# Namibia Census of Agriculture 2013/2014 (NCA 2013/2014)

## Methodology

### Introduction

The agricultural sector in Namibia consists of the commercial and the communal agricultural sectors. Commercial agricultural sector covers large farms which are mainly concentrating on the livestock production for the commercial purposes. In the case of communal sector a large part of the population in the northern regions (Kavango East, Kavango West, Ohangwena, Omusati, Oshana Oshikoto, Zambezi, and Kunene) are involved in subsistence farming activities which include crop and livestock farming. In the remaining regions also one may find subsistence farming activities mainly livestock farming but not in the same magnitude as the northern regions. There is a possibility that some subsistence farmers may have converted themselves to commercial farming activities but the numbers are unknown and may be still small. Government also has a program where commercial farms are being bought for resettlement. The resettled farms may contain small scale commercial farming activities and /or subsistence farming activities.

### Census of Agriculture design

#### Objectives

The *immediate* objective is to obtain baseline agricultural production and structural variables that do not change rapidly from year-to-year in both the communal and commercial farming sectors at the national and regional levels. More specifically, up-to-date and more reliable data on the numbers of agricultural holdings, on land areas, crop production and livestock numbers are to be collected.

The *long-term* objective of the project is to contribute to improved planning and decision-making in the agricultural sector and satisfy the information needs of the socio-economic database being set up by the NSA. Another important objective is to provide a sampling frame for subsequent agricultural surveys and other sample surveys on agricultural holdings. Equally of importance is that the agriculture census will provide data to estimate future trends/changes in agricultural behaviour through statistical projection models.

#### Approach

##### Target population (Universe)

The *universe* for the NCA 2013/2014 is all the agricultural households engaged in both commercial and communal farming activities. Since the purpose and the activities carried out in the two sectors are quite different the universe could be considered as consisting of the agricultural households in commercial and communal sectors separately.

##### Commercial sector agricultural households

Commercial farms in the commercial agricultural sector are operated and managed by farmers who use modern methods for their production activities and because of the commercial nature of their activities they can be identified. Therefore a complete list of such farmers is already available or an available list could be updated so that a complete list of such farmers is available. They can be reached by post or may be through e mail. Hence the agricultural households in the commercial agricultural sector will be covered using a complete enumeration through postal enquiry. Face to face interviewing method will be utilized in the case of non-responding households to reduce the overall non response.

##### Communal sector agricultural households

The agricultural households in this sector are mainly engaged in subsistence agriculture. However emerging commercial farming activities may be found in small numbers. A list of subsistence farmers is not available and to make such a list is also prohibitively expensive. Hence it is impossible to follow the above method to cover the subsistence farmers. Therefore it is necessary to meet the subsistence farmers face to face and collect the required data from them. Two alternatives are open, one to cover all of them or to cover a part of them. Considering the available human and financial resources it was decided to use a larger sample of subsistence farmers to cover the communal sector. Therefore Instead of the full coverage as in the commercial sector, a sample survey will be planned to cover the communal sector.

### Design for the communal sector sample survey

#### Administrative structure of Namibia

Namibia’s administrative structure consisted of 13 political regions and 107 constituencies. Each region is headed by a Governor and a councilor is appointed for each of the constituencies. This was the situation for the most part of the planning process of the NCA2013/14. Recently, based on the recommendations of a delimitation commission, the President declared changes in this structure. According to this presidential declaration there are 14 regions in Namibia and 121 constituencies. Therefore the NCA2013/14 will be carried out under this new administrative structure. Refer appendix 1 for the new regions and constituencies.

#### Unit of analysis and unit of selection

Based on the objectives of the survey the unit of analysis is the agricultural households and the agricultural holders who are operating the agricultural activities within these households. Hence the unit of selection will also be the agricultural household.

#### Target population (universe) for the communal sector sample survey

The target population for the communal sector sample survey is all the agricultural households in the rural communal areas of Namibia including the semi urban areas around the urban centers.

#### Survey population

The households in the developed part of the municipalities, towns, villages and settlements are excluded from the communal sector survey. Such households may have home gardens or small scale livestock farming but for the purpose of this survey they are excluded. The areas around these urban centers which are undeveloped as the urban part but within their town land boundaries will be included in the survey to cover any agricultural activities occurring in those areas.

Out of the 14 regions the 7 northern regions (Kavango East , Kavango West, Ohangwena, Omusati, Oshana, Oshikoto, and Zambezi) predominantly have such agricultural households. In these regions large commercial farming activities are not found except in the Guinas and Tsumeb constituencies in Oshikoto region. These are also the regions where the crop cultivation activities take place to a large extent along with the livestock farming for subsistence purposes.

Kunene region in the North West has both the communal and commercial farming activities where still a larger part is covered by communal farming activities. Livestock farming is the major communal farming activity although the crop cultivation activities could be also found to a lesser extent.

The remaining 6 regions are more concentrated on the commercial farming activities. They also have communal farming activities, mostly on livestock farming but the coverage is relatively less. The crop cultivation activities are very limited in these regions.

#### Sample design

The ultimate unit of analysis of NCA2013/2014 is the agricultural households/holdings and the holders and other members of such households. Therefore it is necessary to select sample agricultural households for the collection of data and such selection procedure needs a list of agricultural households covering the whole country and also different sub population groups of interest. Such lists covering the target population are not available in Namibia. Hence to overcome this difficulty the approach will be to adopt a cluster sample design where the sample will be selected in stages to arrive at a sample of agricultural households. Such designs are generally called multi stage sample designs and the specific design to be used in NCA2013/2014 is a two stage sample design where the first stage is a selection of clusters of households followed by a selection of agricultural households within the selected clusters. Further stages may have to be used to arrive at the sample plots for crop cutting experiments for the purpose of estimating the crop production. Hence the sample design for the NCA2013/2014 is a stratified two stage cluster sample design where the first stage units are geographical areas designated as the PSUs and the second stage units are the agricultural households.

