

DATA MANAGEMENT AND DATABASE SYSTEMS AT WORLD FISH CENTER DHAKA, BANGLADESH.



Mission report

**Gertjan de Graaf
Nefisco foundation**

March 2004

WorldFish, Dhaka, Bangladesh

SH
207
P251
G72
2004
c.2



#278

Table of Contents

1. Introduction	3
2. Data sources OF DSAP AND CBFM.....	4
1.1. DSAP	4
1.2. CBFM-2.....	8
3. Present Data processing and analysis.....	10
1.3. Data storage.....	10
1.4. Data entry.....	10
1.5. Data analysis, SPSS vs MsAccess	10
4. Improved database structure and Design.....	11
5. Recommendations and plan of action.....	13
1.6. Building the database system	13
1.7. Improvement of the data	13

Table of Figures

Figure 1: Proposed structure of the database for DSAP and CBFM-2 at WorldFish Center in Dhaka, Bangladesh	11
Figure 2: Example of a users friendly front-end of an MsAccess based system	12
Figure 3: Example of a users friendly front-end of an MsAccess based system with GIS.....	13

Table of Tables

Table 1:Data sets generated by DSAP.....	5
Table 2 Sampling Framework for monitoring farmers supported by Partner NGOs	6
Table 3: Sampling framework for monitoring farmers supported by associate partner NGOs.....	7
Table 4: Data generated by CBFM-1 and CBFM-2	9

SH
207
P251
G72
2004
C-2

0000016426

1. INTRODUCTION

Since the late 90's, the WorldFish Center ¹ in Dhaka has two major research programs ongoing in Bangladesh:

1. **Development of Sustainable Aquaculture Project (DSAP).** This project is aimed at improving food security and eradicating poverty by introducing small –scale fresh water aquaculture as an element into the economic activities of resource poor households in rural areas of Bangladesh. The target groups are poor producers and consumers who can benefit from the better use and management of aquatic resources
2. **The Community Based Fisheries Management project, Phase 2 (CBFM2).** The goal of this project is to improve the livelihoods of poor people dependent on inland aquatic resources. This will be achieved by developing, testing and assessing arrangements for user-based (community-based) fisheries management across the diversity of inland fisheries in Bangladesh, and by informing and facilitating appropriate changes in policy.

Over the year both projects, and their predecessors, have gathered large numbers of field data. All the data are stored in computers in Dhaka. With the ongoing of both project data storage, management and analyses gradually developed at WorldFish. In 2003 it was realized that the data management system was not coherent and easy accessible and should be improved.

This report presents recommendation on the improvement of the data management system².

A detailed ToR of the mission is provided in Annex I.

¹ Former ICLARM

² The mission was covered by the DSAP project and in the ToR emphasis was given to the DSAP data. During the mission it was realized that this was not making sense and consequently CBFM-2 were included.

2. DATA SOURCES OF DSAP AND CBFM

This chapter describes the different data sources and the way they are presently stored and analyzed

1.1. DSAP

Although WorldFish Implemented a number of aquaculture research programmes in Bangladesh previous to DSAP, no raw data of these projects could be traced in Dhaka during the present mission. DSAP has implemented a number of surveys since 2000/2001 and details are presented in Table 1. DSAP has generated 22 datasets since 2000/2001, however, the different datasets can be generalized to three common types implemented for partner and associated NGOs in the project area. The basics of the three different types are presented in Table 2 and Table 3 encompass;

1. **Pond book verification:** During the year the involved farmers enter data on their inputs, out puts, yields, etc in a pond book. The pond book remains with the farmers and each year DSAP sample a number of pond books to get information on the results of the involved farmers.
2. **Impact survey:** In order to study the impact of the project, one or two years after the farmers have received their training of the project, A number of farmers are visited and data on income, production, costs, inputs, etc are collected.
3. **Restore approach:** This is a separate research programme looking in a more integrate approach at the impact of the project and the position of aquaculture income generation in the overall livelihood pattern of the individual farmers.

