size. CPUEs are indicative of earnings, profits and economic rent. This makes CPUE or catch rates very useful indicators of the general health of the fishery.

2.5 Recommendations

Based on a review of policy, cost of data collection and consultations between decision-makers and stakeholders, performance indices and reference points need to be developed for the management of resources. This should form the subject of a high-level workshop to agree an approach and initiate the data collection programme.

The performance indices should inexpensive to estimate and simple to calculate. Average CPUE is an obvious choice, and for many of the small-scale fisheries is appropriate. For more highly valued fisheries, such as shrimp, indices that are more complex maybe used, although the improved management should still justify the greater expenditure on data collection and analysis.

Performance indices should cover the status of the stock and the status of the industry. For this purpose, a minimum of three indices are probably required: CPUE, earnings (or profits) and employment. Whereas CPUE and earnings should be available for each fishery, employment can be calculated overall. Earnings may require information external to the fishery for its reference point (e.g. the opportunity cost for fisher's labour). In the absence of biological information, a reference point for CPUE would probably rely on historical data. Barahona data may be the only historical information available, and therefore analysis of the early Barahona data would be important in setting up appropriate reference points for many fisheries.

3. DATA VARIABLES REQUIRED

3.1 Overview

I assume three performance indices are required. More information may be needed for others, however these represent the basic approaches which may be used to start the monitoring and decision-making process. In particular, at least some fisheries will require biological data collection, and this is covered.

It is important not to overload the system with data. While more data may be collected, it must be born in mind that this involves a cost. Not only will enumerators spend more time, it may be more onerous for the fisher, the encoder, and (less importantly) increase the size of the database. In general, there must be a justification for any data that is collected, but that justification can be scientific research or future monitoring needs.

3.1.1 Commercial Group Definitions

When landed, the catch is sorted to meet the request from the buyers. The grading of the catch will be dictated primarily by commercial requirements. Hence, commercial groups are related to market demands and often attract different prices.

The commercial groups will also reflect taxonomy. The catch will nearly always be separated at least into fish, shrimps, cephalopods and others.

3.1.2 Landings

Landings are estimated by normal means as the fish can be inspected at the landing site, where measurements on numbers and weight of fish can be made. It is important that, where possible, real measurements are taken rather than guesses made. This requires providing appropriate equipment to take measurements.

The data most easily accessible for the enumerator is the weight by commercial group and the price per kilogram. Sometimes a commercial group is the same as a species, but in general, the enumerator must take a sample to estimate the species composition of a commercial group. If a commercial group is a size class of one species, the total weight of the species must be calculated from the sum of all size classes.

Catch consists of landings, which means the part of the catch that is actually brought to the land, and discards, that which is dumped at sea. Catch is the variable of primary interest to biologists.

3.1.3 Estimation of Discards from Limited Data

The safest, but also most costly, method for estimating discards is to place observers (enumerators) onboard a representative selection of fishing vessels, and then let the observers record the total catch (landings and discards) and take samples from the discards. The use of observers may have many more objectives than collecting discard data, and in general, they would be required to monitor the fishing operations. Observers onboard the vessel during the fishing trip is only recommended if the vessels are large, and the budget and personnel allows for this expensive activity.

Using research vessels to simulate commercial fishing is a sound approach to collecting samples of detailed data. To what degree experimental fishing should be used is very dependent on the resources available to the data collection programme. Extensive experimental fishing is expensive.

A third simpler but less dependable method to estimate discards is to use a sub-sample of hauls of a trip where the discards are retained for examination. These samples allow CPUE, including discards, to be estimated. However, it is unlikely this can be relied upon in artisanal fisheries.

For artisanal fisheries, where the catches are expected to be small in size, a fisher can be asked to land his entire catch, including discards. This may require the data collection programme to purchase the entire catch, use it to gather biological information, then sell on the commercial fish. This will require a revolving fund, which will need to be topped up regularly.

Discards can be estimated inexpensively through trip interview. The approximate proportion of fish returned to the sea dead can be provided by the fisher. By itself, this information may not be accurate, and should be corroborated by other data. However, as a monitoring index of relative discarding, it may be acceptable.

The information collected can be extrapolated to the entire fleet using either linear models (based on auxiliary information) or assuming an underlying species abundance in the exploited ecosystem (Medley 2001).

3.1.4 Fishing Effort

One of the main aims of fisheries science is to link the mortality of fish with fishing effort. For example, we may wish to find the link between the number of vessel-days of each fleet and the proportion of the stock that was harvested. This not only allows the assessment of the state of the fishery, but can also provide advice on how effort should be controlled to protect the stock.

For a bio-economic analysis, effort is often the link between the biological and economic models. It is related to production through fishing mortality, as well as to variable and fixed costs. Effort expressed as fishing days or days away from port is the most important variable for the bio-economic assessment of fisheries, as the number of active days is often assumed to be related to the variable costs of fishing.

Fishing effort can be measured in many different ways (Table 3). The effort measurement may be selected to fit a specific type of vessel and gear. For example, for a trawler you might use the number of trawling hours, for longlines the “number of hooks per line” and for gill nets the “number of gill nets set per night”. What can actually be used as measure for fishing effort, of course, depends on which data are available from the fishing operations.

The sampling of activity observations may be done with trip interviews, but it may also be collected through the frame survey. From the estimate of the average activity level and the number of vessels, the total number of effort units is estimated as the product:

Table 3 List of possible effort measures, in order of priority according to the ability of measure to provide a relationship between fishing effort and fishing mortality.

FIRST	**PRIORITY**	
Fishing Gear	**Effort Measure**	**Definition**
Fishing with FAD Fish Attracting Device, including pesqueros/casitas	Number of hours or days since last fished	Number of hours or days in which FAD (Fishing Attracting Device) is left since it was fished last time if cyclical fishing. For trolling, the number of hours trolling around each FAD and number of lines set.
Beach seines	Number of sets	Number of times the gear has been set or shot, and the number in which a catch was made.
Castnet	Number of casts	Number of times the gear has been cast, and whether or not a catch was made.
Trawls	Number of hours fished	Number of hours during which the trawl was in the water (midwater trawl), or on the bottom (bottom trawl), and fishing.
Gill nets (set or drift)	Number of effort units	Length of nets expressed in 100-metre units multiplied by the number of sets made (= accumulated total length in metres of nets used in a given time period divided by 100).
Gill nets (fixed)	Number of effort units	Length of net expressed in 100-metre units, the number of times the net was cleared and soak time.
Traps (uncovered pound nets)	Number of effort units	Number of days fished and the number of units hauled.
Covered pots and fyke nets	Number of effort units	Number of lifts and the number of units (= total number of units fished in a given time period) and estimated soak time.
Longlines (set or drift)	Numbers of hooks	Number of hooks set and time of setting and hauling.
Pole-and-line	Number of days fished	The number of days fishing (24-hour periods, reckoned from midnight to midnight) including days searching. Similar to purse seine, in that schools are searched for then fished, except no net is set.
Rod-and-reel (recreational)	Number of line-hours	Number of hours during which the lines were in the water times number of lines used.
Troll	Number of line-days	Total number of line hours/days in the given time period.
Surrounding nets (e.g. purse seines)	Number of sets and Searching time	Number of times the gear has been set or shot, and whether or not successful. This measure is appropriate when school is related to stock abundance or sets are made in a random manner. This represents time on the grounds, less time spent shooting net and retrieving the catch etc. This measure is complicated by the use of aircraft spotting as well as by the dissemination of information from vessel to vessel. Ideally, it should include the area searched as well. The measure is appropriate when a set is only made when a school has been located.
Jigs, (hand and mechanical)	Number of line-days	Total number of line days in the given time period.
Other small scale net gears	Number of operations	Number of fishing operations, whether or not a catch was made. These include push net, scoop net, drive-in net etc.
Other small scale stationary gears	Number of hours fished	Number of hours during which the gears were in the water for fishing, whether or not a catch was made. Those gears include guiding barriers, bag net, stow net, portable net, etc.
Harpoons / spears	Number of days fished	The number of days fishing (24-hour periods, reckoned from midnight to midnight) including days during which searching took place without fishing. If more than one spear-fisher operates from a vessel, the numbers of fishers (spears) need to be recorded as well.

SECOND	PRIORITY	
Fishing Gear	**Effort Measure**	**Definition**
Active gears, such as trawls	Number of sets made	Number of times the gear has been set or shot (either in mid-water or to the bottom), whether or not a catch was made
All gears	Number of days fished	The number of days (24-hour period, reckoned from midnight to midnight) on which any fishing took place. For those fisheries in which searching is a substantial part of the fishing operation, days in which searching but no fishing took place should be included in "days fished".

3.2 Biological Information

Biological information is the composition of the catch, including discards. As information about landings is usually easier to obtain than information about the catch, biological information is often based on landings.

3.2.1 Species List

A species list of the catches has already been compiled (see PROPESCAR-SUR, 1994). This forms the basis for biological data collection.

