# Details of the creation of input tables for SWITCH Mexico Model at University of California, Berkeley

Aldo Sayeg Pasos Trejo. Facultad de Ciencias, Universidad Nacional Autónoma de México.

July 27, 2016

## 1 Loads

### 1.1 Hourly demand table creation

Inside of the "Loads/Data Analysis" directory of the "Switch_mexico_data" repository there is a directory for each of the projections of hourly load in Mexico. There is a high, mid and low projection. Each one of this projections is made by the CENACE ("Centro Nacional de Control de Energia", National Center for energy control).

This projections consist of a table with the hourly demand for every Control Region of the SEN ("Sistema energetico nacional", nacional energy system) from 2016 to 2030. On the other hand, the SENER has also a table of the porcentual assigment of load for each Node of every Control Region of the SEN. With this information, we were able to make a table consisting of the hourly demand for all the SEN Nodes from 2016 to 2030 in each of the 3 projections.

Due to the amount of data, colecting all of this in a single cvs file would have been too large for efficient manipulation inside Python. Also, as these tables were made manually due to the fact that the original format was an xlsx file with several sheets, it would have been really long to create a single file with all the data at these point.

As a result of this, we splitted this hourly demand per Node table in 15 cvs files contained in the "tables" folder of each projection directory, one for each year from 2016 to 2030. Each of the files has the name of the year of which it contains data. All of this tables were made manually in Excel, and then saved to the desired format.

### 1.2 Python scripts

For each projection, we made a Python script called "Table creation" that creates some desired tables. First of all, the script created a single data frame called "df" that contained the data of every year (a concatenation of the yearly-named csv files in tables). It exported this file to the "tables" folder as a large csv file called "HourlyLoadPerNode".

The same script creates another table, called "LoadHighlightsPerNode" that calculates, for every node at each month and year from 2016 to 2030, the peak day of the month ("PeakDay" column), the peak day average load ("PeakDayValue" column) and the average of the month ("MonthlyAverage" column). We will procede to explain how were these columns calculated

#### 1.2.1 "LoadHighlightsPerNode" data details

For every Node in the columns of the "LoadHighlightsPerNode" and a given row of year and month, the "PeakDay" column corresponds to the day of that month and year for which the average load per hour (the sum of each hour load divided by 24) is maximum compared to the other days.

The "PeakDayValue" column shows this value, the mean of the hourly load in that day.

On the other hand, the "MonthlyAverage" value calculates the daily average of the daily load for every day in that month and year, but excluding the peak day. For example, if at node "Hermosillo" at year 2016 and month 09, the peak day is 03, the Monthly average column will correspond to the sum of the daily load of days 01,02,04,05,...,n (n is the number of days in month 09 at year 2016) divided by n-1 (as we have excluded the peak day, we hace only summed n-1 days, so for the average of these sum we hace to divide by n-1)

The Script called "Graphical Analysis" displays graphics corresponding to this data. The first part displays the average hour load for every day in a year, for each year from 2016 to 2030, in a separate graphic ("figure", in matplotlib's termns) for every node in the SEN.

The second part asks you for an input consisting of a Node, a Month and a year. For that informations, it displays two graphics: one consisting of the hourly load for every day in the month and another one of the hourly load for selected days.

This selected days are the following: the "PeakDay" from the "LoadHighlightsPerNode" table, the hourly average median day of the month (the day of the month whose hourly average load is the median compared to all the other days hourly average load), the day whose hourly average is closest to the "MonthlyAverage" value from "LoadHighlightsPerNode" table and a random day

The last script called "Switch input creation and SQL import" creates a table called "la_hourly_demand_*" where * is high, low or medium depending of the projection directory we are working in. These table is saved at the "tables" folder of each projection directory. It also writes this table to the SWITCH database, that it is written in SQL, using a previously created SSH tunnel.

This table is used as input data for the SWITCH model, hence the importance of writing it to the database