Final Project - Hydro Power Generation and Streamflow Analysis

Kaitlyn Elliott and Jon Nichols

Rationale and Research Questions

Hydroelectric power accounts for 52 percent of renewable electricity generation and 7 percent of total electricity generation in the US. These plants are a critical clean energy resource for grid operators and utilities due to their operational flexibility, low maintenance costs, and non-intermittent production. However, hydroelectric generation is susceptible to climate change, and changes in temperature, precipitation patterns, glacial melt, and the frequency of extreme weather events such as floods and droughts may have a direct negative effect on electricity production. Thus, the impact of climate change on the variability and availability of streamflow is critical for current hydroelectric generation and future operational forecast plans.

There are three types of hydropower plants: impoundment facilities that use a dam to store water, run-of-river facilities that generate electricity by channeling water through a turbine and do not use a reservoir, and pumped storage plants. For our analysis, we investigated the impact of climate change on both impoundment facilities and run-of-river facilities. We expect that run-of-river facilities will experience more significant declines in generation given their unique vulnerability to variations in streamflow. In general, lower stream discharges reduce the amount of electricity a hydropower plant can generate.

We examine 20 years of streamflow and electric generation data for 9 facilities across the US and assess the impact of climate change on stream flows and consequently on the hydropower generation of these plants. We specifically investigate the monotonic trends for streamflow and electric generation at each hydropower plant, and then analyze whether stream discharge can explain the variation in observed generation data. The results obtained will allow us to better understand the relationship between hydropower generation and streamflow, while also identifying hydropower plants that are vulnerable to climate change.

Background and Methods

In this project we set out to explore the relationship between hydro power generation and stream generation. To do this we looked at power generation data from the EPI and stream gage data from the USGS. For generation data we selected generation sites/dams so that half were from the western United States and half from the East. Stream gage data was selected upstream from dams and reservoirs to get an accurate representation of stream flow over time. Both power generation data and stream flow was analyzed for seasonality and monotonic trends using a seasonal mann kendall test. This was to help understand if generation was increasing or decreasing as a result of stream flow. Although outside factors could also contribute increases or decreases in power generation.

Set Up

Data Retrieval

Hydroelectric generation data was collected from the EIA, and streamflow data was collected from the USGS. Generation data was obtained through the EIA's Electricity Data Browser, which is an interactive query tool. Nine separate queries - one for each hydropower plant- were used to download the data needed for the analysis. For each query, we used the date range January 2001 through December 2021. Data was downloaded as a "Chart (CSV)", which presented four rows of descriptive information and then two columns for the Month and "All primemovers - All fuels (ALL) megawatthours". The descriptive info of each generation dataset was cleaned in Excel to streamline data import into R. Further information on the generation data is provided in the table below.

Table 1. Data Information for Hydropower Generation

Detail	Description
Data Source	EIA Net Generation Data
Retrieved From	https://www.eia.gov/beta/electricity/data/browser/#/plant/10550/?freq=M&pin=
Variables Used	Month, "All fuels (ALL) megawatthours"
Date Range	January 2001 - December 2021

Table 2. Data Information for Stream Discharge

Detail	Description
Data Source	USGS Stream Discharge Data
Retrieved From	USGS using readNWISdata
Variables Used	Date, Discharge
Date Range	January 2000 - December 2021

Data Wrangling

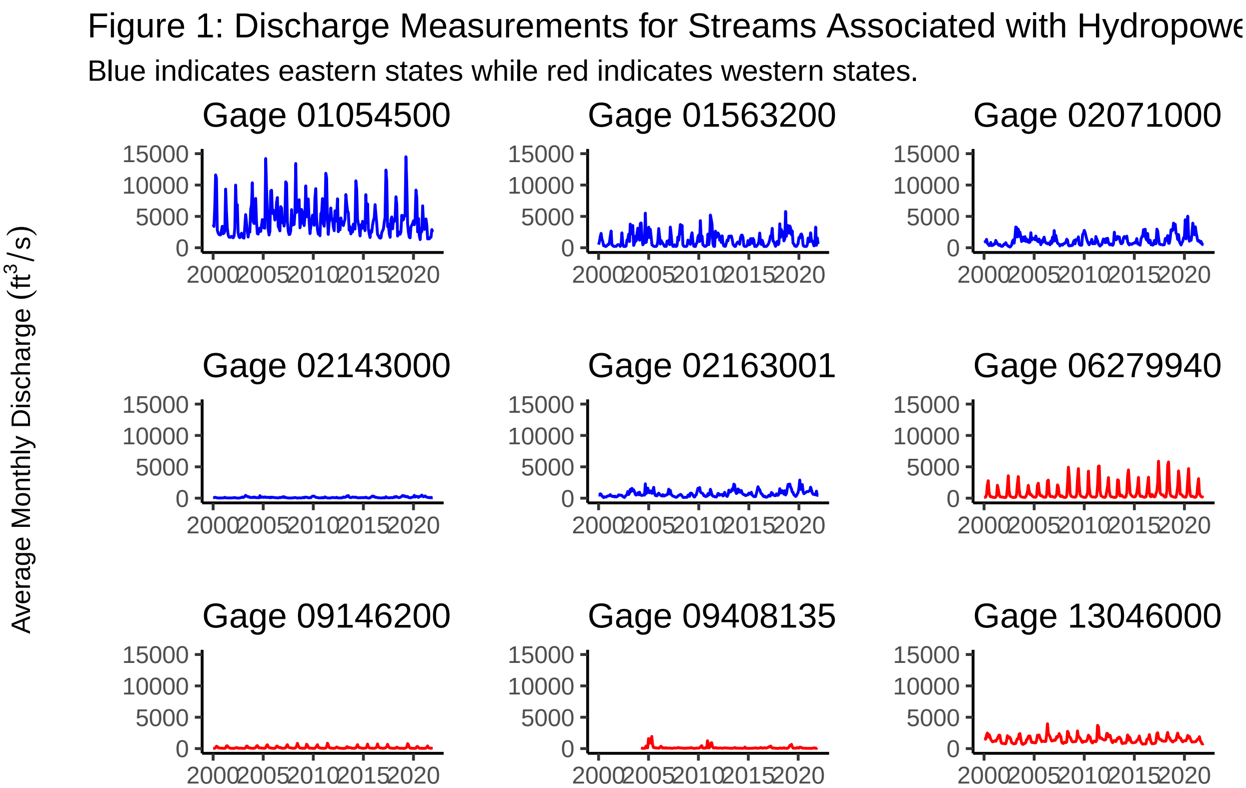
After streamflow data was imported, the column containing discharge data was mutated and titled "Discharge". Next, three columns were selected: Discharge, Datetime, and Site Number. The Datetime was mutated to produce a Year column and a Month column. We then grouped the dataset by the year and month, and created an "Average Monthly Discharge" column. This process was repeated for each stream gauge dataset. Additionally, an arbitrary day - the first of each month - was added to each imported dataset to enable proper formatting.

After hydroelectric generation data was imported, the Date column had to be wrangled and reformatted to allow for further analysis. Month and year columns were added to each dataset, generation data was transformed to be numeric, the date column was converted to a date variable in R, and data was arranged in ascending fashion. Empty or "NA" values were then dropped to have complete sets of data for each hydroelectric plant.

After importing and wrangling hydroelectric generation and streamflow data separately, we used a left join function to combine the datasets for each specific facility. Given that some generation data was missing for the 20-year period, we used left join to match the streamflow data to the generation data that did exist for each site.

Discharge Visualization

Following our initial wrangling, we created plot grids to visualize the streamflow and generation data. Both visualizations compared sites in the Western US to sites in the Eastern US. Streamflow data for each site was graphically visualized using the cowplot plot_grid function. Figure 1 displays the discharge measurements for the selected stream gauges over the last 20 years.



Hydroelectric Generation Visualization

Hydroelectric data for each power plant was also graphically visualized using the cowplot plot_grid function. Figure 2 displays hydropower generation for the selected power plants over the last 20 years.

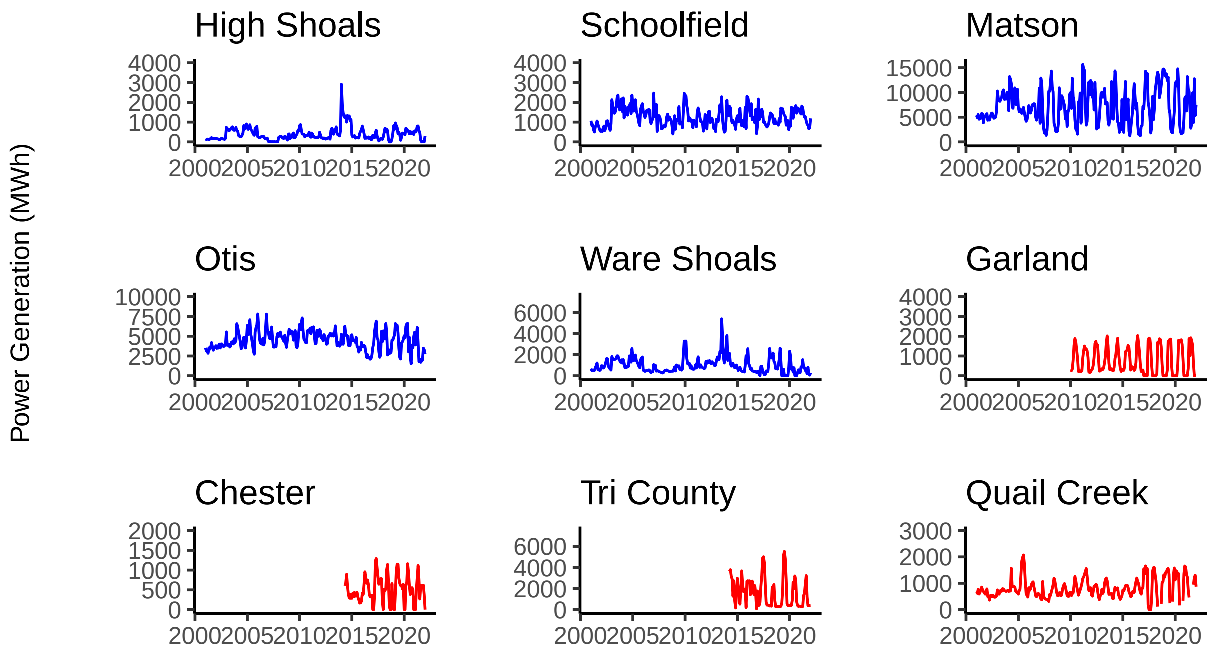
```
## Warning: Removed 108 row(s) containing missing values (geom_path).
```

```
## Warning: Removed 160 row(s) containing missing values (geom_path).
```

```
## Warning: Removed 159 row(s) containing missing values (geom_path).
```

Figure 2: Hydropower Generation for Sampled Power Plants

Blue indicates eastern states while red indicates western states.

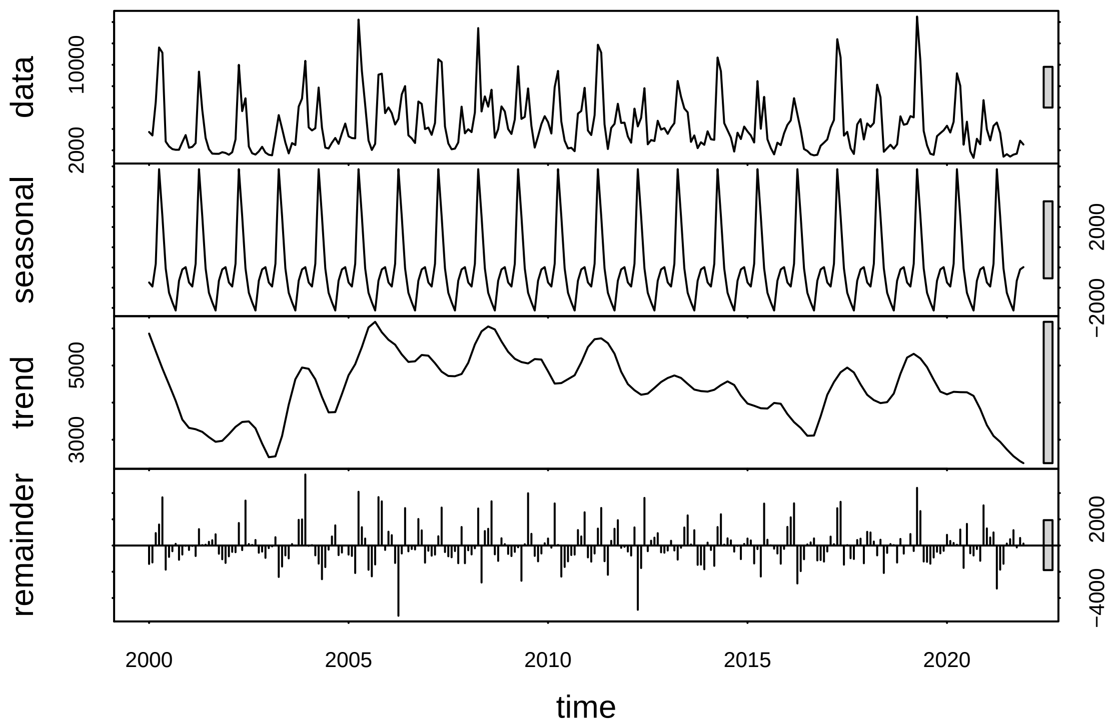


Data Analysis

Question 1: Are there trends in the streamflow data?

After visualizing the data, we investigated trends in streamflow. We created a time series for each stream gauge, and then decomposed each time series into its seasonal, trend, and remainder components. We plotted each decomposed time series, and observed that each site displayed strong seasonal trends in streamflow.

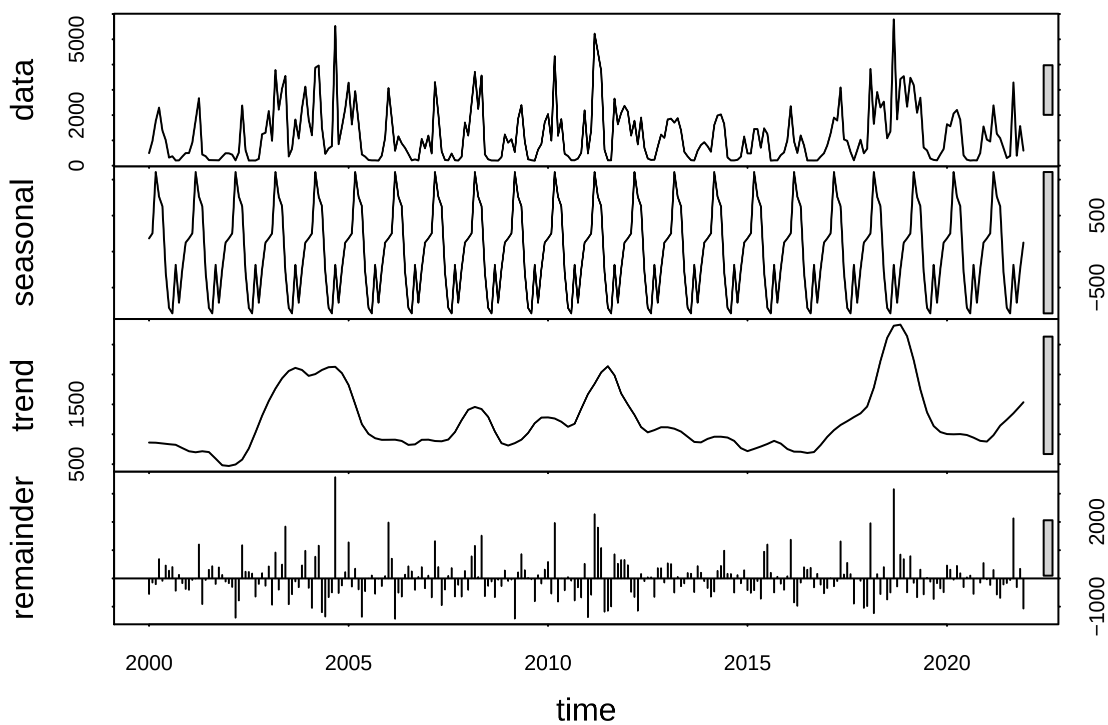
We ran a Seasonal Mann-Kendall test on each time series to test whether there was a monotonic trend in streamflow over time. The results are displayed in Table 1. Six of the nine stream gauges demonstrated trends. Four gauges, which were all located in the East, showed increasing trends. Two gauges, which were located in the West, showed decreasing trends. The results are consistent with broad climate predictions regarding drier conditions in the Western US and wetter conditions in the Eastern US.




```
##
##  Seasonal Mann-Kendall trend test (Hirsch-Slack test)
##
## data:  gage_01054500_ts
## z = -0.57794, p-value = 0.5633
## alternative hypothesis: true S is not equal to 0
## sample estimates:
##     S  varS
##   -72 15092
```

```
##
##  Seasonal Mann-Kendall trend test (Hirsch-Slack test)
##
## data: gage_01054500_ts
## alternative hypothesis: two.sided
##
## Statistics for individual seasons
##
## H0
##                      S   varS    tau      z Pr(>|z|)
## Season 1:   S = 0   49 1257.7  0.212  1.354 0.175896
## Season 2:   S = 0   65 1257.7  0.281  1.805 0.071127 .
## Season 3:   S = 0   39 1257.7  0.169  1.072 0.283935
## Season 4:   S = 0  -21 1257.7 -0.091 -0.564 0.572782
## Season 5:   S = 0  -15 1257.7 -0.065 -0.395 0.693012
## Season 6:   S = 0  -45 1257.7 -0.195 -1.241 0.214713
## Season 7:   S = 0   21 1257.7  0.091  0.564 0.572782
```

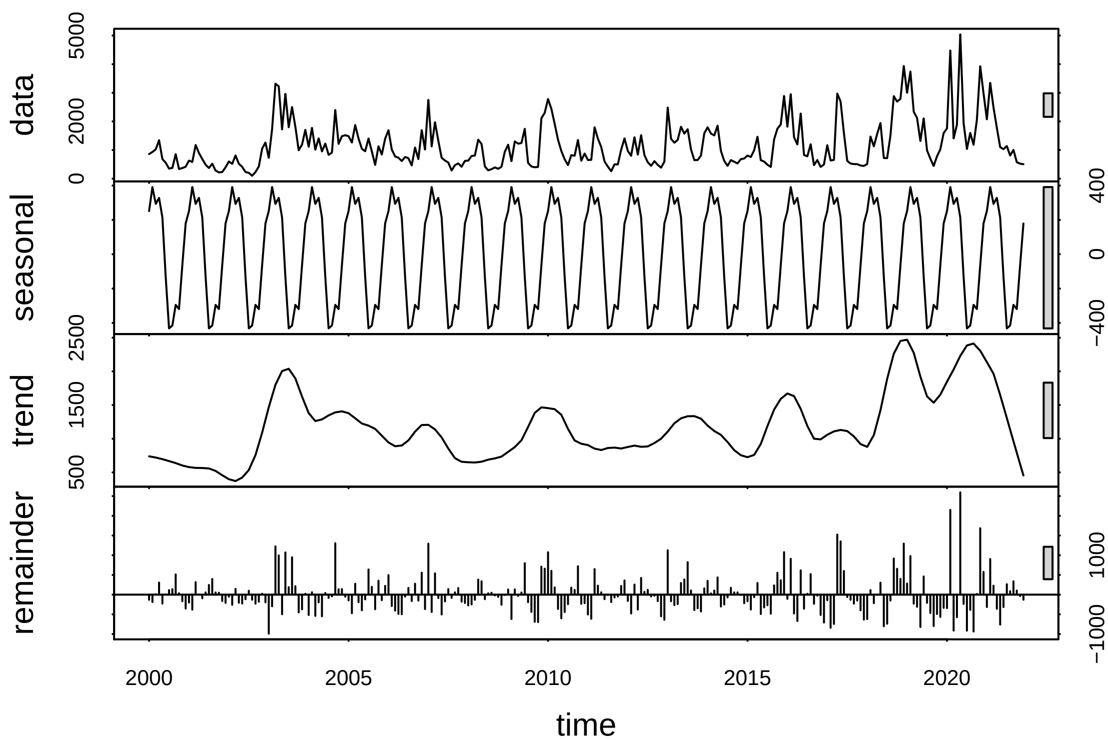
```
## Season 8:   S = 0  -29 1257.7 -0.126 -0.790 0.429795
## Season 9:   S = 0  -65 1257.7 -0.281 -1.805 0.071127 .
## Season 10:   S = 0 -19 1257.7 -0.082 -0.508 0.611760
## Season 11:   S = 0 -41 1257.7 -0.177 -1.128 0.259355
## Season 12:   S = 0 -11 1257.7 -0.048 -0.282 0.777959
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```




