SDP

SMALLHOLDER DAIRY PROJECT

A Manual for Estimating Cattle Populations

DESIGNED FOR THE HIGHLANDS AND HIGH POTENTIAL DISTRICTS OF KENYA

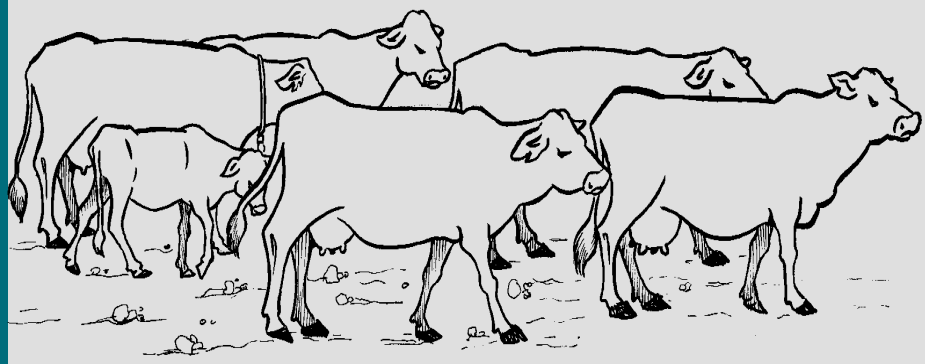
J. Nyangaga, F. Wanyoike, D.M. Mwangi, A. Wokabi,
T. Lore, M. Kembe and S. Staal

M oLFD

KARI

ILRI
INTERNATIONAL LIVESTOCK RESEARCH INSTITUTE

DFID Department for International Development



MAY 2005

The Smallholder Dairy Project

The Smallholder Dairy Project (SDP) carries out research and development activities to support sustainable improvements to the livelihoods of poor Kenyans through their participation in the dairy sub-sector. SDP is jointly implemented by the Ministry of Livestock and Fisheries Development (MoLFD), the Kenya Agricultural Research Institute (KARI) and the International Livestock Research Institute (ILRI). The project is led by the Ministry with primary funding from the UK Department for International Development (DFID). The three organisations work with many collaborators, including government and regulatory bodies, the private sector and civil society organisations.

Key areas of SDP research and development activities are:

- Analysis of factors constraining the competitiveness of smallholder dairy farmers, including farm constraints, markets and infrastructure, and information services.
- Analysis of policies and institutions affecting the dairy sub-sector, and provision of resulting information to support planning needs of stakeholders and policy-makers in the dairy sub-sector
- Analysis of social benefits of smallholder dairy production, including income, employment and child nutrition
- Participatory development of improved dairy farm technologies, such as improved fodder plants and feeding strategies
- Development of appropriate technologies and strategies for small scale milk and dairy product traders
- Development of extension and training materials to support smallholder farmers and small milk traders, and the development agencies serving them
- Spatial analysis of dairy systems for improved targeting of technology and investment

By combining the research capacity of KARI and ILRI with the experience and networks of the Ministry, SDP has been providing high-quality and wide-ranging research information to support smallholder dairy farmers, market agents, stakeholders and policy-makers since 1997.

A MANUAL FOR ESTIMATING CATTLE POPULATIONS

DESIGNED FOR THE HIGHLANDS AND HIGH POTENTIAL DISTRICTS OF KENYA

MAY 2005

J. Nyangaga, F. Wanyoike, D.M. Mwangi,
A. Wokabi, T. Lore, M. Kembe and S. Staal

Smallholder Dairy (R&D) Project
PO Box 30028, Nairobi, Kenya

Ministry of Livestock and Fisheries Development (MoLFD)
PO Box 30028, Nairobi, Kenya

Kenya Agricultural Research Institute (KARI)
PO Box 57811, Nairobi, Kenya

International Livestock Research Institute (ILRI)
PO Box 30709, Nairobi, Kenya

Department for International Development-UK (DFID)
PO Box 30465, Nairobi, Kenya

A manual for estimating cattle populations
Designed for the highlands and high potential districts of Kenya

J. Nyangaga, F. Wanyoike, D.M. Mwangi, A. Wokabi, T. Lore, M. Kembe and S. Staal

ISBN 92-9146-170-9

Correct citation: Nyangaga J., Wanyoike F., Mwangi D.M., Wokabi A., Lore T., Kembe M. and Staal S. 2005. *A manual for estimating cattle populations: Designed for the highlands and high potential districts of Kenya*. Smallholder Dairy (R&D) Project, Nairobi, Kenya.

ACKNOWLEDGEMENTS

The authors and the entire project team wish to extend their gratitude to all farmers and extension staff for their co-operation during the various field visits to collect data that has been used to develop this manual.

Appreciation goes to the numerous field staff who participated during actual surveys and in various feedback forums where their contributions of real-life experiences and possible scenarios were significant inputs to the practical approach that this manual strives to adopt.

