Electric Power Systems Research 111 (2014) 177–184

Contents lists available at ScienceDirect

Electric Power Systems Research

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

Determining spatial resolution in spatial load forecasting using a grid-based model

Joel D. Melo [a,*], Edgar M. Carreno [b], Aida Calviño [c], Antonio Padilha-Feltrin [a]

[a] Department of Electrical Engineering, University of the State of Sao Paulo – UNESP, Ilha Solteira, SP, Brazil
[b] Center for Engineering and Mathematical Sciences – CECE, West Parana State University – UNIOESTE, Foz de Iguaçu, PR, Brazil
[c] Department of Applied Mathematics and Computational Sciences, University of Cantabria, Santander, Spain

ARTICLE INFO

Article history:
Received 14 November 2013
Received in revised form 19 February 2014
Accepted 21 February 2014
Available online 16 March 2014

Keywords:
Electrical distribution planning
Grid-based clustering approach
Spatial load forecasting
Grid-based models
Spatial resolution

ABSTRACT

This paper presents a grid-based model that aims to find a suitable spatial resolution to improve visualization and inference of the results of spatial load forecasting for feeders and/or distribution transformers. This approach can be considered as an unsupervised learning approach to cluster the input data (i.e., the power rating of the distribution transformers) in cells (clusters) to find a cell size that gives high internal homogeneity in the cells and high external heterogeneity of each cell with respect to its neighbors in order to reduce the inference errors that can affect the results of spatial load forecasting methods. The proposal was tested considering the spatial distribution of transformers installed in a real distribution system for a medium-sized city. Using the resolution determined by the grid-based model, it is possible to build a map of the spatial distribution of load density in a service area with a low relative local dispersion and a high relative global dispersion. To demonstrate the efficacy of the approach, spatial electric load forecasting of the study zone is performed using different spatial resolutions; the grid size determined via the proposed model represents the equilibrium between spatial error and computational effort, which is the main original contribution of this work. The techniques of spatial electric load forecasting are beyond the scope of this paper.

1. Introduction

Spatial load forecasting provides information about the spatial distribution of the future loads in a service area. This information helps to identify the subzones with the highest anticipated load growth in the city, allowing the planners of electrical distribution networks to verify that the elements of distribution system will be able to meet the increased load density at each location [1]. The use of this class of forecasting has grown in recent years in countries with large developing economies (i.e., those with rapid urban growth [2,3]) in response to the redevelopment of different cities around the world [4,5]. Moreover, laws have been instated in some countries that mandate the use of spatial load forecasting studies for expansion planning of electric distribution systems [6].

Reference [1] explains the previous steps that must be performed by planners of electrical distribution networks in order to carry out forecasting in spatial form. Among these steps, the determination of the spatial resolution is the most relevant as it determines the amount of information needed to implement the spatial load forecasting studies. The determination of the spatial resolution is the aim of this paper.

All spatial load forecasting methods divide the utility service zone into a number of small areas. The most popular technique used to define these areas is a grid-based approach as it provides enough spatial resolution and detail for distribution system planning [1].

Distribution utilities have increasingly used geographic information systems (GIS) to better characterize the service area in studies of operation and planning of distribution networks [7,8]. Besides, the GIS information permits high-resolution modeling to be incorporated into spatial load forecasting studies [9,10]. While it is possible to use any resolution in spatial–temporal models of spatial load forecasting, the challenge becomes determining which resolution best fits the data available for each system, as the same resolution yields different results for different patterns of load density.

Spatial statistics researchers have highlighted two issues that introduce inference errors that can affect the grid-based approach results. The first one is the term ‘ecological fallacy’ (see [11]), which describes errors caused by the use of results obtained from

* Corresponding author. Tel.: +55 18 37431000x1639.
E-mail addresses: joeldavidm@gmail.com, joel_melo@aluno.feis.unesp.br (J.D. Melo).

http://dx.doi.org/10.1016/j.epsr.2014.02.019
0378-7796/

aggregated data to make inference about local phenomena. The second issue depends on how the data are grouped because the estimates realized by these models can change when the data are grouped at different resolutions; this problem is known as the 'modifiable area unit problem' (MAUP) [11]. These issues are unsolved problems in spatial analysis area [12] and have not been considered before in models to determine a spatial resolution for spatial load forecasting methods. However, they are considered in this work to improve visualization and inference of the results of spatial load forecasting for feeders and/or distribution transformers.

Some models have been developed to explain the problems of ecological fallacy and MAUP [11,12]. These models indicate that planners should strive for a spatial resolution with high internal homogeneity in the chosen grouping and high external heterogeneity of each group with respect to its neighbors. These recommendations can be achieved using a grid-based clustering approach [13].

Grid-based clustering models have been used to summarize data and speed up clustering in order to obtain useful information for exploratory analyses in very large spatial databases [14,15]. This clustering approach can be used to identify a suitable spatial resolution for grid-based models, as will be shown in this paper.

