HIGH SCHOOL EDITION
Journal of Student Research
Volume 13 Issue 2 (2024)

Using a Predictive Model to Reduce Emissions/Energy Costs with Virtual Power Plants in North India

Ansh Aggarwal[1] and Mauricio Hernandez[#]

[1]William Fremd High School, USA
[#]Advisor

ABSTRACT

This project aims to reduce the emissions and energy costs in a pollution riddled country like India by optimizing the energy output from power plants. The emissions/costs are minimized using Mixed Integer Linear Programming (MILP), an optimization tool that accounts for linear constraints and objective function. Utilizing the MILP results allows for the creation of Unit Commitment and Economic Dispatch (UC/ED). UC determines a schedule for which power plants should be on or off at which times while accounting for constraints such as startup/shutdown costs and ramping flexibility. Economic Dispatch optimizes power generation levels for each specific power plant while considering constraints such as meeting the energy demand and minimum/maximum generation limits to be found. Virtual Power Plants (VPPs) are the real-world application of our optimized UC/ED findings as multiple decentralized energy sources, power plants in this case, can aggregate resources and function as a singular plant. It can recalibrate energy usage based on data such as hourly power demand, weather, fuel source, etc., and apply UC/ED by using constantly updated real-time data. A Random Forest Regressor machine learning model which predicts the gains in terms of CO2 emissions of having VPP assets in the power grid was used. The accuracy at 70F, 73F, and 76F cooling points were 98.89%, 97.76% and 87.67% respectively. A machine learning model is used instead of the MILP model as it works much faster and can be feasibly used for day-to-day operations.

Introduction

Air Pollution is a world-wide problem which causes 7 million premature deaths each year (World Bank, 2023). According to IQAir in 2022, India was ranked 8th in the most polluted countries in the world (IQAir, 2022). Power plants account for 37% of GHG emissions in India, a major source of energy, but also a major source of pollution (Chateau, 2023). Most power plants don't run on renewable energy sources, in fact, most of them run on non-renewable energy sources. Harboring a majority of coal power plants and the largest population, India has a high pollution rate and costs from the hundreds of power plants due to the high demand. Reducing pollution and costs from power plants is a starting point in helping India to reduce its GHG emissions.

Virtual Power Plants (VPP), a niche technology, can be used to minimize the emissions and costs of power plants in India and across the world. A VPP is a concept where multiple decentralized energy sources, power plants in this case, aggregate resources and function as a singular plant (U.S. Department of Energy). VPPs optimize the consumption and output of energy sources by making schedules and power outputs frequently based on constantly changing daily data. Additionally, VPPs can decrease the reliance on non-renewable energy sources by aggregating and coordinating resources from multiple energy sources, including renewable energy, to have a more efficient output.

In terms of current work on VPPs, the United States has also been incorporating Virtual Power Plants in states like California, Hawaii, Massachusetts, Utah, Vermont, and more through companies like Sunrun and SunPower corp (Inside Climate News, 2023). They are also currently being implemented across the world in places like South Australia, Japan, and Europe by companies such as Statkraft, Shell-owned, Next Kraftwerke, and Tesla (Reuters, 2023).

Machine learning can be used to solve this problem by creating a schedule of optimal power output for power plants. This method is called Unit Commitment and Economic Dispatch (UC/ED) which can optimize the high emissions and costs in India. More specifically, UC helps determine a schedule for which power plants should be on or off at which times while accounting for several constraints such as startup/shutdown costs and ramping flexibility. ED allows for the optimal power generation levels for each specific power plant while also considering constraints, such as meeting the energy demand and minimum/maximum generation limits to be found (Watson, 2019).

This project is meant to extend the use of VPPs to a country that can greatly be benefitted by this technology, as India struggles with poverty and pollution, not to mention the extreme weather during summer and monsoon seasons. Focusing on helping the individual person first, the project targets the residential sector of northern India. By starting out in northern India, there is hope for future work that can apply VPPs to the entire country and to other sectors and cut the bills and emissions not only for residents, but for big industries as well. Emissions and costs will be minimized using MILP and the predictive accuracy will be measured from the VPP as implementation in real life will require real-time evaluations as data changes from day-to-day operations. The efficiency and practical use of VPPs are shown by the results of the machine learning model as it proves the real-time capability of VPPs to coordinate multiple power plants.

Literature Review

The Indian power sector is composed of mostly non-renewable energy sources with plants split into three categories: thermal, hydro, and nuclear. Due to the heavy use of non-renewable resources, pollution and system costs are high and need to be solved in an impoverished country like India. VPPs are the bridge to optimization and efficiency. The nameplate capacity of the plants was found through listing values from an interactive map of power plants in India (National Informatics Centre, Government of India, 2023).

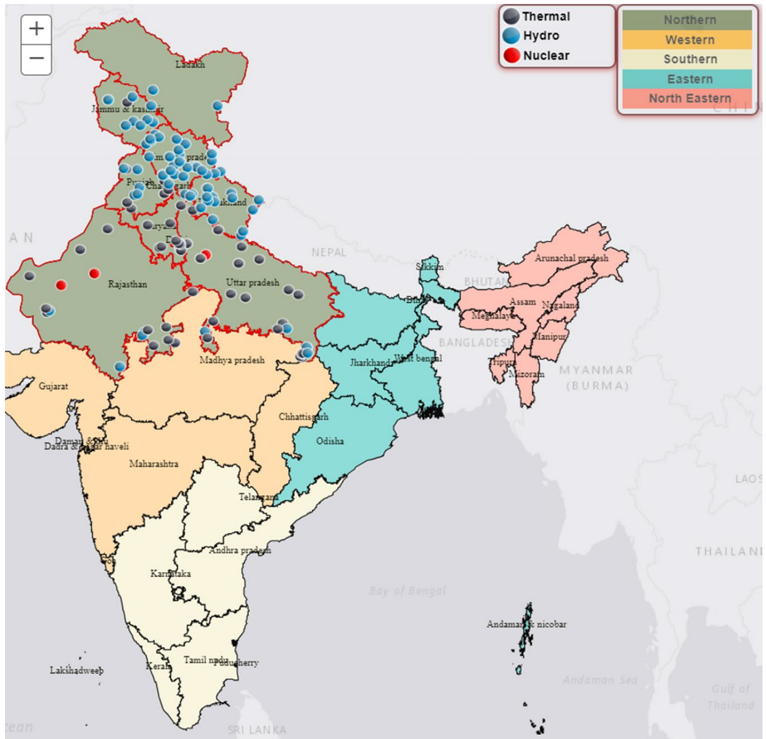

Figure 1. Location and nameplate capacity of power plants by type (Hydro, Thermal, Nuclear). This project only uses data from the Northern Region.

