Chapter – III

Methodology and Data Collection

Chapter-III

Methodology and Data Collection

3.1 Introduction

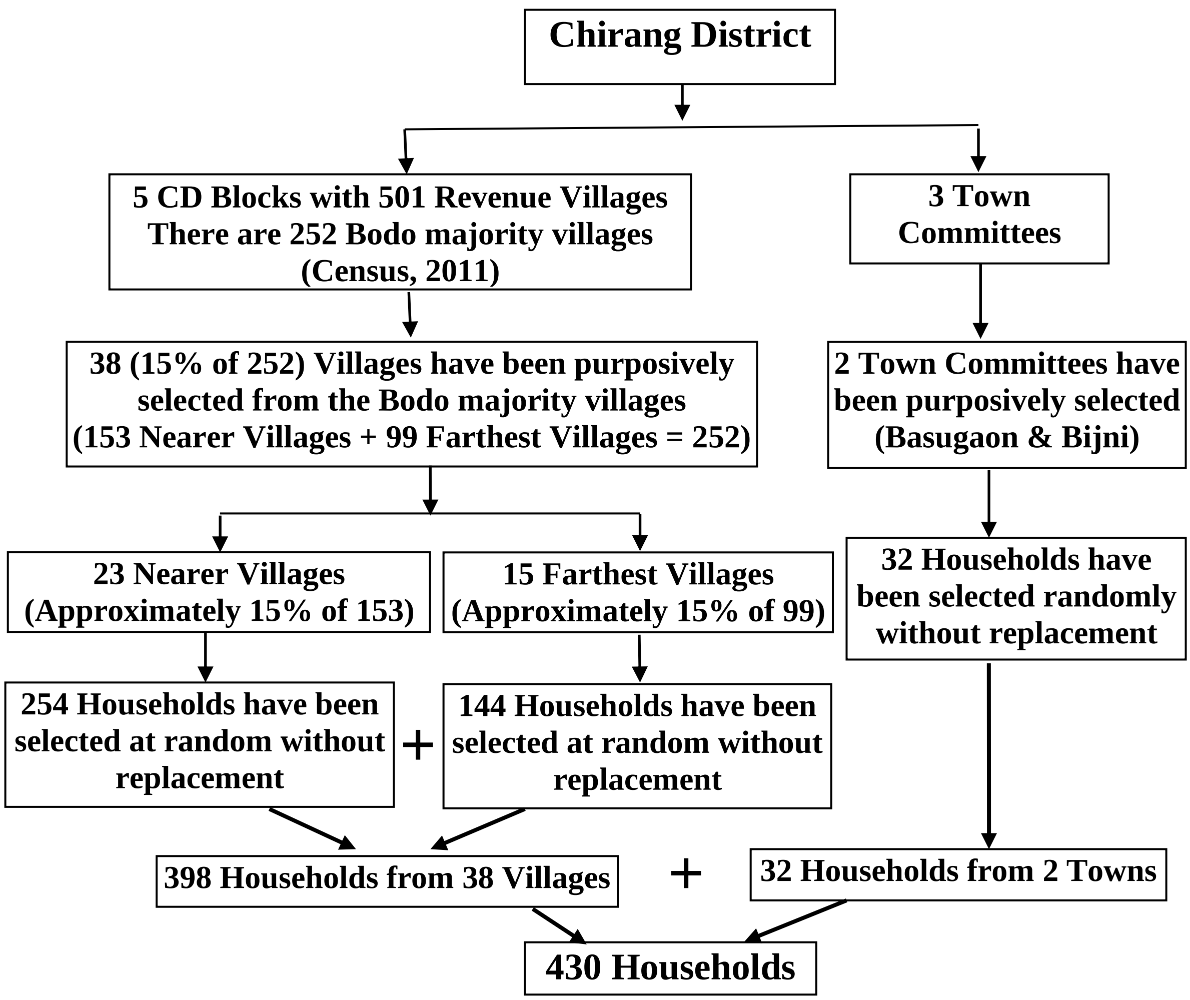
The methodology is the application of a scientific method or procedure for searching for answers to meaningful questions. The methods applied for investigation make a study scientific, reliable and precise. The research methodology includes specific procedures or techniques used for identifying, selecting, processing, and analysing information about a topic. This section explains the methodology used for this study. It includes the nature and the types of data collected for it, sample design, the tools and models used for analysing data.

3.2 Data and Sample Design

Though this study is based on primary data, secondary data is used to investigate the variations in the socio-economic status of all the districts of Assam. The people of the Bodo community in the Chirang district are only the targeted group of this study. Bodoland Territorial Area Districts (BTAD) of Assam came into existence under the BTC Accord in February 2003. In any study related to the Bodo community, in general, BTAD comes into the centre of references. A comparison of the socio-economic status of the districts of Assam is done based on secondary information for specifying the case of the BTAD area as well as of Chirang district. For this purpose, the direct observation method has been applied. Accordingly, secondary information on numbers of institutions, total population, literacy rate, life expectancy rate, crime rate, sex ratio etc. and all other required secondary information have been collected from various published sources like books, magazines, newspapers, journals, official circular, different statistical report, internet, website etc. References of the secondary information have been given at the time of discussion when they are necessary.