The most common recommendations regarding spatial resolution can be found in [1]. Furthermore, this reference presents a spatial filter to analyze the stability of the frequency spectrum of the error generated by the spatial load forecasting method used. The author recommends that the spatial resolution should be the one that produces less impact on the planning (i.e., high efficacy at the level of spatial prediction). This may be determined by considering the cutoff frequency of the filter. Although this mathematical model is a good method to find the spatial resolution, it does not take into account MAUP and the ecological fallacy.

The spatial resolution used in the grid-based spatial load forecasting method should allow planners to make appropriate inferences from the spatial load forecast results. The model proposed in this paper finds a resolution where the cells present high homogeneity at the local level and high heterogeneity at the global level. Homogeneous cells (i.e., cells with low dispersion) at the local level are defined in this work as cells with consumers that have similar electrical consumption. On the global level, the heterogeneity is used to find a spatial pattern that permits identifying sub-areas that change from a state without load to a state with load in the spatial load forecasting simulation process. To consider the dispersion for both levels, the coefficient of variation (*CV*) is used in the proposed model as a measure of relative dispersion.

The proposed grid-based model can be considered as a process of data analysis that provides an appropriate resolution for grid-based methods used in spatial load forecasting. A map using the calculated resolution will meet the criteria of high internal homogeneity in the grouping formed (i.e., low local dispersion) and high external heterogeneity of each group with respect to its neighbors (i.e., high global dispersion).

To facilitate the construction of the model, the locations of transformers were used as aggregation points. These data were selected because they are available to almost every distribution utility. However, if low-aggregation load data were available, the proposed method could be adjusted without additional modification. Each transformer is considered as a point, and these points are grouped into cells for each regular grid considered in the process of dividing the spatial area. The *CV* is calculated at the local and global levels. The partitioning cost is defined as the ratio between the global and local *CV*. The size of the grid is increased at each iteration of the proposed model and, hence, the cell size decreases. This process is repeated until the partitioning cost reaches a maximum value established by the planner, in order to minimize the computational effort of the grid-based algorithm.

In summary, this paper presents a method to find a suitable spatial resolution using a grid-based model that can help to improve inference for the results obtained in grid-based models for spatial electric load forecasting algorithms. This approach considers clustering the input data (i.e., the power rating of the distribution transformers) in cells (clusters) to find a cell size that gives high internal homogeneity in the cells and high external heterogeneity of each cell with respect to its neighbors in order to meet the recommendations to reduce the effects of MAUP and the ecological fallacy.

2. Spatial resolution for grid-based methods

Reference [1] explains that the behavior of load density varies according to the spatial resolution considered, and presents recommendations for choosing a spatial resolution according to the level of electric planning. For example, a small area size appropriate for distribution substation planning is 0.63 km^2. However, these recommendations cannot be used in all electric distribution utilities due to the different spatial patterns of load of each city.

Furthermore, reference [1] presents a spatial filter to determine the spatial resolution in spatial load forecasting methods. Finding a resolution via the filter operator may lead to errors due to non-natural city growth caused by infrastructure or projects that generate a local growth effect [16]. As explained in [17], the spatial pattern is modified by such projects, and the effects – which can vary temporally – can either be damped quickly or over the course of several years, as explained in [5]. So, these disturbances can modify the spatial pattern, thereby introducing errors in the convolution filter operator as the disturbances are not cyclic [18] (i.e., they are not performed repetitively over time).

In the specialized literature there are very few models focused on determining spatial resolution for simulation models used in spatial load forecasting. Researchers and electric distribution utilities have traditionally selected spatial resolutions for spatial load forecasting according to the level of aggregation of the available information (e.g., zoning maps [19] or maps of land use [20,21]). Grid-size calculators are also available in some statistical software packages [22], but these require specialists with knowledge of remote sensing and the ability to execute the programs.

The problems associated with the ecological fallacy and MAUP cannot be completely avoided in studies relying on aggregated data [11]. Thus, researchers in the field of spatial analysis recommend considering the least aggregated level possible and using clustering techniques or combinatorial optimization to aggregate areas as appropriate [11]. In this work, the least aggregated level available is the level of distribution transformers; this level is grouped in cells of decreasing area until a partitioning cost reaches a maximum value established by the planner (as will be shown in Section 3).

This paper presents a model to find an appropriate spatial resolution for grid-based models by creating a larger grouping that is as homogeneous as possible at the local level (cell) and an appropriate heterogeneous grouping at the global level (grid), in order to permit better inferences from the results of spatial load forecasting. The data used to implement the proposed model have been extracted from [16].

3. Grid-based clustering model

There are many cluster algorithms that differ significantly in their notion of what constitutes a cluster and how to efficiently find them. For example, spatial load forecasting techniques have