

# **Comparison of Official Data Sources and Construction of a Sampling Frame for Household-based Livestock Surveys in Nueva Ecija, Philippines**

**Anna Ma. Lourdes S. Latonio**

*Department of Statistics, Central Luzon State University*

**Isidoro P. David**

*Institute of Statistics, University of the Philippines Los Baños*

**Zita V.J. Albacea**

*Institute of Statistics, University of the Philippines Los Baños*

A basic issue in designing a procedure on selecting a sample is how well the sampling frame corresponds to the target population from where the sample will come from, making the construction of the sampling frame a vital aspect in the development of a sampling design. The main purpose of this paper is to evaluate and compare different official data sources to be able to combine them into a sampling frame appropriate for use in household-based livestock operations surveys for major livestock types (carabao, cattle, goat, and swine) in the Province of Nueva Ecija.

The data values of different official data sources available such as the Philippine Carabao Center Inventory (PCCI), Livestock and Poultry Survey (LPS), Barangay Agricultural Profiling Survey (BAPS), and Local Government Unit (LGU) records during 2007-2010 for the province were compared. Scatter plots, strengths of relationships and relative differences between observations on the same variables were analyzed. Specifically, relevant barangay level data common in the different surveys were used to perform the comparisons.

Based on the assessment of the different data sources, a barangay level Household Based Livestock Frame (HBLF) was constructed. The constructed HBLF consists of 849 barangays as basic sampling units, each with attached livestock related information, and auxiliary information such as total rice area (TRA) and number of rice growers (NRG). The constructed HBLF can be used in the development of separate or combined sampling designs for household-based livestock surveys in the Province of Nueva Ecija.

*Keywords: Agriculture, livestock, sampling, sampling frame, livestock inventory*

## 1. Introduction

The picture of any economy will not be complete without the important and current information about the livestock industry. Livestock, as defined by the Philippine Statistics Authority (PSA, 2020) refers to “domesticated animals with four legs kept or reared in captivity mainly for agricultural purposes.” For purposes of census and surveys, livestock covers only those that are tended and raised by an operator.

How good a sample is judged not in terms of its characteristics, but in accordance to the process by which it was selected, so that a basic issue in the design of the procedure of selecting the sample is how well the sampling frame corresponds to the target population from which the sample was to be chosen. The sampling frame is the list of elements from where the sample is actually drawn (Cooper and Schindler, 2014). Since any conclusion based on the sample is applicable supposedly only to the population from where the sample came from, it is a crucial aspect of the sample design. According to Kish (1965), the nature of available or feasible frames is an important consideration in the sample design, including factors such as the type of sampling units, extent of coverage, the accuracy and completeness of the list, and the amount and quality of auxiliary information on the list. Turner (2003) in addition mentioned that the sampling frame also influences the cost and the quality of a survey and is a common source of non-sampling error especially if important subgroups of the population are not covered.

There are several data sources available for use in the construction of a sampling frame for household based livestock surveys in the province of Nueva Ecija. However, these data sources should be evaluated for completeness and reliability. The main purpose of this paper is to evaluate, compare and combine different official data sources to be able to construct a sampling frame appropriate for use in developing household-based livestock survey designs in the province.

## 2. Materials and Methods

The Philippines is geographically subdivided into 17 administrative regions, each composed of several administrative provinces. The province of Nueva Ecija is in Region III. Provinces are composed of municipalities which are further composed of barangays. Barangay is the smallest administrative unit in the country.

### 2.1. *The official sources and their data*

The barangay level data sources on livestock and agricultural characteristics evaluated for possible use in the construction of the sampling frame for the province of Nueva Ecija are the following: (1) 2007 Barangay Agricultural Profiling Survey (BAPS), (2) 2010 Livestock and Poultry Survey (LPS), (3) 2010

compiled records of Local Government Units (LGU) from the City and Municipal Agricultural Offices (C/MAOs), City/Municipal Veterinary Offices (C/MVOs) and City/Municipal Planning and Development Offices (C/MPDOs), and (4) the 2010 Philippine Carabao Center Inventory (PCCI).

***The Barangay Agricultural Profiling Survey (BAPS).*** BAPS was conducted by then the Bureau of Agricultural Statistics (BAS, now Philippine Statistics Authority (PSA)), simultaneously in selected regions in the Philippines during 2006 - 2007 with data reported as of April 1, 2007 for Region III which includes the province of Nueva Ecija. BAPS, a barangay level census by key-informant (KI) method was designed to collect barangay level information on basic agriculture and fishery structure, to be used as inputs by the government in the identification, design and implementation of appropriate development programs and identification of target beneficiaries. The survey was intended for the identification of areas suitable to the production and marketing of priority commodities in the provinces (BAS, 2007a).

BAPS employed Contractual Data Collectors (CDCs) to interview face to face or by groups, KIs such as Barangay Chairman, Barangay Official-in-Charge of agriculture, Heads or representatives of C/MAOs, C/MPDOs, and C/MVOs or Livestock Inspectors (LIs). The BAPS data used for the study were the latest validated values of the relevant variables for 2010, obtained in July 2011 from the BAS provincial office in Nueva Ecija.

***The Livestock Population Survey (LPS).*** The LPS was a nationwide barangay level census implemented by BAS for the PCC in 2011. Other livestock agencies of the Department of Agriculture which collaborated on the planning and implementation of the activity were the Livestock Development Council, Bureau of Animal Industry, and the National Dairy Authority. The purpose of the activity was to generate basic information on the population and number of raisers of major livestock commodities at the barangay level, both for the backyard and commercial livestock farms (BAS, 2010b). The LPS reference date was as of March 2010.

