Energy Economics Letters
ISSN(e): 2308-2925
DOI: 10.55493/5049.v12i2.5665
Vol. 12, No. 2, 126-141.

URL: www.aessweb.com

Measuring and analysing multidimensional energy poverty with unequal weights: A logistic PCA and AI approach

Krishnendu Das

Department of Economics, Lalbaba College, University of Calcutta, India.
Email: krishnendudas20@gmail.com

Article History
Received: 5 September 2025
Revised: 10 October 2025
Accepted: 20 October 2025
Published: 7 November 2025

Keywords
Artificial neural networks
Logistic principal component analysis
Multidimensional energy poverty
Multilayer perceptron
West Bengal.

JEL Classification:
K32; O13; I32; C45; C38; R20; Q48.

ABSTRACT

This study develops a robust methodological framework for measuring and analyzing multidimensional energy poverty using unit-level household survey data. The approach integrates logistic principal component analysis (Logistic PCA) to construct a composite index that assigns unequal weights to diverse energy deprivation indicators, thereby capturing the heterogeneity and complexity of energy poverty more accurately than equal-weight methods. The index is further disaggregated into moderate and severe categories, enabling a nuanced assessment of deprivation intensity. To complement the measurement stage, artificial intelligence techniques specifically multilayer perceptron (MLP) and artificial neural networks (ANN) are employed to model the socio-demographic and economic determinants of energy poverty. This dual-stage design allows for both explanatory and predictive insights: the statistical modeling validates the significance of key predictors such as household wealth, family size, and access to basic amenities, while the AI models enhance predictive accuracy for identifying high-risk households and regions. By combining unequal-weight composite measurement with AI-driven predictive modeling, the framework offers a scalable and transferable tool for researchers and policymakers. It facilitates targeted, data-driven interventions aimed at reducing energy poverty and promoting equitable energy access. The methodological innovations presented here are adaptable to diverse contexts, making them valuable for comparative studies and policy applications beyond the specific dataset used.

Contribution/ Originality: This study contributes to multidimensional energy poverty research by applying Logistic Principal Component Analysis (LOGPCA) to binary variables for unequal indicator weighting. It further employs a Multi-Layer Perceptron (MLP) artificial intelligence model, making it one of the few studies to do so and the first using NFHS-5 exclusively.

1. INTRODUCTION

Energy poverty refers to the lack of access to affordable, reliable, and clean energy, and it significantly impacts various dimensions of human well-being, including health, education, and economic productivity. It manifests through inadequate electricity access, absence of modern cooking facilities, and limited heating and cooling solutions. Conceptually, energy poverty can be understood through either a unidimensional or multidimensional lens. While unidimensional measures focus on singular aspects such as electricity availability or energy expenditure multidimensional approaches incorporate a broader range of factors, including the availability, reliability, affordability, and efficiency of energy services (Alkire & Foster, 2011; Nussbaumer, 2011). Although unidimensional measures offer simplicity, they often fail to capture the complex and interconnected nature of energy deprivation (Sen, 1999).

In India, energy poverty remains a persistent challenge, especially in states like West Bengal (Acharya & Sadath, 2019; Manasi & Mukhopadhyay, 2024; Teja, 2024). Despite national-level progress in electrification, many households continue to face unreliable supply, unaffordable tariffs, and inadequate infrastructure (Rajić et al., 2020). Rural regions, in particular, are disproportionately affected due to infrastructural gaps, leading to continued dependence on traditional biomass fuels such as firewood, kerosene, and cow dung (Bhide & Monroy, 2011; Lan, Khan, Sadiq, Chien, & Baloch, 2022). Moreover, socio-economic barriers including poverty, unemployment, and limited awareness of clean energy technologies exacerbate the issue. Gender disparities further compound energy poverty, as women and girls often bear the burden of fuel collection and household energy management (Abbas et al., 2022).

Addressing multidimensional energy poverty thus requires an integrated strategy that combines energy access with broader social and economic development. Policy interventions such as subsidies for clean energy technologies, expansion of decentralized renewable energy systems, and promotion of energy-efficient appliances can significantly enhance energy security for marginalized populations (Bhide & Monroy, 2011). Moreover, community-level initiatives focusing on capacity building and awareness generation can support the long-term sustainability of these efforts.

To effectively identify and address the drivers of energy poverty, advanced analytical tools are essential. One such method is the use of artificial intelligence, particularly Multilayer Perceptrons (MLPs), a class of artificial neural networks capable of modeling complex, non-linear relationships. MLPs are increasingly used in the social sciences for classification, regression, and predictive analytics. In the context of energy poverty, MLPs can process large datasets to uncover hidden patterns and forecast deprivation levels based on historical energy usage and socio-economic variables (Abbas et al., 2022). This makes them powerful tools for supporting evidence-based policymaking.

Given this context, the present study aims to assess the extent of multidimensional energy poverty across the districts of West Bengal using unit-level data from NFHS-5 (2019–21). We employ logistic principal component analysis to construct a weighted multidimensional energy poverty index, which is then categorized into moderate and severe levels to better capture the intensity of deprivation. Furthermore, we utilize artificial intelligence techniques, particularly MLP models, to identify the key socio-demographic and economic factors influencing energy poverty in the region. By combining statistical rigor with machine learning, this study seeks to provide deeper insights into the structure of energy poverty and offer actionable recommendations for targeted interventions.

2. LITERATURE REVIEW

P. F. F. Nussbaumer (2011) emphasized the importance of developing and applying metrics to measure energy poverty in order to inform policy-making is emphasized. They introduced the Multidimensional Energy Poverty Index (MEPI-equal weight) as a tool to evaluate energy poverty at different levels. The MEPI focuses on the incidence and intensity of energy poverty, utilizing detailed household survey data for analysis. It is noted that the MEPI is just one instrument among others in monitoring progress and designing effective policies to address energy poverty (Nussbaumer, 2011).

Pelz, Pachauri, and Groh (2018) the measurement of energy poverty has evolved from a binary perspective to a multidimensional approach, with the development of frameworks like the Multitier Framework (MTF). However, there are challenges in operationalizing the MTF at the global tracking level and gaining acceptance in national energy planning. The accurate measurement of access to basic household energy services at the global level requires further work, drawing lessons from debates on defining and estimating basic needs in other domains. There is a need for clearer separation of global and national tracking objectives, capturing dynamic movements in and out of energy poverty, and addressing intrahousehold distributions and vulnerabilities, particularly for women and children. Collecting sex- and age-disaggregated data is important for better understanding and addressing these vulnerabilities. Overall, the value of any metric lies in its ability to inform policies that deliver the greatest welfare benefits, and multidimensional frameworks like the MTF are making progress in this regard (Pelz et al., 2018).

Halkos and Aslanidis (2023) analyzed energy poverty (EP) is a complex issue that requires a multidimensional approach for effective monitoring and eradication. Here are key aspects they considered across different bases like theoretical basis, environmental basis, institutional basis, technical basis, socioeconomic basis, and multidimensional basis. They stressed that consensual and expenditure-based approaches are crucial for framing energy poverty. Defining what energy poverty truly means is essential for fostering public participation, strengthening democratic decision-making, and promoting public participation. Institutional cooperation is necessary to integrate Green Deal goals and combat corruption. Addressing social exclusion and inflation due to war is crucial to minimize the impacts on energy poverty. Considering energy-related apeirophobia in discussions is important. Energy poverty creates vicious circles impacting affordability and accessibility. Monitoring these dimensions is critical in both developed and developing countries. Collaboration and Green Deals can offer a cohesive policy framework to address energy poverty effectively (Halkos & Aslanidis, 2023).