Primary data have been used for calculating the Socio-economic Index, Human Development Index and Multidimensional Poverty Index for the Bodo people of Chirang district. The method of multistage random sampling is applied to collect the primary data. There are 5 (five) Community Development blocks and 3 (three) Town Committees (TC) in the Chirang district. The CD blocks have 501 revenue villages. In 252 revenue villages, Bodo people are either more than 50 percent or cent percent and they are assumed to be Bodo majority inhabited villages in Chirang district. Out of these 252 villages, 38 (15 percent of 252) villages have been selected purposively considering the percentage of Bodo people living there (Table: 3.1).

Table 3.1 Sample Design



Location distance of a village from its nearest town is an important factor of socio-economic status of the people living there. A village located nearby a town gains more infrastructural advantages than that of the farthest village. So, the distance of a village from its nearest town has been considered at the time of purposive sample collection. A sample village is located at a distance of up to 25 km. the roadway from its nearest town may be considered as a nearer village from its nearest town (Kachari & Maity, 2015). A village nearer to a town may expect a positive impact on the socio-economic status of the people of the village and so we may assign 1 as a dummy value against this village. Accordingly, the categories of farthest villages will be assigned 0 values as their dummy value. But, the Chirang district is located at the international border area of the neighbouring country Bhutan. And it has lots of forest villages with poor quality connecting roads of at least 20km. towards semi-urban areas like Basugaon, Bijni, Bongaigaon, and Shorbhog etc. which are the highest infrastructure accommodated areas situated within or nearest to Chirang district. Therefore, in the case of Chirang district, a maximum 20 km roadway distance of a village from its nearest towns may be considered as the nearer village and otherwise, it is the farthest village.

Thus, Category 1 includes nearer villages (roadway distance less than 20 km towards nearest town) and category 0 includes farthest villages (roadway distance 20 km or more towards nearest town) as stated in Table: 3.2. This sample has 15 farthest villages, 23 nearer villages and two towns.

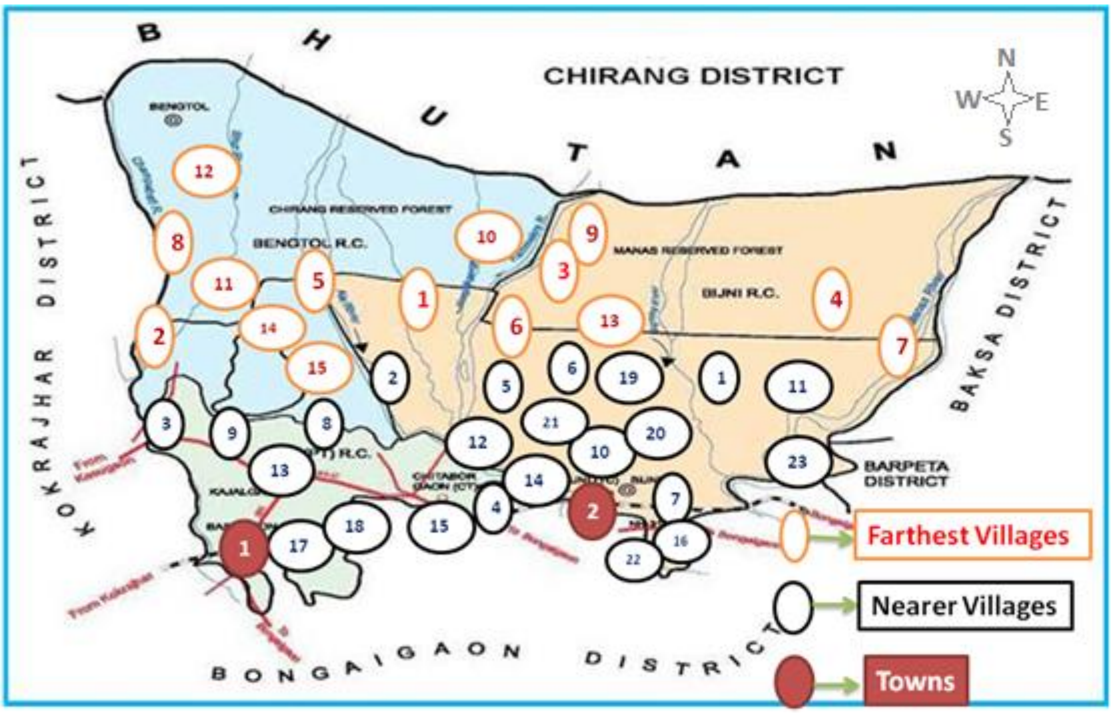
A map indicating sample villages and towns is given below where orange-coloured ovals indicate the farthest villages and the black coloured ovals indicate the nearer villages (Map 2). On the other hand, the red-coloured ovals stand for sample towns.