3.2.2 Species Composition of Commercial Groups

The starting point for sorting catch into species is usually the commercial group. A commercial group may consist of one or more species. Species will be separated only if the buyers of the landings and their customers appreciate species as different commercial products.

For the scientific definition of stocks, species composition is a minimum requirement, and in practice is often all that might be achieved. A fish stock is a sub-set of a species, so before stocks (or management units) can be identified, the species must be identified.

For low value products, the number of species may be very large. A commercial group “Mixed low value small fish” may contain small specimens of the high value species as well. In any case, the enumerators must take samples from some trips (not necessarily all trips) to record the species composition of the commercial groups.

There is no way manpower and funds will allow for all species to be recorded. The sampling programme must prepare a list of the most important species. Other species found in the species composition samples would then appear as “Other”, although recording the number of species (or “recognisable taxonomic units”) in a sample may prove useful in future assessments. The number of species will depend on the ability of technicians to identify species accurately.

3.2.3 Data Requirements for Fish Stock Assessment

Ideally, three principal types of data are needed, broken down by stock:

- Total catches in weight for each commercial group, fleet and period
- Size and species frequencies within fleet commercial groups, based on sampling.
- Total effort by fleet and by period (for example number of fishing days/month).

Usually, it is not a great problem to collect samples of length frequencies. The funds for buying the samples must be available (although the fish can be sold after measurement), or the fishers/buyers may allow measurements without charge. The necessary scientific and technical manpower to collect and analyse the samples must also be available.

Although it may be not possible to get some important types of data, whatever data is available should be obtained for monitoring purposes. For example, if length frequencies cannot be obtained, other types of data, such as catch (in weight) per day, should still be adequate for many types of fisheries assessment.

3.3 Economic Information

Bio-economics is a combination of resource evaluation (i.e. fish stock assessment) and a cost and earnings analysis of the harvesting sector. Ideally, the management and development of fisheries should be based on data and analyses representing all major aspects of the fisheries sector. Thus, biological, economic and sociological information should be collected and processed so the combined output can be used for rational decision-making.

The bio-economic methodology is structured by biological, technical and economic sub-models. The economic sub-model incorporates prices, costs and a range of economic performance criteria. The technical sub-model establishes a physical relationship between fishing effort and fishing mortality, and the biological sub-model describes how the fish population changes. As the exertion of fishing effort incurs costs, and as different species and sizes of fish realise different prices in the market, maximum sustainable yield is not, in most instances, a desirable objective of fishery management from an economic point of view.

3.3.1 Fish Prices

Bio-economics operates with two principal types of prices: ex-vessel prices and wholesale prices. Ex-vessel prices are those received by the fisher/vessel-owner at the landing site. Wholesale prices are those received by the first hand buyer when selling fish to either the domestic or export market. Fish sold for export will earn foreign exchange, which may have additional economic benefit.

Prices are always given by commercial group, and the commercial group is therefore a very important element linking biology to economics.

The value of the landings is often a more useful quantity than landings weight as it takes into account changes in price per kilogram between species groups.

3.3.2 Costs of Fish Harvesting

There are three different types of harvesting costs, namely

(i) Costs depending on number of fishing effort units
(ii) Costs depending on the number of fishing vessels
(ii) Costs depending on the value of the landings (i.e. ex-vessel value)

The first category of costs is often assumed to be proportional to fishing effort, and is also referred to as “variable costs”. This category of costs comprises elements such as fuel and oil, repairs and maintenance, ice, crew wages (independent of yield in value), etc. Information on these costs can be collected routinely on a "per trip" basis.

The costs that depend on the number of vessels refer to costs that arise even if a vessel does not go out fishing. They are also referred to as "fixed costs" and comprise mainly of maintenance and depreciation of the engine and vessel. For many fishers, they may take the form of loan repayments. Routine data on these types of costs are currently not collected.

The third category consists of those costs that depend on the ex-vessel value of the landings. In fishing, crews are often paid a share of the value of the landings. Although the details of the share system may vary from place to place, usually some variable (effort dependent) costs, such as fuel, are deducted from the ex-vessel value prior to sharing the proceeds between the crew and the owner of the fishing vessel. These data are currently collected in Barahona.

The total harvest costs of a particular fleet is obtained as the sum over all three cost categories. Clearly, total harvest costs will not increase linearly with fishing effort because some costs depend on the value of the landings produced by that effort, which changes with the intensity of exploiting the fishery resource. The total costs of the entire harvesting sector are given as the sum of the costs of each fleet.

4. DATA COLLECTION METHODS

4.1 General Principles

It is important to separate scientific data collection from monitoring for enforcement purposes. Where enforcement and science are mixed, the scientific data will be compromised and fishers are less likely to co-operate. However, due to the lack of resources, it is often difficult to separate the two completely.

Data can be collected in more than one way to allow checks on data collection procedures. For example, trip interview data can be compared to logsheets or buyer's forms to find significant differences. However, if there are few errors, this activity may be better directed at collecting new data. Double collection may only be desirable at the start of the programme and thereafter as random infrequent checks on data quality.

4.1.1 The Fishing Trip

An interview at the end of a trip forms the basis for most data collection. In the simplest case, a fishing vessel leaves the port, steams to the fishing grounds, fishes for a while and then steams back to the port where it lands all the catch. Together these events are called a “fishing trip”. The duration of the fishing trip is the time between leaving port and returning to port. This is often referred to as “days away from port” as a 24 hours day is often used as unit for time. For single day or night trips, hours may be a more appropriate measure. The “time fished” (days fishing) is the trip duration minus the time spent steaming to and from the fishing grounds. The duration of the trip and time fished are important variables for many different types of analyses.

The important feature about the “fishing trip” is that it should be possible to allocate catch and fishing effort to a fishing trip unambiguously. In practice, this may not always be easy. For many fishing trips, the allocation of catch and effort is straightforward, but there are also many types of trips that are more complex. Some of the most common deviations include group fishing, where vessels collaborate during the fishing operation, sharing catch and effort, transshipment where vessels offload catch at sea, and occasions when vessels land at multiple landings places. These deviations were not encountered in the Dominican Republic, but may become apparent as the data collection programme expands.

4.1.2 Fishing Vessels

A ”fishing trip” is carried out by one or more ”fishing vessels”. Thus there is a ”one-to-many relationship” between the vessels and the fishing trips (”for one vessel there are many fishing trips”).

The data about the fishing trip is combined with the data about the fishing vessel when the landings data are processed. For example, a set of trips can be grouped according to which fleet the vessels belong to and the total effort estimated for each fleet separately.

The set of data collected from a fishing trip cannot be processed in isolation, but only in combination with a vessel information. If a vessel register is not available, some other set of data, which can partly replace the vessel register, must be made available. In this case, the 1990 census is only data available, although a register is currently being developed.

4.2 Collection Methods

4.2.1 Vessel Registration

The vessel register is important to fisheries statistics. The creation and maintenance of a vessel register has first priority among the different fisheries data. The vessel register should contain the characteristics of each vessel, such as a unique registration code, the dimensions of the hull and the engine etc. These data have a wide variety of uses including bioeconomic and stock assessment analysis and the control of fishing activities.

It is of utmost importance that the vessel registration codes are unique, so no two vessels have the same code. Each administrative centre must have a unique code forming part of the registration code.

4.2.2 Fisheries Census

The vessel codes in each fleet can easily be extracted from a vessel register at any time. If the vessel register is updated with a short time lag, the number of vessels by fleet can be calculated precisely. If a complete vessel register is not available, an alternative way of counting the number of vessels in each fleet is required.

A census or frame survey is an inventory list of fishing units at a specific time, sometimes combined with an indication of their activity levels (e.g. the number of active fishing days per month). A census is usually a complete enumeration as far as the number of units is concerned, whereas the activity data are often sampled.

The frame survey should be updated every fishing season or within some other time unit which forms the basis for raising samples to total landings. A full frame survey can be repeated less often, every five years or so. The local fisheries authorities will usually assist the sampling programme with the counting of vessels in each port.

The inland fishery would benefit from a frame survey. This would improve the design of a sampling programme and estimates of total catch. While considerable information exists on the inland fishery, it is not in the quantitative form of a census. The methods from the 1990 census could be repeated for inland fishery. If full vessel registration will take a long time to complete, an update census of the marine fishery may also be considered.

For the marine fishery, vessels do not migrate seasonally. For the inland fishery they do, which will complicate the sampling procedures. This will require a different sampling programme at different times of year.

4.2.3 The Trip Interview at the Landing Place

To collect information about the fishing-trip the crew of the vessel and/or the buyer are interviewed at the time of landing. The data in question are catch and effort and fishing trip details. Data on vessel details may also be collected if a dependable vessel register is not available.

The different types of data to be collected are indicated in Table 4, which gives a complete set of “trip-data”. The complete set of “trip-data” is, however, not common. Most often only a sub-set of the data will be collected.

The fishers or the buyers will nearly always sort the catch into commercial species and size groups. The species and size groups are determined by the market prices and the appearance (size, quality etc.) of the species in the commercial groups.

