

# New River Light and Power Data: A Data Ecosystem for Electric Power Systems Research

Ed Hassler<sup>1,\*</sup>, Joseph Cazier<sup>1</sup>, Preston MacDonald<sup>1</sup> and Jamie Russell<sup>2</sup>

<sup>1</sup> Center for Analytics Research & Education, Appalachian State University, USA

<sup>2</sup> Appalachian Energy Center, Appalachian State University, USA

\* Corresponding Author: [hasslere@appstate.edu](mailto:hasslere@appstate.edu)

## Abstract

This article provides data useful for electric power systems research by students, faculty and the community to download or to use in a cloud-based analytics environment with state of the art analytics tools. The work includes carefully cleaning, documenting, preparing and enriching 12+ years of non-personally identifiable substation data for the 5 substations managed by New River Light and Power (NRLP), a non-profit electric utility operated by Appalachian State University's Division of Business Affairs. This data includes bulk summary data for the aggregate 8,000+ meters NRLP manages with a record of the total energy used per substation for each hour of day. This data has been enriched with matching hourly weather and solar radiation data, town daily water and sewer flows to measure town occupancy, a list of university and community events likely to influence power consumption, and a description of each substation region.

## Indexing Table

<b>Supported UN SDGs</b>	<b>Goal 7:</b> Ensure access to affordable, reliable, sustainable and modern energy for all
<b>Type of Data/Article</b>	Archival Data
<b>Class of Analytics</b>	Descriptive, Diagnostic and Predictive
<b>Data Tables</b>	Five tables: Substation, Event, Hourly_Demand, Weather, Water_Sewer
<b>Key words</b>	Energy, Demand Modeling, Teaching Energy Systems, Sustainable Energy

## Introduction

This paper presents hourly substation data from the town of Boone, North Carolina collected over a 13 year period from 2007 through 2020. Each of the 5 substations is somewhat unique, comprising a business district, medical district, university, sub-urban and rural areas. Additional data was collected for hourly weather, solar radiation, water and sewer flow, and information about the university schedule and local events. Together this data forms a data ecosystem useful as a living learning laboratory for studying energy usage, demand, and the influence of population and university occupancy, weather, solar radiation, time of day, season and other factors.

This resource will be particularly useful for the following:

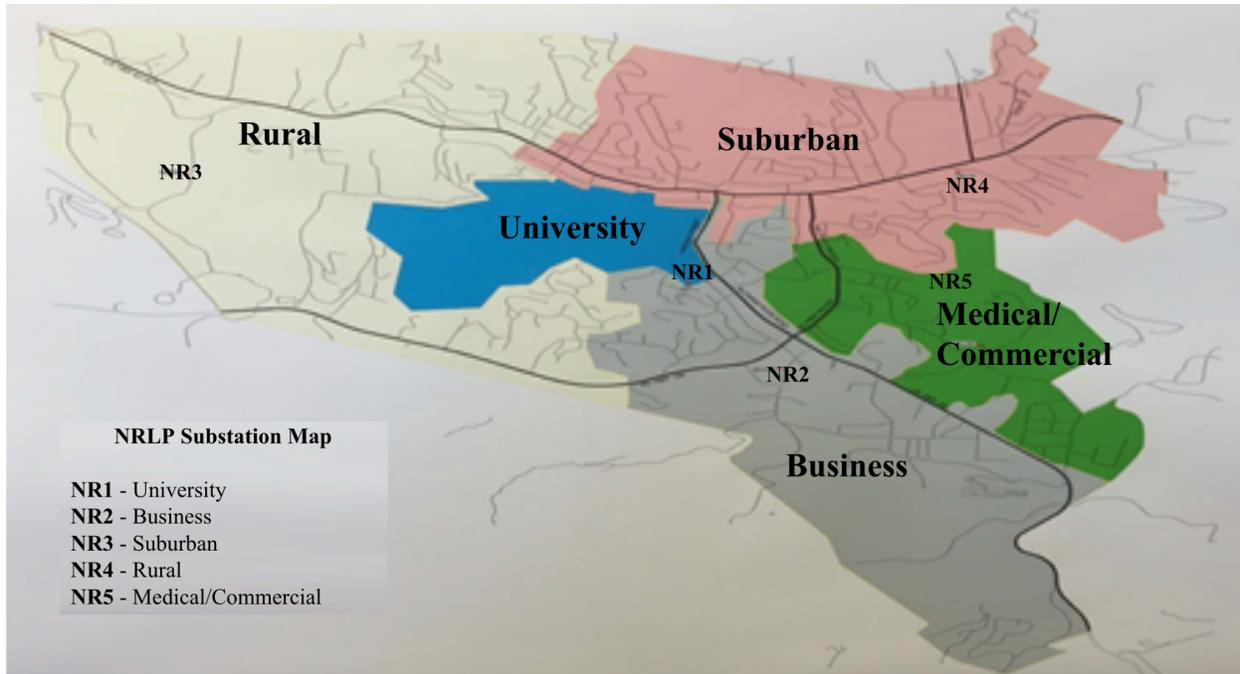
- *Teaching* - The data and tools could be used in the classroom for exploring the impact of weather, development and substation characteristics on energy usage. For example, they could explore how energy patterns vary on campus and in the community based on class schedules, breaks and special events.
- *Research* - This is a great resource for those researching energy systems. They can see the impact of energy use over time and the impact of weather on each substation with its unique characteristics. Additionally this will help model peak load for NRLP and similar communities.
- *Outreach* - This can be used to educate NC residents about the impact of weather and how power consumption patterns change based on insights and as a tool for demonstrations with nice visuals and dashboards on the system. Additionally, this can help the 100 utilities in NC better model their demand and load based on insight from NRLP and the factors affecting energy demand.