Sadath and Acharya (2017) assessed energy poverty in India using household-level primary data from the India Human Development Survey-II (IHDS-II), 2011-12. From this, they revealed extensive energy poverty in rural areas where traditional biofuels are heavily relied upon. Their key points from the assessment are: Dalits, Adivasis, and marginalized sections are most affected; women spend significant time and energy on collecting and using solid biofuels, impacting labor market participation and causing health issues due to indoor pollution among women and children; energy poverty is multidimensional and should be evaluated using a theoretical framework like Amartya Sen's capability approach with a Multidimensional Energy Poverty Index (MEPI) (Sadath & Acharya, 2017).

Acharya and Sadath (2019) discussed Access to modern energy resources plays a crucial role in promoting the welfare of society globally. Electrified households enable children to study comfortably at night, contributing to better educational outcomes. Electrification enhances security, especially for women and girls, while access to LPG improves health outcomes by reducing health risks associated with traditional biofuels. Transitioning to clean energy sources like LPG not only ensures healthier cooking practices but also frees up time for women and girls to engage in other productive activities. Energy security is vital for sustaining life, particularly in cold regions, and is essential for industrial economies to promote manufacturing and generate employment opportunities. The study examined the extent of energy poverty in India and its relationship with economic development between 2004–05 and 2011–12. Despite progress, a significant portion of the population, especially in poor states, still lacks access to modern energy services, suggesting the need for further interventions (Acharya & Sadath, 2019).

Bhide and Monroy (2011) pointed out that access to clean energy is crucial for sustainable development, particularly in India, where a significant percentage of the global poor reside. Given that most of India's poor live in rural areas, enhancing access to energy sources in these regions is paramount for development and poverty reduction. Despite various government programs and initiatives, there are challenges in meeting development goals and reducing energy poverty. Unrealistic targets highlight the need for increased efforts, especially in promoting renewable energy technologies (Bhide & Monroy, 2011).

Lan et al. (2022) examined the energy poverty in five Asian countries and revealed insights into energy access and its implications. Sri Lanka and Bangladesh face higher levels of energy poverty compared to India and Pakistan. Pakistan and India rely significantly on imported oil, while Bangladesh has reduced its dependence on imported oil over the past decade. A significant portion of India's population lacks access to modern cooking methods, contributing to energy poverty. Factors like access to electricity, clean fuels for cooking, and private investment in energy show significant relationships with the energy poverty index in all countries (Lan et al., 2022).

Nussbaumer's (2011) analysis contributes significantly to the field in two main ways: the MEPI methodology and the global application of MEPI. The study conducted a cross-country analysis of energy poverty at an aggregated level, emphasizing the diversity of energy poverty landscapes between countries. By analyzing the evolution of energy poverty over time using the latest data, the study identifies positive trends in reducing energy poverty intensity in selected sub-Saharan countries. The MEPI not only assesses the incidence of energy poverty but also provides insights into the intensity of energy poverty, showcasing improvements in all countries analyzed. Exploring the connection between MEPI results and policy analysis can help identify effective policies and strategies to address energy poverty (Nussbaumer, 2011).

Abbas et al. (2022) measured and analyzed extreme multidimensional energy poverty in developing countries using the Multidimensional Energy Poverty Index (MEPI) and identified key socioeconomic determinants through supervised machine learning. The MEPI highlighted widespread severe energy poverty across Asian and African countries, with specific nations like Afghanistan, Yemen, Nepal, India, and others being particularly susceptible. Machine learning algorithms identified significant socioeconomic factors influencing multidimensional energy poverty, including household wealth, housing characteristics, marital status of breadwinners, and their residential location. The study emphasized that a combination of various socioeconomic variables, such as income, education, and employment nature, collectively contributes to multidimensional energy poverty rather than a single factor. In conclusion, the study underscores the importance of addressing severe energy poverty by targeting basic energy services, enhancing household socioeconomic status, and promoting universal access to modern energy sources. Implementing these policy directions can lead to significant reductions in energy poverty and improve overall well-being for communities affected by energy deprivation (Abbas et al., 2022).

3. DATA AND METHODOLOGY

In this present study, we used unit-level household data from the National Family Health Survey 5 (NFHS-5) for the period 2019-2021 for the districts of West Bengal. We also used reports such as Energy Statistics 2020 by the Government of India (India, 2020) and Key World Energy Statistics 2020 (International Energy Agency, 2020). West Bengal's Energy Transition (Teja, 2024) for the study. Statistical packages used for data preparation and estimation include R-Programme 4.3.3, SPSS 26, and Microsoft Excel 365.

For the estimation of Multidimensional Energy Poverty (MEP), we followed the methodology proposed by the Oxford Poverty & Human Development Initiative (OPHI) (Nussbaumer, 2011). We considered five dimensions of energy poverty, which are Cooking (D-1), Lighting (D-2), Telecommunication (D-3), Education/Entertainment (D-4), and Household Appliances (D-5). Then, we used multiple indicators to acknowledge these dimensions. The indicators are Type of Fuel used for cooking (I), Separate Kitchen (II), Electricity Access (III), Mobile/Landline (IV), Radio (V), Television (VI), Computer (VII), Fan (VIII), Bike (IX), Fridge (X), and Water Pump (XI) (Manasi & Mukhopadhyay, 2024) (Table 1).

Table 1. Dimensions and indicators used for estimating MEP with unequal weights by logistic PCA.

Dimensions	Indicators	Cause of deprivation	Unequal weights by LOGPCA
Cooking (D-1)	Type of fuel used for cooking (I)	If unclean fuel is used	0.0979
	Separate kitchen (II)	If no such kitchen is available	0.1051
Lightning (D-2)	Electricity access (III)	No access to electricity	0.0597
Telecommunication (D-3)	Mobile/ landline (IV)	Not having mobile or landline	0.0597
Education/ Entertainment (D-4)	Radio (V)	Not having radio	0.1104
	Television (VI)	Not having television	0.0849
	Computer (VII)	Not having computer	0.0950
Household appliances (D-5)	Fan (VIII)	Not having electric fan	0.0849
	Bike (IX)	Not using motorbike	0.1025
	Fridge (X)	Not having refrigerator	0.0998
	Water Pump (XI)	Not having water pump	0.1004

Most of the studies used equal weights for the estimation of MEP, which is quite logical but not always effective due to its heavy dependence on standard theory, which may or may not be perfectly applicable to different regions. Additionally, some studies have used an unequal weighting system to achieve region-specific accuracy, primarily utilizing Principal Component Analysis (PCA) (Manasi & Mukhopadhyay, 2024). PCA is generally suitable for continuous variables, not categorical variables. As in our study, most variables were categorical (especially with binary outcomes); hence, PCA is not very effective for obtaining weights of the variables. For this study, we used Logistic Principal Component Analysis (LOGPCA).

Logistics Principal Component Analysis (Das, 2021) is a multivariate generalization of the so-called Bernoulli distribution. The Bernoulli distribution for a univariate binary random variable:

$$P(x|p) = p^x(1-p)^{1-x} \quad (1)$$

Where $x \in \{0,1\}$ with p as the mean. We can write this distribution in terms of the log-odds, where the parameter.

$$\theta = log\left(\frac{p}{1-p}\right) \quad (2)$$

So, the logistic function.

$$\sigma(\theta) = \left[1 + e^{-\theta}\right] \quad (3)$$

Now the Bernoulli distribution becomes.

$$P(x|\theta) = \sigma(\theta)^x\sigma(-\theta)^{1-x} \quad (4)$$

The above expression is considered as the Bernoulli distribution expressed as a member of the exponential family. A generalization of the above equation gives us the Logistic Principal Component Analysis model (Landgraf, 2016; Schein, Saul, & Ungar, 2003).

Assuming there is a d-dimensional binary data set with n observations, the matrix notation would be: $(x_{ij})_{n\times d}$. If x_{ij} is an element with Bernoulli probability p_{ij}, then the parameter θ_{ij} is the logit of the probability.

$$\theta_{ij} = log\left(\frac{p_{ij}}{1-p_{ij}}\right) \quad (5)$$

There are three major types of logistic principal component analysis, which are Exponential Family PCA, Logistic PCA, and Convex Logistic PCA.

