# Big Data for Demand Management Programs Designing for Colombia's Industrial Sector

Jairo Pineda Agudelo, Sandra Ximena Carvajal Quintero, Daniela Valencia Lopez
Universidad Nacional de Colombia
Manizales, Colombia
e-mails: {jpinedaa, sxcarvajalq, davalencialo}@unal.edu.co

***Abstract*—Demand Management Programs is not a new concept; moreover, the key technologies for their implementation are already successful. However, Demand Management Programs applications in a worldwide context have been slow, especially in the industrial sector. Despite this, emerging countries like Colombia have great opportunities to internalize these policies in their energy planning and economic growth programs, so as to maximize their use as a tool integrated to energy markets. Demand Management Programs may allow to deal with the risks associated with system demand and to satisfy the reliability needs of an active and dynamic energy market. For this to take place, one should migrate to active and dynamic demand response, under reliability criteria based on the smart grid paradigm. This paper uses a big data analysis for planning industrial demand-management programs, based on the mechanisms and instruments of demand management and integration processes in smart grids.**

***Keywords- All Data; Big Data; Demand Management Programs; Industrial demand; Energy Efficiency; Open Data.***

## I. INTRODUCTION

Global consumption of electricity has increased 45% since 1980 [1] and is projected to grow 70% by 2030; this is mainly due to growth in emerging countries like China and India [2]. Meanwhile, mature markets, such as North America, Europe and Japan will also face increasing demand and limited fossil resources [3].

In Colombia, industrial demand is the largest consumer of electrical energy from fossil fuels [4] and this causes problems to the environment by greenhouse - gas emissions.

The creation of methodologies to efficiently manage electricity consumption is internationally known as Demand Management Programs (DMP). DMP have environmental benefits, allow postponing investment in electricity infrastructure and increase electricity service coverage, because an efficient management increases transformers capability near the end user and improve voltage profiles [5].

DMP design depends on the accuracy in predicting the power consumption behavior of the target demand [5]. The characterization of the demand allows a prediction that might ensure a balance between generation and demand in order to achieve a more efficient operation of the electrical power system. Therefore, it is possible to avoid cost overruns in the generation, transmission and distribution as well as a resources optimization by the trader that provides the end user with quality standard conditions [6].

The objective of this research is to select industrial users to implement DMP in Colombia. Such response is estimated from data found in a virtual platform, since deregulated users have smart metering devices that send usage information via web to the network operator [7].

The paper is organized as follows: Section II describes the proposed methodology, including response and analytical prediction, Section III presents the case study and Section IV summarizes the main conclusions and proposes future work.

## II. DEMAND SIDE MANAGEMENT AND BIG DATA ANALYTICS: OVERVIEW

DMP have been implemented to achieve a better participation in demand compared to electrical market prices or the need to improve reliability levels of the electric power system. Participation of demand in DMP seeks to mitigate the constraints of power grid and yield economic benefits for stakeholders [8]. Historically, the possibility of increasing the efficiency of the system and the existing investment in generation and transmission of electricity has been the key to the introduction of DMP [9].

DMPs in mature markets, like USA and Europe, show consumption reductions up to 40% in peak hours and a reduction in the need for generation reserve up to 50% [8]. The effectiveness of DMPs in USA and Europe is due to their technology implementation, which allows them to exert active demand control [9]. In the regulation, they have also incorporated various DMPs, which are based on studies of user's behavior and habits with assistance of methodologies related to Analytics.

Latin America has slowly begun integrating policies related to the efficient use of energy and the integration of demand management programs in residential, commercial and industrial sectors. Chile, Brazil and Colombia have been referring to the region for the implementation of innovative energy policies.

Particularly, in Colombia, there are initiatives that allow, in some way, implementation of some basic concepts about the possibilities of managing electricity demand in energy markets; such initiatives are advances in demand management in the country, although not are demand management programs. Among these initiatives there are the hourly pricing for the case of industrial demand and the recent decision of the regulator regarding contracts load shedding, disconnecting Voluntary Demand [10], as well as Law 1715 2014 [11], which considers the participation in energy efficiency, including renewable energy.

However, ignorance of the laws and regulations, as well as the lack of adequate human and financial resources, plus the difficulty of designing realistic and clear goals with measurable results, makes it difficult to quantify the potential benefits of these policies [12].

