

CROP PROTECTION PROGRAMME

Archiving data from integrated pest and disease management projects within the Uganda Banana Research Programme

R 8301 (ZA 0565)

FINAL TECHNICAL REPORT

1 May 2003 – 31 May 2004

Project Leader: Savitri Abeyasekera

Statistical Services Centre, The University of Reading
P.O. Box 240, Whiteknights Road, Reading RG6, 6FN, U.K.

Date FTR completed: 27th May 2004

"This publication is an output from a research project funded by the United Kingdom Department for International Development for the benefit of developing countries. The views expressed are not necessarily those of DFID." [R8301 Crop Protection Programme]

Executive Summary

The project sought to assist the National Banana Research Programme (NBRP), Uganda, in improving the quality of their research through a systematised approach to the collection, organisation and management of their research data and related information. Initial steps towards achieving this goal were completed through the archiving of all data, meta-data, study protocols and other information of the three CPP-funded cluster of banana projects, the setting up of guidelines and procedures necessary for maintaining a good data management system, and the development and documentation of an appropriate data management strategy for all NBRP research activities.

The archiving work provided on-the-job training in writing protocols for NBRP staff leading activities under the CPP project, while a 4-day *Research Data Management* workshop provided 28 scientists, research assistants, technicians and data entry persons to learn about procedures for minimising errors at stages of data collection and computerisation, and for further data validation at a later stage. Two members of staff from the Biometrics Unit at Kawanda Agricultural Research Institute received on-the-job training in data management aspect throughout the project's duration and additionally received short-term training in the UK on research data management and statistics. The Unit is now better skilled at providing advice on data management to other NBRP staff.

A considerable number of NBRP scientists, research assistants and technicians participated in meetings to discuss the development of guidelines for good data management. Several informal discussions also took place. Their contributions led to the production of a manual on *Guidelines and Procedures for Effective Data Management*. With endorsement by the Director General of the National Agricultural Research Organisation (NARO) in Uganda, 250 copies of this manual are shortly to be circulated to research scientists in NARO's nine agricultural institutes for their use, or for adapting and adopting the guidelines to suit their own research disciplines.

Discussions with senior scientists of NBRP led to the production of an *NBRP Policy for Research Management with particular emphasis on Research Data Management and Statistical Analysis*. This has been accepted by NBRP senior staff and is expected to assist the programme management in monitoring banana research projects and related activities more effectively. The policy is expected to be adopted by NBRP with immediate effect.

Contents

	Page
Background	1
Project Purpose	1
Research Activities	2
1.1 <i>Preparation of protocols.</i>	2
1.2 <i>Preparation and checking of project data files</i>	2
1.3 <i>Training Biometrics Unit staff in data management and statistics.</i>	3
2.1 <i>Training NBRP research staff in research data management</i>	3
2.2 <i>Development of guidelines on data management</i>	3
2.3 <i>Dissemination of the “Guidelines” manual</i>	4
2.4 <i>A case study to test recommended guidelines</i>	4
2.5 <i>Participation in the follow-on data management project</i>	4
3.1 <i>Consultations regarding a data management strategy for NBRP</i>	5
3.2 <i>Development of a strategy for data management</i>	7
Outputs	7
<i>Output 1. Data archive for CPP Projects R7567, R7529, R7972</i>	7
<i>Output 2. A manual on Guidelines for Data Management</i>	9
<i>Output 3. Development and documentation of</i>	9
Contribution of Outputs to development impacts	9
Appendices	
Appendix 1	Activity Protocols of the IPM Project R7567
Appendix 2	Activity Protocols of the BSV Project R7529
Appendix 3	Activity Protocols of the Weevils Project R7972
Appendix 4	UK-based short term training – Trainees’ Reports
Appendix 5	Mapping Spatial Data – An Example
Appendix 6	Research Data Management Workshop Report
Appendix 7	Data Management Guidelines Manual
Appendix 8	Data Management Proposal approved for Rockefeller funding
Appendix 9	Feedback from researchers regarding data management
Appendix 10	NBRP Policy for Research Management
Appendix 11	Project Protocol for IPM Project R7567
Appendix 12	Project Protocol for BSV Project R7529
Appendix 13	Project Protocol for Weevils Project R7972
Appendix 14	Abstract of a NARO conference paper

Background

Information should include a description of the importance of the researchable constraint(s) that the project sought to address and a summary of any significant research previously carried out. Also, some reference to how the demand for the project was identified.