### Sample frame(s) and the size measure

#### Primary sampling frame

A prerequisite of a multi stage design is a complete list of small geographical areas (clusters of households) consisting of a set of dwelling units/households of a given size covering the whole country without duplication and omission. Such small geographical areas called Enumeration Areas (EAs) were developed for the national population and housing census 2011 (PHC 2011). Hence such a list of EAs can be used to select sample EAs as the first stage sample units through which a set of sample households could be accessed. This list of EAs which serves as the starting point of sample selection for national household surveys is called the National Sample Frame. These first stage units comprising of EAs are called Primary Sampling Units (PSUs). Therefore this list of PSUs is known as the sample frame.

When a multi stage sample design is adopted, the first stage sample (area sample) is determined based on their sizes. In general, in the household surveys, the size measure used is the number of households or number of dwellings within the given PSU based on the recently concluded population and Housing Census (2011 Population and Housing Census – 2011PHC). Three questions were included in the 2011 PHC questionnaire about the own account agricultural activities carried out by the household. Based on these questions it is possible to classify a household as an agricultural household. Hence for the purposes of agricultural censuses and surveys, the number of agricultural households within a PSU can be taken as the size measure instead of all households. This will be a better size measure because it is much more closely related to the agricultural activities than the normal households.

Since some PSUs have small number of agricultural households they have to be merged with the adjoining ones to make PSUs of a reasonable size (more than 40). If there are any large ones they may have to be split also to make them smaller to be reasonable in size. Hence this sampling frame will be different from the general sample frame. This frame will then be used for NCA 2013/14. It will also serve as the basis for all the agricultural surveys in future.

In the southern regions there will be problems to apply this principle since some PSUs will be quite small in size. A reasonably sized PSU could be very large in geographic size. Because of this reason smaller PSUs (smaller in size) may have to be accepted in the southern regions.

There are recent developments (delimitation commission report), which has affected the administrative boundaries of the Kavango region and a larger number of constituencies in some other regions. It should be noted here that these changes will affect the sampling frames and the necessary changes had being implemented.

#### Secondary sampling frame(s)

The first field activity of the survey is to list the households of the selected PSU. Based on the information collected, agricultural households within the PSU can be identified. The list of agricultural households within the PSU is the secondary sampling frame where the sample agricultural households will be selected. A third frame also has to be developed for the purpose of selecting the plots for the crop cutting experiment. This will be a list of all the plots under a same crop within all the agricultural households in a given sample PSU. Three separate lists have to be prepared for the crops Mahangu, Maize and Sorghum for which the crop cutting experiments will be conducted. Three sample plots will be selected randomly from each of the 3 crops from the list of plots prepared for the PSU.

### Stratification

#### Primary sampling frame

Each region will be made a separate stratum and the constituencies implicitly stratified within the region. Since the population living around the urban centers and the settlements are also part of the target population, PSUs falling within the town land boundary but not developed will also be part of the frame. These PSUs will be grouped together and with the other rural PSUs will make the frame of the regiob. PSUs can be further stratified using a geographic order within the constituencies.

#### Secondary sampling frames

The list of agricultural households could be ordered within the PSU using the additional data collected during the listing process before the selection of the sample. This will enhance the representation of the different type of agricultural households in the sample.

### Sample size

The total sample size of about 10500 agricultural households is derived based on the following facts.

1. The lowest domain of estimation is the region
2. The expected level of precision for most of the variables at the regional level should be around 5% to 10% of CV.
3. The standard errors/CVs and the design effects of the variables from the previous Annual Agricultural Surveys (AAS) were used as the estimated measures of variation in the population
4. Available human and financial resources
5. A minimum sample of 800 agricultural households are fixed as the initial sample size for each region.

The sample size within a PSU is fixed at 10 agricultural households. This will contribute to have a larger spread of PSUs which will enhance the representation of the population thus covering for any larger design effects. In the southern regions where the number of PSUs with communal agricultural households is less, more households per PSU may have to be taken.

It is expected that these sample sizes will yield precise estimates around 5% to 10% CV for almost all variables at the regional level. Constituency level estimates are not expected from this sample size. But estimates with their sampling errors could be produced for the constituencies with a sample size of about 250 agricultural households. If there CVs falls within an acceptable range then such figures could be used.

Constituency level estimates with their sampling errors need to be produced (not for publication) for the important variables and their CVs should be examined as an input for the future agricultural surveys especially in the 8 northern regions including Kunene.

Refer Appendix 2 for the details of the sample size calculations

During the phase 2, a crop cutting experiment will be conducted to estimate the yield of major crops –Mahangu, Maize and Sorghum- which will be used to estimate the crop production. Three crop cutting experiments per crop within a selected PSU will be carried out. That means altogether 9 crop cutting experiments per PSU will be carried out.

### Allocation of the sample to strata

The overall sample of 10500 agricultural households needs to be allocated to the 14 regions to derive the regional sample sizes. Three options are available.

1. Equal allocation
2. Proportional allocation
3. Compromise allocation

#### Equal allocation

In this allocation procedure the sample will be allocated equally to each region. This approach does not take into account the variation of different variables across the regions. Hence some variables will be estimated precisely while some others may not. If the important variables have high variation then their estimates could be less reliable. Hence this method is not adopted.

