A generic multi-period economic dispatch model for the power plants operated by Duke Energy and Solar & Biomass IPP in the Carolinas region built using publicly available datasets

Abbhijith Hari Gopal

Abstract:

This paper involves developing a generic multi-period (or dynamic) economic dispatch model to model the power plants operated by Duke Energy and Solar & Biomass IPP in the Carolinas region built using publicly available datasets, making appropriate assumptions. The main objectives are to minimize system cost, and CO_2 footprint, and to increase reliability. The analysis was carried out on December 2, 2020. The 24-hour time period was divided into two time zones namely 7 AM to 6 PM & 7 PM to 6 AM. In the first period, 7 AM to 6 PM, solar IPP was introduced whereas their power o/p was set to 0 MW for the second period (7 PM to 6 AM). Preliminary data analysis was done in MS Excel, Python's Pyomo library, and 'cbc' solver were used to code & solve the model, and TABLEAU was used to visualize the generation mix.

1. Introduction:

1.1 Literature Review

Shen et al. studied the cost minimization economic load dispatch model with LMP elastic demands. The way they modeled their study is as follows: First, they formulated the expression for the demand function of an elastic load which was a function of locational marginal price. Power usage was then defined. Since LMP is a dual variable and its value is unknown before the optimal power flow problem (OPF) is solved. However, given the unclear nature of the elastic demand, setting up the OPF model is not possible. The authors address this problem by re-defining LMP at discontinuous points to obtain an equilibrium solution. Thus, the MILP ED problem was set up and the solutions to this model give insights into dispatch and LMP (Shen, Wei, Wu, Shafie-khah, & Catalão, 2021). Considering the characteristics of heat recovery system generators, Chen et al. built an ED model for combined heat and power (CHP) systems. The authors were trying to solve for the contradictory nature of CHP systems where there is an inverse relationship between accuracy and solvability. They proposed a novel model to balance the two problems. Using the concepts from heat exchange theory, they integrated HRSGs into the traditional ED model. They had a cost-minimizing objective function where the costs included total operation cost, cost of thermal plants, and a penalty term for wind power. The constraints included power balance, heat balance, CHP plant constraints, thermal plant constraints, and network constraints. The convex optimization problem was solved in MATLAB 2017b with IPOPT which is an open-source solver that works on the primal-dual interior-point method. To verify their results i.e, checking for the deviation between the proposed model and results from simulation, the authors used SimuWorks which is commercial software for thermal system simulation. They concluded that the optimal dispatch results indicated that HRSG characteristics have a large influence on the participation of the CHP plant in the economic dispatch. Also, in terms of optimal dispatch results, the proposed model outperformed the conventional polygon. Finally, the computational time for solving the proposed model on the test system is within the scale of seconds (Chen, Guo, Chen, & Sun, 2020). To model for high penetration of wind power, Lin et al. proposed a mean tracking model based on stochastic ED dispatch for power systems. The forecasted for wind generation were taken as a reference for the SED model. The authors contended that the existing methods seldom factor in the preschedule of wind power and the resulting solution from the ED model largely differs from the preschedule. To re-adjust this, the authors used the mean tracking algorithm where the objective is to minimize the generation cost and tracking error. A Quasi-Monte Carlo method was used to factor in the uncertainties in wind power. The multi-objective problem was solved by GSOMP. The results showed that the proposed method fully explores the value of the base pre-schedule and the obtained dispatch solution becomes more practical in the actual operation of power systems (Lin et al., 2020). Similarly (considering high renewable penetration), Velasquez et al. modeled a distributed predictive control (DPC) ED model with high renewable penetration. The authors contended that the system operator must make full use of forecasted values of renewable power generation in the decision-making process to reduce the unpredictable generation changes. They used a closed-loop algorithm for solving economic dispatch at runtime while reducing potential deviations of generation schedules. The authors showed initially that by using MPC techniques, the ED model could be enhanced. Then, they used DDMPC to show a) most benefits of MPC and b) verify behavior in a case study with high renewable penetration. They concluded by asserting the better performance characteristics of the proposed approach since it was able to constantly anticipate changes in the system. Advantages of DDMPC over MPC were also highlighted (Velasquez, Barreiro-Gomez, Quijano, Cadena, & Shahidehpour, 2019). Patino-Echeverri et al. studied the economic and environmental implications of different approaches to hedge against wind production uncertainty in two settlement electricity markets in PJM. Basically, they had 4 two-settlement electricity market clearing designs. The first design was the DMC (deterministic market clearing) design that is currently used across the US for wholesale electricity markets. For the other three designs, the authors had day-ahead wind production uncertainty in the market mechanisms. To account for sufficient generation from VERs, they used an augmented deterministic market clearing price mechanism that uses the concept of ramp capability products. The last design incorporated a stochastic market clearing (SMC) mechanism to explicitly account for uncertainty in wind production. The test system had a 12% capacity of PJM's fossil power fleet. The data collected spanned from 2010 to 2014. Simulations were done on an hourly basis for the whole year. Their results highlighted that SMC is superior as its cost reductions are more than two times the improvements attained by ADMC and HDMC. The SMC results in electricity prices that are better aligned with operation costs and cuts the spread between the day-ahead and real-time prices by > 40% (Daraeepour, Patino-Echeverri, & Conejo, 2019). Alqahtani and Patino-Echeverri studied the combined effects of policies to increase energy efficiency and distributed solar generation in the Carolinas. They estimated changes in the cost of electricity, reliability, and atmospheric emissions resulting from large penetration of residential roof-top

Photovoltaic (PV) and end-use energy efficiency (EE) within the service areas of Duke Energy in the Carolinas. They solved for day-ahead unit commitment model and real-time economic dispatch model. They concluded by saying that 8.7–10.2% of 2015 electricity consumption could have been avoided by upgrading all residential units to comply with Energy Star standards. They also highlighted a reduction in system costs and carbon emissions (Alqahtani & Patiño-Echeverri, 2019). Cornelius et al. assessed the environmental, economic, and reliability impacts of flexible ramp products in MISO's electricity market. The authors contended the fact that with high penetration of wind and solar, dispatching conventional generators with sufficient ramping capability is an absolute necessity since it can save the system on last-minute dispatches and various costs associated with it. MISO modified their UC/ED model to take this ramp capability (RC) into account. The authors basically explored the outcomes of the market-clearing process by simulating 10-minute operations of a scaled MISO test system, for three representative months under low and high wind penetration scenarios. Their results showed that adding RC products lowers CO2 emissions and system costs, facilitates wind integration, and improves reliability (Cornelius, Bandyopadhyay, & Patiño-Echeverri, 2018).

