Final Report

Lakeland Electric Revenue and Load Modeling

Capstone

12/11/24

Introduction:

Lakeland Electric has a lot of unique challenges when it comes to forecasting electricity usage and revenue. Standard industry models work as a solid baseline, but they do not capture Lakeland's specific behavior. This project is designed to create a model for improving the revenue and load modeling relating to Residential customers and then expand if time permits. The current Lakeland Electric models focus on monthly aggregates and then create a 1 year short-term forecast and a 10 year long-term forecast. Our goal is to create models that utilize hourly/daily data to obtain more specific forecasts. From there, if the models work well, they can be expanded and modified in future iterations to eventually handle real-time forecasting and other building classes. Lakeland Electric plans on this being a multiple year project so our goal is to create a solid starting point that can be built upon by future work. The priority is creating something reliable and scalable that can grow and adapt as Lakeland Electric's needs change.

Customer Needs/Requirements:

Lakeland Electric needs models that can accurately predict electricity load and revenue to improve resource planning and pricing.

- Key Needs:
  - More specific forecasting models that utilize daily/hourly data and local patterns.
  - Initial focus of residential customers since they are by far the biggest new business.
  - Pipeline for expanding into other building classes/customers.
  - Utilizing hourly/daily data in the forecasts to obtain more specificity.
- Functional Requirements:
  - Capable of processing hourly and daily data for energy usage, peak demand, and revenue prediction.
  - Use of historical weather and demographic data (See Appendices C and E).
  - Implementation of the various residential rate structures (RS, RSX-1, RSD)[1] (Appendix C).

- Clear, interpretable outputs.
- Accurate models when compared to ITRON and available Lakeland Electric data (Appendices B-D).

Problem Decomposition/Constraints:

1. Problem Decomposition: (See Appendices A-E for further information)
    - Load Model:
        - Analyze and predict hourly/daily electricity load for residential customers.
        - Use weather (HDD/CDD)[1], customer, and time of day trend variables.
    - Revenue Model:
        - Forecast revenue based on rate classes and customer usage patterns.
        - Account for various residential rate structures within the model.
2. Constraints:
    - Data Availability:
        - Data is obtained directly from Lakeland Electric – we can't obtain some of it ourselves.
        - Will need to fill any missing data with external data sources – e.g weather.gov.
    - Scope:
        - Residential customers/classes only initially – look to expand once this model is capable.
    - Accuracy vs. Understandability:
        - Models need to have interpretable and reliable output.

This approach helps the project move forward step by step, focusing on immediate needs while staying flexible for future growth.

Potential Solutions/Analysis:

We have identified multiple potential models to start out with. Initially, we will be focusing on multiple regression models like the ITRON standard Lakeland Electric currently uses. From there we can look to feature engineering and other model improvements.

1. Literature Review (See Appendices B-E):
    - Examined existing energy forecasting methods, including:
        - Itron's SAE econometric approach for understanding residential energy consumption trends [1][2].
        - Modeling techniques like multiple regression, ARIMA, and exponential smoothing.
    - Incorporates Lakeland data but doesn’t adapt the models to utilize it best.
2. Proposed Modeling Approaches (See Appendices A-E):
    - Load Model (Refer to Appendix A):
        - Start with multiple regression to establish basic relationships between variables such as weather, time-of-day, and load.
        - Incorporate time series methods (e.g., exponential smoothing, ARIMA) for capturing trends and seasonality.
    - Revenue Model (Refer to Appendix A):
        - Use regression-based approaches to link load predictions to revenue.
        - Test additional features (X below) to try and identify ways to improve the models.
3. Alternatives Considered:
    - Machine learning models (e.g., Random Forests, Gradient Boosting) seem promising as well, but time and initial data constraints will make them more of a future work goal.
    - Smart-grid data integration was also identified as a longer-term goal.
4. Initial Focus:
    - Simpler models (multiple regression) to start with and then expand via feature engineering and various forecasting methods (exponential smoothing vs ARIMA).

Applicable Codes/Standards:

There is no safety code or standard required for this project. However, Lakeland Electric uses Itron’s forecasting tools, widely recognized as an industry standard for energy forecasting[1][2]. We will be using their current model and practices as a baseline to initially create and compare our model to.

Industry Standards:

1. Itron Methodologies (Appendix D):
    - The SAE model provides a benchmark for comparing the residential load and revenue models developed in this project.
    - We will utilize their way of handling weather data and add to it (HDD/CDD)[1] (See Appendix C).
2. Best Practices for Data-Driven Forecasting:
    - Proper obtaining and cleaning of the data will be performed.
    - Models created will follow the rules learned within our classes here at Florida Poly.

This project aligns with how ITRON performs their industry standard analysis. By building on these methodologies and incorporating them into the model design, this project ensures that standard practices are followed.

Conclusions/Future Work:

This project focuses on tackling the distinctive challenges of forecasting electricity usage and revenue for Lakeland Electric, beginning with residential customers. This allows us to create models for the customer section they most need it for, and then expanding when/if time allows. Our models will incorporate the Lakeland weather, customer behavior, and usage obtained directly from Lakeland Electric. From there we will investigate potential features that could be useful that are not currently included. Initial efforts emphasize scalable and interpretable methods, with the option to expand into more advanced techniques and additional customer classes in the future. Refer to Appendix E

Future Work:

1. Data Acquisition/Cleaning:
    - Obtain data from Lakeland Electric (currently in process).
    - Ensure the data is clean and ready to use in modeling.
2. Model Creation:
    - Start with multiple regression as shown in Appendix A, and then expand the features and interactions between features as we continue.
3. Evaluation/Benchmarking:
    - Compare our model outputs to ITRON's to see how well our models are performing.

- Cross-validation to ensure our models are consistent.

4. Expand Feature Set:

    - Additional variables such as some of the smart-grid data will be explored.

    - Nonlinear feature engineering to try to identify any interactions not captured by the initial model.

5. Scalability Planning:

    - Pipeline will be created to help Lakeland Electric have an idea of how to proceed with the other customer classes if we don't have time to get to it.

6. Final Deliverables:

    - A complete report will be provided to Lakeland Electric outlining our entire process along with any code or models created.

    - This plan will also include any pipelines created for expansion for the other customer classes or real-time implementation.

Future work will be focused on expanding the models to the other customer types along with trying to identify ways to scale this further into the future for real-time data use.

Citations

1) Tariff Document:

Lakeland Electric. (2024). *Residential Service Rate Schedules RS, RSX-1, and RSD.* Lakeland Electric Tariff Documents. Retrieved from the Lakeland Electric utility services on December 1st, 2024.