Within a commercial species group, the species composition may then be sampled. Sometimes a commercial group (in particular in the case of valuable product for export) will comprise only one species, or even only a size group of one species. Expensive species like groupers are usually gathered into small, medium and large grouper categories, and sometimes by species. However, the most common case is that the commercial group comprises two or more species. If the species composition is required, it is necessary to buy or otherwise obtain a sample, and then separate it into species.

The ultimate level of detail collected is the length frequency samples and/or other biological data. These samples will be the same as the samples taken for species composition, or sub-samples of the species composition samples.

Table 4 The principal data types that should be recorded during a trip interview.

Sample information	Date of sampling Enumerator
Vessel information:	Vessel registration number Gear details
Effort and activity:	Days away from port, days fishing, additional detailed effort data
Spatial information:	Fishing grounds Landing place
Information on total landings:	Weight by commercial group Price per kg by commercial group
Discard information:	Percent discarded
Within commercial group samples:	Sample species composition from selected commercial groups Weight (number) by species
Within species, length frequencies samples:	Sample for length frequencies (and other biological measurements) for selected species Number of specimens by length class

Sometimes the fishers will be able to give information on the amounts discarded. They may in the case of a trawl fishery be able to tell how many hauls were made from which a certain fraction was discarded. They may also be able to tell the composition (species and size groups) of the discards.

While collecting landings data, the enumerator should also collect information on the gear, the effort and the fishing grounds. Two types of data related to effort can be collected:

1. **Effort**: for example, number of fishing hours or traps pulled
2. **Activity level**: Past fishing effort, for example the number of days spent fishing over the previous month.

The most basic effort data you can collect is the number of days away from port. If possible, the number of fishing days will also be recorded. Often this will be the only effort data available. In the case of Barahona, soak time was recorded which is adequate for traps and other passive gears.

4.2.4 Log Books / Sales Slips Filled in by Fishers / Buyers

Logbooks contain the essential fishing trip information, such as start and end date, fishing grounds, gear used, catch and landings etc. Each trip is recorded in the logbook on one or more pages. This sort of scheme is only applicable to larger vessels and processors.

The compulsory logbook / sales slip system, is the ideal system from the data collector's point of view, as it allows for complete enumeration. However, logbooks and sales slips do not automatically imply a very high quality of data. Compulsory filling in of logbooks and sales slips needs to be combined with strict enforcement of fisheries regulations.

If only a proportion of the fleet and processors and buyers fill in these forms voluntarily, the data may be biased. The willingness of fishers and buyers to complete forms is self selecting and not random. To guard against potential bias, at least some data needs to be compulsory and fully enumerated or some forced sampling needs to be conducted. For example, those who do not co-operate could be required to have observers aboard their vessels, or subject to processing plant inspections. However, combined with trip interview data, even a co-operative logbook scheme is likely to be valuable.

4.2.5 Data From Processing Plants

Processing plants may hold data records for monitoring their own commercial activities. These records may also have other uses and be an inexpensive source of information for data collection programmes. In particular, processing plants may grade fish by size, which will be particularly important for shrimp. With co-operation from the processors, this is an inexpensive source of information. Where possible, additional information linking processed landings to individual vessels, and the sex and size composition within categories, should be collected. However, the processor interviewed did not keep records, but appeared willing to provide information if provided with forms.

4.2.6 Data From Other Sources

Other types of data from the fisheries sector, which are not specific to fisheries only, are kept in databases outside the area of responsibility of the Fisheries Department. The types of potential data are very wide, and these information are peripheral to the immediate aims. Nevertheless, it is very valuable to review all data that might be available in monitoring the fishery performance at very little extra cost to the department. The following is an indication of the types of data which may prove useful.

a) **Demographic** data: human parameters of the fishing communities, e.g. statistics and details of fisher families. These data are often available from other departments.
b) **Infra structure** data: roads, housing, transport, ports, processing plants etc. If a Geographical Information System project exists, it is often very easy to incorporate landing sites and mooring facilities to produce maps containing this information. Some GIS facilities do exist in the Dominican Republic, although there is no use by the Fisheries Department.
c) **Institutional and Community** data: laws, regulations, inspection and enforcement of fisheries, education, taxation and subsidies, loan providers. Institutional data outside Government sources may be difficult to obtain. It should form part of the development of the fishing co-operatives to maintain relevant data bases on their own activities.
d) **Market structure**: Market structure has changed with smaller numbers of buyers consolidating purchases for large scale retail outlets. This needs to be monitored so that performance related to value added of the fish products can be considered.
e) **Import / export** data: quantities and qualities of imported/exported marine products. The customs department should monitor exports and imports.
f) **Household** data: this is the sociological information on the fisher families, such as number of family members, their age, education, occupation, income, houses and installations in houses, ethnic background, religion etc. This should be available from national census data, but the Fisheries Department may need to ensure relevant questions are included.
g) **Meteorological and geographical** data: these data may be used for definitions of fishing grounds and seasons and to estimate environmental effects that drive the

stock dynamics or limit fisheries. For inland fisheries and fisheries based around estuaries (e.g. shrimp fisheries) rainfall and river flows may be particularly important.

4.3 Data Collection Forms

Some thought should be given to the design of data entry forms. They should be as simple as possible, but cover a wide variety of vessel and landing types. They should also be flexible, and enumerators should be encouraged to record any extra information they think is important.

In general, the layout and structure of the paper forms should reflect the layout and structure of the database forms, which are used for data entry. This makes data entry much easier and clearer.

Forms often make up a hierarchy. For enumerators recording data at a landing site, for each vessel observed there should always be a trip interview form. Other forms, such as those for recording biological data, may also be filled in. These forms, however, must be linked so that the information can be combined for particular types of analysis.

Forms should be in Spanish, which is clearly understood by the enumerators.

All data forms should have certain attributes in common:

Sample Number: This links forms together and links to any samples collected at that time. So for example, a trip interview form should be linked to any biological sampling forms taken from the same vessel at that time. Sample numbers are unique, and often made up of the date, a code for the enumerator and a sample sequence number.

Date, location of sampling (landing place): This identifies the sampling and is critical for interpreting samples.

Comments: There may occur various unexpected observations, which do fit into the specific fields, and this information may be entered in the “comment-boxes”.

Lists: Whenever possible, the fields are filled in by selecting an item from a list. This is not only clearer, but will link with the database system. For a large number of possibilities, codes may be used to indicate the appropriate choice.

In the following sections, the variables already collected in the Barahona data collection programme are marked with an asterisk (*).

4.3.1 Vessel Activity

The primary aim of the form recording the vessel activity at a site is to ascertain the number of vessels which are active that day (i.e. have gone out fishing) and the number that have not, but have remained in port. This is particularly important if an interview is not made on all trips. Data should include the number of vessels that:

- are based at this site
- out for a one day trip (i.e. must land that day)
- out for more than one day (i.e. may not land that day)
- were sampled (i.e. for which a trip interview form was completed).

4.3.2 Trip

Most of these data are collected in Barahona. For the trip interview form, the basic information should be:

Vessel registration number: Once registration is complete for the fleet, the vessel registration number should be sufficient to get all the vessel details, although the possibility of validating and updating information may be provided.

***Gear**: The form should allow entry of several different gears with their characteristics (e.g. mesh size for nets).

Fishing grounds: Name of fishing grounds as reported by the fishers which might be linked to known areas marked on a map defined in terms of statistical rectangles.

***Landings**: The total quantity of fish being landed in weight by commercial group. Additional information may be required if the catches are processed at sea.

Discards: The total quantity of fish discarded. This is an estimate by the fisher and is likely to be inaccurate for larger quantities of discards. It could be provided as a percentage of the catch.

***Effort**: Effort information requested may depend on the gear used. However, all vessels should be asked for the number of days (or hours) fishing as this forms a shared measure between fleets and gears and is useful for estimating variable costs. Dates and times of departure and arrival at port and fishing grounds are recommended. Not only do they provide this information, but also allow analyses to consider time of day, moon phase and other factors that may affect catches.

4.3.3 Species and Length Composition of a Commercial Group

The sample of the commercial group should reflect the composition. There are a number of ways of doing this, depending upon the logistics of collecting data at the landing site. A common technique is to randomly select boxes of fish within the commercial groups and identify the species and weigh or count them. From within the main species catch, a random sample is taken to obtain length, sex and maturity. The important issue is to avoid non-random selection that will lead to biases in species or size composition information (for example, selecting larger specimens only). Where the catch is small, all the landings should be included (i.e. complete enumeration).

Number of species: The total number of species identified within the commercial group.

Sample weight: The total weight of the all specimens identified of each species.

Then of the sample, for each fish measured the following should be recorded:

***Species name**: The “Species name” should be a code selected from a look-up table, which reflects a full taxonomic (scientific) classification of the species.

***Sex and maturity**: For some species, this is very difficult without considerable laboratory work. Other species are simple to measure. Where the cost of measurement is high, only a small sample should be taken and returned to the laboratory for inspection. This sample must be random.

