Power Outage Prediction

Can Weather Patterns and Customer Energy Usage Predict Forced Outages?

Outline

1. Context
2. Data Sources
3. Main Features
4. Data Wrangling
5. EDA
6. Experiments
7. Conclusion

Problem of Interest

*Daily Average Usage per US home - 30.04 kw

Daily Average PJM Outage - 2,812 mw

Daily Average PJM Mid Atlantic Usage - 1,013,195 mw

Outage to usage ratio - 0.27%

*Quantity of homes equivalence = (2,812 mw * 1,000) / 30.06 kw = 93,609 homes

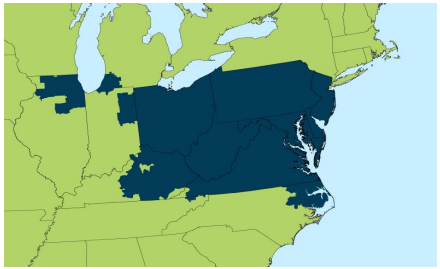
*Comparison to Baltimore Population - 15%

***Calculations are found on the "Calculations" slide**

Data Sources

Obtained 4.5 years (2015-2020) from each data source.

- PJM Interconnection coordinates the movement of wholesale electricity in the Mid-Atlantic U.S.
- Energy usage
- Outage (MW's of energy lost)



- National Oceanic and Atmospheric Association API
- Retrieved weather data from weather stations in PJM service territory
    - MD, VA, DE, PA, DC, NJ, OH, WV, MI, IL, NC, IN, KY, TN = **131 stations**

Types of Energy Outages

- Planned
    - Scheduled removal from service for inspection
    - Predetermined duration* 30 days prior
- Maintenance
    - Scheduled removal from service to perform repairs
    - Shorter than planned outages* 2 to 3 days prior
- Forced
    - Immediate reduction in output by reason of an emergency or unanticipated failure

Raw Data

PJM Outage (Daily Time Series)

forecast_execution_date_ept	forecast_date	region	total_outages_mw	planned_outages_mw	maintenance_outages_mw	forced_outages_mw	execution_date
2019-04-01	2019-04-01	Mid Atlantic - Dominion	28432	20519	6328	1585	4012019

PJM Usage (Hourly Time Series)

datetime_beginning_utc	datetime_beginning_ept	nerc_region	mkt_region	zone	load_area	mw	is_verified
2015-09-03 06:00:00	2015-09-03 02:00:00	RFC	MIDATL	JC	JC	2429.369	True

NOAA (Daily Time Series)

	station_dt_key	date	station	PRCP	SNOW	SNWD	AWND	TMAX	TMIN	elevation	mindate	maxdate	latitude	name
116245	GHCND:USC00466442_2017-12-01	2017-12-01	GHCND:USC00466442	0.17	0.0	0.0	NaN	59.0	32.0	205.7	1892-09-01	2020-07-19	40.5261	NEW CUMBERLAND L D, OH US

Merge

forecast_execution_date_ept	forecast_date	region	total_outages_mw	planned_outages_mw	maintenance_outages_mw	forced_outages_mw
2020-07-01	2020-07-01	Mid Atlantic - Dominion	7577	0	3786	3791

```
import datetime as dt
def convert_dt(df,col_name=None,year=True,month=True):
    df[col_name] = pd.to_datetime(df[col_name])

    df['plot_date'] = pd.to_datetime(df[col_name].dt.strftime('%Y-%m-%d'))
    df['agg_date'] = df[col_name].dt.strftime('%m%d%Y')

    if year:
        df['year'] = df['agg_date'].str[-4:]
    if month:
        df['month'] = df['agg_date'].str[0:2]
        df = df.astype({'month': 'int64'})

    df = df.drop(columns='agg_date')
    return df
```

```
def get_mkt_region(mkt):

    if mkt == 'MIDATL':
        return  'Mid Atlantic - Dominion'

    elif mkt == 'SOUTH':
        return 'Mid Atlantic - Dominion'

    elif mkt == 'WEST':
        return 'Western'
```

datetime_beginning_ept	datetime_beginning_utc	nerc_region	mkt_region	zone	load_area	mw	is_verified
2020-07-01 01:00:00	2020-07-01 05:00:00	RFC	MIDATL	AE	AECO	1027.360	False

station_dt_key	date	station	PRCP	SNOW	SNWD	AWND	TMAX	TMIN	mkt_region	region
GHCND:USC00444128_2020-07-01	2020-07-01	GHCND:USC00444128	0.00	0.0	0.0	NaN	83.0	59.0	MIDATL	Mid Atlantic - Dominion

Merge Continued

PJM Outage

	forecast_execution_date_ept	forecast_date	region	total_outages_mw	planned_outages_mw	maintenance_outages_mw	forced_outages_mw	plot_date
5029	2020-07-01	2020-07-01	Mid Atlantic - Dominion	7577	0	3786	3791	2020-07-01

PJM Usage (Group by Sum)

		mw
region	plot_date	
Mid Atlantic - Dominion	2015-01-01	1049743.647

NOAA Weather (Group by Mean)