2) **ITRON Information:**

Itron. *Itron Energy Forecasting and Analytics.* Retrieved November 16, 2024, from https://na.itron.com/

Appendices

Table of Contents:

Appendix A: Revenue and Load Models Created

Load Modeling:

- **$L(t, c)=\beta 0 + \beta 1 \cdot H(t,c)+\beta 2 \cdot U(t,c)+\beta 3 \cdot W(t)+\beta 4 \cdot X(t,c)$**
  - **L(t, c):** Projected load at time t for residential customers.
  - **β0 :** Intercept (baseline load when all other factors are zero).
  - **H(t, c):** Historical average load at time t for residential customers
  - **U(t, c):** Current usage trend for residential customers.
  - **W(t):** Weather data at time t (such as temperature, humidity, etc.).
  - **X(t, c):** Any additional features that may influence load, e.g:
    - Day of the week
    - Building occupancy levels
    - Economic activity
    - Special events
    - Etc.
  - **β1, β2 , β3 , β4 :** Coefficients for the features

Above represents the first attempt at creating a model for forecasting the electrical load for Lakeland Electrical. Each variable is explained underneath the model above. Starting off the only building class/customer type being considered is Residential. Currently the only weather data Lakeland Electrical uses is HDD and CDD which are explained further in Appendix D below. This provides a way for us to easily expand their existing models by incorporating additional weather data such as humidity, expected rainfall, windspeed, etc. The historical average load, current usage data, and the basic weather data are currently in the process of being provided to us by Lakeland Electric. They are working on gathering data for residential customers for us to utilize. The additional features we plan on exploring are shown above as X. Factors such as building occupancy, number of occupants who work from home, price elasticity for residential customers, and peak hour/day usage will be explored depending on time permitted and data availability.

Revenue Modeling:

- **$R(t,c)=\beta 0 + \beta 1 \cdot L(t,c) + \beta 2 \cdot P(t) + \beta 3 \cdot X(t,c)$**
  - **R(t,c):** Projected revenue at time t for residential customers
  - **L(t,c):** Projected load (electricity usage)
  - **P(t):** Price of electricity at time t

$$P(t) = \begin{cases} \boldsymbol{Ppeak}, & if\ t\ \in peak\ hours \\ \boldsymbol{Poff} - peak, & if\ t\ \in off - peak\ hours \end{cases}$$

  - X(t,c): Any additional features that might affect revenue modeling
    - Seasonal trends
    - Customer class
    - Location
    - Demand response (e.g reduced usage during peak hours)
    - Etc.

Above represents the Revenue Model currently being worked on. It builds off the load model and incorporates the load into the equation. This equation focuses on utilizing the projected load from our load model along with the various pricing options (see Appendix B) to predict the revenue. The additional features are once again shown as X in the equation. This time the additional features such as seasonality and demand response will be incorporated into our forecasting analysis initially. Customer class and location will be utilized as time permits since we are initially focusing solely on residential customers from set locations they are providing the data for.

Appendix B: Lakeland Electric Relevant 10 Year Plan Information

“3.0 Forecast of Electric Demand and Energy

Annually, Lakeland Electric (LE) develops a detailed short-term (1 year) electric load and energy forecast for budget purposes and short-term operational studies. An annual long-term (10 years) forecast is developed for the Utility’s long-term planning studies including LE’s Ten Year Site Plan (TYSP).

Sales and customer forecasts of monthly data are prepared by rate classification. Separate forecast models are developed for inside and outside the City of Lakeland corporate limits for the Residential, Commercial, Industrial and Other (municipal departments and outdoor lighting) rate classifications. Monthly forecasts are summarized annually using fiscal period ending September 30th for the short-term budget forecast and by calendar year for long-term studies and reporting.

Lakeland Electric uses MetrixND, an advanced statistical forecasting software tool, developed by Itron, to assist with the development of LE’s number of customers, energy and demand forecasts. Lakeland Electric uses MetrixLT, another Itron software tool, which integrates with MetrixND to develop the long-term system hourly load forecast.

The modeling techniques used to generate the forecasts include multiple regression, study of historical relationships and growth rates, trend analysis, and exponential smoothing. Lakeland Electric utilizes Itron’s Statistically Adjusted End-Use (SAE) econometric modeling approach for the residential and commercial sectors. The SAE approach is designed to capture the impact of changing end-use saturation and efficiency trends, by building type, as well as economic conditions on long-term residential and commercial energy sales and demand.

Many variables are evaluated for the development of the forecasts. The variables that have proven to be significant and are included in the forecasts are weather, gross regional product, disposable personal income per household, persons per household, number of households, local population, electricity price, building type, appliance saturation and efficiency. Binary variables are used to explain outliers in historical billing discrepancies, trend shifts, monthly seasonality, rate migration between classes and other issues that could affect the accuracy of forecast models.

Weather variables

Heating and cooling degree days are weather variables that attempt to explain a customer’s usage behavior as influenced by either hot or cold weather. Heating Degree Days (HDD) occur when the average daily temperature is less than Lakeland Electric’s established base temperature of 65 degrees Fahrenheit. Cooling Degree Days (CDD) occur when the average daily temperature is greater than 65 degrees. The formulas used to determine the number of degree days are:

HDD = Base Temperature (65) – Average Daily Temperature

CDD = Average Daily Temperature – Base Temperature (65)

These HDD and CDD variables are used in the forecasting process to correlate electric consumption with weather. The HDD and CDD variables are weighted to capture the impacts of weather on revenue from monthly billed consumption.

Lakeland Electric uses weather data from its own weather stations, which are strategically placed throughout the electric service territory to provide the best estimate of overall temperature for the Lakeland Electric service area.

The most recent 20 years of historical normal weather is used as an input into the sales forecast models.

Normal peak-producing weather is also developed using historical 20-years weather. A weighted average of temperatures on both the day of historical monthly peak and day prior to peak is used to create the HDD and CDD variables.

Economic and demographic variables

The economic and demographic projections used in the forecasts are purchased from Moody's Analytics.

Price variables

A real price forecast by month and rate class is created based on Lakeland Electric historical price data, projections from the Lakeland Electric Rates and Fuel teams, the U.S. Energy Information Administration (EIA) Annual Energy Outlook (AEO) forecasted price of electricity, historical and projected Net Energy for Load, and the projected Consumer Price Index. The 12-month moving average of projected real price of electricity is the price variable used in the sales and demand SAE models.

Structural Indices

The end-use saturation and efficiency indices used in the models are purchased from Itron. Itron's Energy Forecasting Group (EFG) offers end-use data services and forecasting support. EFG's projections are based on data derived from the EIA's AEO forecast for the South Atlantic Census Division. Itron is also contracted to further calibrate the indices based on Lakeland Electric's service area using average square feet by building type for the Commercial Sector and average use by dwelling type for the Residential Sector.

Lakeland Electric reviews the forecasts for reasonableness, compares projections to historical patterns, and modifies the results as needed using informed judgment.

Historical monthly data is available and is analyzed for the 20-year period. Careful evaluation of the data and model statistics is performed; this often results in most models being developed using less than a 10-year estimation period.

Lakeland Electric currently does not have any specific energy savings goals through Demand Side Management (DSM) programs; therefore, Lakeland Electric does not assume any deductions in peak load for the forecast period.

Lakeland Electric forecasts the number of monthly electric accounts for the following categories and subcategories:

- Residential, Inside and Outside City Limits
- Commercial, Inside and Outside City Limits
- Industrial, Inside and Outside City Limits
- Other, Inside and Outside City Limits

Lakeland Electric's Energy Sales Forecast is the sum of the following forecasts:

- Residential, Inside and Outside City Limits
- Commercial, Inside and Outside City Limits
- Industrial, Inside and Outside City Limits
- Other, Inside and Outside City Limits

Residential Energy Sales Forecast

The Residential energy sales forecast is developed using the Statistically Adjusted End-Use (SAE) econometric modeling approach.

