2019

MODELING AND FORECASTING GLACIER NATIONAL PARK VISITATION

Michael James Kernan
University of Montana

MODELING AND FORECASTING GLACIER NATIONAL PARK VISITATION

By

MICHAEL JAMES KERNAN

Bachelor of Arts in Economics, University of Wisconsin - Madison, Madison, WI, 2014

Thesis

presented in partial fulfillment of the requirements
for the degree of

Master of Arts
in Economics

The University of Montana
Missoula, MT

August 2019

Approved by:

Scott Whittenburg, Dean of The Graduate School
Graduate School

Douglas Dalenberg, Chair
Economics

Jeffrey Bookwalter
Economics

Jennifer Thomsen
Park, Recreation, and Tourism Management

Kernan, Michael, M.A., Summer 2019 Economics

Abstract

Chairperson: Douglas Dalenberg

National parks have recently seen increased visitation demand. Glacier National Park is located in an area where changes in weather and climate will occur at an accelerated rate. Changes in fire and precipitation regimes are taking place at a time when Glacier National Park is setting new visitation records. This paper uses regression and forecast models to investigate the changing landscape of park visitation.

Findings suggest important impacts to monthly visitation and cycling on Going-to-the-Sun Road in Glacier National Park are associated with forest fire activity and precipitation. The conclusions of this paper support perceptions about the effect natural occurrences have on park use and help give direction to actions that Glacier National Park can take to keep visitation high and promote efficient use of park facilities. Economic prosperity in communities around Glacier National Park relies on successful park operation bringing a seasonal influx of visitors each year. Adapting park planning to unpredictable natural events as well as forces outside the park will facilitate efficient use of resources in and around Glacier National Park.

This thesis is based upon work supported by UM BRIDGES through funding from the National Science Foundation under Grant No. DGE-1633831.

I would also like to acknowledge the support of my committee, Douglas Dalenberg, Jeffrey Bookwalter, and Jennifer Thomsen.

1 Introduction

The United States is the birth place of national parks as the world knows them. Huge swaths of land were set aside “for the benefit and enjoyment of the people” (National Park Service, 2018a). Yellowstone was established as the first National Park in 1872. Later, in 1916 the “Organic Act” was signed into law by President Woodrow Wilson, creating the National Park Service (NPS). It stated their “purpose is to conserve the scenery and the natural and historic objects and the wild life therein and to provide for the enjoyment of the same in such manner and by such means as will leave them unimpaired for the enjoyment of future generations” (National Park Service, 2018a). Since establishment of the parks and passage of this act, millions of people have indeed found enjoyment in national parks.

The United States spans a vast area and each NPS site has unique features which affect visitation. Recent years have set attendance records at some parks and many people make it their goal to visit every US national park, with the National Park Travelers Club even awarding achievement certificates to people as they rack up NPS unit visits (National Park Travelers Club, 2019). Whether a park is designated a national monument or national park also plays a role in visitors’ decision making processes. Weiler and Seidl (2004) find changes of designation from national monument to national park increases annual visits to the new national park while not detracting from the level of visits to other sites of lesser designations. NPS sites received 330,882,751 recreational visits in 2017, which was a slight dip from 2016 when the record was set at 330,971,689 recreational visits. Of these visitors in 2017, 3,305,512 were visitors to Glacier National Park, a record for Glacier and its first time exceeding three million visitors in one year (National Park Service, 2018b).

The Crown of the Continent, Glacier National Park (GNP) was established in 1910. It lies entirely in the US state of Montana and encompasses 4,100 square kilometers. The northern boundary of GNP is the Canadian border. Canada established Waterton Lakes National Park in the area north of the United States' park. Congestion in GNP is an issue during peak season, but new activities, like cycling on Going-to-the-Sun Road (GTSR), are becoming more popular leading to increased visitation demand over a greater part of the year. Natural events, such as forest fires, can cause visitation to deviate from expectations. Models to forecast visitation or bicycle use on GTSR take factors affecting these use numbers into account and give the park an opportunity to better understand how and when visitors will use GNP. The goal of this paper is to model determinants of park visitation and interpret them alongside forecasts of future visitation.

2 Literature Review

2.1 Tourism and National Park Visitation

The tourism industry works with a unique good; hotel rooms unfilled, restaurant tables left open, and seats on an airplane cannot be stored and sold later (Chu, 2009). The US NPS is part of the tourism industry and they must use resources efficiently while facilitating positive experiences for visitors to their sites. When expectations are off and fewer people visit or too many visitors come in a short period of time, effective park management becomes an even more formidable task. Macroeconomic forces, park fees, travel costs, whether GTSR is open to vehicles, bicycle tourism, precipitation, air quality/visibility, and forest fires are factors which influence visitation patterns to GNP.

There are 418 NPS sites and 117 charge an entrance fee, including GNP which has charged entrance fees since 1914. Glacier increased entrance fees twice during the observation period in this paper; in 2015 and 2018. The

2018 increase took effect June 1st and raised entry fees by $5 per vehicle, from $30 to $35 (Alley, 2018). With this fee increase, GNP expects an increase in entrance fee revenue, indicating they think visitation demand is inelastic (Sage, Nickerson, Miller, Ocanas, & Thomsen, 2017). Schwartz and Lin (2006) find an increase in entrance and usage fees enacted in 1995, known as the Recreational Fee Demonstration Program (RFDP), lowered visitation relative to what the level would have been expected to otherwise be. With 1986-1995 data, they fit multivariate models to predict counter-factual visitation and compare this with actual visitation, including a control group of sites which were not part of the RFDP. When actual numbers of visitors are compared to their predicted values, 83% of annual visitation numbers are below the counter-factual (Schwartz & Lin, 2006). This is consistent with the law of demand. Ostergren, Solop, and Hagen (2005) find through factor analysis that people who agree with the statement "entrance fees are too high" also agree with the statements "hotel and food costs are too high" and "service fees are too high." The clustering of these statements indicates people consider it costly to visit NPS sites in general and not solely because of park fees (Ostergren et al., 2005). The debate around park fees can be subjective, but in almost all circumstances a $5 fee increase will not alter the already small proportion of overall cost which the entrance fee makes up. The circumstance where fee increases will be consequential, though, is among visitors who live within a few hours' drive of GNP (Sage et al., 2017).

Tourism as a commodity is different from many other products because the consumer must be transported to the good (Barry & O'Hagan, 1972). National Parks are relatively low cost vacation options, but as costs increase people will substitute away from distant parks and opt for parks closer to home. It is particularly important to control for travel cost when a study looks at multiple parks (Poudyal, Paudel, & Tarrant, 2013). Spurred by high gas prices and gas rationing in the 1970s, Kamp, Crompton, and Hensarling

(1979) surveyed travelers to see how their stated preferences for traveling would change under higher gas prices or more severe rationing regimes. Responses varied depending on the reason for travel and the size of vehicle which the respondent owned, but there was a fairly even split between people saying they would either use alternative forms of transportation, shorten their trip, or stay home (Kamp et al., 1979). Gas availability was so tight at this time that the government even discouraged pleasure travel, resulting in a 13% decrease in National Park visitation relative to the previous year (Kamp et al., 1979). Oh and Hammitt (2011) also used surveys during a more contemporary period of high gas prices in the mid-2000s to ask how people intended to change their travel behavior. As expected, higher gas prices led to less travel or different forms of travel. Gas prices are often used to incorporate travel costs into models. When looking specifically at Glacier, considering how people plan their visits months in advance is important. If fuel prices rise, people who have been planning to visit for months may not cancel their trip. High fuel costs, though, will be factored into decision making during and for some time after the initial trip planning period; when people still have time to evaluate other options and substitute away from traveling a long distance. For this reason, lagged gasoline prices could be used instead of gasoline prices at the time of observation.

Pleasure travel is a discretionary good and financial confidence will impact this sector of the economy. Macroeconomic forces are found to influence how many vacations people take, so they should be considered when forecasting how many people will travel to a national park (Ritchie, Amaya Molinar, & Frechtling, 2010; Papatheodorou, Rosselló, & Xiao, 2010; Smeral, 2010). Poudyal, Paudel, and Tarrant (2013) found parks to be a normal good, whose visitation is adversely affected by slowdowns in economic conditions (Poudyal, Paudel, & Tarrant, 2013). Conversely, Weiler and Seidl (2004) find evidence in their econometric models that (new) national parks may be inferior goods, one which people consume more of when they feel less affluent. The possibility that national parks are inferior goods was first noticed by Johnson and Suits (1983), but they were unable to make this claim with certainty due to a high degree of collinearity between upward trends in visitation and income in their models. Poudyal, Paudel, and Tarrant (2013) look at the effect of recessions on aggregate national park visitation in the United States. They test five macroeconomic variables as proxies for recessions to see which one explains changes in visitation best. In models with monthly park visitation as the dependent variable, the unemployment rate variable was significant at the 10% level and variables for economic output, business cycle, consumer confidence, and consumer's expected inflation were significant at the 1% level. All variables had signs in the expected direction, indicating the macroeconomy is an important part of understanding use of national parks and that parks on the aggregate are a normal good (Poudyal, Paudel, & Tarrant, 2013).

Bicycles became popular in the 1890s and had a strong relationship with tourism. This subsided once automobiles became affordable to most people, but the emancipating ability of the bicycle is now experiencing renewed interest. Recreational cyclists and cycle tourists enjoy the physical exercise, freedom, escapism, relaxation, and peace of cycling; many of the same reasons why it was popular before adoption of the automobile as the main form of transit. GNP features Going-to-the-Sun Road which bisects the park from east to west and traverses the Continental Divide, shown in Figure 1. The challenge of cycling a road which crosses the continental divide is a draw for cyclists. GTSR has been identified as the primary visitor use facility in GNP (1977 Master Plan, 1988 Statement for Management). GTSR is listed on the National Register of Historic Places as a "significant cultural resource." There are limited roadways in GNP and when one becomes overcrowded, there are not alternative routes available. Due to this, bicycle use is prohibited on GTSR during the period June 15th - Labor Day from 11am-4pm when the road is open to automobile traffic (see Figure 1). Battaglia (2016) argues this restriction is unwarranted though.

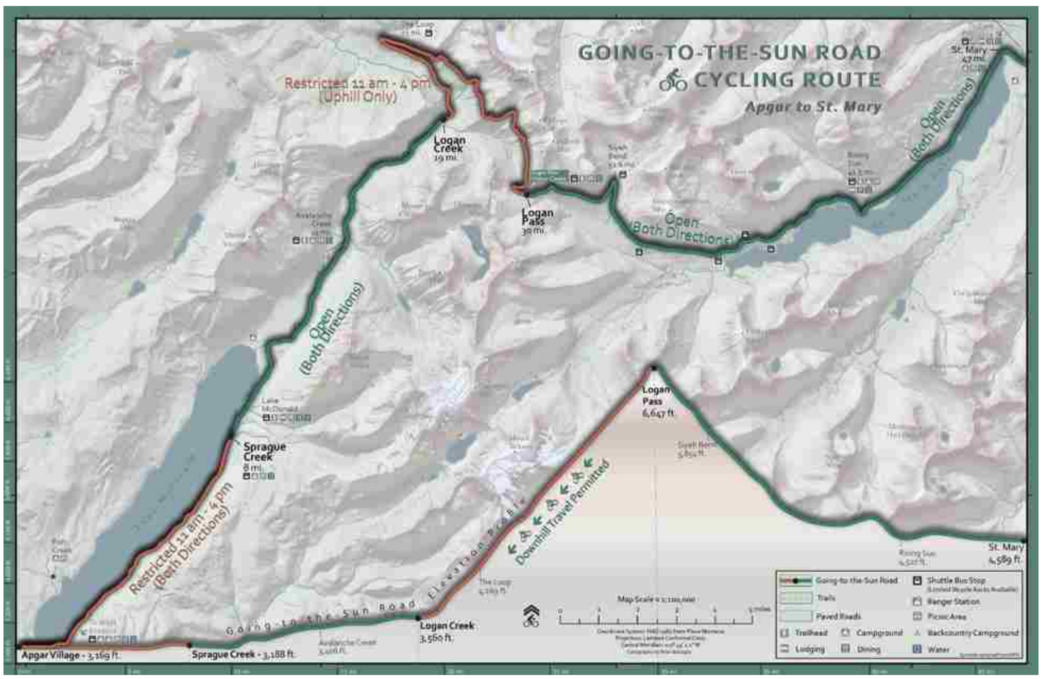


Figure 1: Map of Going-to-the-Sun-Road showing bicycle restriction areas and elevation change

(Battaglia, 2016)

Perceived danger from riding in traffic is found to be a major deterrent to cycling (Davies, Halliday, Mayes, & Pocock, 1997). The level of fear or comfort when riding with automobiles can be heavily based on the cyclist's level of experience (O'Connor & Brown, 2010). Converting old train tracks to bike trails, Rails to Trails, has been a bicycle tourism success in the US and one of the main draws is the comfort afforded from not having to share the route with automobiles (Tracy & Morris, 1998). This hints at a latent demand in the cycling tourism industry and in many ways is analogous to GTSR during the Spring when the road is open to bicycles but not cars. While passionate cyclists will ride regardless, cycling tourism becomes an option for the majority of people when the car/safety perception issue is not a factor. Trips to GNP to cycle GTSR will often still require a car trip to GNP and since GNP is such an attractive location, the park itself remains the prime attraction, which indicates it is not only serious cyclists finding pleasure riding their bikes on GTSR. A car-free GTSR presents the opportunity for leisure cyclists to use the bike as a means for taking in beautiful landscapes at a meaningful and enjoyable pace.

Automobiles are currently essential to the vast majority of people traversing GTSR, but cycling on the road should continue to be encouraged as a sustainable, destination based experience prior to opening of the road for motorized traffic (Lumsdon, 2000). Lamont (2009) evaluates previous ways of defining *bicycle tourism* and synthesizes what he feels best encapsulates this sector of the tourism industry as "trips away from an individual's home region, of which active or passive participation in cycling are considered the main purpose for that trip." In his definition he does not differentiate between same-day and overnight trips. GTSR cyclists would fall under this label of *bicycle tourism* since some camp in the park and surrounding area while others make day trips in from the region around GNP (Battaglia, 2016).

Air quality affects the experience people expect to have at a national park. Reduced visibility hampering one's ability to take in vistas and high levels of particulate matter instead of *clean, mountain air* lowers visitor satisfaction. This is a problem as Mace, Bell, and Loomis (2004) report haze from human-made pollutants in wilderness areas of the western US has already decreased visibility from 140 miles (225 km) to 33 - 90 miles (53 - 145 km) since preindustrial times. In a survey conducted by Great Smoky Mountains National Park, Papadogiannaki, Braak, Holmes, Eury, and Hollenhorst (2009) find 94% of responding visitors expect "viewing scenery/taking a scenic drive" to be an activity done while in the park. Poudyal, Paudel, and Green (2013) estimate the effect of visibility on park visitation at Great Smoky Mountains National Park. Every model tested finds the effect of lagged visibility to be positive and significant, indicating the park had higher visitation during and after months of good visibility.

When GNP's Sperry Chalet burned in 2017, part of the park's history was lost and there was an outpouring of fond memories from visitors to the chalet over the years (Ouellet, Nicky, 2017). Fires in 2018 were also disruptive, causing temporary closure of GTSR along with numerous other roads, trails, and campsites. Climate change is altering ecological regimes in the northern Rocky Mountains. Higher latitudes and elevations are predicted to experience greater degrees of change; the Northern Rockies in which GNP is located will be an area that notices these changes first (Westerling, Hidalgo, Cayan, & Swetnam, 2006). Fire regimes have a rapid response to changes in climate. They will have the most immediate effect on ecosystems, acting as a bellwether for climate change (Flannigan, Stocks, & Wotton, 2000). Another notable detail about the 2018 fires were how early in the season they began. The majority of recreational visitors to GNP come in July and August. These are the park's peak months. The 2018 Howe Ridge and Boundary Fires began in August leading to the first decline in visitation since 2011. Despite this, visitation still reached its second highest level and may have otherwise set another record if it were not for the fires (IRMA, NPS, GLAC annual visits, 2019). Duffield, Neher, Patterson, and Deskins (2013) assess forest fire effects on visitation to Yellowstone Nation Park (YNP). They find that while fire affects visitation, lagged variables for fires of previous years are not significant. Also, they find a negative relationship between fires in GNP and visitation in YNP, showing the two parks are complements (Duffield et al., 2013). McIntosh and Wilmot (2011) also found a complementary effect/positive externality between national parks near one another. Active fires in YNP could affect visitation to GNP.

2.2 Tourism Demand Forecasting

Forecasting tourism demand takes different forms depending on circumstances and desired understandings. Burkart, Medlik, et al. (1981) claim an understanding of tourism demand is needed for three reasons: to form value and significance of destinations, to support infrastructure and services planning, and for effective marketing. An understanding of demand can be achieved with quantitative methods, qualitative methods, or a combination of the two. Uysal and Crompton (1985) recommend using qualitative forecast methods to complement quantitative forecast methods, even when qualitative methods are not the only option available to the researcher. Regardless, Song, Turner, et al. (2006) find the majority of published studies use quantitative methods when forecasting tourism demand. In subsequent work where articles about tourism demand are analyzed, Song and Li (2008) find all but two of the 121 reviewed papers, which were published between 2000 and 2007, use quantitative methods. Song and Li (2008) find no single model to consistently outperform all others in every circumstance. However, qualitative methods may be the only option if the thing being forecast is so new there is not any data or if changes in the surrounding system have shifted the environment around the tourist destination dramatically.

Forecasters are always striving to improve their methods. Box-Jenkins forecasting models, known as ARIMA (autoregressive integrated moving average) models, came into widespread use in the 1970's (Box, Jenkins, Reinsel, & Ljung, 1970). Geurts and Ibrahim (1975) conducted an initial comparison of the Box-Jenkins approach versus an exponentially weighted average, another popular forecasting method. They present a single case study based on a monthly series of the number of tourist visits to Hawaii over the period 1952-1971. In comparing models they used the previous 24 months and forecasts one month ahead as their metric on which to judge the two approaches. Choosing how far ahead to forecast is an important detail when making model selection since some models are better with short-term forecasts and others are better at longer forecasts. One month ahead is practical for tourism management, so the two models are compared at that scale. Based on their choice of forecast periods and lead time the Box-Jenkins approach and exponentially smoothed models did equally well when their forecast errors were compared. Due to this Geurts and Ibrahim (1975) conclude it better to use an exponentially smoothed model for this monthly series because of ease in setting up and lower cost to run. Geurts (1982) follows up on his earlier study by again looking at overnight visitors to Hawaii, this time using only an exponential smoothing model. Forecast accuracy is improved through data modification. Using Hawaiian time-series data from the original Geurts and Ibrahim (1975) study and also using updated data through 1976, Geurts replaces outliers with forecast values. When this approach is used, 16 of 91 periods from May 1969 to January 1977 are deemed atypical and replaced with a forecast value. This reduced forecasting error for the exponential smoothing model from 10% to 7.5% (Geurts, 1982).

3 Data

3.1 Monthly Recreational Visits

Monthly recreational visitor counts come from the Integrated Resource Management Applications (IRMA) portal on the National Park Service website (National Park Service, 2019a). Counts for Glacier National Park (abbreviated GLAC here) are available on the IRMA page in the "STATS (Park Visitor Use Statistics)" section of IRMA-Applications by selecting GLAC as the park and then clicking on "Recreation Visits by Month (1979 - Current Calendar Year)" (National Park Service, 2019b). The NPS makes reports about how they conduct counts available at the bottom of this page as well under the "Visitor Use Counting Procedures" tab. There are inductive loop traffic counters at the West Glacier, Saint Mary, Many Glacier, Two Medicine, and Camas entrance lanes. Pneumatic tube traffic counters are at the Polebridge and Walton/Goat Lick entrances. All traffic counts are multiplied by the persons-per-vehicle (PPV) multiplier of 2.9. Buses, non-reportable vehicles, non-recreation vehicles, and duplicate re-entries are not included as part of this count. Bus counts are later added in from counts made by fee collectors. Some counters are removed during the winter months. At the three entrances where this happens, visitation is estimated at 50 visitors per month during the the months when the counter is not in place. Estimated visits are added in for two additional locations during peak and shoulder season months, too. Belly River Trailhead accounts for another 5,000 in May and September, with 10,000 added on in June, July, and August. Similarly, Cut Bank has estimated visits of 1,500 in May and September, and 3,000 in June, July, and August. Aside from the use of pneumatic tube traffic counters at Polebridge and Walton/Goat Lick, and a higher when-counter-is-removed winter estimate for Walton/Goat Lick (100 per month), the counting methods have remained the same since at least 2003, which is further back than the data I use goes.

