Trend analysis and model fitting of electricity generation and access in Ethiopia

Final Project for the course ENVIRON 790.30: Time Series Analysis for Energy Data

Yared S. Asfaw

2022-03-29

1. Introduction

Ethiopia has a huge potential in electricity generation from renewable sources mainly hydropower (45,000MW), geothermal (<10,000MW), solar (4-6 kwh/m2 per day) and wind (100GW). Of these potentials, the country utilized less than 5% of the hydropower potential and less than 1% of each one of the other three renewable sources of energy (Ethiopian Electric Power, 2021). As a result, with a population of over 114 million (78.31% of which resides in rural areas), according to the World Bank collection of development indicators compiled from officially recognized sources, the country's overall electricity access stands at 48.27% as of 2019. And this access is mainly concentrated in urban and peri-urban areas, only 36.28% of the rural population has access to electricity (World Bank, 2022) leaving the remaining portion of the population in energy poverty which leads to unsustainable utilization of natural resources to fulfil their energy demand.

Recognizing these facts, the governemnt of Ethiopia introduced a Growth and Transformation Plan (GTP) in 2005.The Country's GTP outlined a 15-year plan with three 5-year phases to transform the country from a developing country to a middle income country by 2025. In the plan the governemnt has put a goal of increasing the installed electricity generation capacity from 2000 to 10000MW primarily through hydropower projects (National Planning Commission, 2016). The government has also pioneered in developing the national Climate Resilient Green Economy Strategy in 2011 with the aim of building carbon neutral and climate resilient middle-income economy by 2025 (MoF, 2022) and launched a mulimillion dollar National Electrification Program in 2017 with the aim of achieveing universal electricity access nationawide by 2025 later revised by 2030 (65% of access provision is targeted with grid soutions and 35%$ with off-grid technologies (solar off-grid and mini-grids) (MoWIE, 2019).

2. Research motivation

The main motivation in this research is to lay a foundation for my master's project which aim to investigate the pros and cons of the Grand Ethiopian Renaissance Dam (GERD) to the riaprian countries by looking at the historical flow of the Blue Nile river. So that the findings can inform interested parties as a scientifc evidence to smoothly facilitate the negotiation process and take approrpatie decisions.Thus, this project will serve as a starting point for the longer journey by looking at the historical trend of electricity generation and consumption in the country.

3. Research objective:

With respect to the country's potential how is the country performing in terms of addressing the demand for electricity. To addresse this objective, the following reserach questions are taken into consideration

1. How is the total electricity generation and consumption changing from year to year?
2. How is the trend of electricity generation looks like?
3. How is the trend of access to electricity looks like?
4. What will be the state of electricity generation and access in the coming five years?

4. Dataset information

The datasets used for this project is obtained from the World Development Indicators Database. "The World Development Indicators is a compilation of relevant, high-quality, and internationally comparable statistics about global development and the fight against poverty. The database contains 1,400 time series indicators for 217 economies and more than 40 country groups, with data for many indicators going back more than 50 years" (https://datatopics.worldbank.org/world-development-indicators/).It is the primary World Bank collection of development indicators, compiled from officially recognized international sources. It presents the most current and accurate global development data available, and includes national, regional and global estimates. As compared to other data sources for this specific project, the World Bank database has a long years of data on the country's energy situation, thus, it is used as the sole source of dataset in this project.

Of the number of parameters available, for this project the dataset compiled on the country's yearly perecentage of electricity production from different sources, access to electricity and consumption of electric power from 1971 onwards is used. The dataset can be accessed from the webiste using the link https://databank.worldbank.org/source/world-development-indicators#.

With the objective of effective utilization of the available datasets on the variables of interest, extraction of the data for electricity production from different sources, access and consumption is done independently so that the historical trend of those parameters can be captured efficiently. i.e. the dataset is used separately to avoid unneccessary elimination of data points while removing NAs for the other variable.

```r
# Loading neccessary pacakges
library(openxlsx)
library(lubridate)
library(tidyverse)
library(dplyr)
library(forecast)
library(Kendall)
library(tseries)
library(outliers)
library(zoo)

# Creating a ggplot theme
mytheme <- theme_classic(base_size = 12) +
theme(axis.text = element_text(color = "black"), legend.position ="right")
# Setting up my_theme as default theme
theme_set(mytheme)
```