Next to the basic three surveys DSAP implemented a Household consumption survey and collect basic data of all individual farmers in the different Fish farming group. The last survey is of utmost important as it is in principal the basic data source of the individual farmers covered by the project.

All data collected by DSAP are referenced up to village level, i.e. for each record the District, Thana³, Union and Village/Mauza is known

³ A complication is that recently new Thanas were created by GoB, and the present data system and the GIS management systems still works with the old 490 Thanas

Table 1:Data sets generated by DSAP

Si no	Name	Description	Year	Source	Source type	Data type	No of records	Data group	Remarks
1	Fish farming record book	data from demonstration ponds as filled in by the farmers	2000	ICLAM/RDSAP 2000/01	MsExcel	household description, pond area, input, out put pond	789	Aquaculture	
2	Fish farming record book	data from demonstration ponds as filled in by the farmers	2001	DSAP					
3	Fish farming record book simple format	Data from demonstration ponds as filled in by the farmers	2002	DSAP	left with the farmers			Aquaculture	
4	Monitoring fish farming record book Form 1 (ponds)	Sub sampled data (5%) from Fish farming record book for ponds	2002	DSAP	MsAccess	HH data Inputs out puts	258	Aquaculture	
5	Monitoring fish farming record book Form 2 (rice field)	Sub sampled data (5%) from Fish farming record book for rice fields	2002	DSAP	MsAccess	HH data Inputs out puts	700	Aquaculture	
6	Particulars of Ponds/plots of fish farmer groups form 3	Raw data on ponds, farmers name, used technology	2002	DSAP	not processed	Pond area & ditch area in rice plot		Aquaculture	
7	Summary of fish farmers groups	Summarized data of Si no 5	2002	DSAP	MsAccess	no of new and old ponds, area and technology used and spreading		Aquaculture	Summarized data
8	Impact study of fish culture in ponds, demo farmers RDSAP 2000	Results of pond production & HH data of farmers under RDSAP in 2000/01	2002	DSAP	MsAccess	HH data, Occupation, income, landholding, effectiveness of RDSAP program	257	Aquaculture	
9	Impact study of fish culture in ponds, control farmers RDSAP 2000	Results of pond production & HH data of non trained farmers 2002	2002	DSAP	MsAccess	HH data, Occupation, income, landholding of non trained farmers	254	Aquaculture	
10	Fish farming record book simple format	data from demonstration ponds as filled in by the farmers	2003	DSAP	left with the farmers				
11	Monitoring fish farming record book Form 1 (ponds)	Sub sampled data (10%) from Fish farming record book for ponds	2003	DSAP	MsAccess	HH data Inputs out puts	258	Aquaculture	
12	Monitoring fish farming record book Form 2 (rice field)	Sub sampled data (10%) from Fish farming record book for rice fields	2003	DSAP	MsAccess	HH data Inputs out puts	700	Aquaculture	
13	Questionnaire monitoring performance of the beneficiaries of DSAP			DSAP					
14	Particulars of Ponds/plots of fish farmer groups form 3	Raw data on ponds, farmers name, used technology	2003	DSAP	Not processed	Pond area & ditch area in rice plot		Aquaculture	
15	Summary of fish farmers groups	Summarized data of Si no 5	2003	DSAP	MsAccess	No of new and old ponds, area and technology used and spreading		Aquaculture	Summarized data
16	Restore HH form	See restore manual	2003	DSAP	Restore/MsAccess				Restore software has to be updated
17	Restore profile form	See restore manual	2003	DSAP	Restore/MsAccess				Restore software has to be updated
18	Restore natural resource form	See restore manual	2003	DSAP	Restore/MsAccess				Restore software has to be updated
19	Restore production form	See restore manual	2003	DSAP	Restore/MsAccess				Restore software has to be updated
20	Livelihood Baseline survey	HH, occupation, Income, Holdings	2004	DSAP		HH data, Occupation, income, landholding,		Livelihood	
21	Livelihood survey of restore village	HH, occupation, Income, Holdings	2004	DSAP					
22	HH consumption survey form no 1	Data on the monthly consumption of fish	2004	DSAP					