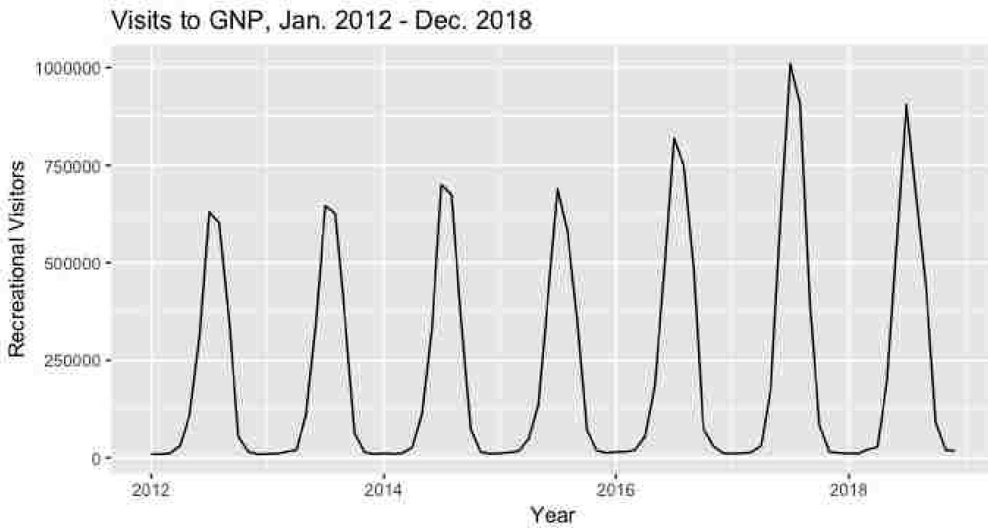


Figure 2: Monthly visits, January 2012 through December 2018

Figure 2 shows a strong seasonal visitation pattern to GNP. Table 1 lists yearly annual recreation visits. Annual recreational visits grew from 2.2 million visitors in 2012 to almost 3 million visitors in 2018. Table 2 gives monthly descriptive statistics.

Table 1: GNP Yearly Visitation Totals 2012-2018

Year	**Recreational Visits**
2012	2,162,035
2013	2,190,374
2014	2,338,528
2015	2,366,056
2016	2,946,681
2017	3,305,512
2018	2,965,309

Table 2: Monthly Summary Statistics GNP Recreational Visits 2012-2018

Month	**Mean**	**Std. dev.**	**Median**	**Min.**	**Max.**
January	12,133	1,718.8	12,087	10,318	15,674
February	12,416	2,332.2	11,847	10,194	16,548
March	16,831	3,690.8	15,758	12,416	21,758
April	35,010	12,131.3	31,594	20,922	55,125
May	146,401	37,972.5	134,741	108,998	195,116
June	438,329	119,297.8	414,671	313,713	620,962
July	771,202	144,239.5	699,650	630,093	1,009,655
August	687,065	112,216.9	667,688	579,007	908,479
September	389,094	51,094.2	356,975	351,388	482,592
October	68,046	21,981.1	72,694	25,965	91,973
November	19,052	5,625.9	16,158	14,924	30,823
December	12,856	2,953.6	12,877	9,862	18,781

3.2 Going-to-the-Sun Road Cyclists

Bicycle count data for Going-to-the-Sun Road is from 2016, 2017, and 2019. Data from 2018 was not available due to malfunctioning counters. The final twelve days before opening of GTSR to automobile traffic in 2019 are also not available because counters were removed for road maintenance work. Data was collected at four locations: Avalanche Creek on GTSR for eastbound and westbound bikers, the Camas bike path, and Camas road. The eastbound/westbound Avalanche Creek GTSR data is the only location used in this analysis. 2016 data spans from May 2nd to June 16th, 2017 data from May 6th to June 27th, and 2019 runs from May 3rd to June 10th. Cyclists were counted using JAMAR Technologies, Inc. TRAX Cycles Plus equipment. Direction of travel was recorded along with the number of bicycles.

The period while GTSR is closed to motorized traffic but open to cyclists is focused on because cyclist numbers are trivial and do not exhibit a pattern once the road opens to automobiles. Weekends see a much higher level of bicycle traffic relative to weekdays and weather is also an important determinant of how many cyclists visit GNP. Figure 3 shows cyclist counts in 2016. The weekly weekend increase of cyclists is noticeable every weekend in Figure 3 except the weekend of May 21-22nd, 2016 when rainy weather interrupted the usual weekend seasonal pattern. Weekend seasonality in 2017 is also noticeable in Figure 4. Average daily eastbound and westbound cyclists were 126 and 119, respectively, in 2016. In 2017, average daily cyclist totals were 186 eastbound cyclists and 180 westbound cyclists. There was a large increase of cyclists between 2016 and 2017 as riding GTSR gained popularity. 2019 saw a number of cyclist similar to that in 2017.

Figure 3: 2016 Cyclists When GTSR Not Open to Motorized Traffic

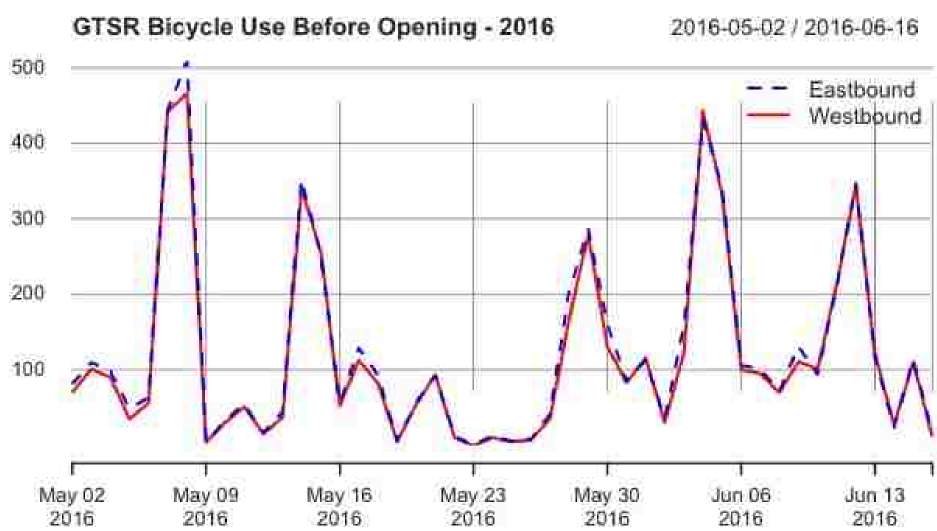


Table 3: Bicycle use on GTSR in 2016 **before** opening to automobiles

	Eastbound Bicycles	Westbound Bicycles
Mean	126	119
Std. dev.	128.7	126.2
Median	93	87
Minimum	0	0
Maximum	509	466

Figure 4: 2017 Cyclists When GTSR Not Open to Motorized Traffic

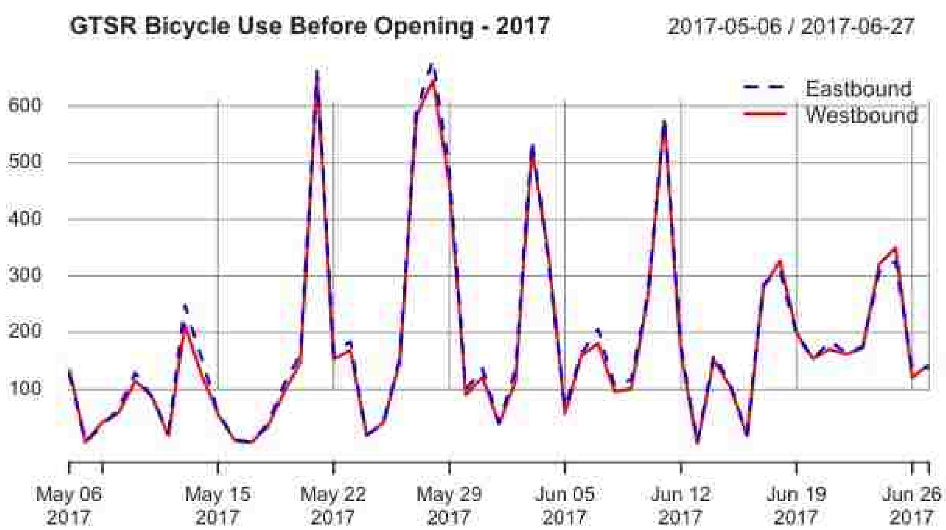


Table 4: Bicycle use on GTSR in 2017 **before** opening to automobiles

	Eastbound Bicycles	Westbound Bicycles
Mean	186	180
Std. dev.	169.4	167.6
Median	153	144
Minimum	4	3
Maximum	681	651

3.3 Fires in GNP, 2012-2018

GNP regularly publishes news releases to their NPS web page. All information on fire activity from 2012-2016 is based on these news releases. The News Releases are supplemented with information from Inciweb for fires in 2017 and 2018. Important fires are described in Table 7 and are those which caused closures of key areas within GNP. Information on the 2012 Avalanche Fire is from the following NPS News Releases, Glacier National Park (2012a, 2012b, 2012c, 2012d, 2012e). There were not any fires of note in 2013 or 2014. A News Release for the 2015 Reynolds Fire states its start date, location, and the restricted park access it caused, including GTSR closure (Glacier National Park, 2015a). Another important fire in 2015 was the Thompson Fire and there are News Releases describing it (Glacier National Park, 2015b, 2015c). There were again no fires of note in 2016, but 2017 was a different story. Some information on the 2017 Sprague Fire is taken from a park News Release (Glacier National Park, 2017a) and other details are found in the Inciweb report about the fire (Incident Information System, 2017a). The other fire which caused closures in 2017 was the Adair Fire (Incident Information System, 2017b). The two large fires in 2018 were the Howe Ridge Fire (Incident Information System, 2018a) and the Boundary Fire (Incident Information System, 2018b).

Table 5: Disruptive Forest Fires in GNP 2012-2018

Name and Year	Period Within Year	Description
Avalanche Fire (2012)	September 1st - 30th	45 acres. Relatively small fire, but it caused closure of the Avalanche Lake trail and campground, both of which are heavily used.
Reynolds Fire (2015)	July 21st - unstated	Counts for fire in my July 2015 data. Caused closure of GTSR and evacuations throughout the park beginning the same day it was first reported.
Thompson Fire (2015)	August 9th - unstated, assumed November 1st	>14,900 acres. Caused closures throughout the park, including closure of GTSR.
Sprague Fire (2017)	August 10th - unstated, assumed November 1st	16,982 acres. Caused closures throughout the park, including closure of GTSR.
Adair Peak Fire (2017)	August 12th - unstated, assumed November 1st	4,074 acres. Caused closure of the Inside North Fork Road from Polebridge Ranger Station to Camas Creek, and the Logging Creek and Quartz Creek Campgrounds.
Howe Ridge Fire (2018)	August 11th - November 1st	14,522 acres. Caused closures throughout the park, including closure of GTSR.
Boundary Fire (2018)	August 23rd - November 1st	2,911 acres. Caused area closures in northern GNP and Waterton Lake Park, Canada.

3.4 Fires in Yellowstone National Park, 2012-2018

Events that affect visitation to Yellowstone National Park (YNP) could also influence visitation to GNP. Duffield et al. (2013) found complementary effects between the these two national parks. YNP publishes news releases similar to GNP. There were not any major fire related YNP closures in 2012 despite there being some large fires in the backcountry (Maughan, 2012; Montag, Hogle, & Pirillo, 2012). Fires in 2013 were similar to 2012, large but mostly in areas that do not experience much human activity. The Alum

Fire in August 2013 did cause a road closure from Canyon Village to Fishing Bridge though, so August 2013 is entered as an active fire month in the data (Gabbery, 2013; Insider, 2013). News Releases on NPS.gov for Yellowstone NP begin in 2014. There are not any news releases pertaining to fire in 2014. 2014 appears to have been a quiet fire year (Insider, 2014). There does not seem to have been any major closures from to the two main fires (at least four total fires) in 2015, but they were large and smokey so September is counted as a fire month (Yellowstone National Park, 2015a, 2015b, 2015c, 2015d, 2015e). 2016 was a big fire year at YNP. That year had the most acres burn since the historic fires of 1988 (Yellowstone National Park, 2016). Inciweb created a report on the largest fire in 2016, the Maple Fire (Incident Information System, 2016). Fires in 2017 were all .1 acre or smaller, burning less than one acre of YNP and causing no closures. Of the fires in 2018, the Bacon Rind fire caused closures in and around the park (Yellowstone National Park, 2018; Incident Information System, 2018c). August through October 2018 are counted as months with fire activity affecting park use and enjoyment. 2012-2018 fires in YNP are listed in Table 8.

Table 6: Disruptive Forest Fires in YNP 2012-2018

Name and Year	**Period Within Year**	**Description**
Alum Fire (2013)	August	>6,150 acres. Closed 13 miles of the Grand Loop Road between Fishing Bridge Junction and the South Rim Drive of the Grand Canyon of the Yellowstone.
Spruce Fire (2015)	September 9th - September 30th	>2,594 acres. No closures caused in the park, but created enough smoke to likely impact visitor experiences.
5L4 Fire (2015)	August 24th - September 30th	∼16 acres. Was not very active, but did cause closure of some backcountry campsites.
Maple Fire (2016)	August 4th - November 1st	>45,000 acres. Largest of the big fires in 2016. Caused various trail and road closures in YNP.
Bacon Rind Fire (2018)	July 20th - November 1st	5,232 acres. Caused trail and road closures in and around YNP.

3.5 GTSR Peak Season Closures

Opening and closing dates for GTSR from 2012-2016 are provided by the park and are publicly available (Glacier National Park, 2017b). 2017 and 2018 dates were found using the park's News Releases page. The number of days in a month when GTSR is entirely open to traffic is what I counted and used as data points. This is summarized in Table 9. Below is detailed when deviations from the open/close dates reported in Glacier National Park (2017b) are made. Deference is given to news releases published by the park on their website when the two sources contradict each other, because the news releases provide a clearer picture of why and to what extend the road was closed.

There was a road closure for 48 hours in July 2012 due to a rock and mud slide (Glacier National Park, 2012f, 2012g, 2012h). Later that summer GTSR was closed from the west side beginning on 9/17 for accelerated rehabilitation work (Glacier National Park, 2012i, 2012j). GTSR may have been closed for a day in July 2014 because of a boulder and snow pile blocking a lane of traffic (Glacier National Park, 2014). The times used in the News Release do not make sense though, unless someone accidentally used *p.m.* for *a.m.* when describing the times of events. Due to this no days were deducted from the number of days GTSR was open in July 2014. Fire caused GTSR closure in 2015 beginning on July 21st (Glacier National Park, 2015d). The west side to Logan Pass reopened July 29th (Glacier National Park, 2015e) and GTSR re-opened completely on August 7th (Glacier National Park, 2015f). In 2015 GTSR closed from the east side after October 4th for road rehabilitation work and closed on the west side after October 18th (Glacier National Park, 2015g). The pdf with all the opening and closing dates lists June 16th, 2016 as the opening date that year, but the park's news release makes it appear that it actually opened one day later on June 17th (Glacier National Park, 2016). GTSR opened completely to motorized traffic on June 28th, 2017 (Glacier National Park, 2017c). GTSR was closed on its west side for most of September 2017 due to fire activity. The closure began on September 3rd (Glacier National Park, 2017d) and lasted until at least September 28th (Glacier National Park, 2017e). There are not any press releases after 9/28/17 detailing the status of the road. Looking at the park's twitter page (where they also publish news releases), it says the road closed on 10/1/17, reopened 10/5/17, then closed again on 10/7/17 and did not reopen that year. GTSR opened on June 23rd, 2018 (Glacier National Park, 2018a). The west side of GTSR was closed for a significant part of the 2018 season due to fire; from August 12th - September 16th (Glacier National Park, 2018b, 2018c). Based on news releases (Glacier National Park, 2018d, 2018e) it appears that Logan Pass closed to visitors starting 9/29/18 due to weather and did not reopen. GTSR opened on June 23rd in 2019.

Table 7: Opening and Closing Dates of GTSR 2012-2018

Year	**Open**	**Close**	**Temporary Closures/Notes**	**Total Days Entire Road Was Open**
2012	June 19th	October 15th from east September 16th from west	July 17th - July 19th (Landslide)	88 days: 12 in June, 29 in July, 31 in August, and 16 in September
2013	June 21st	September 23rd	None to report	94 days: 10 in June, 31 in July, 31in August, and 22 in September
2014	July 2nd	September 22nd	None to report	82 days: 30 in July, 31 in August, and 21 in September
2015	June 19th*	October 5th from east October 18th from west	July 21st - August 6th (Fire) *opened from west side to Logan Pass on June 11th	91 days: 12 in June, 20 in July, 25 in August, 30 in September, and 4 in October
2016	June 17th	October 12th	None to report	117 days: 14 in June, 31 in July, 31in August, 30 in September, and 11 in October
2017	June 28th	October 7th	September 3rd - September 28th (Fire)	71 days: 3 in June, 31 in July, 31 in August, 4 in September, and 2 in October
2018	June 23rd	September 29th	August 12th - September 16th (Fire)	64 days: 8 in June, 31 in July, 12 in August, and 13 in September

3.6 Days in a Month with Precipitation

Precipitation data is provided by the ncdc.noaa.gov website and its Climate Data Online portal. For monthly data, the number of days in a month with any precipitation are counted and used as the precipitation data point for that month. For daily data, each day is a 0-1 dummy variable that equals 1 on days with precipitation. Data for West Glacier were made using observations from the West Glacier weather station. Whenever there was a missing day in the West Glacier precipitation data, data from the Kalispell, MT weather station were used instead. Precipitation days for East Glacier are from the East Glacier weather station. Missing data points for East Glacier were filled in with those from St. Mary. There were still seven observations unavailable from East Glacier or St. Mary, these were taken from Many Glacier. East Glacier had many more missing and replaced precipitation observations than did West Glacier. A day with any amount of precipitation counts towards the monthly tally and the daily dummy value.

All precipitation days for the GTSR cyclist count data set were obtained from the same source and are taken from the West Glacier weather station. There was one day missing, June 7th, 2019, and it was filled in with data for that day from the Kalispell, MT weather station.

3.7 Macroeconomic Indicators

Macroeconomic data for the variables unemployment rate and consumer sentiment were downloaded from the FRED St. Louis Fed website, `https://fred.stlouisfed.org/`. The non-seasonally adjusted civilian unemployment rate was downloaded from `https://fred.stlouisfed.org/series/UNRATENSA` and the seasonally adjusted civilian unemployment rate was downloaded from `https://fred.stlouisfed.org/series/UNRATE`. Consumer sentiment is an index recorded by Michigan State University, also made available from the FRED at `https://fred.stlouisfed.org/series/UMCSENT/`. I have re-indexed their data to make my start year and month (Jan. 2012) the base year/month. These sites were last accessed on September 27th, 2019.

3.8 Average National Gasoline Prices

Gas prices are from the U.S. Energy Information Administration. They are monthly average national prices per gallon in U.S. dollars, including taxes. The monthly average prices used range from November 2011 to December 2018. November and December of 2011 are included so one month and two month lags can be tested. The data is publicly available at `https://www.eia.gov/totalenergy/data/browser/index.php?tbl=T09.04`. This site was last accessed on September 27th, 2019.

4 Empirical Strategy

Forecasting can be broken into sub-categories based on methods used. They are time-series methods and causal econometric models. Time-series methods forecast ahead by using a variable's past values and random error terms to predict what future values will be. This is done by focusing on trends and patterns, such as seasonality, in the time-series data. Time-series models can be straightforward and less costly because they only require previous observations of the variable of interest. Econometric models analyze causal relationships between the variable of interest (the dependent variable) and whatever explanatory variables are believed to explain movements of the dependent variable. They measure how change in one of the explanatory variables affects the dependent variable. Gathering data on the explanatory variables is where econometric models become more difficult than time-series approaches and often the econometric models do not produce better forecasts.

Econometric analysis is good for more than straight forecasting though, since there are things which cannot be forecast but are none-the-less scenarios which might occur (Song & Li, 2008). One example is forest fires, nobody knows when forest fires will happen based on previous trends and seasonality in the data, but an econometric model can be set-up and tested as a scenario, giving insight into the relationship between forest fires and visitation to GNP. Econometric analysis can interpret the change in visitation demand given a set number of days in a month when disruptive forest fires are burning, based on how large fires have previously affected GNP visitation. This is done under the assumption of *ceteris paribus*, holding all else constant. Holding the other explanatory variables constant, how will n days of severe forest fire change visitation in a given month? This perspective is useful both for evaluating existing park policies and making informed policy recommendations.

Monthly visitation data for GNP shows seasonality and could have a trend depending on the time window assessed. Bicycle data exhibits its own form of seasonality, weekend cyclist use of GTSR is much higher than on weekdays. Between 2016 and 2017 bike data shows a trend, but it is hard to draw many conclusions since this is only two years of data. An upward trend was assumed to have continued in 2018, but the bicycle counters malfunctioned so this cannot be said for certain. Once 2019 cyclist counts became available, the upward trend had continued, but not at the same high rate as between 2016 and 2017. Weekday use actually decreased between 2017 and 2019, but the increases on weekends and Memorial Day were enough to keep the overall trend up.