According to the Rocky Mountain Institute, VPPs enhance grid reliability. By 2030, VPPs are estimated to reduce peak demand in the United States by 60 gigawatts and could grow upwards to 200 gigawatts by 2050 (Brehm, et al., 2023). This ties in with VPPs ability to reduce emissions as less power is being released from the power plants. Not only that, but VPPs are said to reduce expenditures by 17 billion dollars in the power sector by 2030 in the United States (U.S. Department of Energy, 2023). India can apply the same principles and save money for millions of people and provide a better quality of life than they are used to in a highly polluted environment. India is currently planning on weaning off of non-renewable resources but is struggling due to a lack of infrastructure (Banerjee, 2023). VPPs are beneficial as they provide real time decisions based on the current supply/demand and power consumption/generation while at the same time keeping the grid reliable.

Flexibility Analysis of Thermal Generation for Renewable Energy Integration in India: This article shows the need for stability by quickly responding to changes in renewable energy with thermal power plants. As reported by (Chand, et al., 2020), the integration of renewable plants in India's National Grid will require that the power system responds quickly to changes in the electricity output.

A conceptual review on transformation of micro-grid to virtual power plant: Issues, modeling, solutions, and future prospects: VPPs can better renewable energy integration, supply/demand, and flexibility of the system using computer models. VPPs are supposed to provide benefits to the environment without sacrificing other necessary components such as stability. There are future prospects in the works to realize the full power of VPP and what it can do. This article highlights the benefits of VPPs in a theoretical sense but mentions that they need to be applied and used to see the extent of their full power (Panda, et al., 2022).

Optimization of energy consumption based on orientation and location of the building: Residential energy consumption in India is increasing and this situation needs to be addressed. By taking in several parameters about the buildings such as height, window material, roof material, etc., we can optimize the energy consumption in India and many other regions and buildings as well. Data given from these parameters can help provide VPPs with valuable information when trying to distribute power across the grid optimally and will lead to a more accurate optimization of costs and emissions for residents (Renuka, et al., 2022).

Comprehensive review of VPPs planning, operation and scheduling considering the uncertainties related to renewable energy sources: Distributed Energy Sources (DERs) have been driven by concerns with the environment with clean energy and filling the demand as well. However, VPPs allow for a smart system which can balance supply/demand, power consumption/generation, etc. VPPs also keep the grid stable with real time decisions and integrating renewable energy sources as well (Ullah et al., 2019).

Data

The data was found from different sources including Kaggle, known for having a variety of datasets, government websites, previous articles published, and technical reports. Kaggle provided data that was the basis of the project such as the power demand, power generation, and power consumption. (National Power Portal, 2017-2020), (POSOCO, 2019-2020)

A list of where all the datasets came from and their purpose is below: There are 129 power plants in northern India in 2023 which are all composed of either thermal, hydro, or nuclear power plants. Thermal power plants were either considered coal or biomass fueled.

1. Although data on emissions of power plants specific to northern India was not available, Consumer Ecology provided source emissions numbers for all fuel source type power plants (Consumer Ecology, 2022).
2. This project focuses specifically on residents in northern India; therefore, the demand is proportional. The constant used was 24.74% as the percentage of residents in northern India (Energy Statistics India, 2023).
3. Power plants take time to startup/shutdown, and each startup/shutdown is associated with some cost. Power plants also have a minimum on and off time as they can't be turned on again immediately. Additionally, causing major changes in power plant power generation at fast rates is not advised, therefore they are ramped up and down. Ramp rate is the rate per minute it takes to get to the maximum output. Data in India was not available for this factor, so ramp rates from the United States are used in the model. Nuclear power plants are on at all times and a ramp rate does not apply to these plants (Xu, et al., 2017).
4. Fuel prices are unique for each source (biomass, gas, coal, and nuclear), therefore different datasets were taken (Sokrethya, et al., 2023), (IEA, 2021), (ET Bureau, 2023), (World Nuclear Association, 2023). Since there was one biomass plant in all of northern India, the specific cost for that plant was found. The nuclear power plant prices are from the United States as Indian prices were unavailable. Hydro power plants do not have any fuel costs and nuclear plants fuel costs were too insignificant and were therefore made 0.
5. Heat rate is the amount of energy used by a power plant to generate 1 Btu (British Thermal Unit) per Kwh (Kilowatt Hour). In minimizing, heat rate is a factor that is multiplied with the fuel costs and power demand for each power plant fuel source type so total costs can be minimized using the machine learning algorithm (U.S. Energy Information Administration (EIA)).
6. Biomass heat rate is not available from this dataset and was alternatively found in another dataset (Wiltsee, 2000).
7. Data from India was not available, therefore power plant ramp heat rate from the United States is used. Since power plants in India are not as efficient as the ones in the United States, a constant is multiplied to the heat rates for all sources of power plants to generalize the data to northern India (IEA, 2009).
8. In order to make sure that units match when multiplying fuel costs, heat rate, and power demand in the minimizing model, an energy conversion factor is needed to convert the costs (based on fuel source) to $/MBtu (Hazel, et al., 2019).
9. For power plants to be viable and worth using, they need to produce a percentage minimum of their nameplate capacity. This percentage is different for all types of power plants (hydro, coal, gas, nuclear, and biomass) (Government of India Ministry of Power, 2023), (Lens, 2014), (World Nuclear Association,

2023), (Deaver, 2019). The value for coal was generalized to India from European data. Biomass data was not available, therefore a recognized baseline of 25% found from California power plants was used.