Table 2 Sampling Framework for monitoring farmers supported by Partner NGOs

Year	2000/01	2001/02	2002/03	2003/04	2004/05	2005/06	2006/07	No of Demo Farmers	Extension Approach	
2000/01	832		300 Gaelle					6248	1-yr Top-down Extension	
2001/02		775	±700			350		6608	3-yr, ± bottom-up ext.	
2002/03		450		±3000			±1200	7654	3-yr, bottom-up ext.	
2003/04			300	420	420	420	420	±11500	3-yr, bottom-up ext.	
2004/05						±6500		±11500	3-yr, bottom-up ext.	
2005/06								±11500		
No hh/ farmers	6248	6608	±5000+8000 =±13,000	±5000+2x8000 +11,500 = ±32,500	3x8000+2x 11,000+11,000 =±55,000	3x11,000+2x 11,000+11,000 =±65,000				
Grant	100% grant	50% with grant money (1x Tk. 2000) and 50% without grant money								
	NGO SUPPORTED EXTENSION		POND BOOK VERIFICATION		IMPACT STUDY		RESTORE STUDY			

Table 3: Sampling framework for monitoring farmers supported by associate partner NGOs

Year	2000/01	2001/02	2002/03	2003/04	2004/05	2005/06	2006/07	No of Demo Farmers
2000/01								-
2001/02				250				5000
2002/03			±270	±700			±320	2721
2003/04					+1300	±1400		3000
2004/05								3500
2005/06								3500
Total hh/ farmers	-	5000	2721	$\pm 1.5 \times 2700 + 3000 = \pm 7000$	$2 \times 2700 + 1.5 \times 3000 + 3500 = \pm 13,400$	$2 \times 3000 + 1.5 \times 3500 + 3500 = \pm 14,250$		

1.2. CBFM-2

Similar to DSAP for CBFM, no data could be traced for previous fisheries programmes implemented by ICLARM/WorldFish in Bangladesh and therefore the inventory is limited CBFM-1 and CBFM-2 only. Since 1997 CBFM 1 and CBFM 2 generated 14 datasets (Table 4) and in principle encompass the following activities:

1. Gear and Household census providing basic data on the household living in the vicinity of the different water bodies
2. Frame survey, monthly monitoring the fishing effort at the different water bodies
3. Catch Assessment Survey, monthly monitoring of the catch of the different gears used

The Gear and household census are important as they provide the baseline for all developments and progress of the project



Table 4: Data generated by CBFM-1 and CBFM-2

Sl no	Name	Description	Year	Source	Source type	Data type	No of records	Sample Location reference	Data group
1	Gear census	Gear census in 15 water bodies	1997-2000	CBFM-1	SPSS	No of gears owned by HH at the different water bodies		Mauza	Fisheries baseline
2	Frame survey	Frame survey on gear use in 15 waterbodies	1997-2000	CBFM-1	SPSS	Monthly fishing effort in different Habitats		Union/Mauza	Fisheries
3	Catch Assessment Survey	Monthly catch data of 15 water bodies	1997-2000	CBFM-1	SPSS	CPUE of different gears in the different habitats		Union/Mauza	Fisheries
4	Kata Kua survey	Data on catch Of Kata Kuas in 15 water bodies	1997-2000	CBFM1	SPSS	Catch and other data of FADs such as Kata and Kua		Union/Mauza	Fisheries
5	House Hold census	Full house hold census on gears owned by HH at 115 water bodies	2001/2002	CBFM-2	SPSS	Gear ownership and social data of individual households	130000	Union/Mauza	Fisheries baseline
6	Baseline survey	Social economic data of HH	2002	CBFM-2	SPSS	HH status land ownership, income, etc	6000	Union/Mauza	Social baseline
7	Impact survey CBFM-1	Study on the impact of CBFM-1 on the HH at 15 waterbodies	2001/2002	CBFM-2	SPSS		1700	Union/Mauza	Fisheries/Socio
8	Frame survey	Frame survey on gear use in 59 waterbodies	2001-present	CBFM-2	SPSS	Monthly fishing effort in different Habitats		Union/Mauza	Fisheries
9	Catch Assessment Survey	Monthly catch data of 59 water bodies	2001-present	CBFM-2	SPSS	CPUE of different gears in the different habitats		Union/Mauza	Fisheries
10	Kata Kua survey	Data on catch Of Kata Kuas in 15 water bodies	2001-present	CBFM-2	SPSS	Catch and other data of FADs such as Kata and Kua		Union/Mauza	Fisheries
11	Sanctuary catch	survey on Catch in sanctuaries	1997-present	CBFM-1/2	?	Catch data from sanctuaries		Union/Mauza	Fisheries
12	Market survey	Survey on the location of markets	1997-present	CBFM-1/2	?	Market location		Union/Mauza	Social/Econ
13	Fish consumption survey	Data on fish consumption of HH	1998	CBFM-1	?	Bi-monthly data on fish consumption		Union/mauza	Social/Econ
14	Fish species Survey/Fish code survey	Data on fish biodiversity	1997	CBFM-1/2	not entered	Fish species occurrence		Union	Fisheries