The residential sales models are estimated with historical monthly energy sales data. They are average use models based on the following equation:

$$AvgUse_{y,m} = b_0 + b_1 XCool_{y,m} + b_2 XHeat_{y,m} + b_3 XOther_{y,m} + \varepsilon_{y,m}$$

Where $XCool_{y,m}$, $XHeat_{y,m}$ and $XOther_{y,m}$ are explanatory variables constructed from weather data, end use equipment efficiency and saturation trends, economic and demographic data, dwelling type (single family, multi family or mobile home) and square footage."

* See attached 10 year Plan (Lakeland Electric) for additional information

Appendix C: Lakeland Electric Relevant Tariff Document Information

“RATE SCHEDULE - RS

RESIDENTIAL SERVICE

Available:

In all territory served by Lakeland Electric.

Applicable:

To all electric service provided to single family homes, mobile homes, apartments, condominiums, or cooperative apartment buildings where such energy usage is exclusively for residential purposes subject to the following requirements:

1. 100% of the energy used is exclusively for the customer’s benefit.

2. None of the energy is used in any endeavor which sells or rents a commodity or provides service for a fee.

3. Each point of delivery will be separately metered and billed.

4. A responsible legal entity is established as the customer to whom a bill can be rendered.

5. Beginning January 1, 2016, new solar electric systems interconnected with Lakeland Electric shall take service under Rate Schedule Residential Service Demand (RSD). Existing customers as of this date may maintain service under this rate scheduled through December 31, 2025.

Character of Service:

A-C; 60 Hertz; single-phase, 3-wire; 120/240 volts or 120/208 volts.

Limitation of Service:

Standby service or resale not permitted under this rate schedule.

Net Rate per Month:

Customer Charge: $14.00

Energy Charge: 0 to 1,000 kWh $0.05392 per kWh

1,001 to 1,500 kWh $0.06161 per kWh

Above 1,500 kWh $0.06816 per kWh

Minimum Bill:

Customer charge, plus Adjustments.

Adjustments:

Fuel charge, as contained in Schedule BA-1

City Utility Tax or Surcharge, taxes, surcharges, and fees as contained in Schedule BA-2

Environmental Compliance Cost Charge as contained in Schedule BA-3

Smart Grid Project Implementation as contained in Schedule BA-5

Payment:

Net bills are due when rendered and are delinquent thirty (30) days after the billing date.

Terms and Conditions:

1. All service hereunder will be supplied at one location through one point of delivery and measured through one meter.

2. Service hereunder is subject to the rules and regulations for electric service as adopted by Lakeland Electric from time to time and on file with the City Clerk.

RATE SCHEDULE RSX-1 RESIDENTIAL SERVICE SHIFT TO SAVE

OPTIONAL TIME-OF-DAY

Available:

In all territory served by Lakeland Electric.

Applicable:

To all electric services provided to single family homes, mobile homes, apartments, condominiums, or cooperative apartment buildings where such energy usage is exclusively for residential purposes subject to the following requirements.

1. 100% of the energy used is exclusively for the customers' benefit.

2. None of the energy is used in any endeavor which sells or rents a commodity or provides a service for a fee.

3. Each point of delivery will be separately metered and billed.

4. A responsible legal entity is established as the customer to whom a bill can be rendered.

5. After January 1,2016, service is no longer available for a customer with a solar electric system interconnected with Lakeland Electric. Service shall be moved to Rate Schedule Residential Service Demand (RSD).

Character of Service:

A-C; 60 Hertz; single-phase 3-wire; 120/240 volts or 120/208 volts.

Limitation of Service:

Resale not permitted under this rate schedule.

Net Rate per Month:

Customer Charge: $ 14.00

Energy Charge:

On-Peak: $0.13709 per kWh

Off-Peak: $0.01544 per kWh

Definitions of the Time-of-Day Periods:

All time periods are stated in prevailing time.

Summer Winter

On-Peak Hours Apr. 1 - Oct. 31 Nov. 1 - March 31

(Monday-Friday) 12:01 PM - 9:00 PM 6:01 AM 10:00 AM, and, 6:01 PM - 10:00 PM

Off-Peak Hours: All other weekday hours, and all hours on Saturdays, Sundays, New Year's Day, Memorial Day, Independence Day, Labor Day, Thanksgiving Day, and Christmas Day shall be off-peak.

Minimum Bill:

Customer Charge, plus Adjustments.

Adjustments:

Fuel charge, as contained in Schedule BA-1

City Utility Tax or Surcharge, taxes, surcharges, and fees as contained in Schedule BA-2

Environmental Compliance Cost Charge as contained in Schedule BA-3

Smart Grid Project Implementation as contained in Schedule BA-5

Payment:

Net bills are due when rendered and are delinquent thirty (30) days after the billing date.

Terms and Conditions:

1. All service hereunder will be supplied at one location through one point of delivery and measured through one meter.

2. Service hereunder is subject to the rules and regulations for electric service as adopted by Lakeland Electric from time to time and on file with the City Clerk.

3. Customers who select to take service hereunder and subsequently terminate service shall be prohibited from returning to service hereunder for twelve (12) months unless approved by Lakeland Electric.

RATE SCHEDULE RSD RESIDENTIAL SERVICE DEMAND

RESIDENTIAL PEAK DEMAND

Available: In all territory served by Lakeland Electric

Applicable:

To all electric services provided to single family homes, mobile homes, apartments, condominiums, or cooperative apartment buildings where such energy usage is exclusively for residential purposes subject to the following requirements:

1) Residential customer with solar electric systems interconnected with Lakeland Electric shall take service under this rate schedule either:

a) Beginning January 1, 2016, for new customers and existing customers under Rate Schedule Residential Service Shift to Save Optional Time-of-Day (RSX-1), or,

b) January 1, 2026, for customers who were on Rate Schedule Residential Service (RS) before

January 1, 2016.

2) Residential customer with solar electric systems interconnected with Lakeland Electric shall receive a Value of Solar Credit beginning October 1, 2018.

Otherwise, service hereunder is available at the customer's option.

3) One-hundred percent (100%) of the energy used is exclusively for the customer's benefit.

4) None of the energy is used in any endeavor which sells or rents a commodity or provides service for a fee.

5) Each point of delivery will be separately metered and billed.

6) A responsible legal entity is established as the customer to whom a bill can be rendered.

Character of Service: A-C; 60 Hertz; single-phase 3-wire, 120/240 volts or 120/208 volts.

Limitation of Service: Standby service or resale not permitted under this rate schedule.

Net Rate Per Month:

Customer Charge: $14.00

Demand Charge: $6.49 per kW of Billing Demand

Value of Solar Credit: $0.92 per kW of Billing Demand

Energy Charge: $0.02592 per kWh

Definition of the Time-of-Day On-Peak Period:

All time periods are stated in prevailing time.

Summer Winter

On-Peak Hours Apr. 1 - Oct. 31 Nov. 1 - March 31

(Monday-Friday) 1:01 PM - 8:00 PM 6:01 AM - 10:00 AM

Off-Peak Hours All other weekday hours, and all hours on Saturdays, Sundays, New Year's Day, Memorial Day, Independence Day, Labor Day, Thanksgiving Day, and Christmas Day shall be off-peak.

Minimum Bill:

Customer charge, plus Adjustments.