The seasonal component plays the biggest part in both GNP data sets that are used. Seasonal components can be dealt with in two ways. Monthly dummy or weekend dummy variables can be included which show and separate the seasonal effect generated by each month or the weekend. Seasonal and Trend decomposition using Loess Smoothing (STL) breaks the data into the three parts in Equation 1. These are the trend (T), seasonality (S), and remainder (R). Once the seasonal part of each time period is known, it can be subtracted from observations of the corresponding time period so data can be seen minus its seasonal component. The data are de-seasonalized once this is done.

$$y_t = T_t + S_t + R_t{}^1 \tag{1}$$

Goodness of fit of forecasts can be judged in many ways and most are based on seeing which model keeps prediction error the smallest or which one explains the most variation in movement of the dependent variable. When making model selection, the model which fits best is found by checking which one has the lowest AIC (Akaike's Information Criterion) (Akaike, 1974), AICc (an AIC corrected for small sample size), MAPE (Mean Absolute Percentage Error), or RMSE (Root Mean Square Error). R^2 is often used when comparing regressions, it shows what percent of variation in the dependent variable is explained by the model with an R^2 close to 1 indicating good fit. When comparing which model to use, these measures should be compared and factored in to the model selection process. Other background knowledge should be considered by the forecaster, but minimizing the AICc, for example, is a good way to find the model which will likely forecast best because it is the model which fits its movements most closely to the data being used. Goodness of fit shows which models are likely to perform best, so it aids in model selection.

[1] Hyndman, Rob J and Athanasopoulos, George (2018)

4.1 Regression Models

Time-series regressions estimate y at time t, y_t, through changes in explanatory variables $x_{1t}, \ldots, x_{kt}$. Here the dependent variable y is monthly visitation or daily bicycle use on GTSR. The magnitude of a predictor variable's effect is determined by its coefficient, $\beta_1, \ldots, \beta_k$. β's are estimated by choosing those which minimize the sum of squared errors. This all takes the form seen in Equation 2. Regressions do require some assumptions about their error terms. They must have zero mean, not be autocorrelated, and not be related to the predictor variables. Variables thought to explain variation in monthly visitation and GTSR bicycle use are selected for each model. Those used to explain monthly visitation to GNP are a ranking of local air quality, the unemployment rate, consumer sentiment, gas prices, days that month with precipitation in East Glacier, days that month with precipitation in West Glacier, active forest fire in GNP dummy, active forest fire in YNP dummy, and days that month GTSR was open to automobiles. Bicycles on GTSR explanatory variables are a dummy variable for the year 2016, a dummy variable for the month of May, a dummy variable for if it is a weekend day, a dummy variable for Memorial Day, the number of weeks it is away from opening of GTSR to motor traffic, a day-with-precipitation dummy, a measure of a day's precipitation, and maxiumum and minimum temperatures for that day.

$$y_t = \beta_0 + \beta_1 x_{1,t} + \beta_2 x_{2,t} + \cdots + \beta_k x_{k,t} + \varepsilon_t \quad (2)$$ [2]

Predicting y is done using the estimated coefficients while letting the error term be zero. The regression equation for estimating y_t now gets a hat, $\hat{y}$, to show it is an estimate. The β's get hats since they are estimates as well, $\hat{\beta}$'s. This is shown in Equation 3. After fitting the model with coefficient estimates, error term assumptions are assessed by plotting model residuals.

[2] Hyndman, Rob J and Athanasopoulos, George (2018)

This includes looking at the autocorrelation function (ACF) plot of residuals. It is common to find autocorrelation in the residuals when using time series data, an assumption violation that can be corrected.

$$\hat{y}_t = \hat{\beta}_0 + \hat{\beta}_1 x_{1,t} + \hat{\beta}_2 x_{2,t} + \cdots + \hat{\beta}_k x_{k,t} \quad (3)$$

[3]

4.2 Exponential Smoothing

Exponential smoothing models generate forecasts from a representation of the trend and seasonality in the data. This is done by calculating averages where the weight of each average decreases exponentially as the observations go further back in the past. Equation 4 is an example of simple exponential smoothing, which leaves out trend and seasonality, but shows how older observations are weighted less and less because $0 \leq \alpha \leq 1$.

$$\hat{y}_{T+1|T} = \alpha y_T + \alpha(1-\alpha) y_{T-1} + \alpha(1-\alpha)^2 y_{T-2} + \cdots \quad (4)$$

[4]

The GNP monthly data displays seasonality and in recent years it also appears to show an upward trend. Some changes to the simple exponential smoothing model are required to bring seasonality and trend into the equation. Using Holt (Holt, 1957) and Winters's (Winters, 1960) methods, researchers incorporate trend and seasonality into the exponential smoothing models. In the equations below, ℓ is the level, b is the trend, and s is the seasonal component. These each have their corresponding smoothing parameters α, β^*, and γ, respectively. The smoothing parameters are estimated by minimizing root mean squared error (RMSE).

The Holt-Winters method brings seasonality and trend into models in additive or multiplicative ways. Sensitivity analysis is performed to determine

[3] Hyndman, Rob J and Athanasopoulos, George (2018)
[4] Hyndman, Rob J and Athanasopoulos, George (2018)

what combination of these fits the GNP data best. An additive model is best when the seasonal variations are fairly constant throughout the series, since as seen in Equation 5 the same seasonal swing from m periods ago is added on, m being based on how the data is structured. For example, with monthly data $m = 12$.

$$\hat{y}_{t+h|t} = \ell_t + hb_t + s_{t+h-m(k+1)} \quad (5)$$

[5]

$$\begin{aligned}
\ell_t &= \alpha(y_t - s_{t-m}) + (1-\alpha)(\ell_{t-1} + b_{t-1}) \\
b_t &= \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)b_{t-1} \\
s_t &= \gamma(y_t - \ell_{t-1} - b_{t-1}) + (1-\gamma)s_{t-m},
\end{aligned}$$

Multiplicative models are used when seasonal variations change relative to the level of the series and one is shown by Equation 6. The monthly data for GNP seems to do this so I expect a multiplicative model will perform better than an additive one on untransformed GNP monthly visitation data.

$$\hat{y}_{t+h|t} = (\ell_t + hb_t)s_{t+h-m(k+1)} \quad (6)$$

[6]

$$\begin{aligned}
\ell_t &= \alpha\frac{y_t}{s_{t-m}} + (1-\alpha)(\ell_{t-1} + b_{t-1}) \\
b_t &= \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)b_{t-1} \\
s_t &= \gamma\frac{y_t}{(\ell_{t-1} + b_{t-1})} + (1-\gamma)s_{t-m}
\end{aligned}$$

A model predicting a linear trend into the future would eventually lead to under or over prediction, particularly in a place like GNP that has limited

[5] Hyndman, Rob J and Athanasopoulos, George (2018)
[6] Hyndman, Rob J and Athanasopoulos, George (2018)

capacity. There might be an upward trend in visits to GNP, but any upward trend can only be extrapolated so far before its growth explodes and becomes unrealistic. Trends can be damped to prevent this by imposing an asymptote. In Glacier's case, a damped trend would imply a ceiling to the possible number of visitors to the park. Damped trends can be used with additive and multiplicative models. Equation 7 shows a damped multiplicative model. The damping parameter is ϕ and it is between 0 and 1.

$$\hat{y}_{t+h|t} = \left[\ell_t + (\phi + \phi^2 + \cdots + \phi^h)b_t\right] s_{t+h-m(k+1)}.^{7} \tag{7}$$

$$\begin{aligned}
\ell_t &= \alpha(y_t/s_{t-m}) + (1-\alpha)(\ell_{t-1} + \phi b_{t-1}) \\
b_t &= \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)\phi b_{t-1} \\
s_t &= \gamma \frac{y_t}{(\ell_{t-1} + \phi b_{t-1})} + (1-\gamma)s_{t-m}.
\end{aligned}$$

The m in the above models is for the seasonal component. When looking at monthly data m equals 12. If using daily data, m is 7. Tests will be run to determine whether the trend should be none, multiplicative, additive, or a damped version. A similar process is used for the seasonal component. Should it be none, multiplicative, or additive? Lastly, the error term, will it be additive or multiplicative? Linear examples of both error term options are below. Whatever way the error term is modeled, it is assumed to be normally distributed with constant variance, $\varepsilon_t \sim \text{NID}(0, \sigma^2)$.

[7] Hyndman, Rob J and Athanasopoulos, George (2018)

$$
\begin{aligned}
Additive\ error: y_t &= \ell_{t-1} + b_{t-1} + \varepsilon_t \\
\ell_t &= \ell_{t-1} + b_{t-1} + \alpha\varepsilon_t \\
b_t &= b_{t-1} + \beta\varepsilon_t,
\end{aligned}
$$

$$
\begin{aligned}
Multiplicative\ error: y_t &= (\ell_{t-1} + b_{t-1})(1 + \varepsilon_t) \\
\ell_t &= (\ell_{t-1} + b_{t-1})(1 + \alpha\varepsilon_t) \\
b_t &= b_{t-1} + \beta(\ell_{t-1} + b_{t-1})\varepsilon_t
\end{aligned}
$$

4.3 ARIMA (Box-Jenkins)

Autoregressive integrated moving average models (ARIMA) aim to describe the autocorrelations in the data. GNP monthly data contains seasonality and perhaps a trend. They will be seasonally differenced and, if necessary, first differenced to be made stationary. A first differenced series gives the change between two consecutive observations, $y'_t = y_t - y_{t-1}$, and a seasonally differenced series uses the change between the current observation and the corresponding observation from m periods ago, $y'_t = y_t - y_{t-m}$.

In an ARIMA model, the 'AR' stands for autoregressive, which means a regression of the variable on past values of itself. The 'MA' is for the moving average part of the model. Moving averages are taken on the error terms, ε_t which are normally distributed white noise with mean zero and variance equal to σ^2. The two parts are combined to give the autoregressive integrated moving average model; an ARIMA(p, d, q) model where p is the order of the autoregressive part, d is the degree of first differencing, and q is the order of the moving average part. Equation 8 shows an ARMA model where the ϕ's are the AR component and θ's the MA.

$$y'_t = c + \phi_1 y'_{t-1} + \cdots + \phi_p y'_{t-p} + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (8)$$

[8]

A seasonal component grows the model to an ARIMA (p, d, q) $(P, D, Q)_m$ where m equals 12 for monthly data and 7 for daily data. The non-seasonal parts are shown with lower-case letters and the seasonal parts are shown with upper-case letters. The seasonal part of the model allows it to run autoregressive and moving average equations on observations m periods ago. For data with strong seasonality like the GNP data, observations m periods ago might be more useful than the previous observation.

5 Regression Modeling of GNP Visitation

5.1 Introduction

Located on the Continental Divide at the Canadian border, Glacier National Park has a relatively short tourist season where visitation is strongly seasonal. As seen in Figure 5 there are large visitation spikes during summer months. Seasonal ebbs and flows hold the most sway over GNP monthly visitation. However, a host of other factors beyond seasonality affect visitation. Estimating their effects with regression models helps build an understanding of what the drivers of visitation are and that can then help park management better allocate resources. Efficient allocation of resources allows park management to provide better care to park assets and thereby increase visitor satisfaction. Gateway-communities around a park such as Glacier have economies that rely heavily on tourism. The park itself is the central part of a nested system, a grasp on what the park can expect will also influence the decisions made by local government officials around Glacier. Gateway-

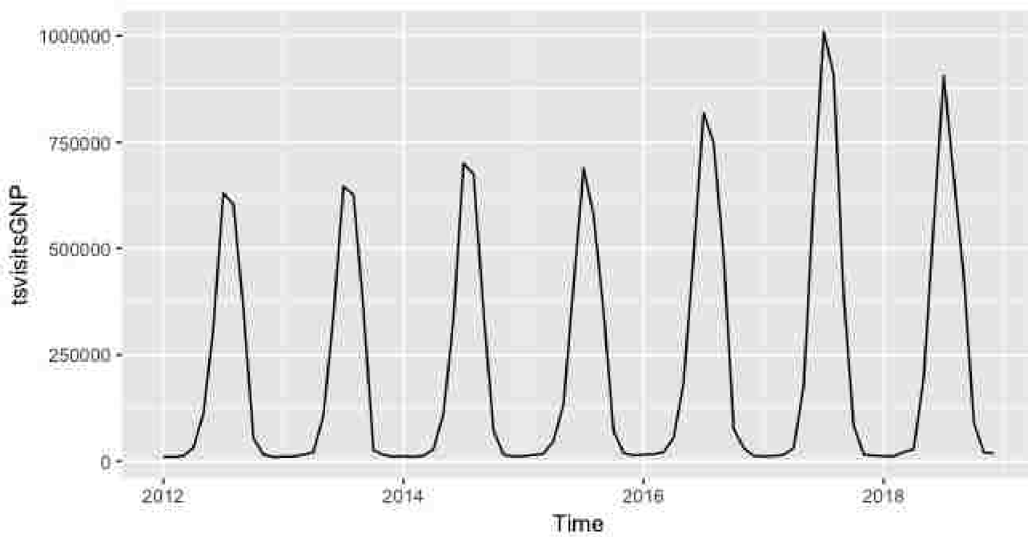
[8] Hyndman, Rob J and Athanasopoulos, George (2018)

community governments need to understand and prepare for tourists just like Glacier does.

Visitors expect to enjoy being in the park and the park is an almost entirely outdoor setting. They visit Glacier to go on hikes, camp, observe wildlife, take pictures, and see beautiful vistas. People's ability to engage in these activities is heavily contingent upon good weather, clear skies, and areas of the park being open. The factors I test which directly impact enjoyment and use of the park are air quality, precipitation, forest fire, and days Going-to-the-Sun-Road is open.

Getting to Glacier NP in the first place is inherent in visitation. Variables for the unemployment rate and gas prices are included to represent macroeconomic challenges faced by GNP visitors. People have limited resources and need some form of employment to earn money. An ability to earn money for things like trips to national parks is represented by the proportion of those in the labor force without a job. When gas prices increase, it becomes more expensive to visit GNP. If this happens, according to the Law of Demand, fewer people will visit the park. Gas prices in the model are lagged one month since trips are planned months ahead. Whether people want to go to Glacier is one thing, whether they can go to Glacier is another.

Figure 5: Monthly Visitation to Glacier National Park, 2012-2018



5.2 Model Specification

The seasonal pattern of the data in Figure 5 indicates the need to control for seasonality. This will be done with monthly dummy variables. Another standout feature in Figure 5 is the surge in visitation during summer 2017 which creates a non-constant variance. To deal with this, log of monthly visits is used in the model. Log transformations help reduce the spread of the variance. Figure 6 shows the spread of the variance has been reduced and more likely meets a constant variance assumption required by regression.

There was a government shutdown in October 2013. It caused a sharp decline in visitation relative to other Octobers in the data set. Figure 7 shows how the reported number of visits in October 2013 are not representative of what visitation in October normally is. This month was low due to the park being partially shut-down and minimally staffed during this time period.

The reported number of visitors in October 2013 is 25,965. Using a robust STL decomposition a replacement value of 63,275 visitors in October 2013 is estimated. This value is used in all following analysis since it represents a number closer to what visits would have been in the absence of a shutdown.

Figure 6: Logged Monthly Visitation to GNP, 2012-2018

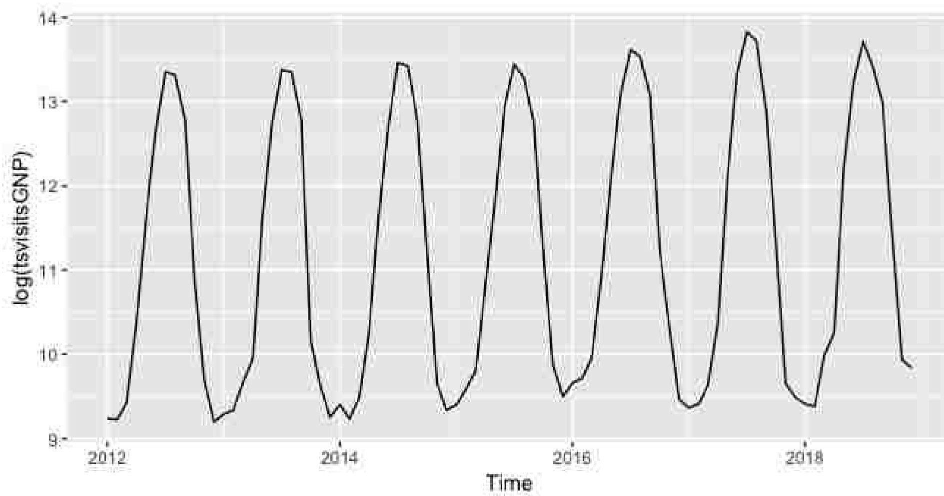
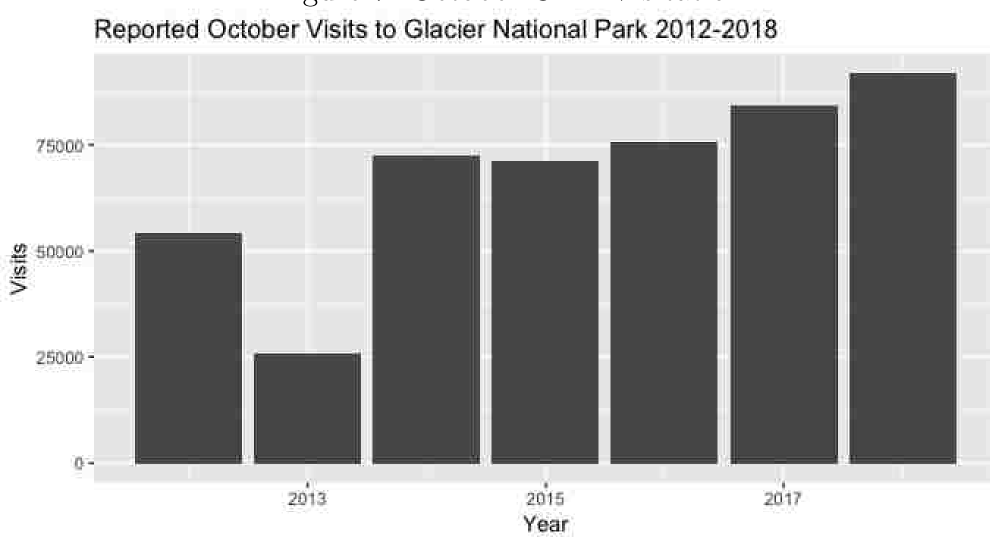


Figure 7: October GNP Visitation



Descriptions of variables in the data set are in Table 8. These data are organized as monthly time-series running from January 2012 to December 2018. Monthly visitation to GNP is the dependent variable in all models. July 2017 is the month which had the maximum value and December 2012 had the lowest recorded visitation over this period. Monthly visitation to Yellowstone National Park and a fire dummy for YNP were added since Duffield et al. (2013) found the two parks to be complements . There were not any months without reported precipitation on the eastern side of Glacier, the minimum precipitation days being two, but there was at least one month on the western side which did not get any precipitation.

There was, on average, half a day per month with air quality registering over 100 on a local index, when the cutoff is lowered to 75 the average number of days increases to 1.2. Even though the average days of bad air quality per month changes by a relatively large amount depending on index cutoff level, the maximum days over the cutoff is 15 when the cutoff is an AQI of 100 and 19 days when the cutoff is an AQI of 75. Both of these maximums are from August 2018, so lowering the cutoff might bring in many more days to be counted as 'bad air quality,' but if a month has bad enough fires to seriously affect visitation like in August of 2018, there will still be many days that have bad enough air quality to go over 100 on the AQI index. This can be seen in Figures 8 and 9. It is suspected that some of the winter months showing bad air quality readings are due to inversions happening in the Flathead valley, but the three big spikes from forest fires in 2015, 2017, and 2018 stand out clearly in both graphs. Having the AQI read 100 or more appears to get rid of some noise in the data, actually. Using an AQI level of 100 or more detects serious forest fires while cutting down on chatter.