Collins suggested exponential family PCA mainly for binary variables. The assumption made by Collins was that the logit of the probability matrix is presumed as a matrix factorization.

$$logit(P) = k_n\mu^T + AB^T \quad (6)$$

Here A & B are the lower rank, k & μ are vectors of dimension d of main effects (Collins, Dasgupta, & Schapire, 2001; Landgraf, 2016).

Landgraf extends Pearson's Principal Component Analysis. Pearson's idea was to find a rank-k projection that minimizes the mean squared error of the data, which may be very close to the original data. In notation, it minimizes.

$$\frac{1}{n}\sum_1^n \|(x_i - \mu) - UU^T(x_i - \mu)\|^2 \quad (7)$$

Over μ and $n\times d$ orthonormal matrix U., The Logistic PCA extends PCA for binary data, using the projection of the natural parameters from the Bernoulli saturated model, which minimizes the Bernoulli deviance. According to Landgraf, letting the d-dimensional vector of natural parameters from the Bernoulli saturated model that is $\underline{\theta_i}$ estimated by

$$\hat{\theta}_i = \mu - UU^T\left(\underline{\theta_i} - \mu\right) \quad (8)$$

Where μ and U are solved for minimizing the Bernoulli deviance.

$$D\left(\hat{\theta}\right) = \sum_{i=1}^n \sum_{j=1}^d -2x_{ij}\hat{\theta}_{ij} + 2log\left(1 + exp\left(\hat{\theta}_{ij}\right)\right) \quad (9)$$

The logistic PCA model has three important benefits over the exponential family PCA: these are:

(i) The number of parameters does not increase with the number of observations.

(ii) The principal component scores can be interpreted easily as linear functions of the data.

(iii) Lastly, only the matrix multiplication is needed when using new data for PCA (Landgraf, 2016; Landgraf & Lee, 2020).

The Convex Logistic PCA is very similar to Logistic PCA. Here, in the Convex Logistic PCA, minimization is performed over the convex hull[1] of rank k projection matrices rather than minimization over rank k projection matrices. Convex Logistic PCA can be solved more quickly and reliably (Landgraf & Lee, 2020). It is primarily developed for categorical variables (Das, 2021). By using LOGPCA, we have estimated the unequal weights for different indicators under different dimensions (Table 1).

After obtaining the unequal weights, cut-offs are to be fixed. Using the unequal weights, we obtained the Multidimensional Energy Poverty Score (MEP Score) for each individual. Then, we proceed as follows:

$$Wt_{Mean(MEP\ Score)} = \frac{\sum_{i=1}^{n}\{MEP\ Score_i \times (HS_i)\}}{\sum_{i=1}^{n} HS_i} \quad (10)$$

This is the weighted mean of MEP Score and the weights we used is the family size of the households (HS_i).

$$Wt_{Sd(MEP\ Score)} = \sqrt{\frac{\sum_{i=1}^{n}\left\{(HS_i)\times\left(MEP\ Score_i - Wt_{Mean(MEP\ Score)}\right)^2\right\}}{\sum_{i=1}^{n} HS_i}} \quad (11)$$

Now we have the weighted standard deviation of MEP Score, here also used the family size of the households as weights. We used $Wt_{Mean(MEP\ Score)}$ & $Wt_{Sd(MEP\ Score)}$ to determine Moderate Multidimensional Energy Poverty & Severe Multidimensional Energy Poverty cut-offs as follows:

$$MODMEP_{\propto} = Wt_{Mean(MEP\ Score)} - 0.5 \times Wt_{Sd(MEP\ Score)}$$

$$SEVMEP_{\propto} = Wt_{Mean(MEP\ Score)} - 0.5 \times Wt_{Sd(MEP\ Score)}$$

Here $MODMEP_{\propto} = 0.4529$ (Moderate Multidimensional Energy Poverty $\propto$ cut-off) & $SEVMEP_{\propto} = 0.6300$ (Severe Multidimensional Energy Poverty $\propto$ cut-off) are the estimated MEP cut-offs and the value of $\propto$ is 0.5. Now for identification, we used the cut-offs as follows: if the Multidimensional Energy Poverty Score (MEP Score) of a household is greater than or equal to 0.4529, then we consider that household to be experiencing Moderate Multidimensional Energy Poverty. Conversely, if the MEP Score is greater than or equal to 0.6300, then the household is considered to be experiencing Severe Multidimensional Energy Poverty.

After cut-offs, we estimated the Moderate and Severe MEP Headcount Ratios (MEPHCR) and Average Intensity (MEPAVGINT) for the districts of West Bengal. We followed the methodology proposed by OPHI for Multidimensional Poverty, as outlined by Alkire & Foster (2011).

$$MODMEPI = MODMEPHCR \times MODMEPAVGINT$$

$$SEVMEPI = SEVMEPHCR \times SEVMEPAVGINT$$

Where MODMEPI and SEVMEPI are the Moderate and Severe Multidimensional Energy Poverty Indices; MODMEPHCR and SEVMEPHCR are the Moderate and Severe Multidimensional Energy Poverty Headcount Ratios; MODMEPAVGINT and SEVMEPAVGINT are the Moderate and Severe Multidimensional Energy Poverty Average Intensities.

Now for the identification of socio-demographic determinants of MEP, we used different socio-demographic and economic variables: House (V1), Wealth (V2), Education (V3), Family Size (V4), Marriage (V5), Status (V6), Residence (V7), Sex (V8), Age (V9), Caste (V10), and Smoking (V11) (Abbas et al., 2022) (Table 3). To estimate the importance of these socio-demographic and economic variables in energy poverty we used Multilayer Perceptron

[1] Convex hull is the intersection of all convex sets containing a given subset of a Euclidean space, or equivalently as the set of all convex combinations of points in the subset. In PCA for clustering of points a convex hull is the smallest polygon that includes all the points of a given level.

Model (MLP) under artificial neural network (Abbas et al., 2022; Bagheri, Taridashti, Farahani, Watson, & Rezvani, 2023). A neural network is a computer program that can identify patterns in data sets and perform tasks faster than average programs. It learns to predict data outcomes by analyzing thousands of examples. Multi-layer perceptrons (MLPs) are composed of interconnected perceptrons, each with weights that affect the network's outcome. MLPs are also known as feedforward artificial neural networks (Rajić et al., 2020). For our MLP, we have a total of eleven variables, of which eight are factors (categorical) and the remaining three are covariates (continuous). We used two hidden layers, with the activation functions being hyperbolic tangent (for inputs) and sigmoid (for output). We considered the MEP score as our output variable (Table 6). The architecture of the Multi-Layer Perceptron (MLP) model used to estimate the Multidimensional Energy Poverty (MEP) score consists of 31 input variables, including categorical indicators such as wealth, education, marital status, residence, caste, smoking habits, household characteristics, and demographic attributes. These inputs are processed through two hidden layers one with 11 neurons and another with 9 neurons both activated using the hyperbolic tangent function. The final output layer generates the MEP score, with a sigmoid activation function ensuring values remain bounded between 0 and 1. This design enables the model to capture complex, nonlinear relationships among socioeconomic and demographic variables in predicting energy poverty (Figure 4).

4. RESULTS AND DISCUSSION

We subdivided this section into two parts: the first part presents the energy poverty estimates of the districts of West Bengal, and the second part provides a predictive analysis of the socio-demographic and economic factors influencing energy poverty in West Bengal using a multilayer perceptrons model under an artificial neural network.

As mentioned in the data & methodology section, we used two types of multidimensional energy poverty (MEP) estimates: moderate and severe. For both cases, we estimated the headcount ratio and average intensity. For moderate MEP, Puruliya (0.8838) and Kolkata (0.3948) were the two districts with the lowest and highest headcount ratios among the districts of West Bengal, respectively. Meanwhile, for the entire state of West Bengal, approximately 74% of households were multidimensionally energy poor at the moderate level. Similarly, for severe MEP, Puruliya (0.6382) and Kolkata (0.0671) were the two districts with the lowest and highest estimates of headcount ratios among all districts of West Bengal. About 34% of households experienced severe multidimensional energy poverty.

