Case Studies Guide
(and related resources)

Ric Coe and Roger Stern

July 2016

This guide is to inform students, staff and researchers about a range of case study materials that can be used within statistics courses. It is an update from a 2009 guide prepared for a new regional MSc course based in Kenya.

STATISTICAL
SERVICES CENTRE

UNIVERSITY of
Reading

Introduction

We propose that statistics courses become more problem-driven, and this guide gives information about some problems that could be used in teaching.

The proposed approach has been used in other contexts, and hence some sets of problems/case studies have already been developed and can be used here. We provide the full materials for these sets of problems, so staff and students have access to the supporting information, as well as the individual case studies

This guide provides an overview of the sets of studies. Each set is described briefly, to provide sufficient information for resource persons and students to choose the individual studies of interest. Within these sets more detailed information is provided for a subset of the case studies, to support their immediate use.

This document and the case studies will be updated on our Moodle and our web site. If you have a good internet connection, then check for the latest versions before using. Go to the Moodle site, on www.statistics-training.org and log in. Resource staff without a good internet connection should download a snapshot of the same information.

Open Data

A good case study

A case study is a real life example of a problem and the research done to address it. Parts of a case study may be simulated or hypothetical but the larger problem and research should be real

Some case studies could be problems yet to be researched, but most will have data

Within the set of case studies we propose some should be as follows:

- They address or involve big cross-cutting and emerging issues (HIV, gender, poverty, climate, environment degradation,…)
- They focus on a wide range of different areas – not all be limited to animals or climate or agroforestry
- They integrate social and biophysical research
- Some are 'live', i.e. address a current problem. Hence the data next year will not the same as the data this year.
- They have some published components (results) available. And there would be some unpublished components, particularly describing elements that did not work out as planned.
- They ideally have a researcher involved who would like to present aspects of the case study to the course. Otherwise a video can sometimes be used, and can become a component of the case study.

A case study needs documentation giving:

- problem description
- research protocols

- raw (uncleaned) data and/or the data in a form that is ready for analysis
- copies of instruments
- references to published results, with ideally a copy (pdf) of the publications
- permission to use it in the teaching (and the conditions if any)

The case studies can be used in a variety of ways within the teaching. They can provide examples with a context that students can understand and even care about. Within this context students can be asked how they would proceed, possibly to design a study, or to conduct the analysis etc.

The publications provide a source of papers that can be used as material to understand/discuss/critique.

They can reduce the 'overhead' of learning about the context and background to examples. If a few large case studies are used often in the same or different courses, students get familiar with them and the background does not need repeating.

They can be used to discuss how results allow conclusions to be reached and hence contribute to solving the problem

We started by compiling and describing existing case studies. These are described here. Lecturers should then also look for cooperating and open-minded staff in each university who is prepared to provide one.

Version and Evolution

This version is prepared to accompany the first release of the R-Instat software. We expect them to change reflecting the interests and real studies of individual lecturers. We hope that with these contributions, the collection will grow and become more relevant. These are open educational resources and have the same philosophy as open source software.

The Collections

Four existing collections are described. Then there are new collections being assembled to address current and emerging thematic areas – participatory research and climate change. Finally we have a collection of other studies that broaden the scope.

Existing Collections

ILRI

This set was initially released in 2006 with an updated version in 2011. It was led by ILRI (International Livestock Research Institute), though involved many staff from the region, and is not limited to livestock examples. Its main components are a set of six articles that cover the research process from conceptualization of the problem to the presentation of

the results. These six articles are linked to materials from 17 case studies, which include surveys, experiments and one on processing climatic data.

Each case study has a report that is presented in a consistent format, making it easy to see the consistency in the research process for a wide variety of tasks. Further information is in the special "ILRI guide".

ICRAF

This is a set of resources prepared by the ICRAF (World Agroforestry Centre) research methods group. It has with the following sections:

1. Problem analysis
2. Current understanding
3. Objectives and hypothesis
4. Research designs
5. Handling data
6. Formal analysis
7. Publication and beyond
8. Closing the loop

Each section links to a description of the corresponding stage in the research process and to useful resources for training and support.