Table 3.2 Sample Collection

Sl. No.	Village/Town	No. of Households	ST Population in %.	Distance from the nearest town	Category of village	No. of Sample Households
Sample Villages beyond 20 km. from their nearest town (Farthest villages)						
1	Amguri	293	100	27	0	15
2	Aminpara	249	98.08	46	0	10
3	Amteka	531	68.26	30	0	15
4	Bhatarmari	120	100	35	0	8
5	Bikrampur	223	96.08	25	0	10
6	Dimajhora	223	72.45	22	0	8
7	Kahitama	381	80.14	32	0	13
8	Khungring	209	97.39	45	0	8
9	Koila Moila	102	70.04	30	0	6
10	Ouguri	177	95.14	23	0	8
11	Patabari	213	96.47	37	0	9
12	Salbari Bhurpar	249	85.67	42	0	8
13	Subaijhar	154	100	22	0	6
14	Tangabari	255	72.55	25	0	9
15	Uttar Runikhata	253	84.21	32	0	11
Sub Total						**144**
Sample Villages within 20 km. of their nearest town (Nearer villages)						
1	Betnapara	237	86.54	15	1	10
2	Baldi No.2	190	100	18	1	9
3	Chamugaon	178	92.55	15	1	8
4	Chapaguri	669	57.11	3	1	19
5	Dahalapara	200	94.98	8	1	9
6	Dakhin Makra	411	70.18	16	1	13
7	Dangaigaon	227	91.04	5	1	12
8	Deulguri	212	73.49	7	1	10
9	Duttapur	910	68.52	17	1	30
10	Gargaon No.1	161	96.89	4	1	7
11	Kachubil No.1	168	83.31	12	1	5
12	Kahibari	128	97.59	8	1	6
13	Kashikotra	366	72.47	5	1	14
14	Khamarpara	235	98.06	15	1	10
15	Kukurmari	761	50.99	3	1	21
16	Kumargaon	185	91.24	7	1	9
17	Maigaon	190	81.09	15	1	7
18	Nilibari	423	65.74	9	1	13
19	Oxiguri	530	70.45	15	1	15
20	Patkiguri	136	99.7	17	1	6
21	Silbari Abadipara	290	71.87	13	1	9
22	Sukhanipara	136	99.48	12	1	6
23	Uttar Burikhamar	147	90.43	15	1	6
Sub Total						**254**
Sample Town Committees						
1	Basugaon	3039	5.12	0	1	13
2	Bijni	2926	10.91	0	1	19
Sub Total						**32**
Sample Size						**430**

Source: Primary data survey.

Map 2 Sample Villages and Towns



Source: www.mapesofindia.com

Regarding the location of the town committees in Chirang district, Basugaon TC is located in the Southwest corner of the district and is located at the border area of two districts Kokrajhar (Westside) and Bongaigaon (at the Southside). Chatiborgaon TC is the nearest to district headquarter Kajalgaon. Bijni TC is the only sub-division of the district located in the east part of the district. All of these TCs are not fully town but may be considered as semi-urban areas. Basugaon and Bijni are taken for a sample survey to get a full representation of the district subject to Bodo inhabitancy in the Chirang district.

Bodo people belong to Scheduled Tribe (ST) in Assam. According to Census-2011, the ST population in Chirang district is 178688 and they are 37.06 percent of the total population of the district. As per the "PCA CDB-1821-F-Census.xlsx" sheet (downloadable from http://censusindia.gov.in/pca/cdb_pca_census/Houselisting-housing-Assam.html), about 148196 ST people live in the 252 Bodo majority villages including Basugaon TC, and Bijni TC. Again, the 252 Bodo majority villages, Basugaon TC and Bijni TC have 38513 households which are 39.67 percent of the total households 97092 in Chirang district. Finally, 430 Bodo households have been selected at random proportionately assigning the percentage of ST people of the sample villages/ towns as weightage for calculating the number of sample households (Table 3.2). The number of ST households in the Chirang district is 360502. The sample size is taken for this study is 430 Bodo households.

3.3 Methodology for Calculating Socio-Economic Index (SEI)

There are numbers of indices devised by the researchers over the years, including Duncan's index for occupational categories according to education and income (Oakes & Rossi, 2003), Townsend's index designed to study the link between health and material deprivation (Morris & Castairs, 1991), and the Living Conditions Index developed by the Social and Cultural Planning Office of the Netherlands to measure inequities in housing, health, etc. (Boelhouwer & Stoop, 1999). This study has adopted the methodology of calculating Socio-Economic Index (SEI) for Bodo people in the study area which was developed and used for constructing area-based as well as a community-based socio-economic index using the technique of Principal Components Analysis (PCA) by various researchers and scholars. As the initial approach, Boelhouwer and Stoop (1999) used PCA for the first time to combine socioeconomic indicators into a single index. Later on, Lai (2003) modified the UNDP Human Development Index by using PCA to create a linear combination of indicators of development. Now, the PCA approach to measuring socio-economic index has been using frequently by researchers in socio-economic research (Fotso & Kuatedefo, 2005; Rygel, O'Sullivan, & Yarnal, 2006; Vyas & Kumaranayake, 2006; Antony & Rao, 2007; Fukuda, Nakamura, & Takano, 2007; Havard, et al., 2008; Messer, et al., 2008; Krishnan,2010; Maity, Haobijam and Sen, 2014; Kachari, 2015). The methodology for calculating the socio-economic index of Bodo people in Chirang district applied for this study is based on the methodologies particularly adopted by Vyas & Kumaranayake (2006), Maity, Haobijam and Sen (2014) and Kachari (2015).

3.3.1 Socio-economic Variables Influencing Socio-economic Status of the Districts of Assam

Depending on the availability and accessibility of data at the district level, we consider the following socio-economic variables as per their sources of information to study the district wise variation in socio-economic status. The variables have been classified as demographic variables, social variables and economic variables (Table 3.3).