This data supports United Nations Sustainable Development Goal # 7 *Ensure access to affordable, reliable, sustainable and modern energy for all* (United Nations, 2021). By better understanding the factors that cause or impact energy demand, steps can be taken to build a more sustainable and reliable energy grid.

### **Data Summary**

New River Light and Power (NRLP) is a 100+ year old public utility located in Boone, a mountainous region of North Carolina. The town is home to Appalachian State University, a public university consisting of approximately 20,000 students located in a classic college town consisting of another ~25,000 residents. NRLP was founded and is still owned and operated by the university to provide power to the university and region. It operates just over 8,000 meters spread over 5 substations in town. Additional data was collected from NOAA, the town of Boone, the university schedule and the internet for dates of important regional events that might influence town occupancy and hence power consumption. The data includes more than 250,000 records over a 13 year period.

Figure 1 shows a map of the town, approximate location of each substation and a visualization of each of the 5 districts.



**Figure 1: Substation Map of Boone, North Carolina**

*NR# = Approximate physical location of each substation followed by our name for the region*

While there are no clean breaks for the region, we estimate the following breakdown for each region in the NRLP area.

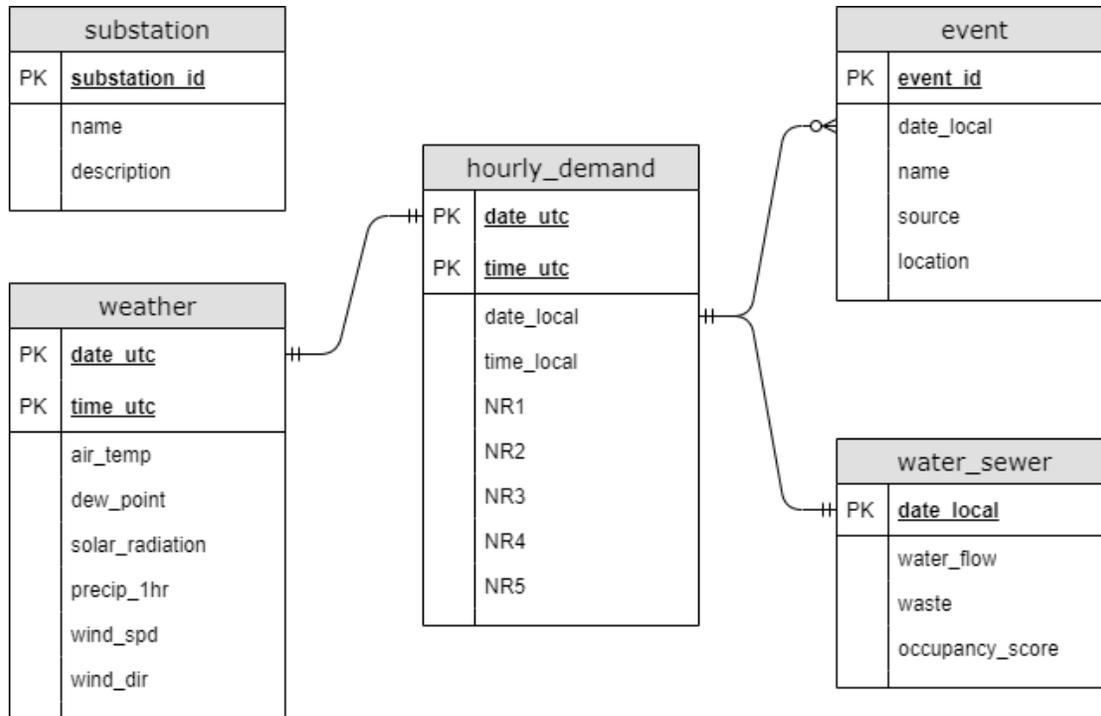
- *University (NR1)* - This substation primarily serves the university, including dorm rooms, classroom space, admin buildings and parking.
- *Business (NR2)* - Primarily consists of commercial business, fast and slow food, hotels, apartments and tourism.
- *Rural (NR3)* - This is the area adjacent to town. While there are a few small businesses in the area, it is mostly town adjacent mountain residential living.
- *Suburban (NR4)* - This area consists of the nicest, most established and convenient locations in town that are generally seen as the most desirable places to live inside Boone.
- *Medical/Commercial (NR5)* - Consists of the area near the local hospital. Though the hospital does not get its power from NRLP, most of the medical offices in that district do with a mix of light commercial power usage.

## Data Description

The data consist of 5 tables, as described below.

- *Substation* - The substation data contains information about each region similar to what was described in the last section. Unlike traditional relational tables, each row in this table is associated with a column in the *Hourly\_Demand* table.
- *Hourly\_Demand* - This table holds the core data for the project consisting of hourly demand for the 13 year period for each of the 5 substations. Note that it contains both Coordinated Universal Time (UTC) and Local Time. UTC is needed to link the data to the NOAA weather data which is also in UTC time. Be careful with local time as it needs to be adjusted for Daylight Savings Time.