1.2 Research Gap and focus of the current study

While previous research articles focus on various pathways for developing the economic dispatch model using a plethora of algorithms, all these works use datasets from utilities that are not public & offer very little flexibility for users to fine-tune the model to fit their case. As a result, novice professionals face innumerable challenges to model the economic dispatch for their region of interest.

In this study, we will come up with a generic multi-period (or dynamic) economic dispatch model to model the power plants operated by Duke Energy and Solar & Biomass IPP in the Carolinas region built *using publicly available datasets*, making *appropriate assumptions*. This is a cost minimization problem having the following objectives:

i) minimize system cost,
ii) minimize CO_2 footprint, and
iii) increase reliability.

To replicate the analysis, users need to modify the demand sheet, generation sheet, lines sheet, and reserves sheet in the excel workbook. Python code needn't be modified, should the assumptions discussed in 2.3.2 remain the same.

2. Materials and Methods

2.1 Data Preparation:

For the generation data, plant names, fuel types, and capacity were obtained from Duke Energy's public webpage & EIA's eGrid. Since we are using only publicly available datasets, it is not possible to get exact values for all power plant parameters. Therefore, industrial averages were used from reliable/trusted sources such as government webpages, and EIA. Parameters obtained by such methods include efficiency, fuel cost ($/kWh), and annual net. generation (MWh/y), Annual CO_2 emissions (tons/y), the social cost of carbon (SCC), and the capacity factor. Nodes were assigned to plants by the K-Means Clustering algorithm, which will be discussed in section 2.1.1

2.1.1 K-Means Clustering:

K-Means Clustering is a very popular, simple, unsupervised machine learning algorithm that basically groups the data points to identify patterns. The way the algorithm works is as follows: First, a random no of 'k' clusters is selected. This 'k' is specified by the user. Centroids are then randomly selected. If there are 2 clusters, for instance, two centroids are selected. Then, by measuring the distance from each point to the centroid, clusters are formed/arranged in such a way that the sum of distances between the points and the respective cluster centroid is minimized. Centroids are again re-computed and this is an iterative process, the termination of which occurs when the previously-mentioned condition is met. For our model, K-Means clustering analysis was done in python. Once the plants of interest were obtained, it was geographically mapped. K-Means clustering involves having data points on the Cartesian plane rather than on a map. Therefore, corresponding latitudes and longitudes were mapped to X and Y values. From the elbow plot shown in **Figure 1**, a linear decrease occurs from k = 4. Therefore, 'k' was chosen as 4, meaning we have a 4-node power system. We have 999 plants in total. From K-Means Clustering, 247 plants were clustered in node 1, 6 plants in node 2, 73 plants in node 3, and 673 plants in node 4.

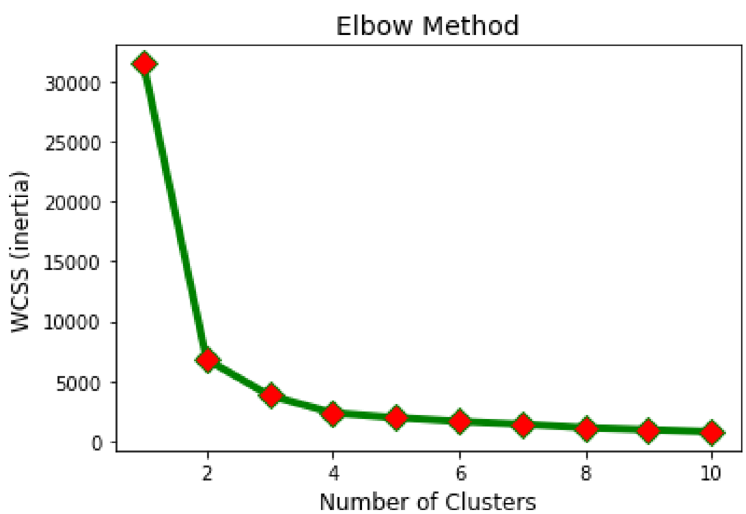


Figure 1: Elbow method plot (K-Means Clustering)

The demand data was obtained from the EIA website. Since the latest generation data available was for the year 2020, to maintain the same level of granularity, hourly demand data for Dec 2, 2020, for regions operated by Duke Energy was obtained. ED model is developed for 24 hours, split into two time zones namely 7 AM to 6 PM & 7 PM to 6 AM. In the second time period - 7 PM to 6 AM – solar power production was set to 0 and the demand is met by the other plants.

The variation of total demand with the time plot is seen in **Figure 2.** Since we are dealing with a 4-node system, demand was apportioned per the number of plants in each node (24.7% demand in node 1, 0.6% demand in node 2, 7.3% in node 3, 67.4% in node 4).