```
##
##  Seasonal Mann-Kendall trend test (Hirsch-Slack test)
##
## data:  gage_01563200_ts
## z = 1.4408, p-value = 0.1496
## alternative hypothesis: true S is not equal to 0
## sample estimates:
##      S  varS
##   178 15092
```

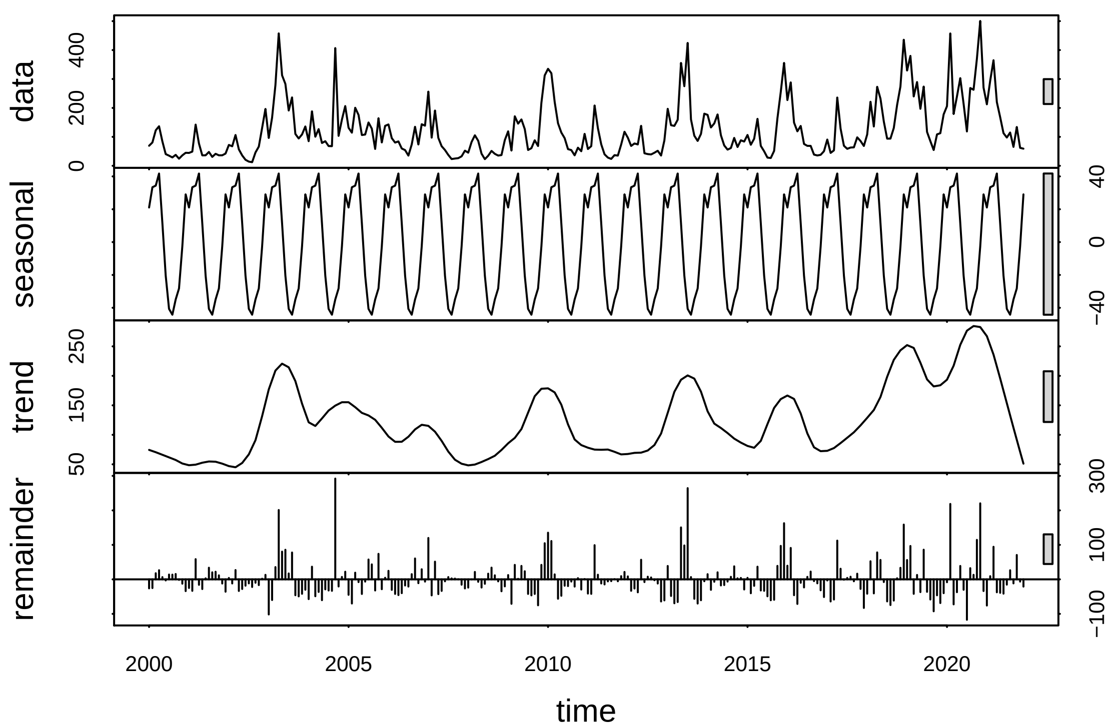
```
##
##  Seasonal Mann-Kendall trend test (Hirsch-Slack test)
##
## data: gage_01563200_ts
## alternative hypothesis: two.sided
##
## Statistics for individual seasons
##
## H0
```

```
##                         S   varS    tau      z Pr(>|z|)
## Season 1:   S = 0   -3 1257.7 -0.013 -0.056 0.955026
## Season 2:   S = 0   63 1257.7  0.273  1.748 0.080417 .
## Season 3:   S = 0   -1 1257.7 -0.004  0.000 1.000000
## Season 4:   S = 0  -25 1257.7 -0.108 -0.677 0.498564
## Season 5:   S = 0   27 1257.7  0.117  0.733 0.463469
## Season 6:   S = 0   45 1257.7  0.195  1.241 0.214713
## Season 7:   S = 0   45 1257.7  0.195  1.241 0.214713
## Season 8:   S = 0   -5 1257.7 -0.022 -0.113 0.910196
## Season 9:   S = 0   27 1257.7  0.117  0.733 0.463469
## Season 10:   S = 0  17 1257.7  0.074  0.451 0.651869
## Season 11:   S = 0  15 1257.7  0.065  0.395 0.693012
## Season 12:   S = 0 -27 1257.7 -0.117 -0.733 0.463469
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```



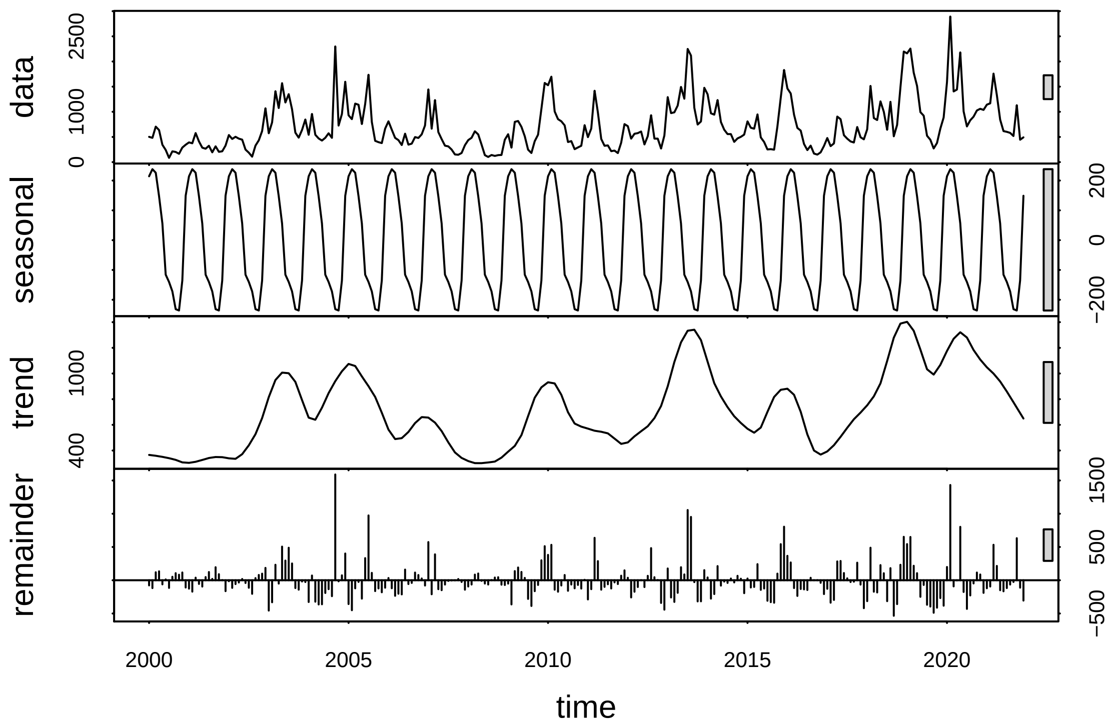

```
##
##  Seasonal Mann-Kendall trend test (Hirsch-Slack test)
##
## data:  gage_02071000_ts
## z = 5.8038, p-value = 0.000000006481
## alternative hypothesis: true S is not equal to 0
## sample estimates:
##     S  varS
##   714 15092

##
```

```
##  Seasonal Mann-Kendall trend test (Hirsch-Slack test)
##
## data: gage_02071000_ts
## alternative hypothesis: two.sided
##
## Statistics for individual seasons
##
## H0
##                         S   varS   tau     z  Pr(>|z|)
## Season 1:   S = 0   67 1257.7 0.290 1.861 0.0627352  .
## Season 2:   S = 0   77 1257.7 0.333 2.143 0.0321097  *
## Season 3:   S = 0   25 1257.7 0.108 0.677 0.4985644
## Season 4:   S = 0   79 1257.7 0.342 2.199 0.0278468  *
## Season 5:   S = 0  109 1257.7 0.472 3.045 0.0023239 **
## Season 6:   S = 0   61 1257.7 0.264 1.692 0.0906697  .
## Season 7:   S = 0   67 1257.7 0.290 1.861 0.0627352  .
## Season 8:   S = 0   61 1257.7 0.264 1.692 0.0906697  .
## Season 9:   S = 0   33 1257.7 0.143 0.902 0.3668796
## Season 10:   S = 0  69 1257.7 0.299 1.917 0.0551796  .
## Season 11:   S = 0  37 1257.7 0.160 1.015 0.3100460
## Season 12:   S = 0  29 1257.7 0.126 0.790 0.4297953
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```




```
##
##  Seasonal Mann-Kendall trend test (Hirsch-Slack test)
```

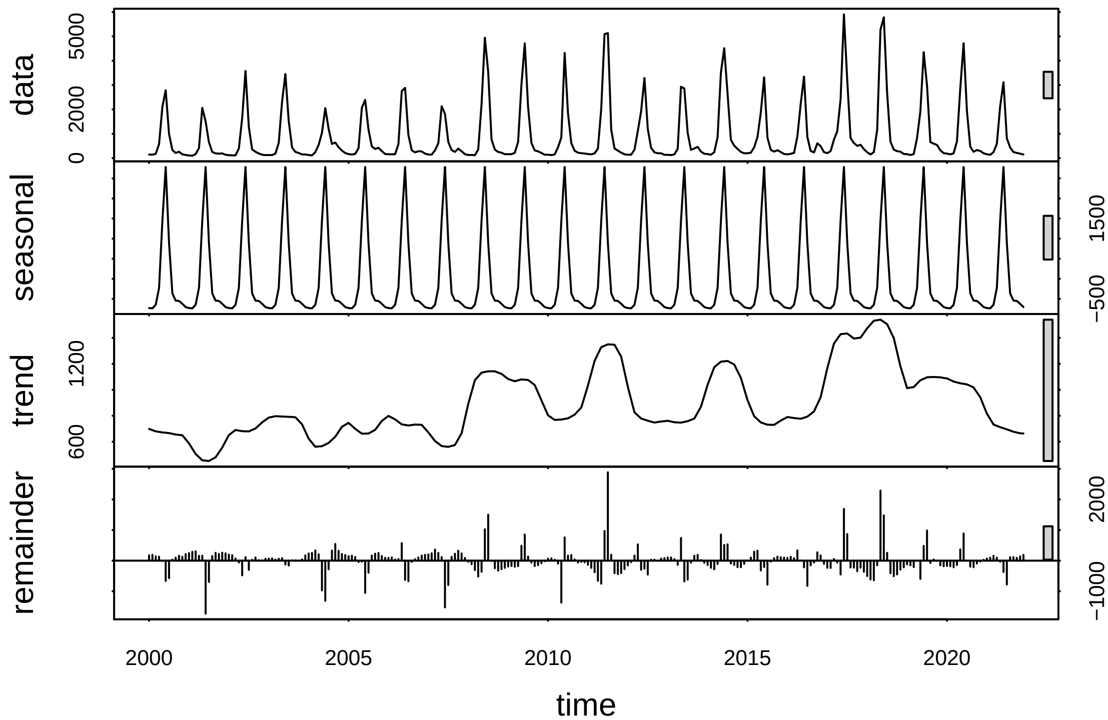
```
##
## data:  gage_02143000_ts
## z = 6.4388, p-value = 0.0000000001204
## alternative hypothesis: true S is not equal to 0
## sample estimates:
##     S  varS
##   792 15092

##
##  Seasonal Mann-Kendall trend test (Hirsch-Slack test)
##
## data: gage_02143000_ts
## alternative hypothesis: two.sided
##
## Statistics for individual seasons
##
## H0
##                     S   varS   tau     z  Pr(>|z|)
## Season 1:   S = 0  95 1257.7 0.411 2.651 0.0080348 **
## Season 2:   S = 0  73 1257.7 0.316 2.030 0.0423310  *
## Season 3:   S = 0  21 1257.7 0.091 0.564 0.5727823
## Season 4:   S = 0  99 1257.7 0.429 2.763 0.0057203 **
## Season 5:   S = 0  95 1257.7 0.411 2.651 0.0080348 **
## Season 6:   S = 0  71 1257.7 0.307 1.974 0.0483982  *
## Season 7:   S = 0  67 1257.7 0.290 1.861 0.0627352  .
## Season 8:   S = 0  71 1257.7 0.307 1.974 0.0483982  *
## Season 9:   S = 0  41 1257.7 0.177 1.128 0.2593549
## Season 10:   S = 0 83 1257.7 0.359 2.312 0.0207650  *
## Season 11:   S = 0 53 1257.7 0.229 1.466 0.1425687
## Season 12:   S = 0 23 1257.7 0.100 0.620 0.5350245
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```




```
##
##  Seasonal Mann-Kendall trend test (Hirsch-Slack test)
##
## data:  gage_02163001_ts
## z = 6.8946, p-value = 5.401e-12
## alternative hypothesis: true S is not equal to 0
## sample estimates:
##      S  varS
##    848 15092
```

```
##
##  Seasonal Mann-Kendall trend test (Hirsch-Slack test)
##
## data: gage_02163001_ts
## alternative hypothesis: two.sided
##
## Statistics for individual seasons
##
## H0
##                       S   varS   tau     z  Pr(>|z|)
## Season 1:   S = 0   91 1257.7 0.394 2.538 0.0111547  *
## Season 2:   S = 0   95 1257.7 0.411 2.651 0.0080348 **
## Season 3:   S = 0   53 1257.7 0.229 1.466 0.1425687
## Season 4:   S = 0  111 1257.7 0.481 3.102 0.0019237 **
## Season 5:   S = 0  117 1257.7 0.506 3.271 0.0010718 **
## Season 6:   S = 0   71 1257.7 0.307 1.974 0.0483982  *
## Season 7:   S = 0   61 1257.7 0.264 1.692 0.0906697  .
```

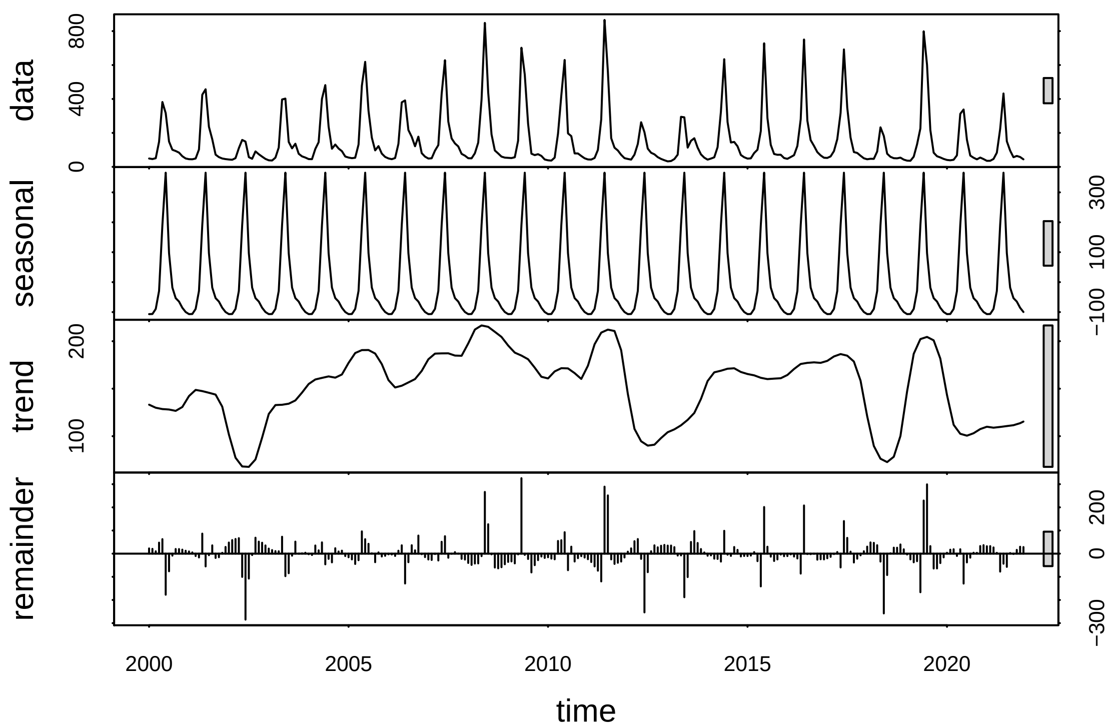
```
## Season 8:   S = 0   69 1257.7 0.299 1.917 0.0551796  .
## Season 9:   S = 0   25 1257.7 0.108 0.677 0.4985644
## Season 10:   S = 0  79 1257.7 0.342 2.199 0.0278468  *
## Season 11:   S = 0  49 1257.7 0.212 1.354 0.1758957
## Season 12:   S = 0  27 1257.7 0.117 0.733 0.4634693
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```