Figure 8: AQI Days Greater Than 100

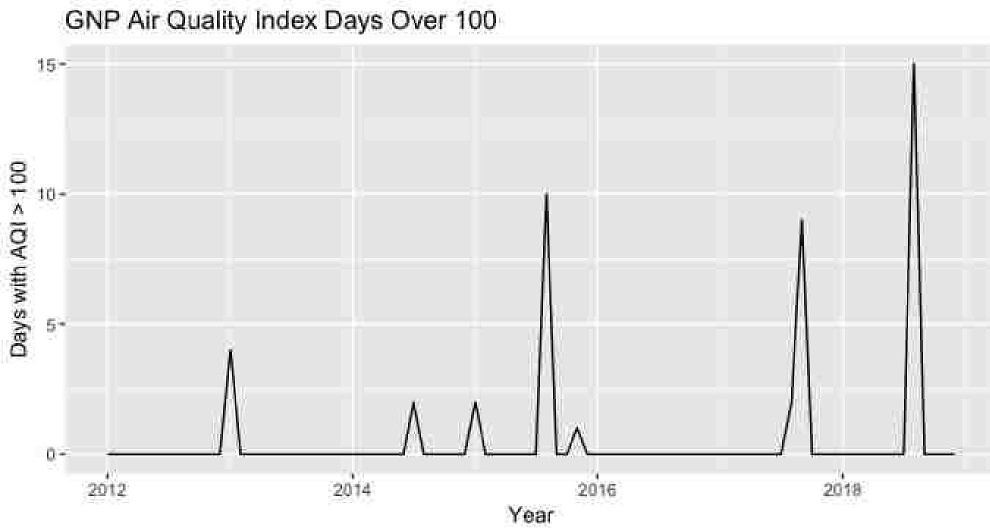
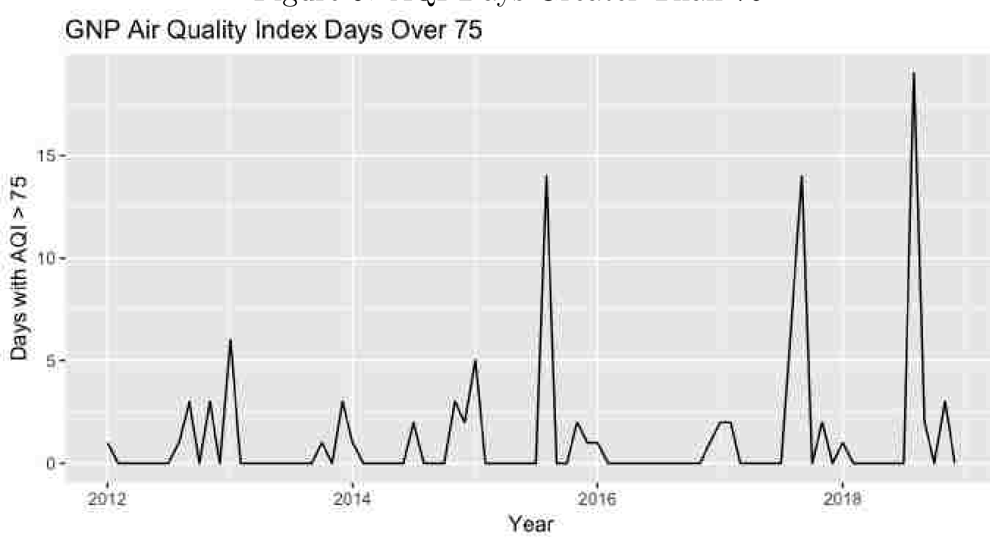


Figure 9: AQI Days Greater Than 75



Seasonally adjusted and non-seasonally adjusted unemployment rates have the same mean and standard deviation, but seasonal adjustment reduces the overall range of values taken. Lagging gas prices does not bring in any prices which alter the overall descriptions of gas price data. Descriptions of gas price data in the bottom three rows of Table 8 are equal regardless of lag. This is not surprising, though, since they share either 83 of 84 data points or 82 of 84 data points.

Table 8: Descriptive Statistics

	obs.	mean	sd	min	max
Year	84	2015.0	2.0	2012.0	2018.0
Month	84	6.5	3.5	1.0	12.0
Glacier	84	217813.8	277761.2	9862.0	1009655.0
Yellowstone	84	318281.2	349177.6	10468.0	995917.0
GNP.fire	84	0.1	0.3	0.0	1.0
YNP.fire	84	0.1	0.3	0.0	1.0
GTSR.open	84	7.2	11.6	0.0	31.0
WG.Precip.Days	84	13.4	6.5	0.0	29.0
EG.Precip.Days	84	12.0	5.1	2.0	24.0
aqi100	84	0.5	2.2	0.0	15.0
aqi75	84	1.2	3.1	0.0	19.0
UnEmpSA	84	5.7	1.5	3.7	8.3
UnEmpNSA	84	5.7	1.5	3.5	8.8
ConSent	84	118.1	11.4	96.4	135.2
Gas	84	2.9	0.6	1.8	3.9
gas.lag1	84	2.9	0.6	1.8	3.9
gas.lag2	84	2.9	0.6	1.8	3.9

Note: Oct. 2013 visitation replaced by estimated value of 63,275.

Table 9: OLS Regression: Full Model and Reduced Model

	log(GNP Monthly Visitation)	
	(1)	(2)
AQI100	−0.018*	−0.017**
	(0.010)	(0.008)
GNP Fire	−0.075	
	(0.069)	
YNP Fire	−0.008	
	(0.066)	
GTSR Open	−0.005	
	(0.006)	
WG Precip. Days	0.001	
	(0.005)	
EG Precip. Days	−0.016***	−0.015***
	(0.006)	(0.004)
UnEmpSA	−0.038*	−0.039**
	(0.019)	(0.018)
Gas lag1	−0.154***	−0.144***
	(0.048)	(0.045)
feb	0.002	0.001
	(0.075)	(0.073)
mar	0.313***	0.315***
	(0.075)	(0.073)
april	0.966***	0.968***
	(0.078)	(0.076)
may	2.414***	2.412***
	(0.079)	(0.078)
june	3.597***	3.553***
	(0.092)	(0.076)
july	4.181***	4.028***
	(0.193)	(0.086)
aug	4.181***	4.001***
	(0.207)	(0.083)
sep	3.518***	3.375***
	(0.159)	(0.082)
oct	1.793***	1.746***
	(0.088)	(0.077)
nov	0.400***	0.402***
	(0.077)	(0.075)
dec	0.023	0.026
	(0.076)	(0.074)
Constant	10.282***	10.260***
	(0.106)	(0.097)
Observations	84	84
R^2	0.994	0.994
Adjusted R^2	0.993	0.993
Residual Std. Error	0.138 (df = 64)	0.136 (df = 68)
F Statistic	603.767*** (df = 19; 64)	791.100*** (df = 15; 68)

Notes: ***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

5.3 Model Selection

Column (1) of Table 9 has several variables of interest, but serial correlation and multicollinearity are suspected which would invalidate standard errors. Breusch-Godfrey tests are run on the full OLS model and they do not find significant serial correlation in the errors ($\chi^2 = 9.4998, p = 0.6598$, fails to reject the null hypothesis of no serial correlation). Serial correlation would impact the standard errors of estimates, but it is not serious enough to merit techniques for correcting it here.

Another concern is that variables exhibit multicollinearity, meaning their movements are highly related. This is checked by looking at levels of correlation among variables. Figure 10 is a visualization of how correlations match up across this data set. Larger circles and dark shades in the Figure 10 correlation matrix indicate a correlation close to 1 or -1. The scale on the right side of the the correlation matrix can be referenced to get an approximate idea of where various correlations land. Correlations near 1 suggest strong positive correlation, movements are related and in the same direction. Correlations near -1 denote a strong negative correlation, movements are related and move in opposite directions (when one increases, the other decreases). There does appear to be multicollinearity. In particular, air quality shows strong correlation with two other variables, the fire in GNP dummy and West Glacier precipitation. Both make sense since there is a positive correlation between bad air quality and forest fire, while there is a negative correlation between precipitation and fire/ bad air quality. Fire's effect on park visitation is of particular interest to the park and communities near parks throughout the western United States, but fires vary so much in their size and effect on visitor perceptions of park enjoyment. Capturing this fire heterogeneity with a dummy variable signifying significant fire activity in a given month misses important nuances within fire activity itself. For these reasons using poor air quality as a proxy for forest fire is better for capturing the effect on visitation. The fire variable is therefore dropped and air quality kept in the model.

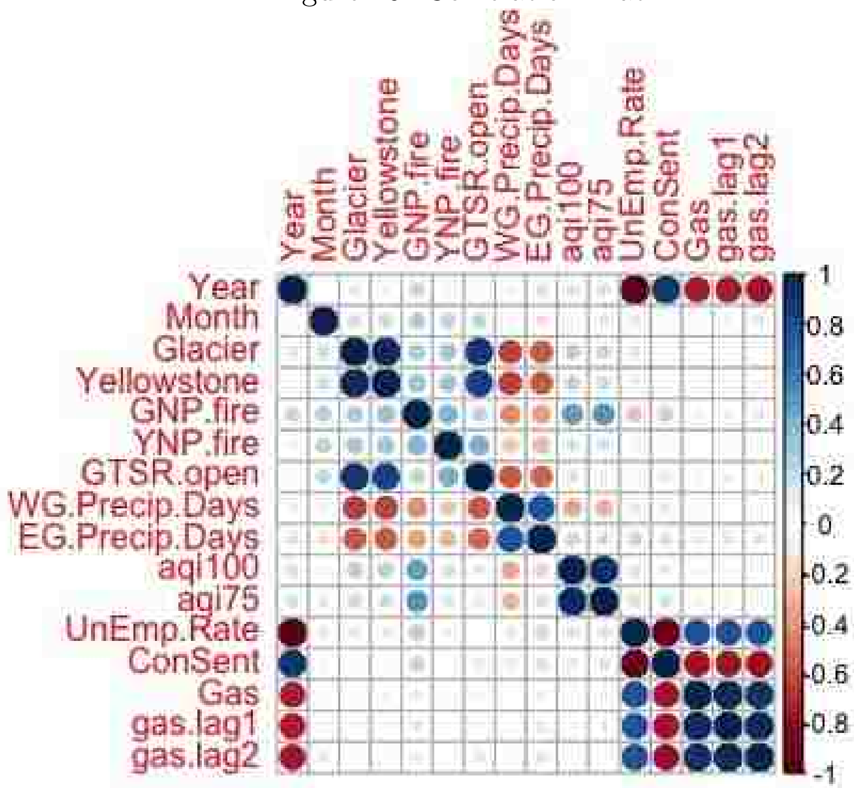
YNP monthly visits ended up being left out of all models due to a similar seasonal pattern causing collinearlty. The YNP fire dummy was dropped for similar reasons to the GNP fire dummy. It was not then proxied for by taking air quality index reading from a nearby weather station. This could be done with future park research. While forest fires were particularly detrimental to park operations in GNP over the 2012-2018 period, they were not as disruptive in YNP so a different or wider observation period might be needed.

Precipitation affects visitation in its own specific ways. It makes sense air quality and West Glacier precipitation exhibit strong correlation since air quality readings come from the western side of the park. Fortunately the model can use a less correlated measure of precipitation from East Glacier to still include the important effect precipitation has on visitation. The West Glacier monthly precipitation counts are not included in the main model while those from East Glacier are.

Another variable left out of the main regression is the number days in month Going-to-the-Sun Road is open. GTSR is an important feature in GNP which helps draw huge crowds in summer months. Its being open, though, is strongly tied to other parts of the park being open which relate to the strong seasonality caused by the park's northern location and high elevation destinations, Logan Pass on GTSR being one such destination. Although GTSR itself is an important feature within GNP, its affect on visitation is absorbed by the monthly dummies since the road almost always opens in June and closes in October. There were times in the period of observation when GTSR was closed due to forest fire when it would otherwise be open. That effect is captured by the air quality variable. Days in a month that GTSR is open is removed from the model.

The main model includes air-quality-index days over 100, East Glacier precipitation days, the unemployment rate, gasoline prices lagged one month, and monthly dummy variables. The reduced version of the regression accounts for the main forces which explain how many people visit Glacier National Park. Seasonality being the main driver is represented through monthly dummies. Forest fire is proxied through a measure of air quality. Precipitation is a direct count of days per month with precipitation from weather stations on the east side of the Park. Macroeconomic forces are included via the unemployment rate. Travel cost is incorporated by using the average national price of gasoline lagged one month.

Figure 10: Correlation Matrix



5.4 Main Regression Results

Table 9 Column (2) shows results of the main regression. All non-monthly variables are statistically significant. Air quality and the unemployment rate are significant at the 5% error level while lagged gasoline prices and East Glacier precipitation are significant at the 1% error level. All variables affect monthly visitation in the expected directions. R^2 is 0.994 and adjusted R^2 is 0.993 which indicates a good fit. An R^2 of 0.994 says the model explains 99.4% of variation in log of monthly visitation to GNP. Neither R^2 nor adjusted R^2 decreased when variables were dropped from the full model to this reduced model.

Quick coefficient interpretation can be done by multiplying the coefficient by 100. This gives a relatively close approximation of the effect.[9] Actual coefficient (β) interpretation for log dependent variable models in what I report are found by using the following formula: $(e^{\beta} - 1) * 100$ equals the estimated percent change in monthly visitation. The four following interpretations of main regression results are done under the assumption of ceteris paribus, holding all else in the model constant. An extra day per month where the air quality index on the western side of Glacier (Flathead Valley Monitoring Station) exceeds a value of 100 is associated with a 1.76% decrease in that month's visitation. The unemployment rate increasing by 1 percentage point reduces visitation, on average, by 3.82%. The price of gasoline increasing by $1 in the previous month will on average reduce visitation the following month by 13.41%. A one dollar increase in gas prices is a large increase, so perhaps a more digestible interpretation would be to say that if gas prices increase by 10 cents, the following month will likely see a 1.34% drop in visits. Each additional day of precipitation recorded on the eastern side of GNP

[9] The coefficient on the Unemployment Rate is -0.039, so $(e^{-}0.039 - 1) * 100$ equals a negative 3.82% change in visitation for a 1 percentage point increase in the unemployment rate, while quickly interpreting the coefficient would indicate a negative 3.9% change.

lowers that month's visits by about 1.49%. A model using logged monthly visits provides the benefit that all percent changes are easily applied to whatever month is in question. This is important since 1.49% is a small actual drop in visitation during winter months, but an extra rainy day represents a decrease of over 10,000 people in July.

Monthly dummy variables can be interpreted similarly with one important difference. $(e^{\beta} - 1) * 100$ equals the estimated percent change in monthly visitation *relative to January*. One month must be left out to avoid perfect collinearity. Since January is the month for which there is no dummy variable, the coefficients for monthly dummies in the model are measures of the difference between those months and January. January has the lowest average visitation so all monthly dummy coefficients are positive. Average January visitation from 2012 to 2018 is 12,133.29. Shoulder and peak months from May-September are of most interest so interpretation of their dummy coefficients follows. The quick interpretation technique does not work for any of these months because their coefficients exceed 1. How May visitation is relative to January is given by $(e^{2.412} - 1) * 100 = 1015.625$, so May visitation is 1015.625% higher than January on average. Further interpretation would be to say that if January visits average 12,133.29, then May averages 135,362.02 visitors over the 2012-2018 period; $(12,133.29 * 11.15625)$. June visitation is 3391.791% above January, so it being June corresponds with $(12,133.29 * 34.91791) = 423,669.13$ visitors from 2012-2018. About 5514.850% more people visit GNP in July relative to January, giving it an average of 681,266.03 visitors $(12,133.29 * 56.1485)$. August, on average, sees an uptick of 5365.278% more visitors which translates to around 663,118.03 people $(12,133.29 * 54.65278)$. September visitation is 2822.428% over January's, which comes out to 354,586.66 visitors $(12,133.29 * 29.22428)$.

5.5 Regression Robustness Testing

Some literature found macroeconomic indicators other than the unemployment rate to be better at describing fluctuations in national park visitation (Poudyal, Paudel, & Tarrant, 2013). Consumer sentiment performed better in their models, but when looking specifically at GNP it appears consumer sentiment has no effect and does not fit the bill for bringing macroeconomic forces into these models. Table 10 shows the comparison of results. For assessing monthly visits to Glacier National Park, the unemployment rate fits the data better.

Most visitors drive to Glacier. Travel cost must be included and this is done with a time-series of average national U.S. gasoline prices. It was reasoned that gas prices might not have an immediate effect on travel decisions so they are lagged one month when used in the regressions. The model was also run with no lag and a two month lag. Table 11 shows there is not much difference between the three options. The significance of the unemployment rate is diminished, though, when no lag and a two month lag are used. Gas prices and unemployment are correlated with each other more than with any other parts of the model as national gas prices are an indicator for the macro-economy too. Unemployment and gas prices have their lowest standard errors in the model where gas prices are lagged one month. There does not seem to be reason for changing model specification away from using a one month lag.

It was judged that air quality from a Flathead Valley monitoring station would work better as a proxy variable for forest fire in GNP than the forest fire in GNP dummy variable which was created based on news releases from the park. Table 12 shows how the restricted model looks with the GNP fire dummy included and then shows results when either one are included. When the fire dummy is added to the reduced model it has no statistical effect and when it is added without accounting for air quality it barely has a statistical effect.

The cutoff for counting a day as having poor air quality is for the Flathead monitoring station to have recorded enough particulate matter in the air to put the indexed level for that day at or over 100. Table 13 shows results when the index level cutoff is lowered to 75. There are not any major changes to the results when the air quality index cutoff is lowered from 100 to 75.

Data could have been de-seaonalized and the regressions run on that data without monthly dummies. This is done in Table 14. The results come out essentially the same. This indicates the monthly dummies are adequately accounting for seasonality and the results of the main model are not spurious.

Table 10: Robustness Testing: Unemployment Rate / Consumer Sentiment

	log(GNP Monthly Visitation)	
	(1)	(2)
UnEmpSA	-0.039^{**}	
	(0.018)	
ConSent		0.003
		(0.002)
EG.Precip.Days	-0.015^{***}	-0.015^{***}
	(0.004)	(0.004)
aqi100	-0.017^{**}	-0.015^{*}
	(0.008)	(0.008)
gas.lag1	-0.144^{***}	-0.173^{***}
	(0.045)	(0.041)
Monthly Dummies	Yes	Yes
Constant	10.260^{***}	9.716^{***}
	(0.097)	(0.354)
Observations	84	84
R^2	0.994	0.994
Adjusted R^2	0.993	0.993
Residual Std. Error (df = 68)	0.136	0.138
F Statistic (df = 15; 68)	791.100^{***}	765.717^{***}

Notes: ***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Table 11: Robustness Testing: Gas Price Lags

	log(GNP Monthly Visitation)		
	(1)	(2)	(3)
aqi100	-0.017^{**}	-0.017^{**}	-0.017^{**}
	(0.008)	(0.008)	(0.008)
EG.Precip.Days	-0.015^{***}	-0.015^{***}	-0.014^{***}
	(0.004)	(0.004)	(0.004)
UnEmpSA	-0.042^{**}	-0.039^{**}	-0.033^{*}
	(0.018)	(0.018)	(0.018)
Gas	-0.135^{***}		
	(0.045)		
gas.lag1		-0.144^{***}	
		(0.045)	
gas.lag2			-0.162^{***}
			(0.044)
Monthly Dummies	Yes	Yes	Yes
Constant	10.255^{***}	10.260^{***}	10.279^{***}
	(0.098)	(0.097)	(0.096)
Observations	84	84	84
R^2	0.994	0.994	0.995
Adjusted R^2	0.993	0.993	0.993
Residual Std. Error (df = 68)	0.137	0.136	0.133
F Statistic (df = 15; 68)	776.651^{***}	791.100^{***}	821.971^{***}

Notes:

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Table 12: Robustness Testing: Air Quality Index > 100 / GNP Fire Dummy Variable

	log(GNP Monthly Visitation)		
	(1)	(2)	(3)
aqi100	-0.013	-0.017^{**}	
	(0.009)	(0.008)	
GNP.fire	-0.062		-0.102^{*}
	(0.064)		(0.059)
EG.Precip.Days	-0.015^{***}	-0.015^{***}	-0.015^{***}
	(0.004)	(0.004)	(0.004)
UnEmpSA	-0.042^{**}	-0.039^{**}	-0.039^{**}
	(0.018)	(0.018)	(0.018)
gas.lag1	-0.142^{***}	-0.144^{***}	-0.148^{***}
	(0.045)	(0.045)	(0.045)
Monthly Dummies	Yes	Yes	Yes
Constant	10.275^{***}	10.260^{***}	10.261^{***}
	(0.099)	(0.097)	(0.099)
Observations	84	84	84
R^2	0.994	0.994	0.994
Adjusted R^2	0.993	0.993	0.993
Residual Std. Error	0.136 (df = 67)	0.136 (df = 68)	0.137 (df = 68)
F Statistic	741.156^{***} (df = 16; 67)	791.100^{***} (df = 15; 68)	774.567^{***} (df = 15; 68)

Notes: ***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Table 13: Robustness Testing: Air Quality Index Cutoff

	log(GNP Monthly Visitation)	
	(1)	(2)
aqi75		-0.014^{**}
		(0.006)
UnEmpSA	-0.039^{**}	-0.040^{**}
	(0.018)	(0.018)
EG.Precip.Days	-0.015^{***}	-0.015^{***}
	(0.004)	(0.004)
aqi100	-0.017^{**}	
	(0.008)	
gas.lag1	-0.144^{***}	-0.142^{***}
	(0.045)	(0.044)
Monthly Dummies	*Yes*	*Yes*
Constant	10.260^{***}	10.282^{***}
	(0.097)	(0.098)
Observations	84	84
R^2	0.994	0.994
Adjusted R^2	0.993	0.993
Residual Std. Error (df = 68)	0.136	0.135
F Statistic (df = 15; 68)	791.100^{***}	802.380^{***}

Notes: ***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Table 14: Robustness Test Using De-seasonalized Data

	Seasonally Adjusted Log of GNP Monthly Visits	
	(1)	(2)
aqi100	−0.012	−0.015**
	(0.008)	(0.007)
UnEmpSA	−0.042**	−0.039**
	(0.017)	(0.016)
gas.lag1	−0.144***	−0.153***
	(0.039)	(0.038)
EG.Precip.Days	−0.014***	−0.010***
	(0.005)	(0.003)
WG.Precip.Days	0.001	
	(0.004)	
GNP.fire	−0.059	
	(0.056)	
YNP.fire	0.021	
	(0.053)	
GTSR.open	−0.001	
	(0.002)	
Constant	12.002***	11.972***
	(0.086)	(0.076)
Observations	84	84
R^2	0.619	0.609
Adjusted R^2	0.579	0.589
Residual Std. Error	0.131 (df = 75)	0.129 (df = 79)
F Statistic	15.252*** (df = 8; 75)	30.704*** (df = 4; 79)

Notes: ***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

5.6 Conclusions: Problems and Possibilities with Regression Results

Regression results give an insightful look into what drives visitation to GNP and to what magnitude. This can be great for park management understanding how to react under certain scenarios. If a particular month is experiencing more rain than usual or there is a sudden spike in gas prices, the park can use knowledge about how visitation previously moved in response to these forces and gauge how the rest of that or the following month will shape up. Regressions can be used with forecasting techniques, but the explanatory variables must be forecast as well since their future values are unknown. Again, this can be great for seeing how particular scenarios would likely play out. Actual forecasting techniques though, as will be used in the following section, will work better for projecting what visitation will be over the coming year. One or two months of the year may face unusual circumstances which the regression results can help park management adapt to, but over the course of a year this generally smooths out and forecasting using past counts of visitation will give the best big picture view of where visitation is headed.