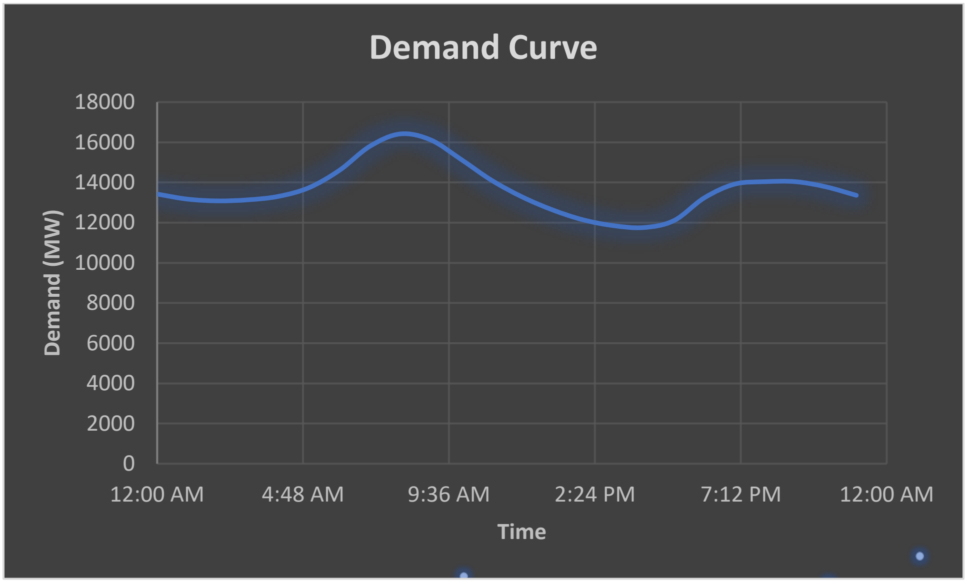


Figure 2: Demand variation (Dec 2, 2020)

Considering our third objective, i.e, maintaining system reliability, it is imperative to have reserves running. In this model, reserves are assumed to be 5% of the demand. Thus, reserve data were obtained accordingly.

Finally, appropriate values were chosen for line capacity and reactance data for the lines between the 4 nodes after perusing the literature. More of this is discussed in assumptions in section 2.3.2

2.2 Software(s) used:

2.2.1 MS Excel

MS Excel was used to collect, clean, and manipulate data. This was the first part of the analysis – getting the input in the right format that could be fed into Python. The main sheets include generation, demand, reserves, and lines data. The README sheet subsumes links from where the data is sourced.

2.2.2 TABLEAU

Once the generation mix was obtained as outputs from python, they were fed into TABLEAU to obtain pie charts depicting their distribution across the time periods, per fuel type (discussed in Section 3). Python could also be used for the same. TABLEAU was preferred to python for a couple of reasons: i) user-friendly, the intuitive environment offered by TABLEAU, and ii) Drag and drop interaction over code-based in python.

2.2.3 Python

Python was used to code for the dynamic ED model. Pyomo package was used for optimization. The excel file mentioned in 2.2.1 was fed in as input. Inline comments are written to enhance code readability. The code generates 3 output files which include 2 excel files and a PDF. The excel output files are i) a Generation mix of all power plants at different times and ii) a file subsuming the locational marginal prices and power flows across the transmission lines at different times. The PDF doc shows the plot of the variation of locational marginal price with time.

Figure 3 shows the data flow diagram which encapsulates (2.2.1-2.2.3)

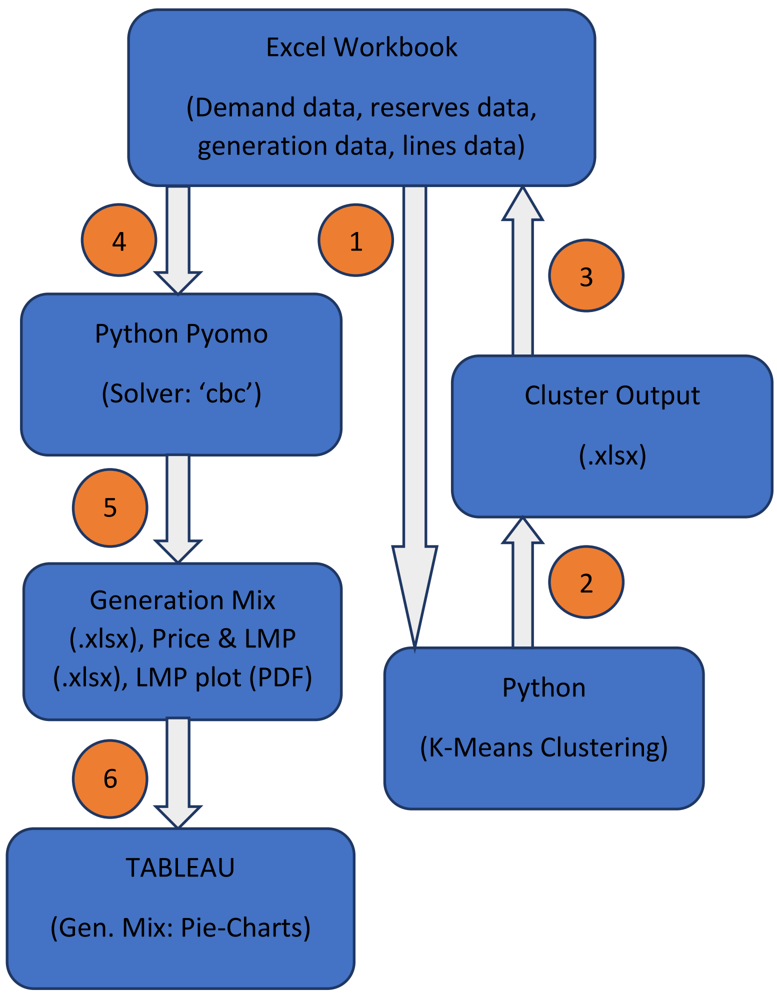


Figure 3: Data Flow diagram

2.3 Building of the model

2.3.1 Model description:

Decision Variables: $p_{g,t}$ be the power generated by generator g at time t
$flow_{l,t}$ be the power flow across line l at time t

Bounds: t ϵ {0, 1, 2, 3, … ,12}
t_1 ϵ {1, 2, 3, … ,12}

Objective function: $\sum_t^T \sum_g^G (p(g,t) * (marginalcost_g + CO_2cost_g))$

Constraints:

a. *Capacity Constraint:* $p_{g,t} \leq capacity_g * capacity\_factor_g \; \forall\, g \; \forall\, t$
b. *Nodal Power Balance:*
$\sum_g^G Generation_g * p_{g,t} + \sum_l^L Powerin_{l,n} * flow_{l,t} - \sum_l^L Powerout_{l,n} * flow_{l,t} = demand_{n,t} + reserves_{n,t} \; \forall\, n \; \forall\, t$
c. *KVL:* $\sum_l^L Line\_Reactance_l * flow_{l,t} = 0 \; \forall\, t$
d. *Power-Flow Constraints:* i) $flow_{l,t} \leq Line\_capacity_l \; \forall\, l \; \forall\, t$
   1. ii) $flow_{l,t} \geq - Line\_capacity_l \; \forall\, l \; \forall\, t$
e. *Ramping-Up Constraint:* $p_{g,t1} - p_{g,t1-1} \leq Ramp\_up_g \; \forall\, g \; \forall\, t_1$
f. *Ramping-Down Constraint:* $p_{g,t1-1} - p_{g,t1} \leq Ramp\_down_g \; \forall\, g \; \forall\, t_1$