5. Time series analysis of electricity production and consumption

```r
# Importing the dataset and selecting relevant columns
ETH_ele_pro_cons <- read.xlsx("../ENV790_TimeSeriesAnalysis_Sp2022/Project/Data/Data_Extract_From_World_
  select(Series.Name, '1971.[YR1971]':'2020.[YR2020]')

# Transposing the rows and columns for easy of work
ETH_ele_pro_cons <- as.data.frame(t(ETH_ele_pro_cons))

# Renaming row names
rownames(ETH_ele_pro_cons) <- c("Year",seq(1971, 2020,1))

# Removing two columns which have no values
ETH_ele_pro_cons <- ETH_ele_pro_cons[-1,1:12]

#Renaming column names
colnames(ETH_ele_pro_cons) <- c("Access to electricity (% of population)",
"Access to electricity, urban (% of urban population)",
"Access to electricity, rural (% of rural population)",
"Electric power consumption (kWh per capita)",
"Electricity production from hydroelectric sources (% of total)",
"Electricity production from renewable sources, excluding hydroelectric (% of total)
",
"Electricity production from nuclear sources (% of total)",
"Electricity production from oil, gas and coal sources (% of to",
"Electricity production from oil sources (% of total)",
"Electricity production from natural gas sources (% of total)",
"Electricity production from coal sources (% of total)",
"Electricity production from renewable sources, excluding hy")

ETH_ele_pro_cons <- ETH_ele_pro_cons %>%
  select('Electricity production from hydroelectric sources (% of total)',
'Electricity production from renewable sources, excluding hy',
'Electricity production from oil sources (% of total)',
```

```
'Access to electricity (% of population)',
'Electric power consumption (kWh per capita)')

year <- seq(as.Date("1971/1/1"), as.Date("2020/1/1"), "years")
Date <- as.data.frame(year)
ETH_ele_pro_cons_clean <- cbind.data.frame(Date, ETH_ele_pro_cons)

# Changing the format from character to numeric
ETH_ele_pro_cons_clean <- ETH_ele_pro_cons_clean %>%
mutate(ele_pro_hydro=as.numeric(`Electricity production from hydroelectric sources (% of total)`)) %>%
mutate(ele_pro_ren=as.numeric(`Electricity production from renewable sources, excluding hy`)) %>%
mutate(ele_pro_oil=as.numeric(`Electricity production from oil sources (% of total)`)) %>%
mutate(Total_Access = as.numeric(`Access to electricity (% of population)`)) %>%
mutate(Power_Consumption = as.numeric(`Electric power consumption (kWh per capita)`)) %>%
  select(year, ele_pro_hydro, ele_pro_ren,ele_pro_oil, Total_Access, Power_Consumption)

# To effectively utilize the available historical data, data extraction for
# percentage of electricity production from hydropower, other renewables and oil from 1971 to 2015
ele_pro_source <- ETH_ele_pro_cons_clean %>%
  select(year, ele_pro_hydro, ele_pro_ren, ele_pro_oil) %>%
  filter(year<='2015-01-01')
```

5.1. Trend analysis of electricity generation by source

The focus for this time series anlysis will be electicity generation from hydropower systems (more than 90% of the electricity production relay on hydropower systems), renewable sources and oils, and per capita electicity consumption.

5.1.1. Electricity production from hydropower

Hydropower is the main and dominant source of electricity in the country as the country has a conducive topography as well as ample water sources. Approximately 90% of the installed generation capacity is from hydropower (US ITA, 2022). As a result the country has around 45,000 MW potential in electricity from hydropower, of which only less than 5% is being used so far. Despite these potentials

```
# plotting the trend of electricity production from hydropower over time
ggplot(ele_pro_source, aes(x=year, y=ele_pro_hydro))+
  geom_line(color="blue")+
  xlab("Year")+
  ylab("Percent of electricity production")
```