The authors wish to express their appreciation for the collaboration within the Smallholder Dairy Project involving the three implementing institutions–the Ministry of Livestock and Fisheries Development, the Kenya Agricultural Research Institute and the International Livestock Research Institute. This publication is an output from a research and development project funded by the United Kingdom Department for International Development (DFID) for the benefit of developing countries. The views expressed here are not necessarily those of DFID.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND ACRONYMS

AEZ, AEZs	Agro-ecological zone(s)
ASAL	arid and semi-arid lands
CD	compact disc
CV	coefficient of variation
DFID	Department for International Development, UK
GDP	Gross Domestic Product
hh, hhs	household, households
KES	Kenya Shillings
MoLFD	Ministry of Livestock and Fisheries Development
SDP	Smallholder Dairy Project
Sub-loc	Sub-location

INTRODUCTION

There is continued and ever-increasing need for accurate information on cattle populations in Kenya. Livestock population data are the basis for calculating milk and meat production, understanding the relative importance of livestock in the rural economy and estimating the contribution of livestock to the national gross domestic product (GDP). Cattle population and distribution data are also critical for helping to inform decisions on where and how to target public and private investment in livestock development. In spite of these needs for information, it is not clear if there has ever been a national livestock census and studies have indicated that the currently quoted cattle populations may be inaccurate. The official source of livestock population figures is the Ministry of Livestock and Fisheries Development (MoLFD). The figures are provided by extension officials who use diverse and sometimes imprecise methods to estimate cattle numbers in their assigned areas. The increasingly few frontline extension staff in MoLFD are constrained by limited resources that hinder their ability to travel widely within their locations to establish the actual situation on the ground. This increases the probability of errors in their approximations of cattle numbers.

Surveys by the Smallholder Dairy Project (SDP) suggest that in some zones, dairy cattle populations are actually four times the numbers reported by official sources. For detailed information on these surveys and their implications on estimates of Kenyan cattle populations, see the SDP Policy Brief *"Counting cows: how many really are there in Kenya?"* and the associated SDP report *(Cattle numbers in Kenya: the need for a good method to determine the population)*. Because of these apparent discrepancies, SDP was requested by the MoLFD to produce a simple, practical manual to guide field staff in estimating cattle populations, hence this document.

Livestock production in the highlands and high potential districts of Kenya occurs mainly on numerous small farms with more or less permanent settlement due to the prevalent land tenure and livestock production systems. Counting cattle in these areas requires numerous individual farm visits, demanding logistical support beyond the resources normally allocated to the local government offices. This manual presents a simple and generally inexpensive method for investigators (interested government officials) to estimate cattle numbers in such areas by using statistical sampling procedures. The method entails counting cattle in fewer farms and using the results to derive the population for a larger area. Cattle figures derived using these procedures are expected to be reliable. Because households and cattle in arid and semi-arid lands (ASAL) of Kenya are often not sedentary, different methods may be required for estimating cattle populations in those areas. This manual is thus not recommended for application in ASAL zones.

THE PROCEDURE

This manual is designed for the Kenyan administrative boundary system where the smallest area units are sub-locations, which make up locations, then divisions, districts, provinces and finally the whole country. Generally the procedure involves selecting sub-locations that are representative of a larger target area and collecting data from a random sample of farms or households in the chosen sub-locations. The data are then analysed and the results used to project the cattle numbers to the larger target area.

For the time being, it is recommended that projections from the sub-location sample data be made to the divisional and district levels only. Beyond this, data from a few sub-locations may not adequately capture and effectively represent the diversity expected in a very large target area. The larger the target area, the greater the expected variation in factors that influence distribution of cattle, and the presence of unstable systems such pastoral lifestyles and transhumance will make the estimation even more unreliable.

The method of estimating cattle population presented in this manual involves three main steps:

- sampling (selecting representative area units and sample farms)
- actual data collection and
- using the sample statistics to project the cattle population of the target area.

The whole process requires the participation of frontline field staff as they are well-placed to obtain and provide information about the sample area units.

PART 1

SAMPLING

1 SAMPLING

Sampling should ensure a representative and sufficiently large sample size for computation of estimates (mainly the mean number of cattle per farm) needed to reliably predict the cattle numbers of the target area. The number of sample farms should however not be unnecessarily large, so as to economise on resources. It is important to consider the variation in cattle numbers in the larger target area. The types and geographical distribution of cattle in Kenya vary mainly according to factors such as agro-ecological zones (AEZ), human population density, production systems and milk market access. The difference in cattle populations in the target area caused by these factors should be considered when selecting the sample sub-locations.

Selecting sample sub-locations

Step 1
Obtain a list of sub-locations to be sampled based on the variation in cattle numbers in the target area.

All the sub-locations in the target area should be distributed so that the sample accounts for most of the variation in cattle numbers in the larger area. This requires that all the sub-locations be grouped into sets with similar features (clusters) based on the above-mentioned factors affecting cattle distribution. Table 1 shows how to define the clusters based on three levels of each of three factors. The levels refer to relative measure of each factor. Using a higher number of clusters will account for most of the variation in cattle numbers in the larger target area. Each cluster should consist of sub-locations that have the same level of the specified factors.