Data management issues have been a primary concern to the National Banana Research Programme (NBRP)¹ in Uganda, for several years, following difficulties faced when attempting to produce results from the NBRP Diagnostic Survey conducted in the late 1990s. There have been substantial increases in the number of projects handled by NBRP since then, but making comprehensive use of the data gathered from NBRP's research has been compromised by the lack of an effective system for data management.

During discussions in June 2001, July 2002 and November 2002 between NBRP researchers and members of the team of UK personnel working on the cluster of banana projects funded by DFID's Crop Protection Programme (CPP), it was clear that there was an urgent need to have a system in place for managing the data effectively, avoiding duplication of effort in data collection and computerization work across the numerous studies undertaken. Such a system would ensure ready access to comprehensible and usable data. The most interesting analyses in the future are likely to be those that integrate data from different component studies, and it was agreed that this would certainly require high quality data management and the availability of a properly maintained database management system.

Since good management of data is a much neglected task worldwide, the Research Support Unit of the World Forestry Centre in Nairobi has developed a software tool called Logbook to handle a wide range of study and data types. It was recognised that this tool would be invaluable in assisting NBRP to develop a database management system for their research, but a primary task before the commencement of such work was to assemble NBRP data sets from a number of projects so as to first establish the most appropriate ways in which the database development work can commence.

Project R8301 was therefore conceived as the first stage of a larger proposal (part approved now for Rockefeller funding) to develop a comprehensive database management system for NBRP. Discussions with many NBRP scientists during a scoping study visit in February 2003 indicated they needed guidelines on how they may contribute towards the setting up of such a system. It was agreed that this would be best achieved through training of research staff in research data management, upgrading the skills of staff in NBRP's Biometrics Unit, and the development of guidelines and procedures for effective management of data. It was also clear that a data management strategy was needed for NBRP, in particular to resolve issues relating to data security and access with a centrally maintained data archive. Progress on these aspects, together with staff involvement in procedures needed to archive data from three CPP funded projects² during this 13-month project was expected to enhance the quality of NBRP's research outputs and thereby help the wider NBRP goal of raising banana productivity in Uganda for the benefit of the poor.

Project Purpose

The purpose of the project and how it addressed the identified development opportunity or identified constraint to development.

The primary purpose of the project was to enhance the quality of research within the banana research programme through the availability of well-documented and good quality data and to set up procedures that NBRP staff can understand, appreciate and follow, as constituting essential ingredients of a good data management system. The aim was also to reach consensus concerning the most appropriate data management strategy for NBRP. Most staff are now well aware of the need to pay attention to data quality and effective data management, and have achieved reasonable skills in systematising their approaches to data collection and the management of their computerised data files. The availability, through activities of this project, of detailed guidelines on data management procedures, and the development of an NBRP policy for research data management will serve to enhance the quality of future research undertaken by NBRP.

¹ The National Banana Research Programme is based at Kawanda Agricultural Research Institute (KARI), Kampala, Uganda.

² The three CPP projects whose data were archived within this project were R7567 on "Integrated management of banana diseases in Uganda", R7529 on "Management strategies for Banana Streak Virus", and R7972 on "Integrated management of the banana weevil in Uganda".

Research Activities

This section should include detailed descriptions of all the research activities (research studies, surveys etc.) conducted to achieve the outputs of the project. Information on any facilities, expertise and special resources used to implement the project should also be included. Indicate any modification to the proposed research activities, and whether planned inputs were achieved.

1.1 Preparation of protocols

The first output concerned the archiving of raw data, meta-data, protocols and reports associated with the CPP-funded cluster of banana research projects in Uganda, namely R7567, R7529 and R7972. Achieving this output began with the preparation of protocols for all individual activities associated with each of these projects. For convenience, these projects will hereafter be referred to as the IPM project (for R7567), the BSV project (for R7529) and the weevils project (for R7972).

The project leader's previous involvement in the IPM project had already resulted in the preparation of some of the associated activity protocols, and others were prepared in consultation with NBRP and UK team members associated with the project. Protocols for the BSV project were prepared by Ph.D. students Charles Murekezi (registered at the University of Reading, UK) and Jerome Kubiriba (registered at the University of Greenwich, UK), who were primarily responsible for the BSV research work in Uganda. Protocols for the weevils project were the responsibility of M.Sc. students Magara Evarist and Venansio Tumuhaise, registered at Makerere University, Uganda, and senior scientist Caroline Nankinga Kukiriza.