Of particular relevance are detailed notes and case studies for a training course called "Data analysis of agroforestry experiments". (This is within Section 6 above) They are more general than the name implies, with the 15 case studies covering surveys, as well as on-station and on-farm studies.

There are four guides for the data analysis training course that give a) an overview, b) lecture notes, c) suggested exercises and d) the protocols for the case studies.

Each dataset is a well-organized Excel file.

SADC

This is a comprehensive set of resource materials for a basic course in statistics, developed for the National Statistical Systems, and for the partner Universities in the 14 SADC countries.

This set of training materials was also designed largely to be problem driven, and includes information on 11 case studies. They include a large subset of the Tanzania agricultural "census" and one year of data from the annual crop cutting survey to estimate maize yields in Swaziland.

SSC

The SSC has a set of resources, partly prepared as part of its work for DFID (Department for International Development). They include a set of good practice guides and a set of case studies that illustrate statistical good practices.

On farm and participatory studies

Thirty years ago most agricultural research was "on-station", with the exception being the work of economists and social scientists. It is now very different, with much more of the research being "on-farm" and (thankfully) less of a divide between those who call themselves "physical scientists" and "social scientists". Not only is research 'on-farm', but farmers are participating in all stages of research from planning and conceptualization, though design and data collection, to evaluation. This has profound implications for research methods, including statistical methods.

Statistical support for agricultural research is therefore a much broader subject than before. Materials are provided both for on-farm and participatory work in its own right, and to help the provision of support for those who move on farm, but try to do their research through out-of-date on-station eyes.

Climate change

There will always be emerging issues. Currently - and for the foreseeable future – climate change has become an important issue in many areas of research.

Many scientists know they must link their work to climate change, to secure funds, and so include phrases in research proposals that their work is "highly relevant to issues of climate change". But they are then unsure how to translate this sort of general statement into real activities, once funding is provided.

We include case studies with climatic data and other resources on climate variability and climate change for two reasons. The first is the general importance of this topic to all our futures, as well as to agricultural research. The second is that climatic examples are useful in explaining many statistical ideas, as described in Kurji and Stern (????).

Other studies

We do not want to limit the collection to any topics. Research and development problems will provide a growing collection of other case studies. We hope these will reflect current trends in research methods. It is useful if some of the problems are ones that are currently being undertaken. These are likely to change each year. In future years the use of current problems can relate to issues found by students during their projects or assignments.

The ICRAF set

The full set of resources, prepared by ICRAF cover the research process stage by stage. Within these resources a training course called "Data analysis of agroforestry experiments" made use of 15 case studies. The four guides are also available. They are as follows:

- an overview,
- lecture notes,
- suggested exercises
- protocols for the case studies.

These guides are available as pdf files and also as Word files, for staff who would like to adapt them. For the MSc course, we expect that the first part of the course (Sessions 1 to 9 in the lecture notes) would be familiar materials for the students after the first semester.

The case studies are as follows (Those in bold with a * are used in Sessions 1 to 9 of the lecture course; Study 7** is also used in the "Continuing" set):

Case study	Title	Type and size
1	Relay planting of Sesbania Sesban and maize	On-station 4 by 3 Factorial trial
2*	**Effect of Tithonia and lantana mulches**	**On-farm trial with 28 farmers from 2 districts, having either 2 or 3 plots**
3*	**Screening of species for three year fallow**	**10 treatments in randomised block design, 2 years**
4*	**Tree management trial**	**Mixed cropping trial, 10 treatments, sort of factorial**
5*	**Seed production trial**	**Incomplete block design, 20 treatments in blocks of size 4**
6	Fruit tree survival	On-farm trial with 5 farmers, 72 plots and survival measured over 3 years
7*	**On-farm cropping with sesbania and gliricidia**	**On-farm trial with 42 farmers having either 2 or 3 plots**
8	Roots and competition	On station trial, randomised block, 11 treatments
9	Hedgerow intercropping systems	Small on-station trial
10	Mulches as P sources for maize	On-station trial with 10 treatments and unequal replication
11	Feeding trial	On-farm study with 12 cows, crossover design.