Table 1. Identifying clusters based on three levels each of three factors that influence cattle distribution.

Production system	AEZ	Milk market access		
		Good	Medium	Poor
Fully zero-grazed	High potential	Cluster 1	Cluster 2	Cluster 3
	Medium potential	Cluster 4		Cluster 5
	Low potential	Cluster 6	Cluster 7	Cluster 8
Semi zero-grazed	High potential	Cluster 9	Cluster 10	Cluster 11
	Medium potential			
	Low potential	Cluster 12		
Fully open-grazed	High potential		Cluster 13	
	Medium potential			Cluster 14
	Low potential		Cluster 15	Cluster 16

Indicate names of sub-locations that belong to each cluster

Some cluster descriptions may not have representative sub-location

Production system broadly refers to the method of keeping cattle: fully zero-grazed, semi zero-grazing and extensive systems characterised by fully grazing in open spaces and rangeland.

Agro-ecological zone (AEZ): This is a combination of agro-climatic potential (rainfall, temperature etc.), soils and altitude.

Milk market access refers to the ease with which farmers are able to sell their milk. Market factors to consider include the distance to nearest buying, urban or shopping centres, or how long the farmer takes to deliver milk from farm to buyer. Milk price can also be an indicator.

Write the names of sub-locations in the spaces provided in Table 1 or make a list of sub-locations for each cluster description using the guide shown in Table 2. Based on the example in Table 1, sub-locations in cluster 1 would be described as: high-potential with good market access where all the cattle are fully zero-grazed; cluster 7 will contain sub-locations with low agro-climatic potential, medium market access and cattle that are fully zero-grazed etc.

Table 2. Grouping sub-locations based on cluster descriptions

Cluster number or name:	Cluster 1	Cluster 2	Cluster 3	Cluster ...
Description:	High potential, good market access, fully zero-grazed cattle	High potential, medium market access, fully zero-grazed cattle	High potential, poor market access, fully zero-grazed cattle	***
Sub-locations:	Sub-location 1 Sub-location 2	Sub-location 3 Sub-location 4 Sub-location 5	Sub-location 6	Sub-location 7 Sub-location 8 Sub-location 9 Sub-location 10 Sub-location 11

Selection can either be random or based on the ease of operation (logistics)

Before selecting the sub-locations to be sampled, verify that each cluster has at least one sub-location. Since the sub-locations in a specific cluster are similar with respect to the cluster description, the sub-locations to be sampled from a cluster can be determined either randomly or based on the logistical ease of operation (actual data collection). The number of clusters–and hence the number of sub-locations–sampled should be weighed against resources available; the greater the number of clusters, the more representative the sample but the more expensive the exercise.

Selecting sample farms

The terms 'individual farms' and 'households (hhs)' are used interchangeably in this manual. Data on household numbers can be obtained from the national census reports but these are compiled once every 10 years and current figures will need updating. On the other hand, most agricultural field offices base their agricultural extension activities on farm holdings or units. These offices have some data on current numbers of these units down to village level. The investigator should choose either 'households' or 'farm holdings/units' and consistently use the respective total baseline numbers in the sub-locations, sample selection and subsequent calculations.

Step 2
Determine the number of sample farms to be visited in each sub-location selected

The number of households or farms to be visited in a sub-location should be based on the expected variation in herd sizes (number of cattle) per farm in that area. The expected level of variation usually depends on the prevalent production systems and is best provided by staff familiar with the sub-location. The expected disparity in these two factors will be used to derive a coefficient of variation (CV). To obtain the CV, the field staff will score the expected relative uniformity of the production systems and herd sizes for each sub-location. Using Table 3 as a guide to deriving the variation score, indicate the degree of uniformity as follows:
1 = uniform, 2 = moderately different, 3 = diverse.

Table 3. Scoring of variation in a sub-location based on two factors affecting cattle distribution

Factors	Production system*	Expected number of cattle per farm	Variation score (Sum of scores)
Score			

* as defined in Table 1

Step 3
Use the variation score to select the coefficient of variation (CV) from Table 4

Table 4. Coefficients of variation to be used to calculate the number of sample farms based on the variation score

Variation score (Total)	Coefficient of Variation
2	50
3	60
4	70
5	80
6	90

Step 4
Calculate the minimum number of sample farms from each sub-location

Use the value of the CV in Formula 1 to calculate the sample size.

Formula 1:

$$n = 2 \times \left[\frac{A \times B}{C}\right]^2$$

Where:

n = The minimum number of sample farms required in each sub-location for a good estimate of the mean number of cattle kept per farm

A = Desired level of confidence when estimating the mean number of cattle kept per farm. For our surveys we use 1.96 for the 95% level of confidence.

