Electrical Power and Energy Systems 97 (2018) 363–371

Contents lists available at ScienceDirect

Electrical Power and Energy Systems

journal homepage: www.elsevier.com/locate/ijepes

A new stratified random sample customer selection for load research study in distribution networks

Yaser Raeisi-Gahrooei, Amin Khodabakhshian*, Rahmat-Allah Hooshmand

Department of Electrical Engineering, University of Isfahan, Isfahan, Iran

ARTICLE INFO

Keywords:
Distribution network
Energy consumption
Load research
Sampling

ABSTRACT

Distribution network decision makers need accurate and reliable information about load characteristics to plan, estimate and control the system properly. Load information is generally extracted from the collected load data of selected sample customers and, therefore, a proper sample customer selection is the pillar of any load research study. In this regard, this paper presents a new stratified sampling technique which includes three stages of sample size determination and customer stratification, sample size assignment to determined strata (subgroups), and sample customer selection. The sample size assignment to determined strata is done by considering the compatibility between load research objectives and sampling design and preparing the way to use the data and information provided by previous load research studies. Furthermore, sample customers are selected by considering the energy consumption (kWh) range of customers, their activity classification, and their locations in distribution feeders. The numerical results from a real data of an electric power distribution system in Esfahan-Iran verify the efficiency of the proposed technique when compared to the conventional method.

1. Introduction

Load research is the process of measuring, collecting and studying the customers' electric load characteristics in order to provide the reliable and thorough information for any company related to the production, distribution and management of electricity [1,2]. One of the major applications of load research analysis is the design and setting of retail tariffs for electricity supply [3,4]. This information can improve the accuracy of forecasting the future demand [5]. The other applications are the usage of load profiles of customers in capacitor placement and the reconfiguration of distribution networks [6], state estimation [7], and load modelling [8]. In this regard, the proper operation and control of distribution system requires all the information available on customers' patterns. This in turn helps the operator take proper action for both normal and atypical cases [9].

Since a large number of customers in different electricity tariffs are usually connected to a distribution network, the only possible way to study load characteristics of these customers is to use the statistical analysis. The development of the load research study consists of three steps; 1-sampling selection, 2-metering, and 3-analyzing the collected load data to extract information. The reliability and accuracy of load research results depend on how well these steps are taken. Among these steps the metering part has been facilitated by using the automated meter infrastructure (AMI) systems with a high degree of confidence. Regarding the third step, after collecting the load data in the first two steps, well-developed data mining approaches including pattern recognition methods and clustering algorithms [10,11] are used to extract the load information from the collected load data. However, few researches are done on sampling technique design, and in this regard this paper presents a new method.

Customers in distribution networks have a number of distinct traits that can be divided into some homogeneous, independent and non-overlapping subgroups. These are known as stratification variables which are electricity tariffs, contract power, geographical position and region type. Stratifying customers into subgroups increases the precision of the estimates and reduces the overall required sample size. In this regard, the stratified sampling technique can be used and includes three stages of sample size determination and customer stratification, sample size assignment to determined strata, and sample customer selection [12]. Sample size used in load research is limited based on company's budget to spend on metering and collecting load data. In this way, an optimal method to allocate samples to subgroups should be adopted to increase the accuracy and reliability of the collected data.

Refs. [13–16] give the implementation of the load survey system to identify the load characteristics of customers, by using the stratified sampling concept, to support system planning and operation. In [13], the load profiles of low voltage customers are analysed and a clustering algorithm based on billing data is proposed. Ref. [14] investigates the

* Corresponding author.
E-mail addresses: yaser_raisee@eng.ui.ac.ir (Y. Raeisi-Gahrooei), aminkh@eng.ui.ac.ir (A. Khodabakhshian), hooshmand_r@eng.ui.ac.ir (R.-A. Hooshmand).

https://doi.org/10.1016/j.ijepes.2017.11.029
Received 27 June 2016; Received in revised form 24 July 2017; Accepted 20 November 2017
Available online 04 January 2018
0142-0615/

Nomenclature

Symbol	Description
N_p	population size of customers
N_i	subpopulation size of subgroup *i*
L	number of subgroups
M_p	sample size
M_i	assigned samples to subgroup *i*
w_i	weighting factor of subgroup *i*
σ	standard deviation
μ	mean value
E_{ij}	equivalence coefficient of the consumed energy by subgroup *i* to that by subgroup *j*
$\widehat{CV_i}$	estimated coefficient variation of subgroup *i*
F_{ji}	assigned sample to consumption range *j* of subgroup *i*
$C_{\kappa j}$	assigned sample to activity class κ in the consumption range *j*
LMi	load moment at load point i
Pi	average daily active power at load point i
Li	distance from the load point i to the substation
GLMi	Global load moment at load point i