3. PRESENT DATA PROCESSING AND ANALYSIS

1.3. DATA STORAGE

Since 1997, the project developed from small interventions and small databases into large interventions generating substantial data sets. Within this period also the staff of WF developed their computer skills and encountered the limitations of the different type of software used. This all reflects itself in the present system of data processing and analysis at WF in Dhaka. At present data are entered in either MsExcel or MsAccess and then further analyzed in SPSS or in MsExcel, using pivot tables. This system makes the process of analysis cumbersome and as the majority of the staff is not experienced with SPSS, they have no direct access to the data.

1.4. DATA ENTRY

The inventory of the different data sets indicated the following :

1. Records are not coded uniformly between DSAP and CBFM
2. the geographical location of the sampled areas or the farmers, i.e. District, Thanas, Unions, Mauzas/villages is not following the Bangladesh standards or the geocodes of the Bangladesh Bureau of Statistics (BBS) which makes the combining of WF data with other existing data in Bangladesh difficult
3. Data for the Thanas/Unions are not entered with 'Pop-up' menus resulting in a number of different ways Thanas and Mauzas are spelled
4. For CBFM-2 the exact location of the different water bodies or the Mauza location is not yet provided and for DSAP the village or Mauza name is insufficiently covered during data entry.

1.5. DATA ANALYSIS, SPSS vs MsACCESS

For CBFM-2 most of the data is stored in SPSS and the consultant understands the reluctance of WF to change the system. However it should be realized that there are some fundamental differences between SPSS and MsAccess:

1. is not database management software, it is a Statistical package designed for the sophisticated statistical analysis of large data sets (ANOVA, Cluster analysis, Path analysis, multi linear regression, etc). However this type of analysis are not carried out by WF on their data
2. SPSS is relatively expensive software and not commonly used in Bangladesh and less users friendly as MsAccess or Dbase
3. MsAccess is database management software, it is used in most project of Bangladesh and it is standard of the National Water Sector Water Base in Bangladesh

Considering the above, it is strongly recommended to use MsAccess as the storage system and management for DSAP and CBFM data at WF in Bangladesh.

4. IMPROVED DATABASE STRUCTURE AND DESIGN

The data collected by DSAP and CBFM-2 are relatively straightforward and analysis is not that complicated and can be carried out almost completely in MsAccess.

However, if we take into consideration the different users of the collected data, then a more proper data management platform and design will be needed. In principle there are three users groups of the collected data, each with a different level of computer skills:

1. The scientific staff of WF, DoF and other partners who want to carry out statistical analysis on the collected data.
2. The partner NGO interested in the progress and results of their work
3. DoF interested in production levels and in observed trends

In order to accommodate the three different users groups three different front-ends of the central database are proposed (Figure 1).