Adjustments:

Fuel charge, as contained in Schedule BA-1

City Utility Tax or Surcharge, taxes, surcharges, and fees as contained in Schedule BA-2

Environmental Compliance Cost Charge as contained in Schedule BA-3

Smart Grid Project Implementation as contained in Schedule BA-5

Payment:

Net bills are due when rendered and are delinquent thirty (30) days after the billing date.

Determination of Billing Demand:

The billing demand for the month shall be the maximum 30-minute integrated kilowatt on-peak demand in the month.

Where charges specified in a rate schedule are based upon the measured maximum 30-minute integrated demand, it is intended that such demand shall fairly represent the capacity which Lakeland Electric is required to stand ready to supply. In case of installations which use this service in a manner such that measurement over a 30-minute interval does not result in a fair or equitable measure of the supply capacity required to service the customer's load, then the measured demand may be adjusted taking into account the known character of use and the rating data of the equipment connected, or from special tests, the intent being that the demand so determined shall fairly represent the customer's capacity requirement.

In cases where Lakeland Electric elects to use connected load instead of demand measurement, as the method for determining demand, it will take into account the known character of use and the rating data of the equipment connected, the intent being that the demand so determined shall fairly represent the customer's capacity requirement.

Terms and Conditions:

1. All service hereunder will be supplied at one location through one point of delivery and measured through one meter.

2. Service hereunder is subject to the rules and regulations for electric service as adopted by Lakeland Electric from time to time and on file with the City Clerk."

* See Attached Tariff Document for further information.

Appendix D: ITRON Research Links

What Itron is about:

Itron is a proven global leader in energy, water, smart city, IIoT and intelligent infrastructure services. For utilities, cities and society, we build innovative systems, create new efficiencies, connect communities, encourage conservation and increase resourcefulness. By safeguarding our invaluable natural resources today and tomorrow, we improve the quality of life for people around the world.

Company website: https://na.itron.com/

Stock market information:

- https://investors.itron.com/stock-information
- https://companiesmarketcap.com/itron/revenue/
- https://investors.itron.com/node/23631/pdf
- https://investors.itron.com/static-files/af39412c-220f-4c29-94cf-eda9fcc115d6
- https://investors.itron.com/static-files/f81951f0-e7d8-4ea2-b207-2e4effc81c5e

PDF containing a press release and links to their social media: https://investors.itron.com/node/24536/pdf

Itron YouTube Channel: Welcome to the Active Grid

All about Itron's Distributed Energy Resource Optimizer (Has very nice resources and information pertaining Evs): https://na.itron.com/products/der-optimizer

Appendix E: Weather Information and Python Code for Obtaining Weather Data.

Weather Research Links:

- https://www.timeanddate.com/weather/usa/lakeland/historic
- https://www.wunderground.com/forecast/us/fl/lakeland
- https://www.weather.gov/wrh/timeseries?site=KLAL
- https://climatecenter.fsu.edu/climate-data-access-tools/downloadable-data
- https://fawn.ifas.ufl.edu/data/
- https://graphical.weather.gov/sectors/florida.php
- https://www.wunderground.com/history/daily/us/fl/lakeland/KLAL

Links to Understand GFS, NAM, NBM, HRRR

Comprehensive PDF containing details about GFS, NAM, NBM, HRRR

https://mag.ncep.noaa.gov/docs/NCEP_PDD_MAG.pdf

GFS Links

https://www.emc.ncep.noaa.gov/officenotes/newernotes/on442.pdf

https://www.ncei.noaa.gov/products/weather-climate-models/global-forecast

https://registry.opendata.aws/noaa-gfs-pds/

https://developers.google.com/earth-engine/datasets/catalog/NOAA_GFS0P25

https://www.rdocumentation.org/packages/rNOMADS/versions/1.2.0/topics/gfs.model.data

https://vlab.noaa.gov/web/gfs/documentation

https://microsoft.github.io/AIforEarthDataSets/data/noaa-gfs.html

https://stackoverflow.com/questions/42298981/documentation-for-noaa-gfs-output-files

https://registry.opendata.aws/noaa-gfs-bdp-pds/

https://www.ncei.noaa.gov/access/metadata/landing-page/bin/iso?id=gov.noaa.ncdc:C00631

NAM Links

https://ams.confex.com/ams/pdfpapers/154114.pdf

https://www.ncei.noaa.gov/products/weather-climate-models/north-american-mesoscale

https://registry.opendata.aws/noaa-nam/

https://www.weather.gov/rah/namnomogram

https://www.ncei.noaa.gov/access/metadata/landing-page/bin/iso?id=gov.noaa.ncdc:C00630

http://wxmaps.org/pix/nam.fcst

https://windy.app/blog/what-is-nam-weather-model-and-how-it-works.html

https://luckgrib.com/models/nam/

https://emc.ncep.noaa.gov/emc/pages/numerical_forecast_systems/nam.php

https://catalog.data.gov/dataset/north-american-mesoscale-forecast-system-nam-12-km2

NBM Links

https://www.weather.gov/news/200318-nbm32

https://vlab.noaa.gov/web/mdl/nbm

https://theweatherguy.net/blog/basic-info-about-the-nbm-and-icon-models/

https://registry.opendata.aws/noaa-nbm/

https://luckgrib.com/models/nbm_oceanic/

https://nomads.ncep.noaa.gov/txt_descriptions/BLEND_txt.html

https://vlab.noaa.gov/web/mdl/nbm

https://aws.amazon.com/marketplace/pp/prodview-dn3lfyvesyh42#links

HRRR Links

https://learningweather.psu.edu/node/90#:~:text=The%20RR%20model%20provides%20data,more%20simply,%20the%20HRRR).

https://rapidrefresh.noaa.gov/hrrr/

https://www.facebook.com/NWSBoulder/videos/hrrr-model-forecast-of-radar-reflectivity-one-of-our-forecasting-tools-indicates/2448958835132461/

https://windy.app/blog/what-is-hrrr-weather-model-and-how-it-works.html

https://journals.ametsoc.org/view/journals/wefo/37/8/WAF-D-21-0151.1.xml

https://www.weather.gov/news/200210-rapid-model

https://repository.library.noaa.gov/view/noaa/53029/noaa_53029_DS1.pdf

https://www.pivotalweather.com/model.php?m=hrrr

https://www.tropicaltidbits.com/analysis/models/?model=hrrr&region=seus&pkg=ref_frzn

Summaries/Explanations to Understand GFS, NAM, NBM, HRRR

GFS (Global Forecast System):

- A global model run by the US Government (NOAA) and its subsidiary agencies
- Funded by American taxpayers, making its forecast output freely available
- Hydrostatic model, meaning it assumes a constant atmospheric pressure with height
- Runs every 6 hours, predicting conditions out to 16 days, with highest accuracy in the 1-4 day range
- Used by many weather websites and apps

NAM (North American Mesoscale Forecast System):

- A limited-area model, focusing on North America and a portion of the northern Pacific, Arctic, and Atlantic Oceans
- Run by NOAA/NCEP, with a 12-km resolution and a nest covering the continental US at 3-km resolution
- Non-hydrostatic model, allowing for more accurate representation of atmospheric phenomena
- Runs every 6 hours, predicting conditions out to 84 hours (12-km domain) and 60 hours (3-km nest)
- Often overly-aggressive with initiating convection over the US, leading to differences in forecasts compared to other models

NBM (North American Mesoscale Blend):

- An ensemble approach, combining multiple models (including GFS, NAM, and others) to produce a single forecast

- Aims to reduce uncertainty by leveraging the strengths of each individual model
- Not a single model, but rather a blended forecast

HRRR (High Resolution Rapid Refresh):

- A regional model, focusing on the contiguous US and a portion of Canada
- Run by NOAA/NCEP, with a 3-km resolution and hourly initialization
- Non-hydrostatic model, allowing for high-resolution representation of atmospheric phenomena
- Runs every hour, providing rapid updates and detailed forecasts for short-term weather events
- Often used in conjunction with other models, such as GFS and NAM, to improve forecast accuracy

Key Differences Between the Models:

- GFS is a global model, while NAM and HRRR are regional models focused on North America.
- NAM and HRRR are non-hydrostatic models, while GFS is hydrostatic.
- HRRR runs hourly, while GFS and NAM run every 6 hours.
- NBM is an ensemble approach, blending multiple models, while the other models are individual forecasting systems.