Each protocol underwent close scrutiny to ensure all details had been captured. This necessitated discussions concerning the actual experimental or survey procedures undertaken, raising queries related to the activity descriptions, and filling gaps that had been omitted in the first draft. Several iterations of this process took place over a period of time, also involving a study of the corresponding data files to ensure that the protocol information were consistent with the contents of the data file. Components of the protocol which help in clarifying information residing in the data files³ were copied across to an additional sheet in the MS-Excel spreadsheet data files. In a few cases, the researcher preferred to make reference to the protocol rather than duplicate the information that already resided in the protocols. This was agreed as being a reasonable alternative approach to follow.

Attempts were made to have a consistent format for all protocols across the three projects. The full set of protocols appears in Appendices 1, 2 and 3.

1.2 Preparation and checking of project data files

During visits to the banana programme, data files within each of the projects were scrutinised, queries raised where data errors were found, and comments given (often in written form) to the researchers concerned to assist them in producing clean data files. Special attention was given to ensuring that the contents of all data files could be clearly understood through the provision of associated meta-data.

Discussions were also held with team members of the BSV and weevils project as to the best format, i.e. computer directory structure to use in archiving information from their own project.

The archiving format for the IPM project was decided in consultation with the UK project leader since the Ugandan lead scientist for this project was not available for discussions, having been moved from the banana programme to work on coffee. Because of this, the organisation, data tidying and archiving of information relating to the IPM project was done by Yusuf Mulumba (statistician) working in the Biometrics Unit at Kawanda Agricultural Research Institute (KARI).

³ This is part of the meta-data, i.e. information that allow the project data to be understood by a person not associated with the research.

1.3 Training Biometrics Unit staff in data management and statistics

The Biometrics Unit currently includes three members of staff, namely 1 biometrician (IITA⁴ funded, but who has responsibility also for helping NBRP staff), 1 statistician (Yusuf Mulumba) and 1 research assistant (Allan Rwakatungu). Short term training in the UK was given to Yusuf (4 weeks) and Allan (12 days) in order to upgrade their skills in research data management and for Yusuf, his statistical knowledge. Their visit reports⁵ given in Appendix 4 show details of the training received and their own comments regarding the training.

Both the trainees have progressed considerably in the past 15 months and have acquired skills necessary to maintain the data archive for DFID-funded CPP projects, now available centrally within the Biometrics Unit. They are also now in a better position to guide NBRP researchers in improved approaches to preparing data collection sheets, procedures for data computerisation and data organisation. Procedures have also been set up for doing regular back-ups of the archive and for ensuring data security is not compromised. Details are set out in Output 3 (Appendix 10), i.e. in the *NBRP Policy for Research Management* document.

Management of the archive is now largely Yusuf's responsibility. Directories for archiving data from other donor projects have also been set up, e.g. Rockefeller funded projects. The completion and submission of his Master's degree thesis has also been assisted greatly through advice from the R8301 project leader.

Allan deals with more computer-based issues. An example is shown in Appendix 5 prepared by Allan to demonstrate learning skills in mapping spatial data. Allan has also been responsible for the development of a new website for NBRP, and is being encouraged with suggestions regarding the structuring of the website contents by a Gatsby funded consultant⁶ dealing with the current outbreak of banana bacterial wilt in Uganda. The new website can be accessed through www.banana.go.ug.

2.1 Training NBRP research staff in research data management

A four-day in-country training programme was held in June 2003 for NBRP staff and a few IITA staff involved in banana research work. The workshop was based on training materials produced by the World Agroforestry Centre in Nairobi in collaboration with the Statistical Services Centre at the University of Reading, UK. Feedback from course participants was generally very positive, but the programme was felt to be too short to give participants full confidence and practice in the numerous procedures available for preparation of effective data collection sheets and data validation. The Training Workshop Report is given in Appendix 6.

2.2 Development of guidelines on data management

Many discussions were held with NBRP staff from time to time to seek their opinions on procedures that could be followed to improve NBRP's data collection, validation and computerisation work. Inputs were received from two field assistants, and several technicians, research assistants and scientists, as well as from staff of the Biometrics Unit. The discussions enabled an initial draft of guidelines for data management to be prepared. This was discussed with several NBRP staff at a meeting on 9th February 2004 (Guidelines Meeting No.1), details of which are given in Abeyasekera (2004). Following suggestions made at this meeting, an updated version was prepared and sent in March 2004 to Charles Murekezi (PhD student, BSV project), who had previously agreed to lead this activity.