10. To minimize system costs, a tax of $35 per ton of CO2 is added to the model (Parry, 2019).

Daily data is not sufficient because an optimal and accurate schedule would require hourly data for the unit commitment to provide a day/night schedule. To get more accurate results and a better understanding of how daytime/nighttime may affect power demand, costs, and power plant schedule, BEOPT (Building Energy Optimization Tool) was used. BEOPT provides data based on a 3D design of a house from the users and parameters such as weather (EnergyPlus), roof materials, neighbor distance, window material, cooling point, and much more. Figure 2 shows the BEOPT model. To simulate northern India best, data from 5 different locations in northern India were averaged to get the most accurate result. The 5 locations were Allahabad, Jaipur, Lucknow, New Delhi, and Amritsar. These were the most populous cities in northern India with available weather data. BEOPT values were found at three different cooling setpoints (70F, 73F, and 76F) to see how thermostat temperatures affect the power demand throughout the year and model accuracy in predicting emissions/costs. After following parameters and the house design of a previously done model of India, data on the hourly kilowatt usage, outdoor bulb temperature, the dew point, etc., values were found and later manipulated in either the minimizing or predicting model. Figures 3 and 4 show examples of the hourly energy demand.

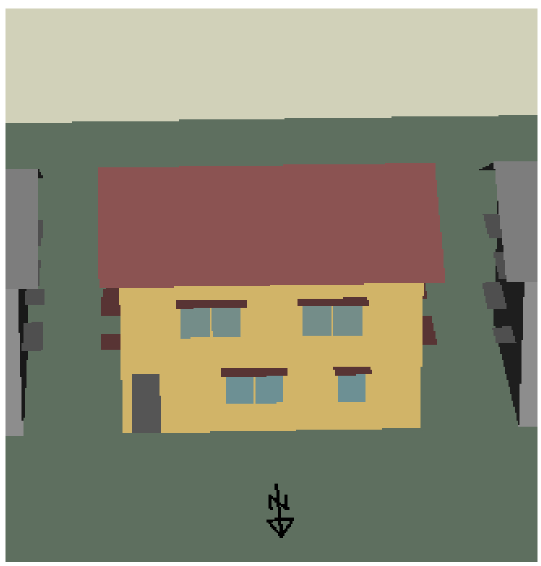

Figure 2. Model house created with the BEOPT tool to model Indian dwellings.

Although the datasets encompass several years of data in each dataset, focusing on 2019 was beneficial because it was common among all data found outside of Kaggle as well. Additionally, 2019 is before the Covid-19 pandemic, which could have heavily affected the data due to a large percentage of the population staying at home, which could have increased the power consumption/demand due to more appliances. For factors that are specific to the fuel source of a power plant, values were randomized for each power plant based on fuel source. The data described above is used to either minimize the emissions and costs or predict them.

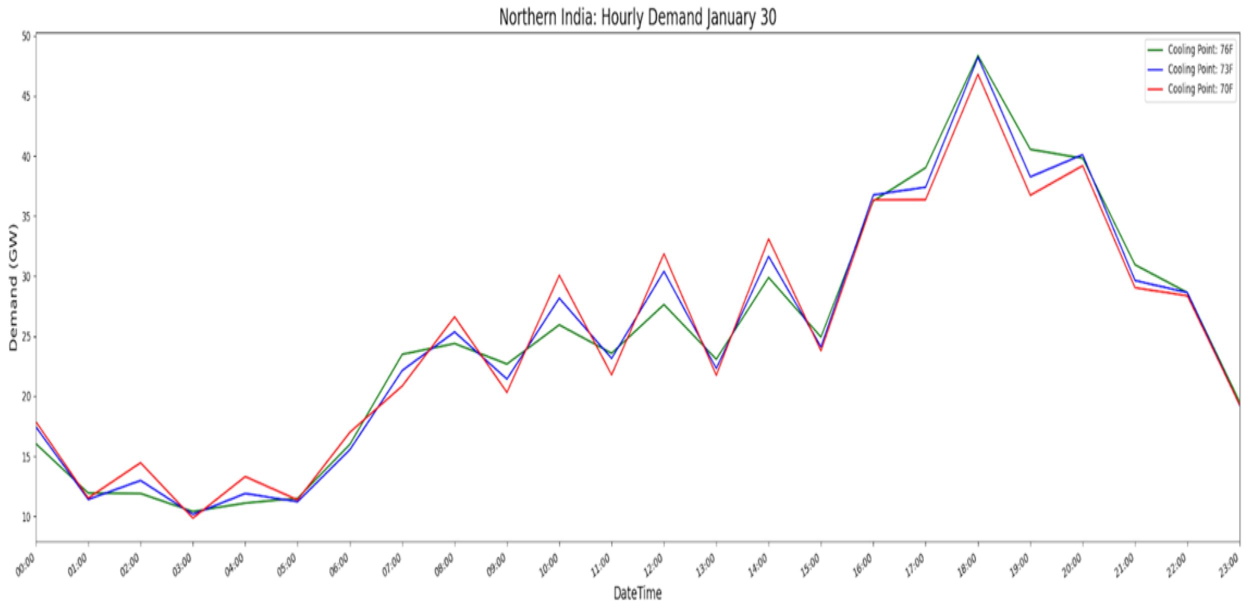

Figure 3. Hourly power demand of the Northern India Region, January 30th. The three lines signify the three different cooling points, showing us the difference in hourly demand based on the thermostat.