2.3.2 Assumptions

The assumptions of the model include the following:

a) The line capacity was assumed to be 2500 MW and the reactance for each line to be 0.1 p.u
b) Having reserves was assumed to be a sufficient condition to maintain reliability.
c) The demand is apportioned per the number of plants in that node and reserves at a particular node were assumed to be 5% of demand in that node.
d) A 4-node power system would be ideal to economically dispatch the load.
e) Exogenous variables that influence the power output by renewable plants such as solar irradiance and inverter efficiency of solar plants, the volume flow rate for hydropower plants, etc. were assumed to be time-invariant.
f) The industrial averages for variables that we couldn't get the actual value of are being, to a large extent, to reflect the real-time values.
g) Transmission losses were ignored.
h) The CO_2 cost in $/MWh was computed assuming a social cost of carbon in the range of 0 $/ton to 100 $/ton with an option to increment by 10 $/ton.
i) Fuel costs of plants, per fuel type, are assumed constant.

j) All plants involved in the study are “committed” by the unit commitment model across the time period.

2.3.3 Limitations

Although the model built helps us to understand the generation mix, power flow along the lines, and the locational marginal prices throughout the time period while minimizing costs, carbon footprint, and increasing reliability; there are many limitations involved which are discussed below –

a) Since the model was built from public datasets and many assumptions in choosing values of certain variables, although carefully chosen from literature and reliable sources, they don’t reflect the real-time dispatches accurately. In short, this model serves as a generic prototype and should be fine-tuned with more accurate values for variables as and when needed.
b) The addition of exogenous parameters for renewable sources must be taken into account since these factors play a pivotal role in dispatching power from renewable plants at a set time period.
c) Transmission line losses are higher than distribution line losses and therefore, should be taken into account.
d) Although reserves help in making sure that a set % power is available to be dispatched during black-outs, to better understand how much capacity should be installed to meet the desired reliability target, a loss of load expectation (LOLE) analysis must be performed. Although it is not quite clear how to factor LOLE in our constraints, it does have an impact in planning for reserves and dispatching them when needed, and therefore must be taken into account.
e) Depending on the economics of the region, fuel costs tend to vary and should be modeled as a non-linear function to reflect real-time scenarios. For simplicity, our model assumes a constant fuel cost.
f) Our model does not include electrification, rooftop PV penetration, or energy efficiency to meet demand. We are operating on plants that are operated by Duke Energy and solar & biomass IPP in the Carolinas region without taking into account the aforementioned factors. In real-time, there is a growing shift to these demand-side response management systems, and therefore should be considered.
g) Only plants that are “committed” by the UC model are economically dispatched. In our analysis, at all time periods, we are considering the same plants being “committed”. This is quite a quixotic scenario because plants usually have a minimum uptime and downtime and can’t run continuously.

3. Results and discussion

Our analysis includes 6 nuclear plants, 26 Hydroelectric plants, 6 Coal-Fired plants, 15 Gas-Fired plants, 2 Pumped-Storage Hydro plants, and 915 Solar plants (most of which are IPP), 28 Biomass plants (IPP), and 1 Fuel-oil operated plant. The economic dispatch model was run for Dec 2, 2020.

All plants were dispatched from 7 AM to 6 PM. Solar Plants, however, from 7 PM to 6 AM, were set to produce no power. **Figure 4** shows the variation of LMP from 7 AM to 6 PM across all nodes. From the LMP plot (**Figure 4**), it was observed that for nodes 2, 3, and 4; the LMP increases from 7 AM to 9 AM and then starts to decrease at 10 AM, remains constant at 5 PM, and increases till 6 PM. The LMP for node 1 is constant at 11.96 $/MWh, which has the lowest LMP. Node 4 had the highest LMP at 8 AM & 9 AM with 82.62 $/MWh. While LMP reflects the values of electrical energy at different locations, load and generation mix, and transmission line capacity of the system, the trend of the LMP plot discussed above cannot be generalized. Higher the demand at a particular node, the higher the LMP (**Figure 2**). **Figure 5** shows the variation of LMP from 7 PM to 6 AM across all nodes. A similar trend was obtained for all nodes with node 4 having the highest LMP while node 1 had the lowest LMP (constant = 11.96 $/MWh). To get LMP for a different date, say Dec 5, 2020; the user has to feed in the demand and reserves data for Dec 5. Other sheets in the Excel Workbook needn't have to change, should assumptions remain the same. It was also observed that for the time period 7 AM to 6 PM, maintaining 5% reserves gave an infeasible solution due to very high demand at 8 AM. On the other hand, reliability could be maintained from 7 PM to 6 AM with 5% reserves. An interesting observation to make here is that we were able to maintain reliability without solar (7 PM to 6 AM) while we weren't able to have additional reserves from 7 AM to 6 PM despite having solar. Therefore, other energy sources such as storage and DERs should be considered.