6 Forecasting GNP Visitation

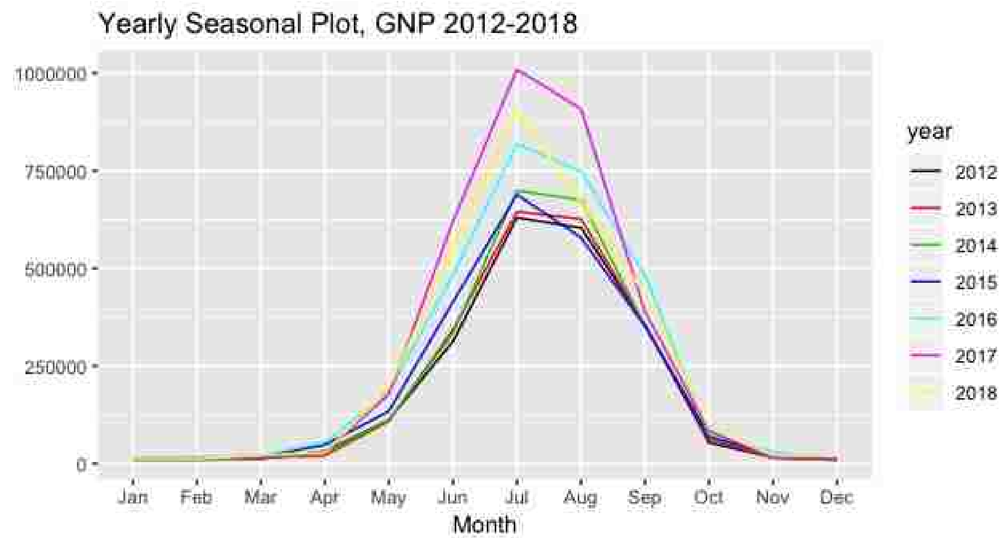
6.1 Introduction

Two of the most widely used forecasting techniques are exponential smoothing (ETS) and autoregressive-integrated-moving-average (ARIMA) models. Exponential smoothing models utilize past descriptions of the trend and seasonality while ARIMA models forecast by describing autocorrelations present in time-series data.

6.2 Exponential Smoothing Models

Glacier National Park monthly visitation data is not stationary since it exhibits strong seasonality. There is an uptick in visits beginning in 2016 which continues in 2017, but it is not a long enough period or sustained in 2018 so that it can be called a trend. Seasonality in GNP visitation is further shown by Figure 11, which puts the years 2012-2018 on top of each other. ETS models handle seasonality, and trend when there is one, through additive and multiplicative decomposition. An additive method works better when the seasonal fluctuations are consistent over the period of observation. A multiplicative method is more favorable when the seasonal variations change relative to the level of visits, as seems to be the case in Figure 11.

Figure 11: Seasonal Plot of GNP Visitation 2012-2018



A log model performed better in the regression model and the same is expected with ETS models, but both untransformed and logged models are run so their performances can be compared. The October 2013 outlier, caused by the government shutdown, is replaced with the same value as was used in the regression model. Decompositions of both styles are done on visits and log visits to see if there is a noticeable reason to choose an additive or multiplicative decomposition.

Looking at Figures 12 and 13 there is less of a pattern in the remainder with the multiplicative decomposition. Seeing the remainder part of the decomposition appear as white noise is a good indication that multiplicative ETS models will work better on the untransformed visitation data.

Remainder patterns appear similar with the logged data in Figures 14 and 15, but the remainders hover near zero when additive decompostion is used. This suggests additive ETS models will forecast more accurately when the visitation data is logged. Both of these outcomes make sense since changes in seasonality will be proportional with the number of visits when the raw data is used and the changes in seasonality will be fairly constant when the variance is smoothed by logging.

Figure 12: Additive Decomposition of Monthly Visits to GNP

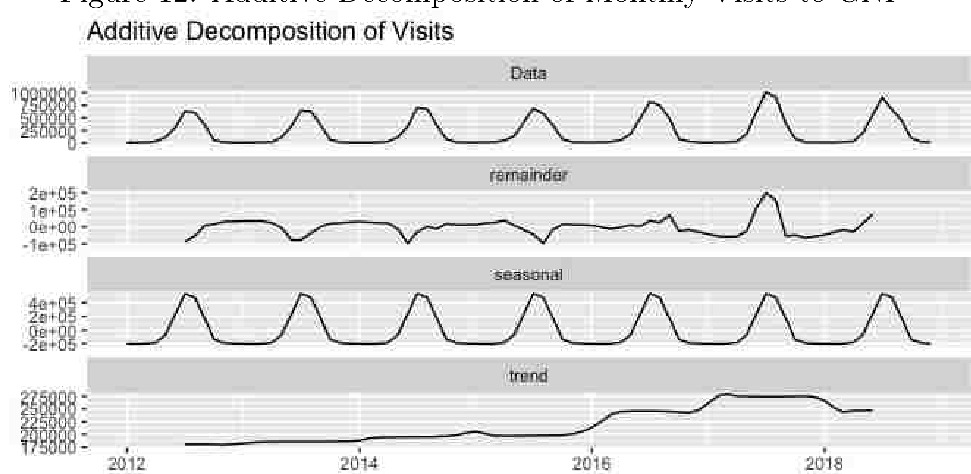


Figure 13: Multiplicative Decomposition of Monthly Visits to GNP

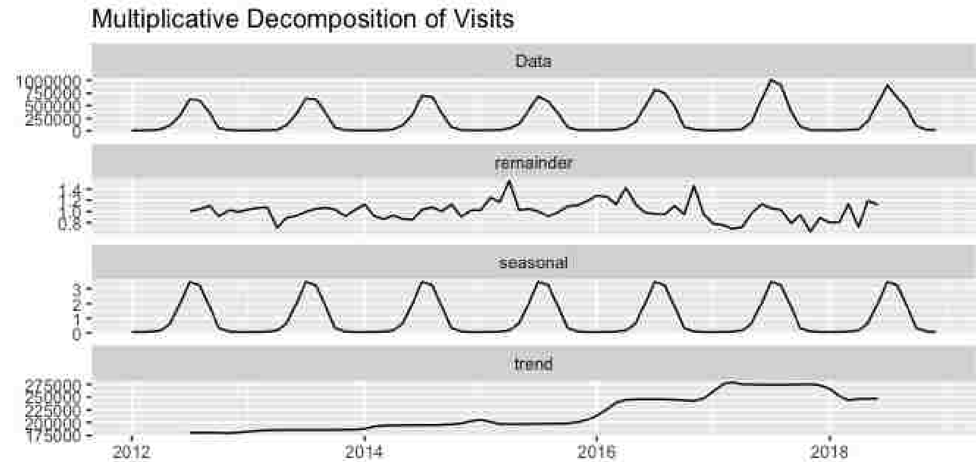


Figure 14: Additive Decomposition of Log Monthly Visits to GNP

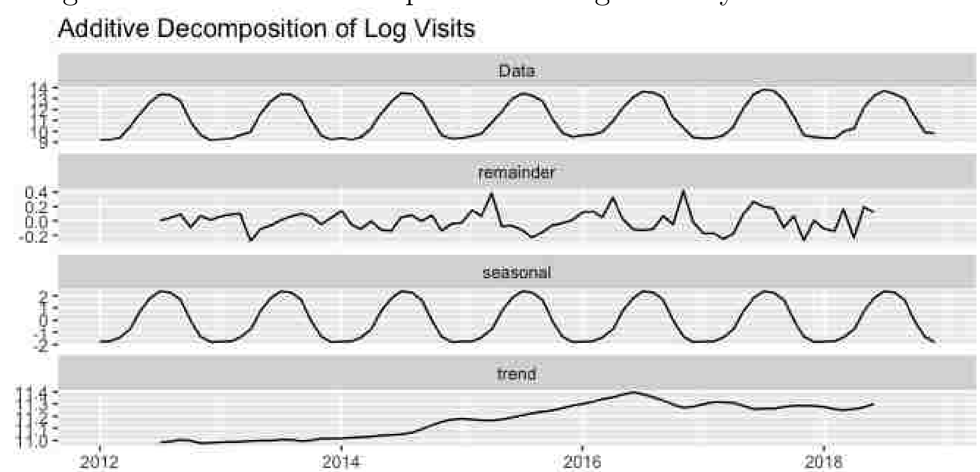
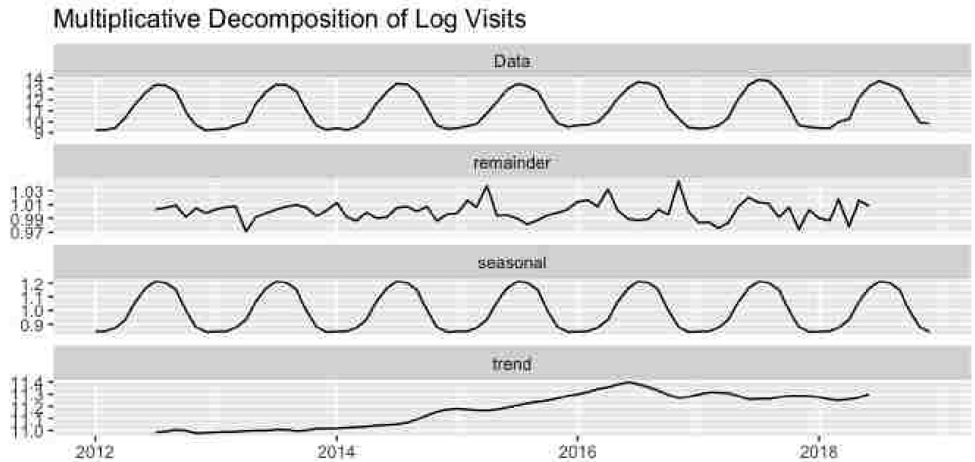


Figure 15: Multiplicative Decomposition of Log Monthly Visits to GNP



After deciding on multiplicative errors and seasonality for untransformed data (ETS(M,N,M)) and additive errors and seasonality for logged data (ETS(A,N,A)), statistical software is used to find model parameters which give the lowest corrected Akaike's Information Criterion (AICc). The smoothing parameter for how much weight is given to older observations is denoted by α. A higher α gives more weight to recent observations, causing older ones to decay at an increased rate as they move farther back from time T. α can be between 0 and 1. The α which minimizes the AICc in the untrasformed model is 0.479 and 0.398 for the log model. Since α is fairly low in both models, observations decay slowly meaning past months retain some impact for a longer period of time. The seasonal smoothing parameter, represented by γ, is .0001 for both models. γ is so small because the seasonal component changes very little over time. There is not a strong enough trend for a trend term to be included. It would be represented by β if it were needed. Table 15 gives the point forecasts of each model alongside actual visitation numbers. Test-sets are run from January 2012 to December 2017, leaving 2018 to be forecast and then used as a check of forecast accuracy.

Table 15: ETS Test Set Point Forecasts and Actual Visitation Numbers

	Untransformed	Log	Actual
Jan. 18	13,121.70	13,866.56	12,222
Feb. 18	13,354.88	14,135.32	11,847
March 18	16,852.16	18,284.50	21,758
April 18	38,837.80	39,058.62	28,404
May 18	146,722.38	157,789.52	195,116
June 18	448,906.08	462,824.83	556,304
July 18	795,967.65	838,971.77	905,959
Aug. 18	736,333.70	766,599.36	667,688
Sep. 18	428,200.53	429,784.51	434,600
Oct. 18	75,289.16	78,290.39	91,973
Nov. 18	20,473.30	20,323.15	20,657
Dec. 18	12,424.60	12,951.39	18,781
Total	2,746,483.94	2,852,879.92	2,965,309

The ETS(A,N,A) model using logged data generates point forecasts closer to actual visitation. Comparison of the Test Sets' Mean Error (ME), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Percentage Error (MPE), Mean Absolute Percentage Error (MAPE), Mean Absolute Scaled Error (MASE), and First-Order Autocorrelation Coefficient (ACF1) further show logged data allow for more accurate forecasts. In all of the accuracy measurements, shown in Table 16, values closer to zero are better and indicate which model will provide more reliable forecasts. In all measures the ETS(A,N,A) model run on logged visitation data wins.

Table 16: Goodness of Fit/ Accuracy Measurements

ETS(M,N,M)
Untransformed Data

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training Set	-1279.2	42,020.9	22,332.5	-1.237	11.994	0.779	0.468
Test Set	18,235.4	50,975.8	31,816.6	5.503	16.687	1.110	0.160

ETS(A,N,A)
Log Data

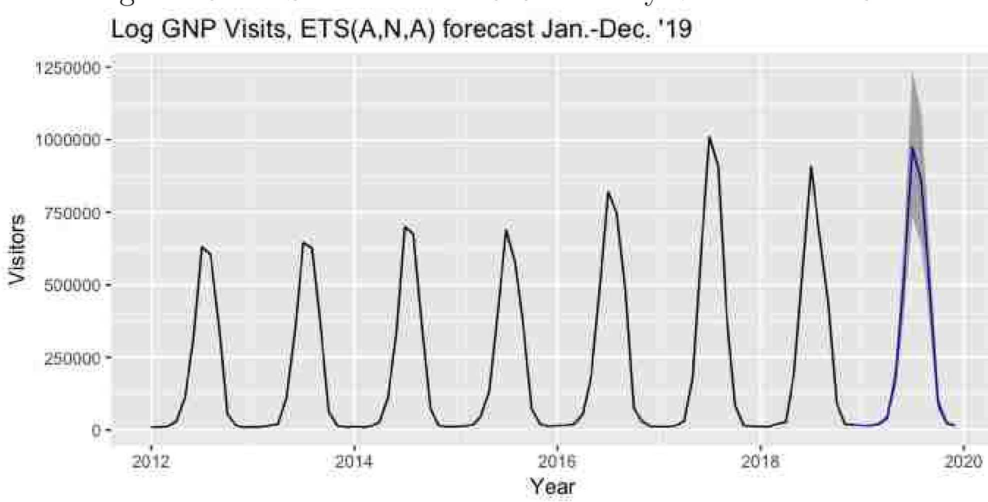
	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training Set	641.5	43,534.5	23,321.0	-1.321	11.669	0.814	0.512
Test Set	9369.1	45,442.6	28,285.6	1.903	16.086	0.987	0.050

Having found the best form for the data to be in, logged visitation data is focused on and numerous specifications are compared when the full data set is used. The model and parameters which result in the lowest AICc is an ETS(A,N,A) model with smoothing parameters $\alpha = 0.2836$ and $\gamma = .0001$. Including 2018, when the possible trend that began to show in 2016/2017 is broken, lowers α and the seasonal component continues to change very little so γ remains very low. Point forecasts for 2019 with 80% and 95% confidence-level prediction intervals are given in Table 17. The narrow prediction intervals over most months suggest visitation is usually easy to forecast due to the strong seasonality. The prediction intervals grow in the summer months when there can be larger swings in visitation, but they are still relatively tight. Figure 16 graphically shows how these point forecasts and prediction intervals look relative to the years used to generate them. In Figure 16, the blue area and grey area surrounding the black line representing 2019 point forecasts are visualizations of the 80% and 95% confidence intervals, respectively.

Table 17: ETS(A,N,A) Point Forecasts with Prediction Intervals

	Point Forecast	Low 80	High 80	Low 95	High 95
Jan. 19	15,579.55	12,387.86	19,052.38	11,053.827	21,351.71
Feb. 19	15,712.94	12,374.47	19,357.65	10,992.373	21,791.53
March 19	21,520.19	16,791.07	26,699.81	14,851.255	30187.24
April 19	42,383.78	32,773.04	52,942.92	28,865.921	60,108.97
May 19	182,647.74	139,999.40	229,646.49	122,810.834	261,787.74
June 19	541,680.07	411,670.47	685,372.39	359,711.895	784,370.99
July 19	971,889.35	732,504.61	1,237,223.61	637,614.622	1,421,347.57
Aug. 19	855,522.99	639,582.05	1,095,537.88	554,665.776	1,263,258.69
Sep. 19	491,346.93	364,419.69	632,807.19	314,894.541	732,332.16
Oct. 19	91,769.45	67,535.65	118,849.53	58,151.722	138,028.24
Nov. 19	23,450.20	17,126.57	30,534.85	14,696.076	35,584.81
Dec. 19	15,773.85	11,434.31	20,647.94	9,778.552	24,144.17
Total	3,269,277	2,831,118	3,699,210	2,488,610	4,201,562

Figure 16: ETS Forecast of 2019 Monthly Visitation to GNP



6.3 ARIMA Models

Autoregressive Intrergrated Moving Average (ARIMA) models also require data be stationary, meaning it has a constant mean and variance. Visitation to GNP has a strong seasonal pattern and is not stationary. This is visible in previous graphs of the visitation time-series and in Figure 17. Figure 17 separates all months into their own subseries, showing that regardless of cross-year monthly variation, the seasonal effect dominates. In Figures 18 and 19, the cosine wave pattern of the autocorrelations further iterates the need to remove seasonality before forecasting can be performed. ARIMA models address non-stationarity through differencing. Differencing comes in two forms. First differencing uses the difference between consecutive observations to reduce a trend and make the series stationary. Seasonal differencing uses the difference between observations m seasonal periods apart. When looking at monthly visits to GNP, $m = 12$. For GNP monthly visitation seasonal differencing is applied.