```
##
##  Seasonal Mann-Kendall trend test (Hirsch-Slack test)
##
## data:  gage_06279940_ts
## z = 6.4876, p-value = 8.721e-11
## alternative hypothesis: true S is not equal to 0
## sample estimates:
##     S  varS
##   798 15092
```

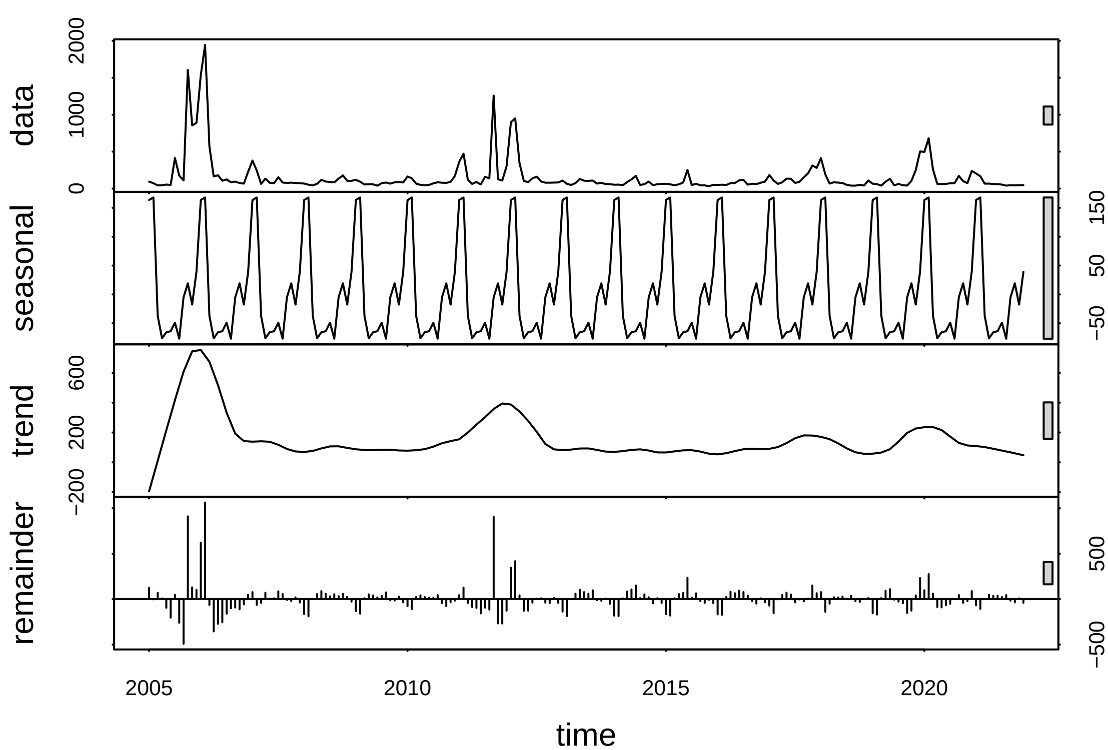
```
##
##  Seasonal Mann-Kendall trend test (Hirsch-Slack test)
##
## data: gage_06279940_ts
## alternative hypothesis: two.sided
##
## Statistics for individual seasons
##
## H0
```

```
##                        S   varS   tau     z  Pr(>|z|)
## Season 1:   S = 0  113 1257.7 0.489 3.158 0.0015876 **
## Season 2:   S = 0   65 1257.7 0.281 1.805 0.0711267  .
## Season 3:   S = 0   61 1257.7 0.264 1.692 0.0906697  .
## Season 4:   S = 0   79 1257.7 0.342 2.199 0.0278468  *
## Season 5:   S = 0   47 1257.7 0.203 1.297 0.1945951
## Season 6:   S = 0   81 1257.7 0.351 2.256 0.0240810  *
## Season 7:   S = 0   35 1257.7 0.152 0.959 0.3376950
## Season 8:   S = 0   49 1257.7 0.212 1.354 0.1758957
## Season 9:   S = 0   47 1257.7 0.203 1.297 0.1945951
## Season 10:   S = 0  67 1257.7 0.290 1.861 0.0627352  .
## Season 11:   S = 0  67 1257.7 0.290 1.861 0.0627352  .
## Season 12:   S = 0  87 1257.7 0.377 2.425 0.0153075  *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```



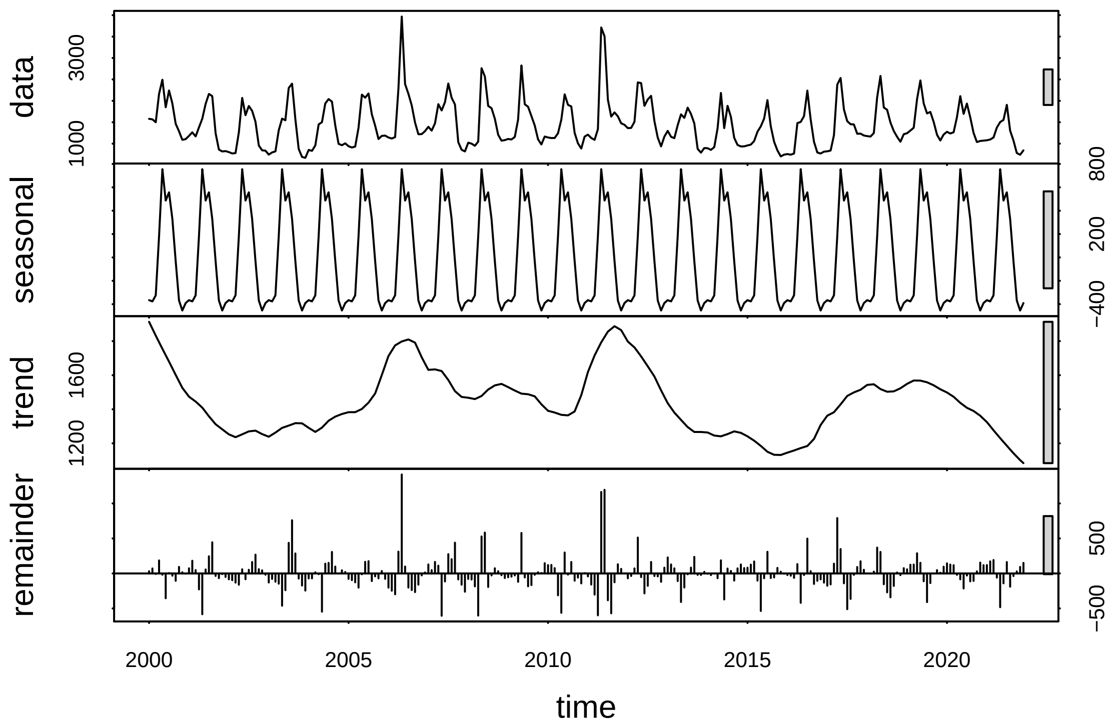

```
##
##  Seasonal Mann-Kendall trend test (Hirsch-Slack test)
##
## data:  gage_09146200_ts
## z = -2.6455, p-value = 0.008157
## alternative hypothesis: true S is not equal to 0
## sample estimates:
##     S  varS
##  -326 15092

##
```

```
##  Seasonal Mann-Kendall trend test (Hirsch-Slack test)
##
## data: gage_09146200_ts
## alternative hypothesis: two.sided
##
## Statistics for individual seasons
##
## H0
##                        S   varS    tau      z Pr(>|z|)
## Season 1:   S = 0  -51 1257.7 -0.221 -1.410 0.158570
## Season 2:   S = 0  -11 1257.7 -0.048 -0.282 0.777959
## Season 3:   S = 0   -1 1257.7 -0.004  0.000 1.000000
## Season 4:   S = 0  -49 1257.7 -0.212 -1.354 0.175896
## Season 5:   S = 0  -91 1257.7 -0.394 -2.538 0.011155 *
## Season 6:   S = 0   51 1257.7  0.221  1.410 0.158570
## Season 7:   S = 0   33 1257.7  0.143  0.902 0.366880
## Season 8:   S = 0    1 1257.7  0.004  0.000 1.000000
## Season 9:   S = 0  -53 1257.7 -0.229 -1.466 0.142569
## Season 10:   S = 0 -55 1257.7 -0.238 -1.523 0.127837
## Season 11:   S = 0 -41 1257.7 -0.177 -1.128 0.259355
## Season 12:   S = 0 -59 1257.7 -0.255 -1.635 0.101948
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```




```
##
##  Seasonal Mann-Kendall trend test (Hirsch-Slack test)
```

```
##
## data:  gage_09408135_ts
## z = -2.5804, p-value = 0.009868
## alternative hypothesis: true S is not equal to 0
## sample estimates:
##    S varS
## -218 7072

##
##  Seasonal Mann-Kendall trend test (Hirsch-Slack test)
##
## data: gage_09408135_ts
## alternative hypothesis: two.sided
##
## Statistics for individual seasons
##
## H0
##                      S  varS    tau      z Pr(>|z|)
## Season 1:   S = 0   -6 589.3 -0.044 -0.206 0.836820
## Season 2:   S = 0   -4 589.3 -0.029 -0.124 0.901650
## Season 3:   S = 0    8 589.3  0.059  0.288 0.773080
## Season 4:   S = 0    8 589.3  0.059  0.288 0.773080
## Season 5:   S = 0    2 589.3  0.015  0.041 0.967142
## Season 6:   S = 0   14 589.3  0.103  0.536 0.592301
## Season 7:   S = 0  -60 589.3 -0.441 -2.430 0.015084 *
## Season 8:   S = 0  -62 589.3 -0.456 -2.513 0.011979 *
## Season 9:   S = 0  -34 589.3 -0.250 -1.359 0.174034
## Season 10:   S = 0 -36 589.3 -0.265 -1.442 0.149375
## Season 11:   S = 0 -34 589.3 -0.250 -1.359 0.174034
## Season 12:   S = 0 -14 589.3 -0.103 -0.536 0.592301
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```




```
##
##  Seasonal Mann-Kendall trend test (Hirsch-Slack test)
##
## data:  gage_13046000_ts
## z = -1.6932, p-value = 0.09042
## alternative hypothesis: true S is not equal to 0
## sample estimates:
##     S  varS
##  -209 15091
```

```
##
##  Seasonal Mann-Kendall trend test (Hirsch-Slack test)
##
## data: gage_13046000_ts
## alternative hypothesis: two.sided
##
## Statistics for individual seasons
##
## H0
##                       S   varS    tau      z   Pr(>|z|)
## Season 1:   S = 0    17 1257.7  0.074  0.451 0.65186928
## Season 2:   S = 0    23 1257.7  0.100  0.620 0.53502445
## Season 3:   S = 0    33 1257.7  0.143  0.902 0.36687960
## Season 4:   S = 0    31 1257.7  0.134  0.846 0.39758739
## Season 5:   S = 0    -7 1257.7 -0.030 -0.169 0.86564910
## Season 6:   S = 0   -43 1257.7 -0.186 -1.184 0.23628916
## Season 7:   S = 0   -86 1256.7 -0.373 -2.398 0.01649488   *
```

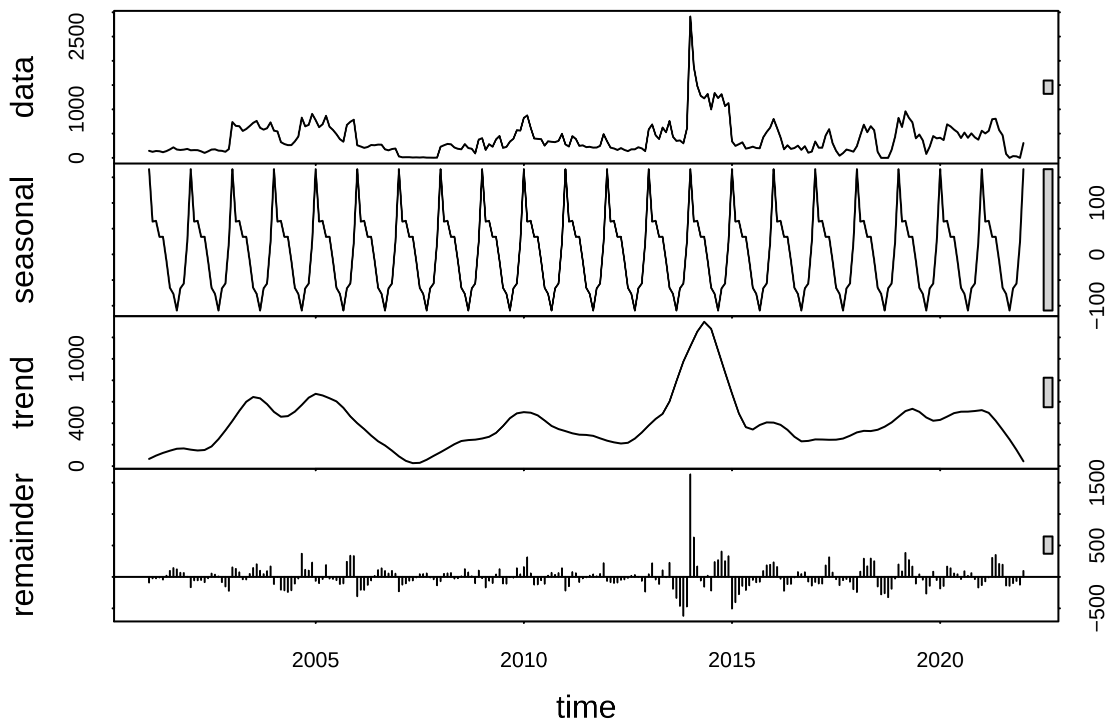
Table 3: Table 3: Seasonal Mann Kendall Tests for Streamflow