Table 3.3 Sources of Socio-Economic Variables at District Level, Assam

Variables	Sl. No.	Variables at District Level	Sources
Demographic Variables	1	The area in Sq. Km. of the districts	Statistical Hand Book of Assam 2019
	2	Population of 2011	
	3	Decadal Growth Rate of Population 2001-11	
	4	Share of District Population to State Population in %	
	5	The Density of Population per sq. km.	
	6	Rural Population in %	
	7	Urban Population in %	
	8	Life Expectancy at Birth	Assam Human Development Report
	9	Infant Mortality rate	
Social Variables	10	Literacy Rate (Census 2011)	Statistical Hand Book of Assam 2019
	11	School Dropout Rate	
	12	Transition Rate from Lower Primary to Upper Primary Level	
	13	% of Households Accessing Electricity for Lighting	National Family Health Survey-4, 2015-16
	14	% of Households Accessing Safe Drinking Water	
	15	% of Households Accessing Improved Sanitary Facility	
	16	% of Households Accessing Clean Fuel for Cooking	
	17	% of Households Accessing Pucca House	
	18	Crime Rate Per-Lakh Population	Statistical Hand Book of Assam 2019
	19	Road Length Per-Lakh Population at District Level	
Economic Variable	20	Per-Capita Income (PCI)	Economic Survey Assam, 2017-18

To estimate the Socio-Economic Index (SEI) for the districts of Assam, we use the method of Principal Component Analysis (PCA). Principal Component Analysis (PCA) is a technique for transforming a large number of correlated variables into a smaller set of uncorrelated (orthogonal) factors called principal components. Through an extensive review of the literature, the following 11 socio-economic variables have been selected to estimate the SEIs for the districts of Assam.

1. DP = Density of Population
2. PUP = Percentage of Urban Population
3. LR = Literacy Rate
4. HAEL = Households Accessing Electricity for Lighting
5. HASDW = Households Accessing Safe Drinking Water
6. HAISF = Households Accessing Improved Sanitary Facility
7. HACFC = Households Accessing Clean Fuel for Cooking
8. HAPH = Households Accessing Pucca House
9. CR = Crime Rate per-lakh population
10. RL = Road Length per-lakh population at district level
11. PCI = Per-Capita Income

3.3.2 Factors Influencing the Socio-economic Conditions of Bodo Households

From the review of literature, we got lots of factors influencing the socio-economic conditions of Bodo households. The entire set of variables may be classified into three categories namely,

(i) Social variables
(ii) Demographic variables
(iii) Economic variables

This study has considered five socio-economic variables for calculating the socio-economic index of Bodo households in the study area. We consider two social variables, two demographic variables and the other is an economic variable.

These five variables are taken into consideration keeping in mind the case of avoiding the possibility of a multicollinearity problem.

3.3.2.1 Social Variables

The two social variables considered for this study are,

a. Literacy status of the Bodo households
b. Whether the Bodo households of a village is near to a town or otherwise.

By following the Alkire and Foster (2007, 2011a) method of identifying households' literacy status, we consider that a household is not literate if no household member has completed five years of schooling. They have assumed that the benefits of any literate member are enjoyed equally by all household members. Therefore, a Bodo household is said to be not literate if no one household member has completed schooling up to class-V. Literacy or education is one of the variables positively influencing the socio-economic status of any individual. This variable is considered on the assumption that a literate household gains better socio-economic status than an illiterate household. Here, we write this variable as LITERATE and assign dummy literacy status as,

LITERATE (Literacy Status) = 0 if no household member has completed five years of schooling up to Class-V.

= 1 if at least one household member has completed schooling up to the Class-V.

The second social variable is the distance (DISTANCE) between the village where households live and its nearest town. A village located nearby a town gains better infrastructural facilities than the village farthest away from its nearest town. Similarly, a village with good infrastructure must have better socio-economic status. Therefore, a village nearer to a town, higher will be the socio-economic status of the Bodo households living there. The Chirang district is located at the foothills of the Himalaya nearby the international border of neighbouring county Bhutan and it has lots of forest villages, many rivers and tributaries. Considering geographical features and observing road and communication and other facilities at the time of field survey in Chirang district, the roadway distance of 20 km will be the maximum limit beyond which a village may be assumed as a farthest village to its nearest town. Therefore, a household is said to be nearer to the nearby town if its village is located within 20 km roadway distance from its nearest town. We assign 0 as a dummy for nearer village and 1 for otherwise.

DISTANCE (Distance) = 0, if the village is nearer to town

= 1, otherwise.

3.3.2.2 Demographic Variable

This study considers two demographic variables- family size (FSIZE) and family members of the age group of 15-59 years (FAGE). Family size is the number of family members of a household. In general, there is an inverse relationship between the size of the family and the socio-economic status of the respondent (Rao and Rao, 2010). On the other hand, family earnings depend on family members of the age group of 15-59 years.

3.3.2.3 Economic Variable

Here we consider one economic variable and it is the annual per head household income. Income is another important factor that determines the socio-economic conditions. A household with higher income enjoys better socio-economic conditions compared to a neighbouring household with lower income. Information about the household's annual income was collected at the time of the sample survey which took place from October 2018 to March 2019. The household income has been converted into a constant price at 2013-14. This study has taken annual per head income at the household level to estimate the household socio-economic index.