Further modifications to the guidelines were made by Charles and other banana staff at a meeting at KARI on 8th April (Guidelines Meeting No. 2). Substantial inputs to the document were also made by Charles to provide additional examples that illustrated recommendations made within the guide. An integrative protocol and a series of activity protocols drawn from NBRP's own work were also put together by him, with assistance from two others, to form appendices demonstrating the level of detail required in protocols to ensure that the research work progressed smoothly. The updated document

⁴ International Institute of Tropical Agriculture

⁵ The training visit reports were re-formatted to use fewer pages

⁶ Gatsby Charitable Foundation consultant is Dr. Simon Eden-Green

was reviewed in early May and discussed with Charles on 8th May during a 3-hour meeting (Guidelines Meeting No. 3) to reach consensus about the contents of the document and further changes that need to be made. This resulted in draft No. 4 of the guide which was circulated to several staff involved in banana research and discussed at length and final amendments agreed and accepted at a half-day meeting on 12th May (Guidelines Meeting No. 4). The meeting was attended by 18 NBRP staff, all members of the Biometrics Unit, the IITA database manager, and the visiting database consultant from the World Agroforestry Centre in Nairobi. There was agreement and acceptance at this meeting of the “Guidelines” manual, subject to the inclusion of a few further suggestions made at the time of the meeting.

A further small group meeting⁷ was held on 13th May (Guidelines Meeting No. 5) to agree on additions suggested at the previous day’s meeting. Further work was subsequently carried out to improve the appendices illustrating protocols for a range of different studies. The final version of the manual on *Guidelines and Procedures for Effective Data Management* appears in Appendix 7.

2.3 Dissemination of the “Guidelines” manual

At a meeting of the senior staff to finalise NBRP Policy for Research Management (see 3.2 below), it was agreed that the manual on *Guidelines and Procedures for Effective Data Management* be shared with other NARO⁸ researchers. This would be done through (i) the inclusion of the manual in the banana website www.banana.go.ug; (ii) inclusion of a mention of the manual in their Annual Report to NARO; (iii) launching the manual at KARI.

Later however an alternative to (iii) above was proposed. NBRP staff were of the view that the support of the Director General of the National Agricultural Research Organisation (NARO) in Uganda would be useful in encouraging other researchers in NARO to adapt and adopt this manual for their own research programmes. A visit to see the Director General – Dr. William Otim-Nape, was therefore undertaken with Charles Murekezi and Jerome Kubiriba. The visit was very successful. Dr. Otim-Nape was clearly pleased with NBRP’s efforts towards the production of the manual and immediately suggested that 250 copies be made and distributed to all NARO scientists, with his endorsement in the foreword to the manual. He was also in favour of a meeting being organised with a few invited persons from each of the nine⁹ NARO research centres so that they could be introduced to the benefits of following key messages in good data management proposed in the manual. He was interested in seeing the manual subsequently adopted for other research areas like livestock and fisheries.

Dr. Otim-Nape was also keen that the work should not stop there. He suggested there should be training programmes organised to improve research staff skills in study design and in data analysis techniques. He requested that a proposal be written (inclusive of a budget) for such a training, and suggested that the proposal should also include a facility for providing research staff with some follow-on back-up support.

2.3 A Case Study to test recommended guidelines

During the course of project R8301, a research project on the *Economics of Soil Fertility Management in Banana Production*, funded by the Rockefeller Foundation, and begun in 2003, was used to test out good practice procedures on data management. The lead scientist for this work was Charles Murekezi, who made every effort to follow procedures suggested during the workshop on data management (see 2.1 above) and to ensure a smooth process was followed in collecting and organising research data. One course of action taken as the result of R8301 activities appears in Box 5, section 6 of the guide in Appendix 7.

⁷ Present at this meeting were Charles Murekezi, Jerome Kubiriba and Savitri Abeyasekera

⁸ National Agricultural Research Organisation

⁹ NARO’s research centres are: Kawanda Agricultural Research Institute (KARI), Namulonge Agricultural and Animal Production Research Institute (NAARI), Serere Agricultural and Animal Production Research Institute (SAARI), Livestock Health Research Institute (LIRI), Fisheries Resources Research Institute (FIRRI), Forestry Resources Research Institute (FORRI), Food Science and Technology Research Institute (FOSRI), Agricultural Engineering and Appropriate Technology Research Institute (AEATRI), and Coffee Research Institute (CORI).