#### Proportional allocation

In this approach the regions with larger size will get larger samples while the smaller regions will get small samples. Although this is the most efficient allocation procedure for the national estimates the regional estimation is going to suffer, especially the smaller regions. The small samples may not be good enough to give reliable estimates for the smaller regions. Smaller regions may also have variables with higher variation. Hence this method is also not considered.

#### Compromise allocation

As the name suggests this is a compromise between the two above approaches. This method will redistribute the sample from the larger regions to the smaller regions while keeping the overall sample fixed. Thus it moves away from the proportional allocation towards the equal allocation but not completely. In this approach one can also look at the variations of the important variables and choose the best allocation procedure, which will satisfy most of the requirements. If this is not possible then probably the overall sample size may have to be increased.

Hence this is the method adopted for the allocation of the sample to the regions.

Refer the Appendix 3 for the allocation table.

### Sample selection

First stage sample (PSU) is selected from the sampling frame using the Probability Proportional to Size sampling procedure (PPS) together with systematic sampling. The size measure is the number of agricultural households as at the previous population and housing census (2011 PHC).

The second stage sample –ie the agricultural households- is selected using systematic sampling procedure from the ordered list of agricultural households prepared as the first step of the data collection.

### Estimation

In the communal sector survey, data is collected from a sample of agricultural households drawn from the defined population. This sample data then needs to be expanded to cover the entire defined population of agricultural households/holdings. The process of this expansion is generally known as estimation and the expanded figures are estimates. The results to be published in the form of tables and other indicators will then be based on these expanded figures. Refer appendix 4 for the details of estimation procedure.

#### Base weight

The factors needed for this expansion (probabilities of selection) are found in the sample selection procedure. Since the sample design used is a two stage sample design, there are two stages of sample selection. First stage occurs when the PSUs are selected while the second stage is at the selection of agricultural households. The samples are drawn in such a way that the probabilities of selection are known. Factors needed for the expansion are the inverse of the probability of selection.

Further weighting components need to be added to this base weight if there is further sample selection stages in the process of data collection. In the agricultural census there is such a stage where the crop cutting experiments are to be conducted from selected crop fields/plots for the estimation of yield under major crops. This information will be used to estimates the production of the major crops in an objective way. Therefore when estimation of these characteristics are carried out, the base weight computed from the first two stages need to be adjusted based on the probabilities of selection of the additional stages.

Refer appendix 4 for the calculation of base weights.

### Non response adjustment

Estimates using the proper base weight as shown above would be unbiased if all the selected households participate in the survey. But as we have experienced not all the selected units respond. There will be certain amount of non-response in any survey operation. Therefore an adjustment has to be applied to the base weights to account for the non-response

### Post stratified adjustment

Once the estimated totals are prepared for sub groups of the population based on the non-response adjusted weights, they can be compared with any existing information from other sources which are **much more reliable**. If there are differences (normally there will be) then using these information another adjustment factor can be applied to the non-response adjusted weight so that the sub group totals from the survey conform to the outside source.

### Quality of the survey estimates

Since this is a sample survey, the estimates produced will differ from the actual figures which would have been obtained if a properly controlled complete enumeration of the agricultural households was carried out. Estimated results are subjected to sampling errors and non-sampling errors.

#### Non sampling errors

Non sampling errors include the biases resulting from inaccurate reporting, measurement and processing and also from the non-response. **These errors are difficult to measure but they can be controlled by proper detail planning and implementing such plans at every stage of the survey.** It is also important to report all aspects of the survey which did not happen according to the survey plan so that the users of the results know about such lapses when using the data.

#### Sampling errors

Sampling errors on the other hand can be estimated and should be calculated at least for all the main variables and should be presented to the users in the reporting. Sampling errors are based on the variances of the estimated figures which in turn depend on the sample sizes and the variation observed in the population units.

##### Variance estimation

The survey design is complex, which involves the cluster sampling, stratification and weighting. Hence the variance estimation formulae will also be complex. Using the correct formulae they can be computed for the totals and means in MS excel. This will be a bit cumbersome. Instead specialized software (such as STATA) can be used to compute the variances. Variances for the ratios are much more complex and hence software needs to be used.

Most commonly used methods applied in such software are the Taylor- series linearization and the replication methods such as Jackknife replication method.

##### Other measures of precision

Based on the sampling variance calculated above, other measures are also presented. Most important measures are

1. Standard error
2. Coefficient of variation (CV)
3. Confidence intervals and
4. Design effects (deff)

The software calculates all the above 4 measures at the same time including the sample size which was involved in the computation.

## Data collection

### Method

It was discussed earlier that the mode of data collection will be the face to face interview in this communal sector sample survey. Traditionally we have used the paper and pencil to collect the data from the selected respondents using a questionnaire. At the operational level this method requires the printing of questionnaires and sending the questionnaires to the field as the first step. Second step will be to collect the completed questionnaires back from the field and transporting them to the processing center. Third step will be to capture the data collected in the questionnaires into the computer system using a process such as scanning. Then load the captured data into the data bases and apply the cleaning procedures to finally get a cleaned data set.

A new method of data collection is going to be used in the NCA 2013/14. This method which is known as the Computer Based Personal Interviewing (CAPI) Method uses a laptop with an application programme where the questionnaire similar to the paper questionnaire is installed. Even some of the skip patterns are already in the application and hence the interviewer cannot go wrong in filling the questionnaire. Some further checks are also included in the application. This means that the time involved in the above mentioned processes are avoided except the questionnaire filling up time. As soon as the day’s work is completed the data could be uploaded into the system in the processing center where further checks can be carried out and the data then can be loaded into the data base.

This method is much more efficient in time management and accuracy. Only thing needed is the equipment and proper training.