Let us focus on the average intensities of MEPs. We found a similar picture to headcount ratios. Here, also, Puruliya and Kolkata were the two districts with extreme levels of unequal energy poverty for both moderate and severe categories. An average energy-poor household in the state experiences approximately 63% and 72% intensities for moderate and severe energy poverty, respectively.

We estimated the MEPIs of moderate and severe levels. The same results were obtained as with the headcount ratios and average intensities (Table 2).

Table 2. District-wise different multidimensional energy poverty estimates of West Bengal during 2019-2021.

Districts	Moderate MEPHCR	Moderate MEPAVGINT	Severe MEPHCR	Severe MEPAVGINT	Moderate MEPI	Severe MEPI
Bankura	0.8369	0.6695	0.4994	0.7465	0.5603	0.3728
Bardhaman	0.6732	0.634	0.301	0.733	0.4268	0.2207
Birbhum	0.8091	0.6323	0.4056	0.708	0.5116	0.2872
Daskshin Dinajpur	0.8076	0.624	0.3533	0.7117	0.5039	0.2514
Darjeeling	0.7038	0.6099	0.2461	0.7267	0.4292	0.1788
Howrah	0.6037	0.5999	0.1779	0.715	0.3622	0.1272
Hugli	0.6626	0.6195	0.2623	0.7224	0.4104	0.1895
Jalpaiguri	0.704	0.6177	0.272	0.7238	0.4349	0.1969
Koch Bihar	0.8391	0.6149	0.3543	0.7002	0.516	0.2481

Districts	Moderate MEPHCR	Moderate MEPAVGINT	Severe MEPHCR	Severe MEPAVGINT	Moderate MEPI	Severe MEPI
Kolkata	0.3948	0.5731	0.0671	0.7064	0.2262	0.0474
Maldah	0.8112	0.6281	0.3842	0.7071	0.5095	0.2716
Murshidabad	0.8364	0.6368	0.4329	0.7136	0.5326	0.309
Nadia	0.7563	0.6201	0.3362	0.6994	0.4689	0.2351
North 24 PGNS	0.5552	0.5941	0.2089	0.692	0.3298	0.1445
Paschim Medinipur	0.878	0.6626	0.5198	0.735	0.5818	0.382
Purba Medinipur	0.8786	0.6391	0.4708	0.7044	0.5615	0.3316
Puruliya	0.8838	0.7278	0.6382	0.7928	0.6432	0.506
South 24 PGNS	0.7492	0.6121	0.3064	0.7005	0.4585	0.2146
Uttar Dinajpur	0.8134	0.6267	0.3905	0.7059	0.5098	0.2756
West Bengal	0.7438	0.6324	0.3465	0.7232	0.4704	0.2506

Source: Authors own calculation from NFHS-5 (2019-21) unit level household data.

We observed that districts (Howrah, Hugli, North 24 PGNS, South 24 PGNS) sharing borders with Kolkata had low energy poverty. As we move further from Kolkata, energy poverty increases. We also found that the extreme northern districts of West Bengal, such as Darjeeling, Jalpaiguri, and Cooch Behar, have less energy poverty (Figure-1, 2, 3).

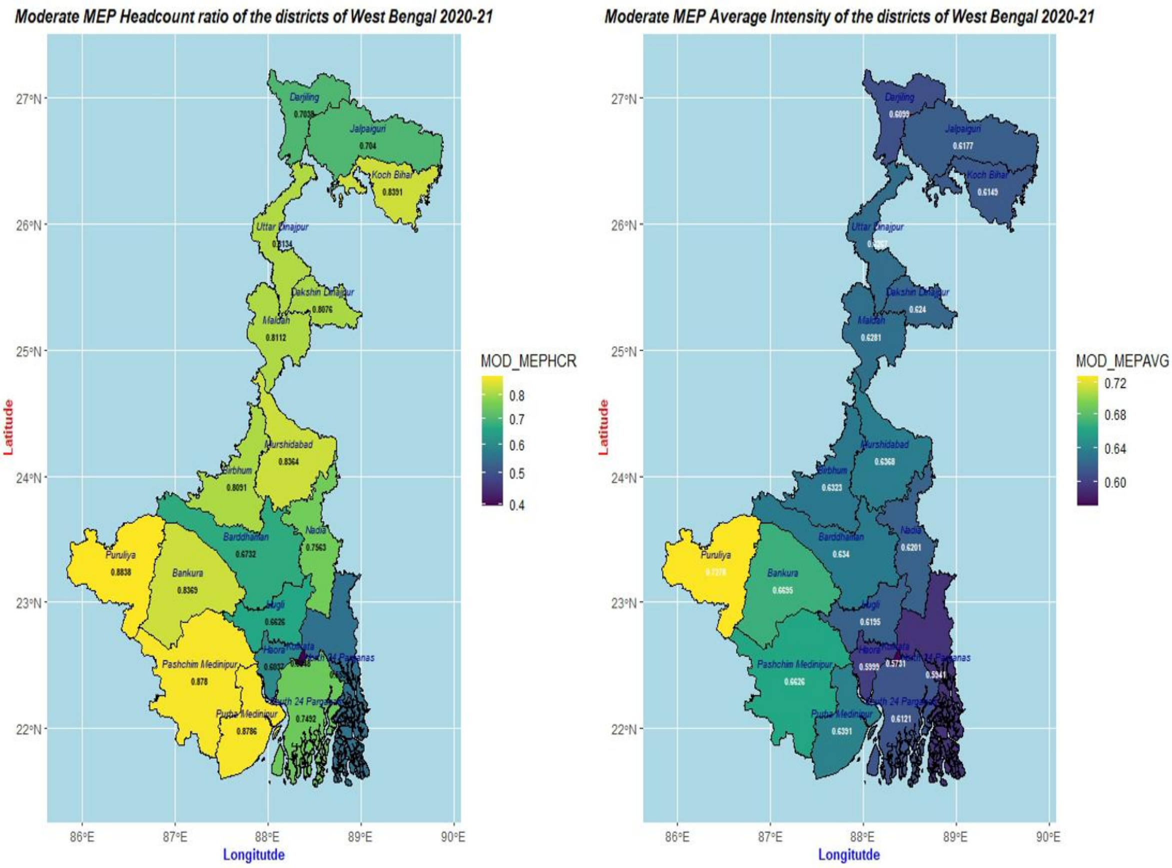

Figure 1. (a) Moderate MEPHCR & (b) Moderate MEPAVGINT of the Districts of West Bengal for the period 2019-21.

Source: Authors own calculation from NFHS-5 (2019-21) unit-level household data.