2.4 Participation in the follow-on data management project

As a result of R8301, assistance was provided to NBRP in preparing a project proposal for continuation of work on data management to the next level i.e. the development of a proper database for banana research data, based on the software tool Logbook which can be run on pcs supporting MS-Access. This proposal (see Appendix 8) received \$74,000 of Rockefeller funding for a two-year period with an initial activity being a 2-week consultancy visit in May 2004 to NBRP by Peter Muraya, a database expert working for the World Agroforestry Centre in Nairobi.

Work on the development of a database system which enables easy retrieval of data has now begun. As originally envisaged (see assumptions under output 3 of R8301's logframe) the R8301 project leader participated in a workshop held on 10th and 11th May to ensure smooth progression of the work under R8301 with the subsequent Rockefeller project on data management. The workshop entitled *Tools and Methods for Data Management* was largely aimed at informing NBRP research staff on the benefits of the Logbook system, and demonstrating how the system worked with examples drawn from NBRP's own work. Figure 1 shows an example.

An initial component of the workshop involved a short presentation by R8301's project leader on achievements in the past year under R8301 activities, followed by small group discussions to ascertain research staff's views on the impact that R8301 has had on research data organisation and management. Appendix 9 (Part A) presents questions given to different groups of participants and their responses. It is clear that there has been good progress, but a few weaknesses were still identified. This is not surprising, given the space of just 11 months since the Research Data Management workshop (Appendix 6) and the participants' comments at the time that the workshop duration was too short. With the Biometrics staff's acquired skills in effective ways of ensuring good data management, together with the availability of agreed guidelines (Appendix 7), it is expected that the weaknesses identified will be overcome shortly in the near future. Indeed when some of the identified weaknesses were discussed a day later, ways of overcoming them (e.g. imperfect data sheets at data entry) were forthcoming from the participants themselves.

Appendix 9 (Part B) shows the results of a short questionnaire given to the four persons primarily involved in supporting data related work. Their views are rather varied, possibly because they would generally advise different persons, but on the whole their comments mirror the findings above. It is also clear that those who attended the training workshop are less weak in their data management skills than those that didn't.

The above small exercises on the impact of the project on research data management within NBRP, later provided valuable contributions towards finalising the manual on *Guidelines and Procedures for Effective Data Management*. The suggestions made by the senior staff (see Box 4, Appendix 9) were also useful and resulted in additions being made to the draft document on *NBRP Policy for Research Management*, discussed fully on the afternoon of 12th May with senior staff of NBRP.

3.1 Consultations regarding a data management strategy for NBRP

Numerous discussions with NBRP staff and Head, NBRP, helped in decisions concerning best strategies to be adopted within NBRP for research data management. A visit to Uganda in January 2004 for another (DFID-funded) training workshop in Jinja, enabled two short inputs to NBRP before and after the workshop. The first input allowed assembling together short communications prepared by NBRP staff (both scientists and data managers) during discussions in 2003 about possible strategies to ensure an effective data management system (Abeyasekera, 2004). This included discussions and presentations made by NBRP staff during the June 2003 workshop on *Research Data Management*. (See Abeyasekera and Dale, 2003).

Figure 1. Logbook example for the IPM project R7567

Data items

title	measurement units	name	datatype
Is flowering date?	()	Isfldate	text
Total leaves at flowering	(count)	Ttlevflg	numeric
Youngest leaf spotted	(count)	yglevspot	numeric
Girth	(cm)	Girth	numeric
Height at flowering	(cm)	Height	numeric

3.2 Development of a strategy for data management

On the basis of communications described above, a draft structure for an *NBRP Strategy for Data Management* was formulated, circulated amongst all staff and discussed during a meeting (Strategy Meeting No. 1) of senior staff on 10th February 2004. An updated version (Draft 2) was sent to Head, NBRP in mid March 2004 for consideration. At this stage, it was also realised that a strategy for data management alone was not possible without due consideration being given to the management of research work within NBRP, and so the strategy document was retitled *NBRP Policy for Research Management with particular emphasis on data management and statistical analysis*. The 2nd draft of this document was further discussed on 7th May in a small group meeting¹⁰ (Strategy Meeting No. 2), then further modifications considered during a meeting with Jerome Kubiriba on 8th May (Strategy Meeting No. 3), and draft 4 circulated for discussion during a meeting¹¹ (Strategy meeting No. 4) with NBRP senior staff on 12th May.