3.3.3 Principal Components Analysis (PCA)

Principal Component Analysis (PCA) is a useful technique for transforming a large number of correlated variables into a smaller set of uncorrelated (orthogonal) factors called principal components. The principal components account for much of the variance among the set of original variables. Each component is a linear weighted combination of the initial variables. The components are ordered so that the first component accounts for the largest possible amount of variation in the original variables. The second component is completely uncorrelated with the first component and accounts for the maximum variation that is not accounted for in the first. The third accounts for the maximum that the first and the second not accounted for and so on.

In general, the factor analysis encompasses both the techniques PCA and principal factors analyses. In most cases, these two methods yield similar results. However, PCA is preferred for data reduction while principal factor analysis is preferred for detecting the structure of the data set. The PCA is an approximation to principal factor analysis when components are rotated. The matrix of the scores of principal components is called the Rotated Component Matrix.

Before going to factor analysis, it is necessary to carry some statistical tests for the verification of appropriateness of the data set. Otherwise, the factor analyses on findings may become misleading.

First of all, there may have some extreme values in the data set called outliers which affect the normalcy of the data set. Different statistical tools like histogram, normal Q-Q plot, box-plot and 5 percent trimmed mean are applied to detect the outliers present in the distribution of a data set. The histogram of the data set is applied to examine the actual shape of the distribution. The normal Q-Q plot is used to plot the observed value for each score against its expected value. The box-plot of the distribution of scores is used to identify any score which lay beyond an outer fence and a 5 percent trimmed mean helps us to exclude the extreme values. Various SPSS procedures on such statistical tools may be carried out at the time of factor analysis. Moreover, some of the descriptive statistics like mean, skewness, kurtosis and range help us to detect the type of distribution of the data set.

The multicollinearity problem arises when there are strong correlations among the variables. Multicollinearity may increase the standard error of factor loadings causing less reliability. Multicollinearity problems may be reduced by either combining collinear variables or doing eliminate them. Some researchers use factor analysis if the variables show multicollinearity and some others forgo factor analysis altogether. But, the Kaiser-Meyer-Olkin (KMO) is a Measure of Sampling Adequacy (MSA) which helps us to handle multicollinearity problems so that the appropriateness of the data set for carrying out a factor analysis can be detected. More specifically, sampling adequacy predicts if data are likely to factor well, based on correlations and partial correlations. The KMO measure compares the magnitudes of the observed correlation coefficients to the magnitudes of the partial correlation coefficients. The formula for the KMO test is as given below,

$$\mathrm{KMO} = \frac{\sum_{i=1}^{p}\sum_{\substack{j=1 \\ i\neq j}}^{p} r_{ij}^2}{\sum_{i=1}^{p}\sum_{\substack{j=1 \\ i\neq j}}^{p} \rho_{ij}^2 + \sum_{i=1}^{p}\sum_{\substack{j=1 \\ i\neq j}}^{p} r_{ij}^2} \qquad (3.1)$$

Where,

$$\rho_{ij} = \frac{R_{ij}}{\sqrt{R_{ii} \, . \, R_{jj}}}$$

$$r_{ij} = R(X_i, X_j)$$

Here, ρ_{ij} stands for partial correlation coefficient and r_{ij} stands for the Pearson correlation coefficient.

KMO test assumes that lower the partial correlation coefficients compared to total correlation coefficients indicate more sampling adequacy. The KMO measure ranges from 0.00 to 1.00. Therefore, the higher the value of the KMO test indicates that the data is more adequate for factor analysis. It suggests six ranges of values for deciding sampling adequacy. It suggests that the value lies 0.00 to 0.49 indicates unacceptable, 0.50 to 0.59 indicates miserable, 0.60 to 0.69 indicates mediocre, 0.70 to 0.79 indicates middling, 0.80 to 0.89 indicates meritorious and 0.90 to 1.00 indicates marvellous (Antony & Rao, 2007; Planning Commission, 1993).

Bertlett's (1954) Test of Sphericity is a test for determining the strength of the relationship among variables. This test was done taking the null hypothesis that the population correlation matrix was an identity matrix or the variables are uncorrelated in the population correlation matrix. According to Bartlett's test of sphericity, a small value of significance level less than 0.05 rejects the null hypothesis at a 5 percent level of significance.

The basic principle of PCA is to extract a set of new uncorrelated variables (principal component) Z_i (i=1, 2, ……,k) as the linear combinations of original variables X_j (j=1, 2, ……,k). Here, a new variable Z_i (i=1, 2, ……,k), is known as i^{th} Principal Component and it is given by the linear combinations of X_j's as given in equation 3.2.

$$\left.\begin{aligned} Z_1 &= b_{11}X_1 + b_{12}X_2 + b_{13}X_3 + \dots\dots\dots\dots + b_{1k}X_k \\ Z_2 &= b_{21}X_1 + b_{22}X_2 + b_{23}X_3 + \dots\dots\dots\dots + b_{2k}X_k \\ &\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ &\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ Z_k &= b_{k1}X_1 + b_{k2}X_2 + b_{k3}X_3 + \dots\dots\dots\dots + b_{kk}X_k \end{aligned}\right\} \qquad (3.2)$$

This method is applied mostly by standardizing the variables using the formula defined by equation no 3.3.

$$Z = \frac{X - \text{Mean}}{\text{Standard Deviation}} \qquad (3.3)$$

Where, the 'b_{ij}' s are called the factor loadings. The b_{ij}s are determined in such a way so that

a) Principal components are uncorrelated, that is, orthogonal, and

b) The first principal component has the maximum variance followed by the second, third and so on.