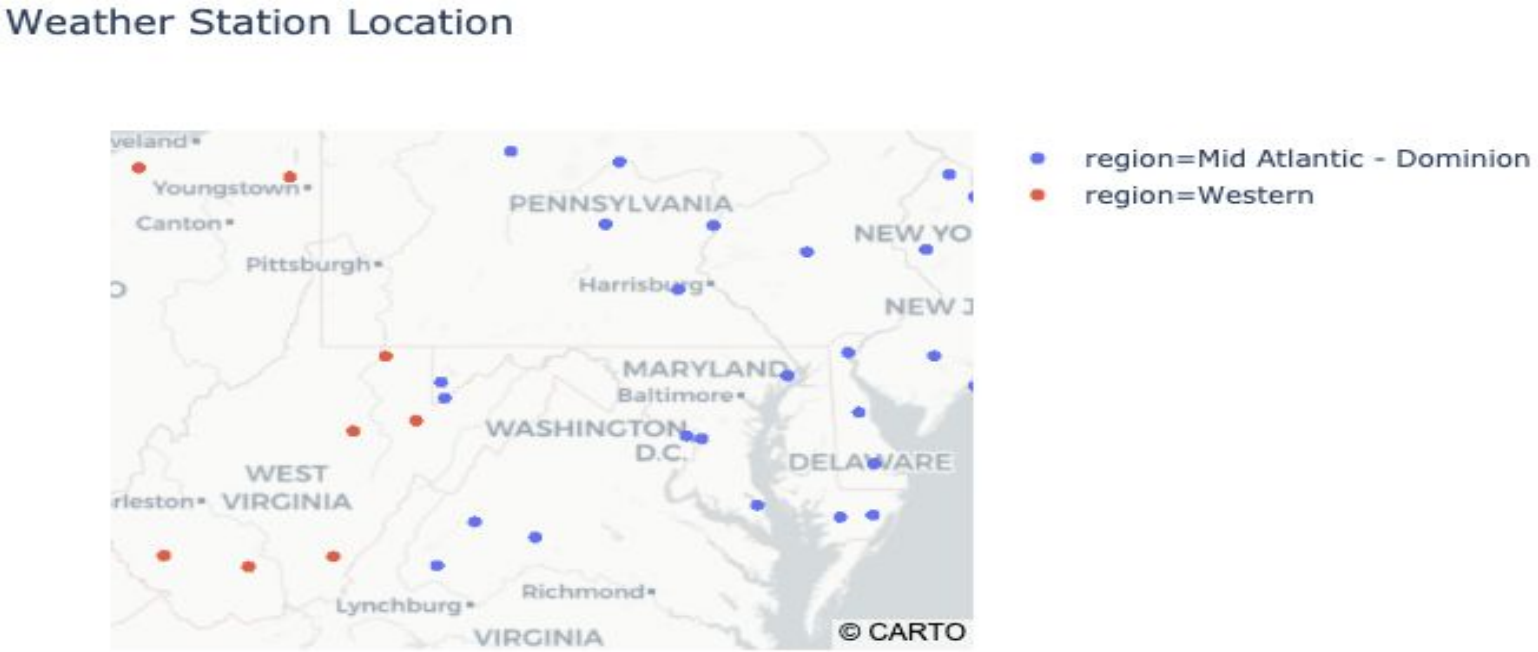
		PRCP	SNOW	SNWD	AWND	TMAX	TMIN
region	plot_date						
Mid Atlantic - Dominion	2015-01-01	0.000000	0.000000	0.000000	10.725000	33.288889	18.911111

Merged Dataset

plot_date	region	forced_outages_mw	Usage_mw	PRCP	SNOW	SNWD	TMAX	TMIN
2015-05-26	Mid Atlantic - Dominion	1472.0	1077796.932	0.004255	0.0	0.0	83.595745	58.266667

Feature	Definition	Scale	Data Source
Plot Date	Time Series of when an Actual outage was recorded	Daily Time Series	PJM Outage Dataset
Region	Region where outage loss occurred	-	PJM Outage Dataset
Forced Outages	Immediate reduction in output by reason of an emergency or unanticipated failure	MW	PJM Outage Dataset
Usage	Load in MW	MW	PJM Usage Dataset
PRCP	Precipitation	Inches	NOAA
SNOW	Snowfall	Inches	NOAA
SWND	Snow Depth	Inches	NOAA
TMAX	Maximum Temperature	Fahrenheit	NOAA
TMIN	Minimum Temperature	Fahrenheit	NOAA

Map of Weather Stations



Project Focus

Can weather patterns and energy usage data be used to predict future forced outages in the Mid Atlantic - Dominion PJM Territory?