The methods of collecting and processing the backyard livestock data for LPS stated in the survey's operational manual (BAS 2010b) were basically similar to BAPS, that is, employing CDCs to conduct individual or group interviews of KIs. Among the KIs listed that may be selected for the interview were the Barangay chairman/councilors, Municipal/City agriculturist, Livestock technician, Cooperative leader, President/Chair of livestock associations, Veterinarian, Livestock raiser/trader, and Meat inspector. BAPS and LPS were similar except for data coverage, since LPS gathered data only on livestock related characteristics of the barangays.

***The Local Government Unit Data (LGU).*** At the local government level, specifically at the city/municipal level, offices such as the C/MPDO, C/MAO and the C/MVO, different monitoring activities are being conducted for purposes of

providing the local government bases for projects and programs. The C/MAO is responsible for the updated agricultural profile of the city/municipality. Part of this profile are the livestock inventories as well as rice area, irrigated and rainfed rice areas, number of farmers and poultry inventory, which are usually reported at the barangay level and summarized to the overall municipal level. The barangay level livestock inventories of the city/municipality on the other hand are the responsibility of the C/MVO or the Livestock Inspector (LI) stationed at the C/MAO. Not all municipalities have a separate Veterinary Office. In case there was no C/MVO, the C/MAO employs one or several LIs whose responsibility includes the keeping and updating of all livestock records of the city/municipality. The Veterinarian/LI deals personally with the livestock raisers in the locality through visitations, vaccinations and other services, and keeps records of these services that are usually free or requires minimal fee and rarely but might cover institutional livestock farms.

The compiled 2010 LGU data from the different local offices have different reference dates. Some of the inventory reflected in the reports may refer only to the inventory of livestock which were serviced by the LI, instead of the actual total inventory of livestock in the area. Nevertheless, these local government offices affirmed that they conduct house to house monitoring of livestock among their constituents to complete the inventory as these figures are required by higher officials for the implementation of local government programs and for annual report submission to the Provincial Agriculturist Office (PAO) and Provincial Veterinary Office (PVO) as well.

The 2010 LGU data used for the study composed the 32 municipal local governments' record gathered and assembled, with different reference dates noted to be submitted by the concerned offices towards the end of 2009 to early 2010.

The BAPS and LPS, and the assembled LGU data have something in common, in that the Agriculturists, Veterinarians and the Planning and Development Officers of the city/municipality had a contribution on the content of these census/survey data. The BAPS and LPS had the Agriculturists, Veterinarians and Livestock Inspectors as KIs and other participants in the data validation stages such as the BAS supervisors/staff at the provincial, regional and national levels, while the compiled LGU data were exclusively a result of the compilation of the inventory reports from the C/MPDOs, C/MAOs, and C/MVOs.

***Philippine Carabao Center Inventory (PCCI).*** The Philippine Carabao Center is “the agency mandated to conserve, propagate, and promote the Carabao as a source of milk, meat, draft power and hide to benefit the rural farmers” (PCC, 2018). The agency therefore needs to generate reliable and timely statistics to monitor the progress of the carabao industry in the country. The Center conducted a census of carabao and carabao raisers in Nueva Ecija during the last quarter of 2010. The coverage was confined to household based livestock raisers only, i.e. excluding the carabao inventories of establishment/institution/government based

carabao farms. The PCC station located at CLSU took charge of the census. A workforce was organized to perform the house to house data collection activity from November to December of 2010 for the entire province.

## *2.2. Comparisons made among data from different sources for the needed sampling frame*

The data values of the different official data sources for the province were compared. Specifically, relevant barangay level data common in BAPS, LPS, LGU, and PCCI were used in the following comparisons:

1. BAPS, LGU on barangay level agricultural and demographic information
2. LPS, BAPS, LGU and PCCI on barangay level backyard carabao inventory estimates
3. LPS and BAPS on barangay level backyard cattle, goat and swine inventory estimates

For (1), the LGU data was used as the reference to calculate relative differences between the two data sources. The BAPS and LGU data sources include agricultural and demographic barangay level information aside from livestock inventory estimates. The relevant barangay level characteristics which were both available for comparison were the barangay total land area (TLA), total agricultural area (TAA), total rice area (TRA), population size, number of households (NHH) and number of rice growers (NRG). The data on these characteristics were available for all barangays and municipalities in BAPS, while for the LGU, data on the same characteristics for some municipalities were not available. Strengths of relationship between data on common variables from the two data sources were evaluated. A third data source for barangay level population size is the latest Census of Population and Housing (CPH). The strength of relationship of population size between CPH and each of BAPS and LGU were also compared.

For (2), the comparisons were carried out using barangay level backyard carabao data of the different data sources with the PCCI as reference data, being a complete enumeration at the household level, and is assumed to have zero sampling error and lower non-sampling error. The LGU data on carabao was checked and was found to compose solely of backyard/household carabao operations.

For (3), comparisons between LPS and BAPS barangay backyard livestock inventory estimates were performed for cattle, goat and swine. The LGU data was not included in the comparison for evidences of inclusion of large livestock operations in the data.

Scatter plots were constructed to examine and visualize the presence of outliers and magnitude of discrepancies between pairs of data sources. Data on a common barangay characteristic gathered by two equally efficient techniques

are supposedly the same in value, or at least strongly linearly correlated, and the points of the scatter plot of the two data sets are expected to lie close to a 45-degree straight line through the origin. Scatterplots and strengths of linear relationships of the same barangay level characteristics by pairs of data sources using Pearson's  $r$  were evaluated and compared. The strengths of associations of  $|r|$  was described as: Very strong to perfect linear association | 0.90 - 1.00 |, Strong | 0.80 - 0.89 |, Moderate | 0.50 - 0.79 |, Weak | 0.25 - 0.49 |, and from No to very weak linear association | 0.00 - 0.24 |.