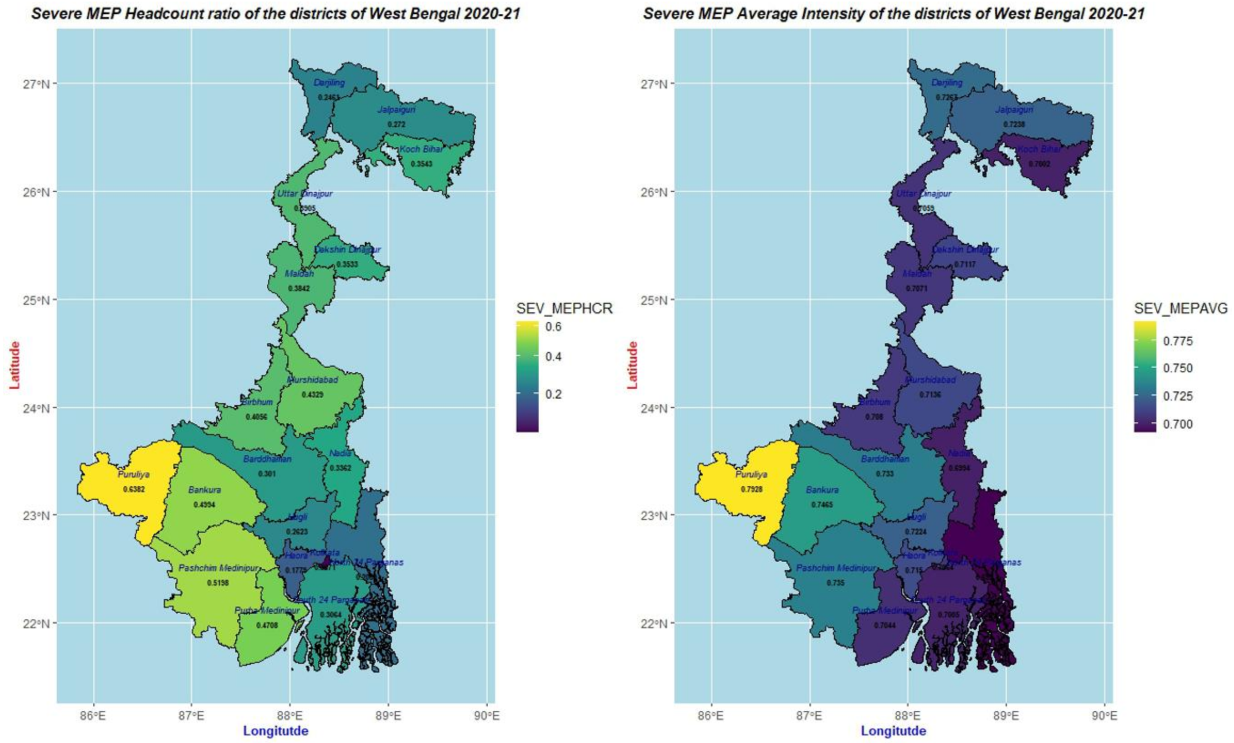

Figure 2. (a) Severe MEPHCR & (b) Severe MEPAVGINT of the Districts of West Bengal for the period 2019-21.

Source: Authors own calculation from NFHS-5 (2019-21) unit-level household data.

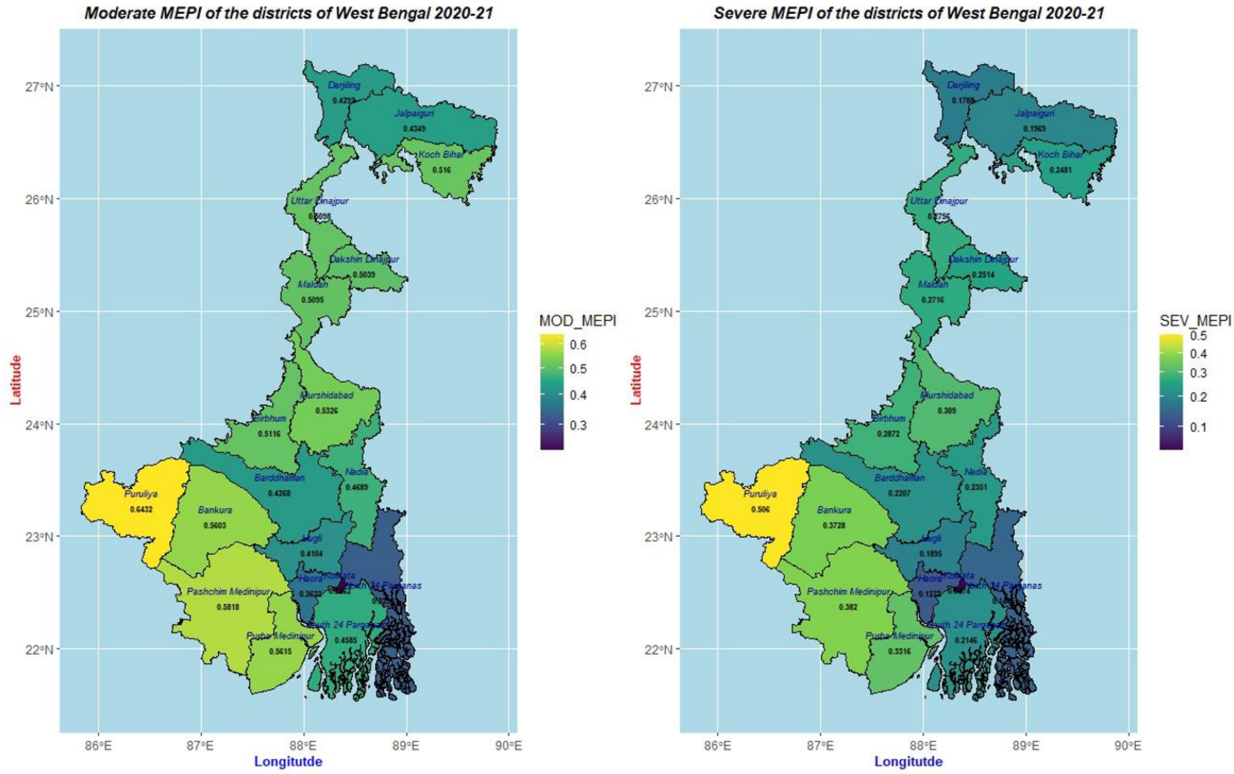

Figure 3. (a) Moderate MEPI & (b) Severe MEPI of the Districts of West Bengal for the period 2019-21.

Source: Authors own calculation from NFHS-5 (2019-21) unit-level household data.