Currently, Colombia has data provided by concentrators of measurements, but no studies have been developed based on this information. A next step is to exploit the billions of data rows to work on transforming useful knowledge in order to provide answers to operational and market issues with active participation of demand [13].

The active participation of demand refers to the study of the users consumption behavior and the changes in those habits regarding tariff incentives that may include hourly-based price discrimination on a daily basis and by use periods [14][15].

The relation between the study of the target demand active participation and the regulatory incentives that can be implemented is the basis for the design of possible successful DMPs. An analytics method that provides guidelines on the best conditions for the development of DMPs is necessary to determine such information from data obtained by the meters [16].

Big Data Analytics is the application of advanced analytics techniques to operate on large data sets. Predictive analytics, data mining, statistics and artificial intelligence are among these advanced data analytics to examine data [16].

The management of such tools to analyze giant datasets requires identifying, combining and managing multiple data sources, and the ability to build advanced analytics models for the prediction and optimization of results. The most critical component is related to the ability to transform the passive contribution of energy customers into an active contribution in real-time to achieve the purposes of DMPs for industrial users.

The information demand for these smart grids involves the consumption quantification and characterization, including the study of clusters within industrial consumers to determine, for example, economic similarities in users with similar loads. This fits the broader definition of Big Data (large data or macro data according to Fundéu BBVA Foundation), since the meters owned by the unregulated user generate an avalanche of data that must be exploited to dramatically improve DMPs performance. According to Harvard Business Review [17], the evidence is clear: “decisions controlled by data tend to make better decisions”. Particularly, electrical system operating practices need to give greater emphasis to effective real-time operation with accurate and timely information and state-of-art technology to facilitate effective contingency planning.

Currently, Management of system security needs to keep improving to maintain reliable electricity services in this more dynamic operating environment. The challenges raise fundamental issues for policymakers [18]. Big data promises to have success in to design energy policy, since the evidence is clear: decisions controlled by data tend to make better decisions.

Along with this disruptive force associated with the data growth, data analysis has evolved into what is now known as the analytics, visualization and, in particular, data mining, from the traditional disciplines of statistical data analysis. It is actually a complement of tools caused by the evolution rather than a dilemma in terms of a break since the statistical criteria remain valid in the sense of preventing failed predictions to make Big Data in Big Data Winter. For Michael Jordan [19], this may be "Due to simple-minded and statistically unsound approaches which will produce too many false positives."

True, the immense amount of data poses unprecedented challenges in terms of its analysis, due inter alia to qualitative change that implies an increase in the amount. However, a statistical point of view is essential as it contributes to consider the component of uncertainty in predictions and errors quantification. This is missing in much of the current literature of Machine Learning [19].

The massive data production growth in all modern society dimensions and the demand for prompt responses for decision-making is a powerful challenge for data scientists who are at risk of an inappropriate statistical coverage of their work and provide models with high randomized bias. Language pushed by the urgency of the decisions ends up being the result of what Douglas Merrill called the "button effect" [20], which refers to a simple expression of the pitfalls of chance and not sufficiently controlled in the model. "Button effect" occurs when surface data analyses are made and this causes erroneous results. In this paper, we propose to reiterate the rigorous use of Big Data tools to ensure correct results.

The next step after data production is to focus on analysis in order to overcome the situation, as described by Graham Williams who suggests that we are very rich in data but very poor in information [21]. This concern coincides with IDC International Data Corporation, predictions for the year 2017 [22], including the role of the CIO Chief Information Officer, that focus 80% of their time on analytics, cyber security and new revenue sources creation through digital services. This does not include recent cybersecurity issues and expanding revenue sources that, despite their importance, are not the subject of this article; the interest is to focus the discussion on the elements of Analytics and its benefit in DMPs for an industrial user.

The study of time series originated in the periodic consumption data collection is the basis for the construction

 ISBN: 978-1-61208-445-9 

of the demand curve, consumption modelling and demand prediction. Advanced analytics also offers the use of decision trees, neural networks and support vector machines to predict time series problems.

The Multilayer Perceptron (MLP) [23] appears to be the most used architecture of artificial neural networks for prediction of nonlinear time series. Forecasting electricity prices and demand are counted among its many applications.

Autoregressive neural networks (ARNN) are obtained by considering the merging of a linear autoregressive model with an MLP [23]. Their initial conceptual development is based on the development of a statistical test for nonlinearity to compare the two previous models. However, ARNN is an important alternative to the use of MLP in predicting time series due to the incorporation of the linear autoregressive component [23].