- *Weather* - The weather table holds data about the local weather, collected from the Boone Airport located about 3 miles from the university. Airport data was used for weather due to its consistency and quality which tends to be better at airports due to their dependence on this data. It also includes data added in for solar radiation.
- *Water\_Sewer* - This table contains data from the Town of Boone regarding daily gallons of water and sewer measured for the years 2008 through 2020 (2007 data was not available). It is thought that this will be a nice measure of the occupancy of the town (more people = more flushes) which could help determine the impact of various holidays and special events. Note that sewer rates can also be impacted by stormwater from rain, and water usage can spike when people wash cars or water their lawns, so thought should go into how to adjust for these factors. Additionally, the University is on its own water/sewer system for which we were not able to get reliable data, therefore this data is only good for residential and commercial areas. Note that NRLP primarily serves homes in city limits, therefore almost all of them are on town water/sewer.
- *Event* - Two types of events were collected. Those that would bring people to town, and those that would send them away. Both are expected to impact energy use. Events that bring people to town include graduation, football games, local and regional events such as the Woolly Worm Festival or Merlefest, and leaf season. The primary events sending people away consists of university breaks (i.e. spring and winter break) when most students leave. Note that university and community events are likely to most impact different substations. Due to the difficulty of finding this data for all 13 years, and the change in events, efforts were focused on gathering this data for the most recent 5 years, excluding the COVID-19 pandemic lockdown.

Figure 2 shows an entity relationship diagram for the data, and Appendix A gives a detailed description and definition of each field in each table.



**Figure 2:** Entity Relationship Diagram of tables.

## Data Collection Methods

### Substation

Each substation was analyzed using Google maps and interviews with local residents familiar with the area to develop the descriptions for each substation. The map was provided by NRLP and labeled by our team.

### Hourly Demand

Hourly demand data for each substation was extracted from operational monitoring systems and associated archival storage. Readings were taken at the top of each hour at the source for each substation. Thus, downstream loss in transmission is included in the readings. Additionally, timestamps on all readings were originally in the local timezone and included a mix of standard and daylight savings time. These timestamps were converted to UTC for accounting purposes and to enable precision joining with the retrieved weather data.