***Length and Weight**: Standard length measurements are usually straightforward to obtain. These should be standardised by species (e.g. fork length for finfish) and use consistent units. Some additional non-routine sampling may be useful to build morphometric models to convert between measures (e.g. length weight relationship).

4.3.4 Variable Costs & Earnings Per Trip

These data are closely related to the other trip interview data, but they are largely used for a different purpose.

A time budget: The time of leaving port and arriving at the fishing grounds, and the time leaving the fishing grounds and arrival at port, and other times steaming between fishing grounds. These data are often collected as part of the effort data.

Expected trip duration: This can be compared to the actual trip duration and deviations may indicate unfavourable conditions for fishing.

Ice used: The quantity of ice used during the trip.

Target species group: The expected main catch composition, species (size group of a species) or species group.

Price by commercial group: The price received for the landings broken down. Fish that is not sold should also be recorded. For example, some low value fish may be retained by the crew for their own families.

Buyer: The name of the buyer (merchant, processing plant, exporter etc.)

Damage/loss (estimate): For example, estimated value of lost or damaged gear, damage of engine etc.

Some variable costs depend on the effort (time fishing). Other trip related variable costs depend on the weight or value of the landings.

Ice: Cost of ice used during the fishing trip (e.g. amount used and price per litre).

Fuel and Lubrication: Cost of fuel (oil or gasoline) and lubrication used during the fishing trip (e.g. amount used and price per litre).

Water: Cost of water used during the fishing trip (e.g. amount used and price per litre).

Food: Cost of food consumed by the crew during the fishing trip.

Bait: Cost of bait used during the fishing trip (e.g. amount used and price per kilo).

In many cases, the crew will receive a share of the value of the landings. The method of allocation and amount allocated to each share should be recorded.

***Total value of landings minus cost deductions**

***Divisible earnings**: The amount to be shared by crew and vessel.

***Shares**: The percentage or amounts of the divisible earnings allocated to each stakeholder. In general, the share is split between the vessel owner, the skipper and the crew.

4.3.5 Vessel Registration

The vessel register should be considered a part of the fisheries statistics, and often it forms a central part component on which most other information depends. The basic data are fishing trips carried out by a vessel. The identification and details of each vessel are stored in the vessel register.

***Owners name and address**: Full name and address of the primary owner.

***Home port**: Port of registration.

Association: Membership of fishers co-operative, if applicable.

Credit scheme: Financing body (moneylender) of the vessel.

***Vessel registration number**: The unique vessel registration code. No two vessels should have the same vessel registration code.

Radio call sign: Radio call sign used to uniquely identify a vessel.

***Name of vessel**: The name of vessel painted on the vessel.

***Type of registration**: Fishing vessel (full time/part time/foreign/recreational etc.)

Main fishing grounds and species: The intended main fishing grounds and target species.

Main Gear: Types of gear fitted or available.

Construction material: Wood, steel, glass fibre, cement etc.

Year vessel built: The year the vessel was delivered from the shipyard.

First year of fishing: The year the vessel started to fish and land the catch.

***Propulsion data**: Method of propulsion, information on all engines, such as engine make, age, and horsepower

Crew number: The average number of crew on an average fishing trip.

***Registered length for regulation purposes (m), Maximum length (m), Overall length (m), Breadth (m), Depth (m), Gross tonnage**, **Net tonnage, Hold capacity**

Remarks by the Registration Authority: Any additional information relevant for the vessel registration.

Additional linking information may be required if the vessel changes hands. In this case, the vessel must be re-registered.

Preceding vessel registration number: The former vessel registration code/number and name of vessel as given in the registration box.

Date of change of ownership and new registration: The dates when the vessel changed hands.

The vessel at some point will be scrapped. It is very important that the register maintains a list of active vessels and that this list is kept up to date. In addition, the effective life of a vessel is valuable for calculating depreciation.

Date and reason for termination: The reason may be scrapping of vessel due to old age, decommission, capsize, sale of vessel for non-fishing purpose, etc.

Economic data may be collected at registration. However, much of these data may be sensitive and may require a separate sampling programme to obtain an unbiased estimate.

Insurance: larger vessels might be insured against loss or damage

Loan structure and repayments: Loans for the purchase of vessel, engine and gear may be structure in different ways, but include capital repayment and interest.

Vessel, engine and gear purchase cost: Vessel and/or gear may be purchased outright.

Fixed refit and maintenance costs: The vessel, engine and gear may regular servicing or replacing even if not used.

4.4 Codes

Despite the development of computers, codes are still useful for recording information. There are a number of reasons for this. Codes act as shorthand replacing complex descriptions, and can speed up data entry once enumerators become familiar with them. They are also often more stable than names and descriptions. For example, scientific names for species may be subject to changes by taxonomists. Codes also avoid linguistic problems and aid international data exchange. They also lend themselves to the use of look-up tables in databases, which prevents many data-entry errors.

National codes may be required for the biological species, commercial group, landing site and so on. A number code already exists for the landings sites. Where possible, codes should be clear in their meaning. For example, the vessel registration code could start with two letters indicating the administrative centre where the vessel was registered (e.g. SA for Samána).

Commercial groups may be very different to taxonomic groups and coding commercial groups may therefore prove useful. In classifying groups, three aspects may need to be considered, the type (species), size and treatment or quality. The code may include quality grade (e.g. freshness) and processing (e.g. fresh, iced, salted, dried and “for export”).

A code system used for international exchange of data needs to meet standards different from those of a national code system. This may need to be considered if a regional database system is established. International codes for species and gear types exist, and may avoid the need to develop local codes. Often, however, due to the complexities of fisheries and ecosystems, such codes are never quite appropriate and a national code system is deemed more appropriate.

The CFRAMP database system, TIP/LRS, already contains coding consistent with the systems used in the USA. Where relevant, it is recommended these are used unless they cannot fit local data.

4.5 Recommendations

Data collection should centre around trip interviews. This is already done in Barahona. It is important that a form is filled for each interview. Filling out forms based on visits to sites as daily totals or other information requiring processing by the enumerator is not recommended. Only raw data should be recorded, that is the answers to specific questions posed at the interview.

Of those vessels for which there is a trip interview, a random sample should be subject to biological sampling of their landings. It is important that the trip interview and biological sample are linked in the database.

For larger vessels, a logsheet programme should be implemented, where vessels are required to complete and return logsheets at the end of each trip, usually for trips longer than one day. The logsheets should contain daily records of the vessels activity and catch. However, this requires co-operation from the vessel captains.

Buyers and processors should be required to complete purchase forms. These would cover a significant proportion of the catches and be particularly useful for bioeconomic analyses. Based on these data alone, it will be possible to assess shrimp stocks as shrimp will probably be sorted into size classes. Biological sampling at the processing facilities, again particularly for shrimp, would provide valuable information for the assessment of the stocks. However, because purchases may be selective, the trip interviews and sampling landings has higher priority.

5. SAMPLING

5.1 Overview

It is important that a standard random sampling scheme is adopted. Without random sampling, there is no standard objective way to interpret the information collected. This can lead to very poor estimates of the state of the fishery and hence poor decision making. It is important that all staff in the data collection system understand the principles of sampling and the way it is done.

Although sampling is often thought of in terms of vessels, the reality of most sampling schemes is that they revolve around landing sites. For example, a particular region may possess 20 sites where fish are landed, but only 4 enumerators. An enumerator can only visit one site on any one day, so enumerators should be allocated randomly to the sites. They then record either all landings or a sample of landings at that site on that day.

5.2 Three Stage Design

If sampling is carried on at different levels, they are referred to as stages. For example, landing sites may be sampled and then vessels landing at each site may also be sampled. The Dominican Republic fishery requires three stages.

Stage 1: Landing Site

In many cases, this will be a visit of an enumerator to one landings site each day. If there are as many landing sites as enumerators, this stage will be fully covered and its contribution to the variance of estimates will be zero. Otherwise, on each day, a random site is chosen and allocated to each enumerator. Where there are a large number of landings, more than one enumerator should be allocated to monitor that site.

Stage 2: Vessel landings

If an enumerator stays all day and vessels land singly, then an enumerator may be able to cover all vessels, recording their catch and effort. However, if vessels land in groups or the enumerator is only able to stay for part of the day, a number of landings may be missed. A significant problem is that it may not be feasible to obtain a clear random sample, and any estimates assuming the sample is random may be biased.

Stage 3: Biological Sampling

Biological sampling includes identifying species, and measuring size, sex and maturity of a random sample from the catch. This is a time consuming operation and therefore would require an additional team, including a biologist, besides the main enumerator. Of the sample for which trip interviews have been conducted, a smaller set of random trips are selected and measurements are taken from the catch. The catch itself may be sampled if it is large. Such sampling is best conducted by stratifying by commercial group. Again, obtaining random samples may be difficult due to logistics, but every attempt should be made to take individual fish at random.

Note that sampling is not necessary at each stage on every occasion. In many circumstances, total enumeration may be possible and is often desirable. Particularly if there is no extra cost in total enumeration, there is no point in sampling. For some sites, it is recommended that enumerators are permanently based there as the number of active vessels is very large. If the enumerator visits a site where only a few vessels land, the enumerator should monitor all landings.