Figure 17: Monthly Subseries of Visitation to GNP

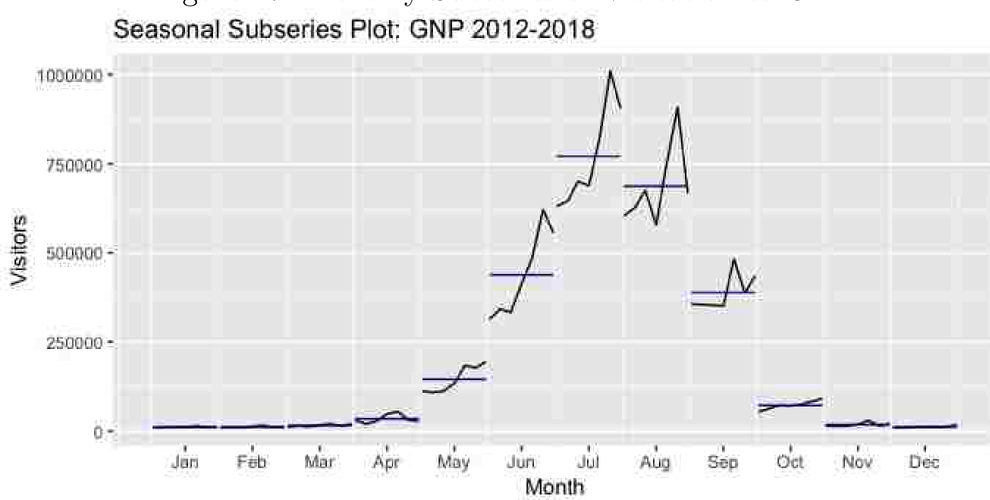


Figure 18: Autocorrelation Function of GNP Visits

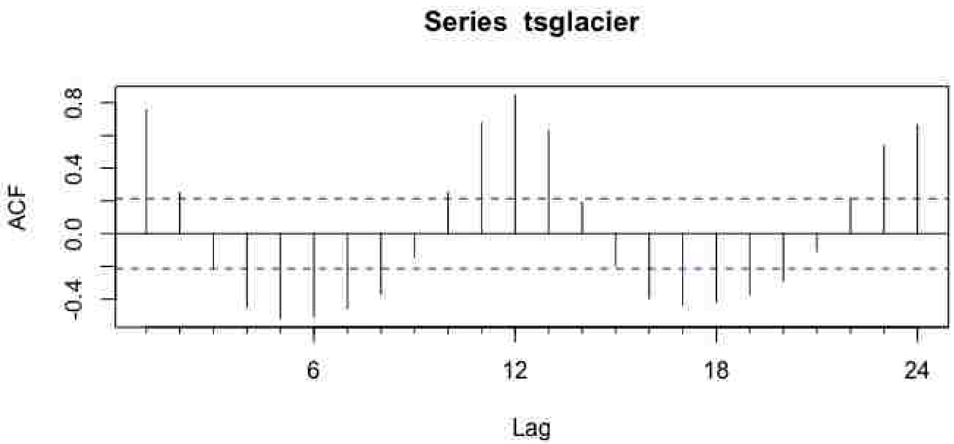
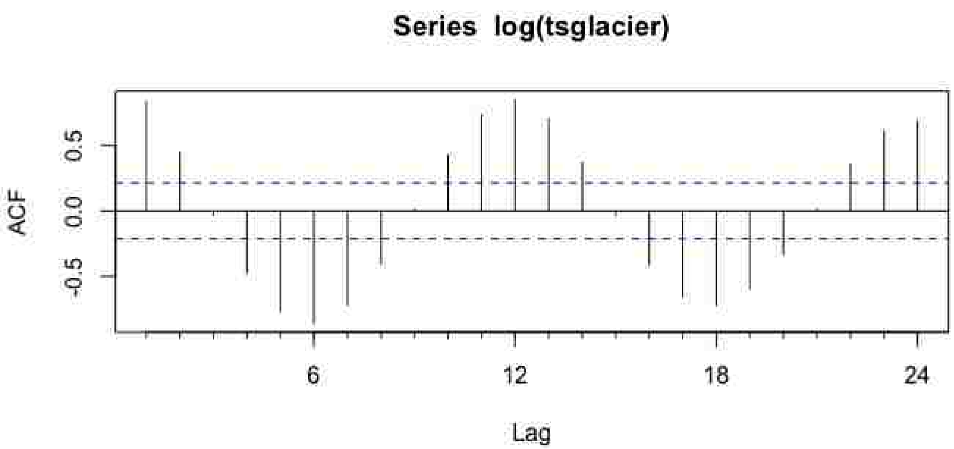


Figure 19: Autocorrelation Function of Log GNP Visits



The ACF graphs are complicated by the interplay of the seasonal terms so model selection based on their characteristics is difficult. Statistical software compares different model specifications and then selects the best model. The "best" model for each is determined by finding which ones have the lowest AICc. The best model for the untransformed data is found to be an

ARIMA(2,0,0)(0,1,0)[12] with drift model and the best model for logged data is an ARIMA(1,0,0)(2,1,0)[12] with drift. Their coefficient weights and standard errors are given in Table 18. These models use data from 2012-2017. Their forecasts for 2018 are compared with actual visitation records to check which is more accurate.

Table 18: Test Model Coefficients and Standard Errors

ARIMA(2,0,0)(0,1,0)[12] with drift

Untransformed Data

	AR1	**AR2**	**Drift**
Coefficients	0.6051	-0.2209	1560.699
Standard Errors	(0.1247)	(0.1241)	(765.916)

ARIMA(1,0,0)(2,1,0)[12] with drift

Log Data

	AR1	**SAR1**	**SAR2**	**Drift**
Coefficients	0.3322	-0.5821	-0.5404	0.0066
Standard Errors	(0.1223)	(0.1605)	(0.1226)	(0.0015)

Monthly point forecasts and yearly totals in Table 19 show the majority of the time using logged data allows for more accurate forecasts. The ARIMA(1,0,0) (2,1,0)[12] with drift using log data forecasts closer to what actual visits were in every month of 2018 except May. May through September are the months when the park really needs to know how many people are likely to visit, so the log model wins in four out of five main season months. Even in May, though, the log model is not far off, it is just that the point forecast for May from the untransformed model is particularly spot on. Visitation in August 2018 was much lower due to forest fires than would have been expected. Had it continued as previous Augusts, visitation would have been much higher and the total number of visitors for 2018 been even closer to what the log model predicted. It is indeed quite impressive how close many point forecasts from the log model are to actual visitation.

The point forecasts suggest using log of monthly visits in the models, but additional comparisons of model accuracy are done by looking at each model's Mean Error (ME), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Percentage Error (MPE), Mean Absolute Percentage Error (MAPE), Mean Absolute Scaled Error (MASE), and First-Order Autocorrelation Coefficient (ACF1). For the AICc which is used to select the best model, the smaller the value the better the model, but the selection criteria in Table 20 are measures of error, so the closer to zero the better. The model using logged data performs better by every measure except ACF1. The following ARIMA forecasts of monthly visitation to GNP have variance reduced by a log transformation and include seasonal differencing.

Table 19: Untransformed & Logged Data Point Forecasts alongside Actual 2018 Visits

	Untransformed	**Log**	**Actual**
Jan. 18	26,820.57	14,077.54	12,222
Feb. 18	32,771.06	16,107.56	11,847
March 18	35,930.37	20,288.12	21,758
April 18	51,075.32	48,630.96	28,404
May 18	196,544.55	183,425.35	195,116
June 18	639,541.67	591,450.43	556,304
July 18	1,028,286.96	974,616.27	905,959
Aug. 18	927,181.90	845,708.43	667,688
Sep. 18	407,871.27	444,824.17	434,600
Oct. 18	103,206.58	91,851.04	91,973
Nov. 18	34,326.65	21,674.90	20,657
Dec. 18	31,996.94	15,920.98	18,781
Total	3,515,553.84	3,268,575.75	2,965,309

Table 20: Goodness of Fit/ Accuracy Measurements

ARIMA(2,0,0)(0,1,0)[12] with drift

Untransformed Data

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training Set	85.64	40,184.7	24,618.7	-32.96	42.87	0.859	-0.012
Test Set	-45,853.7	87,464.1	50,308.5	-54.31	55.33	1.756	0.151

ARIMA(1,0,0)(2,1,0)[12] with drift

Log Data

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training Set	-3485.1	33,708.4	15,142.0	-2.675	10.31	0.528	0.319
Test Set	-25,272.2	56,512.3	27,962.7	-11.84	16.53	0.976	0.243

Returning to model diagnostics, there is nothing to gain by looking at the ACF in Figure 20 which shows a sinusoidal seasonal pattern, but the PACF in Figure 21 does offer some clues. The PACF shows significant spikes through lag 3 indicating three non-seasonal MA terms could be included in the model. Indeed, when the software is told to test a broader set of models on the full data set and select the one with the lowest AICc, it is a model with three non-seasonal MA terms and no AR terms. An ARIMA(0,0,3)(0,1,1)[12] with drift model. The AICc is one of the best measures to choose a model by, but the error measurements shown in Table 20 can be used alongside or instead. While the ARIMA(0,0,3)(0,1,1)[12] model has a low AICc (-36.08), an ARIMA(1,0,0)(2,1,1)[12] with drift model similar to that previously used when deciding between using untransformed or logged data has an AICc that is almost as low (-35.53) and preforms better in every other measure of error seen in Table 21. Its one difference from the test model is that it now includes a seasonal MA term. This is helpful because of the large error introduced by including August 2018 in the training set, recall August 2018 had a large drop in visitation relative to previous Augusts due to forest fire.

Even though it does not have the lowest AICc of any model, I decide to use the ARIMA(1,0,0)(2,1,1)[12] with drift model because I think it will do the best job of reacting to and incorporating a previous year's as well as a previous month's visitation. It seems logical that if one month has above average visitation, the likelihood of the next month having above average visitation is increased. A similar argument justifies the importance of seasonal AR terms, if the last two Septembers had higher than usual visitation, then there is an increased chance the next September will have a greater amount of visitors. The fact that the ARIMA(1,0,0)(2,1,1)[12] with drift model also outperforms the model with a lower AICc in all other measures of model fit, Tabel 21, further indicates it has better potential to generate more accurate forecasts. When using the full data set for forecasting 2019 monthly visitation to Glacier National Park, an ARIMA(1,0,0)(2,1,1)[12] with drift model will be used. A check of the residuals in Figure 22 and 23 confirms this is a suitable model to forecast with since the residuals appear to be white noise and do not have an improbable number of significant spikes in their ACF or PACF.

Figure 20: Autocorrelation Function of Log GNP Visits

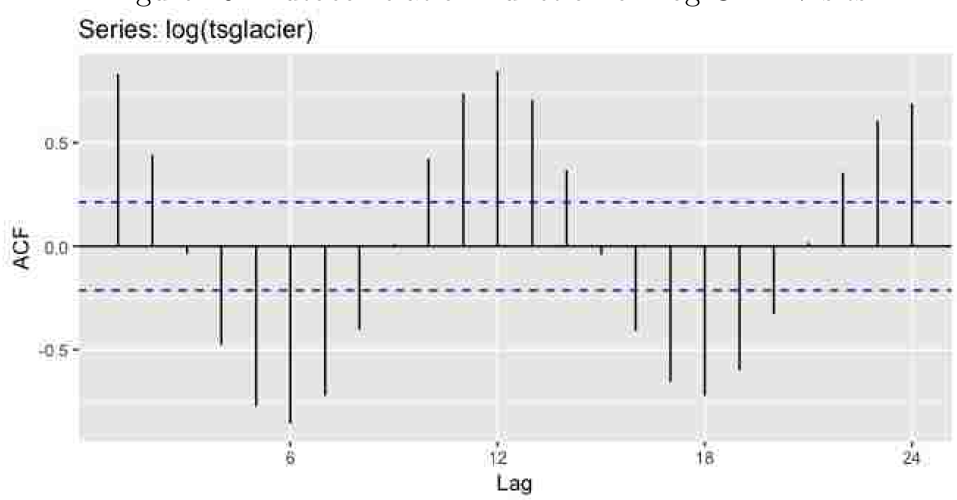


Figure 21: Partial Autocorrelation Function of Log GNP Visits

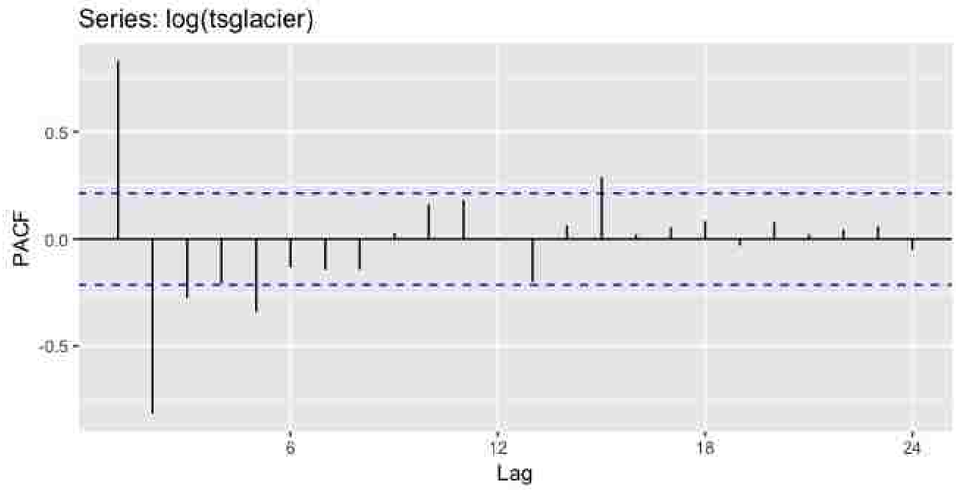


Table 21: Goodness of Fit Measurements: 2012-2018 Data for 2019 Forecast

ARIMA(0,0,3)(0,1,1)[12] with drift

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training Set	-7676.7	53051.7	24155.6	-2.085	11.347	0.784	0.166

ARIMA(1,0,0)(2,1,1)[12] with drift

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training Set	-4171.9	36877.0	15707.2	-1.756	9.607	0.510	0.282

Figure 22: Preferred Model Residuals

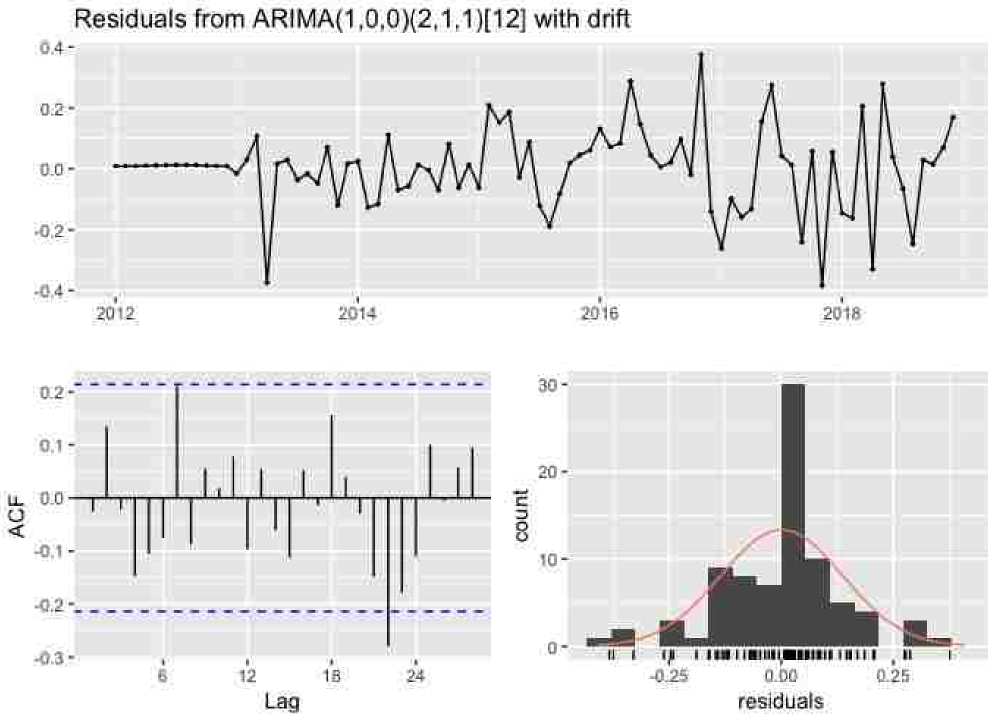
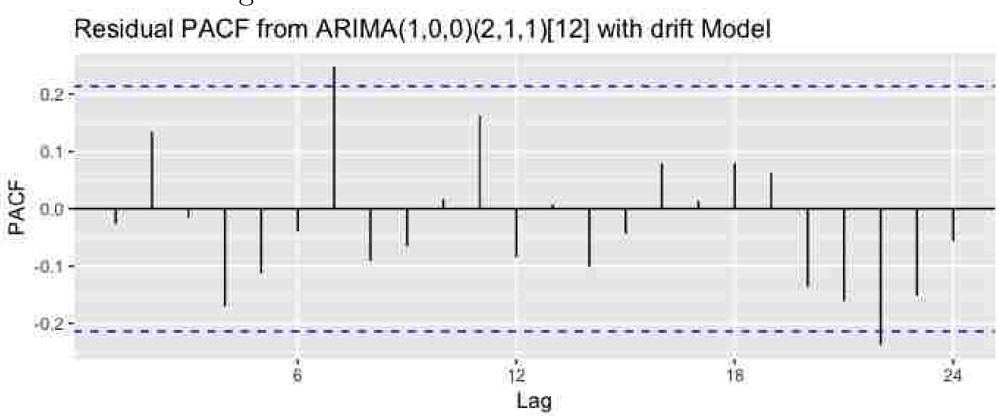


Figure 23: Preferred Model Residual PACF



The selected ARIMA(1,0,0)(2,1,1)[12] with drift model is run on logged monthly visitation counts from Janurary 2012 - Decemeber 2018. The model coefficients with corresponding standard errors are given in Table 22 and point forecasts with 80% and 95% confidence bounds are listed in Table 23. The given point forecasts are all realistic and should provide a good starting point for understanding what level visitation will reach each month. In Figure 24 the black line represents the point forecasts, the blue area around it is the 80% confidence bound, and the larger grey area around both is the 95% confidence bound. The 80% confidence bound stays quite close to the point forecasts, showing there is a high likelihood that the forecasts will be near actual visitation on average. As we are now well into 2019, visitation for Janurary - July 2019 is available, so a comparison of these months is shown in Table 24. January, February, and April 2019 actual visitation numbers do not fall within the 80% bounds, but these are not crucial months. Forecasted visitation for the shoulder month of May, June, and July 2019 are well within the 80% confidence bounds. May 2019 is over-forecast by about 21,000 visitors, June 2019 is under-forecast by about 23,000 visitors, and July 2019 is under by about 79,000 visitors.

Table 22: Select ARIMA Model Coefficients with Standard Errors for Forecasting 2019 GNP Visitation

ARIMA(1,0,0)(2,1,1)[12] with drift	**AR1**	**SAR1**	**SAR2**	**SMA1**	**Drift**
Coefficients	0.2484	-0.0077	-0.4455	-0.8183	0.0061
Standard Errors	(0.1159)	(0.1602)	(0.1429)	(0.4290)	(0.0008)

Table 23: ARIMA(1,0,0)(2,1,1)[12] with drift

	Point Forecast	**Low 80**	**High 80**	**Low 95**	**High 95**
Jan. 19	18,537.36	15,069.50	22,386.17	13,570.78	24,858.45
Feb. 19	18,095.37	14,623.73	21,981.64	13,128.30	24,485.54
March 19	24,994.97	20,192.39	30,373.92	18,124.07	33,840.20
April 19	51,540.81	41,636.74	62,633.85	37,371.43	69,782.44
May 19	188,523.81	152,296.90	229,099.75	136,695.34	255,247.78
June 19	520,882.22	420,788.98	632,991.66	377,682.61	705,237.44
July 19	958,280.71	774,136.55	1,164,531.40	694,832.61	1,297,443.87
Aug. 19	845,126.60	682,726.24	1,027,023.13	612,786.54	1,144,241.25
Sep. 19	551,135.90	445,229.09	669,756.82	399,619.03	746,198.75
Oct. 19	97,678.53	78,908.55	118,701.83	70,825.03	132,249.71
Nov. 19	29,576.52	23,893.18	35,942.11	21,445.58	40,044.22
Dec. 19	17,855.07	14,425.13	21,696.34	12,947.92	24,171.64
Total	3,323,337.87				

Figure 24: Glacier National Park 2019 Monthly Visits Forecast

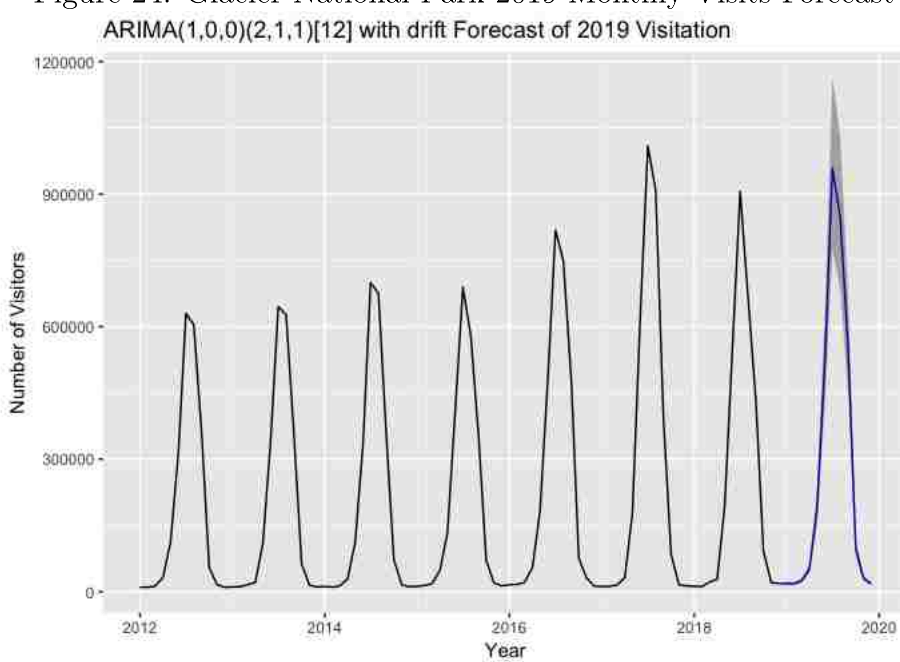


Table 24: January - July 2019 Comparison

	ARIMA Forecast	ETS Forecast	Actual
Jan. 2019	18,537	15,580	13,581
Feb. 2019	18,095	15,713	11,240
March 2019	24,995	21,520	23,989
April 2019	51,541	42,384	35,491
May 2019	188,524	182,648	167,403
June 2019	520,882	541,680	544,088
July 2019	958,281	971,889	879,711

ARIMA models are one of the more complex forecasting methods, but they do a very good job separating and reacting to changes in time-series data. The models used to test and forecast monthly visitation to GNP created point forecasts and corresponding confidence bounds that would give park management a clearer view of what to expect in time periods ahead. The unfortunate part about the ARIMA models being more complicated is that they may need to be re-evaluated and re-specified every so often as patterns in the data change.