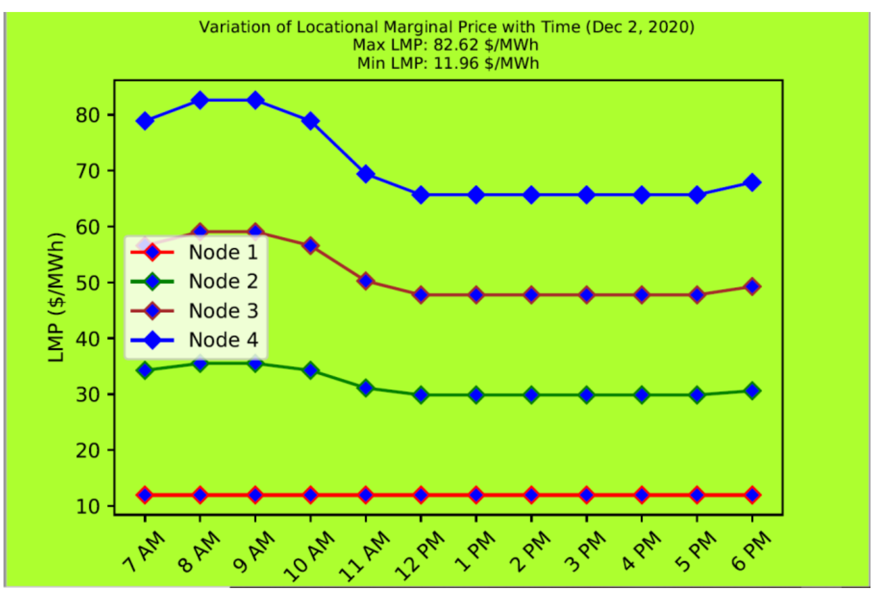


Figure 4: LMP variation with time (7 AM to 6 PM)

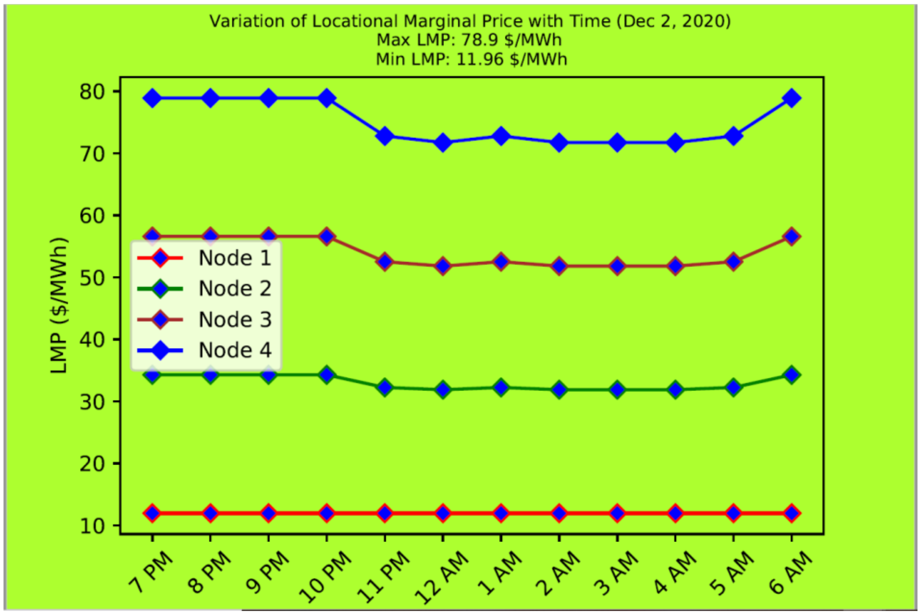
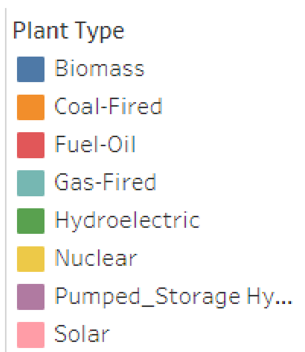
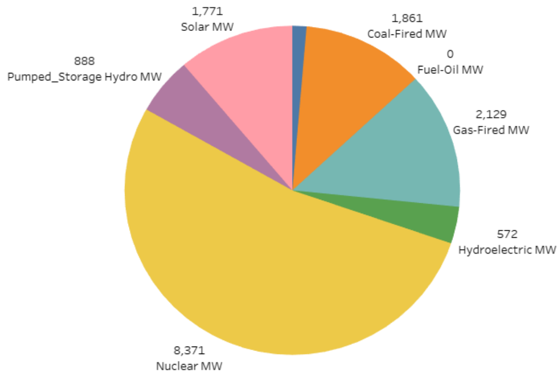


Figure 5: LMP variation with time (7 PM to 6 AM)

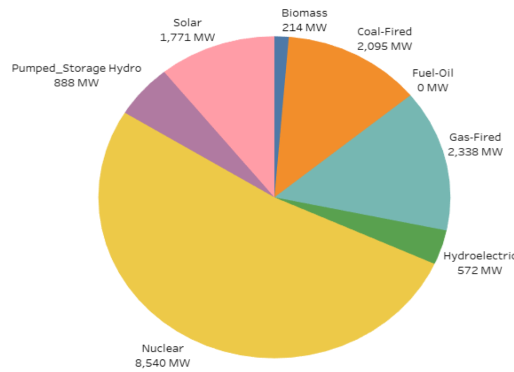
After obtaining the power flows, LMPs, and generation mix from python as output files, the generation mix output file was fed as input to TABLEAU for visualization. Figure 6 (a-l) shows the generation mix (in MW) of different fuel types from 7 AM to 6 PM while Figure 7 (a-l) shows the generation mix (in MW) of different fuel types from 7 PM to 6 AM. Of course, realistically, all these plants will not be dispatched at all times due to minimum up and downtime constraints in the UC model.

Legends for Figure 6 (a-l) & Figure 7 (a-l)