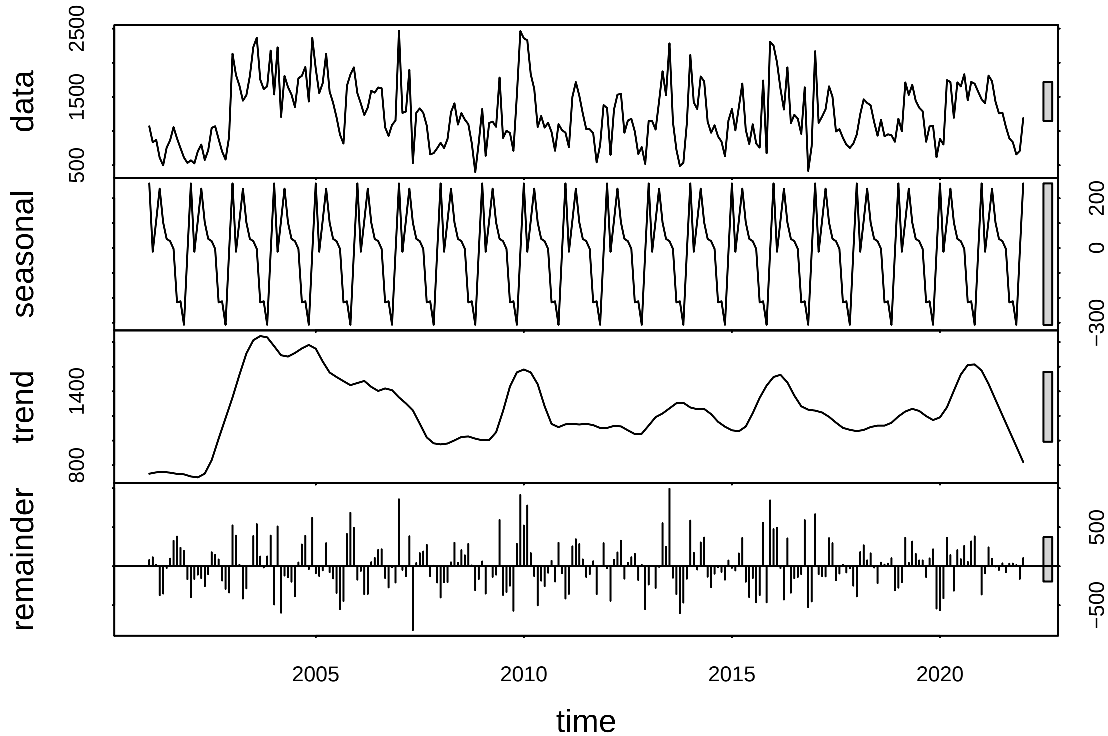
Stream_Gauges	tau	Two_sided_P_value
High Shoals	0.097	0.033
Schoolfield	0.026	0.562
Matson	0.050	0.271
Otis	-0.052	0.249
Ware Shoals	-0.105	0.021
Garland	-0.102	0.115
Chetser	0.075	0.386
Tri County	-0.292	0.001
Quail Creek	0.100	0.028

```
## Season 8:   S = 0  -131 1257.7 -0.567 -3.666 0.00024663 ***
## Season 9:   S = 0   -61 1257.7 -0.264 -1.692 0.09066966   .
## Season 10:   S = 0   -9 1257.7 -0.039 -0.226 0.82152542
## Season 11:   S = 0   -1 1257.7 -0.004  0.000 1.00000000
## Season 12:   S = 0   25 1257.7  0.108  0.677 0.49856439
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

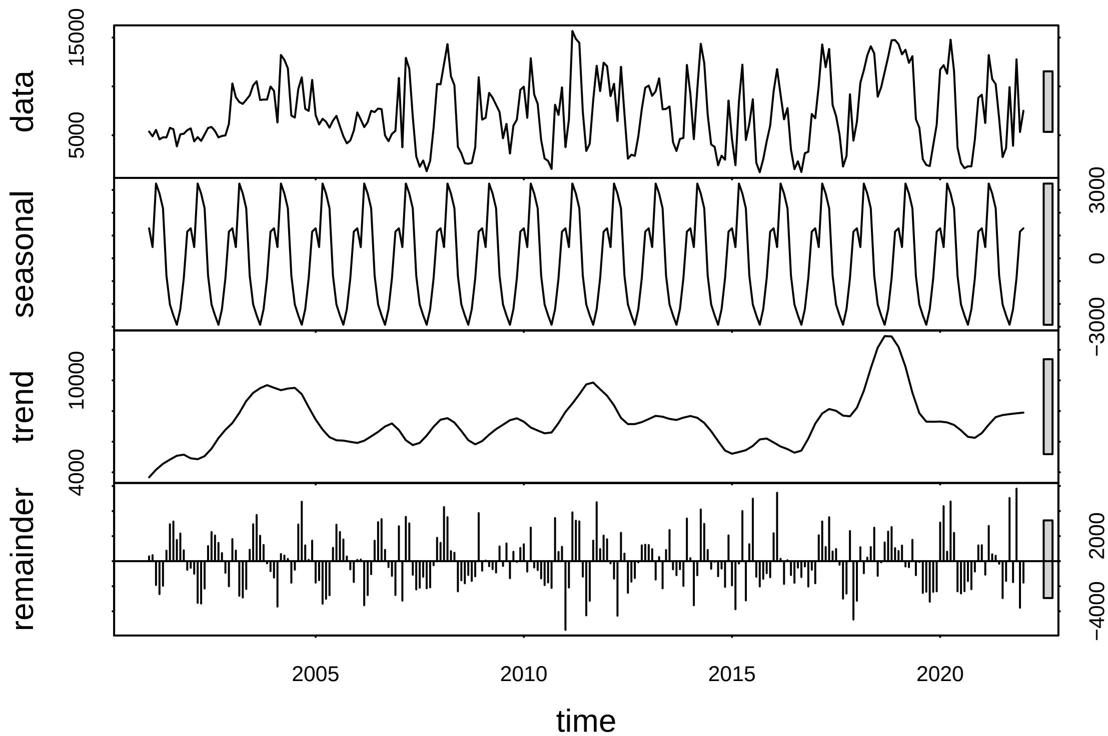
Question 2: Are there trends in the generation data?

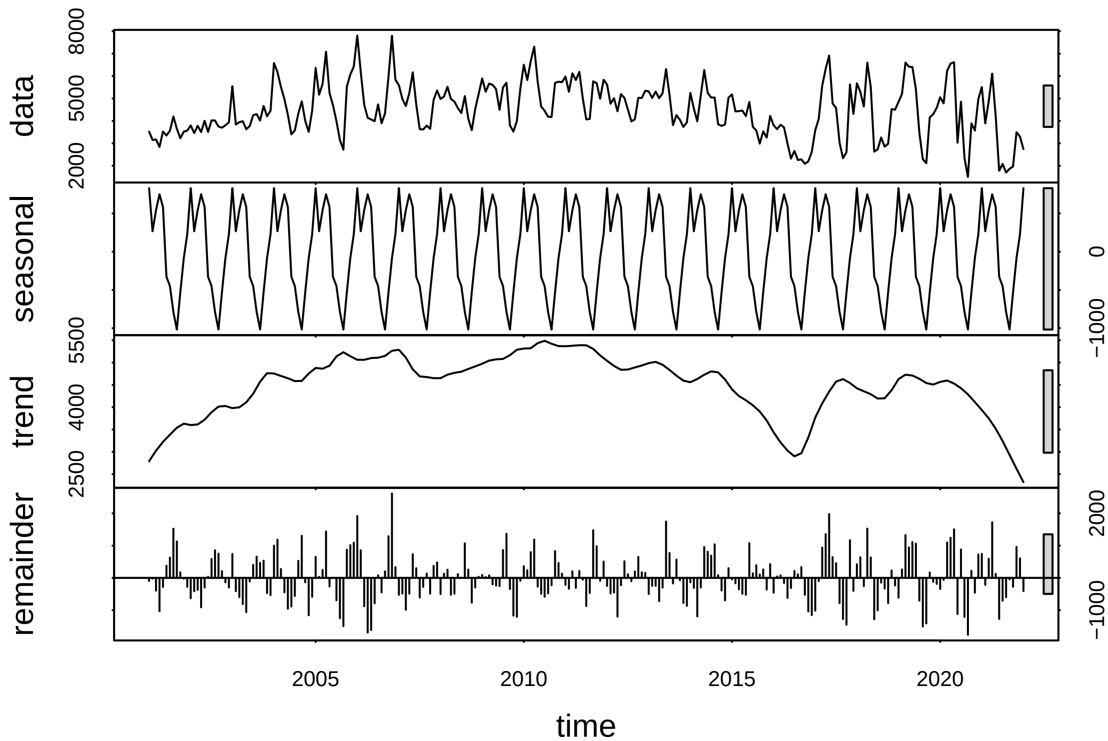
We also created a time series for each hydropower plant, and then decomposed each time series into its seasonal, trend, and remainder components. We plotted each decomposed time series, and observed that each site displayed strong seasonal trends in streamflow. Additionally, we ran a Seasonal Mann-Kendall test on each time series to test whether there was a monotonic trend in hydroelectric generation over time. The results are displayed in Table 2. Two plants, High Shoals and Quail Creek, display a slight statistically significant positive trend over time. Conversely, Ware Shoals and Tri County display a statistically significant negative trend over time. Tri County, which is located in Colorado, has a relatively substantial negative trend. The Tri County result is consistent with the relatively significant drought conditions experienced in Colorado over the last 20 years. The other hydropower plants have 2-sided P-values greater than 0.05, and thus we cannot reject the null hypothesis that no trend exists. Overall, the results indicate that hydropower generation demonstrates no clear trend on the national level, and each power plant is subject to unique conditions that may cause generation to trend in either direction.

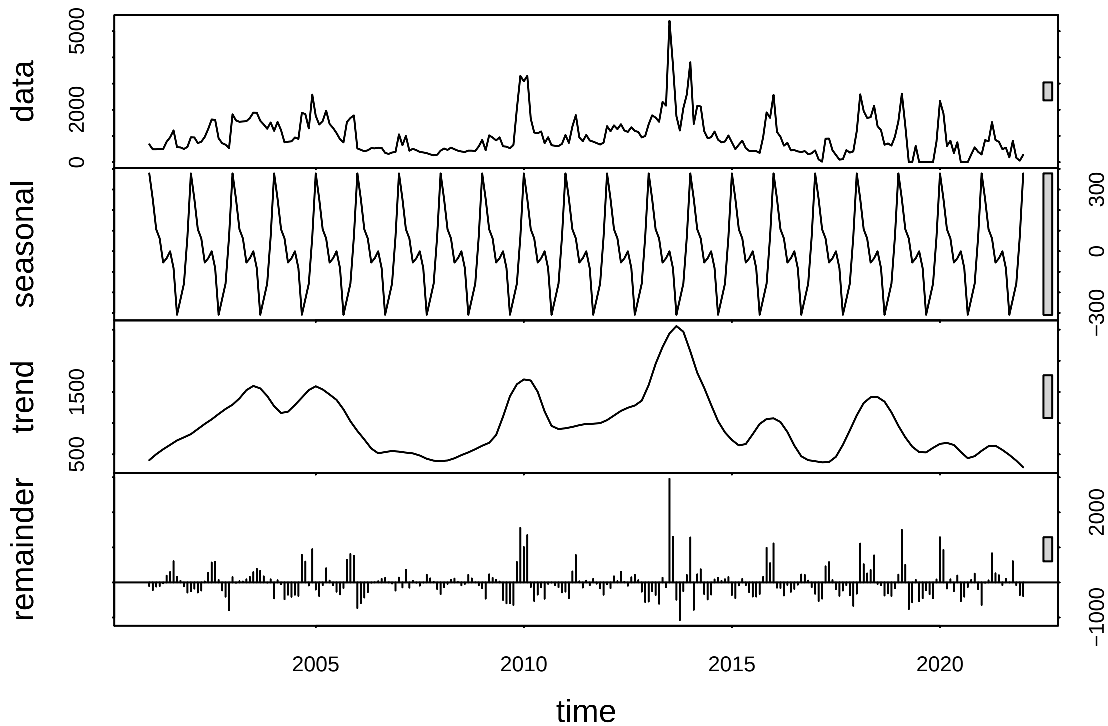

data
seasonal
trend
remainder
0
1000
2500
100
0
−100
0
400
1000
1500
500
−500
2005
2010
2015
2020
time

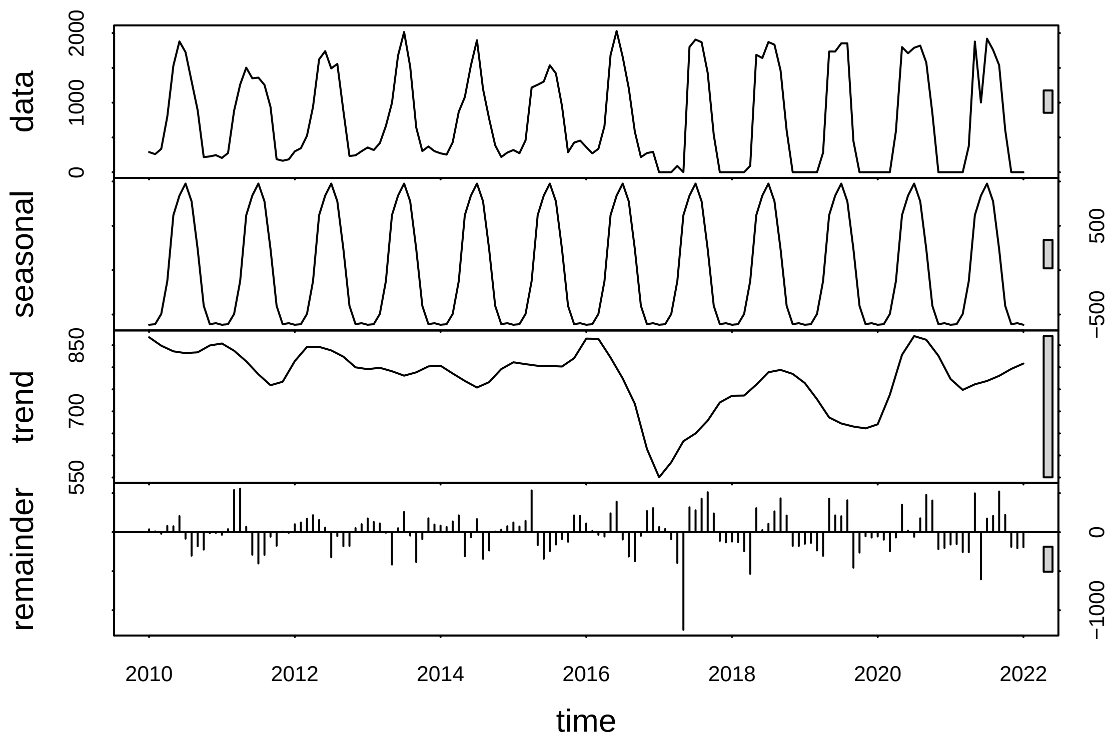


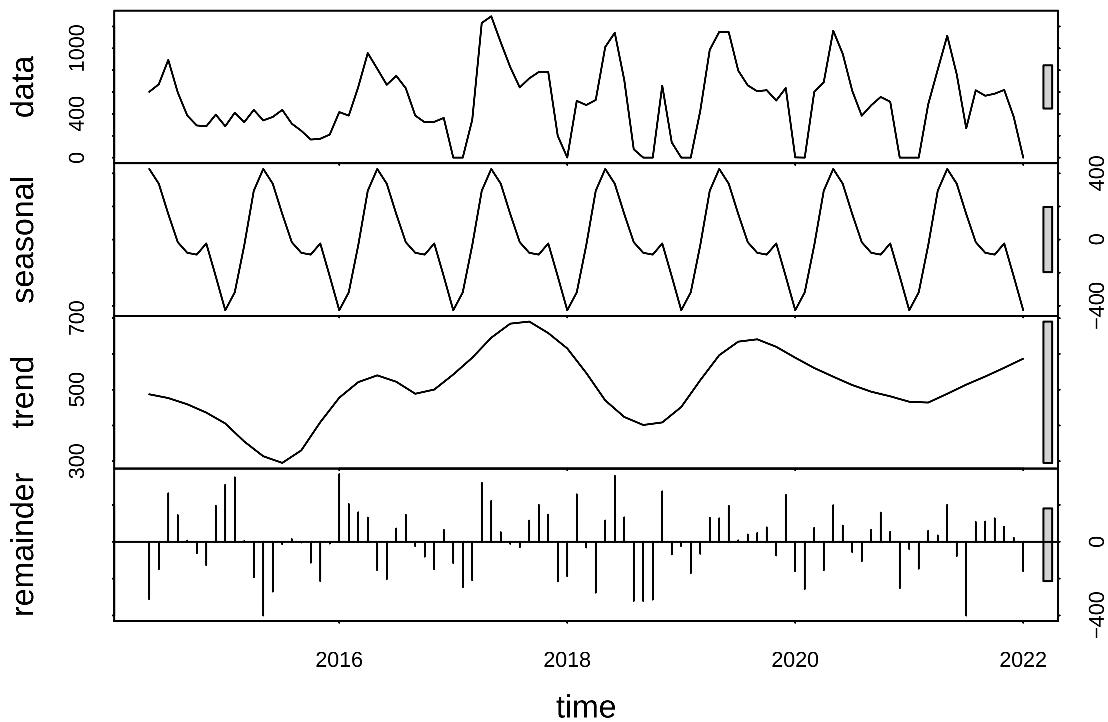
data
seasonal
trend
remainder
2500
1500
500
200
0
−300
1400
800
500
−500
2005
2010
2015
2020
time

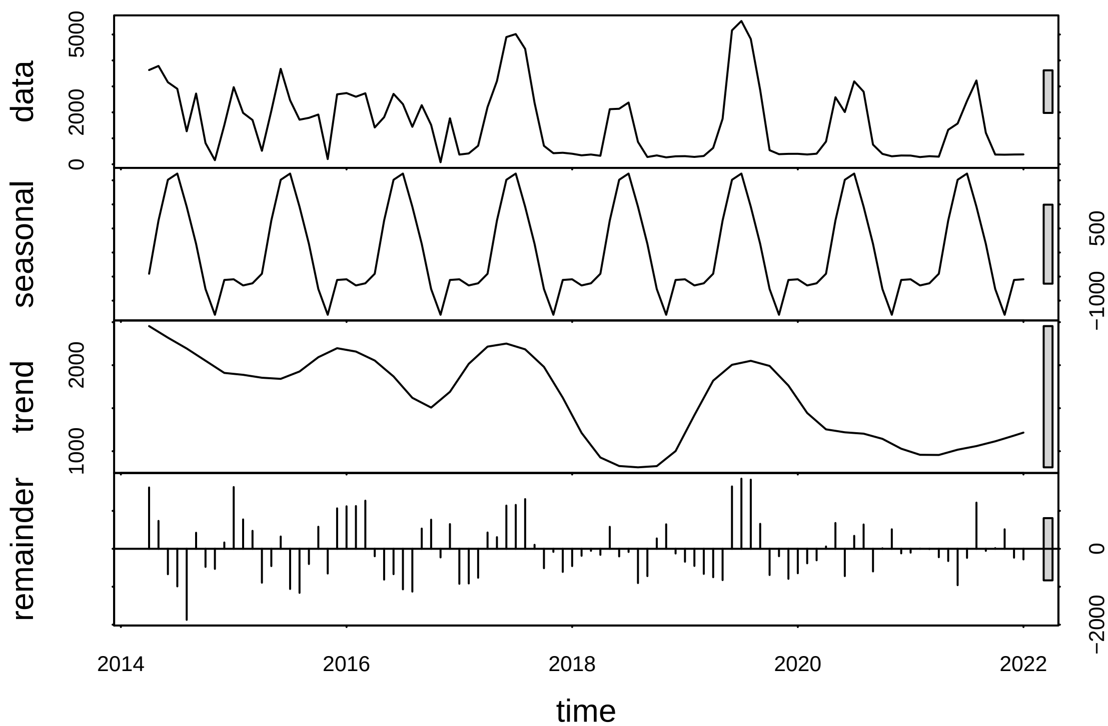

data
15000
5000
seasonal
3000
0
−3000
trend
10000
4000
remainder
2000
−4000
2005
2010
2015
2020
time


data
seasonal
trend
remainder
2000
5000
8000
0
−1000
2500
4000
5500
2000
−1000
2005
2010
2015
2020
time

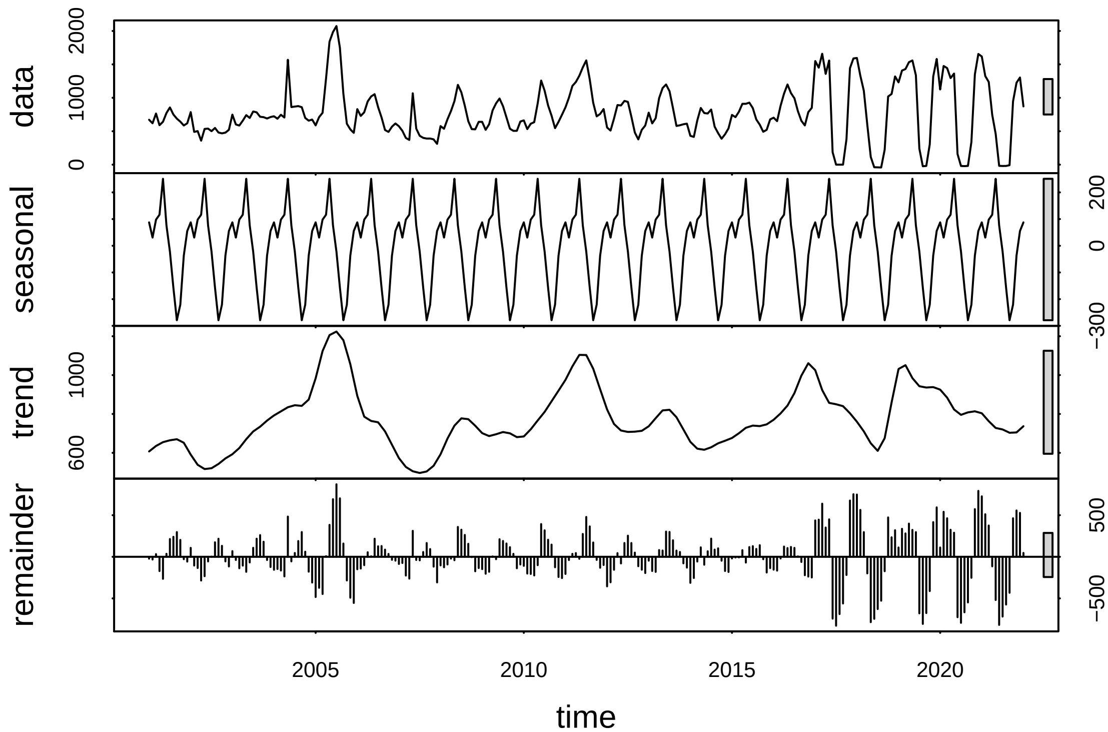

data
seasonal
trend
remainder
0
2000
5000
300
0
−300
500
1500
2000
−1000
2005
2010
2015
2020
time


data
seasonal
trend
remainder
0
1000
2000
500
−500
550
700
850
0
−1000
2010
2012
2014
2016
2018
2020
2022
time


data
seasonal
trend
remainder
time
0
400
1000
700
500
300
400
0
−400
0
−400
2016
2018
2020
2022



data
seasonal
trend
remainder
0
2000
5000
500
−1000
1000
2000
0
−2000
2014
2016
2018
2020
2022
time



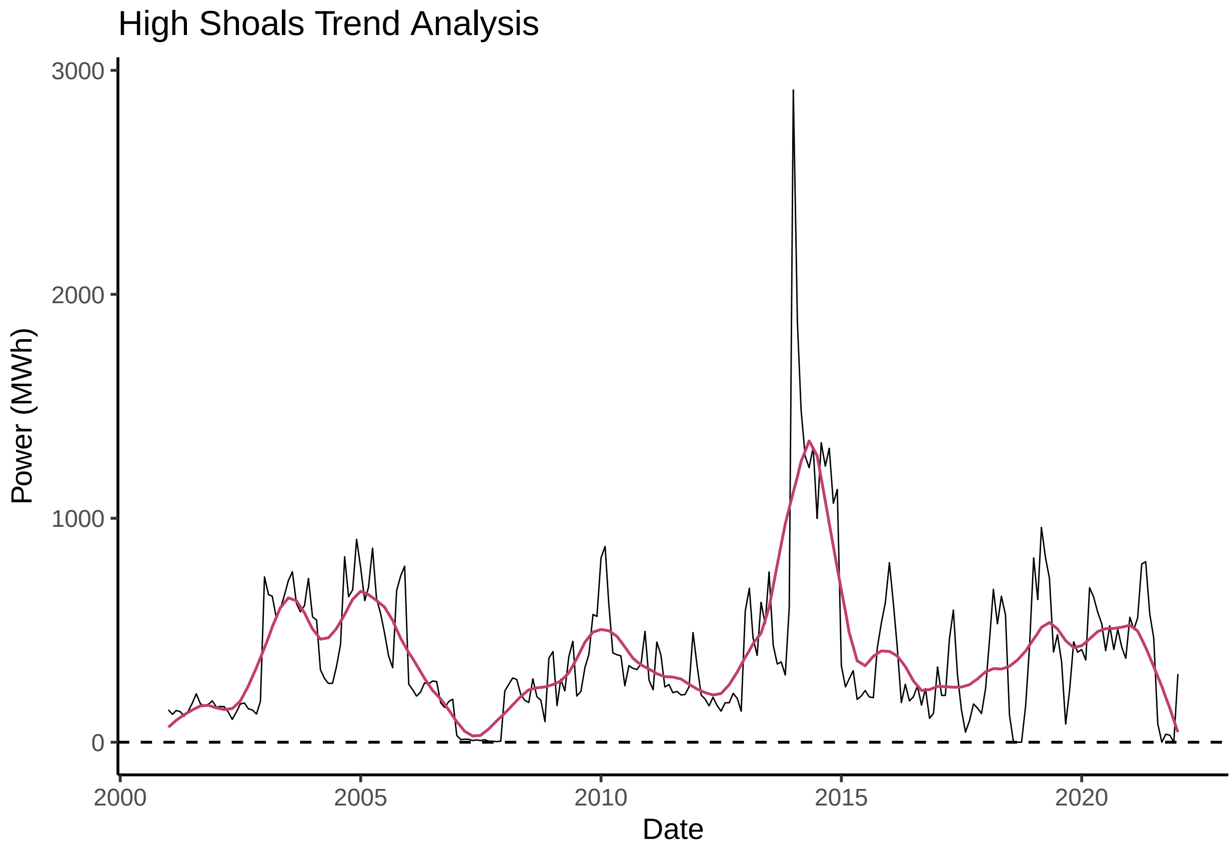