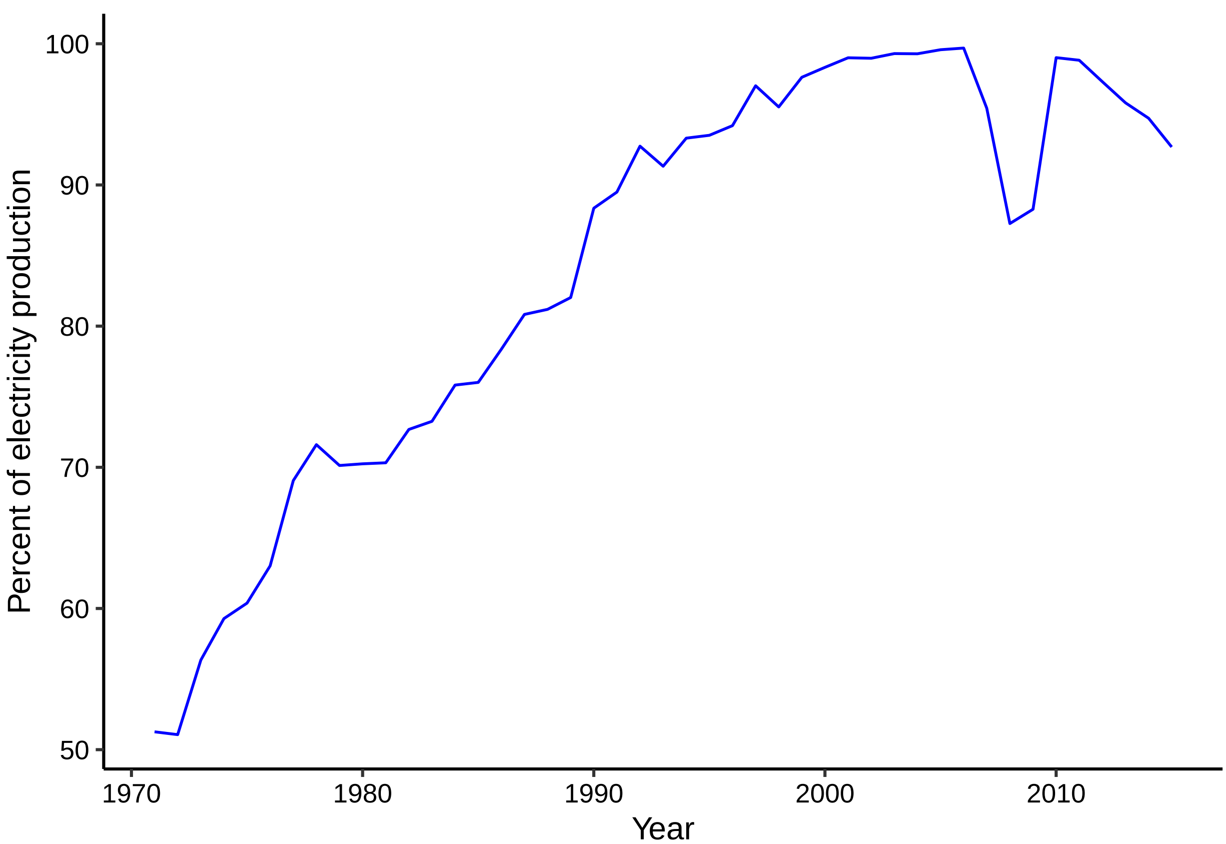
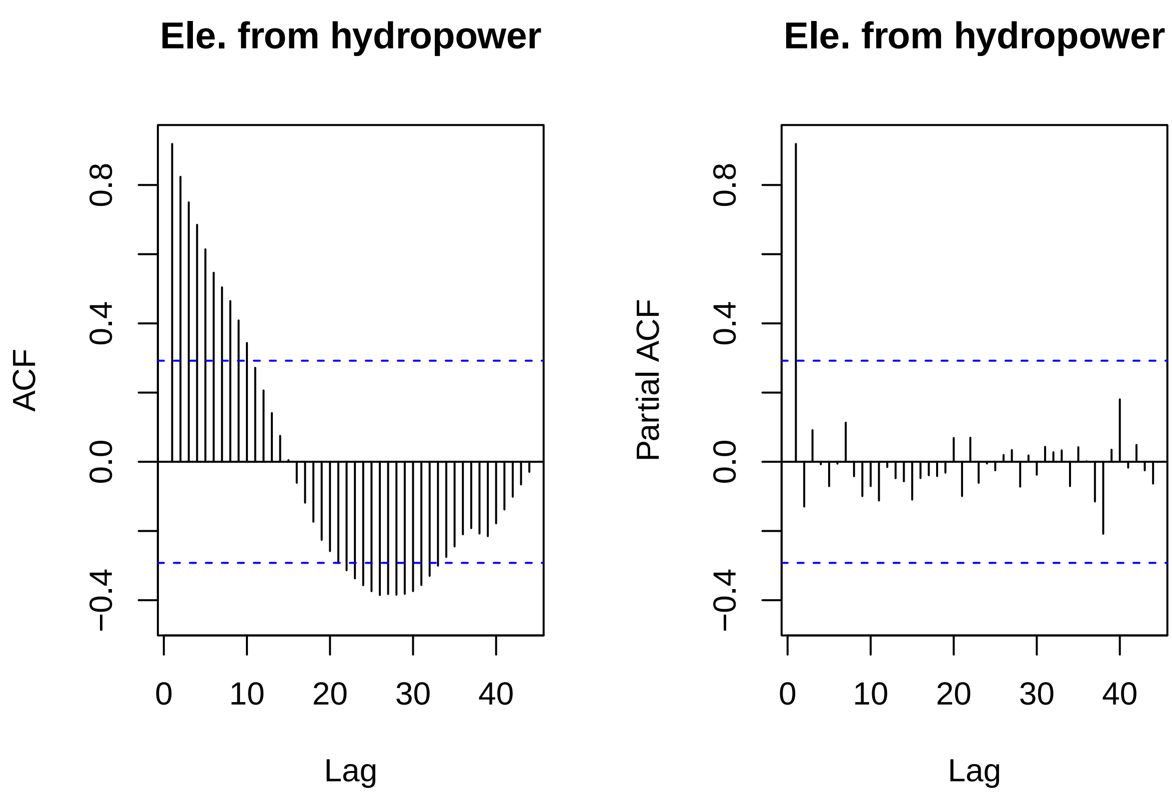