### Field work structure

There are two phases in the data collection process.

Phase 1 deal with the data at the planting time of the crops. This will involve among others, the listing and selection of the agricultural households, completing the demographic characteristics of the households and the information of the holder, identifying the land use types and the area measurement of the plots, fields and the parcels of the holding which is operated by the holder.

Phase 2 covers the livestock numbers, crop areas and the production data including the objective measurement of the yield through crop cutting experiments.

Phase 1

##### Area measurement

Traditionally area measurement was carried out using the tape and the pole method. This method is also replaced in the new approach of the data collection. Instead of the tape and the poles, Global Positioning System (GPS) will be used to measure the area. It was found during a pilot exercise that this method produces the area of the land with the same accuracy as the traditional method but it reduces the measurement time a lot. Therefore this is the method that is going to be adopted for the area measurements in the survey.

#### Phase 2

##### Crop cutting experiment

In the case of estimating the production of crops, two methods – a subjective and an objective method- will be used. Subjective method is to ask the farmer to estimate the production under each crop in the units that is being used in the area. Objective method is to carry out crop cutting experiments in scientifically selected plots so that the yield can be estimated using the design features. Based on these results the production can be estimated.

# Appendices

## Appendix 1 List of new regions and constituencies

## Appendix 2 Sample size

## Appendix 3 Allocation procedure

## Appendix 4 Estimation procedure

## 

## Appendix 1

**List of regions and constituencies (alphabetical) as per gazette notification (09/08/13)**

|  |  |  |  |
| --- | --- | --- | --- |
| Region | Region Code | Constituency | Constituency Code within region |
| !Karas | 01 |  |  |
|  |  | !Nami ≠ Nüs | 1 |
|  |  | Berseba | 2 |
|  |  | Karasburg East | 3 |
|  |  | Karasburg West | 4 |
|  |  | Keetmanshoop Rural | 5 |
|  |  | Keetmanshoop Urban | 6 |
|  |  | Oranjemund | 7 |
| Erongo | 02 |  |  |
|  |  | Arandis | 1 |
|  |  | Daures | 2 |
|  |  | Karibib | 3 |
|  |  | Omaruru | 4 |
|  |  | Swakopmund | 5 |
|  |  | Walvis Bay Rural | 6 |
|  |  | Walvis Bay Urban | 7 |
| Hardap | 03 |  |  |
|  |  | Aranos | 1 |
|  |  | Daweb | 2 |
|  |  | Gibeon | 3 |
|  |  | Mariental Rural | 4 |
|  |  | Mariental Urban | 5 |
|  |  | Rehoboth East Urban | 6 |
|  |  | Rehoboth Rural | 7 |
|  |  | Rehoboth West Urban | 8 |
| Kavango East | 04 |  |  |
|  |  | Mashare | 1 |
|  |  | Mukwe | 2 |
|  |  | Ndiyona | 3 |
|  |  | Ndonga Linena | 4 |
|  |  | Rundu Rural | 5 |
|  |  | Rundu Urban | 6 |
| Kavango West | 05 |  |  |
|  |  | Kapako | 1 |
|  |  | Mankumpi | 2 |
|  |  | Mpungu | 3 |
|  |  | Musese | 4 |
|  |  | Ncamagoro | 5 |
|  |  | Ncuncuni | 6 |
|  |  | Nkurenkuru | 7 |
|  |  | Tondoro | 8 |
| Khomas | 06 |  |  |
|  |  | John Pandeni | 1 |
|  |  | Katutura Central | 2 |
|  |  | Katutura East | 3 |
|  |  | Khomasdal | 4 |
|  |  | Moses //Garoeb | 5 |
|  |  | Samora Machel | 6 |
|  |  | Tobias Hainyeko | 7 |
|  |  | Windhoek East | 8 |
|  |  | Windhoek Rural | 9 |
|  |  | Windhoek West | 10 |
| Kunene | 07 |  |  |
|  |  | Epupa | 1 |
|  |  | Kamandjab | 2 |
|  |  | Khorixas | 3 |
|  |  | Opuwo Rural | 4 |
|  |  | Opuwo Urban | 5 |
|  |  | Outjo | 6 |
|  |  | Sesfontein | 7 |
| Ohangwena | 08 |  |  |
|  |  | Eenhana | 1 |
|  |  | Endola | 2 |
|  |  | Engela | 3 |
|  |  | Epembe | 4 |
|  |  | Ohangwena | 5 |
|  |  | Okongo | 6 |
|  |  | Omulonga | 7 |
|  |  | Omundaungilo | 8 |
|  |  | Ondobe | 9 |
|  |  | Ongenga | 10 |
|  |  | Oshikango | 11 |
|  |  | Oshikunde | 12 |
| Omaheke | 09 |  |  |
|  |  | Aminius | 1 |
|  |  | Epukiro | 2 |
|  |  | Gobabis | 3 |
|  |  | Kalahari | 4 |
|  |  | Okarukambe | 5 |
|  |  | Otjimbinde | 6 |
|  |  | Otjinene | 7 |
| Omusati | 10 |  |  |
|  |  | Anamulenge | 1 |
|  |  | Elim | 2 |
|  |  | Etayi | 3 |
|  |  | Ogongo | 4 |
|  |  | Okahao | 5 |
|  |  | Okalongo | 6 |
|  |  | Onesi | 7 |
|  |  | Oshikuku | 8 |
|  |  | Otamanzi | 9 |
|  |  | Outapi | 10 |
|  |  | Ruacana | 11 |
|  |  | Tsandi | 12 |
| Oshana | 11 |  |  |
|  |  | Okaku | 1 |
|  |  | Okatana | 2 |
|  |  | Okatyali | 3 |
|  |  | Ompundja | 4 |
|  |  | Ondangwa Rural | 5 |
|  |  | Ondangwa Urban | 6 |
|  |  | Ongwediva | 7 |
|  |  | Oshakati East | 8 |
|  |  | Oshakati West | 9 |
|  |  | Uukwiyu | 10 |
|  |  | Uuvudhiya | 11 |
| Oshikoto | 12 |  |  |
|  |  | Engodi | 1 |
|  |  | Guinas | 2 |
|  |  | Nehale lya Mpingana | 3 |
|  |  | Okankolo | 4 |
|  |  | Olukonda | 5 |
|  |  | Omuntele | 6 |
|  |  | Omuthiyagwiipundi | 7 |
|  |  | Onayena | 8 |
|  |  | Oniipa | 9 |
|  |  | Onyaanya | 10 |
|  |  | Tsumeb | 11 |
| Otjozondjupa | 13 |  |  |
|  |  | Grootfontein | 1 |
|  |  | Okahandja | 2 |
|  |  | Okakarara | 3 |
|  |  | Omatako | 4 |
|  |  | Otavi | 5 |
|  |  | Otjiwarongo | 6 |
|  |  | Tsumkwe | 7 |
| Zambezi | 14 |  |  |
|  |  | Judea Lyabboloma | 1 |
|  |  | Kabbe North | 2 |
|  |  | Kabbe South | 3 |
|  |  | Katima Mulilo Rural | 4 |
|  |  | Katima Mulilo Urban | 5 |
|  |  | Kongola | 6 |
|  |  | Linyanti | 7 |
|  |  | Sibbinda | 8 |