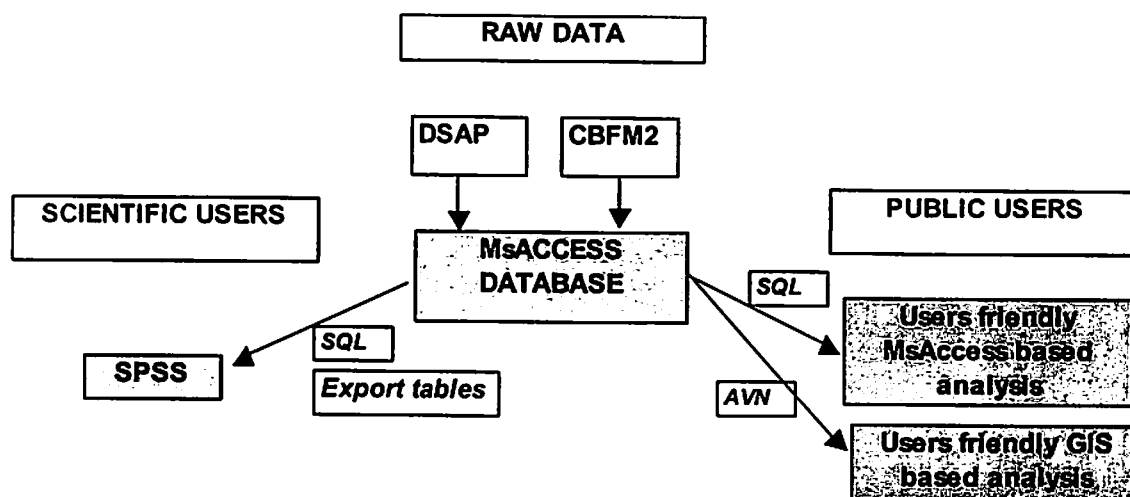


Figure 1: Proposed structure of the database for DSAP and CBFM-2 at WorldFish Center in Dhaka, Bangladesh

For the scientific users, an export facility will be made for tailor made tables or specific data sets to SPSS. In order to avoid that the data sets are

duplicated on the computers it is strongly recommended to use SQL (Microsoft Structured Query Language)

For the non scientific users, a users friendly front-end with simple tables and graphs (MsAccess based) and a users friendly front-end displaying the results in a more geographical context will be made for the latter Visual basic and ArcView Avenue coding will be used.

Figure 2, illustrates an example of a users friendly MsAccess front-end as made for the CPP fisheries data set. By clicking on the different pop-up menus and selecting the different options, Tables or graphs are generated from the raw data.