```
## tau = 0.0968, 2-sided pvalue =0.033049

## Score =  246 , Var(Score) = 13320
## denominator =  2540.499
## tau = 0.0968, 2-sided pvalue =0.033049

## tau = 0.0264, 2-sided pvalue =0.56157

## Score =  67 , Var(Score) = 13321
## denominator =  2541
## tau = 0.0264, 2-sided pvalue =0.56157

## tau = 0.05, 2-sided pvalue =0.27117

## Score =  127 , Var(Score) = 13321
## denominator =  2541
## tau = 0.05, 2-sided pvalue =0.27117

## tau = -0.0523, 2-sided pvalue =0.24918

## Score =  -133 , Var(Score) = 13321
## denominator =  2541
## tau = -0.0523, 2-sided pvalue =0.24918

## tau = -0.105, 2-sided pvalue =0.020693

## Score =  -267 , Var(Score) = 13319
## denominator =  2539.999
## tau = -0.105, 2-sided pvalue =0.020693

## tau = -0.102, 2-sided pvalue =0.11505
```

Table 4: Table 4. Seasonal Mann Kendall Tests for Hydroelectric Generation

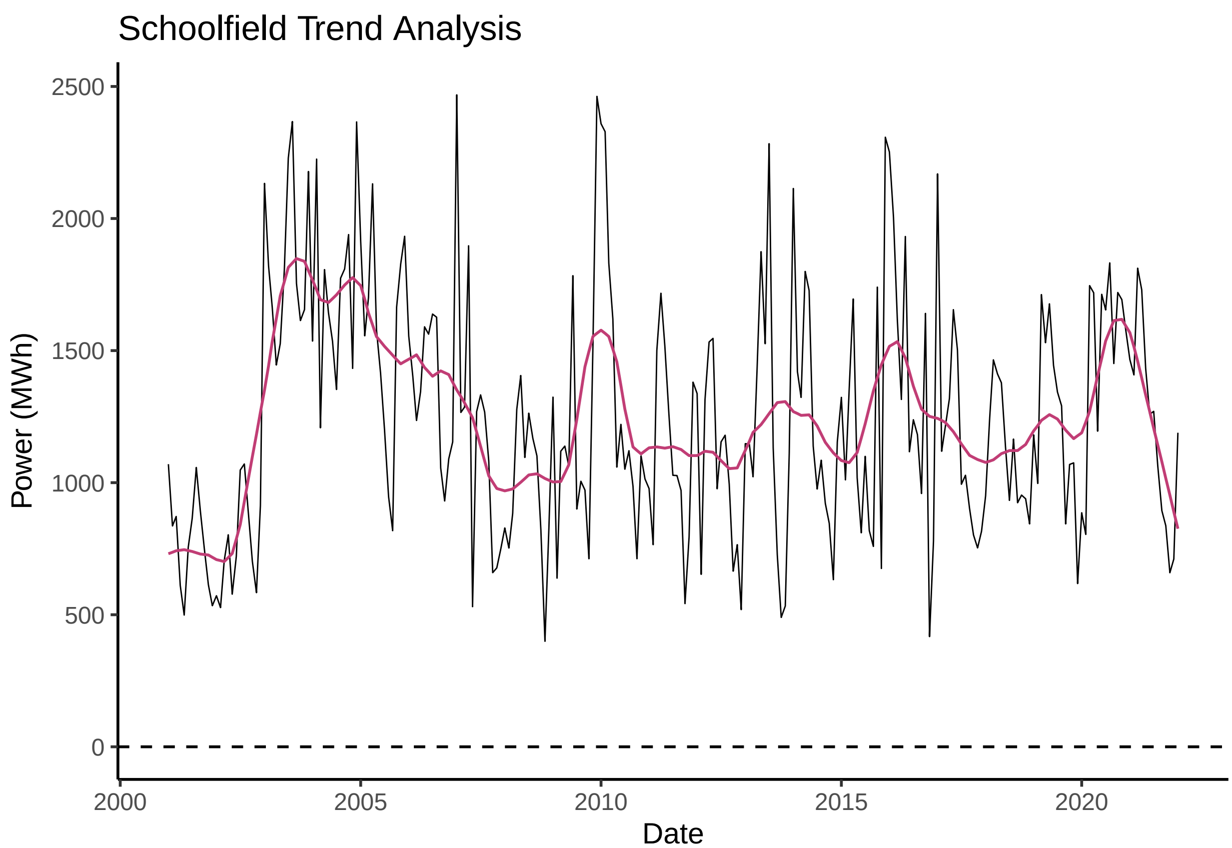
Hydropower_Plants	tau	Two_sided_P_value
High Shoals	0.097	0.033
Schoolfield	0.026	0.562
Matson	0.050	0.271
Otis	-0.052	0.249
Ware Shoals	-0.105	0.021
Garland	-0.102	0.115
Chetser	0.075	0.386
Tri County	-0.292	0.001
Quail Creek	0.100	0.028

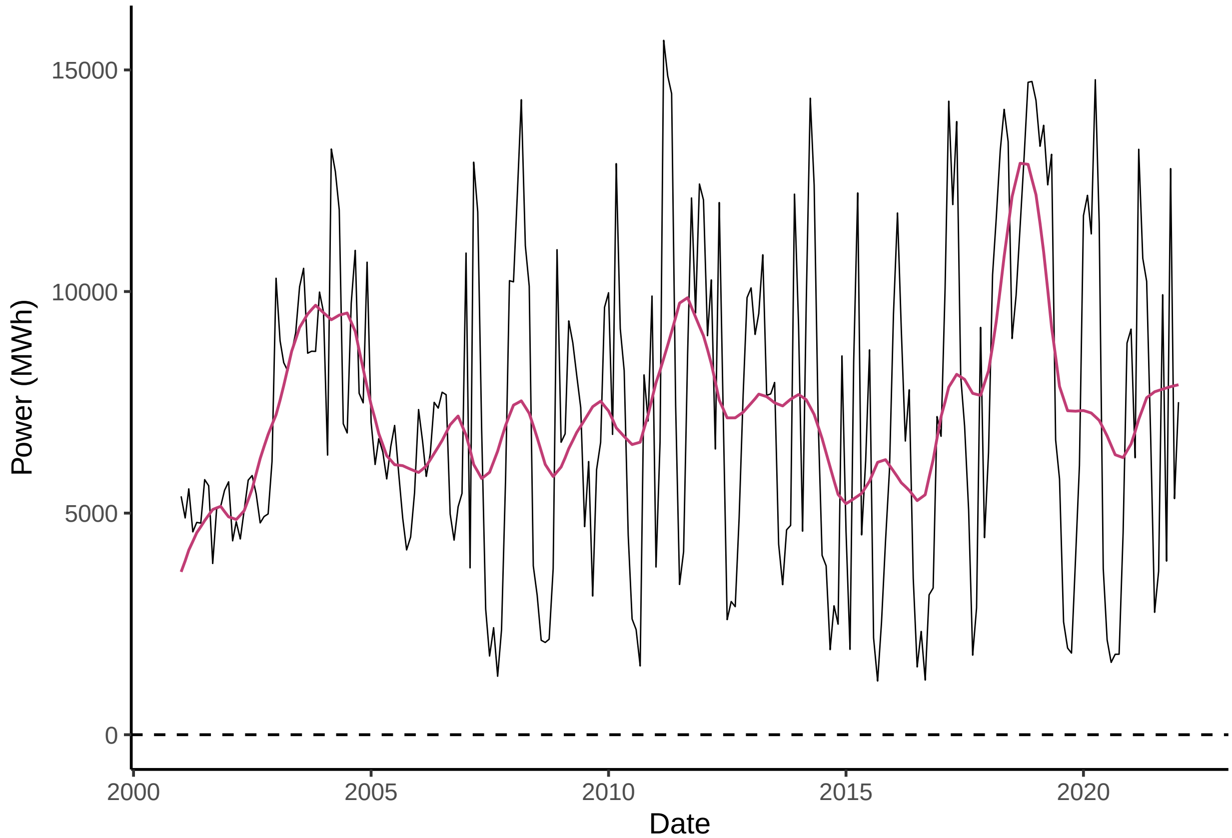
```
## Score =  -79 , Var(Score) = 2513
## denominator =  775.2789
## tau = -0.102, 2-sided pvalue =0.11505