5.3 Sampling Landing Sites

In reality, landings sites are the only units that can be sampled with genuine randomness. Random sampling guards against any bias and ensures estimates represent true landings and effort. Therefore, careful consideration should be given to how these sites are sampled.

5.3.1 Stratifying Landings Sites

Landing sites need to be grouped both for logistical and administrative reasons, and to increase estimate accuracy. The administrative structure is set by government and will not improve sampling estimates. However, for sampling purposes the administrative regions can be split into smaller groups to improve estimates.

The current administration aims to split data collection among 16 stations or administrative centres. I have allocated sites identified in the 1990 census among these stations based on proximity and split landing sites in contiguous areas where appropriate. This allocation is given in the accompanying spreadsheet "DRSamplingAllocation.xls". Sites could be re-allocated easily as the administration sees fit, as the stratification is not organised for sampling purposes.

It is possible to work out optimal allocation, where the sampling effort is distributed according the total number of units (active vessels: see Table 6), the sample unit variance (gear variety: see Table 7) and the cost of sampling (cost per vessel; see Table 8). Three schemes are given as real data on variance and costs are unavailable, so proxies are used instead (see table captions).

These allocation schemes can be compared with Table 5, which gives the actual allocation. Considerable numbers of inspectors appear to have been allocated to the inland fisheries. Without a census of the inland fisheries, it is difficult to comment on the distribution of inspectors, but it does not appear to be compatible with the distribution of marine fishing activity.

Table 5 Current inspector allocation amongst administrative stations. Inspectors are responsible, among other things, for data collection in both marine and inland fisheries. However, it is noticeable that Samaná appears to have few inspectors despite the fishing activity in the area reported in the 1990 census.

Station	Registered Inspectors	% Inspectors
Montecristi	5	1.77%
Santiago (Puerto Plata)	42	14.89%
Rio San Juan	7	2.48%
Hatillo	55	19.50%
Samaná	8	2.83%
Miches	3	1.06%
San Pedro (Romana)	33	11.70%
Santo Domingo	49	17.38%
Bani	58	20.57%
Barahona	5	1.77%
San Juan	17	6.03%
Total	282	

Table 6 Recommended strata and enumerator allocation based on number of active boats recorded in the 1990 census (see Thompson 1992, p107).

District		Landing Sites	Active Boats	Proportion of Sampling
1	Montecristi	8	191	5.87%
2A	Puerto Plata	14	118	3.63%
2B		1	198	6.09%
3&4		17	197	6.06%
5A	Samana	20	234	7.19%
5B		16	378	11.62%
5C		1	315	9.68%
6&7	Miches	10	176	5.41%
6A		1	168	5.16%
8&9	Romana	16	313	9.62%
10&11&12	Santo Domingo	16	339	10.42%
13&14	Bani	15	272	8.36%
15	Barahona	13	169	5.20%
16		12	185	5.69%
Totals		160	3 253	

Table 7 Optimal allocation of sampling effort (i.e. enumerators) taking into account gear variety. The gear variety is measured as the multinomial variance of the gear composition given in the 1990 census. It is only an indicator of possible variation in catch rates due to gear variety. The score is calculated from the standard deviation and number of active boats (see Thompson 1992, p107).

District		Standard Deviation	Score	Proportion of Sampling
1	Montecristi	45.96	8 779	5.17%
2A	Puerto Plata	59.12	6 976	4.11%
2B		15.37	3 044	1.79%
3&4		64.90	12 785	7.53%
5A	Samana	69.96	16 370	9.64%
5B		63.56	24 025	14.15%
5C		21.74	6 848	4.03%
6&7	Miches	49.65	8 738	5.15%
6A		19.54	3 283	1.93%
8&9	Romana	64.86	20 303	11.96%
10&11&12	Santo Domingo	64.95	22 019	12.97%
13&14	Bani	61.34	16 685	9.83%
15	Barahona	57.65	9 742	5.74%
16		55.02	10 179	6.00%

The results including costs and CPUE variance seem to give a reasonable allocation (Table 8), not overly favouring sites with many vessels, but clearly indicates sampling effort should concentrate on these areas. Strata 2B (El Muelle, Puerto Plata), 5C (Sanchez, En El Pueblo) and 6A (Playa Arriba Y Abajo) each are allocation of permanent enumerators to single sites. These sites apparently possess a large number of active vessels, so that sampling should always take place at these locations.

Table 8 Optimal allocation of sampling effort taking into account costs. The cost calculation assumes the cost of one day's sampling is the same for each site. The variation costs is the number of observations that can be made. This depends on the number of active boats at a site, with fewer boats making the visit more costly per observation. The maximum number observations is capped arbitrarily at 30, representing the maximum coverage a single enumerator can undertake.

District		Standard Deviation	Score	Proportion of Sampling
1	Montecristi	0.12	25 486	5.11%
2A	Puerto Plata	0.22	14 792	2.96%
2B		0.05	13 613	2.73%
3&4		0.17	30 595	6.13%
5A	Samana	0.14	44 209	8.86%
5B		0.08	87 044	17.44%
5C		0.05	30 627	6.14%
6&7	Miches	0.26	17 051	3.42%
6A		0.05	14 684	2.94%
8&9	Romana	0.25	40 668	8.15%
10&11&12	Santo Domingo	0.08	78 147	15.66%
13&14	Bani	0.11	50 956	10.21%
15	Barahona	0.19	22 098	4.43%
16		0.12	29 096	5.83%

5.3.2 Within Strata Sampling

Where there are fixed numbers of data collectors at a site, the site should have observations for every day. The strata in this category are fully enumerated and there is no sampling.

The usual approach to estimating total catches (and other variables) from sampled data is to raise the sample using variables that have been accurately censused. For example, with a vessel register we may multiply the average landings per vessel by the number of vessels at sites not sampled to obtain their expected landings. This requires accurate census data, which is currently not available for the Dominican Republic. The best information currently available is the 1990 census, which is already over 10 years old and is probably not accurate.

Although strata will account for some variance in the data, there will be considerable variation within strata between sites, which must be covered by random sampling. There are two options:

Simple random sampling without replacement. This is simple to organise and implement. The estimates of the mean and variance are simple to calculate. On any day, all sites have an equal chance of selected, and available enumerators are allocated each to one site on each day.

Unequal probability sampling without replacement. In the same way we can allocate greater sampling effort to the most important sites by making them more likely to be selected in the random sampling process. On any day, sites have a chance of being selected which is proportional to their importance, and available enumerators are allocated at random each to one site on each day. Although the method is a much more efficient design to simple random sampling (Table 9), estimates are complicated to calculate and impossible for large strata. Software has been provided for the strata defined (see "DRSamplingEstimation.xls").

Table 9 Simulated sampling of eight sites around Montecristi illustrating the superiority of the unequal probability design. Both the estimated standard deviation (SD) and the Root Mean Square Error (Root MSE) are lower for the unequal probability design for exactly the same sampling effort. This method would result in greater confidence in analysis of the data and hence probably allow higher levels of harvesting under the precautionary approach.

Summary Table

	Total	**Variance**	**SD**	**MSE**	**Root MSE[1]**
Unequal Probability	5896	8352	91	11605	108
Random	5908	155373	394	165641	407

[1]Root MSE is the square root of the sum of squared differences between the estimated and true value of the variable being measured. It provides an absolute measure of the performance of the sampling method with a lower value

Both sampling designs are described. Given the much better performance of the unequal probability design, it is the preferred option. However, the method is more complicated, particularly the estimation, and although software has been provided, some training would be necessary in its use. Despite this complexity, it is still recommended that the unequal probability design is used as it makes better use of the sampling effort. In these data collection programmes, most of the work (and cost) goes into collecting the data, while the estimation, with suitable software, takes little time and effort. On the other hand, more important than efficiency is the sustainability of the data collection programme. If a more complex design would jeopardise this, then the simple random sampling approach should be adopted.

Table 10 Summary of the strengths and weaknesses of the two within-strata sampling approaches.

	Strength	**Weakness**
Unequal Probability	Much more efficient for a fixed sampling effort	Calculations are more complex
	Makes intuitive sense to spend more time at most important sites.	Probabilities need to be proportional to variable of interest.
	Sampling and estimation can be automated using the computer	
Simple Random Sampling	Simple to implement and estimate values of interest	Relatively inefficient, so it is effectively more costly

5.3.3 Sampling Logistics

Often logistics and random sampling designs are in conflict. For example, although ideally fish should be selected at random within each commercial class so that the sample will represent the size and species composition, in practice this is not possible. What may be possible is to select boxes in a random sequence or use some other system to avoid bias. For example, count a random number (between 0 and 9) of boxes being removed from the fish hold before selecting a box to measure its contents, thereby measuring approximately 10% of the catch. However, even this might not be possible, depending on how vessels are offloaded of their catch. The only way to develop a practical approach is through negotiation between enumerators and statisticians. Through discussions, both should appreciate the others point of view and a compromise should be reached.