Keiser's criterion indicates that we should consider only those Principal Components for whom the eigenvalues or latent roots are greater than one. The Principal Components so extracted or retained are then rotated from their beginning position to enhance the interpretability of the factors. Communality or h^2 value shows how much of each variable is accounted for by the factors retained in PCA. A high communality value means that not much of the variables are left over after whatever the factors represent is taken into consideration. So,

h^2 of the i^{th}, variable = (i^{th} factor loading of 1^{st} factor)^2 +(i^{th} factor loading of 2^{nd} factor)^2 + ……… (3.4)

The amount of variance explained by each principal factor is equal to the corresponding root. Factor scores (f_{jk}) are obtained by regressing the variables on factor loadings. f_{jk} measures the position of the j^{th} Bodo household with others concerning the k^{th} factor (Singh and Das, 2013).

3.3.4 Calculation of Socio-Economic Index (SEI)

To compute the Household Socio-Economic Index (HSEI), the factor scores and the corresponding weights are used. The formula for calculating the $HSEI_j$ of j^{th} household is,

$HSEIj = \sum w_{kj} \, f_{jk}$ for all j = 1, 2, …………, k (3.5)

Where,

$HSEI_j$ = Household-level Socio-economic Index of j^{th} Bodo household

w_{kj} = the percentage of the variation of the k^{th} factor

f_{jk} = factor score of the k^{th} factor.

This index measures the socio-economic status of one Bodo household relative to the other on a linear scale. The value of the index can be positive or negative, making it difficult to interpret. Therefore, a non-standardised household socio-economic index (NSHSEI) is standardized to a scale of 0-100 using the following formula,

$$\mathrm{SHSEI_i} = \frac{\mathrm{NSHSEI_i} - \mathrm{MinNSHSEI}}{\mathrm{MaxNSHSEI} - \mathrm{MinNSHSEI}} * 100 \qquad (3.6)$$

Where,

$SHSEI_i$ = standardized household socio-economic index for i^{th} Bodo household

$NSHSEI_i$ = non-standardised household socio-economic index for i^{th} Bodo household

MinNSHSEI = minimum value among the non-standardized socio-economic index of the Bodo households

MaxNSHSEI = maximum value among the non-standardized socio-economic index of the Bodo households

The Village Socio-Economic Index (VSEI) of a village is obtained by averaging the Household Socio-Economic Index (HSEI) of the households of that village. The formula for calculating VSEI for j^{th} village is,

$$\mathrm{VSEI_j} = \frac{1}{N} \sum_{1=0}^{N} \mathrm{HSEI_i} \qquad j = 1, 2, \dots\dots\dots.., 40 \qquad (3.7)$$

Where, N is the number of households of the j^{th} village.

3.4 Methodology for Calculating Human Development Index (HDI)

Methodologies for calculating HDI may be classified into two types as old methodology and new methodology. UNDP's methodology used for the calculation of HDI in the years 1990 and 2009 is considered by researchers as an old methodology. On the other hand, there are methodological modifications in the calculation of HDI since the year 2010. It is known as the new methodology for the calculation of HDI (UNDP, 2010). The new methodology has been applied here for calculating HDI for the sample villages to know the human development of Bodo people in the Chirang district. However, in calculating Income Index the methodology used in the Assam Human Development Report, 2014 is followed. In UNDP methodology, the income index is calculated by per-capita income at Purchasing Power Parity (PPP). But AHDR, 2014 calculated the income index by using surveyed household income data directly. It is beyond the scope of this study to explain the methodological differences between old and new methodology. We state both these two only to see preliminary differences in calculating HDI by these two methodologies.

3.4.1 Old Methodology

According to the methodology of UNDP's Human Development Report-1990, HDI is the average of three indices namely- health index, educational index and standard of living index. The health index is for achievements in life expectancy. The educational index is calculated from the child enrolment ratio and adult literacy rate. And the standard of living index or income index is calculated from GDP per-capita at the purchasing power parity of a country. All these three indices are normalised by using the formula,

$$\text{Index} = \frac{\text{Actual Value} - \text{Minimum Value}}{\text{Maximum Value} - \text{Minimum Value}} \qquad (3.8)$$

The formula for calculating HDI is

$$\text{HDI} = \frac{\text{Health Index} + \text{Educational Index} + \text{Income Index}}{3} \qquad (3.9)$$

3.4.2. New Methodology

Instead of using a simple average, the new methodology (UNDP 2010) used geometric mean to estimate the Human Development Index (HDI). According to the new methodology, HDI is the geometric mean of Life Expectancy Index, Education Index and Income Index.

a. Life Expectancy Index

Life Expectancy Index is calculated from life expectancy at birth. Life expectancy is the year of life that a child can expect to live at the time of his/her birth. The formulas for calculating the life expectancy index is,

$$\text{Life Expectancy Index} = \frac{\text{Actual Life Expectancy} - \text{Minimum Life Expectancy}}{\text{Maximum Life Expectancy} - \text{Minimum Life Expectancy}} \qquad (3.10)$$

Where Minimum Life Expectancy is 20 and the Maximum Life Expectancy is 85.