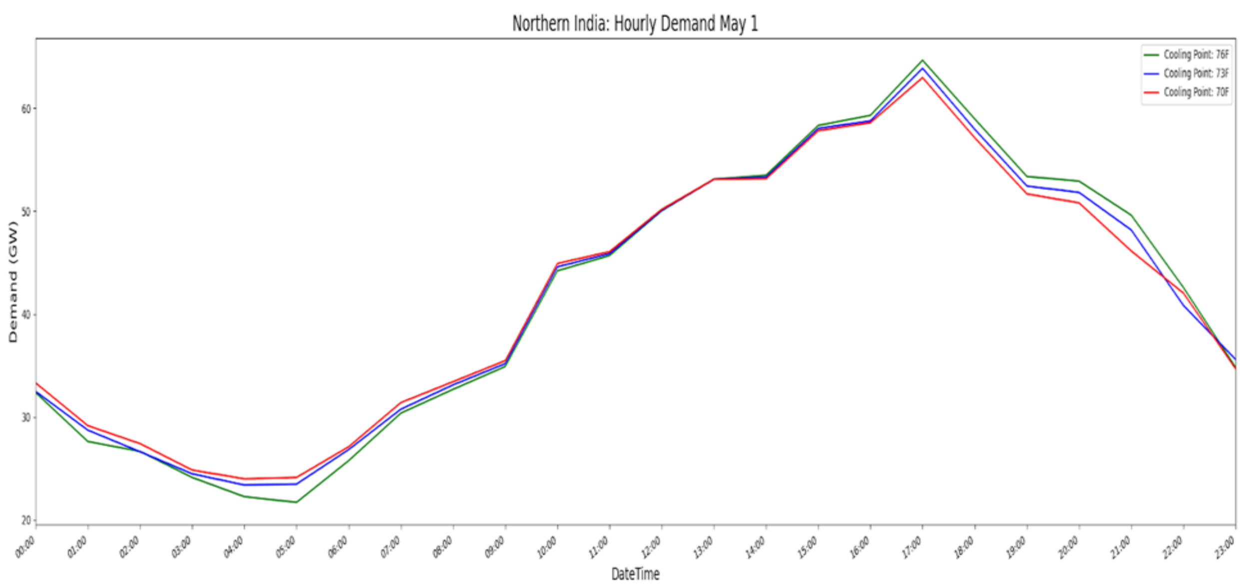

Figure 4. Hourly demand for May 1st, a summer day, when you might expect the AC to be turned more, increasing the demand.

HIGH SCHOOL EDITION
Journal of Student Research
Volume 13 Issue 2 (2024)

unit_nam	Region	typeof_fu	node[stat	nameplat	Breakdow	PERC_RAN	STARTUP_	SHUTDN_(	STARTUP_	SHUTDN_1	MIN_ON_1	MIN_OFF_	PERC_GRI	SOURC_EN	WEIGHT_E	SOURC_EN	WEIGHT_E	HEAT_RAT	FUEL_COSTS
Nimmo BA	Northern	Hydro	Ladakh	45	Hydro: 18	22.43	1.24	0.32	0.33	0.87	0.79	0.33	10.4	20.5	2.13	182	18.9	0	0
CHUTAK H	Northern	Hydro	Ladakh	44	Nuclear: 1	16.28	2.42	0.25	0.69	0.35	0.75	0.16	10.4	20.5	2.13	182	18.9	0	0
Kishangar	Northern	Hydro	Jammu &	330	Coal/Ligni	17.71	1.33	0.4	0.55	0.22	0.19	0	10.4	20.5	2.13	182	18.9	0	0
Pampore	Northern	Gas	Jammu &	175	Natural Ga	9	122.69	3.29	2.09	1.09	0.98	0.34	4.2	486	20.6	845	35.8	8.41	7.43
Upper Sin	Northern	Hydro	Jammu &	105	Biomass:	17.32	3.59	0.07	0.21	0.63	0.79	0.69	10.4	20.5	2.13	182	18.9	0	0
Uri-I HPS	Northern	Hydro	Jammu &	480	Total: 668	21.19	2.8	0.22	0.72	0.86	0.13	0.04	10.4	20.5	2.13	182	18.9	0	0
Salal HPS	Northern	Hydro	Jammu &	690		15.18	4.63	0.04	0.26	0.9	0.58	0.07	10.4	20.5	2.13	182	18.9	0	0
Baglihar H	Northern	Hydro	Jammu &	450		18.73	0.7	0.4	0.97	0.18	0.51	0.56	10.4	20.5	2.13	182	18.9	0	0
Baglihar-II	Northern	Hydro	Jammu &	450		23.4	3	0.45	0.33	0.74	0.39	0.74	10.4	20.5	2.13	182	18.9	0	0
Dulhasti H	Northern	Hydro	Jammu &	390		22.4	3.62	0.07	0.17	0.74	0.9	0.6	10.4	20.5	2.13	182	18.9	0	0
Lower Jhe	Northern	Hydro	Jammu &	105		17.15	3.98	0.15	0.23	0.18	0.75	0.43	10.4	20.5	2.13	182	18.9	0	0
Sewa-II HI	Northern	Hydro	Jammu &	120		23.04	4.26	0.35	0.52	0.16	0.62	0.14	10.4	20.5	2.13	182	18.9	0	0
Mukerian-	Northern	Hydro	Punjab	58.5		15.2	4.98	0.31	0.93	0.47	0.31	0	10.4	20.5	2.13	182	18.9	0	0
Goindwal	Northern	Coal	Punjab	540		6.36	132	14.24	7.86	3.14	2.85	11.56	72.5	863	625.32	4839	3506.3	13.19	1.3

Figure 5. This partial table shows the data points taken from the several datasets and clean/filtered into one. This shows the randomized values for each data factor and its consistency with the power plant fuel source. Note: All columns of data are shown

Methods

The goal of the project is to integrate a smart system into India to optimize the energy to minimize cost and CO2 emissions using a Virtual Power Plant (VPP). The energy assets include wind turbines, hydro energy, solar panels, etc. The VPP optimizes the consumption and output of energy sources. From the 129 power plants in Northern India, ones in the same vicinity can harness a VPP to minimize the emissions and costs which will make a large impact in pollution riddled India that is a leader in greenhouse gas emissions.