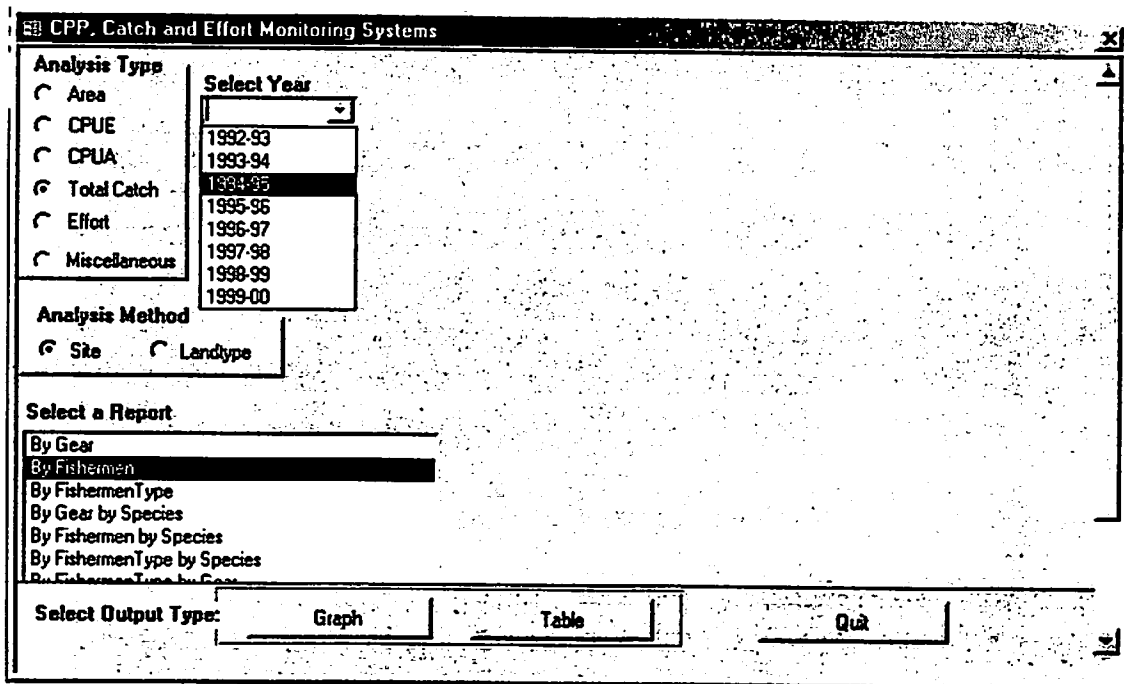


Figure 2: Example of a users friendly front-end of an MsAccess based system

Figure 3, illustrates an example of a users friendly GIS/MsAccess front-end as made by CEGIS for the pilot phase of the National Fisheries Database. Again by clicking on the different pop-up menus and selecting the different options, tables, graphs are made or spatial related data are displayed in the maps.

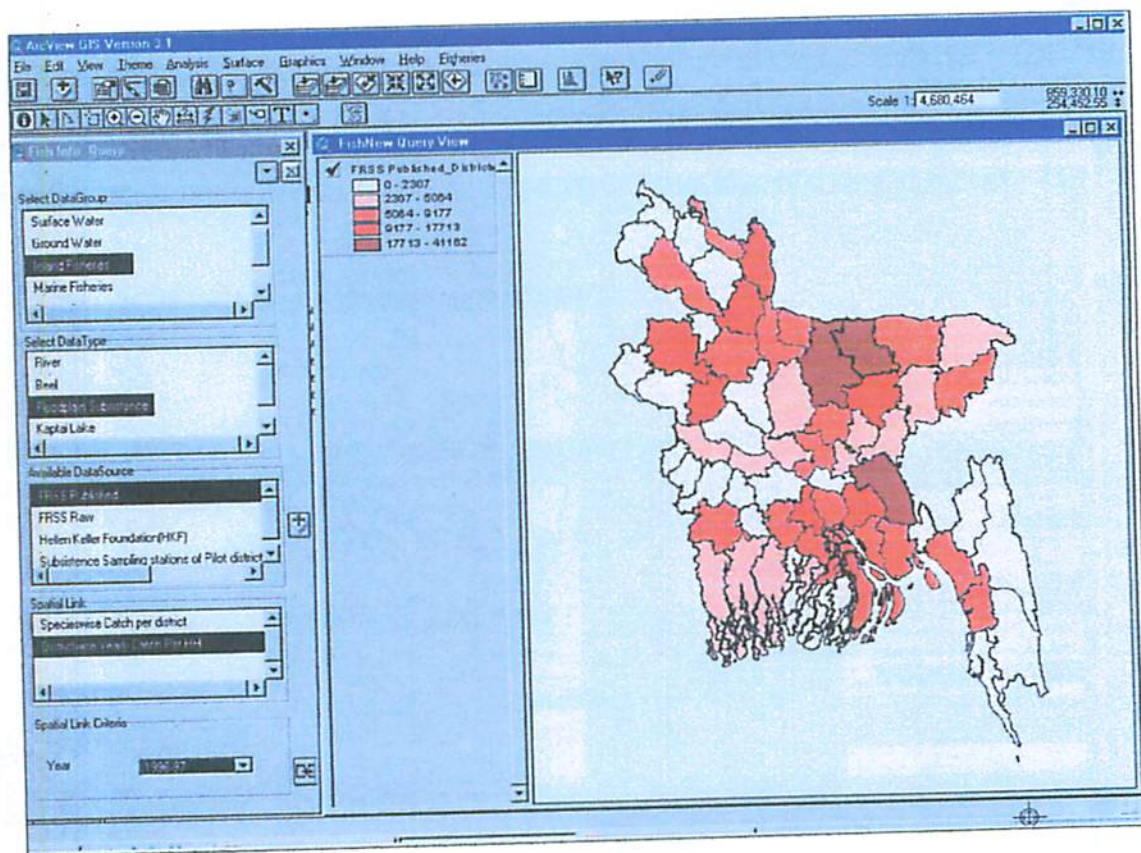


Figure 3: Example of a users friendly front-end of an MsAccess based system with GIS.