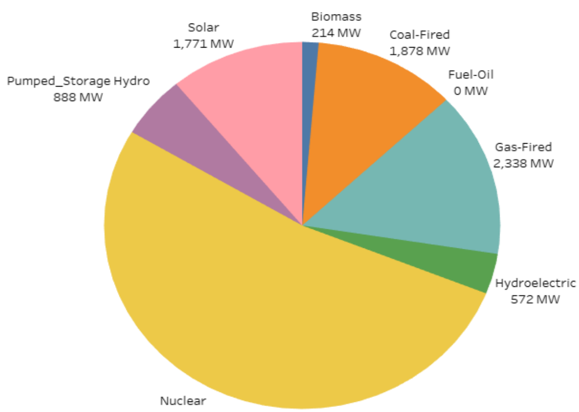




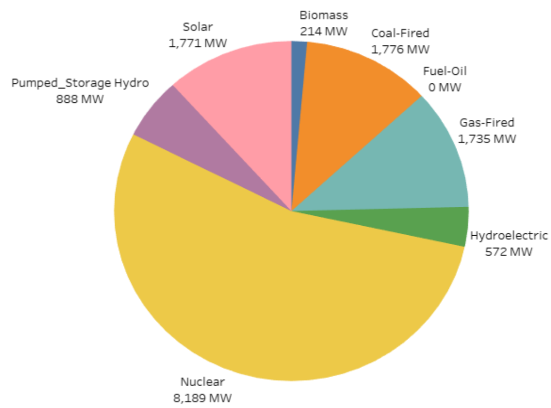
(a) 7 AM



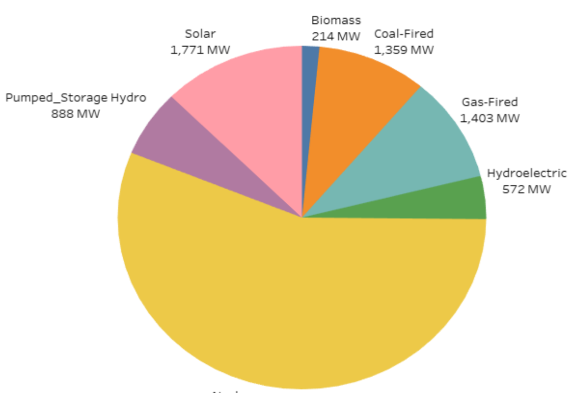
(b) 8 AM



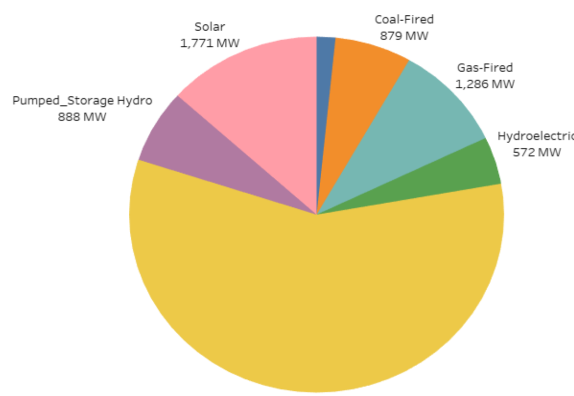
(c) 9 AM



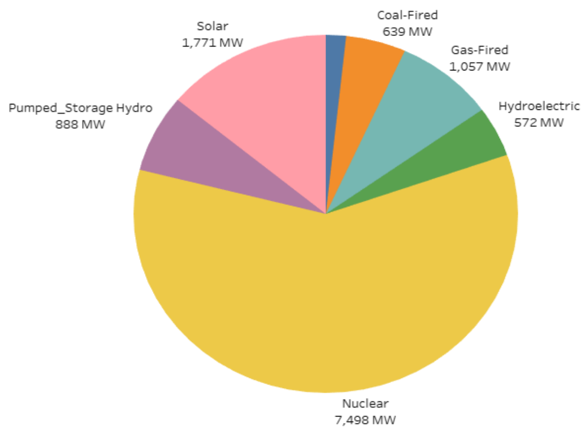
(d) 10 AM



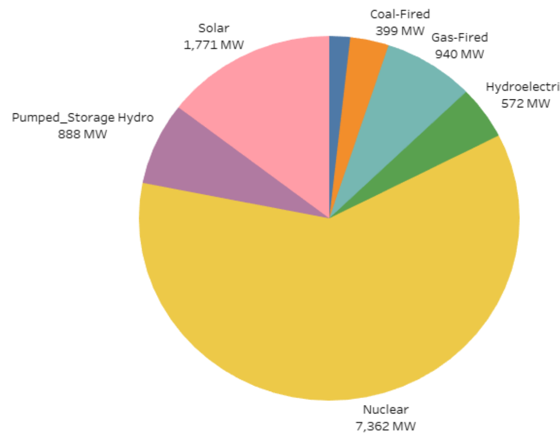
(e) 11 AM



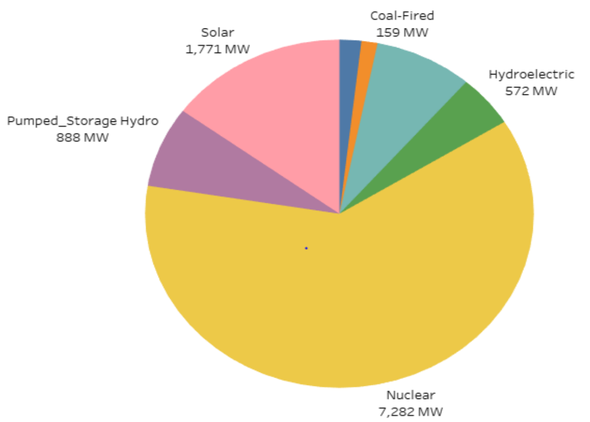
(f) 12 PM