The hourly energy consumption of the 5 sites was estimated with the BEOPT tool and parameters followed by the paper listed in the Literature Review. Then, the hourly demand data was adjusted from the BEopt Tool with the daily demand data of Northern India. This allowed for a visualization of the change in demand in a day/night schedule and a more accurate result given the many more data points for the machine learning model to learn from for its prediction. The data were fitted using a normalization factor. Based on a day with hourly data, the demand percentage of each hour relative to a specific day would be the normalization factor for that hour. An example of the formula for one hour would be $\left(\frac{hour\ \#1\ energy\ demand}{sum\ of\ 24\ hour\ energy\ demand}\right)$. This process was repeated for the entire year. By multiplying the normalization factor for each hour by the daily demand data, the hourly demand for Northern India was estimated. This was the process for the base case, 76F cooling point. The other cooling points are relative to the base case. To find the hourly demand data for the 73F and 70F cooling points, the formula followed for hour 1 was:

$$\left(\frac{Total\ Energy\ 73F}{Total\ Energy\ 76F}\right) * \ Hour\ 1\ Demand\ for\ 73\ F \text{ and}$$

$$\left(\frac{Total\ Energy\ 70F}{Total\ Energy\ 76F}\right) * \ Hour\ 1\ Demand\ for\ 70\ F.$$

This process was repeated for both cooling points and all 8760 hours in a year. Figure 6 shows the shift in demand as the year progresses in the summer months and winter months.

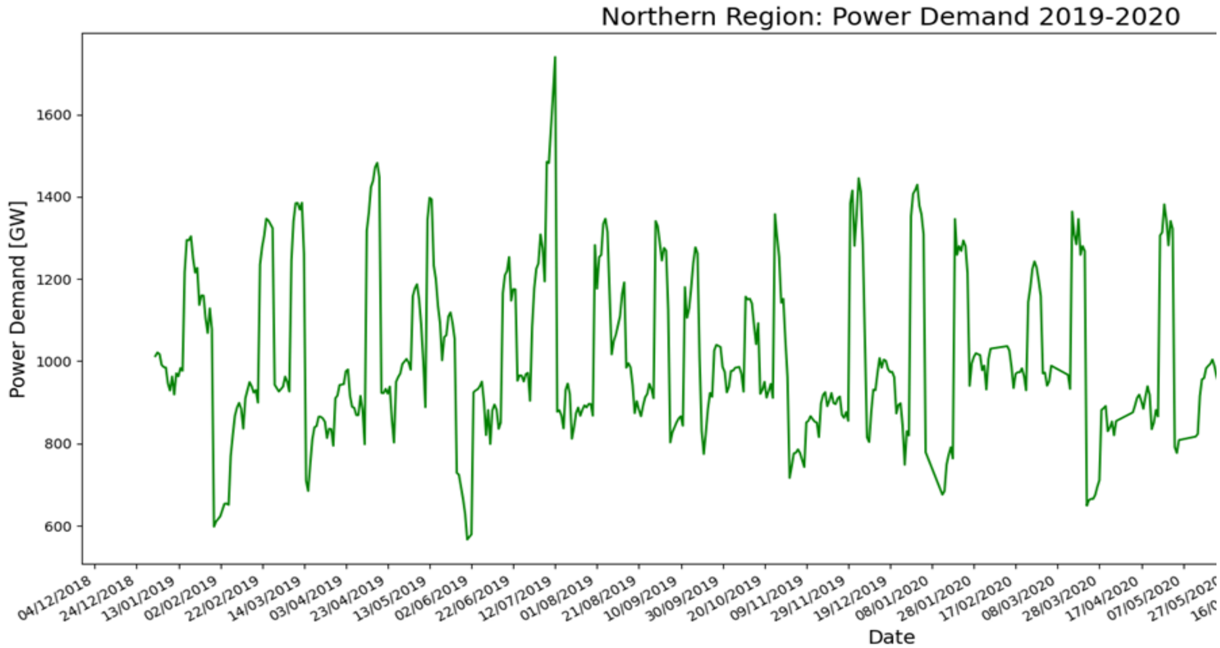

Figure 6. Daily power demand of Northern India. We can see the shift in demand as the year progresses in the summer months and the winter months. Power demand for a day is shown at intervals of 20 days.

Mixed Integer Linear Programming (MILP) is an optimization method using linear constraints and linear objectives. Essentially, MILP finds points on a graph that fit an objective function based on constraints created with the inputs of the function. The MILP model used for this project is meant to minimize the emissions and costs of the power plants and create a schedule/power output for the power plants to meet the demand also known as Unit Commitment and Economic Dispatch.

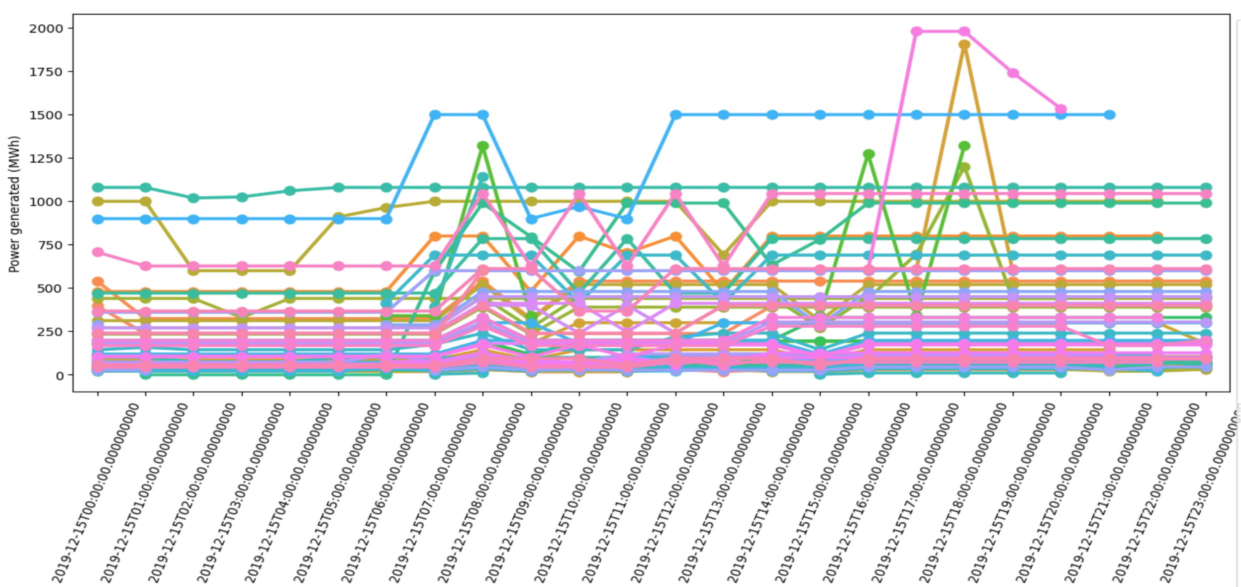

Figure 7. Power dispatch of generators that resulted from the Unit Commitment model. The graph is from Dec 15th. VPPs can do the same and make real-time decisions when making a power output schedule for hundreds of power plants to meet the demand and keep the grid stable.