12	Effects of nutrients on striga, weeds and maize	On-station split plot design
13*	**Improved fallows on soil phosphorus**	**On-farm split plot design with 9 farmers**
14*	**Farmers experience of improved fallows**	**Survey with 2035 households in 17 villages (population), and 473 households in participating sample**
15	Improved fallow and inorganic fertilizer	On-farm trial in 17 villages, with 53 farmers

The ILRI set

This resource was released in 2006 and updated in 2011. It is a rich resource and worth exploring. The front page is as follows:

The case for this type of resource was made in a meeting in 1999, the proceeds of which are included within the resource, under Rowlands, 2000. This includes 29 articles of which the first two are:

The changing nature of agricultural research in Africa, by J. Lynam.

The impact of the changing agricultural research focus on biometric methods and the role of the biometrician, by R. Coe.

The resource was described in a paper by Rowlands et al, 2005.

The main components of the resource are six teaching guides and 17 case studies. The teaching guides are as follows:
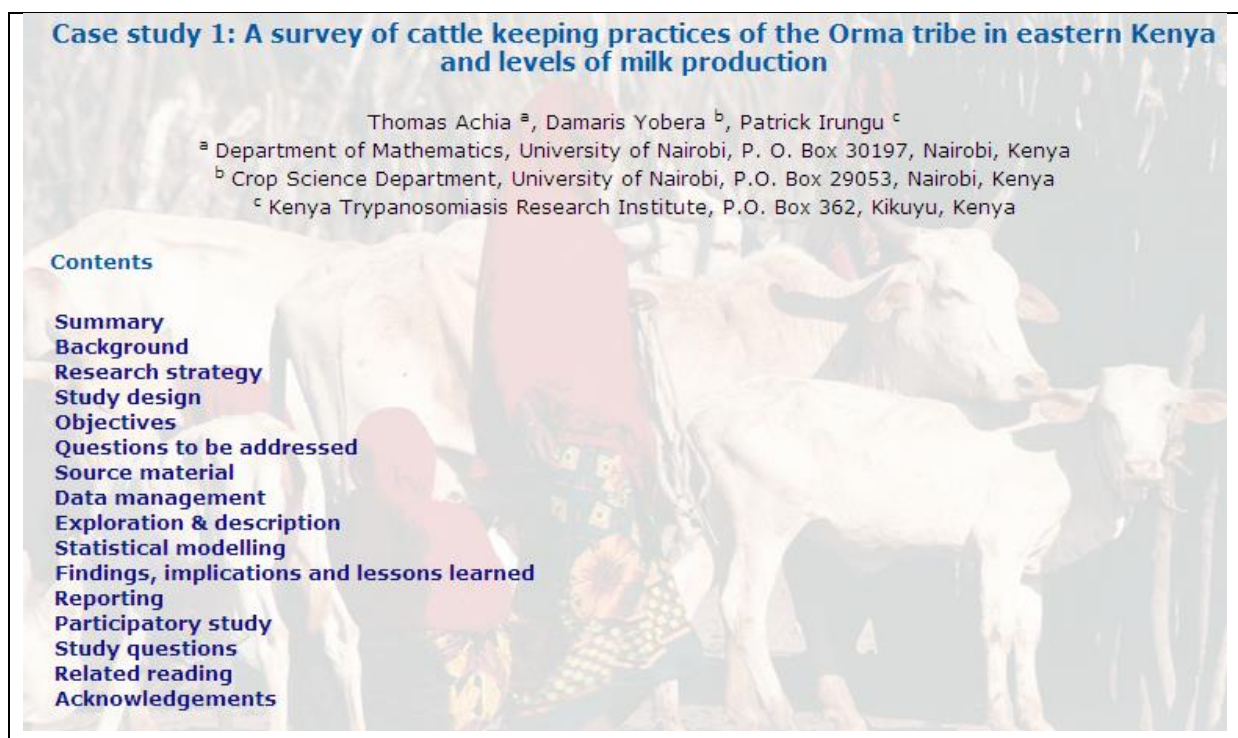
Research Strategy