Using Chiang Method[1], the life expectancy for Bodo people has been calculated at village level based on the records of death and birth obtained from the head man of the village called 'Gaonburha'.

b. Education Index

The Education Index is an average of the index for Mean Year of Schooling (MYS) and index for Expected Years of Schooling (EYS).

$$\text{Education Index} = \frac{\text{MYS Index}+\text{EYS Index}}{2} \qquad (3.11)$$

Where,

$$\text{MYS Index} = \frac{\text{MYS}}{15} \; and \qquad (3.12)$$

$$\text{EYS Index} = \frac{\text{EYS}}{18} \qquad (3.13)$$

----

1. Chiang method is a widely used technique of calculating life expectancy based on life tables and it has been using by leading statistical agencies like WHO, the Office for National Statistics in the UK and UNDP.

The average year of institutional education is known as the mean year of schooling (MYS). Fifteen (15) years MYS is the projected maximum value for 2025. The expected year of schooling (EYS) is the year of institutional education that a child can expect to complete in his/her life. Eighteen (18) years is considered as the maximum value of EYS because it is equivalent to achieve a master's degree in most countries.

c. Income Index

The formula for calculating Income Index is,

$$\text{Income Index} = \frac{\log(\text{Actual Income}) - \log(\text{Minimum level Income })}{\log(\text{Maximum level Income}) - \ln(\text{Minimum level Income})}$$

(3.14)

This study has used household income data obtained from a sample survey. The annual per-head household income has been calculated at village level at a constant price of the financial year 2013-14. Similar to the AHDR, 2014, the minimum level of income of Rs. 5090/- as and the maximum level of income of Rs. 119032/- are used in this study for normalisation of the income index.

d. HDI Calculation Formula

Finally, the formula for calculating HDI is,

$$\text{HDI} = \sqrt[3]{\text{Education Index} * \text{Income Index} * \text{Life Expectation Index}}$$

(3.15)

Thus, according to UNDP-2010 methodology, HDI is the geometric mean of Education Index, Income Index and Life Expectancy Index.

3.5 Methodology for Calculating the Multidimensional Poverty Index (MPI)

The deprivation has traditionally been measured in one dimension, usually income. Poverty is a measure of deprivation. But deprivation is not one-dimensional, it is multidimensional. Similarly, poverty is multidimensional. A single indicator, such as income, can not uniquely able to capture the multiple aspects that contribute to poverty (Alkire, 2011). This is the reason due to which the Multidimensional Poverty Index (MPI) comes into existence as a framework for assessing the capability deprivation. The methodology for measuring MPI was proposed by Alkire & Foster (2007, 2009). The Human Development Report, 2010 estimated MPI for the first time by replacing the Human Poverty Index (HPI). MPI is a measure for multiple deprivations at the household and individual level in the areas of education, health and living standard. As a measure of acute poverty, MPI bears two characteristics. Firstly, it includes people living under conditions where they do not reach the minimum internationally agreed standards in indicators of basic functioning. And secondly, it considers people living under conditions where they do not reach the minimum standards in several aspects at the same time (Santos, 2010). The Multidimensional Poverty Index (MPI) assesses both the nature and intensity of poverty at the individual level.

3.5.1 Dimensions and Indicators of MPI

The MPI is composed of three dimensions namely, education, health and standard of living. These three dimensions are measured by using ten indicators. The dimensions of health, education include two indicators each and the standard of living includes six indicators. Each indicator weighted equally within the dimensions associated with each indicator is a minimum level of satisfaction, which is based on international consensus (such as the Millennium Development Goals). This minimum level of satisfaction is called a deprivation cut off.

3.5.1.1 Education

Education is represented by years of schooling and child school enrolment. Years of schooling is assumed as a proxy for literacy and level of understanding of the members of a household. Basu & Foster (1998) assumed that all the family members of a household are benefited from anyone literate family member. A household is said to be deprived of if no family member has completed five years of schooling. In the case of enrolment indicator, a household is said to be deprived of if anyone school aged child is not attending school in class-I to VIII.

3.5.1.2 Health

Health is represented by child mortality and malnutrition. A household is deprived of mortality if any child has died in the family before the completion of age 14. In the same way, a household is deprived of if at least one undernourished family member is in it.

3.5.1.3 Standard of Living

The living standard is represented by access to electricity, clean drinking water, improved sanitation, flooring (no dirty, sand or dung floor), clean cooking fuel, and asset index. Clean drinking water depends on the safe sources of drinking water. Supplied water, water from deep boar whole, covered and good filtered water are considered as safe sources of drinking water. According to Millennium Development Goals (MDGs), an "improved" sanitation is as the kind of toilets like Flush toilet, Connection to a piped sewer system, Connection to aseptic system, Flush/ pour-flush to a pit latrine, Pit latrine with slab, Ventilated improved pit latrine (abbreviated as VIP latrine) or Composite toilet. A household is deprived of improved sanitation if it has not such type of sanitation. A household is considered as deprived of an asset if it has not at least one asset related to access to information (radio, television or telephone) or having at least one asset related to information but not having at least one asset related to mobility (bike, motorbike, car, truck, animal cart or motorboat) or at least on asset related to livelihood (refrigerator, arable land or livestock)

3.5.2 Weights to Indicators

The MPI is a weighted indicator and weights can be applied in three ways in the multi-dimensional poverty measures; (i) between dimensions (the relative weight of health and education), (ii) within the dimensions(if more than one indicator is used), (iii) among the people in the distribution, for example, to give greater priority to most of the disadvantaged.