## tau = 0.0745, 2-sided pvalue =0.38591

## Score =  23 , Var(Score) = 703.6667
## denominator =  308.5676
## tau = 0.0745, 2-sided pvalue =0.38591

## tau = -0.292, 2-sided pvalue =0.00055883

## Score =  -94 , Var(Score) = 742
## denominator =  322
## tau = -0.292, 2-sided pvalue =0.00055883

## tau = 0.0996, 2-sided pvalue =0.028363

## Score =  253 , Var(Score) = 13319
## denominator =  2539.999
## tau = 0.0996, 2-sided pvalue =0.028363
```

ggplot was used to graphically represent the trend and seasonal components relative to the observed generation data. Additional figures are included in the R code file associated with this report. ggplot was also used to display the relationship between average monthly discharge and monthly power generation, as seen in Figure 3.

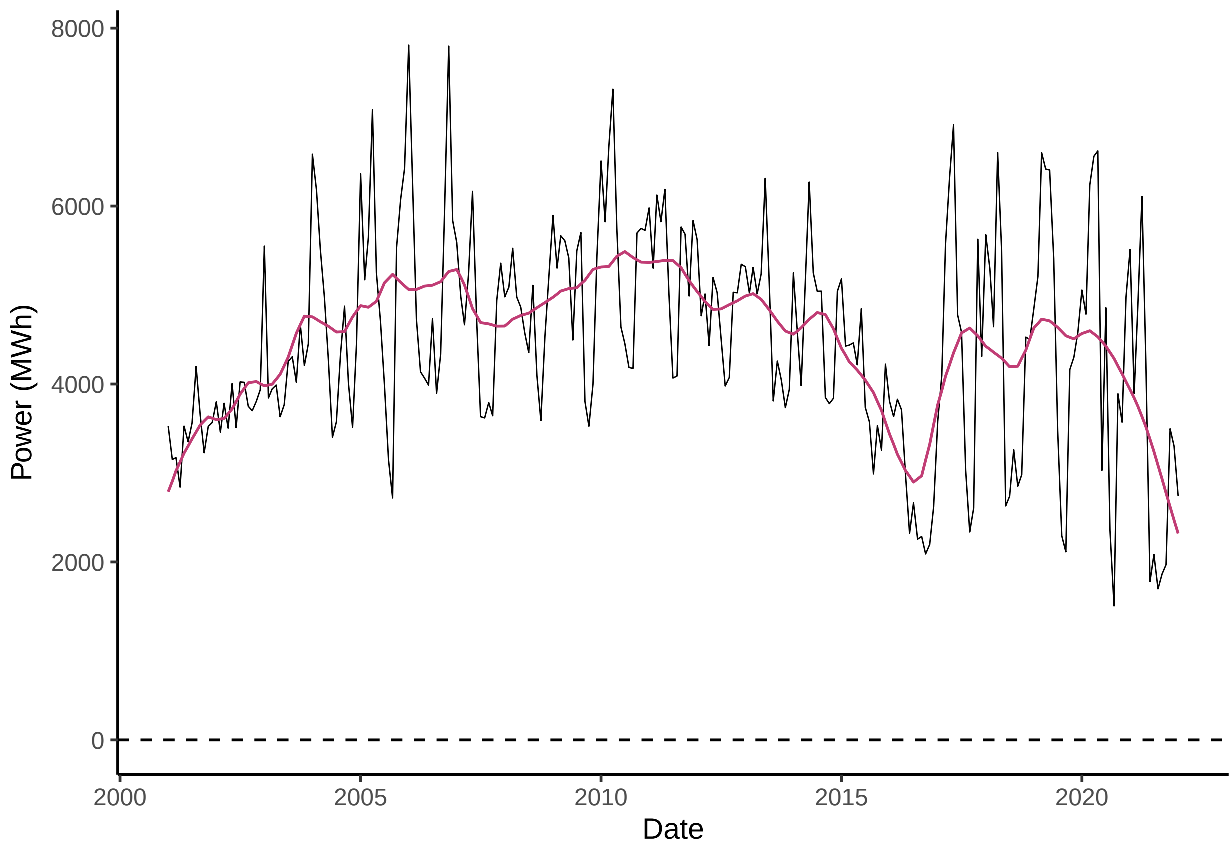


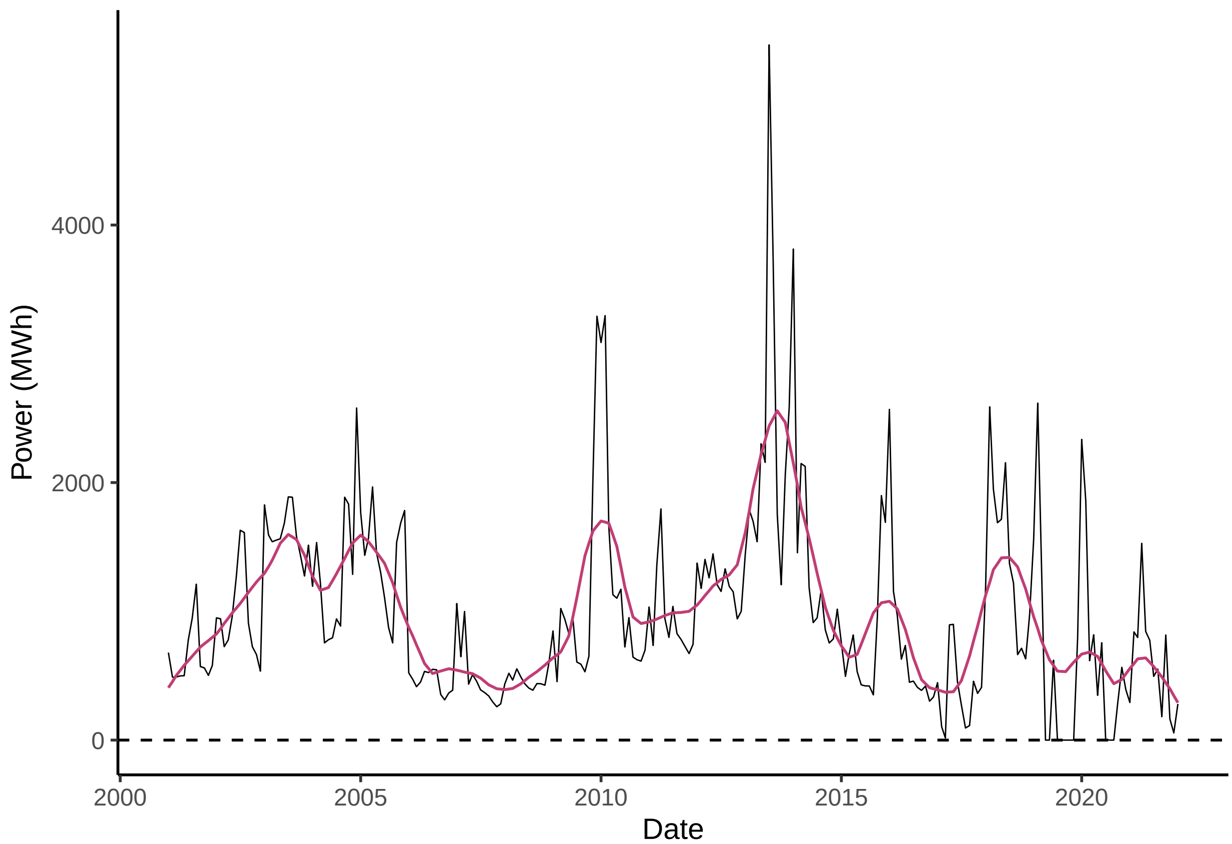
High Shoals Trend Analysis
Power (MWh)
3000
2000
1000
0
2000
2005
2010
2015
2020
Date

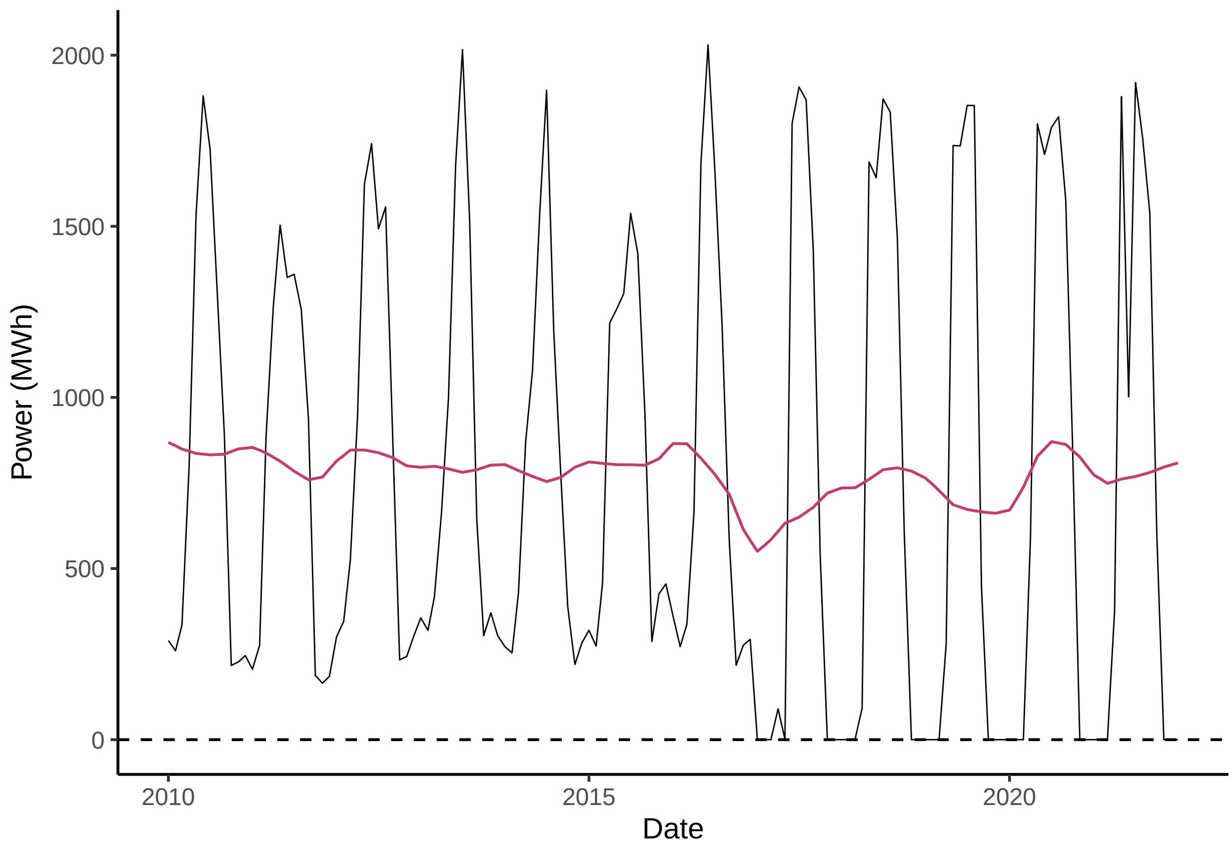

Schoolfield Trend Analysis
Power (MWh)
2500
2000
1500
1000
500
0
2000
2005
2010
2015
2020
Date



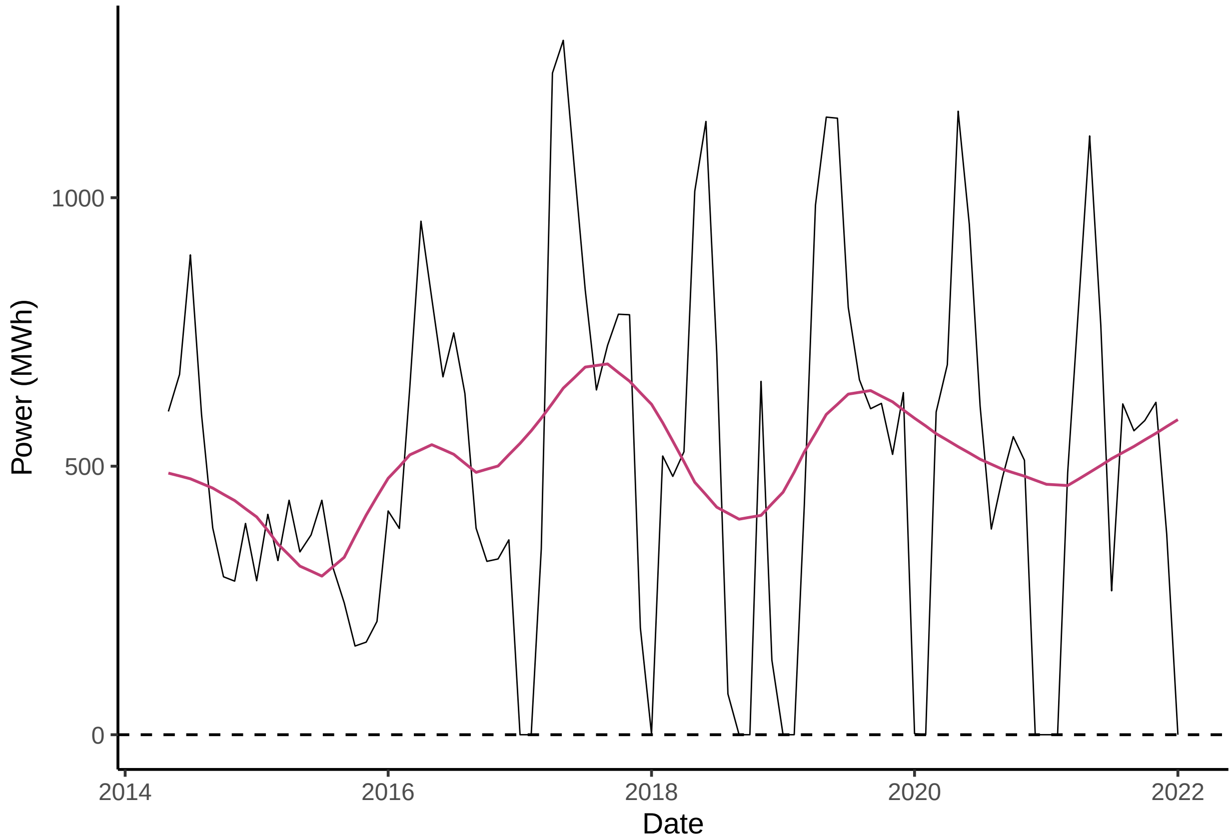
Power (MWh)
15000
10000
5000
0
2000
2005
2010
2015
2020
Date


8000
6000
4000
2000
0
Power (MWh)
2000
2005
2010
2015
2020
Date

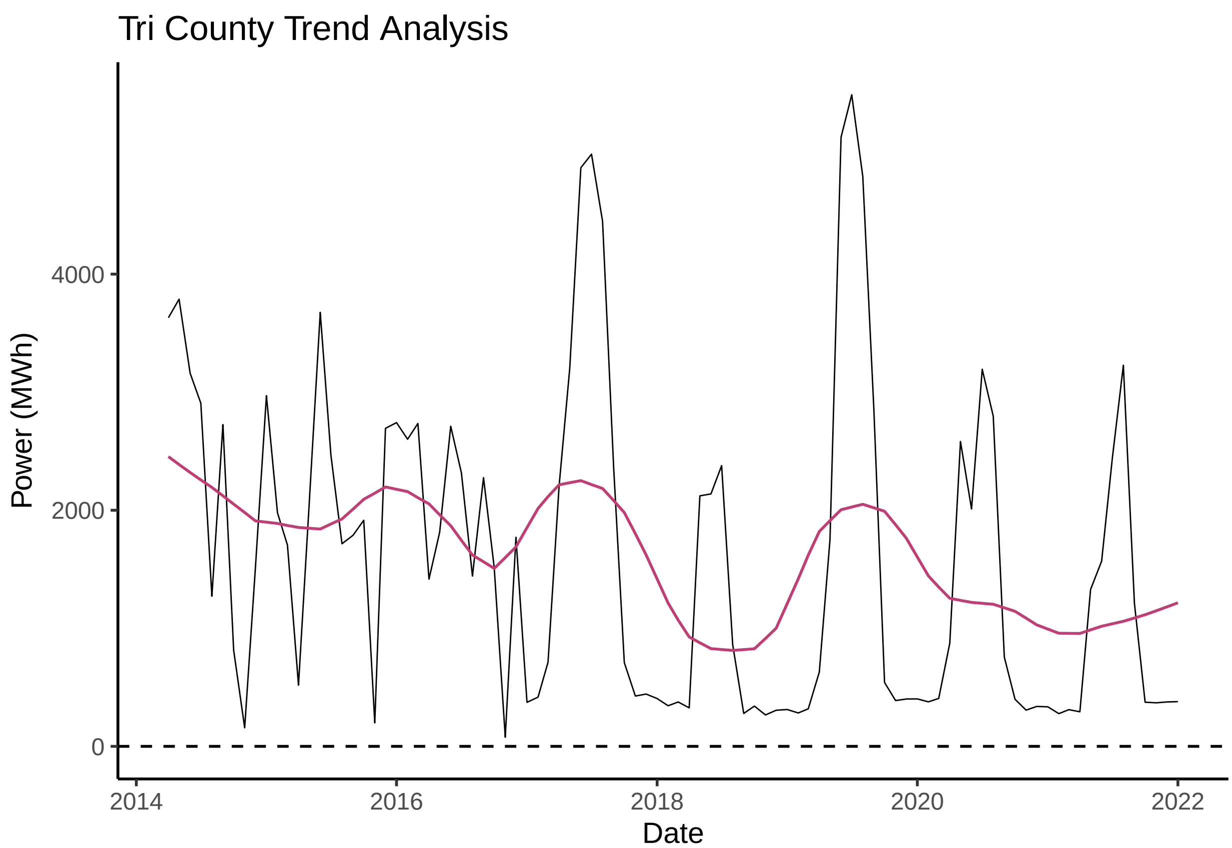