### Weather

Weather data were derived from the National Centers for Environmental Information Integrated Surface Dataset (NOAA NCEI, 2001). Data files for the local aviation weather station covering the period 2007 through 2020 were downloaded. The mandatory reporting fields are extracted yielding observations at 20

minute intervals. Hourly observations were then constructed by utilizing the observations within +/- 5 minutes of the top of the hour and selecting the observation closest to the top of each hour. The solar radiation data was collected from Mt. Jefferson tower, which is located about 18 miles away from the town of Boone.

### Water and Sewer Data

Daily water and sewer data was collected following a Freedom of Information Act (FOIA) request submitted to the town of Boone.

### Event Data

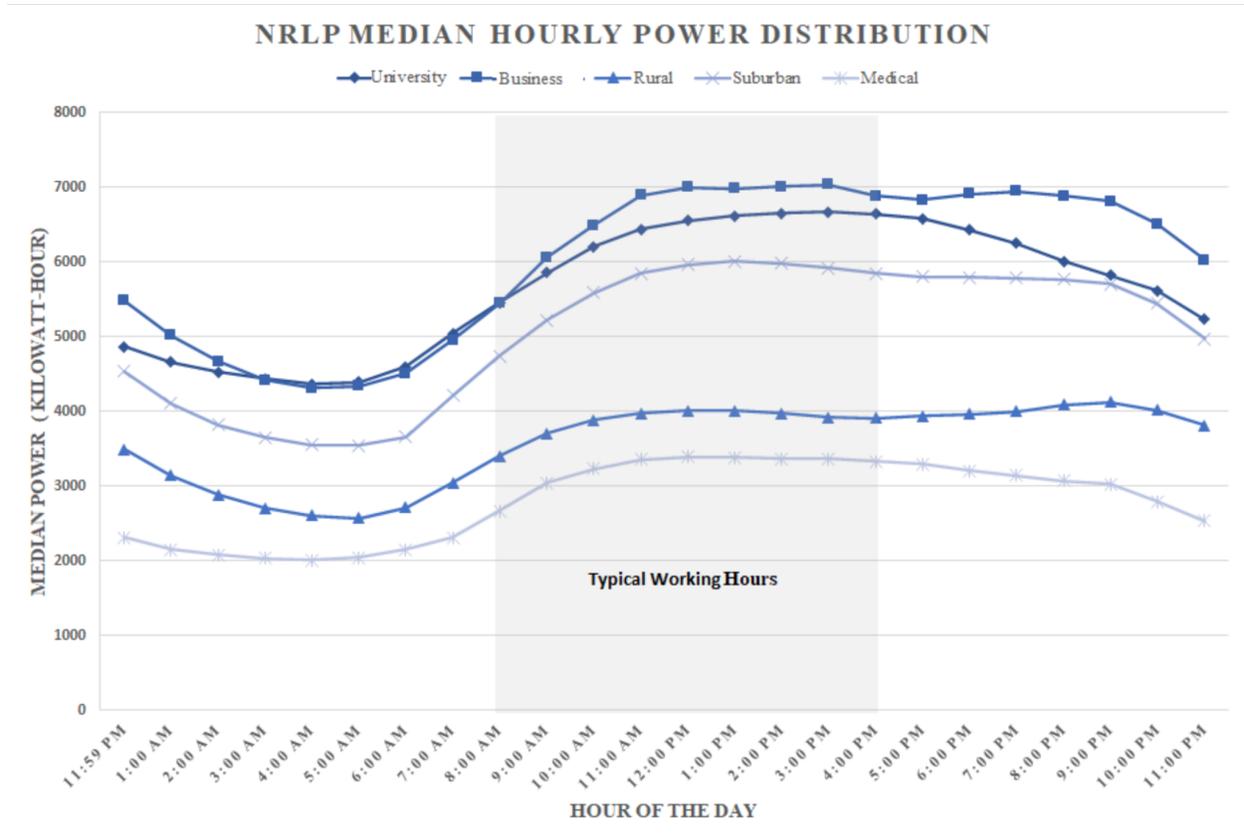
Event data was gathered from the Appalachian State University academic and registration calendars, news articles indicating major events, and the High Country Events Calendar. The data spans from 2015 to 2019 and is consolidated into the dataset with general information such as the location and the source of the event.