It is usual for sampling programmes to undergo revision as always problems will be encountered and need to be resolved after the programme has started. Typically, these are related to the practicalities of data collection.

5.4 Recommendations

It is recommended that the stratification including costs and gear variance is used in the first instance (Table 8). Once two years data has been obtained, the situation should be reviewed and optimal allocation recalculated based on real observed variance and costs. Enumerators can then be redistributed to optimise data collection subject to logistical constraints.

The unequal probability design should used to sample landing sites. This will allow a robust estimates of total landings and average CPUE without a vessel register or up-to-date frame survey. It will be possible to incorporate auxiliary information from the vessel register and elsewhere in the future.

6. ESTIMATION

6.1 Overview

This technical section is provided for completeness. Estimation techniques are dealt with in detail in Thompson (1992).

Data can be provided in sufficient form for stock assessment. Often this means an estimate and its variance rather than raw data. However, in some cases a stock assessment model can be fitted to raw data, so raw data should be available for an assessment if required.

Performance indicators recommended for most fisheries in the Dominican Republic are simply estimates of key variables such as CPUE, catch and profit. These are simple to calculate, but the statistical estimation may be more complicated. Some important estimators are provided.

In general, estimates for the different sampling stages can be scaled upwards directly. This allows us to consider each stage separately in terms of the estimate and its variance. The results from each stage feed up to higher levels as appropriate.

6.2 Estimating Total Landings

6.2.1 Simple Random Sampling

To estimate totals, at least one variable must cover the whole population. For example, a vessel register maintains the total number of vessels. The total catch of a sample of vessels can be raised to the entire fleet based on the number of vessels in the fleet. Furthermore, the register can be used to stratify the sample, and vessel attributes can be used as covariates, both of which will improve the estimate.

This requires two stage sampling by site and by vessel landing. Assuming random sampling without replacement, the estimate of the total catch at a particular site.

$$C_i = \frac{M_i}{m_i}\sum_{j=1}^{m_i} C_{ij}$$

$$C = \frac{N}{n}\sum_{i=1}^{n} C_i$$

$$\mathrm{var}(C) = \frac{N(N-n)}{n(N-1)}\sum_{i=1}^{n}\left(C_i - C/N\right)^2 + \frac{N}{n}\sum_{i=1}^{n}\frac{M_i(M_i - m_i)}{m_i(m_i - 1)}\sum_{j=1}^{m_i}\left(C_{ij} - C_i/m_i\right)^2$$

In the simplest case, all landings at a particular site are recorded, making the total catch estimate exact and eliminating the second term of the variance estimate.

The small sample size makes stratification within an area difficult. If there are 4 enumerators and 20 sites, stratification would have to interfere with the allocation over a number of days to be applicable. For example, combining over a month would mean that we could allocate 120 visits. Sites could be split into three strata based on the numbers of vessels landing at each site, and visits allocated accordingly. As long as visits are randomly distributed amongst days and sites within strata, estimates should be unbiased. Although more efficient than simple random sampling, the approach has problems. Visits cannot be randomly assigned over days and sites as the number of visits per day is constrained by the number of enumerators, so allocation cannot be truly random. Any constraints or time series patterns in the data may lead to bias. For example, fewer enumerators may be available at weekends and there may a distinct weekly pattern in landings.

6.2.2 Unequal Probability Design

The same multistage estimate for the variance applies. However, the estimate for the primary stage, landing sites, is altered to account for the unequal probabilities. In general, the Horvitz-Thompson estimator is appropriate for all unequal probability designs (Thompson 1992):

$$C = \sum_{i=1}^{v}\frac{y_i}{\pi_i}$$

where C is the total (e.g. catch) to be estimated from a sample of v observations, y_i, each with a probability of π_i of being included in the sample. Notice that where the probability of a site being included in a sample is proportional to size of the catch, the calculation is similar to using raising factors. However, in contrast to raising factors, the design is more robust as departures from this assumption do not automatically lead to biased results and the method is self-correcting with adequate sample sizes.

The most robust variance estimator for this stage is positively biased, but always positive (the unbiased estimator can be negative) and easy to calculate:

$$\mathrm{var}(C)=\left(\frac{N-v}{N}\right)\frac{s_t^2}{v}$$

where N is the number of items in the population, and s_t^2 as sum of squares statistic calculated as:

$$s_t^2=\frac{1}{\sum v_j-1}\sum_{j=1}^{V}\sum_{i=1}^{v_j}\left(\frac{v_j y_i}{\pi_i}-C_j\right)^2$$

where v_j is the number of samples taken within a particular unit. Note that this assumes the variance remains constant over samples. This is useful for the sampling proposed here, as the number of samples taken on any particular day is small. Repeat samples may be combined over each month, and the variance estimated from the month combined which should lead to an accurate estimate. In simulations, this variance estimate was found to be higher than the MSE, and a poorer estimate than that derived from the inclusion probability covariance matrix (the unbiased estimate). However, the latter is much easier to calculate and hence recommended for initial use. Further work could be used to compare these approaches to test their behaviour, using for example, bootstrapping.

The method is straightforward once inclusion probabilities are available. If the probabilities of selection vary for every possible item, the probability for each sample permutation will be different. For a large population of items, the direct estimation becomes impossible (the number of calculations is dependent on the factorial of the number of items in the population). Software was developed to calculate these probabilities using two techniques to reduce the problem:

- A dynamic programming technique reduces the problem from one of permutations to one of combinations, so the number of calculations is reduced from $N!$ to $N!/(r!*(N-r)!)$ where N is the population of items and r is the sample size. It ignores the order in which sample items are drawn which is not relevant for calculating inclusion probabilities.
- Groups of items with the same probabilities are treated together. They will share the same probabilities, although individual items will still need to be taken into account in terms of their removal as they appear in samples. In general, when dealing with groups, only the number of group items in the sample needs to be accounted for, significantly reducing the number of calculations.

The limit on the number of items is around 20. Above this, the calculation becomes very slow unless a significant number of items share the same probabilities, and results may be inaccurate. Under these circumstances, an alternative design may be better (Thompson 1992). This problem has been addressed here by stratifying in to smaller groups based on geographical location consistent with the way the data collection system will be administered.

The software is supplied in the dynamic link library "UnequalSampling2.DLL". This file must be in the current path of current directory for the unequal probability calculation software to work in the appropriate EXCEL spreadsheets.

6.3 Landings Composition

The landings composition is based on the total landings by commercial group and uses the same estimate. However, the species composition within the commercial group should be based on random sampling of individuals. This makes the estimated composition dependent on multinomial distribution (see Thompson 1992). The expected landings of a particular species is the proportion of individuals in the sample which belong to that species multiplied by the total catch. The same estimation technique applies to catch in size groups. Note that the larger the number of categories, the larger the sample size should be to get adequate numbers of fish within each category. For this reason, raw lengths should be recorded so that different size categories can be tested.

6.4 CPUE

Unlike total catch, CPUE is an average value. The average CPUE for one day, estimated across sites, is best estimated as the ratio between the estimate of the total catch and total effort. Total effort can be estimated in the same way as the total catch in an unequal probability design.

Auxiliary data, such as gear type and vessel characteristics, could be used to improve estimates of CPUE in particular. Although the 1990 census information is available, it is very old and therefore it is not recommended that it is used unless the information can be verified.

6.5 Other Indices

Other indices can be estimated using the same methods. All indices are totals or averages, so they use the same principles of estimation. The database should contain reports that calculate the estimates and their variance automatically.

7. DATABASES AND DATA MANAGEMENT

7.1 Overview

Currently the Fisheries Department has no database. A priority is to develop a database suitable to the data collection programme, and the uses of the data.

Given what technology promises, it is usually tempting to go for complex solutions. For example, databases connected to the World-Wide-Web would be ideal. Unfortunately this requires skills that are in high demand and not currently available in the Fisheries Department. It would therefore be costly to pursue high-technology solutions immediately. An straightforward single PC based database is recommended initially, with options to upgrade in the future.

7.2 Database Design

A database should be used to store the raw data. All processing of data should be executed by the database system. If processed data are entered, as incurring the greater cost of pre-processing, the validation of data becomes difficult and some types of analyses may become impossible.

It is important to develop the database simultaneously with the development of the data collection programme, as the database acts as a check on the consistency of the data collection programme. The database will show if the programme actually produces the expected output. The database will also help in the search for the best stratification.

A “database” consists of “Tables”, “Forms” and “Reports”. These three main components are linked by aid of “Relations” and “Queries”. Below follows a very brief introduction to these basic concepts. For a more comprehensive introduction, the reader is referred to the textbooks on databases.

The “tables” are the containers of the data, the “forms” are the tools used to interact with the database and the “Reports” are the output produced by the database.

7.2.1 Forms

The “forms” are the tools for all communication with the fisheries database. The “soft-forms” are the screen pictures through which you communicate with the database.

It is important that the forms are designed so that the data entry is “friendly” to the encoder. The paper interview forms should not be too different from the “soft form”. The encoders with their accumulated experience should contribute to the design of the database forms.