For the scatter plots, two dashed lines were plotted on both sides at 0.1 Range units parallel to the 45-degree bold straight line of the plots. A thin line representing the linear regression line generated using Microsoft Excel is also included in the plots. A regression line which extends below the 45-degree line indicates that the data source assigned to the horizontal axis tends to give higher values than the data source assigned to the vertical axis, and interpreted in reverse, if otherwise. A point in the scatterplot noticeably distant from the 45-degree line and outside either of the two dashed lines on the other hand is considered as an outlier, suggesting that the observation from one of the two data sources being compared is either very high or very low compared to the observation from the other data source.

In addition to the strengths of relationships and scatterplots, descriptive measures (mean, total, standard deviation (SD) and coefficient of variation (CV)) of the different barangay level characteristics as measured by the different data sources were evaluated and compared. The ideal situation that the values of a descriptive measure are equal among the different data sources is not probable since the reference dates of data collections are not the same. Nevertheless, the descriptive measures for the same barangay level characteristic, from two equally efficient and reliable data sources, are expected to be close in values.

A reference data source is specified in comparisons (1) and (2). As the mean and total are usually the information of interest in official surveys, this leads to the data source with these descriptive measures closer to that of the reference data source as candidate for the possibly more reliable data source. Aside from closeness to the reference data source in terms of the mean and total, a desirable property likewise would be to have a variability among observations as similar or close to that of the reference data source. For this purpose, the SDs were obtained, to compare, on the average, deviations of the actual observations gathered by the different data sources on the same type of barangay level characteristic each from their respective mean. When several data sources report different mean values on a common barangay level characteristic, and if accompanied by relatively equal SDs, this gives a picture of homogeneity of variances among the observations gathered by the different data sources. The data source with the closest mean and SD to that of the reference data source on that variable is the more reliable, as it manifests closer behavior or characteristics as the reference data source.

The SD expressed as a percentage of the mean, as a measure of variability and reliability, known as the coefficient of variation (CV), has the advantage that it can be used when comparing several data sets on the same or different variables, as it is unitless. The CVs were evaluated as basis of the reliability of the different data sources for the situations when comparison involves not only the same but also includes different types of barangay level characteristics. Overall reliabilities of the different data sources were compared using the CVs.

### *2.3. Construction of the sampling frame for household-based livestock surveys in Nueva Ecija*

Based on the results of the different comparisons described, the sampling frame with barangays as sampling units accompanied with pertinent barangay level information was constructed using the PCCI household level census on carabao inventory accumulated to barangay totals in Nueva Ecija, the LGU barangay level characteristics and livestock inventories, supplemented by LPS and the CPH. As part of the preparation, each of the barangay level information on total land area (TLA), total agricultural area (TAA), total rice area (TRA), number of rice growers (NRG), population (CPH) and number of households (NHH), as well as livestock related characteristics were carefully examined for completeness. Observations for the identified outliers were checked and appropriate substitutions were made using data from respective identified supplementary data sources.

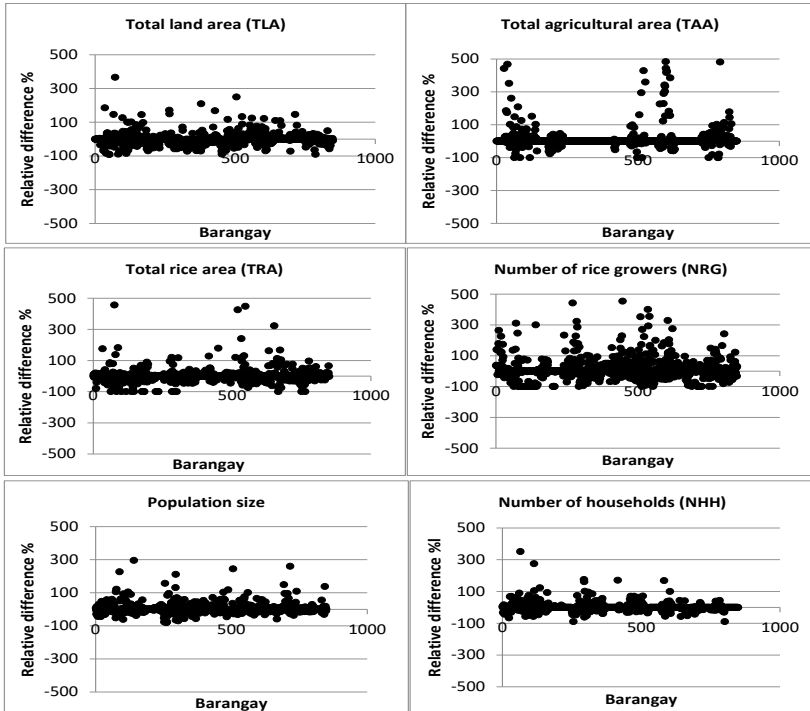
## **3. Results and Discussion**

### *3.1 Comparison of BAPS and LGU data sources on barangay level agricultural and demographic information in Nueva Ecija*

Among the barangay level information available from BAPS and LGU are the total land area (TLA), total agricultural area (TAA), total rice area (TRA), number of rice growers (NRG), population size, and number of households (NHH). Using the LGU as reference data, plotting and examining the relative differences of BAPS-LGU on the different variables show that in each graph, there are substantially more relative differences even higher than 100% (Figure 1). These are evidences that more BAPS estimates tend to be higher than LGU values on all the barangay level characteristics being considered.