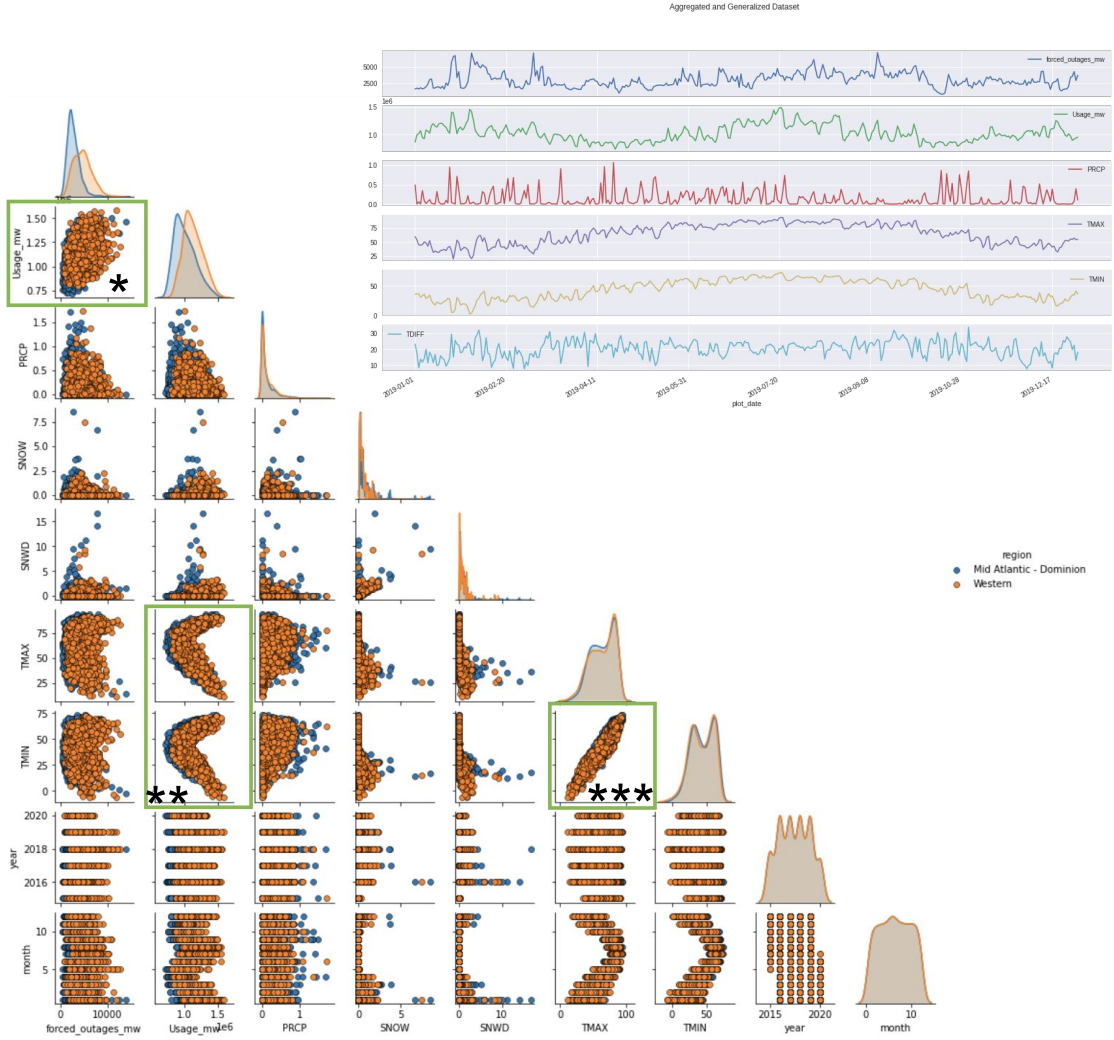
Note: Forced outages are recorded in terms of MW lost during the outage

Feature Comparison

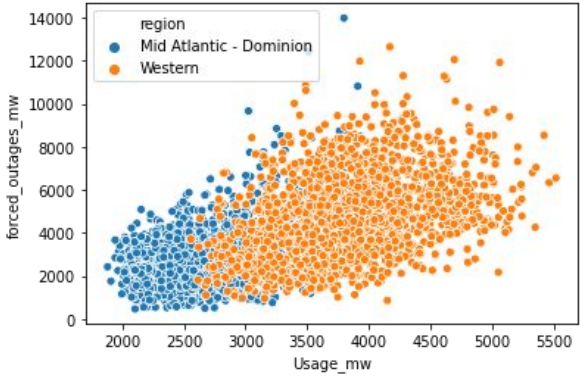
*Moderately strong positive correlation between usage and forced outages

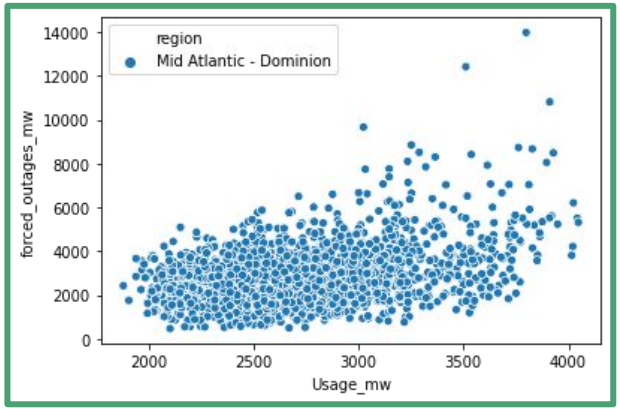
**Strong seasonal relationship between temperature and usage

