Chapter 7

Conclusions

In this work, a powerful technique named bootstrap was presented to estimate confidence intervals containing the missing values of power consumption. Two different studies were carried out. The first one dealt with the estimation of confidence intervals for nodal hourly power consumption. The results of this study are needed in real-time operations to solve load flow problems related to distribution networks. The input parameters for this calculation are the composition of the load and the hour of the year considered. The nonparametric bootstrap was employed in this study. The parameter chosen to characterize the population in that case, was the mode. An estimator of the mode has been proposed and implemented. The second study was about the estimation of confidence intervals for nodal maximum power consumption per customer. These intervals are utilized in planning studies such as the sizing of distribution transformers. The method assumes to be known the composition of the load. The estimation is performed for a given month. It was explained why the original nonparametric bootstrap may provide inaccurate results in that case. The parametric bootstrap was then introduced as a reliable tool to solve the problem. The maximum likelihood estimator was picked in this study.

More precisely, the first step of the work was to process data useful for the confidence intervals calculation. Data processing has been performed by using programs in FORTRAN language. Negative data has been deleted and good data arranged in order to conduct sound statistical study. It was shown that the samples' distributions really depart from a Gaussian distribution. Thus, the use of a classical parametric method to calculate confidence intervals was not acceptable. The nonparametric bootstrap method has been then introduced because it does not assume that the original sample was selected from a probability distribution. This method has been applied to calculate 95% confidence intervals for nodal hourly power consumption. Three different examples have been proposed in Chapter 5 to illustrate it and draw conclusions. It was shown that the size of the confidence intervals was closely related to the composition of the load. The more the load includes customers from the *C*ommercial *L*oad class, the larger the range of the confidence interval is.

In the last part of the study, it was explained that the nonparametric bootstrap may provide inaccurate outcomes in certain case. Notably in the calculation of confidence intervals for maximum power consumption per customer. Here, the parametric bootstrap method has been used. Three examples have been proposed in Chapter 6 to illustrate this approach. In the parametric bootstrap method, a parametric distribution is picked to represent the sample's

distribution. It was shown that, depending on the composition of the load, the parametric distribution that is selected is different. If the load only consists of residential customers, the beta distribution is the best representation for the sample's distribution. The 95% confidence intervals estimated in that case has a very small range. If commercial loads are added, a lognormal distribution best approaches the sample's distribution. The interval estimated is then much larger.

Future Work

As underlined previously, work on distribution networks has received a great deal of attention only in recent years. However, there is still room for improvement, especially in distribution automation. Indeed, an integral part of automation is the real-time monitoring using state estimation techniques. But, up until recently, most of the studies on distribution state estimation have focused on the adoption of transmission state estimator approaches, without accounting for issues specific to distribution networks such as radiality, non-normal statistical behavior of states, load diversity, low ratio of real-time measurements to number of states. Today, some papers have proposed probabilistic approaches to distribution state estimation and paved the way for later studies.

In regard to the specific area considered in this research, a possible way for continued work would be in the testing of the method proposed in Chapter 5 for a real distribution network. By knowing the composition of each load, it would be possible to infer 95% confidence intervals where the values of consumption are not measured. In a second step, load flow studies related to the distribution networks could be carried out. As explained, two load flows may be carried out. The first one would use the lower bounds of each interval as well as the metered values. The second one would utilize the upper bounds and the metered values. From these first calculations, confidence intervals for lines power flows as well as for Ohmic losses would be inferred. These intervals would be finally employed to find the network configuration which would minimize Ohmic losses or avoid lines overload.

Also of interest would be the introduction of the Diversity Factor in the intervals calculation for nodal maximum power consumption. This factor models how the peak demand of a group of customers is not a simple addition of individual maximums demands. It is always superior or equal to one. It is equal to one when all the individual demands reach their maximum at the same time. The upper bound of the interval for nodal maximum power consumption per customer would be picked and multipliate by the number of customers making up the load. This, in order to find the nodal maximum power consumption. One could account for the correlation between loads by weighting this value with the Diversity Factor.