BAPS and LGU data on TLA, TRA, and population size were expected and found to be very strongly linearly correlated ( $r > 0.90$ ) while data on NHH, NRG and TAA were observed to have inferior or moderate strengths of relationship ( $0.50 \leq r \leq 0.79$ , Table 1). The scatter plots in Figure 2 constructed for the different barangay level characteristics showed points (outliers) outside the dashed lines implying distances by more than 0.1 Range units from both the 45-degree bold line and the thin line (the regression line produced by Microsoft Excel). The outlier points represent barangays where there existed very large discrepancies

between the BAPS and LGU, causing weakness in the strength of relationships of the data from the two sources. Thus, more outliers were observed between data on variables with inferior strengths of relationship such as NHH, NRG and TAA, where there are several gross outlier points that brought the correlation down.

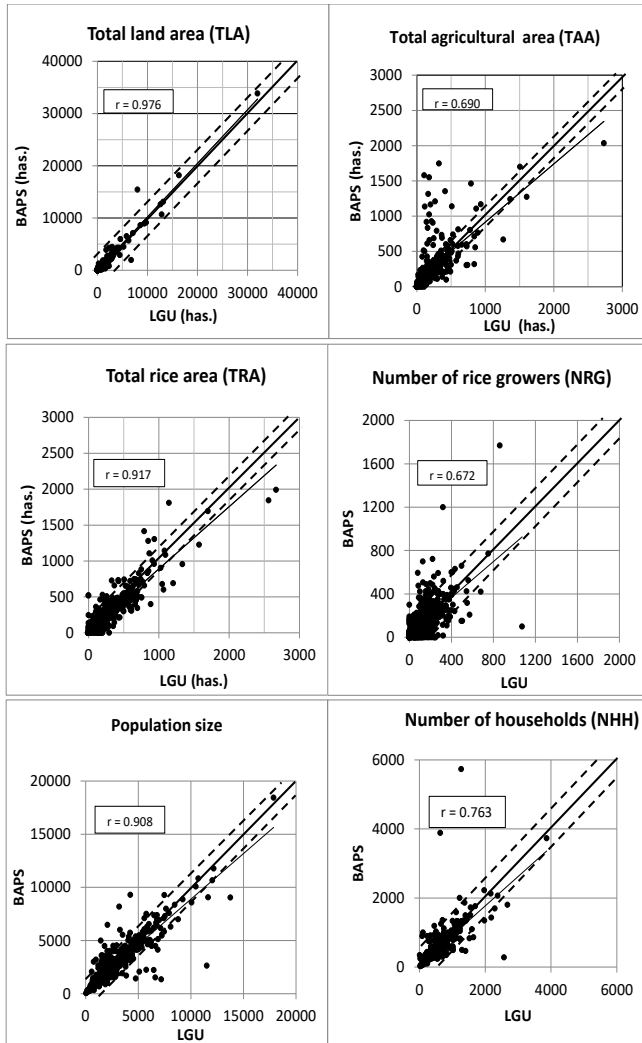


**Figure 1. Relative differences (%) of BAPS over LGU on barangay level characteristics in Nueva Ecija, 2010.**

**Table 1. Correlation between BAPS and LGU Data on Common Barangay Level Characteristics, Nueva Ecija, 2010**

Barangay Characteristic	Pearson's r	Number of barangays
Total land area (TLA)	0.976	845
Total agricultural area (TAA)	0.694	389
Total rice area (TRA)	0.917	849
Number of rice growers (NRG)	0.672	839
Population size	0.908	849
Number of households (NHH)	0.763	632
Between LGU TAA and BAPS TRA	0.862	354





**Figure 2. Scatter Plots of BAPS - LGU Data on Barangay Level Characteristics in Nueva Ecija, 2010**

Which data source is more reliable? For the case that the data between the sources/techniques on the same variable has a strong correlation, the two data sets can be used to supplement each other in the construction of a sampling frame. It is also useful to identify which can serve as the major data source, and which will be used for supplementary purposes. In the case when the data between the sources/techniques on a variable is not strongly correlated, it is important to identify

which is the more reliable between the two data sources. A way to determine which among the two data sources is more reliable for a variable is to compare the strengths of the relationship of the different sources' data on the variable with the data of another variable which is expected to be strongly correlated with it. For instance, NHH is expected to be highly correlated with population size, and NRG with that of TRA. Another way is to use a data on the same variable from a third data source which is known to be reliable to compare with the two current sources. For instance, a third data source for population size is the latest Census of Population and Housing (CPH). Examining the correlations between these pairs of variables gave hint to which data source is more reliable for the variables being considered.

Table 2 shows that in all the pairs of variables considered to be supposedly strongly correlated, those gathered by LGU were more highly correlated than those reported by BAPS. The province of Nueva Ecija, being a major rice producing province, and being the Rice Granary of the Philippines, has most of its agricultural lands as rice areas. The agricultural area and rice area are even reported synonymously by several C/MAOs. Both area measures therefore are expected to be positively correlated with the number of rice growers. Based on the results, the LGU data on TAA and TRA, as well as NRG will be considered as more reliable.