It is worth reporting an addition that was made to the "Strategy" document following Charles Murekezi's involvement with the activity protocols of the *Data Management Guidelines* document. This was his own recognition of the importance of justifying reasons for every action taken during the research process. He emphasised this point strongly with respect to data collection activities undertaken within NBRP at Strategy Meeting No. 4, when he reported that a questionnaire he was developing had initially been 10 pages long, but when he tried to justify reasons for every question in the questionnaire, the length of his questionnaire reduced from 10 pages down to 5 pages! Savings in data collection effort from emphasis given to this issue have been incorporated in section 4.2.2 (ii) in the "Policy" document in Appendix 10.

Other highlights of the "Policy" document include (a) establishing a *Management Monitoring Panel* to ensure that NBRP policy and guidelines for research and data management are followed; (b) setting up procedures for backing-up of archived information and data security; (c) agreeing on the need to allocate budget lines in project proposals for data management and dissemination activities; (d) formulating procedures for predicting and monitoring time needed for different activities by each staff member, and (e) establishing procedures for the publication of project findings and issues concerning data access.

Amendments suggested to the "Strategy" document at Meeting No. 4 were discussed with Jerome Kubiriba and Charles Murekezi at a meeting on 13th May, and a final version of the NBRP Policy agreed. This final version appears in Appendix 10.

Outputs

The research results and products achieved by the project. Were all the anticipated outputs achieved and if not what were the reasons? Research results should be presented as tables, graphs or sketches rather than lengthy writing, and provided in as quantitative a form as far as is possible.

Output 1. Data archive for CPP Projects R7567, R7529, R7972

The first output is the availability of all raw data, meta-data and study protocols of the CPP-funded IPM (R7567), BSV (R7529) and weevils (R7972) projects in a central archive. This output was fully achieved with respect to the raw data and study protocols associated with activities within each of the projects. Although the IPM and BSV projects' Final Technical Reports were submitted to CPP management in mid 2003, some field research activities are still on-going, and hence it was aimed that the data archive for these projects would include all data collected up to December 2003. This was achieved. With respect to the weevils project (R7972), it was decided to include all data collected up to March 2004 to coincide with its end date. This too was achieved, although a few data collection activities are still on-going.

¹⁰ Present at this meeting were Dr. Tushemereirwe, Head, NBRP, Charles Murekezi and Jerome Kubiriba (BSV project scientists), Nora Odoi (Communications Specialist), and Savitri Abeyasekera (SSC, Reading University).