Figure 1: Trend of electricity production from hydropower in Ethiopia from 1971-2015

The plot shows an overall increasing trend of perecent of electicity production from hydropower overtime. In betwen 2008 and 2010, there were a significant decrease in the production of electricity, stayed lower for few years and again significantly returned back to its original level in 2010. This is mainly due to the the severe drought events that occurred in 2008, 2009, 2010 (WFP, 2022) forcing all hydro power dams to produce well below their capacity. However, even after bouncing back from the dent, the percentage of electricity production from hydropower is showing a decreasing trend as a result of the continued impacts of the drought and other drought events. The 2013-2015 marked a period of relative calm with no major drought incidence. However, from 2016-2021, 9 of the 12 rainfall seasons have exhibited mild to severe drought conditions affecting some geographic areas (WFP, 2022.The plot also shows there is no seasonal pattern of electricity genertion from hydropower sources.

```r
# Transforming the data into a time series object
ts_ele_pro_hydro <- ts(ele_pro_source$ele_pro_hydro, start = c(1971,1), frequency = 1)
# Plotting the Acf and Pacf plots
par(mfrow=c(1,2))
Acf(ts_ele_pro_hydro, lag.max = 60, main="Ele. from hydropower")
Pacf(ts_ele_pro_hydro, lag.max = 60, main="Ele. from hydropower")
```



The ACF and PACF plots of the percent of electricty production from hydropower shows an AR process with order p=1, q=0

Running stationarity tests

```
# Conducting Mann Kendall and ADF tests
MKtest_ele_hydro <- MannKendall(ts_ele_pro_hydro)
print(MKtest_ele_hydro)
```

```
## tau = 0.754, 2-sided pvalue =< 2.22e-16
```

The p-value is less than the significance level, thus, we reject the null hypothesis and say the series has a trend. We conduct the ADF test and check whether the trend is stochastic or not.

```
ADFtest_ele_hydro <- adf.test(ts_ele_pro_hydro)
print(ADFtest_ele_hydro)
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  ts_ele_pro_hydro
## Dickey-Fuller = -0.63604, Lag order = 3, p-value = 0.9686
## alternative hypothesis: stationary
```

The p-value is greater than the significance level, thus, we fail to reject the null hypothesis thus the series has a unit root. And conclude that the series has a stochastic trend. Therefore, we use differencing to achieve stationarity.

Modele fitting - electricity production from hydropower

```
# Since the series has a stochastic trend it needs to be differenced
# to achieve stationarity.
# Checking for the number of differences needed
n_diff_ele_hydro <- ndiffs(ts_ele_pro_hydro, alpha = 0.1, test = c("kpss", "adf", "pp"),
max.d = 2)
cat("Number of differencing needed: ",n_diff_ele_hydro)
```