B = The coefficient of variation obtained from Table 4.

C = Desired level of precision in estimating the mean number of cattle kept per farm. For our surveys we use 20% (i.e.± 20%) of the population mean.

Example:

Assume that we wish to estimate the mean number of cows per farm in a sub-location with a precision of ± 20% and a 95% level of confidence. Assume also that the coefficient of variation of cattle herd sizes across farms has previously been estimated to be 80% based on the variation score read from Table 4.

The required number of sample farms from the sub-location can be estimated using Formula 1 as:

$$n = 2 \left[\frac{1.96 \times 80}{20} \right]^2 = 122.93$$

This can be rounded up to 123 farms.

If the number of farms per sub-location is too low or too high, discretion can be used to place a minimum or maximum limit provided that the total from the sub-location is not lower than the calculated minimum sample size.

Note on additional number of sub-locations and farms

The procedure described above shows how to calculate the minimum number of sub-locations or farms necessary to get a specified degree of accuracy. Investigators are at liberty to adjust this number but should only do so upwards, that is, by sampling more sub-locations or farms. However, the decision to increase the sample size must be weighed against resources allocated for the exercise. Any extra precision expected may not be worth the additional cost (staff, days, funds). The coefficient used for the 95% level of confidence (1.96) provides an acceptable margin to accurately estimate livestock population. You could use a higher level of confidence (e.g. 99% with a coefficient 2.56), but the larger sample size, extra farm visits and difference in accuracy may not be worth the resulting additional costs. However, additional households will come in handy if the investigator expects to deduct urban household types from the total during the final analysis (Refer to questionnaire administration and format in Appendix 1).

Step 5
Select the farms to be visited

This is the practical field phase of the survey. Ensure that the farms to be visited within the sampled sub-location are randomly selected. You can use a numbered list of all farms in the sub-location and visit a random sample of those but such lists may not be readily available or contain all farms in existence at the time of the survey. A good alternative is to use systematic random sampling, which involves sampling at specified farm/household intervals. The interval will depend on the total number of farms or households in the sub-location and the sample size (number to be visited). A rough guide is to divide the total number of households in the sub-locations by the number of households to be visited. For instance, if there are 100 households in the sub-location and 20 are to be sampled, then you would sample every fifth household encountered.

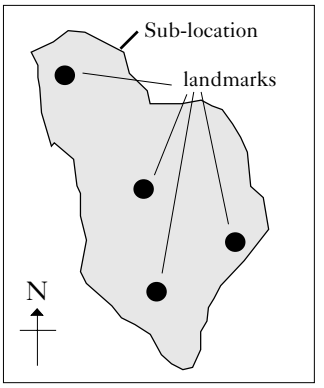
Figure 1 illustrates a guideline for selecting households or farms in a given sub-location.

Note on additional transects:

You may add transects to areas likely to be missed by initial transects. You can deliberately draw additional transects on parts of the map where the households are located. However, make sure the visits randomly target households/farms and do not attempt to specifically target areas where there are cattle. Once identified, the targeted farms should be visited whether they have cattle or not. Likewise, one can choose not to draw transects through strictly urban residences (e.g. towns or cities) containing households that are not likely engage in livestock farming. However, in some districts rural-urban or peri-urban households have a considerable presence of livestock that should be counted. In such cases, ensure that there are transects passing through these areas, visit the households but mark them for possible elimination in the data analysis (Refer to questionnaire administration and format in Appendix 1) depending on factors like proportion of households with cattle, mean number of cows per household etc.

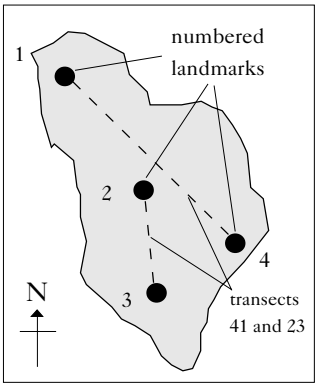
Figure 1. Procedure for random selection of sample farms within a selected sub-location

Draw a map of the selected sub-location and indicate landmarks (e.g. bridges, hills, dams, churches, schools etc.) as evenly and widely distributed over the area as possible

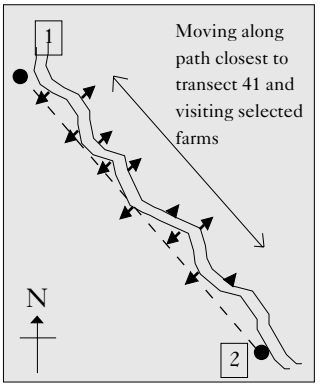


Assign each landmark a unique number

Randomly select a pair of numbers assigned to the landmarks. This can be done using any convenient method: picking pieces of papers in a bowl or hat, random numbers (from tables, calculators) etc.