## Appendix 2 Sample size

## Sample size calculation

### Statistical measure

The statistical measure used in these computations is the CV (Coefficient of Variation). This is a relative measure which is used to indicate the precision levels of the estimated figures, especially the means and the totals

The CVs can be classified into 4 broad groups. These groups can be then taken as the general guideline for the acceptable levels of precision.

|  |  |  |
| --- | --- | --- |
| **Group No** | **CV %** | **Indicator** |
| 1 | 1% - < 5% | Highly precise |
| 2 | 5% - < 10% | Good precision |
| 3 | 10% - < 15% | Acceptable if close to 10% |
| 4 | 15% - < 20% | Less precise and cannot be used for policy matters |
| 5 | 20% or more | Very low precision, cannot be used |

### Domains of estimation

These are the sub population groups for which separate estimates are required. In large scale household surveys the estimates are required at the Namibia level, Namibia urban and Namibia rural level, and the region level. In the case of the agricultural census the domains are Namibia and regions since the urban areas are excluded. Initial plan was to produce the estimates at the constituency level. Then the constituencies should have been the lowest domain. But the human and financial constraints had prevented producing the estimates at this level where a large sample is needed.

Namibia is a higher level domain to the region and constituency while region is at a higher level domain to the constituency. The domains at higher levels will always have larger samples than the lower levels since the lower level sample is accumulated to the next level. Hence Namibia will have the highest sample size followed by the regions and the constituencies (if applicable). Thus the Namibian estimates will always have a better precision than the regions and regions will be better than the constituencies always irrespective of the sample size.

### Estimated population CVs

All large scale surveys collect data on large number of characteristics. Each characteristic has its own variation in the population. Some will have very high variation while others may be moderate or small. It is necessary to know about this variation before the sample size can be determined because this sample size must cover these different variations as far as possible. It is sometimes an impossible task to cover all the variations of all the variables. Instead important characteristics are decided for the computation and their variations in the population need to be known. One way to get an indication of these variations is to use previous survey data and calculate the CVs for them.

For the agricultural census computations it is decided to use the following 3 variables – planted area, mahangu production and number of cattle. The CVs for these variables using the previous Annual Agricultural Survey data for some years (from 1996/97 to 2002/03) have been calculated and is shown below.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Region | Planted area | | Mahangu production\* | | Number of cattle | |
| Average CV% | Sample Households | Average CV% | Sample Households | Average CV% | Sample Households |
| Caprivi | 18.6 | 216 | 43.3 | 211 | 16.8 | 211 |
| Ohangwena | 15.3 | 240 | 31.3 | 239 | 35.0 | 239 |
| Kavango | 11.2 | 240 | 21.7 | 239 | 18.1 | 239 |
| Omusati | 12.7 | 228 | 12.5 | 228 | 15.3 | 227 |
| Oshana | 15.5 | 259 | 21.2 | 263 | 21.7 | 263 |
| Oshikoto | 14.4 | 214 | 20.6 | 215 | 24.7 | 215 |
| All six | 6.6 | 1306 | 25.1 | 1395 | 9.6 | 1394 |

\*Farmer’s self-assessment

Note – The regions shown here are the previous regions defined before the 2013 delimitation commission report. Caprivi is now named as Zambezi and Kavango is divided into 2 regions - Kavango East and Kavango west-.

Above table shows, that the planted area behaves in a uniform way across the regions compared to the other two variables. Mahangu production variable has higher variation in all regions except Omusati region. Variation in the number of cattle is bit more steady than the Mahangu production although the figures are higher than those of the planted area. This highlights the fact that the variation depends not only upon the characteristic under study but also on the geographic domain.

The CVs for the planted area in the regions falls among 11 to 19%. Kavango, Omusati and Oshikoto have the CV between 10 and 15% while the others are between 15 to 20%. It is only the combined estimate for all six regions, which shows an acceptable level of precision (6.6% CV) because of its larger sample size. In the case of number of cattle, CVs fall within 15 to 20% for some regions while for others it is above 20%. The combined estimate CV is 9.6 % again because of its larger sample size.