Support Vector Machines (SVM) [24] are a type of neural network that was originally designed for the solution of nonlinear classification problems; but recently, they have been applied to time series regression and forecasting problems. This is due to their generalization features, which is a direct function of their structure and methodology used for their parameters estimation [24].

Similarly, it is possible to conduct a study of clusters based on the response to different price levels. This information is included as an objective in this first stage of the DMP and it is extremely important to have data available from open source like the traditionally so-called Open Data, more recently called Urban Data [25]. Subsequent claims are related to operations analysis and analytical applications.

## III. CASE STUDY

As mentioned before, the aim of this work is industrial demand. In Colombia, industrial demand is classified within the so-called unregulated users, *i.e.*, those users with consumption above 0.1 MW. The current regulation allows them to purchase electricity at prices agreed freely [6], which usually causes reductions in kWh prices, compared to the price the residential user must pay for the same kWh [26].

To access tariff privileges, industrial user must install a measuring system with telemetry capacity to determine the traded energy hourly [6]. This information is recorded in a web site via the Internet that can be accessed by the user later.

The study was conducted in the west central area of Colombia and industry data pertaining to the metropolitan area corresponding to the municipalities of Manizales, Villamaría, Neira, Palestine, and Chinchiná were used.

Unregulated users in that region were identified according to the operating company of the Colombian electricity system [27] in order to model and characterize the curve of daily demand. Consumption of industrial users will be used since the meters installed send this information to a web site.

After a characterization of industrial users and knowing the main factors that influence their behavior, the next step is to create a DMP appropriate for users, which is crucial because of the change in demand by industrial activity. Another aspect that directly affects the development of a DMP is the knowledge that the end user has about the active participation benefits he gets. It is also important to know the law requirements and its participation in these programs and the technologies that will be used for remote controls and monitoring on consumption in real-time.

This paper considers the influence of the type of activity within the industrial sector in the demand behavior to develop diversified demand graphs that allow observing a typical demand curve for selected industrial activities. Figure 1 shows the graphs of diversified demand corresponding to various activities of industrial production.
The demand factor in a range of a distribution system or a load is the ratio between its peak demand in the range considered and the total installed load.

The demand factor is a dimensionless number; therefore, peak demand and installed load must be considered in the same units, the demand factor is generally less than 1 and it will be unitary when, during the interval, all installed loads are operating at nominal power. Therefore, in the time intervals in which the demand profiles shown in Figure 1 exceed the unity, industrial users are considered to be operating above 100% of permissible value for the electrical installation.

Exceeding safe operating condition means overloading the electrical system of each user. Figure 1 shows that generally this behavior occurs about two standard hours, the first reflected at 8:00 when work activity begins, and the second can be seen at 16:00 before the end of the day.

The generalized overload behavior at certain intervals causes widespread power quality problems [15] since the voltage supply is reduced and there are technical losses in the power system due to electrical current increase. These technical conditions result in economic losses for industrial users since electrical machines can reduce the production and their lifetime might decrease as well [15]. Additionally, Figure 1 shows the trend of stable consumption during the day with exceptions in the activity of cement and iron and steel, which have more marked valleys at 12m throughout the workday. In order to identify the fundamental reason of that behavior, it is necessary to have detailed knowledge about the production process development.

Typical demand curves form the basis of statistical analysis for the corresponding decision-making in terms of DMP. They gather the consumption history of different industrial activities and allow defining the demand profiles through the statistical analysis methodology of time series.

The description of such curves includes identifying the long-term trend in industrial use and its seasonal component, present in all of them, given the regularity of energy consumption behavior, which is typical in any industrial activity, as deduced from visual analysis of Figure1.

The input provided by the description of the curve displays the consumption through historical time in various industrial activities and, therefore, the proposals to manage demand. Table 1 shows an operation typology in which several criteria that must be clear at the time of initiating the process of analysis and data collection are identified.

TABLE I DIFFERENTIATING FACTORS IN OBTAINING DATA

| DEVELOPMENT CRITERIA | OPTIONS | |
|---|---|---|
| User Type | Regulated User | Unregulated User |
| Initiative by | Power system operator | Agents of the liberalized market |
| Consumer type | Low Voltage (small businesses and home users) | High voltage (Industry and high trade) |
| Obtaining information | Installation of smart meters and creation of specialized software to gather information | Telemetry devices owned by the user, to discriminate the value of consumption on an hourly basis |

The target population of this work has the features framed in column 3 of Table 1. Data collected from the target population are used to process first description and then forecasting, so that in the time horizon defined by operators and in accordance with industry, the curves are projected in time and therefore demand. Hence, it is possible to harmonize generation and consumption in relation to the forecasts provided by the data. Of course, it is always desirable to supplement the historical course provided by the demand curve with one related to the prospects for consumption in the short- and medium-term. The final model gathers past behaviors and the immediate prospects.