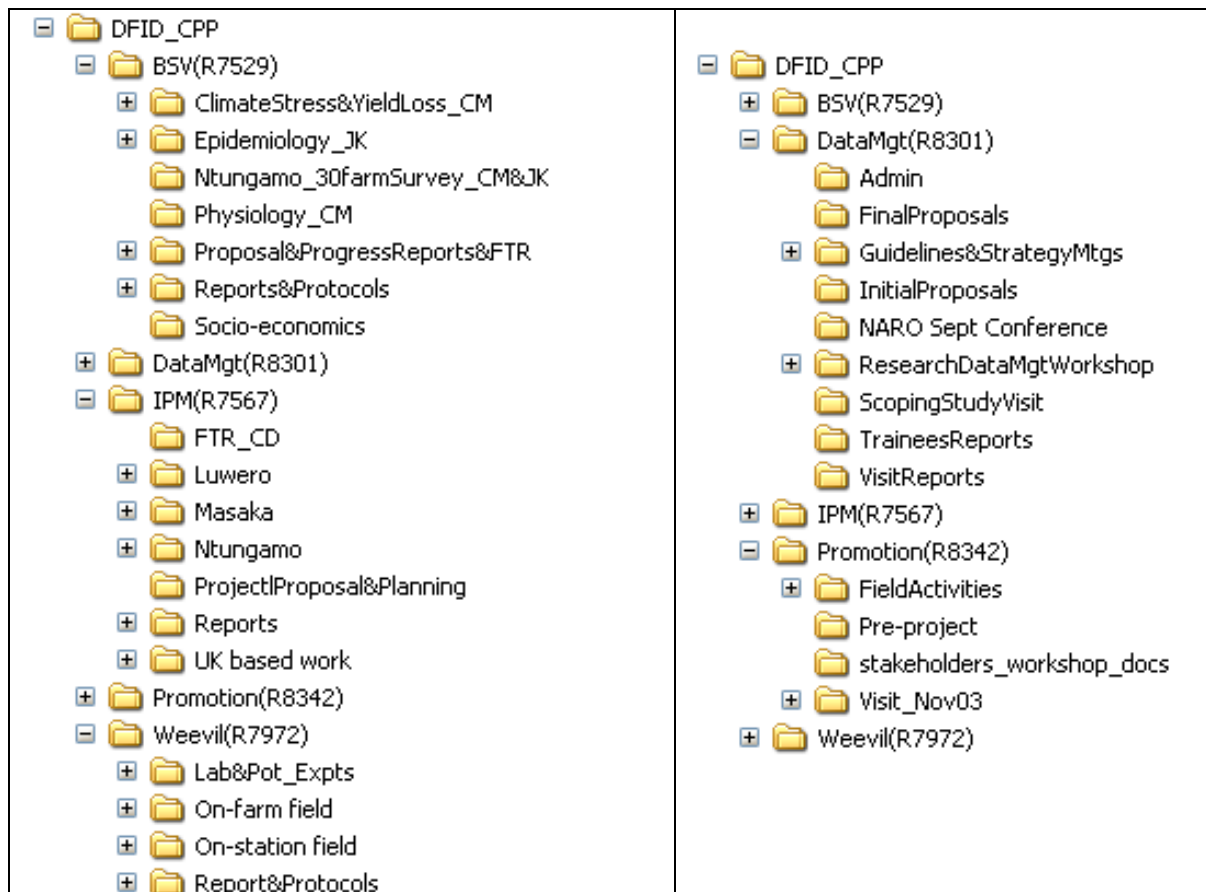
¹¹ Present at this meeting were Dr. Tushemereirwe, Head, NBRP, Dr. Dezi Ngambeki (socio-economist), Nora Odoi (Communications Specialist), Charles Murekezi and Jerome Kubiriba (BSV project scientists), Peter Muraya (Database Consultant from World Agroforestry Centre, Nairobi), Fred Bagamba (Socio-economist), Philip Ragama (IITA, Biometrician), Yusuf Mulumba and Allan Rwakitungu (Biometrics Unit), and Savitri Abeyasekera (SSC, Reading University).

The inclusion of meta-data involves keeping all background information related to the numerical data (e.g. full description of all measurements, maps, questionnaires, etc), together with other associated information such as photographs, reports, talks and other presentation material. The archive attempted to capture all such information and was successful in doing so except in a few instances. For example, in the IPM project, the lead scientist is no longer with the banana programme, and therefore clarification on some issues could not be resolved. In the BSV project, queries with respect to some of the socio-economic information could not be resolved due to some communication difficulties. In the weevils project, the version of the Final Technical Report is still incomplete, and the archive has yet to receive soft copies of the MSc theses written by the two students who worked on this project. These theses are yet to be examined. They will then be placed in the archive.

Overall however, the archive is as complete as can be achieved with project activities to date. Figure 2 shows the directory structure of the archive. The left-hand side shows details for the three projects for which an archive forms the first output of R8301. The right-hand side of Figure 2 shows that the general principles of archiving are now being followed for project R8301 (for which this document forms the Final Technical Report) and for project R8342 on *Promotion of improved IPM practices for banana diseases and pests in Uganda*.

In Figure 2, the IPM, BSV and weevils project directories also included a README file giving a brief overview of the project, with extracts drawn from the project proposal, inclusive of the project's logframe. The README documents corresponding to the IPM, BSV and weevils projects are shown in Appendices 11, 12 and 13. They also show a full list of all computer files available in the archive.

Figure 2. Directory structure for the archive of CPP-funded projects



Output 2. A manual on guidelines for Data Management

The second project output is the production of a manual documenting guidelines and procedures necessary for maintaining a good database management system in consultation with NBRP staff and facilitated through staff training in research data management. This output was achieved. The manual giving *Guidelines and Procedures for Effective Data Management* appears in Appendix 7. It has been agreed and accepted by NBRP staff, as well as by the Director General (DG) of the National Agricultural Research Organisation (NARO) in Uganda, who has suggested its dissemination to all other NARO researchers. DG-NARO is also advocating that researchers from other disciplines such as fisheries and livestock adapt the manual to suit their own requirements.