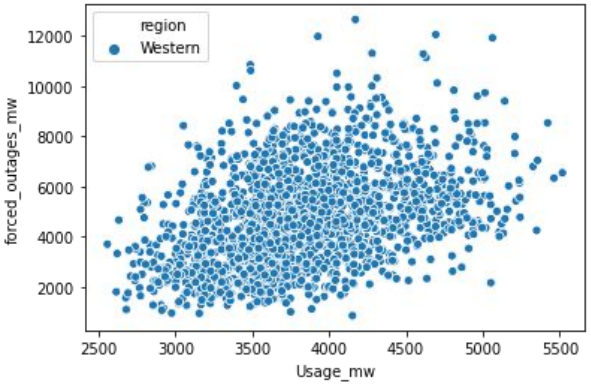
*** Strong positive relationship between TMIN and TMAX



Deciding on the Region to Predict Forced Outages



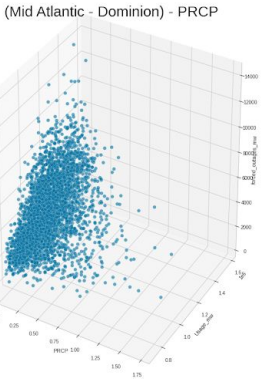


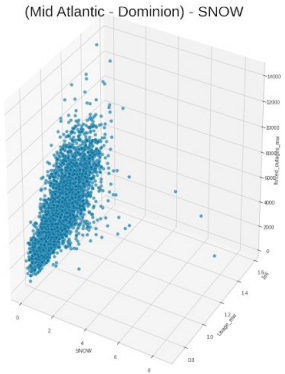


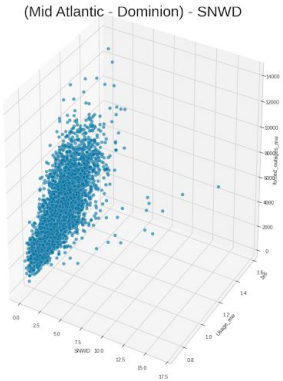
EDA of NOAA and Energy Data

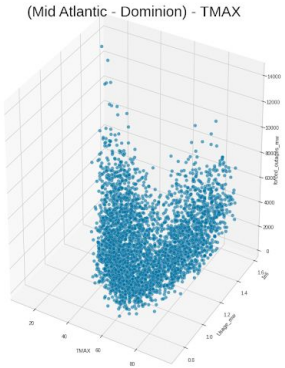
- Extreme temperatures (min and max) show that Forced outages are more common when temperatures are at extreme **lows** or **highs**.

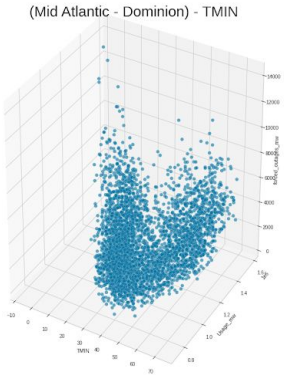
- Indicates that precipitation and snow data will not be good indicators of outages on the regional level.











Experiments - Single Output Time-Series Models

1. Predict TMAX with TMIN, PRCP, SNOW, and SNWD
2. Predict Forced Outages MW with Usage MW
3. Predict Forced Outages MW with Forced Outages MW
4. Predict Forced Outages MW with all features

Models and Parameters

Models

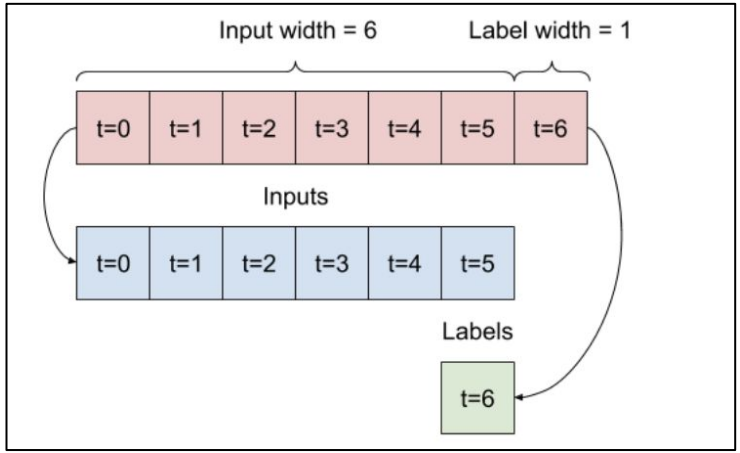
1. Baseline - Predicting "No change"
2. Linear
3. Dense
4. Multi-step Dense
5. Convolution Neural Network
6. Recurrent Neural Network -
    Long Short Term Memory (LSTM)

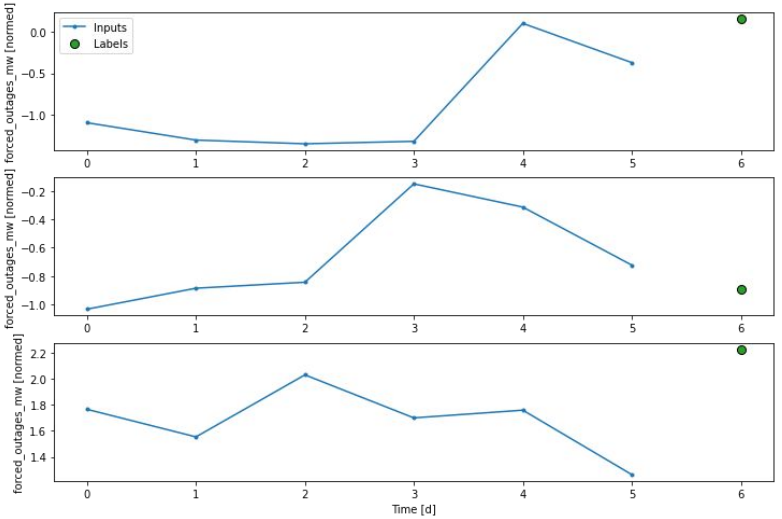
Parameters

1. Target
2. Features
3. Feature Criteria
    4. Window Size - (7, 1, 1)
        a. 6 inputs
        b. 1 label
        c. 1 shift (label)