Power (MWh)
4000
2000
0
2000
2005
2010
2015
2020
Date

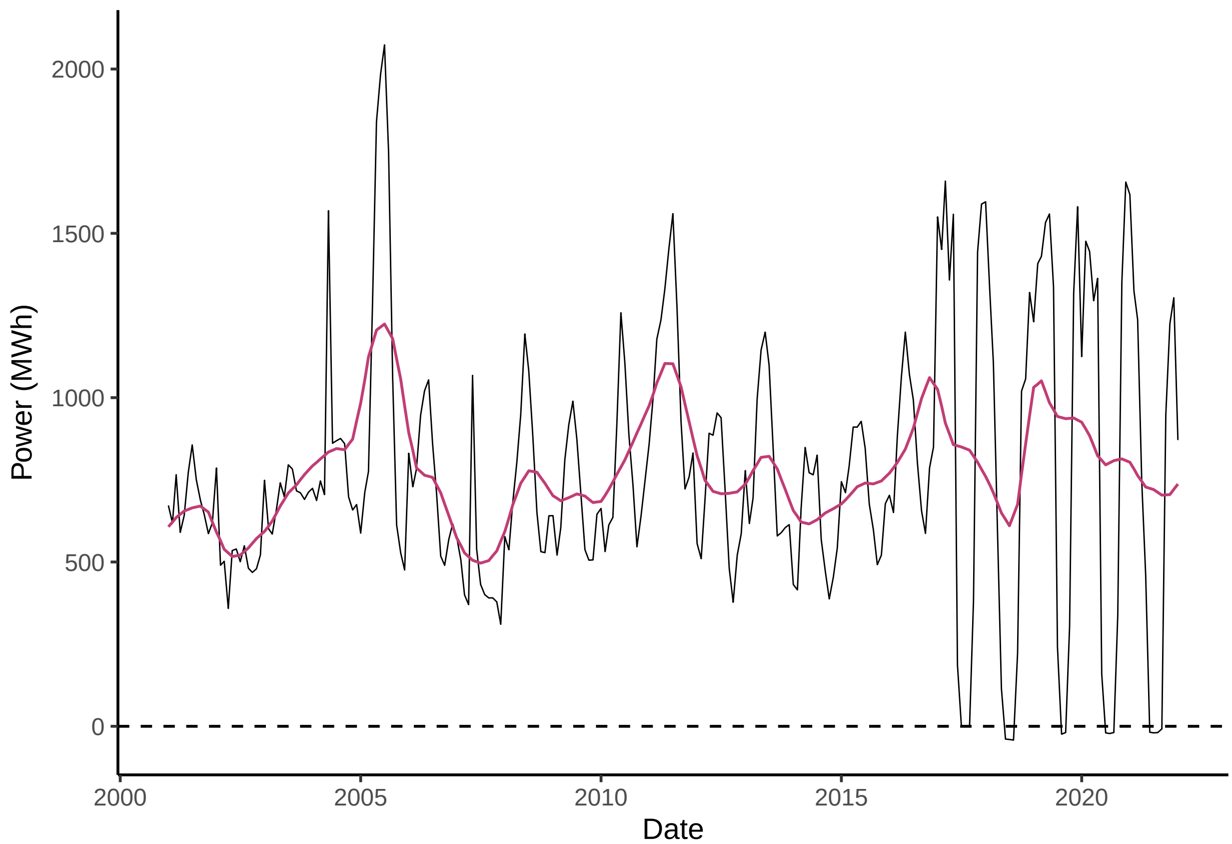


Power (MWh)
2000
1500
1000
500
0
2010
2015
2020
Date

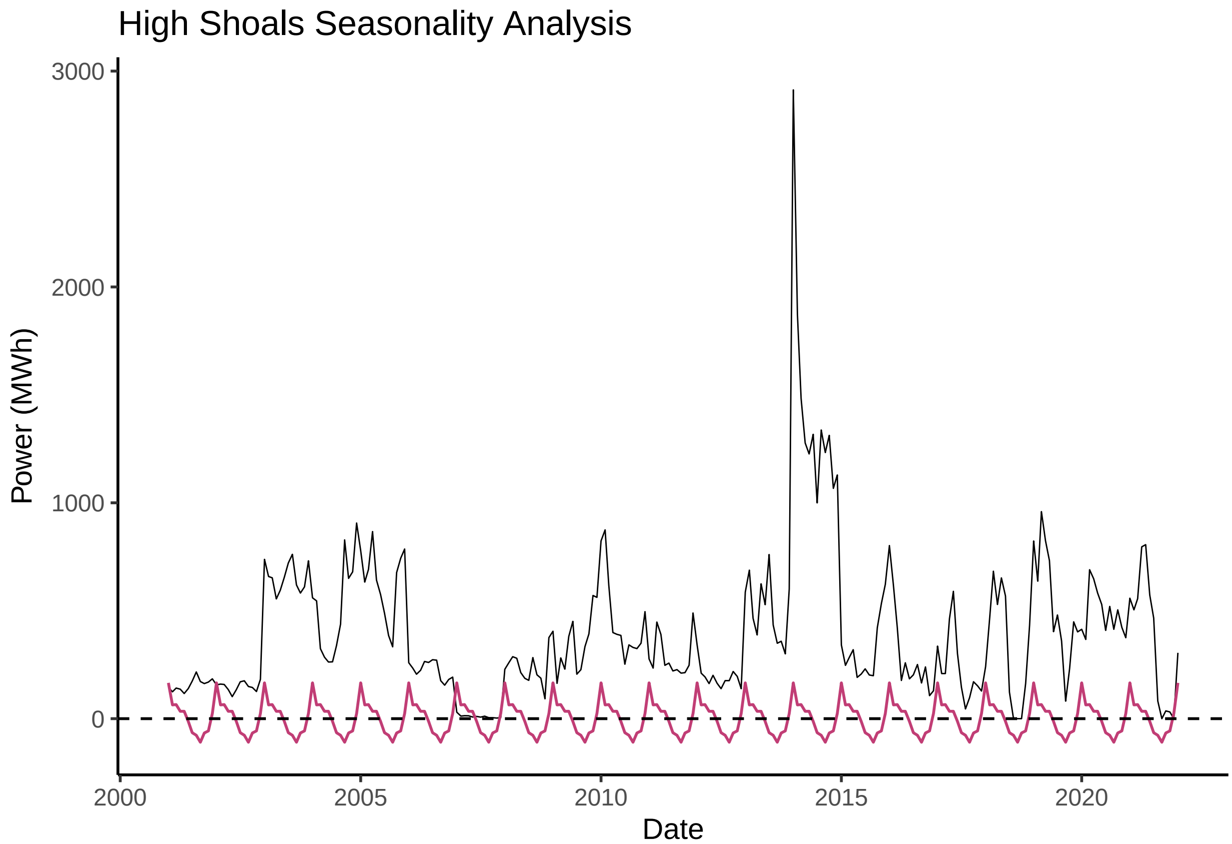


Power (MWh)
1000
500
0
2014
2016
2018
2020
2022
Date

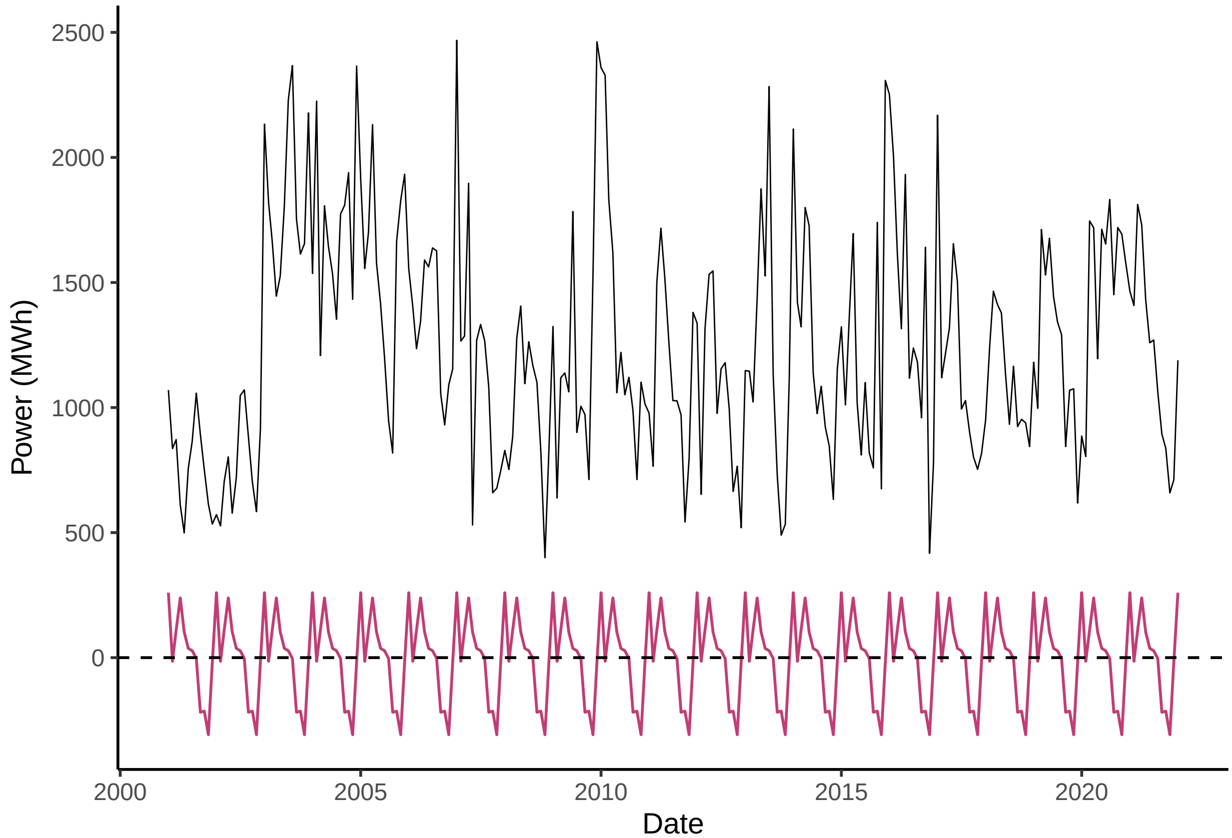

Tri County Trend Analysis
4000
2000
0
Power (MWh)
2014
2016
2018
2020
2022
Date



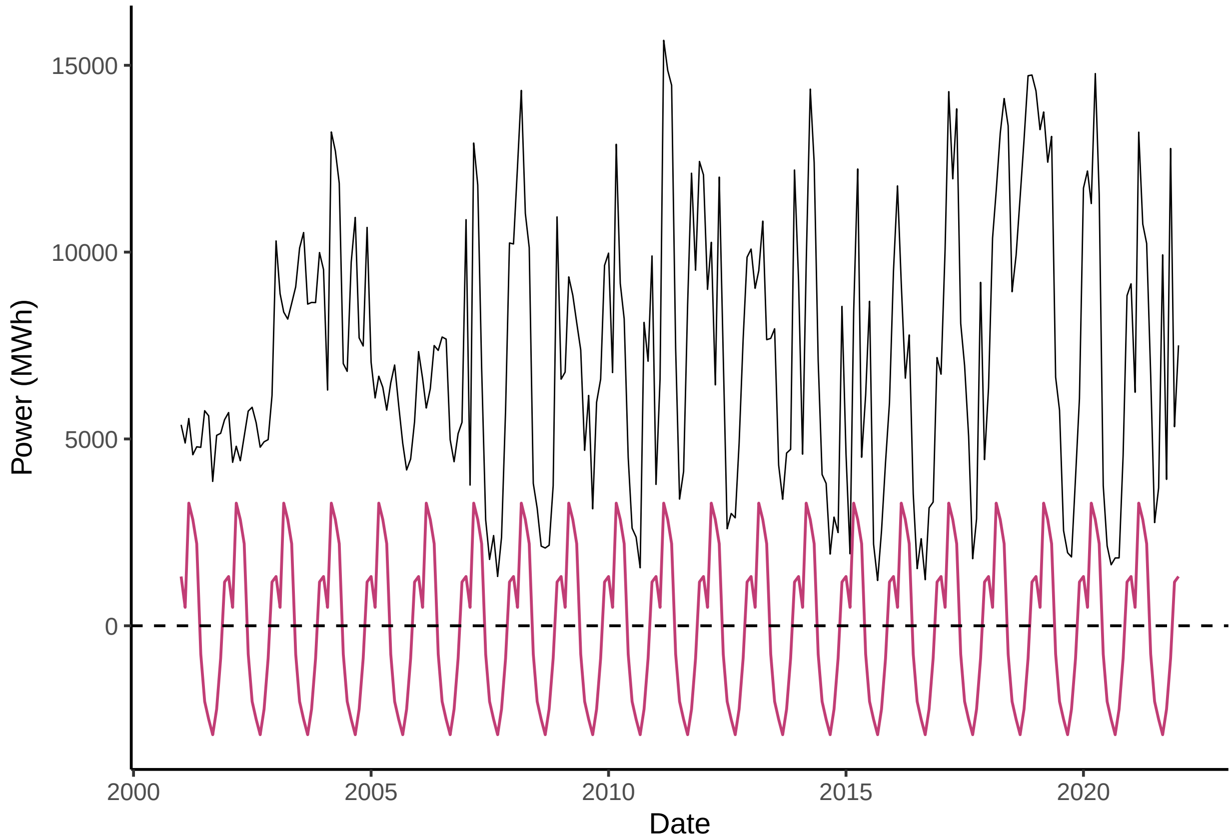
2000
1500
1000
500
0
Power (MWh)
2000
2005
2010
2015
2020
Date

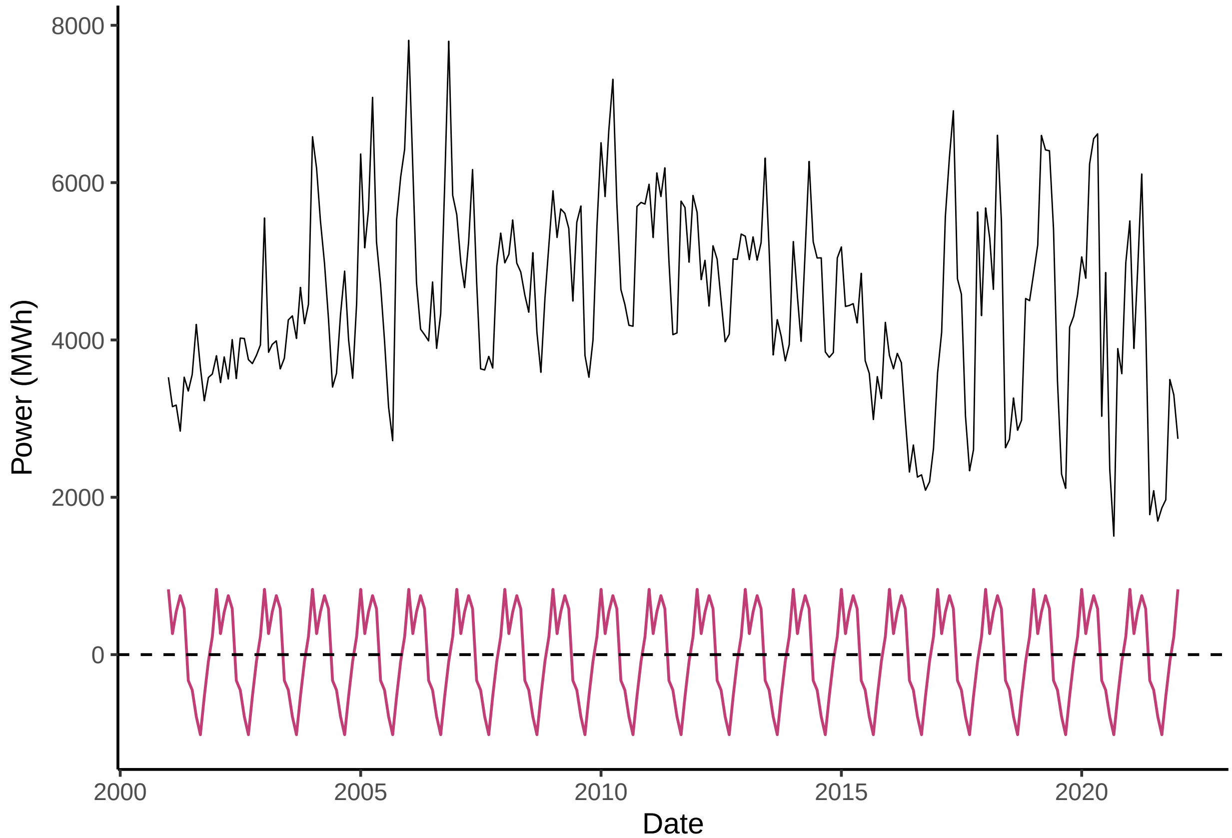


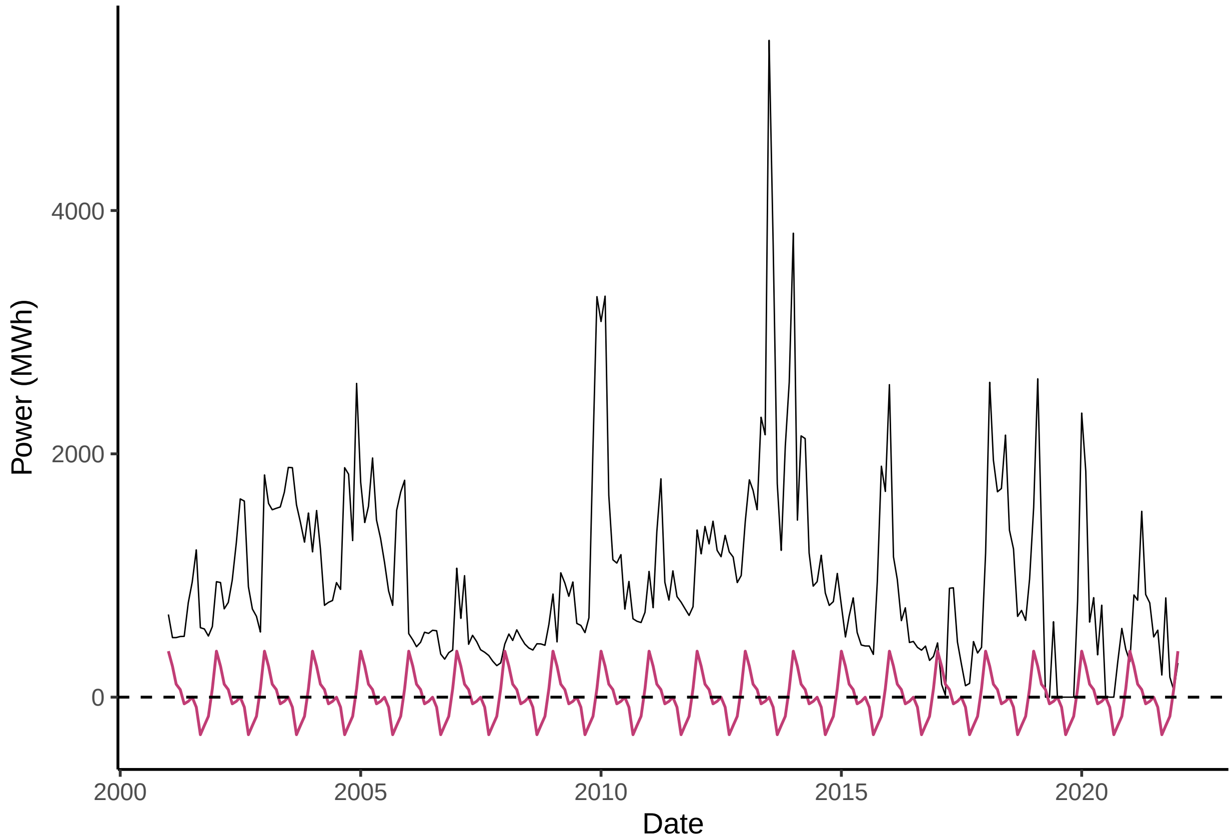
High Shoals Seasonality Analysis
Power (MWh)
3000
2000
1000
0
2000
2005
2010
2015
2020
Date



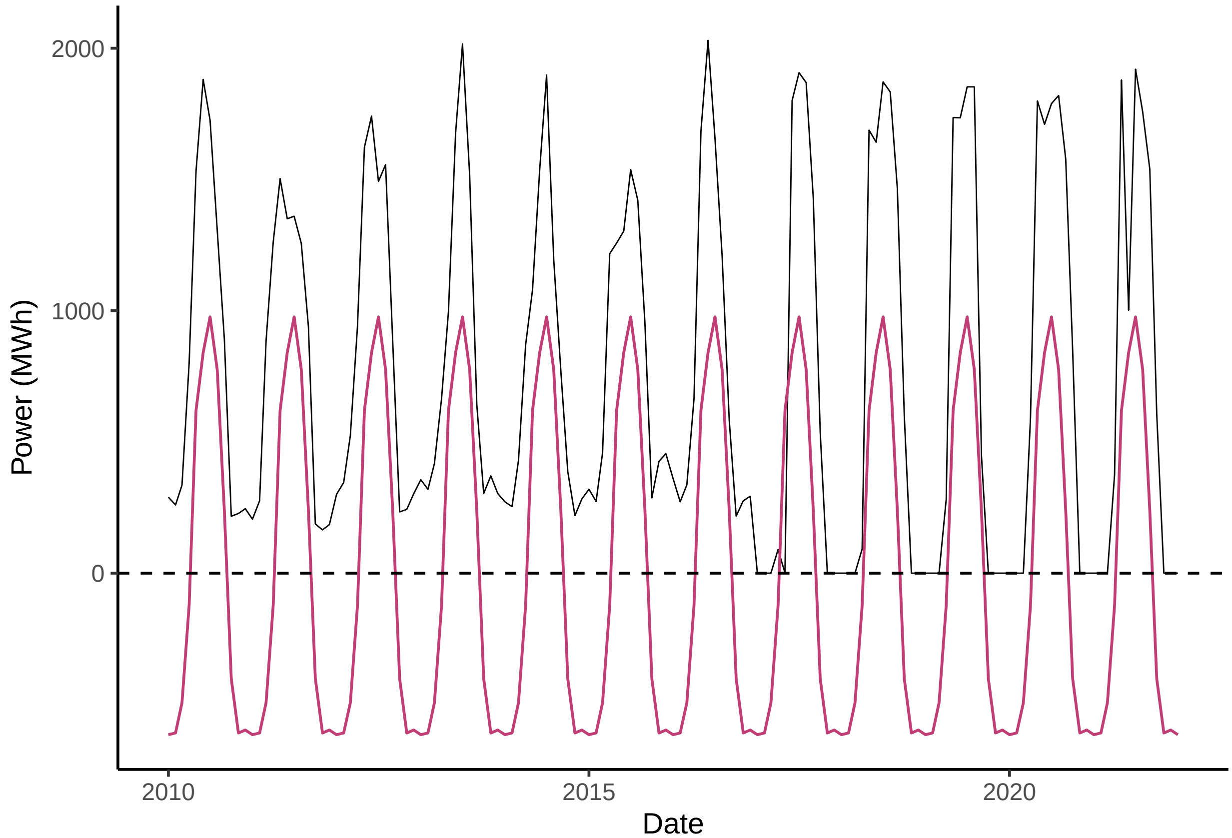
2500
2000
1500
1000
500
0
Power (MWh)
2000
2005
2010
2015
2020
Date

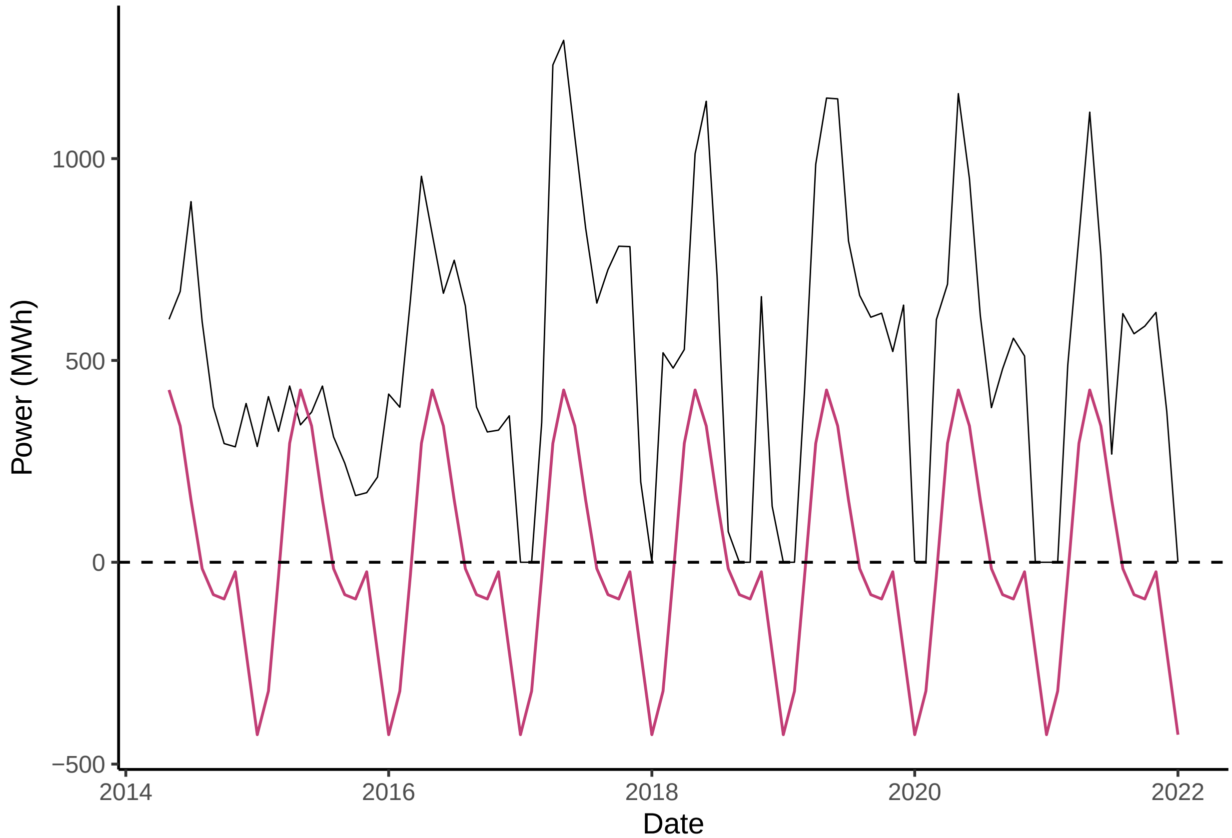

15000
10000
5000
0
Power (MWh)
2000
2005
2010
2015
2020
Date

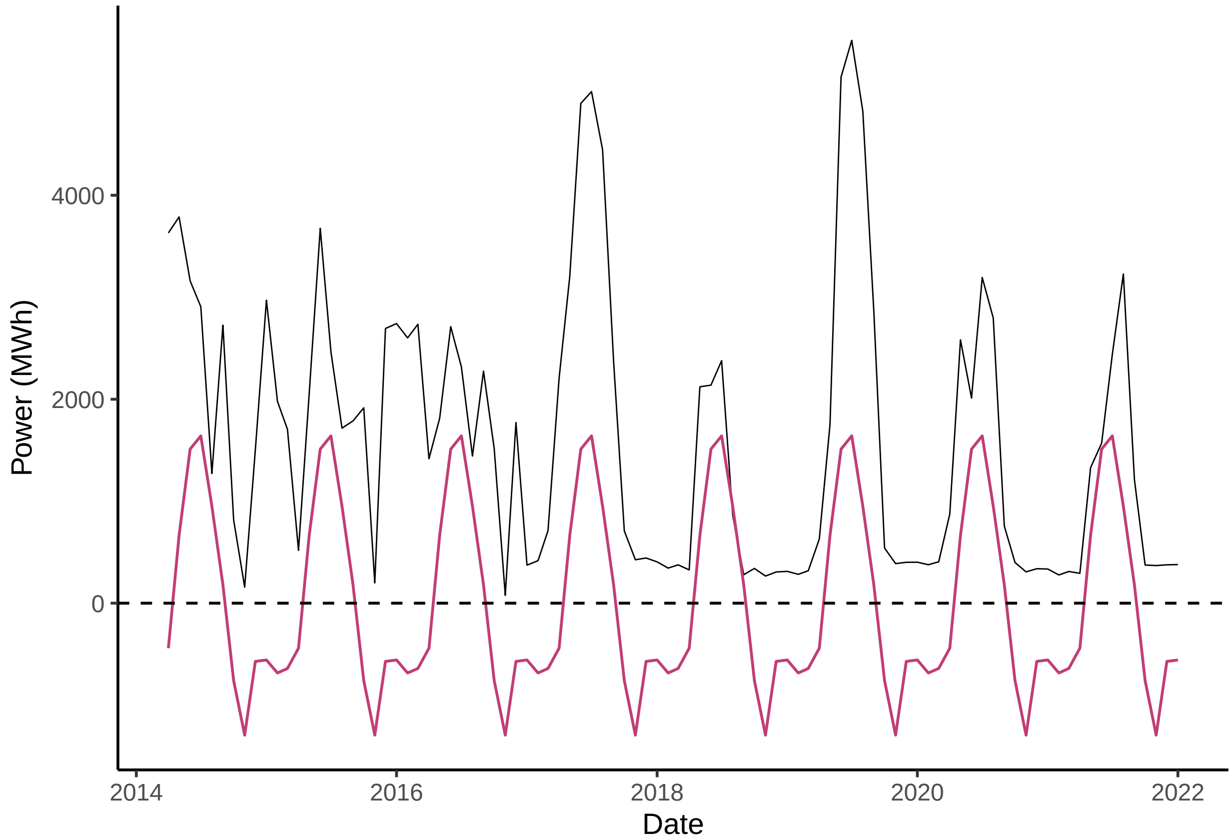

8000
6000
4000
2000
0
Power (MWh)
2000
2005
2010
2015
2020
Date



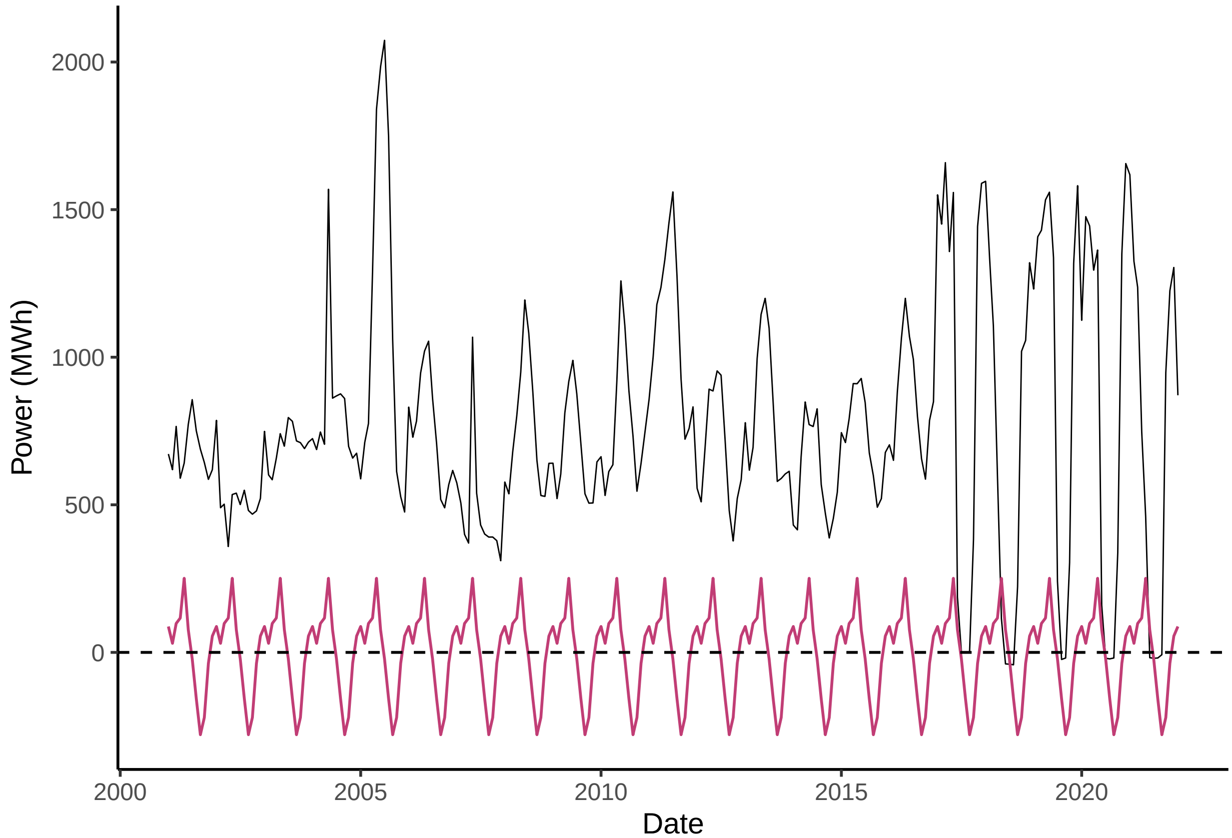
Power (MWh)
4000
2000
0
2000
2005
2010
2015
2020
Date


2000
1000
0
Power (MWh)
2010
2015
2020
Date


Power (MWh)
1000
500
0
−500
2014
2016
2018
2020
2022
Date



Power (MWh)
4000
2000
0
2014
2016
2018
2020
2022
Date



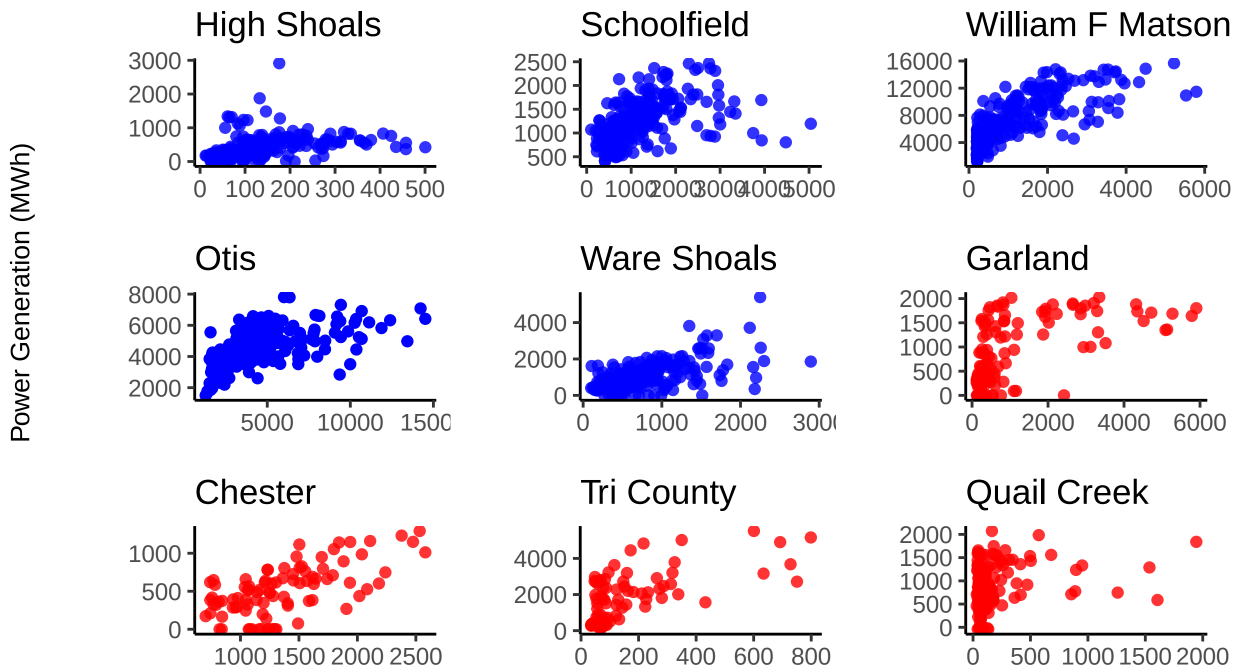
Combining Datasets