Study design

Data management

Exploration and description

Statistical modelling

Reporting

Each of these guides cross links to the materials in the 13 case studies. The first page of one of the case studies is shown below:

Case study 1: A survey of cattle keeping practices of the Orma tribe in eastern Kenya and levels of milk production

Thomas Achia [a], Damaris Yobera [b], Patrick Irungu [c]
[a] Department of Mathematics, University of Nairobi, P. O. Box 30197, Nairobi, Kenya
[b] Crop Science Department, University of Nairobi, P.O. Box 29053, Nairobi, Kenya
[c] Kenya Trypanosomiasis Research Institute, P.O. Box 362, Kikuyu, Kenya

Contents

Summary
Background
Research strategy
Study design
Objectives
Questions to be addressed
Source material
Data management
Exploration & description
Statistical modelling
Findings, implications and lessons learned
Reporting
Participatory study
Study questions
Related reading
Acknowledgements

Each case study is written up in a consistent format and with sections, as shown in the figure above, that follow the same order as the teaching guides.

Each guide includes the data sets, usually as Excel files, with questionnaires, where available. Most of the reports link to articles of published results. A references section from the front screen gives the full list, and there is also an Institutes section that provides the status of biometry at Institutes in Sub-Saharan Africa in 2006.

The 17 case studies are as follows (Those indicated * and in bold include more detailed information about their use):

Study	Title	Type and size
1	Cattle-keeping practices in Eastern Kenya	Survey of 48 households plus quantitative measurements on 164 animals
2	Herd monitoring study in Uganda	Survey (2-stage cluster design) on 467 cows from herds in 8 regions.
3*	**Sheep weaning weight in Kenya**	**Breeding experiment over successive years – standard analyses**
4	Sheep weaning weight in Kenya	Same data as case study 3. Calculation of genetic variation
5	Best-bet accessions of Napier grass. Study in Ethiopia	Series of experiments
6	Access to credit for smallholder farmers in Kenya	Experiment plus cross sectional and a longitudinal surveys
7*	**Climate change in Southern Zambia**	**Monitoring data – daily rainfall records from one site. See also SADC study 7**
8	Goat feeding study	Experiment with three treatments
9*	**Germination study in Ethiopia**	**Laboratory experiment**
10	Relationship between tsetse control and livestock productivity in Ethiopia	Monitoring study
11*	**Animal genetics resources survey in Swaziland**	**Data from a livestock breed survey.**
12	Animal genetics resources survey in Swaziland	Same data as study 11. Further analysis
13	Nutritional study on goats in Ethiopia	Experiment with 3 genotypes and three feeding regimes

The SADC set

This set was produced in 2007. The resource consists of detailed notes including case studies for a basic course in official statistics. It provides resources at 3 levels, "Basic", "Intermediate" and "Higher".

This is also a very "rich" resource and the materials at the first two levels are designed to support a problem-based approach to learning.

Courses could "cherry-pick" items from specific sessions to provide useful resources for some teaching sessions. As an example the figure below shows the sessions for half a module at the basic level. This includes some topics that are not always associated with a first statistics course.

Part C: Methods used for collecting data

Session 11. Methods used for collecting data and the data cycle

Session 12. Asking "good" questions

Session 13. Qualitative methods for data collection

Session 14. Choosing an appropriate methods for data collection

Part D: Special topics

Session 15. Identifying and reducing sources of error in data collection

Session 16. The role of enumerators in statistical data collection

Session 17. Assessing data critically

Session 18. Principles for Drawing and Reading a Map

Session 19. Preparing a structured field report

Session 20. Ethical issues in the collection of statistics

The case studies are as follows:

Study	Title	Type and size
1	Tanzania agricultural survey	Subset for Tanga region, with 3223 households, and other data at many levels
2	Swaziland educational census	Census on all 754 schools
3	Millenium development goals	Information on the MDGs including a 2007 progress chart
4	Fundamental principles of official statistics	Information on the principles including a UN survey of all countries
5	Swaziland crop-cutting survey	Survey data at 3 levels, interview, crop area and crop yields
6	Solar energy	Information, particularly on solar cooking, including (hourly) sunshine data from Botswana
7	Climate change in Zambia	Same data as ILRI Study 7, with more information and reports
8	Climatic data	See the emerging issue set of studies below
9	Kenya poverty survey	Data and other resources from the 1997/8 survey
10	Rice survey	Sample data set from the paddy survey game
11	SADC Breed survey	Same as ILRI Study 11

The SSC set

The SSC website provides a wide range of resources that are relevant as support materials. They include a series of "good-practice" guides and other resources, largely produced as part of DFID funded projects. The figure below indicates some of the resources:

Examples of our involvement in DFID-funded projects

- Bolivia: CIAT, Strategy for Development of Biometric Capabilities
- Ethiopia: Use and Management of Donkeys by Poor Societies in Peri-Urban Areas
- Forest Dependent People – A Feasibility Study
- Malawi: Free Inputs Programmes
- Tropical Forests: Documentation of UK holdings of growth and yield, inventory and other data
- Uganda: Banana Research
- Zimbabwe: Food Security and Humanitarian Relief Programme
- DFID SSRU "Young Lives" project – an international study of childhood poverty in several developing countries

Resources for researchers working in development

These resources have been made available, thanks largely to funding by DFID. They are intended primarily for researchers working on DFID projects.

- Statistical Guidelines
- Case Studies of Good Statistical Practice
- Integrating qualitative and quantitative approaches – a methodological framework

The case studies of good statistical practice are of particular relevance here. They are as follows:

Case Studies of Good Statistical Practice

The SSC took up a suggestion by DFID to review a number of research projects with a view to producing case studies of good statistical practice. Eight case studies have been produced, intended as illustrative examples to help you in your research and development work. These can be viewed as PDF files via the links below:

1. Good practice in well-linked studies using several methodologies – based on a CPP project in the Philippines.
2. Good practice in survey design and analysis – based on a CPP project in India.
3. Good practice in on-farm studies – based on a joint NRSP/CPP project in Bolivia.
4. Good practice in researcher and farmer experimentation – based on an NRSP project in Zimbabwe.
5. Determining the effectiveness of a proposed sample size – based on work for DFID by a private sector research firm in Bangladesh.
6. Good practice in data management – based on a bilateral project in Malawi.
7. Good practice in the preparation of research protocols – based on (i) a Plant Sciences Research programme project in West Africa; (ii) an NRSP project in Nepal; and (iii) an NRSP project in Bangladesh.
8. Developing a sampling strategy for a nationwide survey in Malawi – based on a bilateral project in Malawi.

We hope these examples of "good practice" will help researchers on future projects, and be an antidote to the view that statisticians only say what was not possible or what was done poorly. We are very grateful to the DFID researchers and the private firm in Bangladesh who all agreed to our use of the materials without constraints.

Case studies for on-farm and participatory activities

Supporting research used to be considered as largely support for the design, management and analysis of experiments or surveys.

Agricultural research and research in many other areas now involves a much richer range of methods and tools. The use of a new method usually includes related challenges for statistical support. Hence, supporting research is a much broader subject than before.

In agricultural research, two such areas are on-farm research and participatory approaches to research. They are related, because many on-farm studies include a participatory element. Materials are provided both for on-farm and participatory work in their own right, and to help the provision of support for researchers who move on farm, but try to do their research through out-of-date "on-station eyes".