Ideas About How GFS, NAM, NBM, and HRRR Account for Seasonality, Time trends, Model Validation, Risk Tolerance, Historical Data/Bias

Seasonality:

- GFS and NAM: As global forecast models, they inherently account for seasonal patterns and cycles in their formulations. However, their performance can vary depending on the season and specific weather phenomena.
- NBM: As a blended model, it incorporates multiple models' outputs, including those that account for seasonality. The NBM's bias correction and weighting scheme help to reduce seasonal biases.
- HRRR (High-Resolution Rapid Refresh): As a high-resolution model, it is designed to capture local and regional weather patterns, including seasonal variations.

Time Trends:

- GFS and NAM: They are designed to forecast future weather patterns, incorporating current conditions and trends. Their forecasts are updated regularly, reflecting changes in the atmosphere and evolving weather patterns.

- NBM: By blending multiple models, it captures a range of time scales and trends, from short-term to long-term. The model's weighting scheme helps to reduce biases and improve overall trend representation.

- HRRR: Its high-resolution output allows it to capture rapid changes and trends in weather patterns, making it suitable for short-term forecasting.

Model Validation:

- GFS and NAM: The National Weather Service (NWS) and other organizations regularly validate these models against observational data, such as surface weather stations, radar, and satellite imagery. This process helps identify biases and areas for improvement.

- NBM: The NWS validates the NBM's performance against a range of observational data, including those mentioned above. The model's bias correction and weighting scheme are designed to minimize errors and improve overall accuracy.

- HRRR: Its high-resolution output makes it challenging to validate directly against observational data. Instead, the NWS and other organizations use ensemble forecasting and other techniques to evaluate its performance.

Risk Tolerance:

- GFS and NAM: As global forecast models, they are designed to provide probabilistic forecasts, which inherently account for uncertainty and risk. The models' ensemble forecasting capabilities help quantify the range of possible outcomes.

- NBM: By blending multiple models, it provides a more comprehensive representation of uncertainty and risk. The model's weighting scheme helps to reduce biases and improve overall risk assessment.

- HRRR: Its high-resolution output allows it to capture localized and rapid changes, which can be useful for risk assessment and decision-making.

Historical Data/Bias:

- GFS and NAM: Historical data and biases are incorporated into their formulations through various techniques, such as model parameterizations and initialization procedures.

- NBM: The model's bias correction and weighting scheme are designed to reduce historical biases and improve overall accuracy. The NBM's blended approach helps to minimize the impact of individual models' biases.

- HRRR: Its high-resolution output is designed to capture local and regional weather patterns, which can be influenced by historical conditions and biases. The NWS and other organizations continually evaluate and refine the model's performance to reduce biases and improve accuracy.

Weather Python Code/Outputs for obtaining data from the above sites (See Attached Python Notebook for further information):

Lakeland Electrical Capstone Project

Potential Data Sources - publicly available

- NOAA or National Weather Association - Free weather data

- U.S Census Bureau or American Community Survey - Census Data

    - Try to find a Lakeland Regional one

- Energy Information Administration or OpenEI - Energy Industry Standards

    - national/state level only

Weather Start

- NWS API - https://api.weather.gov/

    - Info: https://www.weather.gov/documentation/services-web-api

    - Lakeland grid - https://forecast.weather.gov/MapClick.php?x=178&y=130&site=tbw&zmx=&zmy=&map_x=178&map_y=130 = Lat 28.05 N, Lon 81.96 W (Rough Center)

- Lakeland Electrical Coverage - https://www.arcgis.com/apps/Viewer/index.html?appid=ffbd05dc2e94406ebcad89e30636877c

FORECASTS - NEXT 12 HOURS AND NEXT WEEK

```
#!pip install requests

import requests

import pandas as pd

from datetime import datetime, timedelta
```

```
# API for Lakeland - chosen rough center of Lakeland Electrical coverage map
lat, lon = 28.05, -81.96
points_url = f"https://api.weather.gov/points/{lat},{lon}"


# Define headers for the API request
headers = {
    "User-Agent": "myweatherapp (contact@myweatherapp.com)"
}
# Fetch the grid information
response = requests.get(points_url, headers=headers)
if response.status_code == 200:
    data = response.json()
    forecast_url = data["properties"]["forecast"]
    forecast_hourly_url = data["properties"]["forecastHourly"]
    grid_url = data["properties"]["forecastGridData"]
    print("12-hour Forecast URL:", forecast_url)
    # 12-hour forecast
    forecast_response = requests.get(forecast_url, headers=headers)
    if forecast_response.status_code == 200:
        forecast_data = forecast_response.json()
        forecast_periods = forecast_data["properties"]["periods"]
        # Convert to DataFrame
        forecast_df = pd.DataFrame(forecast_periods)
        # Precipitation parsing - only output the number
        forecast_df["probabilityOfPrecipitation"] =
forecast_df["probabilityOfPrecipitation"].apply(
```

```
            lambda x: x["value"] if x and x["value"] is not None else 0
        )
        # Generate a new date column that increments every two rows
        start_date = datetime.now().date()
        forecast_dates = [
            (start_date + timedelta(days=i // 2)).strftime("%Y-%m-%d")
            for i in range(len(forecast_df))
        ]
        forecast_df.insert(0, "forecastDate", forecast_dates)
        # Columns to keep
        forecast_df = forecast_df[[
            "forecastDate", "name", "temperature", "shortForecast",
            "probabilityOfPrecipitation", "windSpeed", "windDirection",
            "isDaytime", "detailedForecast"
        ]]

        print("\n12-hour Forecast with Additional Features:")
        display(forecast_df)
    else:
        print("Error fetching 12-hour forecast data")
else:
    print("Error fetching information")
```