Example: Single Output - Window Size

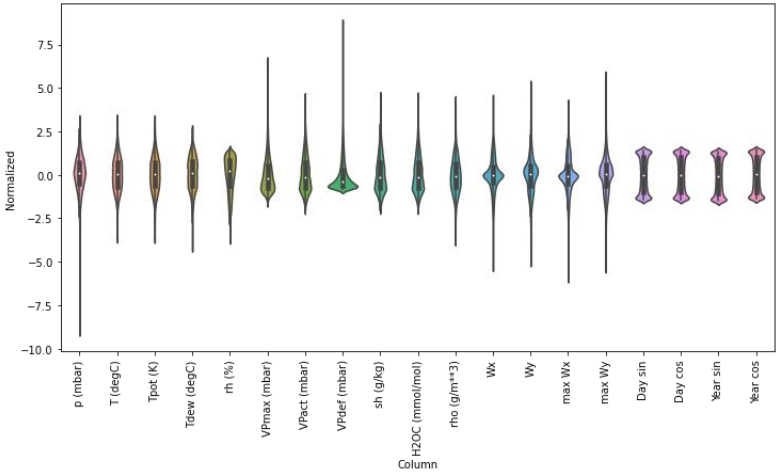
Window Size = 7

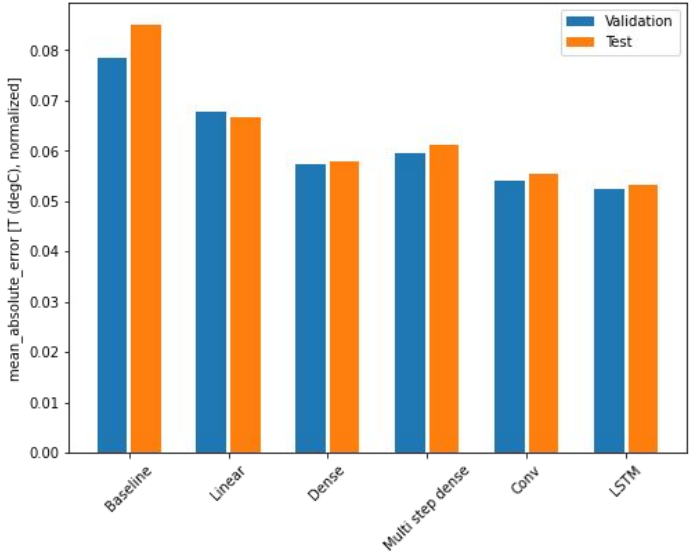




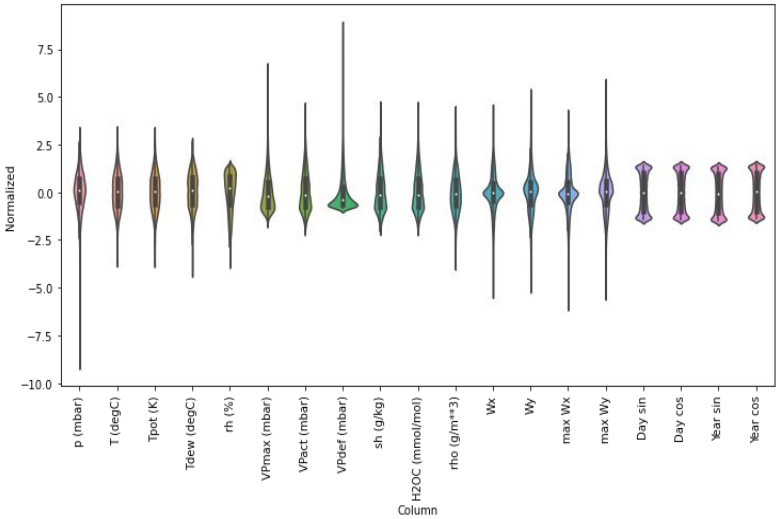
Source: https://www.tensorflow.org/tutorials/structured_data/time_series#linear_model