The MPI explicitly weights each dimension equally and each indicator within the dimension equally. Equal weighting between the dimensions is an outcome of the HDI convention. The maximum score is 10, with each dimension equally weighted. Therefore, the maximum score in each dimension is ⅓. The health and education dimensions have two indicators each, so the weight for each component is (1⁄3) ÷ 2 = 0.167 or 16.7%. The standard of living dimension has six indicators, so each component is worth (1⁄3) ÷ 6 = 0.056 or 5.6% (Table 3.4).

3.5.3 Poverty Cut-off 'C'

The method of poverty identification is based on the dual cut-off method of Alkire & Foster (2011a). According to this method, first, indicators of dimensions of MPI are identified which are called indicator cut-offs. And then MPI assigns equal weights across dimensions and within each dimension indicators are weighted equally. Each of the three dimensions gets an equal weight of 1/3 or 33.3%. Education and health have two indicators in each. Therefore, distributing 33.3% equally into two indicators, 16.7% is weighted in each indicator of education or health dimension (Table 3.4). Similarly, distributing 33.3% equally among six indicators of living standard dimension, each indicator is weighted approximately by 5.6%.

Table 3.4 Dimensions of MPI with Indicators' Weights

Dimensions	Indicators		Indicator weight
Education	I	No one has completed five years of schooling	**(1⁄3) ÷ 2 = 16.7%**
	II	At least one school-age child not enrolled in school	**(1⁄3) ÷ 2 = 16.7%**
Health	I	At least one member is malnourished	**(1⁄3) ÷ 2 = 16.7%**
	II	One of more children have died in the family age	**(1⁄3) ÷ 2 = 16.7%**
Living Conditions	I	No electricity	**(1⁄3) ÷ 6 = 5.6%**
	II	No access to clean drinking water	**(1⁄3) ÷ 6 = 5.6%**
	III	No access to adequate sanitation	**(1⁄3) ÷ 6 = 5.6%**
	IV	House has a dirty floor	**(1⁄3) ÷ 6 = 5.6%**
	V	Household uses "dirty" cooking fuel (dung, firewood or charcoal)	**(1⁄3) ÷ 6 = 5.6%**
	VI	Household has no access to information and has no access related to mobility or access related to livelihood	**(1⁄3) ÷ 6 = 5.6%**
Household deprivation score- 'C' (sum of each deprivation multiplied by its weight)			
A household is multidimensionally poor if C ≥ 33.3 percent.			

Source: UNDP Methodology 2016.

N.B.- Assets: not having at least one asset related to access to information (radio, television or telephone) or having at least one asset related to information but not having at least one asset related to mobility (bike, motorbike, car, truck, animal cart or motorboat) or at least on asset related to livelihood (refrigerator, arable land or livestock)

According to Alkire and Foster method, the MPI cross-dimensional cut-off is one third. Therefore, a household is multidimensionally poor if it's weighted deprivations sum up to one third or more. In other words, if a household's total deprivation score is 33.3 or more (i.e.≥ 33.3 percent), then the household is said to multidimensionally poor. If the deprivation score is 20 percent or more but less than 33.3 percent, households are near multidimensionally poor. Households with a deprivation score of 50 percent or more are said to be severely multidimensionally poor.

3.5.4 Calculation of MPI

Multi-dimensional poverty consists of two numbers, the headcount ratio (incidence) and the intensity (or breadth) of poverty.

3.5.4.1 Head Count Ratio (Incidence of Poverty)

According to MPI, one is poor when he deprives at least one-third of the weighted deprivation. The MPI combines two key pieces of information: (1) the proportion or incidence of people (within a given population) who experience multiple deprivations and (2) the intensity of their deprivation: the average proportion of (weighted) deprivations they experience. The Headcount ratio (H) is the proportion of the population who are multi-dimensionally poor.

$$\mathbf{H} = \frac{\mathbf{q}}{\mathbf{n}} \qquad (3.16)$$

Where **q**= is the number of persons who are multidimensionally poor and

n= is the total population

3.5.4.2 The Average Intensity of Poverty

The intensity of poverty is the average number of deprivation people experience at the same time. The intensity of poverty, **A**, reflects the proportion of the weighted component indicators in which, an average poor person is deprived of.

$$\mathbf{A} = \frac{\sum_{1}^{q} \mathbf{c}}{\mathbf{qd}} \qquad (3.17)$$

Where **c** is the total number of weighted deprivations the poor experience and **d** is the total number of the component indicators considered (10 in this case).

3.5.4.3 Multidimensional Poverty Index (MPI)

The MPI value summaries the information on multiple deprivations into a single number. It is calculated by multiplying the incidence of poverty (headcount ratio) by the average intensity of poverty.

$$\mathbf{MPI = H \times A} \tag{3.18}$$

Where,

H= Percentage of poor people

A= Average Intensity of deprivation in percentage

The MPI reflects the number of deprivation of poor households that they face at the same time. A household is multidimensionally poor if it is deprived of a few indicators at the same time. Concretely, MPI represents two values of the MPI. The variable ‘c’ reflects the sum of weighted indicators in which a household must be deprived of to be considered multidimensionally poor (Alkire, 2010). A person has to be deprived of at least the equivalent of 33.3 percent of the weighted indicators ($c \geq 33.3$) to be considered multidimensionally poor (Table3.4). Thus, a person is multidimensionally poor if the weighted indicators in which he or she is deprived of sum up to 33.3 percent or more.

*****