## **Preliminary Data Exploration**

The following charts and graphs are intended to give the modeller some clues as to the factors influencing power consumption.

### Figure 3: NRLP Median Hourly Power Distribution.

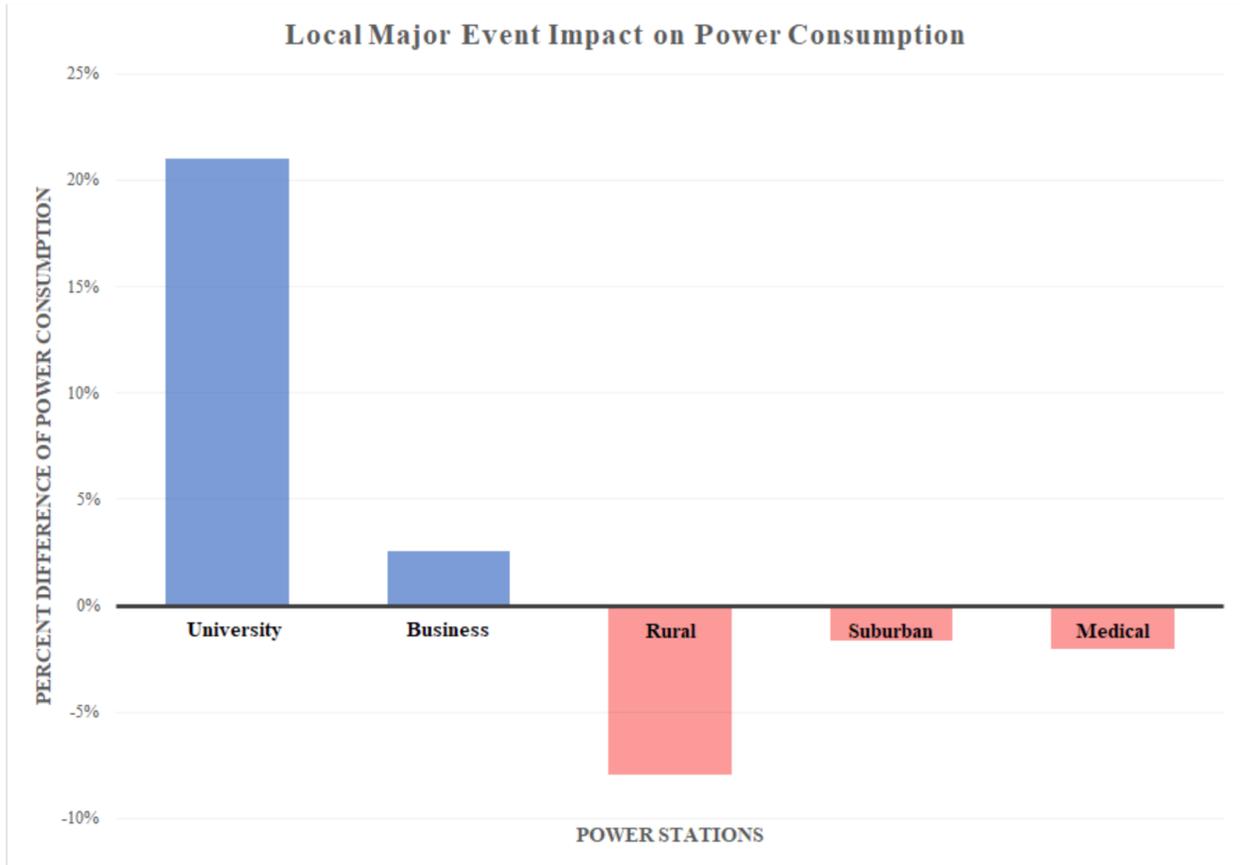
The 5 substation power outputs are plotted across a 24 hour time period with the aggregated median hourly power data from 2015-2019. Note that during the typical working hours between 8:00 AM to 4:00 PM Eastern Time (EST) shows high power usage and how these trends differ between the different substations.