7 Regressions for Going-to-the-Sun Road Bicycle Use

Experiencing GTSR before it becomes congested with automobiles is a great option for people willing to ride their bikes up it. This use of GNP fits with trying to get people to visit during the shoulder season, so GNP began recording cyclist counts in 2016. Between 2016 and 2017 there was a large increase in the number of cyclists riding GTSR as can be seen in Tables 26, 27, 28. Over this period east-bound weekend use increased by 20%, weekday use by 63.08%, and Memorial Day use by 200%. Since 2018 counts were not available due to malfunctioning counters, it was exciting to see if an upward trend continued to form through to 2019. Now that 2019 counts are in, it appears any upward trend has dampened. As May and June bicycle use on GTSR becomes more popular, parking in areas where cyclists depart from has become increasingly congested. The Avalanche campground and trailhead is the preferred starting point for those wanting to ride their bikes up GTSR, but on weekends and Memorial Day parking spaces run out and potential cyclists are forced to park in pull offs farther back toward Lake McDonald or to take shuttles from the visitor center. Under the current regime cyclists on GTSR will not show the growth in numbers that they did between 2016 and 2017, but instead will likely remain near the numbers seen in 2017 and 2019. Memorial Day 2019 shows how large daily cyclist use on GTSR can be, though. Tables 25 and 28 show Memorial Day 2019 had 734 east-bound cyclists, the most cyclists of any day in the data set.

Table 25: Bicycle Regression Summary Statistics

	obs.	mean	sd	min	max	range	se
2016	138	0.3	0.5	0.0	1.0	1.0	0.0
May	138	0.6	0.5	0.0	1.0	1.0	0.0
June	138	0.4	0.5	0.0	1.0	1.0	0.0
aveb_bicycles	138	168.3	171.5	0.0	734.0	734.0	14.6
Weekend	138	0.3	0.5	0.0	1.0	1.0	0.0
Memorial.Day	138	0.0	0.1	0.0	1.0	1.0	0.0
Rain (Dummy)	138	0.4	0.5	0.0	1.0	1.0	0.0
Precip. (Inches)	138	0.1	0.3	0.0	2.4	2.4	0.0
Temp. MAX (Fahrenheit)	138	68.8	9.9	44.0	89.0	45.0	0.8
Temp. MIN (Fahrenheit)	138	42.5	5.7	30.0	56.0	26.0	0.5

Table 26: Weekend Day Bicycle Use Percent Changes on GTSR

Weekend Days		
	East Bound	West Bound
Total Days Counted 2016	12	12
Total Count 2016	3483	3389
Average Weekend Use 2016	290	282
Total Days Counted 2017	16	16
Total Count 2017	5570	5484
Average Weekend Use 2017	348	343
Average Daily Weekend Percent Change '16-'17	20.00% Increase	21.63% Increase
2018 Extrapolation	417.60	417.19
2019 Extrapolation	501.12	507.43
Actual Average Daily Weekend Use 2019	368	364
Actual Percent Change'17-'19	5.75% Increase	6.12% Increase

Table 27: Weekday Day Bicycle Use Percent Changes on GTSR

Weekday Days		
	East Bound	West Bound
Total Weekdays Counted 2016	33	33
Total Count 2016	2129	1950
Average Weekday Use 2016	65	59
Total Weekdays Counted 2017	36	36
Total Count 2017	3823	3579
Average Weekday Use 2017	106	99
Average Weekday Day Percent Change '16-'17	63.08% Increase	67.80% Increase
2018 Extrapolation	172.86	166.12
2019 Extrapolation	281.90	278.75
Actual Average Weekday Use 2019	93	90
Actual Percent Change'17-'19	12.26% Decrease	9.09% Decrease

Table 28: Memorial Day Bike Use Percent Changes on GTSR

Memorial Day		
	East Bound	West Bound
2016	161	131
2017	483	462
Percent Change '16-'17	200.00% Increase	252.67% Increase
2019	734	783
Percent Change '17-'19	51.97% Increase	69.48% Increase

7.1 Model Selection

Models with the full selection of variables from the data set are first run. These models have as their dependent variable east-bound and west-bound counts of cyclists, both focusing only on the time period before GTSR opens to cars. They include as explanatory variables a dummy for the year 2016 which had much lower numbers of cyclist relative to 2017 and 2019, a dummy for the month of June, a dummy for if it is a weekend day, a dummy for Memorial Day, a dummy variable for if there was any rain that day, the high temperature that day, and weeks until GTSR opens to cars. An term interacting weekend days with rain is also added since the data show rain having a greater effect on cyclist numbers when it rains on a weekend day. On a rainy weekend day people may opt to do leisure activities other than biking up GTSR. Breusch-Godfrey tests are run on both models and their results do not indicate serial correlation is a problem (east-bound $\chi^2 = 4.9497$, $p = 0.6661$ and west-bound $\chi^2 = 4.9802$, $p = 0.6624$).

The model results in Table 29 show there is essentially no difference between the east-bound and west-bound count models. Graphs of cyclist counts in 2016 and 2017 before opening of GTSR to motor traffic have also shown east-bound and west-bound counts to be nearly the same. This changes once GTSR opens to traffic, but these regressions focus on the period before GTSR opens so the following models will only use east-bound counts.

A check is done for multicollinearity by looking at variable correlations. June and the number of weeks until GTSR opens are correlated so the dummy for June is dropped from what becomes the main model. The main regression model will include variables for 2016, Memorial Day, weekend days, if it rained, a weekend/rain interaction term, the high temperature, and weeks until GTSR opens.

Table 29: Full Model Run on East and West Bound Cyclist Counts

	Dependent variable:	
	Eastbound Bicycles	Westbound Bicycles
	(1)	(2)
2016	-59.966^{***}	-61.820^{***}
	(19.125)	(19.188)
June	-9.741	-5.916
	(33.401)	(33.511)
Memorial Day	359.315^{***}	366.161^{***}
	(59.052)	(59.246)
Weekend	331.989^{***}	331.999^{***}
	(24.643)	(24.724)
Rain	-31.817	-29.455
	(21.716)	(21.787)
Weekend:Rain	-154.120^{***}	-156.176^{***}
	(37.542)	(37.665)
TempMAX	3.307^{***}	3.258^{***}
	(1.032)	(1.035)
Weeks.Until.GTSR.Opens	-9.502	-9.533
	(7.649)	(7.674)
Constant	-64.668	-68.109
	(80.059)	(80.322)
Observations	138	138
R^2	0.698	0.697
Adjusted R^2	0.679	0.678
Residual Std. Error (df = 129)	97.146	97.465
F Statistic (df = 8; 129)	37.250^{***}	37.097^{***}
Note:		*p<0.1; **p<0.05; ***p<0.01

7.2 GTSR Cyclist Regression Results

Main model results are shown in Table 30 and all variables except for the rain variable show a significant effect on the number of cyclists riding GTSR before it opens to motorized vehicles. Year 2016, Memorial Day, Weekend days, Weekend days interacted with Rain, and High Temperature are significant at the 1% level. Weeks-Until-GTSR-Opens is significant at 10% level.

Of the three years in the data set, 2016 had an average of 58.15 less cyclists per day than did 2017 and 2019, holding all else constant. 2016 having such a negative, significant coefficient suggests it was correct to include a term for that year. Memorial Day generates an average of 363.02 extra cyclists relative to a usual weekday, ceteris paribus. A weekend day sees about 331.80 more cyclists than a weekday, other things equal. A normal weekend day does not draw quite as many cyclist as Memorial Day does, but they are close. Weekend days when it rains do not see as high of an increase of cyclists on GTSR. A one degree Fahrenheit increase in a day's maximum temperature is associated with 3.20 more cyclists riding GTSR that day. One might think Weeks-Until-GTSR-Opens should have a positive coefficient since weather generally improves the closer it gets to when GTSR opens, but weather is held constant through other explanatory variables. Holding all other variables constant, the coefficient on Weeks-Until-GTSR-Opens says fewer cyclist ride GTSR as it gets closer to the time when GTSR opens to cars.

Table 30: GTSR Cycling Regression - Main Results

	Dependent variable:
	Cyclists
2016	-58.150^{***}
	(18.019)
Memorial Day	363.020^{***}
	(57.466)
Weekend	331.798^{***}
	(24.548)
Rain	-33.096
	(21.193)
Weekend:Rain	-154.187^{***}
	(37.409)
TempMAX	3.196^{***}
	(0.956)
Weeks.Until.GTSR.Opens	-7.673^{*}
	(4.362)
Constant	-68.826
	(78.501)
Observations	138
R^2	0.698
Adjusted R^2	0.681
Residual Std. Error	96.803 (df = 130)
F Statistic	42.861^{***} (df = 7; 130)
Note:	*p<0.1; **p<0.05; ***p<0.01

7.3 Cylist Regression Robustness Tests

The dummy variable for June was dropped due to multicollinearity, so a model with June instead of Weeks-Until-GTSR-Opens is run to check how this affects results. Table 31 shows the main model in Column (1) and a model with the June dummy in Column (2). The fit of the main model in Column (1) is better than the model in Column (2) so it seems dropping the June dummy variable was the right choice.

The first year cyclists were counted was 2016 and the counts that year were much lower than in 2017 and 2019. A special term for 2016 was added to the model to control for and show the lower effect of this year. The 2016 year dummy is significant in the main model and Table 32 gives a comparison of it next to one run without the 2016 dummy. Most other coefficients and standard errors do not change much, although Weeks-Until-GTSR-Opens loses its significance and the overall fit of the model as judged by adjusted R^2 goes down. Comparing these two models shows no reason to remove the 2016 dummy variable from the main model.

When Weeks-Until-GTSR-Opens is removed from the regression, as in Column (2) of Table 33, adjusted R^2 goes down. This indicates having Weeks-Until-GTSR-Opens in the model is not over-fitting even though the variable itself is barely significant.

Maximum temperature for a day is included in the main regression. The minimum temperature for a day is also part of the data set used, but people will ride GTSR during parts of the day when the temperature is highest. Due to this it is assumed most will plan according to what high temperatures are forecast to be. Table 34 shows the maximum temperature for a day does appear to have greater influence over how many cyclists ride GTSR. Column (2) of Table 34 shows results when minimum instead of maximum temperature is used. With this model parameterisation, minimum daily temperature does not affect the quantity of cyclists riding GTSR. The variables Rain and Weeks-Until-GTSR-Opens experience increased significance, but this may be caused by the high temperature variable no longer controlling important weather characteristics that the low temperature variable is unable to control for. A day's high temperature explains variation in cyclist GTSR use better than its low temperature does.

Table 31: Robustness Testing: June

	Dependent variable:	
	Cyclists	
	(1)	(2)
2016	-58.150^{***}	-50.813^{***}
	(18.019)	(17.686)
Memorial.Day	363.020^{***}	374.153^{***}
	(57.466)	(57.952)
Weekend	331.798^{***}	331.110^{***}
	(24.548)	(24.684)
Rain	-33.096	-35.360
	(21.193)	(21.573)
Weekend:Rain	-154.187^{***}	-152.714^{***}
	(37.409)	(37.603)
TempMAX	3.196^{***}	3.120^{***}
	(0.956)	(1.023)
Weeks.Until.GTSR.Open	-7.673^{*}	
	(4.362)	
June		24.287
		(19.156)
Constant	-68.826	-107.942
	(78.501)	(72.235)
Observations	138	138
R^2	0.698	0.694
Adjusted R^2	0.681	0.678
Residual Std. Error (df = 130)	96.803	97.349
F Statistic (df = 7; 130)	42.861^{***}	42.175^{***}

Note: *p<0.1; **p<0.05; ***p<0.01

Table 32: Robustness Testing: 2016

	Dependent variable:	
	Cyclists	
	(1)	(2)
2016	-58.150^{***}	
	(18.019)	
Memorial Day	363.020^{***}	365.932^{***}
	(57.466)	(59.488)
Weekend	331.798^{***}	331.512^{***}
	(24.548)	(25.414)
Rain	-33.096	-33.006
	(21.193)	(21.941)
TempMAX	3.196^{***}	3.635^{***}
	(0.956)	(0.979)
Weeks.Until.GTSR.Opens	-7.673^{*}	-4.599
	(4.362)	(4.407)
Weekend*Rain	-154.187^{***}	-145.236^{***}
	(37.409)	(38.623)
Constant	-68.826	-132.974^{*}
	(78.501)	(78.624)
Observations	138	138
R^2	0.698	0.673
Adjusted R^2	0.681	0.659
Residual Std. Error	96.803 (df = 130)	100.222 (df = 131)
F Statistic	42.861^{***} (df = 7; 130)	45.032^{***} (df = 6; 131)
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 33: Robustness Testing: No Week-Until-GTSR-Opens

	Dependent variable:	
	Cyclists	
	(1)	(2)
2016	-58.150^{***}	-51.227^{***}
	(18.019)	(17.724)
Memorial Day	363.020^{***}	367.697^{***}
	(57.466)	(57.861)
Weekend	331.798^{***}	331.275^{***}
	(24.548)	(24.741)
Rain	-33.096	-30.811
	(21.193)	(21.321)
Weekend:Rain	-154.187^{***}	-150.140^{***}
	(37.409)	(37.635)
TempMAX	3.196^{***}	3.689^{***}
	(0.956)	(0.921)
Weeks.Until.GTSR.Opens	-7.673^{*}	
	(4.362)	
Constant	-68.826	-139.695^{**}
	(78.501)	(67.911)
Observations	138	138
R^2	0.698	0.690
Adjusted R^2	0.681	0.676
Residual Std. Error	96.803 (df = 130)	97.574 (df = 131)
F Statistic	42.861^{***} (df = 7; 130)	48.710^{***} (df = 6; 131)
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 34: Robustness Testing: Temperature

	Dependent variable:	
	Cyclists	
	(1)	(2)
2016	-58.150^{***}	-68.395^{***}
	(18.019)	(18.817)
Memorial Day	363.020^{***}	349.920^{***}
	(57.466)	(59.753)
Weekend	331.798^{***}	332.147^{***}
	(24.548)	(25.600)
Rain	-33.096	-47.348^{**}
	(21.193)	(21.886)
Weekend:Rain	-154.187^{***}	-175.375^{***}
	(37.409)	(38.718)
Weeks.Until.GTSR.Opens	-7.673^{*}	-13.262^{***}
	(4.362)	(4.965)
TempMAX	3.196^{***}	
	(0.956)	
TempMIN		-0.964
		(1.771)
Constant	-68.826	228.339^{**}
	(78.501)	(89.427)
Observations	138	138
R^2	0.698	0.672
Adjusted R^2	0.681	0.655
Residual Std. Error (df = 130)	96.803	100.767
F Statistic (df = 7; 130)	42.861^{***}	38.124^{***}
Note:	*p<0.1; **p<0.05; ***p<0.01	

7.4 Conclusions from GTSR Cyclist Regression

Analysis of cyclists riding GTSR started with only two years of data from 2016 and 2017. That data showed a large uptick in cyclists riding GTSR, so the big question with 2018 data missing was if a strong upward trend had continued or if things had leveled off. Growth in cycling on GTSR has leveled off as effects from weekend parking congestion begin to set in. Businesses in communities surrounding GNP rent bikes to people to ride GTSR though, so the park will likely have help facilitating increased bicycle use on GTSR. Prior to GTSR opening to cars, the seasonal weekend effect and Memorial Day have the largest influence on how many people ride GTSR. This changes dramatically if it is raining.

The Avalanche Lake parking lot fills up on Memorial Day and weekends. There may be a ceiling to the number of people biking on GTSR without further accommodations. Shuttles with bike racks/trailers are running, so people need to be aware that the shuttle may be their best option for getting to the Avalanche area because parking is limited. Growth in cyclist use of GTSR could otherwise be capped by parking. This is not be a problem on weekdays though, so weekdays should be treated like shoulder months are when considering monthly visitation. Weekdays see a relatively lower level of cyclists riding GTSR. Weekdays were the only time that saw a drop in overall use, going down by about 12% in the east-bound direction from 2017 to 2019. Increasing weekday use may hold the most potential for allowing more people to experience the beauty of GTSR from the saddle of their bicycle. Cyclists should be encouraged to avoid the crowd and come on a weekday if possible, they will get the full experience with less hassle.

8 Results

Multiple forms of modeling Glacier National Park use allows for unique consideration of when and why use numbers fluctuate. Seasonal movements are so strong that it is easy for other influences to fade into the background, but once seasonal patterns are controlled for it is clear they are not alone.

Regressions run on GNP monthly visitation add to an growing understanding of how forest fires change visitation patterns. As found in previous research, the macro economy and travel cost are important factors when explaining levels of visitation to a national park. Gauging the effect of forest fire at a specific park is what made models in this paper unique though. Over the 2012 to 2018 observation period there were years with severe fires next to years without any major fires, making this a good time period for measuring how many fewer visitors GNP gets when forest fires are burning. Measures of air quality were the preferred method to proxy for forest fires and when this was done a causal relationship could be quantified.

Models in this paper found GNP to be a normal good in economic terms; a good which people consume more of when they have more money. The seasonally adjusted national unemployment rate was used in the regression models to fill the need for seeing how changes in the economy at large play with GNP visitation. All models in this paper found an inverse relationship between unemployment and GNP visitation, indicating GNP visitation is a normal good. Previous literature, mentioned in the Literature Review section, found national parks to be both normal and inferior (people demand more of something when they have less money) goods, but their findings involved the park system as a whole or involved the age/status of the park. The national park system is so large and diverse that it could be some parks are normal goods while others are inferior goods. The population density around Great Smoky Mountains National Park (GSMNP) is far higher than that around GNP, so even though both offer a relatively similar outdoor experience, the pool of people who can inexpensively drive to GSMNP is much larger than those who could just pop over to GNP for a quick visit. Considering the time and distance required to get to GNP, it being a normal good does fit within reason.

Forecast models combined the levels of previous years and months to generate probable ranges of what monthly visitation to GNP will be in 2019. Three important months for visitation have passed by the time this was written. May, June, and July were all within the 80% confidence prediction intervals generated by the ETS and ARIMA models. Projected visitation in May through July 2019 was 1,696,217 with the ETS forecast and 1,667,686 with the ARIMA forecast, while the actual total is 1,591,202. The ETS model has been closest during the first three big months of 2019. July experiences the highest visitation of any month and in July 2019 GNP had the lowest visitation total of any of the last three years. While none of the terms in the ARIMA model account for observations greater than three years old, the ETS model does. The ETS model did not only take into account the last two years, when there was a large spike in visitation, and this allowed it to not generate as high of forecasts. 2019 is showing a dip in visitation relative to the last two years and this has led to the ETS forecasts being more accurate given the way the ARIMA model is specified. With the hindsight knowledge that visitation would cool off, telling the ARIMA model to use observations older than two years could be a safer plan even if the models were not the best as far as the goodness-of-fit measures indicate. This could be a worthwhile insight to consider for those looking to do future forecasts involving national park visitation.

Cycling on GTSR is a fun, new way for people to enjoy GNP. It is a leisure experience that people voluntarily participate in for a variety of reasons.

With this in mind it is understandable that a rainy or cool weekend might change someone's mind if they had been planning to drive to GNP and bike GTSR. If what would be a fun bike ride turns into a damp or cold bike ride, people will substitute away from this activity to something they think will be more enjoyable. On dry, warm weekends in May and June before GTSR opens to automobiles, GNP should expect more cyclist to arrive than there are parking spaces at the Avalanche campground.

9 Conclusions

Visitation estimates with confidence intervals set a level for park management and gateway communities to plan from. Combining this with a quantified causal understanding of what affects visitation demand enables park and business planning for unpredictable events. Forest fires cannot be predicted, but they will become more frequent and severe as the climate warms. Detailed contingency plans based on previous changes in visitation due to fire will aid in park operation and continued economic opportunities in the surrounding areas.

Regression results used with forecasts allow for different scenarios to be tested. The ARIMA point forecast for August 2019 is 845,126.6 visitors and regression results indicate that an extra day of poor air quality, the proxy for forest fire due to all the smoke they cause, leads to a 1.76% decrease in a month's visitation. If August 2019 were to have 10 days of bad air quality from forest fire smoke (this is even less than the 15 days in August 2018), the forecast of visitation in this scenario could be lowered by 17.6% to 696,384.3 visitors. No one will know exactly how severe fires are going to be and how much smoke they will make, but using the regression results along with forecast levels of visitation can lead to better contingency planning by park management.

Visitation increased from 2012 to 2017, then severe fires in August of 2018 lowered that month's visitation enough for 2018 counts to fall below 2017, breaking the streak of increasing yearly visitation. 2018 was still the second highest total visitation ever recorded at GNP, though. This dip in visits from the high set in 2017 seems to be continuing in 2019 as 2019 is on track to see fewer visitors than 2017 and 2018 did. Given this change and that forecasts in this analysis use monthly counts from a time period which almost exclusively saw increasing numbers of people, the forecasts generated in this paper over-shot how many people would make the trip to GNP. So this begs the question, what are the implications of under forecasting and what are those of over forecasting?