The situation is much worse for the Mahangu production except for Omusati region. Even the combined estimate does not have an acceptable level of precision (25.1% of CV) with the sample size of 1395 agricultural households. This variable is based on the subjective assessment of the production by the farmer and hence there will be very high variation here. This should not be considered as an important variable in the calculation of sample sizes.

### Sample size

Since we know the estimated CVs and the corresponding sample size which produced those estimates, it is possible to use the following fact to compute the sample sizes.

***To reduce the observed CV to half of that value, the sample size must be raised 4 times the observed sample.***

*Example*

*The average CV of planted area for Ohangwena region is 15.3%. The sample size which produced this estimate is 240 agricultural households from the above tables. Then to reduce 15.3% to half (ie 7.65%) the sample size needs to be raised to 960 agricultural households (4x240).*

*In the same argument if we want to reduce the CVs for quarter of the original value, the sample size needs to be raised by 16 times.*

This fact is used to derive sample sizes for each of the regions shown in the table for the different characteristics as shown below.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Region | Planted area | | | |
|  | Observed average CV% | Observed sample Households | Half the CV% | Sample size needed |  |
| Caprivi | 18.6 | 216 | 9.3 | 864 |  |
| Ohangwena | 15.3 | 240 | 7.6 | 960 |  |
| Kavango | 11.2 | 240 | 5.6 | 960 |  |
| Omusati | 12.7 | 228 | 6.4 | 912 |  |
| Oshana | 15.5 | 259 | 7.7 | 1036 |  |
| Oshikoto | 14.4 | 214 | 7.2 | 856 |  |
| All six | 6.6 | 1306 | 3.3 | 5224 |

It can be seen that for the variable planted area, the regional CVs can be reduced to the second broad group (ie 5% to 10%) with the derived sample sizes in the last column of the table while the CV for the combined estimate of all six regions is in the first broad group which is a highly precise estimate. This arrangement is quite acceptable.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Region | Mahangu production\* | | | | | |
|  | Observed average CV% | Observed sample Households | Half the CV% | Sample size needed | One third of the CV% | Sample size needed |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Caprivi | 43.3 | 211 | 21.7 | 844 | 14.4 | 1899 |
| Ohangwena | 31.3 | 239 | 15.7 | 956 | 10.4 | 2151 |
| Kavango | 21.7 | 239 | 10.9 | 956 | 7.2 | 2151 |
| Omusati | 12.5 | 228 | 6.3 | 912 | 4.2 | 2052 |
| Oshana | 21.2 | 263 | 10.6 | 1052 | 7.1 | 2367 |
| Oshikoto | 20.6 | 215 | 10.3 | 860 | 6.9 | 1935 |
| All six | 25.1 | 1395 | 12.6 | 5580 | 8.4 | 12555 |

\*Farmer’s self-assessment

This variable does not behave in the same way as the planted area as explained before. The CVs computed for half the value, fall into the second, third and fourth broad groups. They are not very consistent. It is difficult to get the Caprivi and Ohangwena CVs into an acceptable range unless by raising the sample size quite large. If most other variables behave in a similar fashion and one such variable behaves differently, then the acceptable precision level for this variable could be relaxed for the domains where it is high.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Region | Number of cattle | | | | | |
| Observed average CV% | Observed sample Households | Half the CV% | Sample size needed | One third of the CV% | Sample size needed |
| Caprivi | 16.8 | 211 | 8.4 | 844 | 5.6 | 1899 |
| Ohangwena | 35.0 | 239 | 17.5 | 956 | 11.7 | 2151 |
| Kavango | 18.1 | 239 | 9.05 | 956 | 6.0 | 2151 |
| Omusati | 15.3 | 227 | 7.65 | 908 | 5.1 | 2043 |
| Oshana | 21.7 | 263 | 10.85 | 1052 | 7.2 | 2367 |
| Oshikoto | 24.7 | 215 | 12.35 | 860 | 8.2 | 1935 |
| All six | 9.6 | 1394 | 4.8 | 5576 | 3.2 | 12546 |

In the case of number of cattle, one could still use the sample sizes adjusted a bit for the Oshana and Oshikoto regions to achieve half the CV. But it is difficult to get the sample adjusted for Ohangwena region to get this precision level, hence we accept a relaxed precision level for this region. There is a possibility in one of the years the CV was very high.

Looking at these figures, the sample sizes derived for half the CVs of planted area can be taken as the guideline but raised to some extent to cover the variation in other variables.

Based on these observations a minimum sample size of about 800 agricultural households are needed for each of the regions in the northern communal areas (7 regions out of 14 – Zambezi, Kavango east, Kavango west, Ohangwena, Oshikoto, Oshana and Omusati) where crop cultivation is predominant with larger number of (communal) agricultural households. The remaining regions are covered mostly with the large commercial farms but with small pockets of communal farming areas concentrated within possibly one or two constituencies except Kunene region which may have a larger proportion of communal farming activities than the other southern regions. Lesser sample sizes could cover the communal farming activities in these regions adequately. Considering all these factors and the budgetary constraints a sample size of about 10500 agricultural households was considered as the overall sample.

It is expected that this sample size will yield precise estimates at the 5% to 10% level of CVs for most of the variables at the region level. Constituency level estimates could also be produced for larger constituencies where adequate samples (about 250 households) are covered. On the other hand constituency level estimates should be produced ( not to be published) at least for the important variables for all constituencies in the northern 8 regions including Kunene and their CVs examined for the purpose of future planning of agricultural surveys.