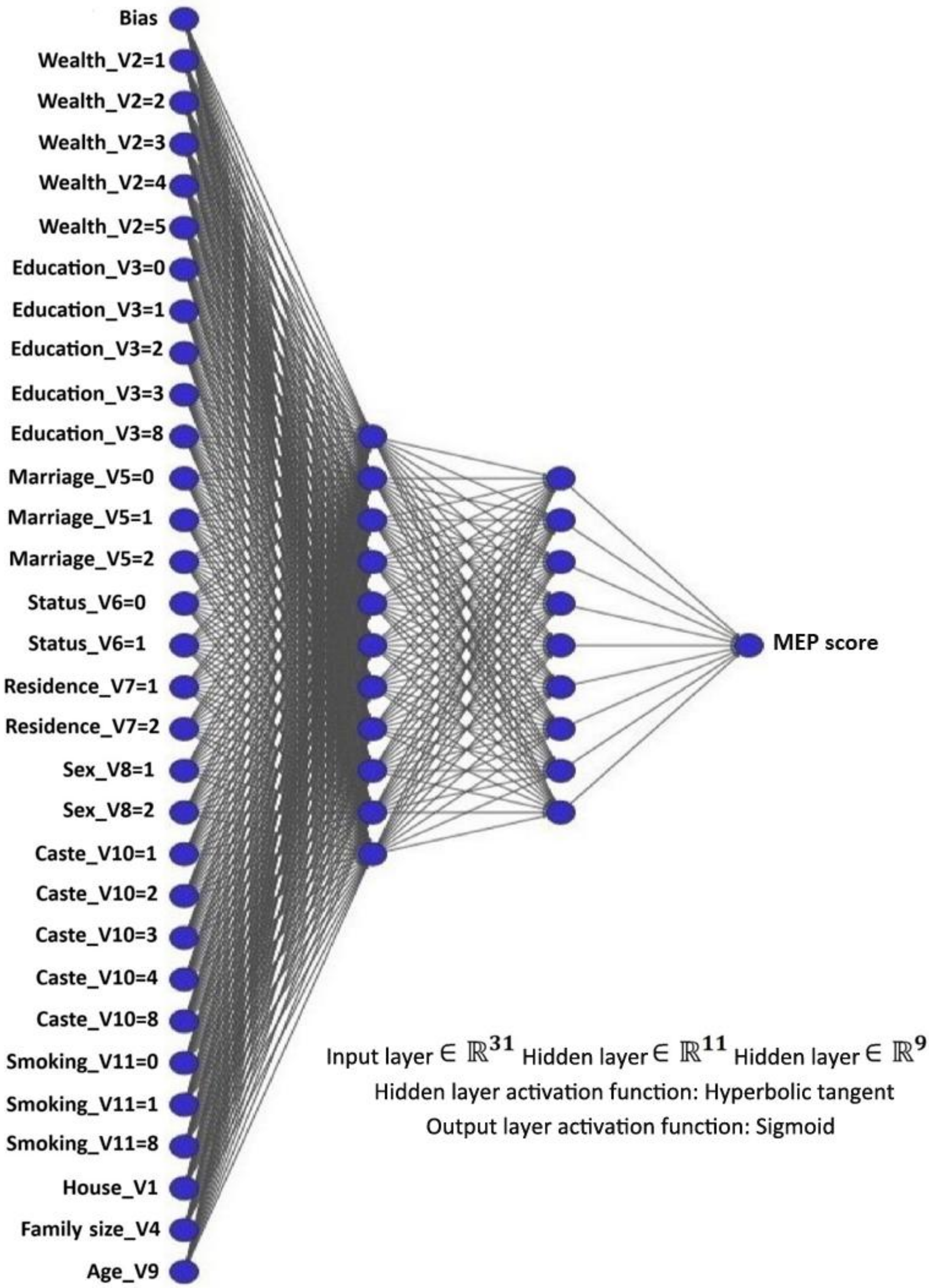

Figure 4. Architecture of artificial neural network for MEP Score by multilayer perceptron model.

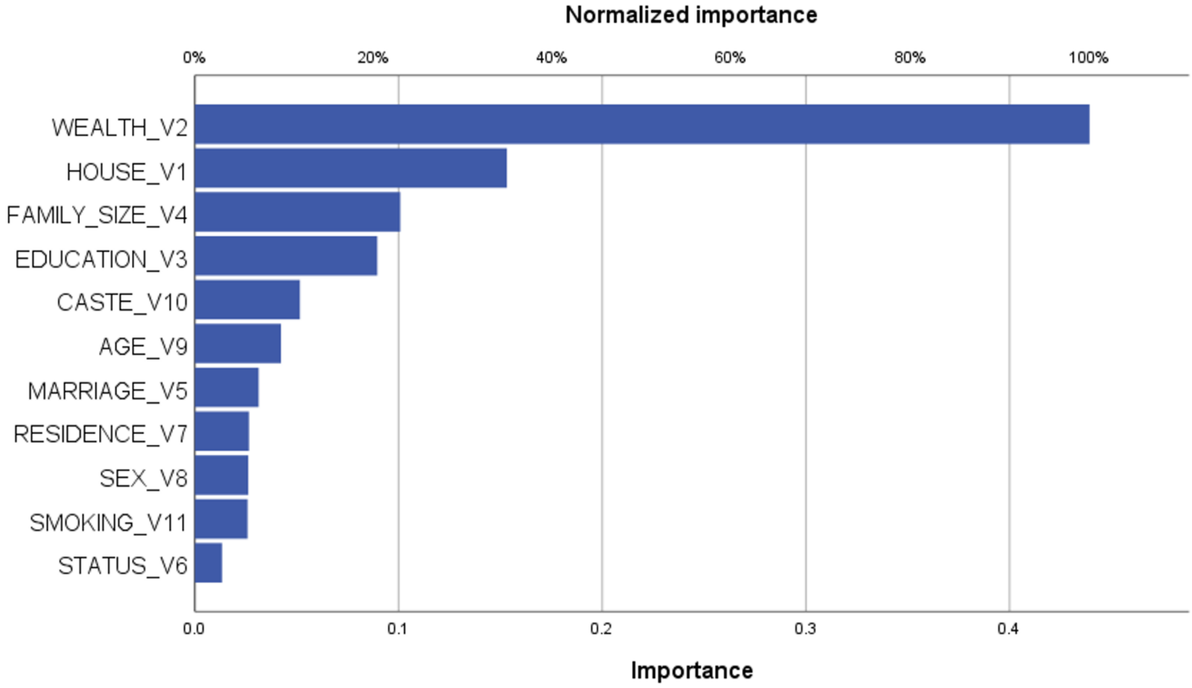

Figure 5. Normalized importance of independent variables by MLP.

Source: Authors own calculation from NFHS-5 (2019-21) unit-level household data.

In continuation of our discussion, we move to the second sub-section of this section, which is the analysis of the predictive model for the determination of socio-demographic and economic factors responsible for energy poverty in West Bengal for the period 2019-20. We used eleven socio-demographic and economic variables as input and the multidimensional energy poverty scores as the output for the multilayer perceptrons model of an artificial neural network (Table 3). Input variables were mostly skewed in their raw nature (Table 4). So, we needed to standardize them to achieve better and more efficient results. The sample size was 18,187 for the entire state of West Bengal. Out of this, we used 12,754 (70.10%) for training and 5,433 (29.90%) for testing (Table 5). We used normalization as a rescaling method for scale-dependent variables (Table 6). The relative error for training was 0.2942, and for testing was 0.3040. These were quite low, implying a high level of significance of the model (Table 7). The plotted graph between the predictive value and the MEP_SCORE is quite significant as the fitted regression line having R2=0.703, means near about more than 70% good fit (Figure 6). Also, the graph of residuals by predicted values shows an insignificant fit, as the R2 value is very low, indicating almost no relation between residuals and the predicted values (Figure 7).

Table 3. Variables for the predictive model by MLP.

Category	Variables	Definition
Input	House (V1)	No. of rooms used for Sleeping
	Wealth (V2)	Wealth Index (grouped)
	Education (V3)	Household head's educational attainment
	Family Size (V4)	No. of Household members
	Marriage (V5)	Is the household head currently married or not?
	Status (V6)	Ownership status of the house
	Residence (V7)	Place of residence Rural/Urban
	Sex (V8)	Household head's Sex
	Age (V9)	Household head's age
	Caste (V10)	The household head belongs to a social group.
	Smoking (V11)	Household head's smoking habits
Output	MEP Scores	Multidimensional Energy Poverty (MEP) scores for each household range between 0 and 1.

Table 4. Statistical summary of the input variables.

	Minimum	Maximum	Mean	Std. Deviation	Skewness
HOUSE_V1	0.0000	25.0000	1.9229	1.0944	5.1247
WEALTH_V2	1.0000	5.0000	2.3069	1.2758	0.6337
EDUCATION_V3	0.0000	8.0000	1.2865	0.9828	0.3346
FAMILY_SIZE_V4	1.0000	19.0000	4.0864	1.8007	1.1802
MARRIAGE_V5	0.0000	2.0000	1.0964	0.3748	1.0428
STATUS_V6	0.0000	1.0000	0.4215	0.4938	0.3179
RESIDENCE_V7	1.0000	2.0000	1.7008	0.4579	-0.8770
SEX_V8	1.0000	2.0000	1.1597	0.3664	1.8578
AGE_V9	15.0000	98.0000	48.4692	13.8949	0.2332
CASTE_V10	1.0000	8.0000	3.9201	2.6542	0.5311
SMOKING_V11	0.0000	8.0000	0.5546	0.5528	2.3819

Table 5. Results of case processing summary.

Case Processing Summary			
		N	Percent
Sample	Training	12754	70.10%
	Testing	5433	29.90%
Valid		18187	100.00%
Excluded		0	
Total		18187	

Table 6. Results of network information.