The Objective Function formula which minimizes the costs/emissions of the power plants is below. All of the constraints added are real constraints applied in power plant facilities and allow for realistic results that may occur in real-life.

$$\sum_{plant\ i}\sum_{hour\ h}\left(f_i\, z_{i,h}\, hr_i \ +\ .035se_i z_{i,h}\ +\ s_i v_{i,h} + t_i w_{i,h}\right)$$

i - power plant
h - hour
f - fuel costs
z - power generated by a plant
hr - heat rate
CO2 emissions tax (.035) - dollar tax amount per kg of CO2 from carbon footprint emissions
se - source emissions
s - startup cost
v -startup the plant (1 if plant i is shut down, 0 otherwise)
t - shutdown cost
w - shutdown the plant (1 if plant i is shut down, 0 otherwise)
m - minimum generations limit
c - maximum power capacity
r - ramp rate percentage
u – plant on or off (1 if plant I is on, 0 otherwise)
d – energy demand
U - minimum on time
D - minimum off time

Constraints seem to be an obvious part of the model, but they need to be coded into the model for the results to accommodate a real-life environment. The constraints for the UC model are listed below:

1. Meet Demand: Make sure that the total power generated by all the power plants in each hour is equal to the power demand in that hour.

$$\sum_{plant\ i} z_{i,h} = d_h$$

2. Maximum and Minimum Generation Levels: Once a power plant is on, it must produce a minimum amount of energy before being turned off for it to be worth the startup. The power plant also cannot produce more power than its capacity.

$$z_{i,h} \le c_i u_{i,h}$$
$$z_{i,h} \ge m_i c_i u_{i,h}$$

3. Nuclear power plants are always on: Turning off nuclear power plants is tough as they take a long time (upwards to 2 days), therefore they are always on in this model.

$$z_{i,h} \geq m_i c_i$$

4. Maximum Ramp Rates: Immediate big changes in power generation for power plants is not recommended, instead, they are ramped up and down when on.

$$-r_i c_i \leq z_{i,h} - z_{i,h-1} \leq r_i c_i$$

5. Link startup/shutdown variables: This restraint gives three outcomes for a power plant, either it is on, off, or on and off. There won't be any overlap in the program.

$$v_{i,h} - w_{i,h} = u_{i,h} - u_{i,h-1}$$

6. Minimum on and off time: Power plants must stay on for a minimum amount of time before being turned on again and vice versa. The variables t in this scenario represents the time when the plant is turned on and T is the total time the plant can operate. Plants cannot be turned on or off immediately in succession, hence the constraint.

a. Minimum On Time

$$\sum_{t=1}^{min\{T,T_i{}^U\}} u_i(t) - min\{T, T_i{}^U\} = 0 \ \forall \ i \in G_{thermal}$$

$$\sum_{t=u_{i,j}}^{t+T_i{}^U,t} u_i(t) \geq T_i{}^U v_i(t), T_i{}^U = min\{T - t + 1, V_i{}^U\} \ \forall \ t = min\{T, T_i{}^U\} + 1, \dots, T$$

b. Minimum Off time

$$\sum_{t=1}^{min\{T,T_i{}^D\}} u_i(t) = 0 \ \forall \ i \in G_{thermal}, t \in T$$

$$u_g(t) + \sum_{j=max\{1,t-V_i{}^D\}}^{t+T_i{}^U,t} w_i(j) \leq 1 \ \forall \ t = min\{T, T_g{}^D\} + 1, \dots, T$$

To minimize the emissions as well, a tax on CO2 emissions is added. The tax is .035 dollars per kilogram of CO2. Effectively, emissions are now part of the objective and are minimized as well. It is easy to see the minimized emissions in C02 as it is just multiplied by a constant factor. All the constraints added are real constraints applied in power plant facilities and allow for realistic results that may occur in real-life.

The Machine Learning model is meant to predict the gains in terms of CO2 emissions of having VPP assets in the power grid of North India. There is an X matrix which contains features from the MILP model such as the hourly demand, nameplate capacity, heat rate, fuel costs, and source emissions, but it also contains unused BEOPT data points such as outdoor dry bulb temperature, outdoor dew point, and outdoor humidity ratio. The purpose of not adding all the data points is that the machine learning model is much faster than the MILP model and so that in a real-life scenario, realistic and attainable data can be used to coordinate resources instead of a lot of tough to find data points like ramp rates for each power plant. The Y matrix contains the minimized values obtained from the MILP model such as the minimized system costs and CO2 costs. The Y matrix also contains the generation of each source, for example, how much energy was produced from power plants fueled by coal. Both X and Y matrices contain data for the months January, May, June, and December so that there is an adequate amount of data for the machine learning model to train from (data for each cooling point 70F, 73F, 76F) and resemble the entire year. The goal is to see how well the machine learning model can predict the data in Y from the data in X. Specifically in this project, the focus is on predicting the CO2 costs as the project minimizes the costs and emissions and this variable provides information on both.

The machine learning model used for predictions was a Random Forest Regressor model as it proved to be more accurate than a neural network, a linear regression model, and a gradient boosting regressor model. Random forests use a collection of decision trees to make predictions.

Results

There are two sets of results: the minimized results from the MILP model used as the Y matrix in the machine learning model and the machine learning model accuracy.