**Figure 3: NRLP Median Hourly Power Distribution.**

Figure 4: Local Major Events Impact on Power Consumption

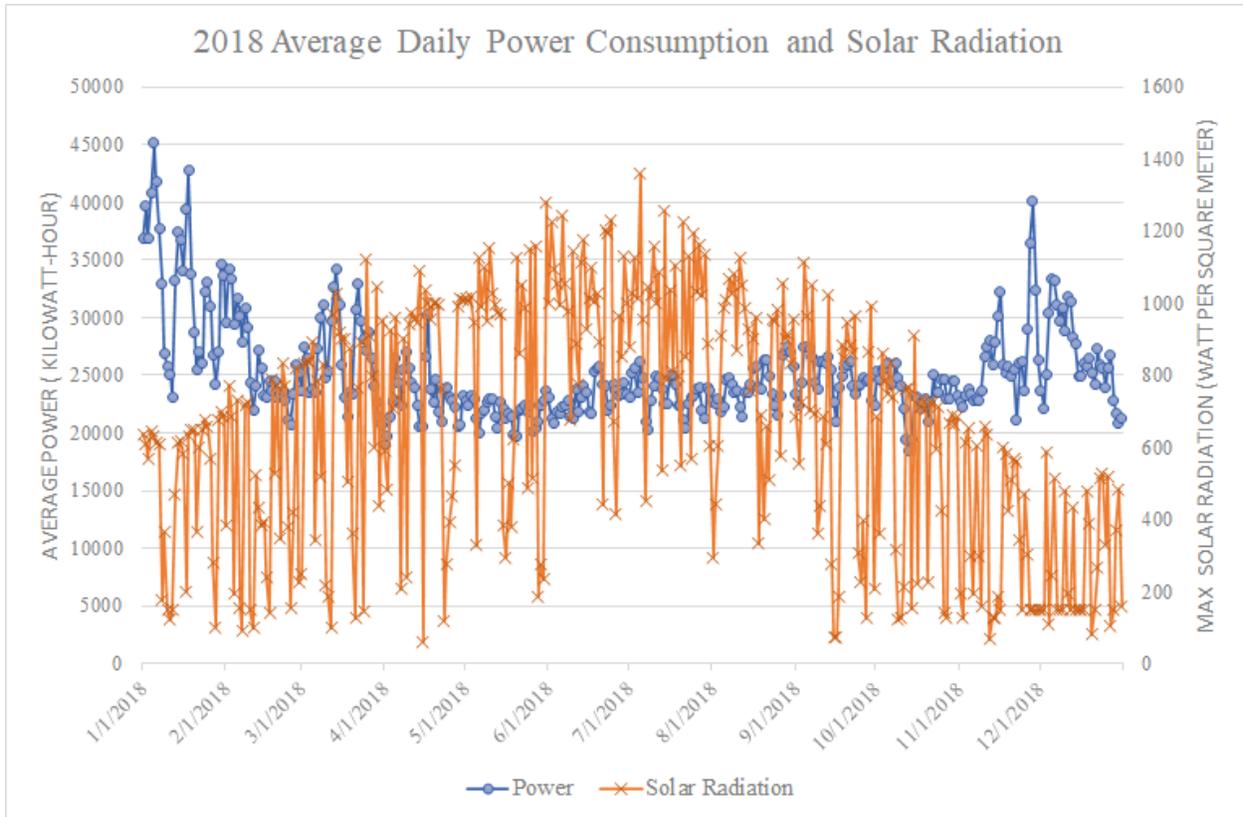
The 5 substation power outputs for major events dates are compared to typical days using medians from 2015 through 2019 in order to calculate the percent change in power consumption. Major events include Appalachian State University graduation dates, home football games, and large events that draw media attention. We can see from the chart below how major events brought a 20% increase in the university area’s power consumption and a modest increase in the business district. This is not surprising given that the university hosts many of the major events and people shop and eat in the business district. Interestingly some of the other substations power consumption drops, possibly due to travel or their attendance at the events.