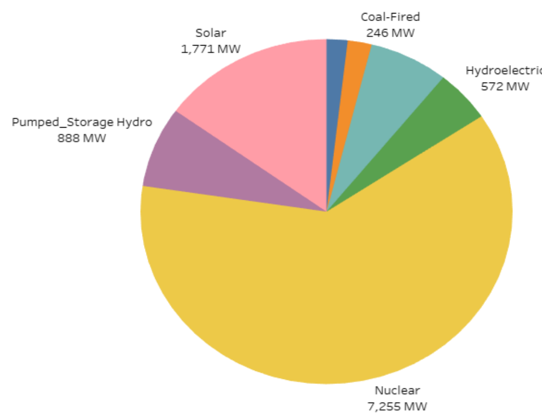
(g) 1 PM



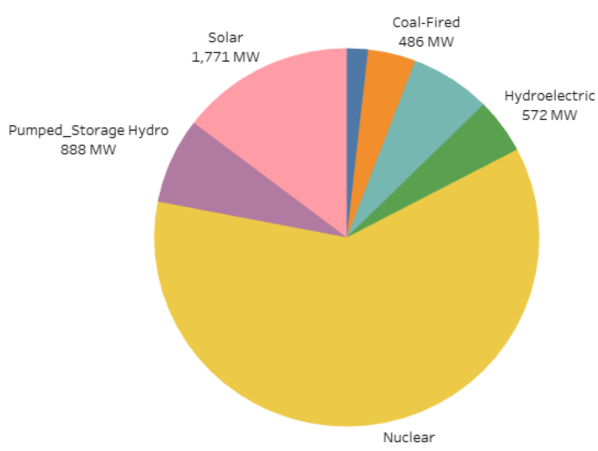
(h) 2 PM



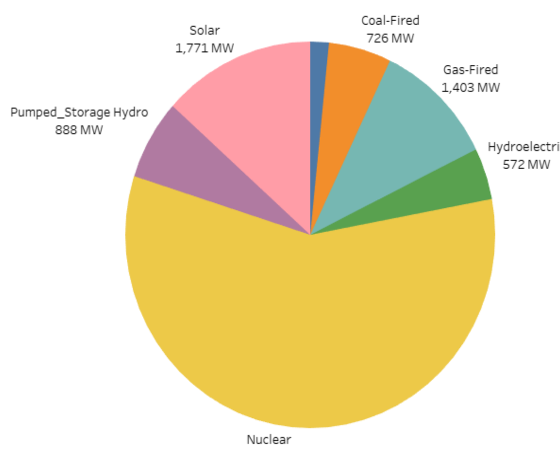
(i) 3 PM



(j) 4 PM

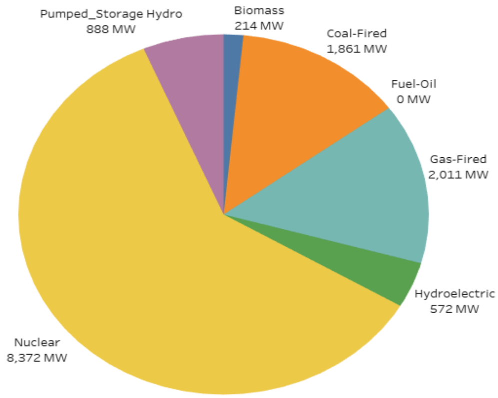


(k) 5 PM

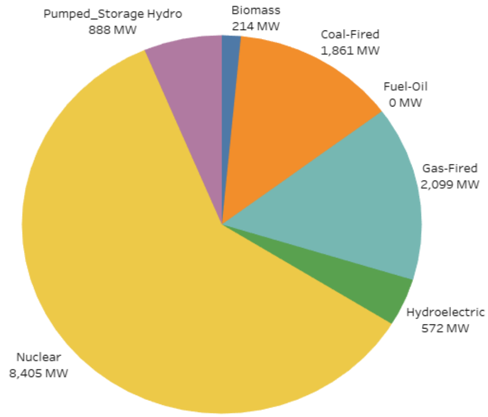


(l) 6 PM

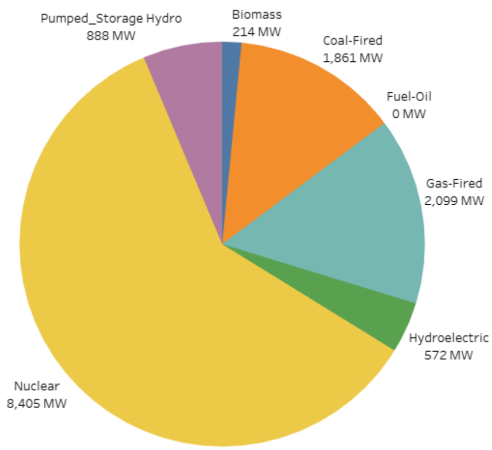
Figure 6: Generation mix from 7 AM to 6 PM