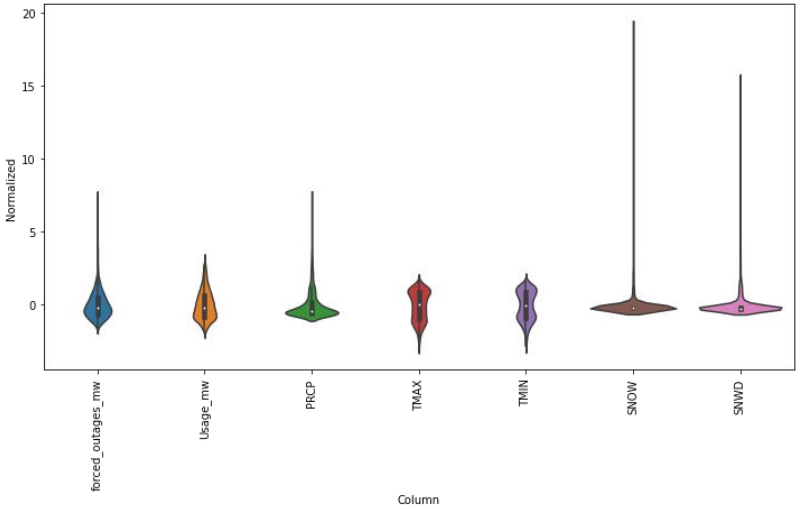
Demo Shape and Model Performance



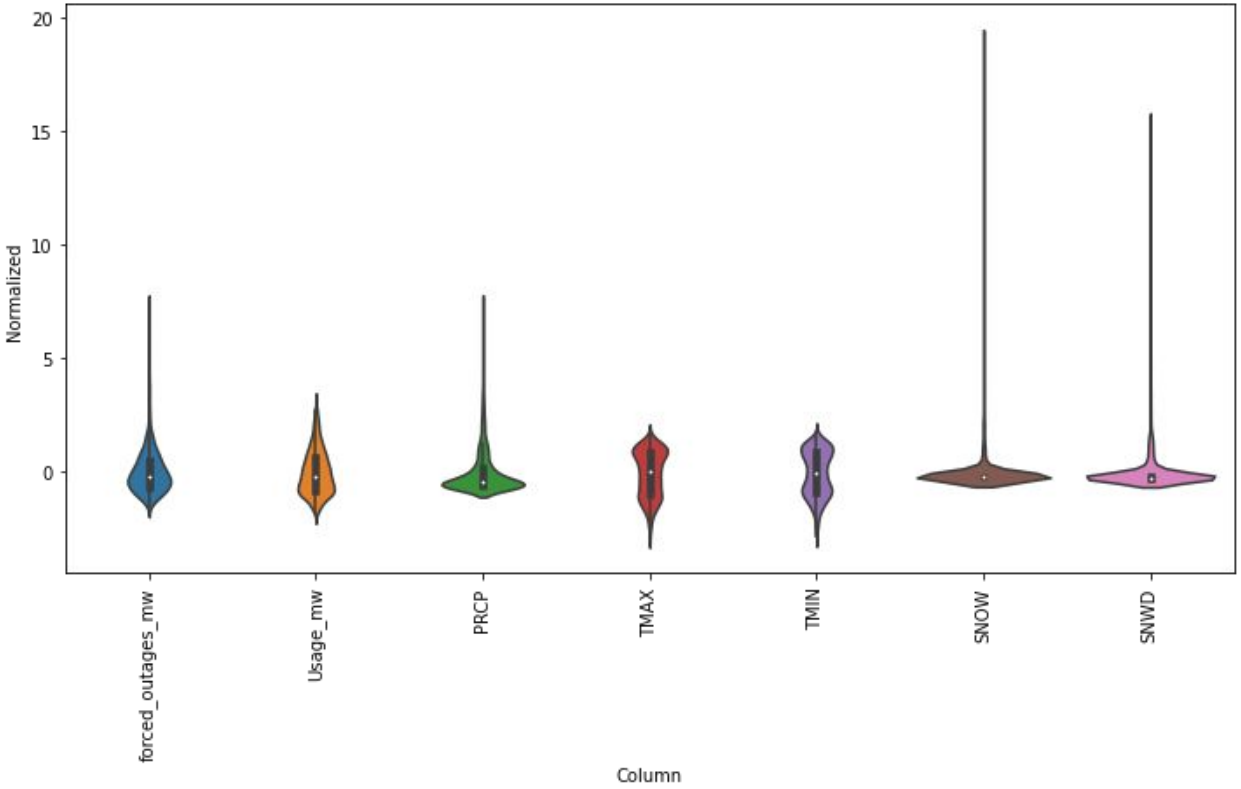


Shape Comparison

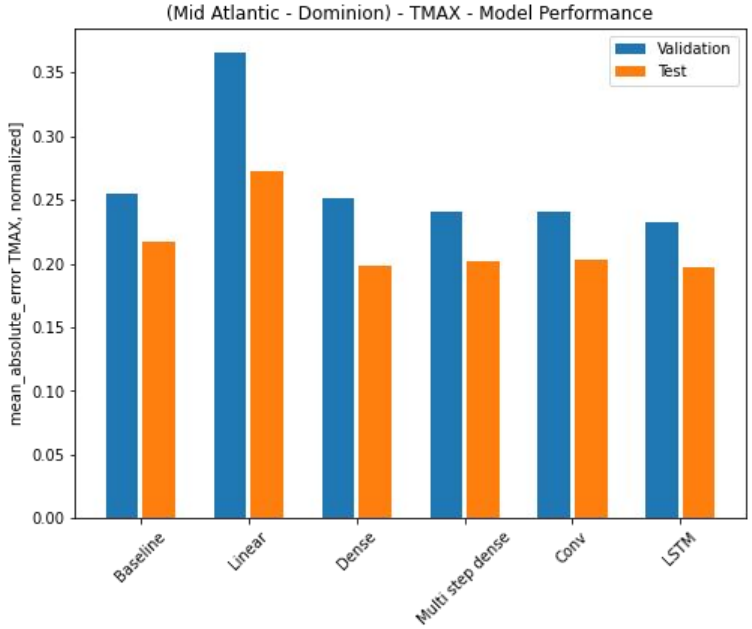


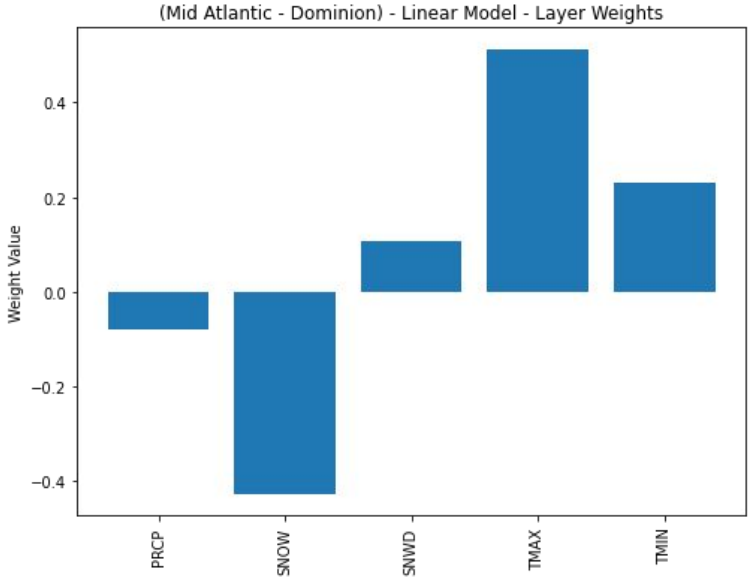


The Shape of Our Data

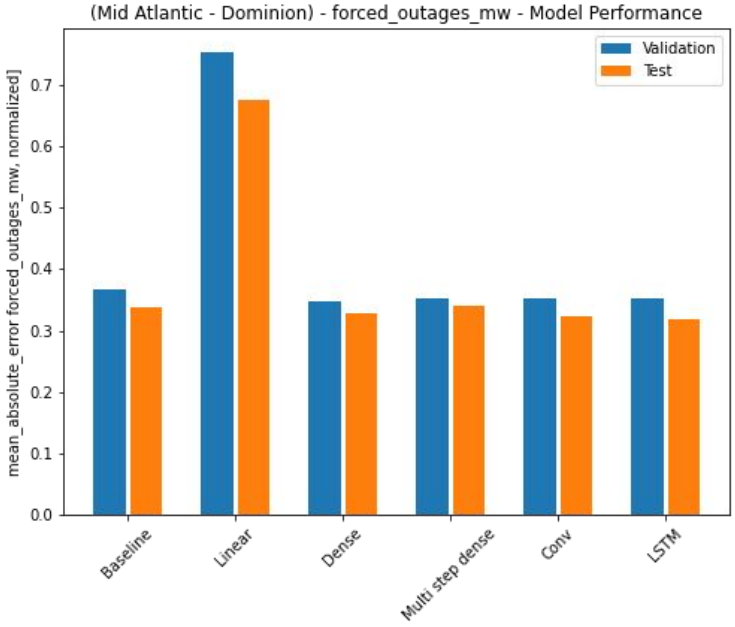


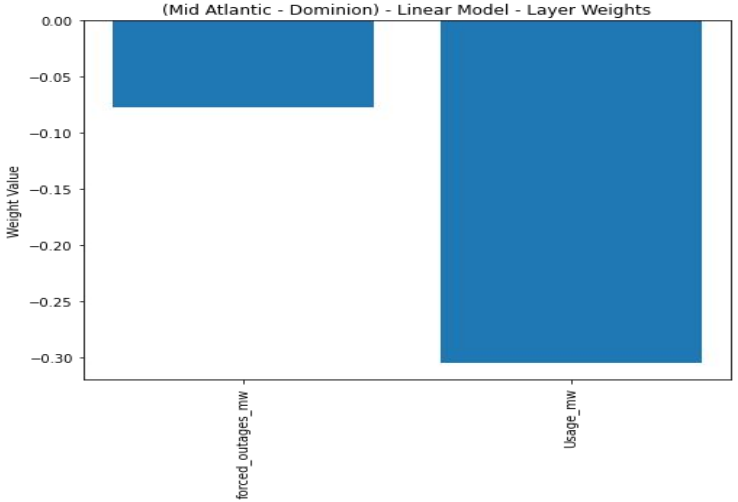
Predict TMAX with all Weather Features



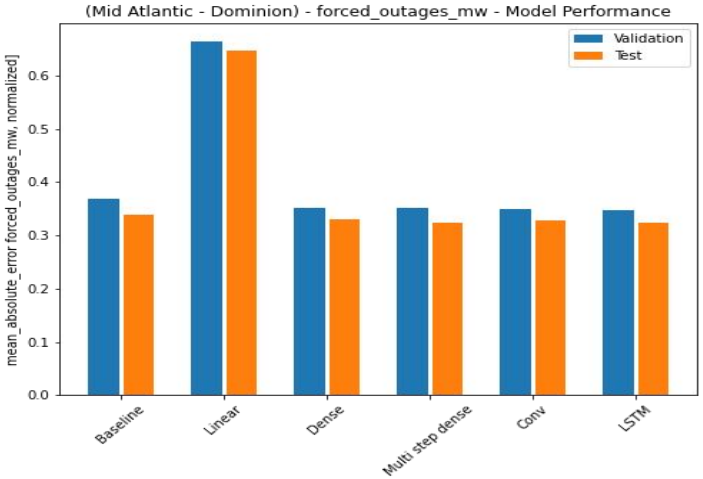


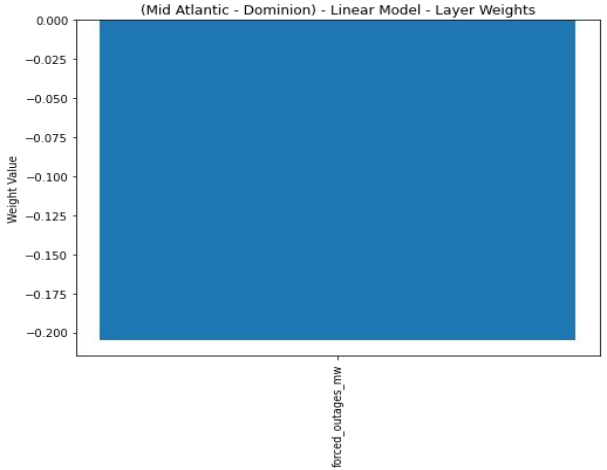
Predict Forced Outages using Forced Outage and Usage MW



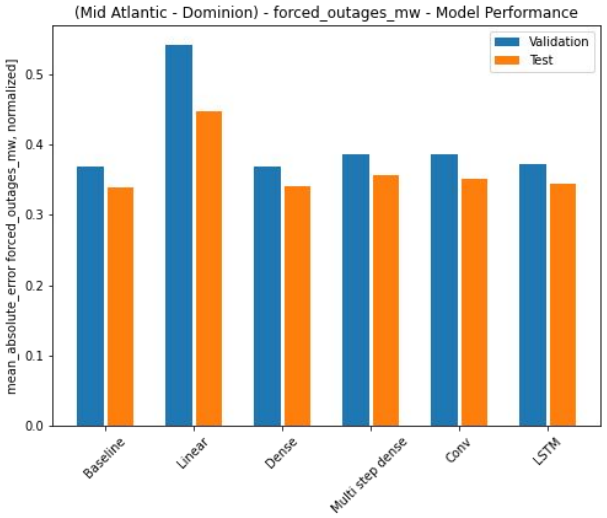


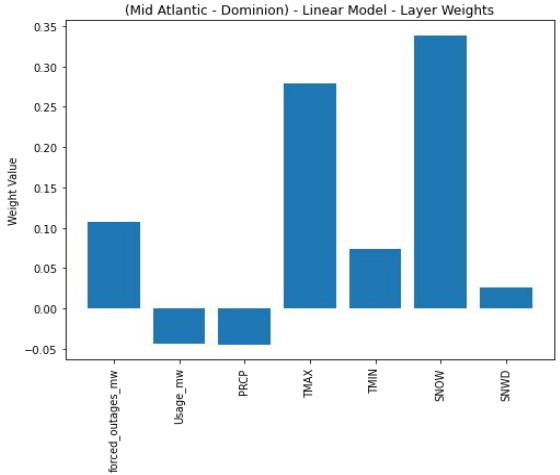
Predict Forced Outages with Forced Outages





Predict Forced Outages MW with All Features





Conclusion

Model

-The linear model performs worse than the baseline model for each