## Appendix 3 Allocation procedure

**Number of all households and the agricultural households by region**

**(Rural and semi urban areas only)**

|  |  |  |
| --- | --- | --- |
| Region | All households | Agricultural households (size measure) |
| !Karas | 9,801 | 1,421 |
| Erongo | 5,186 | 1,832 |
| Hardap | 8,783 | 547 |
| Kavango east | 12,408 | 8,450 |
| Kavango west | 13,566 | 9,612 |
| Khomas | 5,153 | 191 |
| Kunene | 13,213 | 5,529 |
| Ohangwena | 39,680 | 35,138 |
| Omaheke | 11,199 | 2,334 |
| Omusati | 44,276 | 34,107 |
| Oshana | 20,524 | 16,350 |
| Oshikoto | 32,332 | 24,621 |
| Otjozondjupa | 14,215 | 3,267 |
| Zambezi | 14,611 | 9,193 |
| Namibia | 244,947 | 152,292 |

Source – 2011 Population and Housing Census

Note –Developed parts of the municipalizes, towns, villages and settlements are excluded

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Size measure raised to the power (λ) , where λ = 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1 by region** | | | | | | | | | | |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
| Region | λ | | | | | | | | | | |
| 0 | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | 1 |
| !Karas | 1 | 2.06661 | 4.27089 | 8.82629 | 18.2405 | 37.6962 | 77.9034 | 160.996 | 332.717 | 687.598 | 1421 |
| Erongo | 1 | 2.11979 | 4.4935 | 9.52528 | 20.1916 | 42.8019 | 90.7309 | 192.33 | 407.7 | 864.237 | 1832 |
| Hardap | 1 | 1.87845 | 3.52856 | 6.62821 | 12.4507 | 23.388 | 43.9332 | 82.5261 | 155.021 | 291.198 | 547 |
| Kavango east | 1 | 2.46994 | 6.10058 | 15.068 | 37.2171 | 91.9239 | 227.046 | 560.789 | 1385.11 | 3421.14 | 8450 |
| Kavango west | 1 | 2.50197 | 6.25983 | 15.6619 | 39.1855 | 98.0408 | 245.295 | 613.719 | 1535.5 | 3841.78 | 9612 |
| Khomas | 1 | 1.69084 | 2.85895 | 4.83404 | 8.1736 | 13.8203 | 23.3679 | 39.5115 | 66.8077 | 112.961 | 191 |
| Kunene | 1 | 2.36736 | 5.6044 | 13.2676 | 31.4093 | 74.3572 | 176.031 | 416.728 | 986.546 | 2335.51 | 5529 |
| Ohangwena | 1 | 2.84825 | 8.11251 | 23.1064 | 65.8129 | 187.451 | 533.908 | 1520.7 | 4331.33 | 12336.7 | 35138 |
| Omaheke | 1 | 2.17175 | 4.7165 | 10.2431 | 22.2454 | 48.3115 | 104.921 | 227.861 | 494.858 | 1074.71 | 2334 |
| Omusati | 1 | 2.83978 | 8.06434 | 22.9009 | 65.0335 | 184.681 | 524.452 | 1489.33 | 4229.36 | 12010.4 | 34107 |
| Oshana | 1 | 2.63847 | 6.96151 | 18.3677 | 48.4626 | 127.867 | 337.373 | 890.148 | 2348.63 | 6196.78 | 16350 |
| Oshikoto | 1 | 2.74872 | 7.55546 | 20.7679 | 57.085 | 156.911 | 431.304 | 1185.53 | 3258.7 | 8957.26 | 24621 |
| Otjozondjupa | 1 | 2.24603 | 5.04464 | 11.3304 | 25.4484 | 57.1577 | 128.378 | 288.34 | 647.619 | 1454.57 | 3267 |
| Zambezi | 1 | 2.49084 | 6.20428 | 15.4539 | 38.4931 | 95.8801 | 238.822 | 594.867 | 1481.72 | 3690.72 | 9193 |
| Namibia | 14 | 33.0788 | 79.776 | 195.982 | 489.449 | 1240.29 | 3183.46 | 8263.38 | 21661.6 | 57275.6 | 152592 |