In the design of forms, as much as possible should be included in each screen, without making the form too complex. Furthermore, look-up tables should be used and checks should be made on field entry, to ensure data is as error free as possible. For example, the encoder should not enter the name of the province from the keyboard, but should select it from a look-up table.

The “copy record” allows the encoder to change the data that differs between records, rather than type the record in again each time. Many of the data entered are the same for subsets of records. For example, the name of the landing place remains the same, for all interviews from that landing place.

Encoders should enter all relevant information with respect to a trip at the same time. This is how relationships between data records are formed.

7.2.2 Reports

There are three types of reports needed from a database.

- Administrative reports monitoring data collection activities, for example giving amount of data collected from each site in each month.
- Assessment reports, containing processed data for analysis in a stock or bioeconomic assessment.
- Annual reports producing standard statistics used to monitoring the performance of the sector.
- Sector profiles, giving an overall description of the fisheries sector, which may need to be infrequently updated. The annual reports and sector profile should complement each other.

A database system should include a number of pre-prepared reports covering administrative and annual reporting requirements. Annual reports will require a certain amount of interpretation on behalf of managers, but largely be standardised output.

Assessment reports are likely to need specialised queries and, with a exceptions, may need to be extracted by the scientist doing the assessment.

It is useful if the annual reports do not change too much between years, in order to maintain the compatibility between years to allow easy comparison. Once officially reviewed and accepted, annual reports should be archived for easy access and distributed to interested clients.

A fisheries sector profile may be considered an extended census, where it has been attempted to describe all elements of the fisheries sector of an administrative division or the entire country, depending on the size of country.

7.2.3 Software

The database software should be able to deal with large amounts of data. Even smaller fisheries build up large numbers of records over time, and accessing these records can be slow with less expensive software.

Users who want to apply the data for non-standard (e.g. research) purposes is not satisfied with the pre-prepared reports. They will need the data structured a specific way, to allow for some specific processing. The best way to approach this is to train users in Structured Query Language (SQL), which is implemented as standard in many database software packages. However, they implement SQL to different degrees, and it should be ensured that the "SELECT" queries in particular are well catered for before purchasing a package.

7.2.4 Data Validation, Errors and Backup

Raw data entered on the computer needs to be checked and validated in as many different ways as possible.

Some data should be validated automatically at the time they are entered in the database. For example, the order of magnitude can be checked and rejected if a value has been given in kilograms where it should have been tons. In addition, dates can be validated. For example, lower and upper limits for acceptable dates can be specified, so that the database will give a warning when a date is outside the limits. These are examples of the easy-to-detect errors.

Data selected from look-up-tables will contain only values from the look-up-table, so the only possibility for error is that the wrong value is selected.

There are usually a number of consistency checks, which a database can perform automatically. For example, the date for leaving and returning to harbour for different trips of the same vessel can be compared.

Statistical verification may be used to identify unlikely values. If a commercial group appears in a sample, but not in the other similar samples, it may be an error and can be checked against the original record. However, if the observation can not be ruled out through investigation, it must be accepted even if it appears unlikely. Most analyses are robust enough to allow for a small number of such errors without giving incorrect results.

In other cases, two independent sources of information can be obtained. For example, a logbook may report the landings and these can be compared to the buyers reports of their purchases from the vessel. Data from the commercial fishery may be compared with data from experimental fishery, which for example, may reveal erroneous species identification.

Supervision and on-the-job-training of enumerators is a kind of double sampling, as the supervisor will check any data collected by the enumerator, and they will (hopefully) sort out any discrepancy between them. Supervision and discussion between programme-staff is probably the most efficient tool for data validation.

The processed data should be evaluated by comparison with general knowledge and common sense. All computed total catches should be evaluated for their reasonability.

Finally, the only true protection against computer failures or data management errors or natural disasters is to backup data onto robust media (e.g. CD-ROM). Regular backup is probably the most important data management task.

7.2.5 Table Structure

This will not necessarily reflect the data forms or the way the data are collected. Relational databases have their own rules of data structure. Although the basic data is recorded in tables, it is in the relationship between tables that much of the information lies. It is important that the database is well constructed, and that data entry procedures are strictly adhered to. This ensures that the integrity of the data and these relationships are maintained.

This integrity is often the reason for hierarchical nature of data entry and the enforced use of look-up tables. In addition, some information is entered by the computer, and the encoder cannot change it. Hence, the construction and maintenance of the database is not an easy task, and will require expertise and consultation.

7.3 Recommendations

The Fisheries Department should consider using the TIP/LRS system sponsored by CFRAMP. To design and maintain a good database is costly and often beyond the technical capability of the available staff. A regional database, maintained by a number of countries, should reduce costs and allow access to technical support. However, the database must either be flexible or suit the needs of the Dominican Republic fisheries adequately to serve their purpose.

A simple database for Barahona for data entry should be constructed quickly to allow better use of their data. It should be upgraded to the new system once a decision has been made.

8. IMPLEMENTATION OF SAMPLING PROGRAMME

8.1 Pilot Sampling Programme

It is normal to recommend a pilot sampling programme before setting up a full scale system. In this case, Barahona represents a fully operational programme. The main aim should be to share the experience of operating this system with other administrators.

However, there are a number of recommendations to improve the Barahona system. Ideally, these should be implemented before setting up the systems in other districts. However, managers of data collection in other areas should be involved in Barahona system so that they can learn how the system operates.

8.2 Budget

This section lists some possible items for considerations when preparing a budget of a data collection programme. It is not possible to prepare a generic budget for a data collection programme, due to the variability of local conditions.

The budget should be carefully prepared particularly for recurrent costs which will determine whether a programme will be sustainable or not. Recurrent costs will include salaries, maintenance, running costs and renewal of equipment (e.g. motorcycles), communication costs, regular training and workshop costs

8.3 Training

Training is one of the most crucial components in the preparation and successful implementation of a data collection programme and must always be given high priority.

Training during the planning and implementation phases is different from the routine training of an on-going data collection programme. Regular training is an indispensable part of a sampling programme. New staff members on each level obviously have a need for training, but all personnel on all levels need regular training to maintain work quality.

The first training programme should train trainers so that the programme is supported internally and therefore sustainable. Ultimately participants should include fishers, data collectors, supervisors, researchers, computer operators, directly responsible data producers and users, and decision-makers. Training workshops should be regular. This will allow the programme to improve, adapt and maintain a freshness despite its routine nature.

Training courses and workshops are a good means for addressing methodological and operational problems encountered in the implementation of data collection programmes. They provide the opportunity for bringing together staff with different responsibilities and activities, such as data collectors and supervisors, data operators, statisticians and researchers.

The training courses and workshops should address all aspects of the sampling programme, and it is important that staff on all levels understand and support the overall idea behind the programme. Enumerators, for example, are often in a situation where they are questioned by fishers about the data collection programme and they should preferably agree with the programme and be able to explain and justify it.

As well as specific skills training, the key questions should be explained and discussed with all types of staff, although the topics should be dealt with in a way matching the different staff groups. Key questions are:

1. Why collect data? (objectives of data collection programme)
2. Who are the clients? (who should we collect data for)
3. Which data to collect? (say, from commercial fishery or from experimental fishery?)
4. How to collect data? (What are the sources of data?)
5. How to store and process data? (Including databases, bio-economics and fish stock assessment).
6. How to report on fisheries? Which groups of clients should get which type of reports?
7. What are the costs of data collection, and where does the funding comes from?

Training high level decision-makers should not be neglected. While they will have little time available, introduction to the concepts particularly of reference points, performance indices and the objectives of data collection. The decision-makers, such as government ministers, need to understand how data is processed into information that they can use. The choice of performance indices and reference points is clearly a policy choice.

Information is the prerequisite for rational management and development, but that information costs money, as explained in various international agreements, such as the ”Code of Conduct for Responsible Fishing”. Decision-makers' appreciation of the value of these data will help considerably in dealing with problems (usually funding) that arise.

It is important that all staff have a good understanding of other staff’s work. For example, encoders and enumerators should be familiar with the other's activities and have experience in each other's work.

8.4 Consultation

The programme should prepare material to inform the industry and the local administration about the sampling programme. A pamphlet of a maximum of 20 pages with many illustrations, aiming at the public as a whole, should be prepared. The data collection programme may get assistance from professional information experts (e.g. journalists or artists) to formulate the text and the illustrations of the pamphlet to reach a wide audience, but with special emphasis on the target group, the fishing community. The programme should design a logo to make the programme easily recognisable and visible and may prepare other material, like a presentation, T-shirts and note books with the logo of the programme etc. Funds must be allocated to this type of activity, as a bad relationship with the industry may destroy an otherwise well-designed data collection programme.

The fisheries administrators are the primary users of the data collection programme. There must be a good relationship between the fisheries administration and the data collection programme. Therefore, regular meetings and consultations should be arranged with them.