5. RECOMMENDATIONS AND PLAN OF ACTION

1.6. BUILDING THE DATABASE SYSTEM

1. A local consultant has been selected to build the data base system and to convert Restore into a newer version of MsAccess. Building the database will be under guidance of the TL of DSAP with constant feedback of the staff of WF. The consultant will start with the DSAP data, if CBFM-2 decides to include their data in the overall data management system, the contract can be extended,
2. A second mission for the international is planned for testing and delivering of the final data base management system. It is strongly recommended to include a mid term mission, as originally planned, to assist with the progress and testing. This could be taken up once CBFM-2 is included in the programme

1.7. IMPROVEMENT OF THE DATA

During the mission the consultant referenced all the DSAP and CBFM data according to the BBS geocodes up to Union level⁴. However, WF staff should crosscheck the codes.

For the DSAP data all village names should be cross-checked with the Mauza names and given the BBS code of the Mauza. If funds are available in the future, the GPS readings of individual ponds should be taken

For the CBFM data the GPS readings of the individual water bodies should be provided in combination with the Mauza the water body is located in. The longitude and latitude can be also taken from the LGED Thana maps. When used in GIS care should be taken that the proper projection setting are used.

⁴ This was required for the development of a preliminary GIS analysis of the data

Annex 1: ToR

Development of a functional multi disciplinary data management and analysis tool for aquaculture related data generated by the WorldFish centre, Dhaka, Bangladesh

Objective and Activities:

Over the years the different programs of WorldFish in Dhaka, Bangladesh has collected thousands of data related to aquaculture. At present the data are stored in different formats, MsAccess, MsExcel and SPSS in different computers. At present, the data are not easy accessible, there is no uniformity and data are difficult to link with data generated by other institutes. As a follow up of the planned GPG/CSI mission it is planned to develop a fully functional multi layered integrated data management and analysis tool for aquaculture data for the WorldFish centre in Dhaka. The following activities are envisaged:

- Elaborate on the analytical frame work for a multi layered data management system as provided by the GPG/CSI mission
- Develop a detailed work programme and design of analytical tool for the data management system
- Assist with the selection and recruitment of a local programmer
- Testing of the system
- Training of staff in the use and further development of the system

To implement the program the planned GPG/CSI mission will be extended with one week to set up the program and select and instruct the local consultant. A mid term visit of one week is planned for June and final deliver, testing and training is planned for September. Travel in February is covered by the CPG/CSI mission, travel and the second mission will be combined with a mission of the consultant for the Dutch government. If the latter is not the case, then travel will be paid by WorldFish.

Annex 2: Schedule of the consultant

Date	Schedule of the consultant
22-02-04	Amsterdam -Dhaka
23-02-04	Arrival Dhaka, Meeting with staff WF, planning of the mission
24-02-04	Inventory of the database, Meeting with EGIS staff
25-02-04	Inventory of the database, GIS layers, selection of local consultant
26-02-04	Inventory social data base & data preparation
27-02-04	Data preparation
28-02-04	Workshop Fish stock assessment programme and data preparation
29-02-04	Development of GIS layers
01-03-04	Development of GIS layers
02-03-04	Data preparation, briefing local consultant
03-03-04	Database design
04-03-04	Development of GIS layers, Presentation to staff of DSAP
05-03-04	Report Writing
06-03-04	Report writing and meeting with local consultant
07-03-04	Presentation results CBFM staff, Travel to Penang
08-03-04	Discussion of results with mark Prein
09-03-04	Report writing, Installing and hand over of GIS system
10-03-04	Presentation of results to staff of HQ and debriefing
11-03-04	Travel to Bangkok
12-03-04	Bangkok - Amsterdam