Date	system_cost_70	CO2_costs_70	startup_cost_70	shutdown_cost_70	Coal_generation_70	NG_generation_70	Nuclear_generation_70	Biomass_Generation_70	Hydro_Generation_70
1/1/2019	5048325.624	3170756.506	2496.66	280.89	94433.6914	0	38880	117.5	418791.09
1/2/2019	3984855.561	2561109.776	2393.76	288.88	74257.71598	7.00E+02	38880	62.5	402022.5804
1/3/2019	4930318.514	3187960.095	2521.23	305.32	94112.94477	1572.695301	38880	62.5	419125.9945
1/4/2019	4364269.371	2886076.697	2398.82	143.72	84826.53506	620.4178387	38880	72.5	411868.359
1/5/2019	3882769.811	2519133.191	1961.09	226.1	72781.02942	6.12E+02	38880	112.5	407635.3987
1/6/2019	3599365.202	2329530.058	2258.24	435.33	66626.13148	5.82E+02	38880	122.5	403159.2408
1/7/2019	4146100.31	2721113.351	2296.21	281.03	79738.13937	174.0155008	38880	87.5	406708.2894
1/8/2019	3068138.754	2019536.657	1852.7	225.36	56840.08614	116.0425115	38880	15	394411.5287
1/9/2019	2955406.075	1898994.913	1746.96	239.52	52676.20758	99.23166242	38880	100	401881.1603
1/10/2019	3230248.869	2033971.976	2275.69	281.27	56677.39148	971.7524203	38880	62.5	400972.4142
1/11/2019	2404330.48	1531782.125	1907.84	253.87	40853.54056	0	38880	102.5	390137.9541
1/12/2019	2892313.377	1874198.465	1750.28	216.45	52071.0118	0	38880	0	395405.0928
1/13/2019	2696247.661	1715357.787	1891.42	242.88	46815.546	0	38880	122.5	394955.9157
1/14/2019	3043519.132	1948336.466	1949.07	254.54	54413.27676	1.59E-12	38880	100	399876.0156
1/15/2019	3778329.515	2397198.977	2204.05	282.26	69390.85003	222.0490691	38880	12.5	389906.8262

Figure 8. In the partial table above, the results from the MILP model are shown with the optimized values for the costs and the power plant generation for each type of fuel.

Coal and Hydro generation is likely the highest in terms of power generation as most power plants in India are coal and hydro powered. The same idea applies for biomass as there is only one biomass plant in northern India, making it the lowest source generation. Nuclear power plants' values are likely the same because they follow special conditions of always being turned on, therefore there is no difference between their production compared to a coal power plant that may be on at different times as other coal plants.

Table 1. In the first code segment, we see a 98.89% accuracy at a 70F cooling point. In the second code segment, we see a 97.76% accuracy at a 73F cooing point. In the third code segment, we see a 87.67% accuracy at a 76F cooling point

	Model Accuracy
Cooling Point: 70F	98.89%
Cooling Point: 73F	97.76%

Cooling Point: 76F	87.67%

The model uses a standard process of using 20% of data for testing and the other 80% for training. The machine learning model provided high accuracies for all three cooling points, however there is a dip in accuracy at 76F compared to 73F and 70F. This could be due to a variety of reasons. One of them could be that the BEOPT tool didn't provide accurate enough data for the model to predict the minimized emissions/costs, or another could be that the MILP model gave an irregular optimization compared to the other cooling points and didn't match the same pattern of data as well. However, the accuracy is still relatively high based on the amount of data fed to the machine learning model, and it will increase with more and more data.

Conclusion and Future Work

From our results, a Virtual Power Plant (VPP) is a valid technology that can be applied in northern India, and in the future, across the entire country or even the world. The machine learning model has a high accuracy for all 3 cooling points and is exceptionally good for 70F and 73F. A VPP can make real-time decisions from collecting recent data from just days before, providing a better measure for reducing emissions and costs. It can aggregate and redistribute resources most effectively from there. In the future, additional factors can be added to the model to get a more accurate minimization. Operating costs such as labor or health costs could be factored into the model in the future. Data specific to each power plant could be obtained to get completely accurate numbers instead of randomizations of ranges for specific fuel types. Lastly, more data will make the machine learning model more and more accurate until it reaches perfection. This type of model in the future can be applied to countries all around the world, improving quality of life in third-world countries all the way to the number 1 ranked economy.

Pollution is slowly eating away at the planet and this is only a minor stepping stone of many tasks that need to be done to heal our world.

Acknowledgments

I would like to thank my advisor for the valuable insight provided to me on this topic.

References

Banerjee, B (2023). *The Future of Virtual Power Plants in India - A Perspective*. Https://Economictimes.Indiatimes.com/. Retrieved January 21, 2024, from https://energy.economictimes.indiatimes.com/news/power/the-future-of-virtual-power-plants-in-india-a-perspective/99880940

Brehm, K., McEvoy, A., Usry, C., & Dyson, M. (2023). *Virtual Power Plants, Real Benefits*. Https://Rmi.org/. Retrieved July 15, 2023, from https://rmi.org/insight/virtual-power-plants-real-benefits/

Centre for Science and Environment (2023). *Energy Statistics India - 2023*. Http://www.Indiaenvironmentportal.org.in/. Retrieved July 15, 2023, from http://www.indiaenvironmentportal.org.in/files/file/EnergyStatisticsIndia2023.pdf

Chand, L. Sharath, et al. (2020). Flexibility Analysis of Thermal Generation for Renewable Energy Integration in India. *IEEE*. https://doi.org/10.1109/npsc49263.2020.9331830

Chateau, J., et al. (2023). A Framework for Climate Change Mitigation in India. *IMF Working Papers*. https://www.imf.org/en/Publications/WP/Issues/2023/10/20/A-Framework-for-Climate-Change-Mitigation-in-India-535854

Consumer Ecology (n.d.). *India Electricity Carbon Footprint & Environmental Impact*. Https://Consumerecology.com/. Retrieved June 15, 2023, from https://consumerecology.com/india-electricity-carbon-footprint-environmental-impact/

Deaver, P. (2019, January). *Updating Thermal Power Plant Efficiency Measures and Operational Characteristics for Production Cost Modeling*. Https://www.Energy.ca.gov/. Retrieved July 10, 2023, from https://www.energy.ca.gov/sites/default/files/2021-06/CEC-200-2019-001.pdf

EnergyPlus (n.d.). *Weather Data*. Https://Energyplus.net/. Retrieved July 10, 2023, from https://energyplus.net/weather