The core of the Analytics of typical consummation demand curves is the Analysis of Time Series whose objective is to identify those components that are present to detect its causes and to predict future series values. Table II identifies the models for the four industrial activities.

TABLE II MODELS FOR THE FOUR INDUSTRIAL ACTIVITIES

| Economic Activity | Exponential Smoothing (Brown) | ARIMA (0,2,0) |
|---|---|---|
| Chemical | α = 0,844<br>R square = 0,937<br>Statistical Ljung-Box<br>Q(18) = 18,22 (0,375) | |
| Food | α = 0,796<br>R square = 0,933<br>Statistical Ljung-Box<br>Q(18) = 7,936 (0,968) | R square = 0,939<br>Statistical Ljung-Box<br>Q(18) = 7,936 (0,905) |
| Drinks | α = 0,924<br>R square = 0,949<br>Statistical Ljung-Box<br>Q(18) = 13,005(0,736) | R square = 0,955<br>Statistical Ljung-Box<br>Q(18) = 15,908 (0,599) |
| Plastics | α = 0,864<br>R square = 0,901<br>Statistical Ljung-Box<br>Q(18) = 9,455(0,925) | R square = 0,935<br>Statistical Ljung-Box<br>Q(18) = 12,056 (0,844) |

The models of exponential smoothing and ARIMA show similarities for the four activities, R square high and significant Statistical Ljung-Box Q [28] (SPSS V. 22)

Analytics combines a statistical and an algorithmic modelling approach to build the basis of DMP.

## IV. CONCLUSION

Demand management programs, framed in the context of data analysis show great potential and a promising future, especially if it is possible to realize the benefits of demand modelling as a fundamental step in understanding industrial user behavior. This is important firstly to estimate the changes that might occur in the demand curves and secondly to assess the impact of these changes in the power system.

At present, it is necessary to make a use of energy resources because of overall demand growth. Thus, sound knowledge about power consumption will have environmental and social impact. By acquiring knowledge, it is possible to transform the thinking of people for efficient use of energy.

Industrial users have particular behaviors given their productive activity, which can be a downside for DMP design. However, this type of users is important because they can have more active participation in a DMP; they also have the resources and technology to provide the data to be analyzed.

Data Analytics is the answer to the exponential growth of data that the industry acquires through remote sensing in order to provide management tools of power demand.

## REFERENCES


[1] “IEA International Energy Agency 1874-2014 Energy Policiy Highlight”, http://www.iea.org/publications/freepublications/publication/energy_policy_highlights_2013.pdf, [Accesed: Dec, 2014].

[2] “International Energy Agency Word Energy Outlook”, http://www.worldenergyoutlook.org/media/weowebsite/factsheets/WEO2013_Factsheets.pdf, [Accesed: Dec, 2014].

[3] “World Energy Investment Outlook”, International Energy Agency, http://www.iea.org/publications/freepublications/publication/weio2014.pdf, © OECD/IEA, [Accesed: Dec, 2014].

[4] B. Paola and C. Angela Ines, “Benefits of Implementing a Demand Response Program in a Non-regulated Market in Colombia”, Innovative Smart Grid Technologies (ISGT Latin America), 2011 IEEE PES Conference on. October 2011. Medellín, Colombia.

[5] K. Jungsuk and R. Ram, “Demand Response Targeting Using Big Data Analytics”, 2013 IEEE International Conference on Big Data, doi: 10.1109/BigData.2013.6691643, Oct 2013, Santa Clara, CA, USA.

[6] Resolution 131 of 1998. Official Journal No 43.465 of December 31, 1998. Comisión de Regulación de Energía y Gas (CREG). http:www.creg.gov.co, [Accesed: Dec, 2014]..

[7] L. M. Johanna, N. P. Phillip, K. Sila, and P. Mary Ann. “Quantifying Changes in Building Electricity Use, With Application to Demand Response”, IEEE Transactions On Smart Grid, vol. 2, no. 3, September 2011, pp. 507-518.