-There is not a clear indicator that the Dense, Multi step dense, Conv or LSTM models are significantly more effective than the baseline model

Combining Datasets

-Aggregating usage and weather data to match the outage regional classification may have over generalized the weather attributes and the energy usage

-The Mid-Atlantic region had over 30 weather stations

-The weather attributes from these stations were averaged to get a regional value

Next Steps

1. Investigate outlier impact on the models
2. Conduct experiments to see if seasonality will improve model performance

Thank you

Sources

Types of Energy Outages

https://www.pjm.com/~/media/documents/manuals/m10.ashx

PJM Territory

https://www.pjm.com/about-pjm/who-we-are/territory-served.aspx#:~:text=PJM%20Interconnection%20coordinates%20the%20movement,and%20the%20District%20of%20Columbia.

U.S Residential Usage

https://www.eia.gov/tools/faqs/faq.php?id=97&t=3#:~:text=How%20much%20electricity%20does%20an,about%20914%20kWh%20per%20month.

TensorFlow Tutorial

https://www.tensorflow.org/tutorials/structured_data/time_series

Calculations

*https://www.eia.gov/tools/faqs/faq.php?id=97&t=3#:~:text=How%20much%20electricity%20does%20an,about%20914%20kWh%20per%20month.

1. Average Day kw = (Average kwh per month / days_in_month)

Average Day kw = (*10,972 / 365)

Average Day kw ≃ 30.06 kw

2. Quantity of homes equivalence = (2,812 mw * 1,000) / 30.06 kw = 93,546 homes
3. Baltimore City's Population = 619,493
    a. 93,546/ 619,493 = 15%

Predict TMAX with all Weather features