12-hour Forecast URL: https://api.weather.gov/gridpoints/TBW/90,104/forecast

12-hour Forecast with Additional Features:

	forecastDate	name	temperature	shortForecast	probabilityOfPrecipitation	windSpeed	windDirection	isDaytime	detailedForecast
0	2024-11-20	This Afternoon	80	Chance Showers And Thunderstorms	50	9 mph	SW	True	A chance of showers and thunderstorms. Mostly ...
1	2024-11-20	Tonight	55	Slight Chance Rain Showers then Partly Cloudy	20	7 to 12 mph	NW	False	A slight chance of rain showers before 7pm. Pa...
2	2024-11-21	Thursday	67	Sunny	0	12 mph	NNW	True	Sunny. High near 67, with temperatures falling...
3	2024-11-21	Thursday Night	48	Clear	0	8 mph	NW	False	Clear, with a low around 48. Northwest wind ar...
4	2024-11-22	Friday	68	Sunny	0	6 to 13 mph	NW	True	Sunny, with a high near 68. Northwest wind 6 t...
5	2024-11-22	Friday Night	44	Clear	0	5 to 8 mph	NNW	False	Clear, with a low around 44. North northwest w...
6	2024-11-23	Saturday	66	Sunny	0	7 mph	NNW	True	Sunny, with a high near 66. North northwest wi...
7	2024-11-23	Saturday Night	46	Clear	0	1 to 5 mph	N	False	Clear, with a low around 46. North wind 1 to 5...
8	2024-11-24	Sunday	72	Sunny	0	2 mph	ENE	True	Sunny, with a high near 72. East northeast win...
9	2024-11-24	Sunday Night	50	Clear	0	2 mph	ENE	False	Clear, with a low around 50.
10	2024-11-25	Monday	78	Sunny	0	2 mph	ENE	True	Sunny, with a high near 78.
11	2024-11-25	Monday Night	55	Clear	0	2 mph	ENE	False	Clear, with a low around 55.
12	2024-11-26	Tuesday	79	Sunny	0	3 mph	NNE	True	Sunny, with a high near 79.
13	2024-11-26	Tuesday Night	58	Clear	0	3 mph	N	False	Clear, with a low around 58.

```
import requests

import pandas as pd

from datetime import datetime, timedelta

# API for Lakeland - chosen rough center of Lakeland Electrical coverage map

lat, lon = 28.05, -81.96

points_url = f"https://api.weather.gov/points/{lat},{lon}"

# Define headers for the API request

headers = {

    "User-Agent": "myweatherapp (contact@myweatherapp.com)"

}

# Fetch the grid information

response = requests.get(points_url, headers=headers)

if response.status_code == 200:

    data = response.json()

    forecast_hourly_url = data["properties"]["forecastHourly"]

    grid_url = data["properties"]["forecastGridData"]

    print("Hourly Forecast URL:", forecast_hourly_url)
```

```
    print("Grid Forecast URL:", grid_url)

    # Grid data for relative humidity and dew point

    grid_response = requests.get(grid_url, headers=headers)

    if grid_response.status_code == 200:

        grid_data = grid_response.json()

        # Relative Humidity

        if "relativeHumidity" in grid_data["properties"]:

            rh_data = pd.DataFrame(grid_data["properties"]["relativeHumidity"]["values"])

            rh_data["startDate"] = rh_data["validTime"].str.split("T").str[0]

            rh_data["startTime_hour"] = rh_data["validTime"].str.extract(r"T(\d{2})")[0]

            rh_data = rh_data[["startDate", "startTime_hour", "value"]]

            rh_data.rename(columns={"value": "relativeHumidity"}, inplace=True)

        else:

            rh_data = pd.DataFrame(columns=["startDate", "startTime_hour",
"relativeHumidity"])

        # Dew Point (if available)

        if "dewpoint" in grid_data["properties"]:

            dewpoint_data = pd.DataFrame(grid_data["properties"]["dewpoint"]["values"])

            dewpoint_data["startDate"] = dewpoint_data["validTime"].str.split("T").str[0]

            dewpoint_data["startTime_hour"] =
dewpoint_data["validTime"].str.extract(r"T(\d{2})")[0]

            dewpoint_data["dewpoint"] = dewpoint_data["value"].apply(

                lambda x: round((x * 9 / 5) + 32, 1) if pd.notna(x) else None

            )

            dewpoint_data = dewpoint_data[["startDate", "startTime_hour", "dewpoint"]]

        else:
```

```
        dewpoint_data = pd.DataFrame(columns=["startDate", "startTime_hour",
"dewpoint"])

    # Hourly forecast

    hourly_response = requests.get(forecast_hourly_url, headers=headers)

    if hourly_response.status_code == 200:

        hourly_data = hourly_response.json()

        hourly_periods = hourly_data["properties"]["periods"]

        # Convert to DataFrame

        hourly_df = pd.DataFrame(hourly_periods)

        # Precipitation parsing - only output the number

        hourly_df["probabilityOfPrecipitation"] =
hourly_df["probabilityOfPrecipitation"].apply(

            lambda x: x["value"] if x and x["value"] is not None else None

        )

        # Split startTime into startDate and startTime

        hourly_df["startDate"] = hourly_df["startTime"].str.split("T").str[0]

        hourly_df["startTime_hour"] = hourly_df["startTime"].str.split("T").str[1].str[:2]

        # Filter data for 1-week forecast

        one_week_ahead = (datetime.now() + timedelta(days=7)).strftime("%Y-%m-
%d")

        hourly_df = hourly_df[hourly_df["startDate"] <= one_week_ahead]

        # Columns to keep

        hourly_df = hourly_df[[

            "startDate", "startTime_hour", "temperature", "shortForecast",

            "probabilityOfPrecipitation", "windSpeed", "windDirection"

        ]]
```

Merge Relative Humidity and Dew Point with Hourly Forecast

hourly_with_grid = pd.merge(hourly_df, rh_data, on=["startDate", "startTime_hour"], how="left")

hourly_with_grid = pd.merge(hourly_with_grid, dewpoint_data, on=["startDate", "startTime_hour"], how="left")

print("\nHourly Forecast with Relative Humidity and Dew Point:")

display(hourly_with_grid)

else:

print("Error fetching hourly forecast data")

else:

print("Error fetching grid forecast data:", grid_response.status_code, grid_response.text)

else:

print("Error fetching grid information")

Hourly Forecast URL: https://api.weather.gov/gridpoints/TBW/90,104/forecast/hourly

Grid Forecast URL: https://api.weather.gov/gridpoints/TBW/90,104

Hourly Forecast with Relative Humidity and Dew Point:

	startDate	startTime_hour	temperature	shortForecast	probabilityOfPrecipitation	windSpeed	windDirection	relativeHumidity	dewpoint
0	2024-11-20	12	77	Chance Showers And Thunderstorms	46	9 mph	S	NaN	NaN
1	2024-11-20	13	79	Chance Showers And Thunderstorms	34	9 mph	S	NaN	67.0
2	2024-11-20	14	79	Chance Showers And Thunderstorms	30	9 mph	SSW	91.0	70.0
3	2024-11-20	15	79	Chance Showers And Thunderstorms	26	9 mph	W	90.0	72.0
4	2024-11-20	16	79	Slight Chance Rain Showers	20	9 mph	WNW	89.0	73.0
...	...	...	...	...	...	...	...	...	...
151	2024-11-26	19	68	Clear	0	2 mph	NNW	53.0	59.0
152	2024-11-26	20	67	Clear	0	2 mph	NNW	52.0	NaN
153	2024-11-26	21	65	Clear	0	2 mph	N	54.0	58.0
154	2024-11-26	22	64	Clear	0	2 mph	N	58.0	NaN
155	2024-11-26	23	62	Clear	0	2 mph	N	65.0	59.0

Dewpoint values are inconsistent - not every hour has that forecast

Filter the DataFrame for rows where dewpoint is not NaN

filtered_data = hourly_with_grid[hourly_with_grid["dewpoint"].notna()]

Check if any data exists

if not filtered_data.empty:

    print("Rows with non-NaN Dewpoint values:")

    display(filtered_data)

else:

    print("No rows with non-NaN Dewpoint values.")