[8] J. M Victor and R. Hugh, “Design of demand response programs in emerging countries”, Power System Technology (POWERCON), 2012 IEEE International Conference

on doi: 10.1109/PowerCon.2012.6401387, November 2012, Santiago de Chile, Chile.

[9] C. Adela and L. Pedro, "Estimating the benefits of active demand management. Review of the state of art and proposals", Economic Notebooks of ICE, ISSN 0210-2633, Nº 79, 2010 p.p. 187-212.

[10] Resolution 063 de 2010. Official Journal No 47.700 of april 27, 2010. Comisión de Regulación de Energía y Gas (CREG). http:www.creg.gov.co, March 20, 2015.

[11] Law 1715 of 2014. "By regulating the integration of renewable energies non conventional of energy to the system national", May 13, 2014, Gobierno Nacional de la Republica de Colombia, Bogota, http://wsp.presidencia.gov.co/Normativa/Leyes/Documents/LEY%201715%20DEL%2013%20DE%20MAYO%20DE%202014.pdf, [Accesed: Mar, 2014].

[12] B. C. Paola, "Implementation of a program to the demand response for electric energy in an unregulated customers of market in Colombia", Rev. maest. derecho econ. Bogotá (Colombia) vol. 6, no. 6, pp: 259-292, Dec 2010.

[13] Schneider electric, Energy efficiency: Solutions Manual. http:www.schneiderelectric.es/.../eficiencia-energetica/eficiencia-energetica, [Accesed: Dec, 2014]..

[14] P. Jennifer. Analytics at SMUD evolve with the smart grid, Nov 4, 2014, http://www.intelligentutility.com/article/14/11/analytics-smud-evolve-smart-grid, [Accesed: Dec, 2014].

[15] I. Toshifiimi, H. Yusuke, and T. Kiichiro, "Definitions of Power Quality Levels and the Simplest Approach for Unbundled Power Quality Services", Ninth International Conference on Harmonics and Quality of Power Proceedings, vol 2, Oct 2000, doi: 10.1109/ICHQP.2000.897711, Orlando-Florida.

[16] J. A. Luis, "Big Data Analysis of large volumes of data in organizations", Alfaomega Marcombo Ediciones Técnicas, 2013

[17] M. Andrew and B. Erik, " Big Data: The Management Revolution" October, 2012, https://hbr.org/2012/10/bigdatathemanagementrevolution/ar, [Accesed: Dec, 2014].

[18] C. Q. Sandra, S. Jean, and A. Santiago, "Colombian ancillary services and international connections: Current weaknesses and policy challenges", Energy Policy, vol 52, Jan 2013, pp 770‑778, Special Section: Transition Pathways to a Low Carbon Economy, doi:10.1016/j.enpol.2012.10.041.

[19] J. Michel, ¨Big Data Winter ahead – unless we change course¨, By Gregory Piatetsky, Oct 30, 2014. , http://www.kdnuggets.com/2014/10/big-data-winter-ahead-unless-we-change-course.html, [Accesed: Dec, 2014].

[20] Douglas M. "Careful with easy answers of Big Data". HBR, August 5, 2014.

[21] W. Graham. "Data Mining with Rattle and R, The Art the Excavation Data for Knowledge Discovery". Springer Use R!. New York 2011.

[22] The 10 predictions of the CIO's agenda in the coming years, http://www.ticbeat.com/tecnologias/10-predicciones-agenda-cio-proximos-anos-segun-idc/, October 31, 2014.

[23] V. Juan, Z. Cristian, and V. Laura, "ARNN: A packages for time series forecasting using autoregressive neural networks" , Computer systems and informatics breakthrougs magazine, vol.8, no2, Jul 2011, Medellin-Colombia, ISSN 1657-7663.

[24] V. Juan, O. Yris, and F. Carlos, **"**Time series prediction using support vector machines", Ingeniare. Rev. chil. ing. v.18 n.1 Arica abr. 2010, pp. 64-75, http://dx.doi.org/10.4067/S0718-33052010000100008.

[25] B. Luciano, P. Kien, S. Claudio, R. V. Marcos, and F. Juliana. Structured Open Urban Data. Understanding the Landscape. BIG DATA, doi: 10.1089/big.2014.0020. [Accesed: Sep, 2014].

[26] Document CREG-138: Revision of the limit unregulated user of electricity power, December 2009. Comisión de Regulación de Energía y Gas (CREG). http:www.creg.gov.co, [Accesed: Mar, 2015].