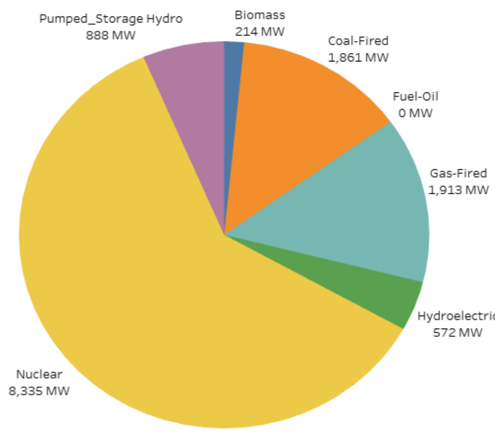
(a) 7 PM



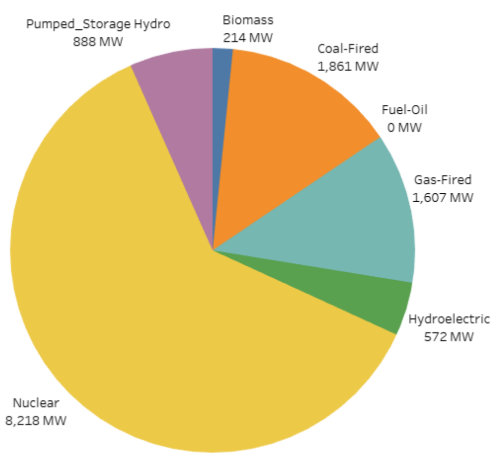
(b) 8 PM



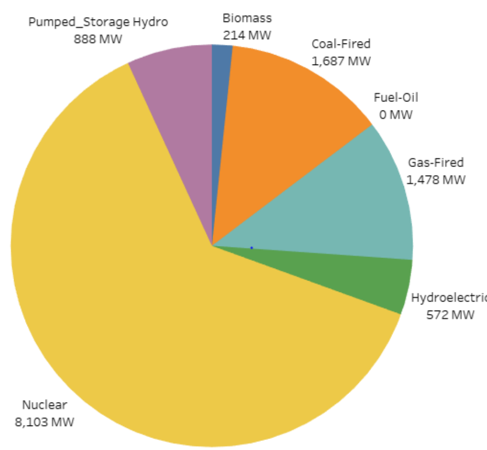
(c) 9 AM



(d) 10 PM



(e) 11 PM



(f) 12 AM

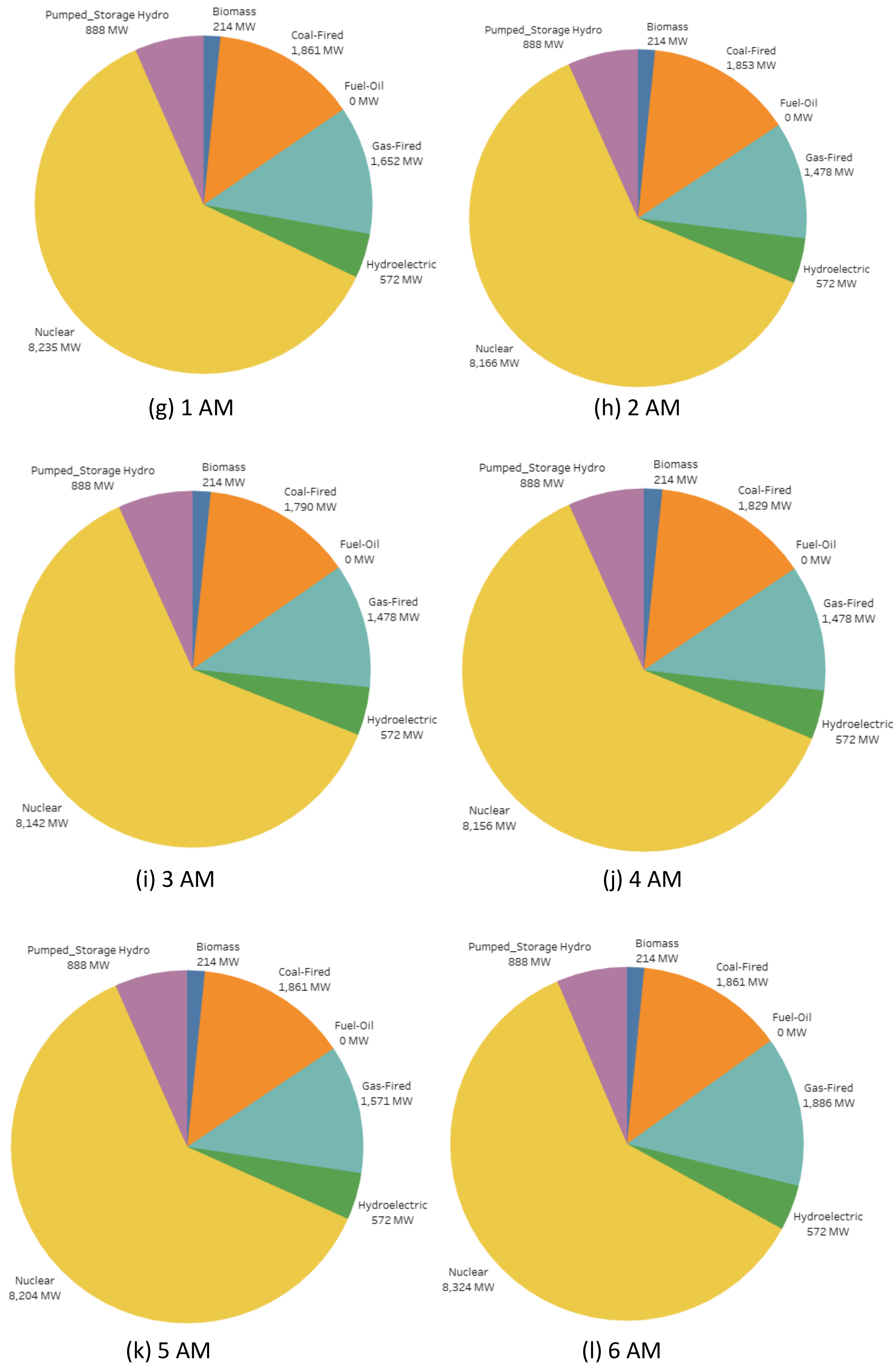


Figure 7: Generation mix from 7 PM to 6 AM

References:

Alqahtani, B. J., & Patiño-Echeverri, D. (2019). Combined effects of policies to increase energy efficiency and distributed solar generation: A case study of the Carolinas. *Energy Policy, 134*, 110936. doi: https://doi.org/10.1016/j.enpol.2019.110936

Chen, B., Guo, Q., Chen, Y., & Sun, H. (2020). An economic dispatch model for combined heat and power systems considering the characteristics of heat recovery steam generators. *International Journal of Electrical Power & Energy Systems, 118*, 105775. doi: https://doi.org/10.1016/j.ijepes.2019.105775

Cornelius, A., Bandyopadhyay, R., & Patiño-Echeverri, D. (2018). Assessing environmental, economic, and reliability impacts of flexible ramp products in MISO's electricity market. *Renewable and Sustainable Energy Reviews, 81*, 2291-2298. doi: https://doi.org/10.1016/j.rser.2017.06.037

Daraeepour, A., Patino-Echeverri, D., & Conejo, A. J. (2019). Economic and environmental implications of different approaches to hedge against wind production uncertainty in two-settlement electricity markets: A PJM case study. *Energy Economics, 80*, 336-354. doi: https://doi.org/10.1016/j.eneco.2019.01.015

Lin, Z., Chen, H., Wu, Q., Li, W., Li, M., & Ji, T. (2020). Mean-tracking model based stochastic economic dispatch for power systems with high penetration of wind power. *Energy, 193*, 116826. doi: https://doi.org/10.1016/j.energy.2019.116826

Shen, Z., Wei, W., Wu, L., Shafie-khah, M., & Catalão, J. P. S. (2021). Economic dispatch of power systems with LMP-dependent demands: A non-iterative MILP model. *Energy, 233*, 121015. doi: https://doi.org/10.1016/j.energy.2021.121015

Velasquez, M. A., Barreiro-Gomez, J., Quijano, N., Cadena, A. I., & Shahidehpour, M. (2019). Distributed model predictive control for economic dispatch of power systems with high penetration of renewable energy resources. *International Journal of Electrical Power & Energy Systems, 113*, 607-617. doi: https://doi.org/10.1016/j.ijepes.2019.05.044