Rows with non-NaN Dewpoint values:

	startDate	startTime_hour	temperature	shortForecast	probabilityOfPrecipitation	windSpeed	windDirection	relativeHumidity	dewpoint
1	2024-11-20	13	79	Chance Showers And Thunderstorms	34	9 mph	S	NaN	67.0
2	2024-11-20	14	79	Chance Showers And Thunderstorms	30	9 mph	SSW	91.0	70.0
3	2024-11-20	15	79	Chance Showers And Thunderstorms	26	9 mph	W	90.0	72.0
4	2024-11-20	16	79	Slight Chance Rain Showers	20	9 mph	WNW	89.0	73.0
5	2024-11-20	17	76	Slight Chance Rain Showers	20	8 mph	WNW	91.0	74.0
...	...	...	...	...	...	...	...	...	...
146	2024-11-26	14	78	Sunny	0	3 mph	NE	79.0	58.0
147	2024-11-26	15	78	Sunny	0	3 mph	N	72.0	60.0
151	2024-11-26	19	68	Clear	0	2 mph	NNW	53.0	59.0
153	2024-11-26	21	65	Clear	0	2 mph	N	54.0	58.0
155	2024-11-26	23	62	Clear	0	2 mph	N	65.0	59.0

HISTORICAL DATA

- NOAA.gov - https://www.ncdc.noaa.gov/cdo-web/webservices/v2

- API:

    - Email: jjackson3465@floridapoly.edu

    - Token: idqwrJvPPATpYzveCAWajOVtSQGvFKpR

Station info: https://www.ncdc.noaa.gov/cdo-web/datasets/GHCND/locations/FIPS:12105/detail - Polk County

import requests

import pandas as pd

from datetime import datetime, timedelta

NOAA CDO API base URL and token

base_url = "https://www.ncdc.noaa.gov/cdo-web/api/v2/data"

token = "idqwrJvPPATpYzveCAWajOVtSQGvFKpR" # NOAA API - DON'T MAKE PUBLIC

```
# Calculate dates

current_date = datetime.now().strftime("%Y-%m-%d")

week_timeframe = (datetime.now() - timedelta(weeks=1)).strftime("%Y-%m-%d")     #
Change weeks value to however many you want

# Define headers with the token

headers = {"token": token}

# Function to fetch data for a specific datatype

def fetch_data(datatype):

    params = {

        "datasetid": "GHCND",         # Daily summaries dataset

        "locationid": "FIPS:12105",  # Polk County FIPS Code

        "startdate": week_timeframe,

        "enddate": current_date,

        "datatypeid": datatype,       # Specific datatype (e.g., TMIN, TMAX, PRCP)

        "limit": 1000,

        "units": "standard"

    }

    response = requests.get(base_url, headers=headers, params=params)

    if response.status_code == 200:

        data = response.json()

        return pd.DataFrame(data.get("results", []))

    else:

        print(f"Error fetching {datatype} data:", response.status_code, response.text)

        return pd.DataFrame()
```

```
# Fetch data for each datatype

tmin_data = fetch_data("TMIN")

tmax_data = fetch_data("TMAX")

prcp_data = fetch_data("PRCP")

# Combine all data into a single DataFrame

historical_df = pd.concat([tmin_data, tmax_data, prcp_data], ignore_index=True)

# Clean and aggregate the data

# Ensure the value column is numeric

historical_df["value"] = pd.to_numeric(historical_df["value"], errors="coerce")

# Pivot table to create separate columns for TMIN, TMAX, and PRCP averages

aggregated = historical_df.pivot_table(

    index="date",              # Group by date

    columns="datatype",         # Use datatype (TMIN, TMAX, PRCP) as columns,
preciptation measured in inches, temps measured in F

    values="value",             # Aggregate the "value" column

    aggfunc="mean"              # Calculate the mean for each date and datatype

).reset_index()

# flatten the column index - for some reason messes up column names if don't do this

aggregated.columns.name = None

aggregated.rename(columns={"TMIN": "tempMin", "TMAX": "tempMax", "PRCP":
"precipitation"}, inplace=True)

# Make lists of the values averaged to help see what's going on

# group by the different datatypes

grouped_values = historical_df.groupby(["date",
"datatype"])["value"].apply(list).reset_index()
```

```
# pivot so they match

lists_pivot = grouped_values.pivot(index="date", columns="datatype", values="value").reset_index()

# rename columns

lists_pivot.columns.name = None

lists_pivot.rename(columns={"TMIN": "tempMin_values", "TMAX": "tempMax_values", "PRCP": "precipitation_values"}, inplace=True)

# Merge dfs

final_df = pd.merge(aggregated, lists_pivot, on="date", how="left")

# Clean up the `date` column to keep only the YYYY-MM-DD part

final_df["date"] = final_df["date"].str[:10]

# Round the `precipitation` column to 2 decimal places

final_df["precipitation"] = final_df["precipitation"].round(2)

final_df["tempMax"] = final_df["tempMax"].round(1)

final_df["tempMin"] = final_df["tempMin"].round(1)

# output

print("Aggregated Daily Weather:")

display(final_df)
```

Appendix F: Meeting Notes Throughout the Semester:

Meeting notes – 10/15/2024

Start:

Mr. Dammer leading things – will put us in touch with others specialized in areas.

Goal: look at not just planning, but revenue mapping – they're rolling out automated smart grid metering system – We're looking at phase 2 – WHAT CAN WE DO WITH MODELING – instead of using assumed values, use real values from data we have from the smart grid system – make a better Lakeland based data set – revenue modeling, load modeling data modeling, etc – make a Lakeland standard

We're driving this – get in and ask him question – generate models and provide him with info – he's not going to be hands-on with this – wants an outside mind viewpoint.

- Multiyear capstone project – hand torch off to next year

Plexos model – load modeling – utility industry is moving into a data driven industry – wants us to take a look at the industry to see if we want to work in it for the future – bring him any frustrations as well – good to know from fresh set of eyes

Key behind this is the process – no real end game goal – framework is satisfactory goal is to learn from it and build off of it. Good to find failures and successes.

They use SQL – we can use whatever

Smart grid team - RNI web based – data comes in through web platform – all that data is ported into an ORACLE meter system – can do some analytics with it but not much - where the SQL comes from.

Separate system for billing – handles revenue information – goal to identify meters out of the smart grid data then pair it with the Kubra (spelling)

They have test environments to play around with most likely

Will get us with the rates people who are using the data and they can help us.