Network information			
Input layer	Factors	1	WEALTH_V2
		2	EDUCATION_V3
		3	MARRIAGE_V5
		4	STATUS_V6
		5	RESIDENCE_V7
		6	SEX_V8
		7	CASTE_V10
		8	SMOKING_V11
	Covariates	1	HOUSE_V1
		2	FAMILY_SIZE_V4
		3	AGE_V9
	Number of units*		30
	Rescaling method for covariates		Standardized
Hidden layer(s)	Number of hidden layers		2
	Number of units in hidden layer 1a		10
	Number of units in hidden layer 2a		8
	Activation function		Hyperbolic tangent
Output layer	Dependent variables	1	MEP_SCORE
	Number of units		1
	Rescaling method for scale dependents		Normalized
	Activation function		Sigmoid
	Error function		Sum of Squares

Note: * Excluding the bias Unit

Table 7. Model summary of an artificial neural network.

Model summary		
Training	Sum of squares error	62.158
	Relative error	0.294
	Stopping rule used	Maximum number of epochs (100) exceeded
	Training time	00:02.3
Testing	Sum of squares error	25.958
	Relative error	0.304

Table 8. Results of the normalized importance of inputs to the outcome.

Independent variable importance		
	Importance	Normalized importance
WEALTH_V2	0.439	100.00%
EDUCATION_V3	0.09	20.40%
MARRIAGE_V5	0.031	7.10%
STATUS_V6	0.013	3.00%
RESIDENCE_V7	0.027	6.10%
SEX_V8	0.026	6.00%
CASTE_V10	0.052	11.70%
SMOKING_V11	0.026	5.90%
HOUSE_V1	0.153	34.90%
FAMILY_SIZE_V4	0.101	23.00%
AGE_V9	0.042	9.60%

Source: Author's own calculation from NFHS-5 (2019-21) unit-level household data.

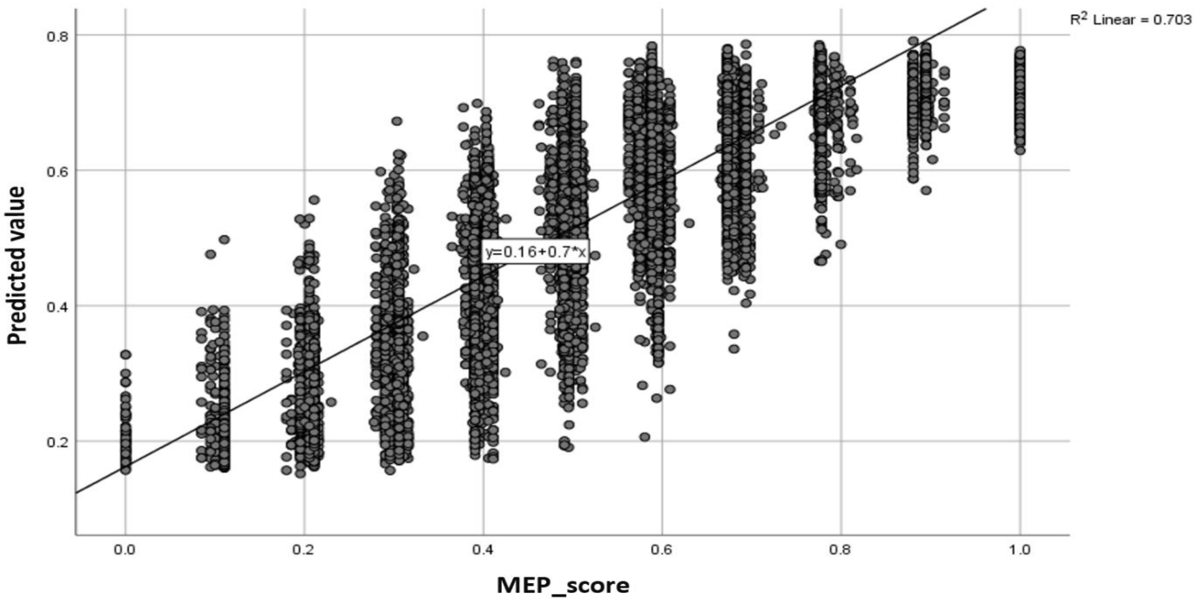

Figure 6. Predicted values based on the observed graph.

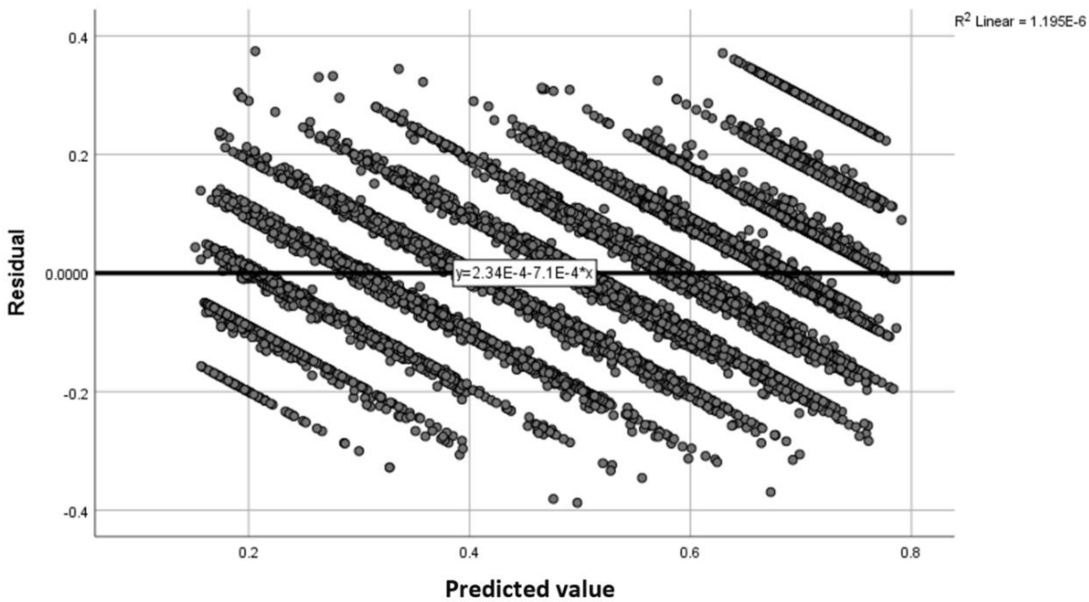

Figure 7. Residual by predicted graph.

We found that wealth (V2) has the highest normalized importance (100.00%), while status (V6) has the lowest (3.00%) among the eleven variables predicting the MEP in West Bengal for the period 2019-21. According to the MLP model, House (V1), Family size (V4), Education (V3), Caste (V10), Age (V9), and Marriage (V5) are significant socio-demographic and economic factors contributing to energy poverty in West Bengal. Conversely, Residence (V7), Sex (V8), and Smoking (V11) have less important contributions to MEP (Figure 5) (Table 8).

Mostly households with a high wealth index are less likely to be energy poor. Wealth plays a significant role in determining MEP, which is very common. However, the number of rooms in the house for sleeping is another important factor for energy poverty. The mean family size variable is 4.0864, indicating that, on average, there are four members in a household (Table 4). So, families with more than 4 (mean value) members are more likely to be energy poor, perhaps. The educational level of the household also seems to be another determining factor for energy poverty. Households with an educated household head are more likely to inspire other members to pursue education. As it is already established, severe income inequality is present among the different social groups (Anand & Thampi, 2016). Hence, caste variable must be a factor for energy poverty. Household head's age is the second least influential determinant among the major factors for the predictive model of energy poverty in West Bengal. The mean value of the head's age is 48.4692, which implies that household heads over 48 may be prone to multidimensional energy poverty. Marriage is a social institution that definitely demands more energy expenditure. Hence, it also determines energy poverty with a normalized importance of 7.10%.

5. CONCLUSION

This study provides a comprehensive and methodologically robust examination of multidimensional energy poverty across the districts of West Bengal, employing both logistic principal component analysis (LogPCA) and artificial intelligence-based predictive modeling through multilayer perceptrons (MLPs). By constructing a weighted multidimensional energy poverty index and categorizing it into moderate and severe levels, the study uncovers not only the extent but also the intensity of energy deprivation experienced by households in the state.