```
## Joining, by = c("Month", "Year")
## Joining, by = c("Month", "Year")
## Joining, by = c("Month", "Year")
## Joining, by = c("Month", "Year")

## Warning: Ignoring unknown parameters: apha

## Joining, by = c("Month", "Year")
## Joining, by = c("Month", "Year")
## Joining, by = c("Month", "Year")
## Joining, by = c("Month", "Year")
## Joining, by = c("Month", "Year")

## Warning: Removed 1 rows containing missing values (geom_point).

## Warning: Removed 1 rows containing missing values (geom_point).

## Warning: Removed 1 rows containing missing values (geom_point).

## Warning: Removed 1 rows containing missing values (geom_point).

## Warning: Removed 1 rows containing missing values (geom_point).

## Warning: Removed 1 rows containing missing values (geom_point).

## Warning: Removed 1 rows containing missing values (geom_point).
```

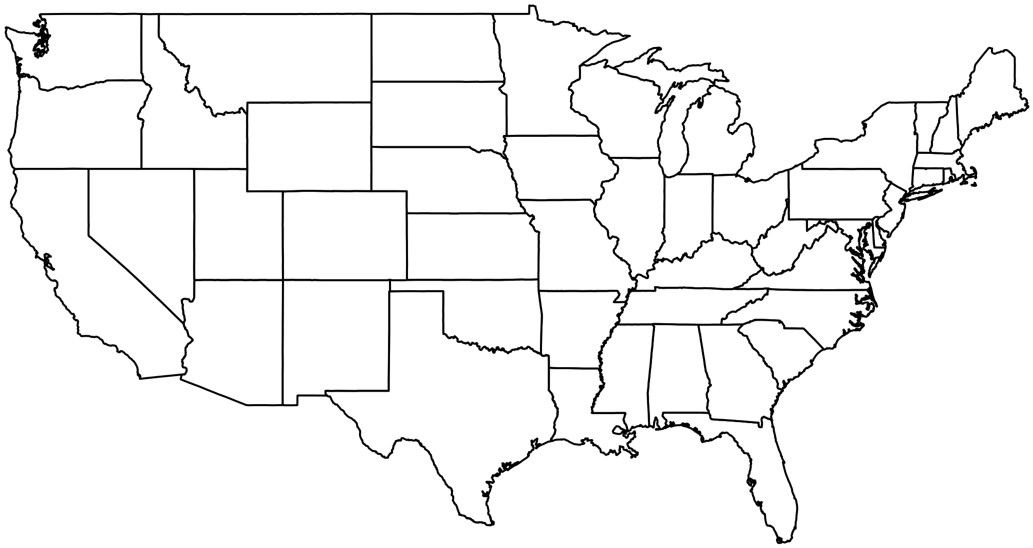
```
## Warning: Removed 1 rows containing missing values (geom_point).

## Warning: Removed 40 rows containing missing values (geom_point).
```

Figure 3: Scatterplot of Streamflow and Power Production

Blue indicates eastern states while red indicates western states.

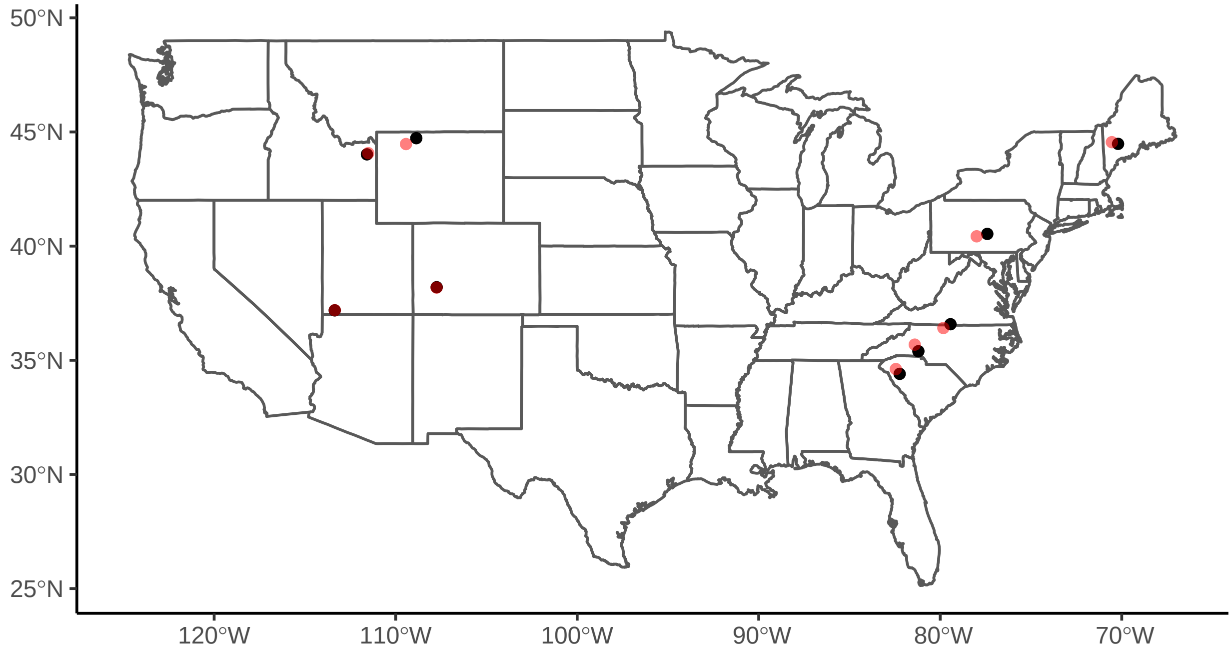


Locations of the streamflow gauges and hydroelectric plants were visualized in two distinct ways. Mapview displays the location of each hydroelectric plant, with each circle size representing the capacity of the individual plant relative to the other plants. Figure 4 uses ggplot and displays the locations of both the stream gauges and hydroelectric plants used in the analysis.



```
## PhantomJS not found. You can install it with webshot::install_phantomjs(). If it is installed, pleas
```

Figure 4: Streamflow Gauges and Hydroelectric Plant Locations



Question 3: Is there a linear relationship between streamflow and power generation, and if so what is the strength of that relationship?

To investigate the linear relationship between streamflow and power generation, the lm function was utilized on each combined dataset. Each analysis resulted in a statistically significant linear regression equation and a statistically significant independet variable. Power generation and stresm discharge displayed a positive linear relationship in every sample analyzed, and the results are shown in Table 3. For some hydroelectric plants, such as Tri County and Matson, greater than 40% of the variation in power generation can be explained by streamflow. However, for other plants, such as High Shoals and Quail Creek, less than 20% of the variation in power generation can be explained by streamflow. The linear relationship between streamflow and generation diminishes at higher streamflow levels due to the maximum generation capacity of the hydropower plants. The analysis suggests that additional explanatory variables are needed to better capture fluctuations in hydroelectric generation.

```
##
## Call:
## lm(formula = Power ~ average_monthly_discharge, data = high_shoals_power_water)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -542.70 -156.79  -67.22   55.86 2449.12
##
## Coefficients:
##                           Estimate Std. Error t value Pr(>|t|)
## (Intercept)               225.2626    32.9931   6.828 6.51e-11 ***
## average_monthly_discharge   1.3549     0.2069   6.549 3.27e-10 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 313.7 on 250 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.1464, Adjusted R-squared:  0.143
## F-statistic: 42.89 on 1 and 250 DF,  p-value: 0.0000000003274
```

```
##
## Call:
## lm(formula = Power ~ average_monthly_discharge, data = schoolfield_power_water)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1318.73  -262.49   -13.59   243.70  1030.45
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)               930.76885   44.85627  20.750  < 2e-16 ***
## average_monthly_discharge   0.26621    0.03076   8.653 6.22e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 403 on 250 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.2305, Adjusted R-squared:  0.2274
## F-statistic: 74.87 on 1 and 250 DF,  p-value: 6.222e-16
```

```
##
## Call:
## lm(formula = Power ~ average_monthly_discharge, data = william_f_matson_power_water)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6867.0 -1670.2    28.8  1684.9  5639.5
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)               4309.1249   223.0483   19.32   <2e-16 ***
## average_monthly_discharge    2.4239     0.1348   17.99   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2371 on 250 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.5641, Adjusted R-squared:  0.5624
## F-statistic: 323.5 on 1 and 250 DF,  p-value: < 2.2e-16
```

```
##
## Call:
## lm(formula = Power ~ average_monthly_discharge, data = otis_power_water)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3043.08  -544.91   -25.68   610.65  2860.38
##
```

```
## Coefficients:
##                             Estimate Std. Error t value Pr(>|t|)
## (Intercept)               3270.30164  120.85331   27.06   <2e-16 ***
## average_monthly_discharge    0.27928    0.02374   11.76   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 960.4 on 250 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.3562, Adjusted R-squared:  0.3537
## F-statistic: 138.3 on 1 and 250 DF,  p-value: < 2.2e-16
```

```
##
## Call:
## lm(formula = Power ~ average_monthly_discharge, data = ware_power_water)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1956.06  -306.51   -38.28   311.84  3032.69
##
## Coefficients:
##                            Estimate Std. Error t value   Pr(>|t|)
## (Intercept)               322.94213   67.51921   4.783 0.00000295 ***
## average_monthly_discharge   0.90850    0.07651  11.875    < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 589.5 on 250 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.3606, Adjusted R-squared:  0.3581
## F-statistic:   141 on 1 and 250 DF,  p-value: < 2.2e-16
```

```
##
## Call:
## lm(formula = Power ~ average_monthly_discharge, data = garland_power_water)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1212.9  -429.9  -186.3   416.1  1229.7
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)               466.93715   56.09438   8.324 6.41e-14 ***
## average_monthly_discharge   0.30864    0.03345   9.227 3.54e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 540.3 on 142 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.3748, Adjusted R-squared:  0.3704
## F-statistic: 85.14 on 1 and 142 DF,  p-value: 3.54e-16
```

```
##
## Call:
## lm(formula = Power ~ average_monthly_discharge, data = chester_power_water)
```

```
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -502.16 -182.22   24.57  171.26  534.19
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)               -125.79465   86.04129  -1.462    0.147
## average_monthly_discharge    0.46996    0.06002   7.830 9.12e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 256.6 on 90 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.4052, Adjusted R-squared:  0.3986
## F-statistic: 61.31 on 1 and 90 DF,  p-value: 9.116e-12
```

```
##
## Call:
## lm(formula = Power ~ average_monthly_discharge, data = tri_county_power_water)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2140.7  -738.2  -360.5   665.0  2833.3
##
## Coefficients:
##                           Estimate Std. Error t value Pr(>|t|)
## (Intercept)               827.8265   142.7765   5.798 9.58e-08 ***
## average_monthly_discharge   5.3617     0.6301   8.509 3.35e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1030 on 91 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.4431, Adjusted R-squared:  0.437
## F-statistic:  72.4 on 1 and 91 DF,  p-value: 3.355e-13
```

```
##
## Call:
## lm(formula = Power ~ average_monthly_discharge, data = quail_power_water)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -888.03 -249.25  -42.72  248.80 1268.56
##
## Coefficients:
##                           Estimate Std. Error t value  Pr(>|t|)
## (Intercept)               728.4747    33.8111  21.545   < 2e-16 ***
## average_monthly_discharge   0.4654     0.1162   4.004 0.0000863 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 419.2 on 211 degrees of freedom
##   (40 observations deleted due to missingness)
## Multiple R-squared:  0.07062,    Adjusted R-squared:  0.06621
```

Table 5: Table 5: Linear Regression Results

Hydropower_Plants	R2
High Shoals	0.146
Schoolfield	0.230
Matson	0.564
Otis	0.356
Ware Shoals	0.361
Garland	0.375
Chetser	0.405
Tri County	0.443
Quail Creek	0.071

```
## F-statistic: 16.03 on 1 and 211 DF,  p-value: 0.00008625
```

Summary and Conclusions

The positive linear relationship determined by our analysis indicates that higher stream discharges lead to higher hydropower generation. This is true until hydropower maximum generation capacity constraints are met, then power generation levels off. Stream discharge on average explained 30-40 percent of the variation in power generation. According to our tests and the data we used, there is no obvious overall trend for hydropower generation currently. However, specific hydropower plants are experiencing trends, which may represent a microcosm of the larger, geographically unique climate change effects that are expected to culminate in the future. We also saw that some streams are experiencing trends in discharge. Climate change will likely create increased discharge for some streams and decreased discharge for others. When discharge is below the level that leads to generation capacity it will affect the output of the hydropower plant. On the other side, if capacity is constantly being met, the plant could be expanded to have a higher capacity and then would be able to generate more power.

Limitations

The overall results of our analysis are limited by our sample size. Nine sites were chosen to streamline the analysis, however the findings, or lack thereof, of hydropower generation and streamflow trends related to climate change may be distorted or inaccurate due to extrapolations from the small sample. The sites chosen were random, however they may not be representative of the true trends and relationships between hydropower and streamflow. Site gauges were selected as close as possible to their associated hydropower plant, but our analysis does not account for potential changes to streamflow between the selected gauge and the generation facilities. Further, certain sites had less than 20 years of data, which limited the credibility of our analysis for those sites. Twenty years of data was selected since it was readily available from the EIA website. However, a longer time period would more accurately capture the effects of climate change on hydropower generation.

Future recommendations

Future research could extend our analysis to account for a greater scope of variables that may explain hydropower transformations due to climate change. These variables may include temperature, glacial melt, a power plant's structural characteristics, and the frequency, timing, and amount of rainfall. Future research could include a wider selection of sites, as well as international facilities. The vulnerability of hydropower plants to climate change could also be analyzed, and forecasting could be used to better understand the future impacts of climate change on hydropower production.