ET Bureau (2023, May 31). *Coal India raises high grade non-coking coal prices by 8%*. Https://Economictimes.Indiatimes.com/. Retrieved June 10, 2023, from https://economictimes.indiatimes.com/industry/indl-goods/svs/metals-mining/coal-india-raises-high-grade-non-coking-coal-prices-by-8/articleshow/100631376.cms?from=mdr

Government of India Ministry of Power (2023, September 7). *FAQs on Hydropower*. Https://Powermin.gov.in/. Retrieved July 20, 2023, from https://powermin.gov.in/en/content/faqs-hydropower

Gurobi Optimization (2023, February 20). Optimizing a Power Generation Schedule. *Gurobi Optimization*. https://www.gurobi.com/jupyter_models/optimizing-a-power-generation-schedule/

Hazel, D., & Bardon, R. (2019, March 15). *Conversion Factors for Bioenergy*. Https://Content.ces.Ncsu.edu/. Retrieved July 10, 2023, from https://content.ces.ncsu.edu/conversion-factors-for-bioenergy

IEA (2009). Energy Efficiency Indicators for Public Electricity Production from Fossil Fuels. *OECD Publishing, Paris*. https://doi.org/10.1787/9789264061996-en

IEA (2021). *Fuels and electricity in India*. Https://www.Iea.org/. Retrieved July 15, 2023, from https://www.iea.org/reports/india-energy-outlook-2021/fuels-and-electricity-in-india

Inside Climate News (2023, June 22). *Virtual Power Plants Are Coming to Save the Grid, Sooner Than You Might Think*. Https://Insideclimatenews.org/. Retrieved July 12, 2023, from https://insideclimatenews.org/news/22062023/inside-clean-energy-virtual-power-plants/

IQAir (n.d.). *World's most polluted countries & regions*. Https://www.Iqair.com/. Retrieved July 11, 2023, from https://www.iqair.com/us/world-most-polluted-countries

Lens, H. (2014). *Mid-Load Operation of Large Coal-Fired Power Plants*. Https://www.Steag.com.br/. Retrieved July 10, 2023, from https://www.steag.com.br/uploads/pics/Power-Gen_Europe_2014_-_Mid_load_operation_of_large_coal-fired_power_plants_06.pdf

National Informatics Centre, Government of India (2023). *All India Status (Category-wise)*. Https://web.Archive.org/. Retrieved July 15, 2023, from https://web.archive.org/web/20230407033716/https:/npp.gov.in/dashBoard/cp-map-dashboard

National Power Portal (2020). *Daily Power Generation in India (2017-2020)*. Https://www.Kaggle.com/. Retrieved June 14, 2023, from https://www.kaggle.com/datasets/navinmundhra/daily-power-generation-in-india-20172020

Panda, Subhasis, et al. (2022). A Conceptual Review on Transformation of Micro-grid to Virtual Power Plant: Issues, Modeling, Solutions, and Future Prospects. *International Journal of Energy Research*, *46*(6), 7021–54. https://doi.org/10.1002/er.7671

Parry, Ian (2019). The Case for Carbon Taxation. *IMF F&D*.https://www.imf.org/en/Publications/fandd/issues/2019/12/the-case-for-carbon-taxation-and-putting-a-price-on-pollution-parry

The World Bank (2023). *Pollution*. Https://www.worldbank.org/. Retrieved July 15, 2023, from https://www.worldbank.org/en/topic/pollution#:~:text=Pollution%20is%20the%20largest%20environmental

Power System Operation Corporation Limited (POSOCO) (2020). *Power consumption in India (2019-2020)*. Https://www.Kaggle.com/. Retrieved June 20, 2023, from https://www.kaggle.com/datasets/twinkle0705/state-wise-power-consumption-in-india

Renuka, S. M., et al. (2022). Optimization of Energy Consumption Based on Orientation and Location of the Building. *Materials Today: Proceedings*, *65*(2), 527-536. https://doi.org/10.1016/j.matpr.2022.03.081

Reuters (2023, January 31). *Explainer: What is a virtual power plant?* Https://www.Reuters.com/. Retrieved June 15, 2023, from https://www.reuters.com/business/sustainable-business/what-is-virtual-power-plant-2023-01-31/

Sokrethya, S., Aminov, Z., Van Quan, N., & Xuan, T. D. (2023). Feasibility of 10 MW Biomass-Fired Power Plant Used Rice Straw in Cambodia. *Energies*, *16*(2), 651. https://doi.org/10.3390/en16020651

U.S. Department of Energy (2023, March 10). *Sector Spotlight: Virtual Power Plants*. Https://www.Energy.gov/. Retrieved July 10, 2023, from https://www.energy.gov/lpo/articles/sector-spotlight-virtual-power-plants

U.S. Department of Energy (n.d.). *VIRTUAL POWER PLANTS*. Https://www.Energy.gov/. Retrieved June 12, 2023, from https://www.energy.gov/lpo/virtual-power-plants

U.S. Energy Information Administration (n.d.). *Average Operating Heat Rate for Selected Energy Sources*. Https://www.eia.gov/. Retrieved June 12, 2023, from https://www.eia.gov/electricity/annual/html/epa_08_01.html

Ullah, Zahid, et al. (2019). Comprehensive Review of VPPs Planning, Operation and Scheduling Considering the Uncertainties Related to Renewable Energy Sources. *IET Energy Systems Integration*, *1*(3), 147–57. https://doi.org/10.1049/iet-esi.2018.0041

Watson, J, et al. (2019). EGRET: Unit Commitment and {Economic. . .} Dispatch. *Sandia National Lab*. https://www.osti.gov/servlets/purl/1645916

Wiltsee, G. (2000). Lessons learned from existing biomass power plants. United States. https://doi.org/10.2172/753767

World Nuclear Association (2023, June 1). *Nuclear Power in India*. Https://www.World-Nuclear.org/. Retrieved June 20, 2023, from https://www.world-nuclear.org/information-library/country-profiles/countries-g-n/india.aspx

Xu, Ti, et al. (2017). Application of Large-Scale Synthetic Power System Models for Energy Economic Studies. *ScholarSpace*. https://doi.org/10.24251/hicss.2017.386