The empirical results highlight stark disparities in energy poverty between districts, with areas like Puruliya and Paschim Medinipur facing the most acute challenges. Conversely, urban centers such as Kolkata show significantly lower levels of deprivation, pointing to a strong spatial divide rooted in infrastructural, economic, and socio-political differences. The analysis of average intensity alongside headcount ratios presents a more nuanced understanding of how deeply energy poverty is felt by affected populations, rather than just how widely it is spread.

The study further employs MLP models to identify the socio-demographic and economic determinants of energy poverty. Key factors such as household wealth, size, housing conditions, educational attainment, and caste emerge as dominant predictors. Among these, wealth holds the highest predictive importance, affirming the well-established link between income levels and access to basic services. Interestingly, family size and housing characteristics (such as the number of rooms available for sleeping) also show strong correlations, suggesting that intra-household resource distribution and spatial adequacy play a significant role in shaping energy deprivation.

The use of artificial neural networks adds a valuable predictive dimension to this research. The relatively high explanatory power ($R^2 \approx 0.703$) of the MLP model demonstrates its potential for policy forecasting and targeted intervention planning. Such models can be employed by government agencies and development practitioners to identify high-risk households and prioritize regions for clean energy initiatives and subsidy programs.

Policy recommendations arising from this analysis include:

- Targeted wealth creation through job-oriented schemes in high-poverty districts.
- Promotion of housing support mechanisms, especially for overcrowded households.
- Educational and awareness campaigns to improve energy literacy and encourage the adoption of efficient appliances.

- Expansion of decentralized renewable energy systems, particularly in geographically remote or infrastructurally weak areas.
- Caste-sensitive and inclusive energy planning that addresses the needs of socially marginalized groups.

In conclusion, addressing energy poverty in West Bengal demands a multidimensional, data-driven, and inclusive approach that integrates technological tools with socio-economic insights. This study not only contributes a novel methodological framework using LogPCA and AI but also provides actionable evidence for tailoring policy interventions. Future research may focus on longitudinal tracking of energy poverty dynamics and incorporate environmental and climate dimensions, thereby enriching the discourse on sustainable and equitable energy access.

Funding: This study received no specific financial support.
Institutional Review Board Statement: Not applicable.
Transparency: The author states that the manuscript is honest, truthful, and transparent, that no key aspects of the investigation have been omitted, and that any differences from the study as planned have been clarified. This study followed all writing ethics.
Data Availability Statement: Upon a reasonable request, the supporting data of this study can be provided by Krishnendu Das.
Competing Interests: The author declares that there are no conflicts of interests regarding the publication of this paper.
Disclosure of AI use: The author used OpenAI's ChatGPT (version 2025) to improve the clarity and grammar of different sections of the study. All content was reviewed and verified by the author.

REFERENCES

Abbas, K., Butt, K. M., Xu, D., Ali, M., Baz, K., Kharl, S. H., & Ahmed, M. (2022). Measurements and determinants of extreme multidimensional energy poverty using machine learning. *Energy, 251*, 123977. https://doi.org/10.1016/j.energy.2022.123977

Acharya, R. H., & Sadath, A. C. (2019). Energy poverty and economic development: Household-level evidence from India. *Energy and Buildings, 183*, 785-791. https://doi.org/10.1016/j.enbuild.2018.11.047

Alkire, S., & Foster, J. (2011). *Understandings and misunderstandings of multidimensional poverty measurement.* Retrieved from Oxford Poverty & Human Development Initiative (OPHI): https://www.researchgate.net/publication/225650144_Understandings_and_misunderstandings_of_multidimensional_poverty_measurement

Anand, I., & Thampi, A. (2016). Recent trends in wealth inequality in India. *Economic and Political Weekly, 51*(50), 59–67.

Bagheri, S., Taridashti, S., Farahani, H., Watson, P., & Rezvani, E. (2023). Multilayer perceptron modeling for social dysfunction prediction based on general health factors in an Iranian women sample. *Frontiers in Psychiatry, 14*, 1283095. https://doi.org/10.3389/fpsyt.2023.1283095

Bhide, A., & Monroy, C. R. (2011). Energy poverty: A special focus on energy poverty in India and renewable energy technologies. *Renewable and Sustainable Energy Reviews, 15*(2), 1057-1066. https://doi.org/10.1016/j.rser.2010.11.044

Collins, M., Dasgupta, S., & Schapire, R. E. (2001). A generalization of principal components analysis to the exponential family. In T. G. Dietterich, S. Becker, & Z. Ghahramani (Eds.), Advances in neural information processing systems. In (Vol. 14, pp. 617–624). Cambridge, MA: MIT Press

Das, K. P. P. (2021). Poverty analysis of the districts of west bengal in 2015-2016: A fuzzy multidimensional approach. *Rabindra Bharati University Journal of Economics, 15*, 47-66.

Halkos, G. E., & Aslanidis, P.-S. C. (2023). Addressing multidimensional energy poverty implications on achieving sustainable development. *Energies, 16*(9), 3805. https://doi.org/10.3390/en16093805

International Energy Agency. (2020). *Key world energy statistics 2020.* Retrieved from International Energy Agency: https://www.iea.org/reports/key-world-energy-statistics-2020

Lan, J., Khan, S. U., Sadiq, M., Chien, F., & Baloch, Z. A. (2022). Evaluating energy poverty and its effects using multi-dimensional based DEA-like mathematical composite indicator approach: Findings from Asia. *Energy Policy, 165*, 112933. https://doi.org/10.1016/j.enpol.2022.112933

Landgraf, A. J. (2016). *An Introduction to the logistic PCA R package.* Raleigh, NC: CRAN / R-Project.

Landgraf, A. J., & Lee, Y. (2020). Dimensionality reduction for binary data through the projection of natural parameters. *Journal of Multivariate Analysis, 180*, 104668. https://doi.org/10.1016/j.jmva.2020.104668

Manasi, B., & Mukhopadhyay, J. P. (2024). Definition, measurement and determinants of energy poverty: Empirical evidence from Indian households. *Energy for Sustainable Development, 79*, 101383. https://doi.org/10.1016/j.esd.2024.101383

Nussbaumer, P., Fuso Nerini, F., Onyeji, I., & Howells, M. (2011). Global insights based on the multidimensional energy poverty index (MEPI). *Sustainability, 5*(5), 2060-2076. https://doi.org/10.3390/su5052060

Nussbaumer, P. M. B. (2011). *Measuring energy poverty: Focusing on what matters.* Retrieved from Oxford Poverty & Human Development Innitiative (OPHI), Working Paper No. 42:

Pelz, S., Pachauri, S., & Groh, S. (2018). A critical review of modern approaches for multidimensional energy poverty measurement. *Wiley Interdisciplinary Reviews: Energy and Environment, 7*(6), e304. https://doi.org/10.1002/wene.304

Rajić, M. N., Milovanović, M. B., Antić, D. S., Maksimović, R. M., Milosavljević, P. M., & Pavlović, D. L. (2020). Analyzing energy poverty using intelligent approach. *Energy & Environment, 31*(8), 1448-1472. https://doi.org/10.1177/0958305x20907087

Sadath, A. C., & Acharya, R. H. (2017). Assessing the extent and intensity of energy poverty using Multidimensional Energy Poverty Index: Empirical evidence from households in India. *Energy Policy, 102*, 540-550. https://doi.org/10.1016/j.enpol.2016.12.056

Schein, A. I., Saul, L. K., & Ungar, L. H. (2003). *A generalized linear model for principal component analysis of binary data.* Paper presented at the International Workshop on Artificial Intelligence and Statistics.

Sen, A. K. (1999). *Development as freedom.* Oxford: Oxford University Press.

Teja, V. (2024). *West Bengal's energy transition: A Rs. 62,000 crore opportunity.* Kolkata, India: Climate Risk Horizons.