**Table 2. Correlations between BAPS and LGU Data on Barangay Level Characteristics in Nueva Ecija, 2010**

Pairs of variables	BAPS (N = 849)	LGU	N for LGU
Population size and Number of households (NHH)	0.797	0.944	632
Total agricultural area (TAA) and Total rice area (TRA)	0.380	0.929	389
Total agricultural area (TAA) and Number of rice growers (NRG)	0.224	0.729	389
Total rice area (TRA) and Number of rice growers (NRG)	0.686	0.832	839
CPH and Population size	0.924	0.969	849
CPH and Number of households (NHH)	0.818	0.914	632

The population size and NHH by LGU are more strongly correlated (0.944) than that of the data on the same variables by BAPS (0.797). The data on these two variables by LGU are also more strongly correlated to the third data source CPH. These evidences again suggest the LGU data on population size and NHH are the more reliable data.

### 3.2 Comparison between BAPS, LPS, LGU and PCCI data sources on barangay level backyard carabao inventories

For carabao, the backyard inventory estimates by BAPS, LPS, LGU and PCCI were compared. The PCCI was assigned as the reference data, being a household level census of backyard carabao related characteristics in the province, accumulated to barangay level totals, and was assumed to be without error. Of interest was to determine which among the three barangay level data sources (LPS, BAPS and LGU) were able to capture data comparably close to the PCCI barangay totals. The LPS and BAPS data were also compared to the LGU data, the data in-charge at the local office being the major KIs of LPS and BAPS.

The descriptive measures presented in Table 3 show that respectively, PCCI and LGU means (37.28 and 36.91), totals (31,500 and 31,337) and variability in terms of standard deviations (46.830 and 43.238) for number of carabao come very close numerically, while the LPS and BAPS on the same variable have higher means (43.13 and 59.63), totals (36,620 and 50,629) and standard deviations (73.11 and 99.34) respectively. These provide information of greater discrepancy between PCCI and the sources LPS and BAPS, and between LGU and the sources LPS and BAPS, compared to that between PCCI and LGU. Further, in terms of variability as a percentage of the obtained mean by each data source, the CV of LGU (118.97%) were also less than the CVs obtained for BAPS (166.59%) and LPS (169.51%), with the CV obtained for LGU even lower than the CV obtained for PCCI. These findings suggest that the LGU data on barangay level backyard carabao inventory estimates manifests closer behavior or characteristics as the reference PCCI data source, supporting the LGU as a more reliable data source than BAPS on barangay level numbers of backyard carabao in Nueva Ecija.

**Table 3. Descriptive Measures on Barangay Level Backyard Carabao Inventory by Different Data Sources in Nueva Ecija, 2010**

Data Source	Number of barangays	Min	Max	Mean	Total	Standard Deviation	CV
BAPS	849	0	1500	59.63	50,629	99.335	166.59
LGU	849	0	528	36.91	31,337	43.238	118.97
LPS	849	0	1466	43.13	36,620	73.111	169.51
PCCI	847*	0	743	37.35	31,635	46.815	125.34

\* two pairs of barangays in the municipality of Aliaga were counted as one each (San Felipe Matanda and San Felipe Bata as San Felipe, and San Pablo Matanda and San Pablo Bata as San Pablo), resulting to 847 barangays. Also, four raisers' carabao inventory were excluded since they exceed the cut-off of 20 animals for a backyard farm.

The LPS and BAPS which employed KI procedure tend to give higher estimates than the PCCI and LGU. Considering LPS and BAPS both have used the subjective FGD method, the obtained figures would usually be adjusted up

if not maintained as they are, and seldom adjusted down, since under regular situations, there is always the question whether the informant(s) had seen all the livestock in the barangay. More often, the adjustment made can be to add a few numbers to take into account possible livestock which were not seen by the informant(s), instead of subtracting from the figure declared by the informant. The reason to adjust data gathered under a KI method can be due to the deemed under-coverage of the available local offices' figures. It is difficult to control the add-up or reduction as applied, owing to the subjective process involved. The LGU on the other hand, was usually found to provide lower inventory than the PCCI due to possibly under-coverage, un-updated data, or differences in reference dates.

In terms of strengths of association, the data between PCCI and each of the LPS, BAPS and LGU on the barangay number of carabao for the province are within moderate strengths ( $0.5 < r < 0.8$ , Table 4). The PCCI-BAPS relationship was the least in strength while that of PCCI-LGU was the strongest. With regards to the relationships LGU-LPS and LGU-BAPS, the latter was lower numerically though within the same moderate degree of association.

**Table 4. Strengths of Relationship between Data from Different Sources on Barangay Level Backyard Carabao Inventory in Nueva Ecija, 2010**

Data Source		LPS	BAPS	LGU
PCCI	r	0.785	0.603	0.787
	N	845	845	845
LGU	r	0.715	0.525	
	N	849	849	

The distribution of strengths of relationships on barangay level total inventory of backyard carabao between pairs of data sources computed for each municipality are presented in Table 5 and Table 6. The strengths of relationship PCCI-LPS and PCCI-LGU were observed to be both of at least moderate degree ( $r > 0.50$ ) in 87.50 % and 93.75% of the municipalities respectively. The relationships PCCI-BAPS were of weaker strengths than that of PCCI-LPS and PCCI-LGU. Also, the relationships LGU-LPS were better than that of LGU-BAPS.