The "case studies" are as follows:

Study	Title	Type and size
1	Poverty and livelihood dynamics	Participatory methodology including "poverty ladder" and case study in 14 villages in Kenya
2	Qualitative module for the Uganda National Household Survey III, 2006/7	Participatory component in 36 villages integrated with Uganda poverty survey
3	Studies from ICRAF set	ICRAF Case studies 2,6,7,11,13,15 are on-farm trials
4	SSC case studies and resources	Good practice case studies 3 and 4 (see above) are on-farm. So are various guides on experimental design
5	Designing participatory on-farm experiments	Training resource (139pps) with extensive supporting materials
6	Analysis of on-farm trials	Two papers with data sets showing alternative methods of analysis for "ordinary" data and for ranks and scores

In this list, studies 1 to 3 are "ordinary" case studies. Studies 4 to 6 provide resources for training courses to support researchers in working imaginatively in their on-farm and participatory work.

Climate change

Ask farmers or researcher about the key problems they face and they will usually include the vagaries of rainfall in their answer. Partly for this reason, most agricultural institutes include a Met station where daily rainfall, temperature and other elements are recorded.

These data often exist for many years. In the past they have often hardly been used, but this is changing with the current recognition of the importance of climate change.

One example of the importance of climatic information is that agricultural experiments are often repeated for two or three years. If asked why, the researcher will usually state that the main reason is the variability of the rainfall, and hence planting dates, so the robustness of the system demands multiple years of experiments.

To a statistician the fact that a sample of two years is deemed fine, while one year is not, should be a surprise. Also a surprise is the seeming lack of interest by the researcher to examine the (say) 50 years of climatic data, to put the one (or two) years of experimental data into a long-term perspective.

We can now capitalize on the current importance of climate change to include analyses of climate variability also. One reason is that strategies adopted by farmers (and researchers) must first cope with the current climate variability, for this to be a springboard for them to adapt to future climate change.

Climatic records are an example of "monitoring data". Unlike a survey or an experiment, the original collection of the data was not for particular research objectives. The data can be used for many objectives and these muse be stated before an effective analysis can be undertaken.

The primary data at many stations is daily. Some stations collect hourly records, and recently automatic stations allow data to be collected at even shorter interval. Daily data is often adequate for many applications.

Data are sometimes summarized on 10-day, or monthly basis. For temperature data these summaries may be adequate but for rainfall this then omits important information on, for example, dry spells. The case study data we have provided is almost all on a daily basis.

Four types of resource are provided:

The first is specific case studies and two are included in this initial set.

- Climate change in Southern Zambia.

- The use of solar cookers, including hourly sunshine data from Botswana

- A third set, on the scope of groundnut production in Malawi is described in the next section

- A current ASARECA project will provide more case studies in 2010.

A second resource is daily climatic data. Records are provided for many countries. In some countries just the daily rainfall data are provided, while others include daily temperatures and some other variables. These are typically for 50 years. We are very grateful to the individual National Met Services, who have provided the data and given permission for them to be used for training purposes.

The third resource is software, articles and other reports that describe how the data can be processed and their importance in agricultural research. Information on these resources is included with the daily data.

The final resource is a teaching session titled "climate variability and change" from the SADC training resource. This was initially intended for trainees in official statistics, but it is specifically concerned with agricultural applications. The learning objectives for this session are stated as follows:

"At the end of this session, you will be able to:

- Explain why climate issues are included in this statistics course

- Relate issues of climate change and climate variability to risk assessment and productivity in agriculture

- Understand the data requirements for a climate variability/change problem concerning planting dates

- Interpret a graph of planting dates to begin to assess the evidence for climate change

- Explain what role, if any, they can play in the climate debate"

Link to files somewhere

Other studies and resources

Case studies are just one type of resource that can enrich statistics courses. We hear mention three others.

- We have collaborated closely with Massey University, New Zealand, who produce a series of interactive statistics textbooks, called CAST. They are freely available and lecturers can adapt the books to suit their own courses.
- GEAR (

We also hope this collection of case studies can build quickly. We would get more than enough if each institute using the R-Instat software provides just one.

We also seek examples of "work in progress".

The one case study provided here is in that category. It is a Canadian-funded project, based in Malawi.