The main purposes of consultations are:

1. To establish good relationships among stakeholders
2. To establish agreements on exchange of data
3. To get advice and information
4. To inform about the data collection programme (objectives and methodology)

In addition to stakeholder consultations, technical issues can benefit from regular expert workshops. Standing committees and working groups on fishery statistics (i.e. stock assessment, statistic standardisation groups) could be set up and play a key role in the co-ordination of data collection programmes, particularly in those cases where different agencies or institutions are involved in various applications sectors and components of an overall survey system. Statistical Committees should meet on a regular basis and their composition and level of authority should allow submission of their recommendations to higher government authorities for consideration and action.

8.5 Exchange of Experiences with Other Countries or International Organisations

Irrespective of differences in type and size of fishing industries, fishery data collection programmes are generally based on certain basic and commonly accepted approaches. The Dominican Republic can benefit from the experience and knowledge of other countries in the region. Exchange can be facilitated by regional workshops, expert consultations and study tours.

Regional meetings can help deal with:

1. Regional workshops on processing of combined data (including fish stock assessment and bio-economic analyses);
2. Standardisation of data format, for exchange of data
3. Establishment of regional databases
4. Co-ordination of data collection from shared resources and vessels not landing in the home country

Although the Dominican Republic should aim to be involved in such regional initiatives, they are often costly and should always be well justified.

8.6 The Iterative Process of Implementing a Data Collection Programme

Once a programme has started, it will need to be regularly revised. Although modifications of programme are needed more or less every year, the modifications should be made so that the compatibility with earlier years is maintained. If there are no major changes in the fisheries sector or the resources, the annual modifications should become gradually smaller. The sampling programme in particular should be reviewed early on in the implementation. As more information comes available, it will be possible to improve the sampling efficiency.

a) **Define objectives of data collection.** What type of general background information is required, which type of resource evaluation or bio-economic analyses is required?

b) **Identify data.** For example, what should be the content of the Fisheries Annual Report? What is needed for resource evaluation or bio-economic analyses?

c) **Identify data sources.** Who can supply information, where? Identify all parts collecting information or are keeping files about the fisheries sector.

d) **Design data collection programme and design database.** Construct a stratification on fleets, landing places, species and season. Select data collection stations, decide on frequency of sampling, allocate tasks to enumerators, design data forms for interview and for data entry, and make the budget.

e) **Execute training courses for enumerators, encoders and support staff.** Create a “team-attitude” – make all staff feel they are important for the programme.

f) **Collect data**. Design interviews and fill in forms, execute frame survey or vessel registration, execute on-the-job training, make contacts to fishers, distribute material of the sampling programme, pamphlet, T-shirts, etc.

g) **Create database.** Select commercial software and create an application for fisheries database in collaboration with enumerators and encoders. Create a database to meet the requirement specified in a, b and c.

h) **Enter data in database.** Enter data into the computer and carry out preliminary validation, train enumerators and check their performance.

i) **Execute pre-processing of data and validate data.** Estimate the first raisings of data, and validate the raised data. Compare samples to identify extreme values and do other validation exercises.

j) **Evaluate data collection programme and database.** Prepare detailed administrative reports, compare plans with actual achievements. Compare actual costs with budget. Identify the problems and their solutions. Invite consultants for independent evaluation. Compare to other data collection programmes, for example, programmes in other countries.

k) **Arrange user meetings with administrators.** Administrators are here the senior staff in the fisheries directorate (central and provincial) who are associated with the data collection programme in the sense that they are the official publisher of the fisheries statistics and they are the supervisors of the data collection programme.

l) Compile relevant recommendations and return to d) if necessary

m) **Process data.** Complete processing, possibly with resource evaluation or bio-economic assessment.

n) **Prepare reports.** Administrative reports, catch/effort reports, resource evaluation report, bio-economic analysis, etc.

o) **Evaluate reports.** Do the reports contain the expected data? Do reports provide the analyses and advice expected? Is the current sampling frequency adequate? Are additional data required?

p) **Consult with managers, politicians and representatives from industry and fishers associations.** Are the primary users satisfied with the reports? Are there additional requests from the users. Do the managers actually apply the reports as a part of the basis for their management decisions? If necessary, return to a) to redesign programme.

Production of manuals should start as soon as possible as a means for communication within the programme staff. Manuals are not permanent and should be revised regularly. Documentation will also be important. Training materials should be obtain prior to starting the programme as implementing training courses and workshops should be one of the first activities.

9. PROPOSED PROGRAMME

9.1 Overview

The main initial aim will be to identify the objectives of the data collection programme. This will ensure that the data collection is well targeted, and more importantly, help ensure these data are used in decision-making.

To do this the Barahona data set needs to be analysed and the results presented at a series of consultative workshops which will decide on what data is required and how it will be used. The results from these workshops will finalise the data collection programme which should be implemented in Barahona. It is proposed that Barahona also be used as the training district, where administrators and trainers visit from other districts to learn the skills they need to implement the programme.

The data collection programme should be modular, so that each administrative centre is able to carry out its main activities independently of the central authority. Following the successful operation of the Barahona data collection programme, similar autonomous systems should be set up at each site. Each administrative centre would be responsible for managing the raw data and reporting to the central database. The administrative centres should not be responsible for analysis. An important advantage of such a distributed system is it is more robust. A break down in one database or collection system will not affect other areas and the smaller scale of the systems will make the databases simpler to maintain. However, the areas will be able to take advantage of expertise held centrally as well as support each other.

9.2 Programme

	Activity	Personnel	Days
1.1	**Create temporary database**. Create a temporary application in MS Access for fisheries database primarily for trip and biological data input. Database forms should mimic data forms. Includes 2 days training.	Database Consultant	10 days
1.2	**Enter data in temporary database.** Enter data into the computer and carry out preliminary validation, train enumerators and check their performance. Trips: 1440 records / 40 = 36 person days Biological data: 240 / 20 = 12 person days	Local Encoders	24 days for 2 persons
1.3	**Data Analysis.** Primary aim of data analysis is to review possible performance indices and reference points using Barahona data. Analysis should include some resource evaluation and bio-economic assessment.	Stock Assessment / Fisheries Management Consultant	14 days
1.4	**Prepare report** suitable for Management Objectives Workshop.	Stock Assessment / Fisheries Management Consultant	6 days
2.1	**Management Objectives Workshop** will aim to identify what data are collected and what outputs are expected. The Workshop should include policy which identify how data is used in decision-making.	Stock Assessment / Fisheries Management Consultant Facilitator Two local counterparts Decision-makers Other staff	4 days 4 days 4 days 2 days
2.2.	**Stakeholder consultation** includes report back and presentation of final results to high-level decision-makers. One travel daya and meetings day would allow 12 sites to be visited.	Facilitator Two local counterparts	24 days 24 days
2.3	**Prepare final report** with recommendations on performance indices, reference points and data variables collected.	Facilitator Two local counterparts	5 days 5 days
3.1	**Implement changes to Barahona data collection**	Facilitator Stock Assessment / Fisheries Management Consultant	5 days 5 days
3.2	**Implement database upgrade** (Assumes TIP/LRS is used as core database, but adaptations required) Reporting should meet management objectives.	Database Consultant	10 days
3.3	**Training**: should extend to key staff not based at Barahona	Database Consultant Selected local trainers	24 days 12 days
4.1	**Extend Barahona programme to other districts**: implement data collection, including forms, database and	Local trainers	ongoing

9.3 Budget

	Days	Rate[1]	Cost (US$)
Stock Assessment Consultant[2]	34	330	11220
Database Consultant[3]	34	330	11220
Facilitator[4]	31	330	10230
Local Encoders	48	40	1920
International Travel	3 trips		6000
Subsistence	71	120	8520
Local Travel[5]	24	50	1200
Computers and software[6]			13600
Communication			1000
Educational Material			1500
Total			66410

Notes

[1]Rates are assumed to be the same for all consultants except data encoders.

[2]The stock assessment consultant will spend 9 days in country. Data analysis can be done at home base.

[3]Database consultant will spend all time, 34 days, in the Dominican Republic. It may be possible to incorporate this consultancy with the TIP/LRS upgrade.

[4]A local facilitator is preferred. If a meeting facilitator is provided by the Fisheries Department, this would eliminate the facilitator honorarium and reduce the subsistence payments.

[5]Local travel would be by hire car, which should cover all meeting participants.

[6]Eight computers with database software.

10. REFERENCES

FAO, 1999. Guidelines for the routine collection of capture fishery data. FAO Fish. Tech. Man. 382. FAO, Rome. 112p.

FAO, 2001. Jamaica Technical Report. Fourth Workshop on the Assessment and Management of the Shrimp and Groundfish Fisheries on the Barzil-Guianas Shelf. 2-13 October 2000, Cumaná, Venezuela.

PROPESCAR-SUR, 1994. Reportes del PROPESCAR-SUR: Contribuciones al conocimiento de las pesquerías en la República Dominicana. Vol 1. 1994)

Medley 2001. Estimating Discards Using Landings Data. Report to Sea Around Us Project, University of British Columbia, Vancouver, Canada.

Sparre, P. 2001. Sampling Manual. Fisheries Technical Report. In prep.

Thompson, S.K. 1992. Sampling. J Wiley and Sons, New York.