Were park management to base their plans for a given year on forecasts of visitation which were too low, what would the consequences be? Parking lots, the number of campsites, and bathroom facilities are fixed resources in the short term. If the park thinks they have an adequate amount of each, there would not be any reason to divert resources away from other uses so as to ready the park for a visitation swell that goes beyond what it is thought to be able to handle. To a certain extent, one could look at July 2017, which was the busiest month ever at GNP, and try draw conclusions from how that level of visitation affected park operations. Parking lots were quickly overflowing, sometimes leading to visitors attempting to park places where they are not allowed (this happens months other than July 2017 as well). Park rangers being forced to deal with increased illegal parking pulls them away from better uses of their time. Campsites reached capacity early in the morning and bathrooms were used at a rate that was difficult for park staff to keep up with. Since many resources are fixed and managing a huge national park with dramatic seasonal visitation swings is – regardless of expectations – a monumental challenge, some congestion issues are unavoidable.

Increasing shuttle access and bringing in portable bathroom facilities are measures that could be taken to compensate for an above average July or August, but what if a month where visitation is expected to be off the charts pans out to be average or even below average? The park will have wasted resources over-preparing for something it was already prepared for. The big problem with being over-prepared is that money is already stretched thin at GNP and across the park system as a whole. Money cannot be spent to alleviate congestion that never comes because that money is desperately needed elsewhere.

Implications from under and over forecasting highlight issues the park is already facing. These problems are exactly what the forecasts and scenarios created with the regression results are intended to help with, although one could imagine a situation where some outside force not included in the models leads to a systematic change in visitation patterns and the findings of this research do nothing but muddy the waters of park management and planning. This is a reminder that all model results should not be taken as truth, but used to enhance the understanding of people who already have an intimate qualitative understanding of GNP.

With 2019 looking to continue the decline in visitation observed in 2018, this is a good chance to consider forces beyond what has been modeled which could drive down how many people come to GNP. Beginning in 2016 the NPS launched the "Find Your Park" campaign to celebrate its 100^th birthday (National Park Service, 2018c). This was an advertising campaign which helped bring extra attention to NPS sites. Could the "Find Your Park" campaign have contributed to a spike in visitation that peaked in 2017? Also, after bad fire years in 2015, 2017, and 2018, are people's perception of GNP changing, leading them to expect fires in the late summer so they plan trips elsewhere? Both of these questions would require further research to answer as they deal with how people perceive GNP.

Congestion will remain an issue in July and August. The more activities that draw people to the park outside these months, the better. Cycling on GTSR in May and June is a great example of this and further increases in this use seems beneficial to GNP and the surrounding communities. Congestion issues often center around the automobile being the primary mode of transportation. Parking issues have already arisen for those driving to GNP in May and June to ride their bikes on GTSR, but finding more ways to facilitate non-motorized use throughout GNP could incentivise people to visit during shoulder months. The popularity that riding on GTSR quickly gained provides a good example of this. Disruptive natural events exacerbated by climate change will continue to affect GNP with increased likelihood and severity. Carbon emissions are not abating and locations at northern latitudes with higher elevations, like GNP, experience the resulting climate changes in dramatic fashion. People want to see the glaciers before they melt, but they do not want to deal with the closures and smoke associated with forest fires. Within this changing landscape, Glacier National Park and its large, historic, wild spaces may have more to offer than ever before.

References

Akaike, H. (1974). A new look at the statistical model identification. In *Selected papers of hirotugu akaike* (pp. 215–222). Springer.

Alley, L. (2018). *Glacier national park changes entrance fees.* NPS - U.S. Department of the Interior. Retrieved 2019-09-27, from `https://www.nps.gov/glac/learn/news/18-11.htm`

Barry, K., & O'Hagan, J. (1972). An econometric study of british tourist expenditure in ireland. *Economic and Social Review*, *3*(2), 143–161.

Battaglia, B. G. (2016). Bicycle mobility in glacier national park: Assessing going-to-the-sun road travelers' attitudes, knowledge, and perceptions of bicycling. *ScholarWorks at University of Montana*. Retrieved 2019-09-27, from `https://scholarworks.umt.edu/etd/10682/`

Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (1970). *Time series analysis: forecasting and control.* John Wiley & Sons.

Burkart, A. J., Medlik, S., et al. (1981). Tourism: past, present and future. *Tourism: past, present and future.*(Ed. 2).

Chu, F.-L. (2009). Forecasting tourism demand with arma-based methods. *Tourism Management*, *30*(5), 740–751.

Davies, D., Halliday, M., Mayes, M., & Pocock, R. (1997). *Attitudes to cycling: a qualitative study and conceptual framework.* Transport Research Laboratory Crowthorne.

Duffield, J. W., Neher, C. J., Patterson, D. A., & Deskins, A. M. (2013). Effects of wildfire on national park visitation and the regional economy: a natural experiment in the northern rockies. *International journal of wildland fire*, *22*(8), 1155–1166.

Flannigan, M. D., Stocks, B. J., & Wotton, B. M. (2000). Climate change and forest fires. *Science of the total environment*, *262*(3), 221–229.

Gabbery, B. (2013). *Fires in yellowstone.* Wildfire Today. Retrieved 2019-09-27, from `https://wildfiretoday.com/2013/08/17/five-fires-in-yellowstone/`

Geurts, M. (1982). Forecasting the hawaiian tourist market. *Journal of Travel Research*, *21*(1), 18–21.

Geurts, M., & Ibrahim, I. (1975). Comparing the box-jenkins approach with the exponentially smoothed forecasting model application to hawaii tourists. *Journal of Marketing Research*, 182–188.

Glacier National Park. (2012a). *Fire fighters responding to wildland fire.* National Park Service. Retrieved 2019-09-27, from `https://www.nps.gov/glac/learn/news/media12-69.htm`

Glacier National Park. (2012b). *Better mapping estimates fire at 45 acres.* National Park Service. Retrieved 2019-09-27, from `https://www.nps.gov/glac/learn/news/media12-70.htm`

Glacier National Park. (2012c). *Fire fighters working fire perimeter.* National Park Service. Retrieved 2019-09-27, from `https://www.nps.gov/glac/learn/news/media12-71.htm`

Glacier National Park. (2012d). *Fire fighters continue containment efforts.* National Park Service. Retrieved 2019-09-27, from `https://www.nps.gov/glac/learn/news/media12-72.htm`

Glacier National Park. (2012e). *Weather slows wildland fire activity.* National Park Service. Retrieved 2019-09-27, from `https://www.nps.gov/glac/learn/news/media12-74.htm`

Glacier National Park. (2012f). *Slides close portion of going-to-sun road.* National Park Service. Retrieved 2019-09-27, from `https://www.nps.gov/glac/learn/news/media12-45.htm`

Glacier National Park. (2012g). *Crews continue to remove debris on road.* National Park Service. Retrieved 2019-09-27, from `https://www.nps.gov/glac/learn/news/media12-46.htm`

Glacier National Park. (2012h). *Entire going-to-the-sun road now open.* National Park Service. Retrieved 2019-09-27, from `https://www.nps.gov/glac/learn/news/media12-47.htm`

Glacier National Park. (2012i). *Going-to-the-sun road rehabilitation pro-*

gresses. National Park Service. Retrieved 2019-09-27, from https://www.nps.gov/glac/learn/news/media12-38.htm
Glacier National Park. (2012j). *Conditions limit access to logan pass.* National Park Service. Retrieved 2019-09-27, from https://www.nps.gov/glac/learn/news/conditions-limit-access-to-logan-pass.htm
Glacier National Park. (2014). *Snow and boulder fall on going-to-the-sun road.* National Park Service. Retrieved 2019-09-27, from https://www.nps.gov/glac/learn/news/snow-and-boulder-fall-on-going-to-the-sun-road.htm
Glacier National Park. (2015a). *Firefighters responding to wildland fire.* National Park Service. Retrieved 2019-09-27, from https://www.nps.gov/glac/learn/news/firefighters-responding-to-wildland-fire.htm
Glacier National Park. (2015b). *Thompson fire creates visible smoke column.* National Park Service. Retrieved 2019-09-27, from https://www.nps.gov/glac/learn/news/thompson-fire-creates-visible-smoke-column.htm
Glacier National Park. (2015c). *Thompson fire spreads east.* National Park Service. Retrieved 2019-09-27, from https://www.nps.gov/glac/learn/news/thompson-fire-spreads-east.htm
Glacier National Park. (2015d). *Firefighters responding to wildland fire.* National Park Service. Retrieved 2019-09-27, from https://www.nps.gov/glac/learn/news/firefighters-responding-to-wildland-fire.htm
Glacier National Park. (2015e). *West-side access to logan pass planned.* National Park Service. Retrieved 2019-09-27, from https://www.nps.gov/glac/learn/news/west-side-access-to-logan-pass-planned.htm
Glacier National Park. (2015f). *East-side going-to-the-sun road*

open. National Park Service. Retrieved 2019-09-27, from https://www.nps.gov/glac/learn/news/east-side-going-to-the-sun-road-open.htm

Glacier National Park. (2015g). *Access to logan pass changes on monday.* National Park Service. Retrieved 2019-09-27, from https://www.nps.gov/glac/learn/news/access-to-logan-pass-changes-on-monday.htm

Glacier National Park. (2016). *Going-to-the-sun road anticipated to open tommorrow.* National Park Service. Retrieved 2019-09-27, from https://www.nps.gov/glac/learn/news/media16-35.htm

Glacier National Park. (2017a). *Sprague fire information: Saturday august 19.* National Park Service. Retrieved 2019-09-27, from https://www.nps.gov/glac/learn/news/fireinformation8_19.htm

Glacier National Park. (2017b). *Logan pass opening and closing dates.* National Park Service. Retrieved 2019-09-27, from https://www.nps.gov/glac/learn/news/upload/Logan-Pass-Open-Close-Dates_Press-Kit-6-26-2017.pdf

Glacier National Park. (2017c). *The going-to-the-sun road is fully open.* National Park Service. Retrieved 2019-09-27, from https://www.nps.gov/glac/learn/news/media17-29.htm

Glacier National Park. (2017d). *Evacuation order in effect between apgar loop road and logan pass.* National Park Service. Retrieved 2019-09-27, from https://www.nps.gov/glac/learn/news/media17-48.htm

Glacier National Park. (2017e). *Sprague, adair peak and elder creek fires.* National Park Service. Retrieved 2019-09-27, from https://www.nps.gov/glac/learn/news/upload/SpragueFactsheet-09282017.pdf

Glacier National Park. (2018a). *Going-to-the-sun road will be fully open on saturday for the 2018 season.* National Park Service. Retrieved 2019-09-27, from https://www.nps.gov/glac/learn/news/going-to-the-sun-road-will-be-fully-open-on-saturday-for

-the-2018-season.htm

Glacier National Park. (2018b). *Sunday evening fire update 2, 8 pm.* National Park Service. Retrieved 2019-09-27, from https://www.nps.gov/glac/learn/news/18-36.htm

Glacier National Park. (2018c). *West side of going-to-the-sun reopens to private vehicles.* National Park Service. Retrieved 2019-09-27, from https://www.nps.gov/glac/learn/news/18-51.htm

Glacier National Park. (2018d). *Going-to-the-sun road closed in logan pass area for season.* National Park Service. Retrieved 2019-09-27, from https://www.nps.gov/glac/learn/news/18-57.htm

Glacier National Park. (2018e). *Weather temporarily closes logan pass area of going-to-the-sun road.* National Park Service. Retrieved 2019-09-27, from https://www.nps.gov/glac/learn/news/18-53.htm

Holt, C. C. (1957). Forecasting seasonals and trends by exponentially weighted moving averages. *International journal of forecasting*, *20*(1), 5–10.

Hyndman, Rob J and Athanasopoulos, George. (2018). *Forecasting: principles and practice.* OTexts. Retrieved 2019-09-27, from https://otexts.com/fpp2/

Incident Information System. (2016). *Maple fire.* National Park Service. Retrieved 2019-09-27, from https://inciweb.nwcg.gov/incident/6111/

Incident Information System. (2017a). *Sprague fire.* National Park Service. Retrieved 2019-09-27, from https://inciweb.nwcg.gov/incident/5510/

Incident Information System. (2017b). *Adair peak fire.* National Park Service. Retrieved 2019-09-27, from https://inciweb.nwcg.gov/incident/5590/

Incident Information System. (2018a). *Howe ridge fire.* National Park Service. Retrieved 2019-09-27, from https://inciweb.nwcg.gov/

incident/6135/

Incident Information System. (2018b). *Boundary fire.* National Park Service. Retrieved 2019-09-27, from https://inciweb.nwcg.gov/incident/6170/

Incident Information System. (2018c). *Bacon rind fire.* National Park Service. Retrieved 2019-09-27, from https://inciweb.nwcg.gov/incident/5992/

Insider, Y. (2013). *Alum fire forces closure of yellowstone road segment.* August Publications. Retrieved 2019-09-27, from https://yellowstoneinsider.com/2013/08/20/alum-fire-forces-closure-of-yellowstone-road-segment/

Insider, Y. (2014). *One acre burned this fire season in yellowstone national park.* August Publications. Retrieved 2019-09-27, from https://yellowstoneinsider.com/2017/10/26/one-acre-burned-fire-season-yellowstone-national-park/

IRMA, NPS, GLAC annual visits. (2019). *Glacier np annual park recreation visitation (1904-last calendar year).* U.S. Department of the Interior. Retrieved 2019-09-27, from https://irma.nps.gov/Stats/SSRSReports/Park\%20Specific\%20Reports/Annual\%20Park\%20Recreation\%20Visitation\%20(1904\%20-\%20Last\%20Calendar\%20Year)?Park=GLAC

Johnson, R. L., & Suits, D. B. (1983). A statistical analysis of the demand for visits to us national parks: travel costs and seasonality. *Journal of Travel Research*, *22*(2), 21–24.

Kamp, D. B., Crompton, J. L., & Hensarling, D. M. (1979). The reactions of travelers to gasoline rationing and to increases in gasoline prices. *Journal of Travel Research*, *18*(1), 37–41.

Lamont, M. (2009). Reinventing the wheel: A definitional discussion of bicycle tourism. *Journal of Sport & Tourism*, *14*(1), 5–23.

Lumsdon, L. (2000). Transport and tourism: cycle tourism–a model for sustainable development? *Journal of Sustainable Tourism*, *8*(5), 361–377.

Mace, B. L., Bell, P. A., & Loomis, R. J. (2004). Visibility and natural quiet in national parks and wilderness areas: Psychological considerations. *Environment and Behavior*, *36*(1), 5–31.

Maughan, R. (2012). *Yellowstone park gets its first big fire of the year.* The Wildlife News. Retrieved 2019-09-27, from http://www.thewildlifenews.com/2012/09/16/yellowstone-park-gets-its-first-big-fire-of-the-year/

McIntosh, C. R., & Wilmot, N. (2011). An empirical study of the influences of recreational park visitation: the case of us national park service sites. *Tourism Economics*, *17*(2), 425–435.

Montag, V., Hogle, A., & Pirillo, D. (2012). *Fires in yellowstone: Past, present and future.* Science Education Resource Center (SERC) at Carleton College. Retrieved 2019-09-27, from https://serc.carleton.edu/NZFires/megafires/Yellowstone.html

National Park Service. (2018a). *Quick history of the national park service (u.s. national park service).* U.S. Department of the Interior. Retrieved 2019-09-27, from https://www.nps.gov/articles/quick-nps-history.htm

National Park Service. (2018b). *National park system sees more than 330 million visits.* U.S. Department of the Interior. Retrieved 2019-09-27, from https://www.nps.gov/orgs/1207/02-28-2018-visitation-certified.htm

National Park Service. (2018c). *Find your park.* U.S. Department of the Interior. Retrieved 2019-09-27, from https://www.nps.gov/subjects/centennial/findyourpark.htm

National Park Service. (2019a). *Integrated resource management applications.* Retrieved 2019-01-24, from https://irma.nps.gov/Portal/

National Park Service. (2019b). *Recreational visitors by month glacier np*

(1979 - last calendar year. U.S. Department of the Interior. Retrieved 2019-09-27, from https://irma.nps.gov/Stats/SSRSReports/Park\%20Specific\%20Reports/Recreation\%20Visitors\%20By\%20Month\%20(1979\%20-\%20Last\%20Calendar\%20Year)?Park=GLAC

National Park Travelers Club. (2019). *Welcome to the national park travelers club.* National Park Travelers Club. Retrieved 2019-09-27, from https://www.parkstamps.org/

Oh, C.-O., & Hammitt, W. E. (2011). Impact of increasing gasoline prices on tourism travel patterns to a state park. *Tourism Economics*, *17*(6), 1311–1324.

Ostergren, D., Solop, F. I., & Hagen, K. K. (2005). National park service fees: Value for the money or a barrier to visitation? *Journal of Park & Recreation Administration*, *23*(1).

Ouellet, Nicky. (2017). *Glacier park's scorched sperry chalet was a haven for the hungry hiker.* National Public Radio - MTPR. Retrieved 2019-09-27, from https://www.npr.org/sections/thesalt/2017/09/09/549561013/glacier-parks-scorched-sperry-chalet-was-a-haven-for-the-hungry-hiker

O'Connor, J. P., & Brown, T. D. (2010). Riding with the sharks: serious leisure cyclist's perceptions of sharing the road with motorists. *Journal of Science and Medicine in Sport*, *13*(1), 53–58.

Papadogiannaki, E., Braak, A., Holmes, N., Eury, D., & Hollenhorst, S. (2009). Great smoky mountains national park visitor study. *Park Studies Unit, Visitor Services Project, University of Idaho, Report*, *205*.

Papatheodorou, A., Rosselló, J., & Xiao, H. (2010). Global economic crisis and tourism: Consequences and perspectives. *Journal of Travel Research*, *49*(1), 39–45.

Poudyal, N. C., Paudel, B., & Green, G. T. (2013). Estimating the impact of impaired visibility on the demand for visits to national parks. *Tourism economics*, *19*(2), 433–452.

Poudyal, N. C., Paudel, B., & Tarrant, M. A. (2013). A time series analysis of the impact of recession on national park visitation in the united states. *Tourism Management*, *35*, 181–189.

Ritchie, J. B., Amaya Molinar, C. M., & Frechtling, D. C. (2010). Impacts of the world recession and economic crisis on tourism: North america. *Journal of Travel Research*, *49*(1), 5–15.

Sage, J. L., Nickerson, N. P., Miller, Z. D., Ocanas, A., & Thomsen, J. (2017). Thinking outside the park-national park fee increase effects on gateway communities.

Schwartz, Z., & Lin, L.-C. (2006). The impact of fees on visitation of national parks. *Tourism Management*, *27*(6), 1386–1396.

Smeral, E. (2010). Impacts of the world recession and economic crisis on tourism: Forecasts and potential risks. *Journal of Travel Research*, *49*(1), 31–38.

Song, H., & Li, G. (2008). Tourism demand modelling and forecasting—a review of recent research. *Tourism management*, *29*(2), 203–220.

Song, H., Turner, L., et al. (2006). Tourism demand forecasting. *International handbook on the economics of tourism*, 89–114.

Tracy, T., & Morris, H. (1998). The experience on 372 trails.

Uysal, M., & Crompton, J. L. (1985). An overview of approaches used to forecast tourism demand. *Journal of Travel Research*, *23*(4), 7–15.

Weiler, S., & Seidl, A. (2004). What's in a name? extracting econometric drivers to assess the impact of national park designation. *Journal of Regional Science*, *44*(2), 245–262.

Westerling, A. L., Hidalgo, H. G., Cayan, D. R., & Swetnam, T. W. (2006). Warming and earlier spring increase western us forest wildfire activity. *science*, *313*(5789), 940–943.

Winters, P. R. (1960). Forecasting sales by exponentially weighted moving averages. *Management science*, *6*(3), 324–342.

Yellowstone National Park. (2015a). *Lightning starts small fire in yellowstone.* National Park Service. Retrieved 2019-09-27, from https://www.nps.gov/yell/learn/news/15066.htm

Yellowstone National Park. (2015b). *Spruce fire increases in size to 425 acres.* National Park Service. Retrieved 2019-09-27, from https://www.nps.gov/yell/learn/news/15068.htm

Yellowstone National Park. (2015c). *Spruce fire continues to grow in yellowstone.* National Park Service. Retrieved 2019-09-27, from https://www.nps.gov/yell/learn/news/15069.htm

Yellowstone National Park. (2015d). *Spruce fire now over 2000 acres.* National Park Service. Retrieved 2019-09-27, from https://www.nps.gov/yell/learn/news/15070.htm

Yellowstone National Park. (2015e). *Spruce fire received significant rain overnight.* National Park Service. Retrieved 2019-09-27, from https://www.nps.gov/yell/learn/news/15072.htm

Yellowstone National Park. (2016). *Significant 2016 wildfire season.* National Park Service. Retrieved 2019-09-27, from https://www.nps.gov/yell/learn/news/16068.htm

Yellowstone National Park. (2018). *Fire updates for yellowstone national park.* National Park Service. Retrieved 2019-09-27, from https://www.nps.gov/yell/learn/news/08132018.htm