Each pair of numbers or landmarks gives a transect line. Make the transects by drawing straight lines on the map between the pairs of landmarks selected



Selection of farms/households:

Using one transect as a guide, move along the nearest passable path or road from one landmark to the other and systematically select and visit farms along the way

Move to another transect and visit more farms until you reach the total number to be sampled in the sub-location. Do not visit the same farm more than once

Use not less than two transects per sub-location

PART 2

DATA COLLECTION

2 DATA COLLECTION

Support from local administration

From the very beginning, the support and assistance from local administration officials (chiefs and their assistants, village elders, councillors etc.) is very useful. These officials and their staff assist as follows:

- They are aware of the latest area boundary changes and locations of useful landmarks hence can provide good references when drawing maps and indicating landmarks
- The local chief's office is invaluable in publicising the objectives of the exercise and the programme of activities in forums such as *barazas*, meetings etc.
- Most rural people are suspicious of strangers and unusual events, especially counting of livestock. A representative from the local chief's office accompanying an enumerator not only eases the people's suspicions but also ensures security to both the enumerator and household being visited. This, however, may have an effect on mode of transport and other resources (See Appendix 2 on budget items).

Household identification

Right from the outset of the survey, the enumerators and the investigator must clearly understand what is meant by households or farm holdings. Data collection and analysis will be based on the numbers of these data units. A household is usually identified as a domestic unit consisting of the members of a family who live together, sometimes with non-relatives such as workers. While in most cases identifying a single household or farm holding unit is easy and straightforward, there is need to clarify some ambiguities. Figure 2 contains some frequently asked questions or situations reported by enumerators in relation to this and guiding concepts on how to tackle each situation.

Figure 2. Questions and situations reported by enumerators during cattle counting surveys and the guiding concepts for decision-making

Question: The transect passes through a place where there are no households. (e.g. through a forest, dam etc.) or through a shopping centre or town. What should I do?

Question: I was instructed to select a farm on one side, then the next farm on the other side of the transect route. But in my area, all the households are positioned only on one side of the road. Do I ignore this transect?

Question: Two transects cross each other in such a way that they share the same routes (roads/paths) and homes. Which transect should be used?

Guiding concept: Transects are supposed to be guides for routes to follow. The routes can be allowed to diverge from the straight transect line but only to a reasonable extent. The investigator is at liberty to guide the selection of household or farms on any side of a transect route. The principle of randomness is to ensure that every household in the sub-location has an equal chance of being selected. When transects cross or share the same route, the enumerator should not collect data from the same farms. If a transect passes through an urban centre the enumerator should mark such households (see questionnaire format in Appendix 1) and the investigator should decide whether or not to increase the number of rural households, or include the urban households in the data analysis.

Question: Suppose the turn-off from the transect road or walk leads to several other entrances to farms, which do I pick?

Question: In some areas, once you take the entrance or the road you come across several households or family units in one compound. This happens when a household head has several wives or sons, each with a household of their own. Which one do I visit?

Guiding concept: Note that households or farms picked are those next to the transect route. Once an entrance off the transect route is identified the closest and most direct household is picked. If the units in one compound are not separate and independent, visit the main household head who should answer questions in relation to all the cattle that belong to the whole compound. If the units are separate and independent, randomly pick one and visit it. The choice must be random and not influenced by the ownership or presence of cattle.

Situation: At one household there was nobody to respond to my questions or the household occupants refused to respond to my questions.

Question: The household I visited did not have any cattle. Should I have interviewed them?

Situation: The cattle were not in the compound and I could not ascertain their breeds or the respondent said s/he had cattle but they were kept in another farm away from the household I was visiting.

Guiding concept: Data must be collected from all households that have been targeted for a visit, whether they have cattle or not. Data is collected from the identified households; if no one is at home, the enumerators can try and obtain answers from the neighbours about the target farm or choose to interview the immediate neighbour. Answers should refer to cattle belonging to the target household and kept in that sub-location even if absent. However, the cattle should not be counted if they are in another farm where they are likely to be counted separately. If the enumerator cannot see the cattle, use the descriptions offered by the respondents to establish the cattle breeds and types. If the respondents refuse to answer questions refer to the above-mentioned section on ‘Support from local administration’.

Questionnaire administration

This procedure has been developed to approximate cattle populations only, and is yet to be tested for other livestock species.[1] Depending on the objective of the survey, the questionnaire format can be altered and the same data analysis used to estimate numbers of types, breeds, sexes and age-groups of cattle in the target area. What is important is that the investigator and enumerators use a different forum or space, beside the manual (e.g. diagrams, photographs, descriptions during training) to agree on which cattle forms they aim to enumerate. However, the questionnaire format should be as simple as possible so that many farms can be covered by an enumerator in the shortest time possible. Collecting data on other livestock species or cattle types will mean more time in individual farm visits, resulting in longer periods for the survey and extra costs.