Output 3. Development and documentation of a Data Management Strategy for NBRP

The project's third and final output is the development and documentation of an appropriate Data Management Strategy for all NBRP research activities, which is accepted by NBRP staff and collaborators. As explained in Research Activity 3.2, this "Strategy" document was retitled *NBRP Policy for Research Management with particular emphasis on data management and statistical analysis*. It was finalised and agreed with NBRP senior staff at a meeting on 12th May. Senior staff not present at this meeting were Caroline Nankinga Kukiriza (on maternity leave), Josephine Namaganda (who was attending a NARS meeting in Jinja), Robert Kalyabara, Kephah Nowakunda and Enid Katungi (who were all out of station at the time) and Gahakwa Rose (who was under pressure from NARO to complete an urgent project proposal). Three of these six persons however participated in previous meetings and/or informal discussions and it is therefore reasonable to claim that the "Policy" document is acceptable to NBRP senior staff.

Time limitations did not permit the policy document to be circulated for comments and acceptance by all NBRP collaborators. However, this is about to be done by Head, NBRP before it is adopted by NBRP as its current policy.

Contribution of Outputs to developmental impact

Include how the outputs will contribute towards DFID's developmental goals. The identified promotion pathways to target institutions and beneficiaries. What follow up action/research is necessary to promote the findings of the work to achieve their development benefit? This should include a list of publications, plans for further dissemination, as appropriate. For projects aimed at developing a device, material or process specify:

- a. What further market studies need to be done?
- b. How the outputs will be made available to intended users?
- c. What further stages will be needed to develop, test and establish manufacture of a product?
- d. How and by whom, will the further stages be carried out and paid for?

Project outputs contribute indirectly towards DFID's development goals in that enhancing research staff capacity to collect reliable data, and to manage and organise the data efficiently and effectively will enable them to produce research results based on solid, defensible, high quality data. Research scientists, research assistants, technicians and data entry personnel in NBRP are now well sensitised to the importance of paying attention to data quality and maintaining the meta-data alongside data of a numerical nature. Together with the emphasis given by the Head of the banana programme to the importance of data management issues, and the availability of an *NBRP Policy for Research Management*, research data quality can now be maintained to a far higher standard than before.

Further work is now progressing towards the development of a proper database management system with assistance from a database expert Peter Muraya from the World Agroforestry Centre in Kenya, and a little support from the Statistical Services Centre at the University of Reading. This work is being funded by Rockefeller Foundation as a direct result of initial work on R8301. Comments from Peter Muraya during his consultancy visit in early May to NBRP were very positive. He expressed surprise at the level of data organisation that already existed and indicated that NBRP were ahead of other groups to whom he had provided advice.

The manual on *Guidelines and Procedures for Effective Data Management* will be printed by NBRP and distributed to NARO research scientists in each of Uganda's nine natural resource based research institutes. The manual has a foreword with a message from the Director General of NARO, urging all scientific staff to use the manual. However, DG-NARO has recognised that data management alone will not improve research quality without parallel improvement in research

methods concerning appropriate study design and methods of data analysis. Discussions are on-going with DG-NARO to determine ways in which these issues can be addressed.

In addition to the above, copies of both the *Guidelines and Procedures for Effective Data Management* and the *NBRP Policy for Research Management with particular emphasis on Data Management and Statistical Analysis* will be placed on NBRP's website at www.banana.go.ug shortly after the guidelines manual has been disseminated.

An abstract of work on this project has also been submitted to organisers of a NARO conference being held in September 2004 in Kampala, Uganda. The conference title is *Integrated Agricultural Research for Development – Achievements, Lessons Learnt and Best Practise*. A copy of this abstract appears in Appendix 14. It is hoped that the abstract will be accepted for presentation and that it will serve to sensitise NARO researchers of the importance of attention to data management issues.

References:

Abeyasekera, S. and Dale, I.C. (2003). Research Data Management Workshop, 2nd and 4th-6th June 2003, Report. (Appendix 6 of this report).

Abeyasekera, S. (2004). Visit to the Uganda National Banana Programme – Visit Report No. 4. Submitted to CPP.

Murekezi, C., Abeyasekera, S., Mulumba, Y., Rwakatungu, A., Kubiriba, J. and Tushemereirwe, W.K. (2004). *Guidelines and Procedures for Effective Data Management (with emphasis on banana research)*. National Banana Research Programme, Kawanda Agricultural Research Institute, Kampala, Uganda, 35 pp.

UNBRP (2004). *Uganda National Banana Research Programme Policy for Research Management with emphasis on Research Data Management and Statistical Analysis*. Kawanda Agricultural Research Institute, Kampala, Uganda, 14 pp. Available at <http://www.banana.go.ug> (Appendix 10 of this report)