[27] Especialistas en mercado XM filial de ISA, "list of users not regulated by voltage levels", http: http://www.xm.com.co/Pages/UsuariosNoReguladosporNivelesdeTension.aspx , November 2014.

[28] P.O. Hermelinda, Estadistics II, Exponential Smoothing (Brown), Chapter 5, tips models, Facultad de Ciencias Exactas y Naturales, Universidad Nacional de Colombia http://www.virtual.unal.edu.co/cursos/sedes/manizales/4030006/lecciones/capitulocinco/5_2_3.html, [Accesed: Mar, 2014].

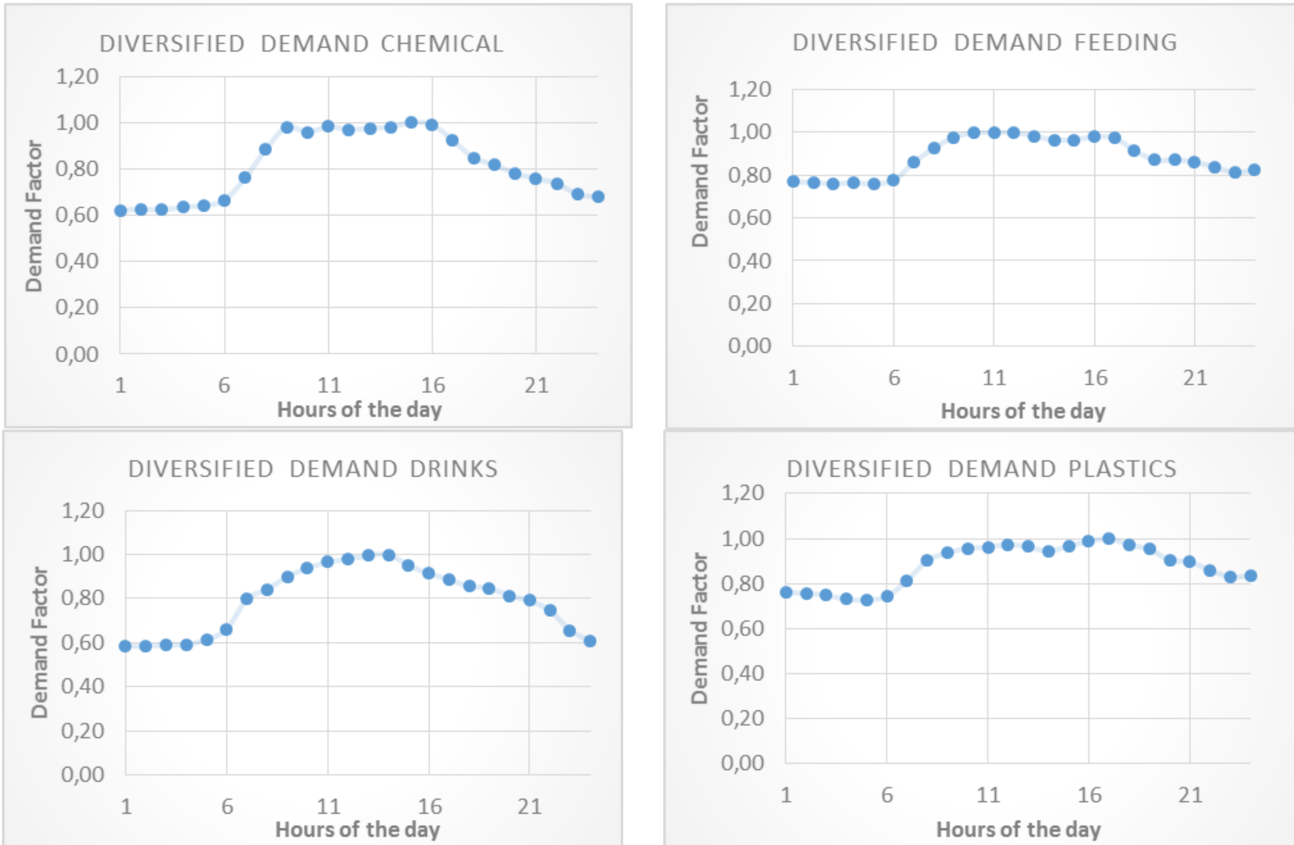


Figure 1. Typical curves of electricity demand by economic activity of industrial users.