Appendix 1 illustrates an example of a simple questionnaire format. The table at the top is useful if the investigator is covering a large area and is likely to end up with very many questionnaires. The transect number refers to the pairs of digits that connect the landmarks. The household number helps ensure all the farms to be visited in a given area have been covered. The name is not necessary but it can be indicated to help the enumerator retrace the household. The third and fourth columns of the lower table are useful if enumerators are visiting households in shopping or urban centres, which may or may not be considered in the final data analysis, depending on their effect on the final average. If a farm has cattle, the enumerator must indicate the number of cows in either of the last two columns. If there are none, the enumerator must indicate a zero (0), not a blank or dash. If the survey aims to determine populations of the different types of cattle (e.g. dairy and zebu in last two columns), the investigator and enumerators must agree on the exact features that describe or distinguish the types.

---

[1] *Although field offices will take advantage of the survey to collect similar data on other livestock species, the sampling procedures and coefficients of variation provided were based on findings related to cattle populations. They may not be appropriate for other species.*

Enumerators should fill the questionnaire in pencil but complete all necessary corrections before submitting the filled questionnaires to the investigator. When collecting the completed forms, the investigator will use discretion to re-visit a few households to confirm the enumerator's visits and verify the data collected. However, this will depend on the resources available. While collecting the completed questionnaires, the investigator must check that all figures are clearly written and clarify any smudges or crosses. In particular, confirm with the enumerator what is meant by any blank spaces that should have been filled.

Logistics

The number of farmers that can be visited by an enumerator in a single day will depend on the distance that must be travelled, the nature of the terrain, weather and mode of transport. One enumerator can visit 30 to 90 farms a day on foot or by bicycle, motorcycle or public transport, depending on resources (transport, fuel etc.) available for the exercise. Arrangements should be made to ensure the exercise takes the shortest time possible or is carried out within a period that does not allow changes over time to affect the results.

PART 3

CALCULATIONS: USING SAMPLE STATISTICS TO PROJECT CATTLE POPULATION

3 CALCULATIONS: USING SAMPLE STATISTICS TO PROJECT CATTLE POPULATION

Note that the overall population in the target area (location, division or district) will depend on the influence of proportion of farms and population means from each cluster on the whole target area. Thus the focus of the calculations is summaries of data carried out for each cluster. The investigator will collect the completed the questionnaires and enter the data in tables that allow for the necessary summaries and analysis. It is important to go through the formulas below to know which summaries should be made. Depending on resources available one can use square ruled forms and summarise the data with the aid of a simple hand-held calculator or work with spreadsheets in a computer (e.g. Microsoft Excel). This manual will also be accompanied with a simple programme (on CD) that can be used to fill in the raw data and carry out the calculations.

Formula 2:

The total number of cattle in a target area (location, division or district) will be projected using:

$$\text{Total cattle number} = DxE/100xF$$

where:

Total cattle number = Projected number of cattle in a target area (location, division or district)

D = Total number of households or farm holdings in the target area

E = Weighted percentage of households keeping cattle

F = Weighted mean number of cattle per farm in the farms with cattle

Formula 3:

To get the weighted proportion of farms keeping cattle (E) use the following process:

1. For each cluster (Ref Table 2) work out the proportion of the total households to the total in the target area (G)
2. From the sub-location representing each cluster, get the proportion of households that keep cattle (H)
3. Multiply G and H

Derive the sum of the products to get E.

This can be summarised as:

$$\text{Weighted proportions of hhs keeping cattle (E)} = \sum_{i=1}^{n} (GxH)$$

where:

E = Weighted proportion of hhs keeping cattle

n = Total number of identified clusters

G = Proportion of total hhs of i^{th} cluster to total in target area

H = Proportion of farms keeping cattle in the i^{th} cluster

Formula 4:

To get the weighted mean number of cattle per farm (F) in the farms with cattle use the following process:

1. For each cluster determine the proportion of households to the total in the target area (G)
2. From the sub-locations representing each cluster, determine the mean number of cattle per farm/household in farms with cattle (J)
3. Multiply G and J

Derive the sum of the products to get F.

This can be summarised as follows:

$$\text{Weighted mean number of cattle per farm (F)} = \sum_{i=1}^{n} (GxJ)$$

where:

F = Weighted mean of cattle per farm

n = Total number of identified clusters

G = Proportion of total hhs of i^{th} cluster to total in target area

J = Mean number of cattle per farm in farms within cluster i

Notes on final calculations

Note that the weighted proportion of households keeping cattle and the weighted mean number of cattle per farm both describe the situation at a glance and can be used to quickly explain the ownership of cattle and numbers in a target area. Collected data can be applied to the same formulas to estimate the numbers of different types, breeds, sexes, ages etc. of cattle in the target area. Considering the final population figures, the investigator should check the effect of including or excluding the urban households to get a more realistic figure.

PART 4

AN EXAMPLE

4 AN EXAMPLE

To better understand the process, this section describes a step-by-step example to help understand the process of sampling and final data analysis. A survey was organised to estimate the total number of dairy and zebu cattle in Maji Matamu division. The division is moderately populated but characterised by small land holdings where most farmers practice zero-grazing. The main factors observed to affect the presence and geographical distribution of cattle in this division are market access and the agro-climate potential (AEZs). The division has 35 sub-locations.