**Figure 4: Local Major Events Impact on Power Consumption**

Figure 5: 2018 Average Daily Power Consumption and Solar Radiation

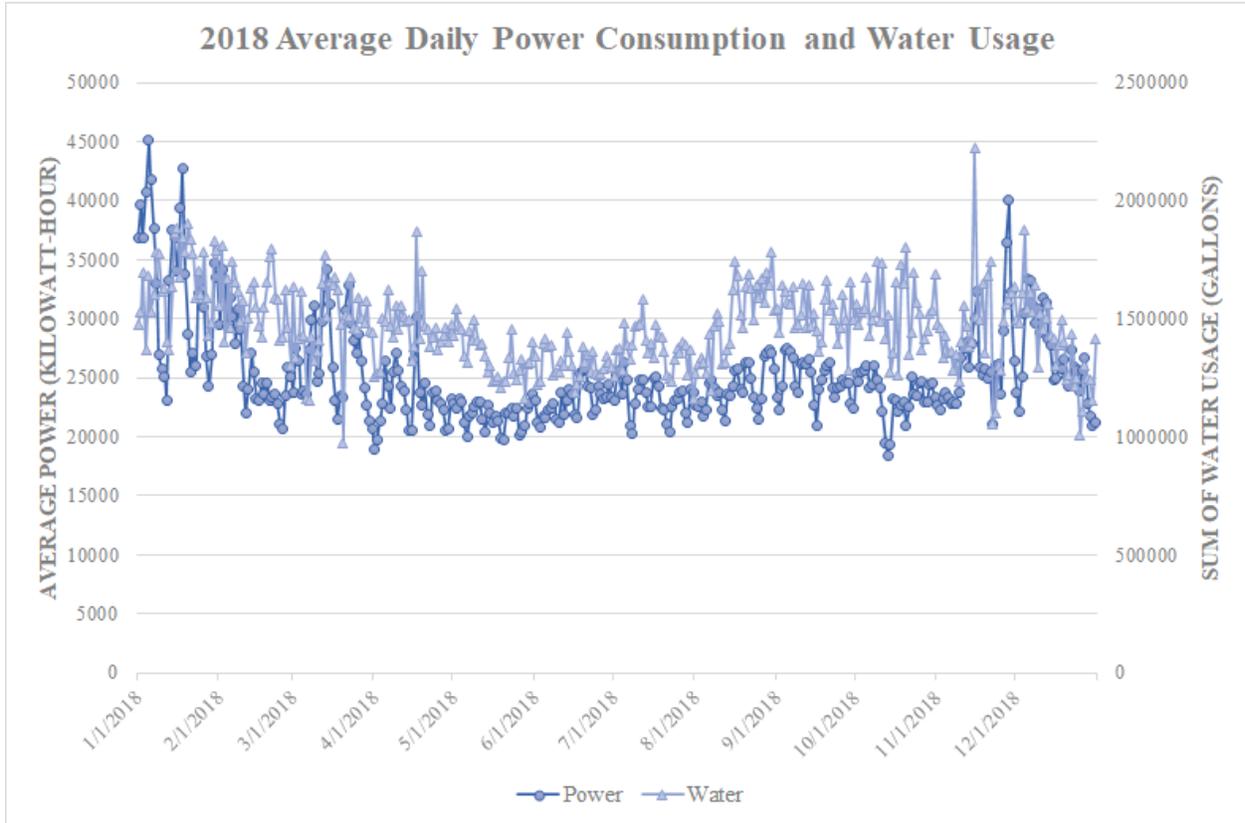
For the 2018 Average Daily Power Consumption and Solar Radiation analysis, we took the average daily total power consumption for all 5 substations and the maximum daily solar radiation readings for 2018. Notice when solar radiation is low, power consumption is high in the winter months (most likely for heating). Additionally in the summer months we see high solar radiation levels and lower power consumption.



**Figure 5: 2018 Average Daily Power Consumption and Solar Radiation**

Figure 6: 2018 Average Daily Power Consumption and Water Usage.

We took the average daily total power consumption for all 5 substations and the total daily water usage for 2018. Using Kendall rank correlation we found a correlation of .35. You will notice that as water usage increases so does power consumption. Also note we had a large outlier in September for water consumption, which we guess is due to a water line break.



**Figure 6: 2018 Average Daily Power Consumption and Water Usage.**

### Data Application, Extensions and Conclusion

This data will be useful for exploring and predicting energy demand across each of the 5 substations based on weather, solar radiation, town occupancy, time of day, and season. Here are a few ways this data might be useful:

- Determining the effect of the university and community calendar, solar radiation and weather on university and town power consumption;
- Looking at seasonal and hourly trends in whole and by substation;
- Exploring the moderating impact of town occupancy, measured by water/sewer flows on power consumption;
- Predictive modeling to predict demand in whole and by substation under various conditions, include peak demand forecasting;
- Finding the best time and place to supplement energy needs with sustainable sources to reduce the carbon footprint and help the town meet its 2050 carbon neutral pledge.