What track is Lakeland on and where can we go

Building new powerplant that will get them to their capacity needs – currently buying some power from Orlando

Florida = vertically integrated state – every area needs to have enough power – not crazy variable swings like you'd see from some other states

Obligated to have 15% + max peak usage on hand at all times

GOAL: REVENUE MODELING THAT IS SUITED FOR LAKELAND SPECIFICALLY

- Developer comes to them looking for X – they need to know how much they will get back Y
    - Want X = Y at the very least to get money back – revenue needs to line up with expenses since government

<u>Notes – 10/23/24</u>

- More on the lines of Revenue (are we getting what we need to be getting)
    - Currently not forecasting well – revenue budget to actuals is way off
    - Pair us with customer service team – Kelly Parish = responsible for rates/forecasting
        - Revenue/load forecasting – currently missing
        - Energy efficiency has plateaued – know what to expect when building a house
            - Won't change until next technological breakthrough
            - Only thing can't currently predict well is human nature – e.g. how much they run AC, etc
            - Currently costs ~30,000$ for new residential homes over 30 years
                - To make back money – need to average ~115$/month over the span
                - Want to break down all the rate customers to get these numbers
                - Kelly can break this down
            - Look at publicly available information for rate classes in Lakeland (for Lakeland electric)
                - Ways_to_pay, tariff – contains the information we'll need to model
                - How many of each type of customer class would be ideal for modeling – 140,000 total meters – most are basic

residential – would be nice to know a good breakdown for modeling

- Want to look at usages from the different rate classes – what are we really selling in terms of power and are we getting back the investment?
    - Are there models we can implement that will help the projections teams?

- Take existing usage/billing data + knowledge of what is coming to make a better crystal ball that specifically works for Lakeland Electric – Use as much science as possible to minimize guessing (make guesses more educated)
    - Some tools being implemented into smart grid but might not necessarily translate to revenue.
    - Only industry that is actively trying to get people to use less of their product (electricity) since usage patterns matter so much

- Can we reasonably craft a time of use rate – e.g from 3-5PM it gets more expensive to try to persuade people to change their usage patterns. – 2 ways to change behavior = carrots and sticks
    - Look at other utility systems – e.g. Austin, Sacramento, salt river project (cutting edge) – look up American Power public Association (APPA) – Lakeland is the $3^{rd}/4^{th}$ largest in Florida – Jacksonville/Orlando good to look at as well (Florida Municipal Electric Authority FMPA)
    - Try to find break-even then normalize the curve (change human behavior) to paint the picture for the people who build policy – identify the patterns about the behavior

- End goal = what info do we have, what can we do with it, what do we do with it, what should we do with it, and what do we do with it (as data scientists/students)

- Next Steps – Kelly Parish – Manager of forecasting/rates – job is to do what we're doing at a higher level – we're doing more of a deep dive
    - She'll let us know what the 'industry standards' they use are so we can utilize and compare with it
    - Smart Grid Team = actual usage data (total Kw/h)

            - Don't know job status of users – might be helpful to know if they're remote vs non-remote or retired vs non, etc.
                - Does demographic matter for revenue or are we better aggregating at a much higher level?
                - Hold various variables constant across similar homogenous areas and identify how changes to others affect those when keeping certain things constant – e.g keep home size constant and look at houses all with same size but different other features
                - Look at house sales and see if the change in users affected things drastically or didn't
                - What would happen to Lakeland Electric if there are drastic changes in energy usage patterns by users – either more efficiently or not
    - FRED Datasets? Any useful ones for Lakeland for us?

Meeting Notes 11-22-24

<u>Meeting with Kelly</u>

- Finance background – 30 years, recently moved into forecasting/budgeting - over the forecasting group – they do the revenue forecasting as well.
    - She will be more on the financial side
    - Dena Prim -
    - Piece that might need other groups -
    - ITRON – linear regression – load data (hourly), economic data, MOOdEY's Analytics, bieber, and woods and poole for economic forecasts, - weather normalization (to predict on a normal weather day what the usage would be)
    - Segmented subdivision for a smaller scale work
    - Dena produces a monthly forecast by class – output = excel
    - Load model: Has past actual data and forward 10 years -
    - Try to pull a new subdivision – commercial
    - Tariff document – can get on their website that defines where customers fall

    - 10 year is filed with Public service commission as well
    - Adoption rate lower in polk city since it's lower income
- Things in the model:
    - Load, sales data (customer stats report – consumption/demand)
    - Billing days
    - Class ratios
    - Price forecasts
    - Economics (moodeys, beaver, woods and poole)
    - Weather
    - Solar – degradation factor
    - ALL PULLED INTO ITRON
- Any proprietary -
    - She doesn't think so
- Getting us – rates tariff, customer statistics report, load model, billing cycle dates, 10 year site plan
- Work with engineering to see if they can get us any more
- Model doesn't handle peak vs off-peak hours – smart meters handle that and then feed into the system
- Traditionally – LE is a winter-peaking unit – going to the same year round probably wouldn't change much
- They send their models to ITRON to be evaluated.
    - Can share the model summary pages with us.

Nick's Version of the 11-22-24 Meeting Notes:

- Kelly Parish uses Itron Software that does linear regression analysis for them. They put in everything form load data to power load data and economic data and analytics.
- They use Moody Analytics for forecasts

- They use weather normalization to predict on a normal weather day what their usage would be.

- Itron deals with these things. Dena produces a monthly forecast and then Kelly takes it and puts it as an excel spreadsheet

- Revenue Model is pretty ugly but they can supply you with what you need. They have hourly revenue data.

- Dena's load forecast has customer amount, kilowatt hours used and kilowatt hours demand. There will be past data in there and forward data in there

- Kelly has commercial, large industrial, and general service demand.

- Go to Documents and then other information to fin the tariff document

- The FRCC is public service

- The FRCC may have information on their site

- Lakeland Electric does an annual meeting with FRCC to discuss EVs and sort etc.

- Polk county is a pretty low income compared to the rest of the state so that's why the adoption rates are low

- The items that go into the itron model: load, sales data with customer account consumption and demand, billing days and billing cycles, they use class ratios, price forecasts, moody's beaver woods and pull, weather, solar information about megawatts.. degradation factor.. pulled into the itron model that pulls their forecast.

- Information that cannot be provided: They are apart of the sunshine law which means that everything is public information. But there isn't really anything proprietary.

- Moody's may or may not be strict about sending data but Lakeland electric will share it via excel docs etc..

- Revenue Model, Monthly Load Forecast, Rates Tariff, Customer Statistics Report, Billing Cycle, Copy of 10 year site plan.. (all of this is specific to Lakeland electric)

- Kelly will work with engineering to find additional information

- Tariff document addresses single residential information. For instance, a residential duplex considering a person an individual customer.

- Once you get to 50 kilowatts you are put into a demand rate. About 400 you get large.. extra large etc

- Florida Public Service and FRCC are good places to research any definitions you need but can always ask Kelly Parish
- Engineers are off on Fridays but work four 10 hour days
- Peak vs Off-Peak hours model: The model does not account for this. It is the metering system (smart grid) that feeds into the billing system.
- Lakeland Electric has been a winter peaking unit from 6am to 10am and 6pm to 10pm
- Going through the same period year round would probably be the same.
- Summer peak is around noon or 1 to 9pm and it can vary with resident zones.
- Dena runs the stats on the model but Itron is the one that sends the model to them
- They can probably share the model summary page about lights, residential, and exogenous models and they look at all of those together and then send models to itron containing all the variables that go into the models as far as weather etc.
- Smart Grid will help you more with the meter capability.

Additional Email correspondence occurred in the meantime – can show if needed.

Appendix G: GANTT Chart for Spring 2025