Selection of sample sub-locations

Step 1
Determine which of the 35 sub-locations will be sampled

The division has a uniform cattle production system (zero-grazing) but there are slightly different AEZ types. The milk market access channels are extremes where for some farmers the milk is collected right at the farm while others have to walk for up to three hours to the nearest collection and shopping centres to sell their milk. Distribute the 35 sub-locations in Maji Matamu according to the cluster types described by the two factors influencing cattle population at three different levels as shown in the table below.

		Milk market access		
		Good	Moderate	Poor
Agro climatic potential (AEZ, etc)		Cluster 1	Cluster 2	Cluster 3
	High	**Sub-loc 1**	Sub-loc 7	Sub-loc 13
		Sub-loc 2	Sub-loc 8	**Sub-loc 14**
		Sub-loc 3	**Sub-loc 9**	Sub-loc 15
		Sub-loc 4	Sub-loc 10	Sub-loc 16
		Sub-loc 5	Sub-loc 11	Sub-loc 17
		Sub-loc 6	Sub-loc 12	
		Cluster 4	Cluster 5	Cluster 6
	Medium	Sub-loc 18	Sub-loc 22	Sub-loc 27
		Sub-loc 19	**Sub-loc 23**	Sub-loc 28
		Sub-loc 20	Sub-loc 24	Sub-loc 29
		Sub-loc 21	Sub-loc 25	**Sub-loc 30**
			Sub-loc 26	
		Cluster 7	Cluster 8	Cluster 9
	Low	Sub-loc 31	Sub-loc 33	(None in this
		Sub-loc 32	**Sub-loc 34**	district)
			Sub-loc 35	

As a result we observe that there are eight possible clusters of similar sub-locations and we select a sub-location from each cluster. Since there is more than one sub-location in each of these eight clusters, we use a random method to select one sub-location we will survey from each cluster. In one cluster we base our selection on the logistical costs or ease of carrying out the survey. The selected sub-location is shown in bold.

Step 2
Determine the number of sample farms

Determine the number of sample farms to be visited in each farm by considering the expected variation in number of cattle kept per farm. The frontline extension personnel indicate how varied the production systems in each sub-location are and hence how diverse we expect the herd sizes per farm to be. We use Table 3 to derive the coefficient of variation (CV) from Table 4. We apply CV value in Formula 1 to derive the number of farms to be visited as shown in the table below.

Selected sub-location	Factors' scores		Variation score (Total)	CV	Number of sample farms to be surveyed (Formula 1)	
	Production system	Expected number of cattle per farm				
Sub-loc 1	1	2	3	60	$2 \times [(1.96 \times 60)/20]^2$	= 69
Sub-loc 9	2	1	3	60		= 69
Sub-loc 14	3	2	5	80		= 123
Sub-loc 20	1	3	4	70		= 94
Sub-loc 23	2	1	3	60		= 69
Sub-loc 30	1	1	2	50		= 48
Sub-loc 32	1	1	2	50		= 48
Sub-loc 34	1	3	4	70		= 94

Step 3
Conduct the survey

This entails publicising the exercise, training the enumerators (very important for data quality), drawing maps and transects of the eight selected sub-locations, visiting and collecting data from all the farms or households indicated, compiling the completed forms and tabulating the collected data.

Step 4
Summarise the collected data and calculate summaries as shown in table below

Refer to Formulas 2 and 3 for meanings of the letters heading the columns.

Cluster	Total households or farms*	G	Selected sub-location	H	J	G x H	G x J
				From collected data		Calculations	
1	755	0.229	Sub-loc 1	72	2.1	6.50	0.481
2	160	0.049	Sub-loc 9	68	1.8	3.30	0.087
3	868	0.264	Sub-loc 14	88	3.5	23.19	0.922
4	150	0.046	Sub-loc 20	72	3.0	3.28	0.137
5	320	0.097	Sub-loc 23	71	2.1	6.90	0.204
6	425	0.129	Sub-loc 30	91	2.5	11.74	0.323
7	278	0.084	Sub-loc 32	69	2.3	5.82	0.194
8	338	0.103	Sub-loc 34	82	1.6	8.41	0.164
	D = 3294					Formula 3, E = 59.35	Formula 4, F = 2.513

*This refers to the sum of farms or households in each cluster and is derived from the sum of farms/ households in each sub-location in the cluster.

Formula 2:

Total cattle number = D x E/100 x F= 3294 x 59.35/100 x 2.513 = 6550.53

The estimated total cattle population in Maji Matamu division is 6550.

PART 5

CONCLUSIONS

5 CONCLUSIONS

The method described involves visiting a sample of households in a small area to collect cattle population data that is then used to approximate the population in a larger target area. The final figure obtained will not be the actual population but it is expected to be more reliable than the various methods often used by the field offices where arbitrary rates of changes in cattle numbers are applied to a baseline figure whose accuracy has not been authenticated by an actual census.