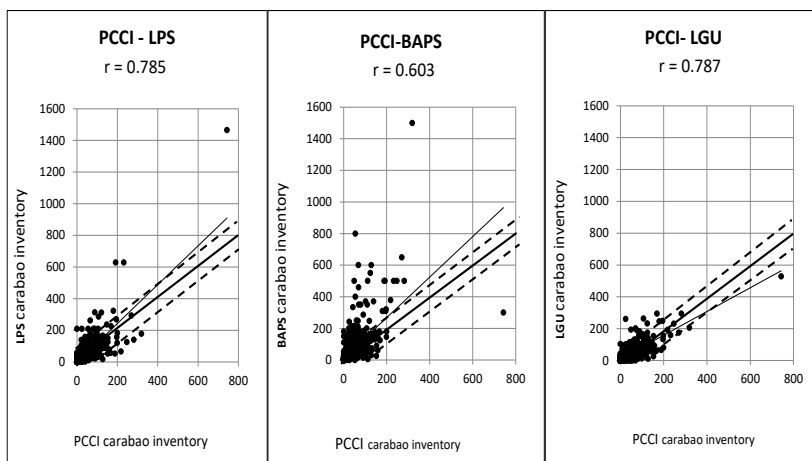
The above distribution of strengths of relationships is supported by scatterplots (Figure 3). The barangay level number of carabao between PCCI and each of the other data sources, that of PCCI-BAPS was found with the greatest number of outliers and the least number of outliers were observed in PCCI-LGU. The scatterplot in Figure 4 between LGU-BAPS also shows more outliers than that between LGU-LPS. The outlier points caused weakness in the strength of relationships between the data from the involved sources. The BAPS data has the tendency to be higher than the rest of the data sources.

**Table 5. Distribution of Municipal Level Correlations Between PCCI and Data Sources LPS, BAPS, and LGU on Barangay Level Number of Backyard Carabao, Nueva Ecija, 2010**

Range of  r	Degree of Association	PCCI – LPS			PCCI – BAPS			PCCI – LGU		
		N	%	Cum %	N	%	Cum %	N	%	Cum %
0.90-1.00	Very strong	3	9.38	9.38	2	6.25	6.25	5	15.63	15.63
0.80-0.89	Strong	11	34.38	43.75	7	21.88	28.13	8	25.00	40.63
0.51-0.79	Moderate	14	43.75	87.50	13	40.63	68.75	17	53.13	93.75
0.00- 0.50	Very weak to weak	4	12.50	100.00	10	31.25	100.00	2	6.25	100.00
Total		32	100		32	100		32	100	

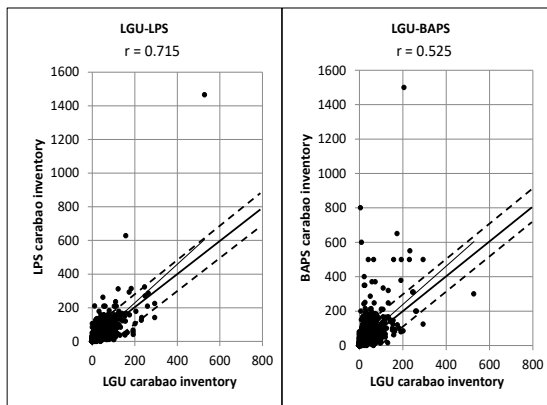
**Table 6. Distribution of Municipal Level Correlations Between LGU and Data Sources LPS and BAPS On Barangay Level Number of Backyard Carabao, Nueva Ecija 2010**

Range of  r	Degree of Association	LGU – LPS			LGU-BAPS		
		N	%	Cum %	N	%	Cum %
0.90-1.00	Very strong	5	15.63	15.63	2	6.25	6.25
0.80-0.89	Strong	5	15.63	31.25	2	6.25	12.50
0.51-0.79	Moderate	17	53.13	<b>84.38</b>	18	56.25	<b>68.75</b>
0.00- 0.50	Low to very low	5	15.63	100.00	10	31.25	100.00
Total		32	100.00		32	100.00	



**Figure 3. PCCI-LPS, PCCI-BAPS and PCCI-LGU Scatterplots on Barangay Level Backyard Carabao Inventory in Nueva Ecija, 2010**

Further, trend lines of the scatter plots of PCCI-LPS and PCCI-BAPS were rising above the 45 degree line, while that between PCCI-LGU was rising below (Figure 3), and the trend lines for the scatter plots of LGU-LPS and LGU-BAPS also rises above the 45-degree line (Figure 4). The positions of the trend lines suggest that although the PCCI is in moderate degree of relationship with each of the three data sources, the LPS and BAPS data usually have higher estimates, while the LGU estimates tends to be lower than PCCI for barangay number of carabao. Results of this section confirmed the order of performance of the LGU, LPS and BAPs data sources. For carabao related information needed for the sampling frame to be constructed, the PCCI data source will be used.



**Figure 4. LGU-LPS and LGU-BAPS Scatterplots on Barangay Level Backyard Carabao Inventory in Nueva Ecija, 2010**

### 3.3 Comparison between LPS and BAPS data on barangay level backyard cattle, goat, and swine inventory estimates

In this section, the barangay level backyard inventory estimates of LPS and BAPS for cattle, goat and swine for the entire province were evaluated against each other by evaluating descriptive measures and strengths of relationships. The compiled LGU data for these livestock types were not included in the comparison for the entire province for evidences of inclusion of large livestock operations in numerous barangays, rendering the LGU data as unfit for the comparison.

Based on the descriptive measures obtained, BAPS barangay level estimates have the tendency to give higher values than LPS, more noticeable especially for small livestock types (goat and swine), coupled with lower reliability (higher CV, Table 7). The relationships between barangay level BAPS and LPS estimates on these livestock types were not expected to result to all weak strengths ( $r < 0.50$ ), having employed the same KI technique (Table 8). Based on these observations,

the LPS data on barangay level backyard number of cattle, goat and swine in Nueva Ecija are considered as more reliable than BAPS.