List of abbreviations

AMI	Automated Meter Infrastructure
CV	Coefficient of Variations
MPCEC	Monthly Per Customer Energy Consumption
LM	Load Moment
GLM	Global Load Moment
EEPDC	Esfahan Electric Power Distribution Company

load profiles to create a load model for the load estimation of distribution transformers. The temperature sensitivities of power consumption of each subgroup are determined in [15]. Moreover, by using the typical load patterns and the temperature sensitivities of subgroups, the effect of temperature change to the power system load demand is studied in [16]. In these studies, Neyman method based on the variation of energy consumption is used to assign samples to each subgroup. This assignment is done proportionally to energy variations and the population size of each subgroup. However, there are some drawbacks that should be taken into account when it is supposed to use this method. First, the goal of the load research is to extract information about load profiles of customers and their variations, while energy consumption variations have been considered here for sampling assignment. Secondly, in this method, the standard deviation is used as a measure of dispersion to calculate the variation of energy consumption of each subgroup. Nevertheless, the mean of energy consumption for different subgroups is widely different. It means that, the lower standard deviation of a subgroup cannot verify the less variable energy consumption of customers and, therefore, a measure of dispersion that is independent from the mean should be used. Thirdly, the accuracy of other steps depends on the sampling step and, therefore, to increase the precision of the load research process, the gathered data of previous works should be used as valuable resources in sample assignment method. However, the method based on the energy consumption variations cannot provide a way to use the prior load research studies and is merely based on the billing data of a distribution company. In order to solve these drawbacks, this paper presents a new method for sample size assignment to the subgroups. The compatibility between load research objectives and sampling design, using an independent mean measure of dispersion of gathered load data and providing a way to use the information and data gathered by the previous load research studies are considered in the proposed method.

Stratifying customers into subgroups and sample size assignment to them are two basic stages and sample selection is the final stage of stratified random sampling technique. In [12] and [17] customers in the residential subgroup are divided into some strata based on their consumption ranges in which the samples are selected randomly. The number of sample customers in each consumption range is considered in proportion to the kWh distribution frequency. In [17] the commercial and industrial customers are randomly selected by considering monthly energy consumption and the number of customers in each industrial and commercial activity codes.

In all previous researches, sample customer placement in distribution network has not been considered in sample selection process. Installed meters for sample customers not only are able to measure the load characteristics but also can measure the voltage level of the connected points in distribution feeders. Consequently, this makes load research study be more appealing for distribution manager. Therefore, in sample selection process, this paper proposes the subject of sample customer placement in distribution network in order to monitor the voltage drops of their connected feeders.

By using the billing information, the collected load profiles and a real low voltage feeder of Esfahan Electric Power Distribution Company in Iran, the proposed new stratified sampling method is implemented. The results clearly show the efficiency of the proposed technique in comparison with the conventional (Neyman) method.

2. Stratified load research random sampling technique

Stratified random sampling technique in load research study is stratifying the customers into some distinct and homogeneous subgroups based on their electricity tariff, contract power and their locations. Then, sample customers are chosen among subgroups randomly. This method includes three stages of determining the sample size and customer stratification, allocating the sample size to obtained strata (subgroups), and selecting the sample customers.

3. Sample size determination and customer stratification

The determination of a proper sample size is an important feature in load research study. In this regard, although a larger sample size yields more accurate and reliable results, it will be costly. A very small sample size provides inconclusive results while a very large sample size wastes resources by providing needless information at a high expense.

In practice, the sample size used in load study is limited and depends on the company budget to spend on metering and collecting load data. Therefore, with the provided metering and collecting instruments, the best sampling design should be adopted to maximize the precision level in estimating load characteristics of all customers.

In distribution networks, customers can be stratified into some separate and non-overlapping subgroups based on the traits obtained from information found in billing records. The traits are available for every customer and are usually electricity tariffs (residential, commercial and industrial), the level of the energy consumption (kWh) and the geographical position. The customer stratification enables researchers to have samples in every subgroup that may be omitted in a more generalized random sampling. Furthermore, stratification can raise the precision of the estimates and decrease the required sample size.

4. Sample size assignment to determined strata (subgroups)

The precision of the stratified design is influenced by how the sample elements are allocated to each subgroup. This provides the optimum number of sample points to give the accurate data. Let consider that there is a population of size N_p divided into L subgroups with