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Allocation table computed for the sample size of 10500 agricultural households (proportional to the size raised to the power λ)** | | | | | | | | | | | |
| Region | λ | | | | | | | | | | |
| 0 | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | 1 |
| !Karas | 750 | 655.9929 | 562.129 | 472.8812 | 391.3084 | 319.1273 | 256.9483 | 204.5726 | 161.2774 | 126.0533 | 97.78036 |
| Erongo | 750 | 672.8717 | 591.4285 | 510.3304 | 433.1634 | 362.3511 | 299.2572 | 244.3877 | 197.6235 | 158.4355 | 126.0617 |
| Hardap | 750 | 596.2638 | 464.424 | 355.1158 | 267.1016 | 197.9979 | 144.9044 | 104.8631 | 75.14292 | 53.38363 | 37.63959 |
| Kavango east | 750 | 784.0169 | 802.95 | 807.2923 | 798.4068 | 778.2072 | 748.8647 | 712.5761 | 671.4034 | 627.1776 | 581.4525 |
| Kavango west | 750 | 794.184 | 823.9103 | 839.1082 | 840.6342 | 829.9917 | 809.0541 | 779.8322 | 744.302 | 704.2906 | 661.4108 |
| Khomas | 750 | 536.7142 | 376.291 | 258.9905 | 175.3456 | 116.9994 | 77.07425 | 50.20592 | 32.38358 | 20.70854 | 13.14289 |
| Kunene | 750 | 751.4574 | 737.6435 | 710.8334 | 673.8142 | 629.492 | 580.6003 | 529.5222 | 478.2065 | 428.1554 | 380.4557 |
| Ohangwena | 750 | 904.102 | 1067.757 | 1237.961 | 1411.862 | 1586.921 | 1760.984 | 1932.304 | 2099.519 | 2261.616 | 2417.879 |
| Omaheke | 750 | 689.3659 | 620.7795 | 548.7873 | 477.2236 | 408.9944 | 346.0587 | 289.5357 | 239.8716 | 197.02 | 160.6047 |
| Omusati | 750 | 901.4136 | 1061.417 | 1226.95 | 1395.144 | 1563.467 | 1729.798 | 1892.439 | 2050.09 | 2201.804 | 2346.935 |
| Oshana | 750 | 837.5129 | 916.2642 | 984.0772 | 1039.654 | 1082.495 | 1112.756 | 1131.082 | 1138.446 | 1136.019 | 1125.059 |
| Oshikoto | 750 | 872.5098 | 994.4394 | 1112.668 | 1224.627 | 1328.372 | 1422.567 | 1506.418 | 1579.584 | 1642.081 | 1694.194 |
| Otjozondjupa | 750 | 712.9427 | 663.9678 | 607.0418 | 545.9353 | 483.8842 | 423.4272 | 366.3835 | 313.9189 | 266.6574 | 224.8054 |
| Zambezi | 750 | 790.6522 | 816.5986 | 827.9631 | 825.7802 | 811.7 | 787.7051 | 755.8778 | 718.2308 | 676.5985 | 632.579 |
| Namibia | 10500 | 10500 | 10500 | 10500 | 10500 | 10500 | 10500 | 10500 | 10500 | 10500 | 10500 |

Looking at the distribution of the sample in the above table, allocation for λ = 0.5 was chosen as the best possible allocation.

**The distribution of the sample based on the compromise allocation of λ = 0.5 and 10 households per PSU**

|  |  |  |
| --- | --- | --- |
| Region | Sample PSUs | Sample agricultural households |
|
| !Karas | 32 | 320 |
| Erongo | 37 | 370 |
| Hardap | 20 | 200 |
| Kavango east | 78 | 780 |
| Kavango west | 83 | 830 |
| Khomas | 12 | 120 |
| Kunene | 63 | 630 |
| Ohangwena | 159 | 1590 |
| Omaheke | 41 | 410 |
| Omusati | 157 | 1570 |
| Oshana | 109 | 1090 |
| Oshikoto | 133 | 1330 |
| Otjozondjupa | 49 | 490 |
| Zambezi | 82 | 820 |
| Namibia | 1055 | 10550 |

Maximum number of communal PSUs in the Khomas region was eight. Hence the allocated sample has to be reduced to 8.

In Erongo and Omaheke there were not enough communal PSUs in the frame to select the above sample. Hence the number of agricultural households per PSU was raised to 16 instead of 10 to reduce the sample PSUs.

Kavango East sample was raised to the minimum level of 800 agricultural households with 80 sample PSUs.

The revised sample sizes are given below.

|  |  |  |
| --- | --- | --- |
| Region | Sample PSUs | Sample Agricultural Households |
|
| !Karas | 32 | 320 |
| Erongo | 24 | 384 |
| Hardap | 20 | 200 |
| Kavango east | 80 | 800 |
| Kavango west | 83 | 830 |
| Khomas | 8 | 80 |
| Kunene | 63 | 630 |
| Ohangwena | 159 | 1,590 |
| Omaheke | 26 | 416 |
| Omusati | 157 | 1,570 |
| Oshana | 109 | 1,090 |
| Oshikoto | 133 | 1,330 |
| Otjozondjupa | 49 | 490 |
| Zambezi | 82 | 820 |
| Namibia | 1,025 | 10,550 |

Note: Erongo and Omaheke sample PSUs has to be multiplied by 16 to get the sample households

## Appendix 4 Estimation procedure

Variables collected during phase 1

Since the sample is selected in 2 stages there will be 2 probabilities of selection, p1 for the first stage and p2 for the second stage. First stage probability is based on the PPS selection procedure and the second stage probability is based on the random sample procedure

First stage probability of selection p1 is given by

Where;

Mhi = Number of Agricultural households in PSU i in stratum h (PSU size as per previous

Population census)

Mh = Number of ag: households in the stratum h (stratum size)

nh = Number of PSUs selected from the stratum h

Second stage probability of selection p2 is given by

Where;

M’hi  = Number of agricultural households in PSU i in stratum h according to survey listing

mhi = Number of agricultural households in the sample from PSU i in stratum h

Therefore the inclusion probability of a holding, p = p1 \* p2

#### Base weight

Since the PPS selection is and unequal probability selection the sample data has to be weighted. These weights which are generally called sample weights or base weights are the inverse of the inclusion probability.

Therefore the base weight W is given by

#### Estimation of a total

A total could be estimated from the sample by the following estimator;

Where;

= value of any characteristic of the jth household in ith PSU of stratum h

L = Number of strata

***Estimation of a ratio***

A ratio is estimated by;

Where is estimated in the same way as .

An *average* is in effect a ratio of two estimates, an estimate of the total and an estimate of the total number of units (agricultural households, individuals etc). An average can thus be estimated in the same way as a ratio, where the variable X takes the value = 1 for all units.

A *proportion* can also be estimated as a ratio. In this case the variable y takes value = 1 if the unit belongs to the specific group and the value = 0 if it doesn’t belong to the group. The variable X takes the value = 1 for all units.

***Variances***

Let;

A simple expression for an estimate of the variance of is;

An estimate of the variance of a ratio is;

Where