**Table 7. Descriptive Measures for LPS and BAPS Barangay Level Data on Backyard Livestock Inventory Estimates for Cattle, Goat and Swine in Nueva Ecija, 2010**

Type of Livestock	Data source	Maximum	Total	Mean	SD	CV
Cattle	BAPS*	673	25,113	29.58	61.119	206.62
	LPS†	304	23,898	28.15	43.413	154.22
Goat	BAPS	2000	101,540	119.60	184.672	154.41
	LPS	500	68,248	80.39	82.088	102.11
Swine	BAPS	5790	151,861	178.87	342.719	191.60
	LPS	1060	49,089	57.82	83.411	144.26

Reference dates: \*BAPS: April 1, 2007 †LPS: March 2010  
 Minimum = 0 for all types of livestock

**Table 8. Relationships between LPS and BAPS Data on Barangay Level Backyard Livestock Inventory Estimates on Cattle, Goat And Swine, in Nueva Ecija, 2010**

Barangay level Characteristic	Correlation
Number of cattle	0.435
Number of goats	0.380
Number of swine	0.300

### 3.5. Summary of the results of the comparisons

In comparing data on common barangay level characteristics gathered by BAPS and LGU, the LGU data were found to be more reliable. Based on the findings, the auxiliary information on barangay characteristics evaluated for inclusion as auxiliary variable(s) of the household-based livestock frame (HBLF) in Nueva Ecija were based on the LGU data source.

For carabao, BAPS was consistently found to have the tendency to give higher barangay level estimates accompanied by greater variability, than LPS, LGU and PCCI data. LPS also have the tendency to give higher estimates than LGU and PCCI. In general, for all types of livestock, the same order of performance of the different data sources were observed. BAPS showed the highest consistent tendency to overestimate livestock inventory, followed by LPS. LPS estimates come close to LGU, however, LPS also tends to provide higher and more variable inventory estimates. The LGU is the reliable data source for barangay level characteristics and livestock (cattle, goat and swine) inventories supplemented by LPS on livestock related inventories, while the PCCI is the reliable data source for carabao related information.

### 3.5. The Household-Based Livestock Frame (HBLF) constructed

**Choice of basic sampling unit.** The barangays being a smaller geographical and administrative unit that are more in number than the municipalities, is assigned as the basic sampling units of the HBLF as it will allow more flexibility in the development of a sampling frame. No combining of small barangays and collapsing of large barangays were performed since the numbers of household-based livestock operations (HBLOs) do not increase with the number of households (NHH) in a barangay. The characteristics of the barangays with the minimum and maximum number of HBLOs for the different types of livestock are presented in Table 9. It was observed that población barangays have high number of households but low in number of HBLOs. The HBLF constructed is an area frame consisting of 849 barangays as basic sampling units, distributed among 32 cities/municipalities in Nueva Ecija.

**Table 9. Characteristics of Barangays with the Smallest and Largest NHH in Nueva Ecija, 2010**

Barangay, Municipality	NHH	Number of HBLOs			
		Carabao	Cattle	Goat	Swine
Mayamot II, Talugtug	32	6	1	12	0
Balante, Gapam	32	8	0	3	2
Abar 1st, San Jose City	3867	6	13	23	12

#### ***Evaluation of barangay level characteristics to be included in the HBLF.***

Barangay level characteristics as auxiliary information are to be included in the HBLF. Among the candidate auxiliary variables are the more reliable data by LGU on total land area (TLA), total agricultural area (TAA), total rice area (TRA), number of households (NHH), and number of rice growers (NRG), and population size (by CPH).

The barangay level TAA is very important, especially since for the 389 barangays with available data on this characteristic, the strength of relationships between LGU's TAA and both PCCI and LGU number of carabao and number of operations are at least moderate (Table 10). However, the TAA data of LGU was found available only for N = 389 barangays and the observations for the other barangays with no TAA data cannot be supplemented by the corresponding BAPS' data since while the LGU-BAPS relationship is moderate (0.694) for the characteristic, the relationship between BAPS' TAA-TRA is weak (0.380) compared to the very strong relationship (0.929) between LGU's TAA-TRA. TAA is one characteristic which meaning is also not very clear in the LGUs since this characteristic is often reported as equivalent to TRA by some C/MAOs, rice being the major crop in the province. For this reason, it is possible that if there is



available data for TAA, results of analyses will be close to the results for TRA. Since the variables TAA and TRA are expected to be strongly correlated, the TAA from BAPS cannot be used.

**Table 10. Relationships between TAA and PCCI and LGU Number of Carabao and Carabao Household-Based Operations in Nueva Ecija, 2010**

Data source	Number of carabao				Number of household-based carabao operations			
	PCCI		LGU		PCCI		LGU	
LGU TAA	r	N	r	N	r	N	R	N
		0.573	389	0.521	389	0.573	389	0.594

In designing a sampling scheme for the estimation of total livestock inventories, the characteristics that shows the stronger relationships with most of the livestock and livestock related inventories are good candidates to use as auxiliary variables in the different stages of the sampling design (stratification, selection and estimation), with sacrifices on the efficiency and reliability of the estimates on the livestock types which are less correlated to the chosen auxiliary variables. For the evaluation of these strengths of relationships data on barangay level cattle, goat and swine inventories were based on LGU supplemented by LPS, while number of household-based livestock operations for these livestock types were based on LPS since LGU data on number of operations is very incomplete. Carabao related inventories were based on PCCI.