Possible extensions to this work and data include:

- Gathering data from nearby power stations to analyze not just NRLP energy patterns, but that of the region, for it is the regional power usage that determines peak demand charges. For NRLP, this includes most of the area in the western third of the state, including Greensboro, NC and Charlotte, NC, which drive much of the demand;
- Data from university food services could give a better count on university occupancy and the extent to which occupancy influences energy consumption;
- Smart meter data from each home would be ideal, and is being collected, but there are legal challenges to sharing it. A compromise approach might be to add detailed summary statistics describing this data by home type, substation locations and other logical clusters that might be useful for analysis.

### **Acknowledgements**

The authors wish to thank Ed Miller of NRLP for his cooperation in sharing and documenting this data and the Appalachian Energy Center for their financial support of this project.

### **References**

- NOAA NCEI. (2001). "Global Surface Hourly Integrated Surface Dataset". *NOAA National Centers for Environmental Information*. Accessed 5/1/2021.
- United Nations (2021). "The 17 goals", United Nations Department of Economic and Social Affairs Retrieved April, 2021, from <https://sdgs.un.org/goals/goal8>

## Appendix A - Data Definitions

### Entity 1: Hourly Demand Data Definitions

Attribute Name	Type	Definition
<i>date_utc</i>	date	Coordinated Universal Standard date
<i>time_utc</i>	time	Coordinated Universal Standard time
<i>date_local</i>	date	Eastern Standard Time or Eastern Daylight Time date
<i>time_local</i>	time	Eastern Standard Time or Eastern Daylight Time
<i>NR1</i>	float	NR1 power consumption over time of observation taken measured in Kilowatt-Hour
<i>NR2</i>	float	NR2 power consumption over time of observation taken measured in Kilowatt-Hour
<i>NR3</i>	float	NR3 power consumption over time of observation taken measured in Kilowatt-Hour
<i>NR4</i>	float	NR4 power consumption over time of observation taken measured in Kilowatt-Hour
<i>NR5</i>	float	NR5 power consumption over time of observation taken measured in Kilowatt-Hour

### Entity 2: Substation Data Definitions

Attribute Name	Type	Definition
<i>substation id</i>	int	Substation identifier
<i>name</i>	string	The substation name
<i>description</i>	string	A short description to describe the substation area

### Entity 3: Event Data Definitions

Attribute Name	Type	Definition
<i>event_id</i>	int	Event identifier
<i>date_local</i>	date	Eastern Standard Time or Eastern Daylight Time date.
<i>name</i>	string	A short description to describe the substation area
<i>source</i>	string	A website URL to indicate where the data point was pulled
<i>location</i>	string	A categorical variable that indicates the general location of the event: Federal Holiday Local Attraction Neighboring County Attraction University Inflow University Outflow

### Entity 4: Water & Sewer Data Definitions

Attribute Name	Type	Definition
<i>date_local</i>	date	Eastern Standard Time or Eastern Daylight Time date.
<i>water_flow</i>	float	Total water flow in gallons per day
<i>waste</i>	float	Total wastewater flow in gallons per day
<i>occupancy score</i>	float	The water flow divided by the season max water flow in gallons per day

### Entity 5: Weather Data Definitions

Attribute Name	Type	Definition
<i>date_utc</i>	date	Coordinated Universal Standard date
<i>time_utc</i>	time	Coordinated Universal Standard time
<i>air_temp</i>	float	The average air temperature in degrees celsius
<i>dew_point</i>	float	The average dew point in degrees celsius
<i>solar_radiation</i>	float	The solar radiation measured in watts per square meter

		(W/m <sup>2</sup> )
<i>precip-1hr</i>	float	The depth of liquid-precipitation that is measured at the time of the observation in millimeters
<i>wind_spd</i>	float	Rate of horizontal travel of air past a fixed point in units of meters per second
<i>wind_dir</i>	float	The angle, measured in a clockwise direction, between true north and the direction from which the wind is blowing in angular degrees