The fact that fewer farms are visited in the cattle-counting exercise than in an actual house-to-house census implies a lower financial requirement for the exercise hence it can be implemented even in cases where limited resources are allocated to field offices. With minimal financial support, the surveys can be incorporated into routine field activities of frontline staff, provided the exercise is completed within a reasonable time. However, the fewer household visits means that the final figures are only an approximation and in some cases, if found at extreme variance with expectations, the difference may call for reasoning and interpretation. The method, however, cannot replace the accuracy expected in an actual census and SDP hopes that the urgent need for this on a national scale is addressed.

PART 6

APPENDICES

APPENDIX 1
A questionnaire format used to count dairy and zebu cattle in a survey by SDP

District:	Enumerator:
Sub-location:	Date: _____ Sheet Number: _____
Supervisor checking:	Any remarks:

Transect number	Household number (or name)	Type of household		Do you keep any cattle?	If cattle present, indicate Total number of ALL cattle (all ages, both sexes)	
		1 = Rural 2 = Rural-urban 3 = Urban 4 = Institution	If 2, 3, or 4 Name of urban centre or institution	1 = Yes 2 = No	Total dairy cattle (>50% dairy phenotype)	Total zebu cattle (<50% dairy phenotype)

APPENDIX 2A
A guide to budget items to be considered when preparing for the survey

* Note that routine operational resources and funds are limited and not all the items appearing in the tables below must be included in the budget. The rates may also be negotiated so that the overall cost is within affordable limits, given what will be allocated either for the exercise only or as part of the investigator's routine expenses. It is therefore important for the investigator to agree with the field staff what financial support to anticipate.

1. Preparation stage: Training, drawing maps and transects

	Number of days	Number of persons	Rate (KES)	Total (KES)
Hall hire				
Stationery: *Pens, pencils, notebooks, flipcharts etc.*				
Stationery: *Photocopying and printing*				
Teas etc. (morning and afternoon)				
Lunch				
Dinner				
Accommodation				
Transport refund*				
Transport repairs and services (*For vehicles to be used*)				
Other expenses				
Sub-total for preparation stage				

* Transport costs and re-funds are monies spent on fuel and oils or bus ticket re-imbursements.

2. Data collection stage

Names and titles of staff involved	Number of days	Night out	Break fast	Lunch	Dinner	Transport*	Total
		Rate in KES					
District supervisor							
Driver							
Enumerator 1							
Enumerator 2							
....							
Chief's assistant 1							
Chief's assistant 2							
....							
Stationery: *Pens*, *pencils*, *notebooks*, *flipcharts etc.*							
Telephone airtime etc.							
Sub-total for data collection stage							
Total							
Contingency (10% of total so far)							
Grand total							

1. Transport costs and re-funds are monies spent on fuel and oils or bus ticket re-imbursements.

APPENDIX 2B
Approximate cost of survey conducted either in Mathira, Nyeri or Kilibwoni, Nandi

* Note that the rates indicated below were those offered by SDP in 2004.

In both divisions, the district supervisor used five enumerators, meeting them thrice (twice before and once after the farm visits). The enumerators visited farms for five days during which the supervisors went to the field for three days, using a government vehicle. The enumerators visited 1103 and 1975 households in Nandi and Nyeri, respectively.

1. Preparation stage: Training, drawing maps and transects

	Number of days	Number of persons)	Rate (KES)	Total (KES
Hall hire	3	7	500	1500
Stationery: *papers, exercise books, pens*				1000
Stationery: *Photocopying and printing*				500
Teas etc. (morning and afternoon)	3	7	100	2100
Lunch	3	7	500	10,500
Transport refund	3	5	500	10,500
Transport repairs and services		5	2000	10,000
Sub-total for preparation stage				36,100

2. Data collection stage

Names and titles of staff involved	Number of days	Night out	Break fast	Lunch	Dinner	Transport	Total
		Rate in KES					
District supervisor	3			500		1500	6000
Driver	3			350			1050
Enumerator 1	5		350	500		500	6750
Enumerator 2	5		350	500		500	6750
Enumerator 3	5		350	500		500	6750
Enumerator 4	5		350	500		500	6750
Enumerator 5	5		350	500		500	6750
Chief's assistant 1	5			200			1000
Chief's assistant 2	5			200			1000
Chief's assistant 3	5			200			1000
Chief's assistant 4	5			200			1000
Chief's assistant 5	5			200			1000
Stationery:							2000
Telephone airtime etc.							1000
Sub-total for data collection stage							48,800
Grand total							84,900

SMALLHOLDER DAIRY PROJECT

S D P
(R&D)

MoLFD - KARI - ILRI - DFID

M oLFD

KARI

ILRI
INTERNATIONAL
LIVESTOCK RESEARCH
INSTITUTE

DFID Department for International Development