Among the different barangay level characteristics, the variable whose relationships with PCCI head count and number of operations which almost got to the moderate degree were those of TRA (0.424) and NRG (0.424) (Table 11). These relationships were expected to be stronger. However, the strength of these relationships might have been obscured by the increasing use of carabao for dairy production in Nueva Ecija, the province being the impact zone of the PCC, on-going programs are conducted by the agency towards the direction of carabao as an alternative source of income through milk production, not to mention the possibility also of the increasing use of hand tractors as substitute for draft carabao. The variables TRA and NRG are the auxiliary information to be included in the HBLF.

**Table 11. Barangay Level Correlations Between Barangay Characteristics and Estimates Of Inventory and Number of Household-Based Operations for Different Types Of Livestock, Nueva Ecija 2010**

Barangay level characteristics		PCCI		LGU/LPS*			
		Carabao	HBCO	Cattle	Goat	Swine	
Total agricultural area (TAA)							
Total land area (TLA)		0.332	0.345	0.119	0.21	0.084	
Total rice area (TRA)		0.424	0.470	0.125	0.201	0.24	
Number of rice growers (NRG)		0.424	0.488	0.247	0.315	0.264	
Population (CPH)		0.248	0.224	-0.062	-0.022	0.266	
Number of households (NHH)		0.230	0.220	-0.05	-0.01	0.227	
PCCI		Inventory	1	0.911	0.089	0.185	0.241
		Operations	0.911	1	0.142	0.245	0.264
LGU/LPS	Cattle	Inventory	0.089	0.142	1	0.399	0.022
		Operations**	0.141	0.213	0.504	0.382	0.010
	Goat	Inventory	0.185	0.245	0.399	1	0.070
		Operations	0.245	0.319	0.320	0.506	0.037
	Swine	Inventory	0.241	0.264	0.022	0.070	1
		Operations	0.264	0.253	0.020	0.083	0.306

\*Data on cattle, goat and swine head counts were based on LGU supplemented by LPS

\*\*Number of household based livestock operations for cattle, goat and swine were based on LPS

With all the above results, the following decisions/features were made in connection with the HBLF:

1. Basic Sampling Units are the 849 Barangays
2. Information about the basic sampling units to be included in the HBLF:
  - a. Primary identifier: Philippine Standard Geographic Classification (PSGC)
  - b. Secondary identifier: The name of the city/municipality in which the barangay belongs
  - c. Auxiliary Variables:
    - i. Total rice area (TRA)
    - ii. Number of rice growers (NRG)
  - d. Livestock related information:
    - i. Inventory and number of HBLO estimates
    - ii. Available only for carabao: within barangay variance of number of carabao

(Sources of information: LGU/LPS 2010 and PCCI 2010)

**Development of a master sample design.** Following Latonio (2012), the constructed HBLF was used in the development of a master sample design for household based livestock surveys in Nueva Ecija, focusing on carabao. Latonio

(2012) combined adjacent barangays to form a sampling frame consisting of primary sampling units each with at least 100 carabao HBLOs, and then performed sampling exercises and to come up with a recommended master sample design for livestock surveys in Nueva Ecija.

#### **4. Conclusion and Recommendation**

The domain of the official livestock sample surveys is the province and the parameters of major concern are the provincial total inventories on different types of livestock. The focus of this paper is the construction of the sampling frame for household-based livestock operations, that is, the household-based livestock frame (HBLF) for the province of Nueva Ecija.

The constructed HBLF has 849 barangays assigned as basic sampling units distributed among 32 cities/municipalities. The choice of barangays as the possible basic sampling unit for the primary stage of sampling design will allow more flexibility in further designs of sampling frames for livestock surveys. Livestock related information such as inventory and number of household-based livestock operations estimates are specified. Auxiliary information such as total rice area (TRA) and number of rice growers (NRG) for each basic sampling unit are included as well. Each basic sampling unit (barangay) is assigned a primary identifier adopted from the Philippine Standard Geographic Classification and a secondary identifier by the name of the city/municipality in which the unit belongs.

The constructed HBLF frame can be used in the development of separate or combined sampling designs for household based livestock surveys in the Province of Nueva Ecija. Adjacent barangays as basic units can be combined to form primary sampling units for a master sample design for livestock surveys for the province as well.

#### **Acknowledgement**

CLSU, DOST-SEI, PCAARRD, PSA(BAS), PSRTI(SRTC), PCC, LGUs of Nueva Ecija

#### **References**

- BUREAU OF AGRICULTURAL STATISTICS, 2007a, Region III BAPS Manual of Operations.
- BUREAU OF AGRICULTURAL STATISTICS, 2010b. Livestock Population Survey Manual of Operations.
- COOPER, D.R. and SCHINDLER, P.S., 2014, *Business Research Methods*, 12th edition. New York. McGraw-Hill Education. p347.
- KISH, L., 1965. *Survey Sampling*. USA: John Wiley and Sons, Inc. p27.
- LATONIO, A.M.L.S., 2012. Development of a master sample design for Philippine livestock surveys, Institute of Statistics, University of the Philippines, Los Baños.

- PHILIPPINE CARABAO CENTER, 2018, Mandate, mission and vision, Available at: <http://www.pcc.gov.ph/mandatevisionmission/>, accessed July 2018
- PHILIPPINE STATISTICS AUTHORITY, 2020, Database of statistical standards and concepts in the Philippines: Official concepts and definitions. Available at: <https://psa.gov.ph/ISSiP/concepts-and-definitions/119562>
- TURNER, A. G., 2003, *Sampling Frames and Master Samples*, Available at: [https://unstats.un.org/unsd/demographic/meetings/egm/sampling\\_1203/docs/no\\_3.pdf](https://unstats.un.org/unsd/demographic/meetings/egm/sampling_1203/docs/no_3.pdf)