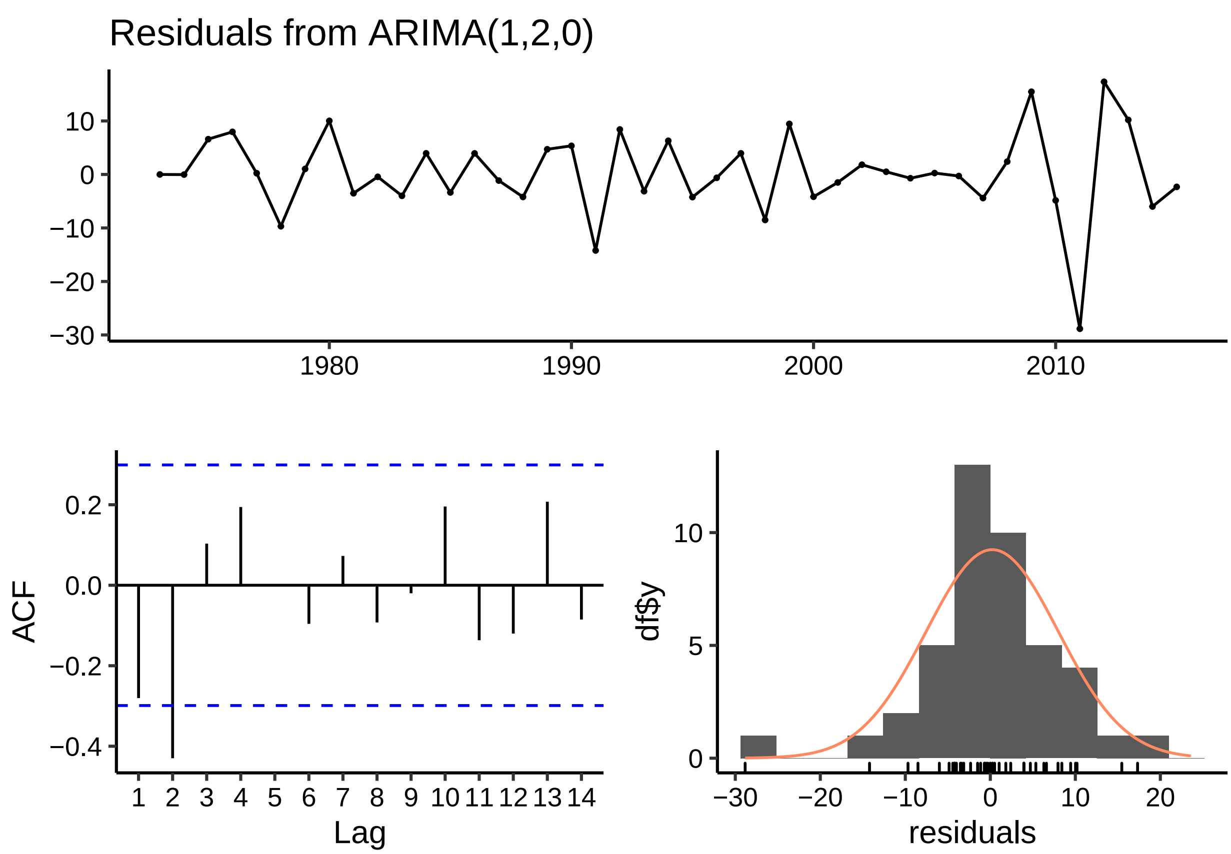
```
## Number of differencing needed:  2
```

Therefore, the ARIMA model for the electricity generation from hydropower is ARIMA(1,2,0)

```r
# Differencing the series to remove the trend.
diff_ts_ele_pro_hydro <- diff(ts_ele_pro_hydro, differences=2,lag=1)
# Fitting an arima model to the differenced series
arima_model_hydro <- Arima(diff_ts_ele_pro_hydro, order = c(1,2,0),
include.mean = FALSE, include.drift = FALSE)
print(arima_model_hydro)
```

```
## Series: diff_ts_ele_pro_hydro
## ARIMA(1,2,0)
##
## Coefficients:
##           ar1
##       -0.6094
## s.e.   0.1207
##
## sigma^2 = 63.74:  log likelihood = -143.08
## AIC=290.16   AICc=290.47   BIC=293.58
```

```r
# Plotting the residual series of the ARIMA fit
checkresiduals(arima_model_hydro, lag=12)
```




```
##
##  Ljung-Box test
##
```

```
## data:  Residuals from ARIMA(1,2,0)
## Q* = 20.258, df = 11, p-value = 0.04193
##
## Model df: 1.   Total lags used: 12
```

The residual series looks like a white noise series with no trend. As shown in the histogram also, the residuals follow almost a normal distribution, and in the ACF plot as well there is no trend that can be traced.

5.1.2.Electricity production from other renewable sources

In addition to the hydropower,the renewable sources of electricty production that are considered for this project are mainly solar photvoaltic system, wind and geothermal energy sources. Given the high potential of the country in solar and geothermal energy sources, looking at the trend of electricity generation from these sources is important to be able to observe the renewable energy resource utilization potential of the country.

```
# Plot of electricity production from renewable sources other than hydro
ggplot(ele_pro_source, aes(x=year, y=ele_pro_ren))+
  geom_line(color="green")+
  xlab("Year")+
  ylab("Percent of electricity production")
```

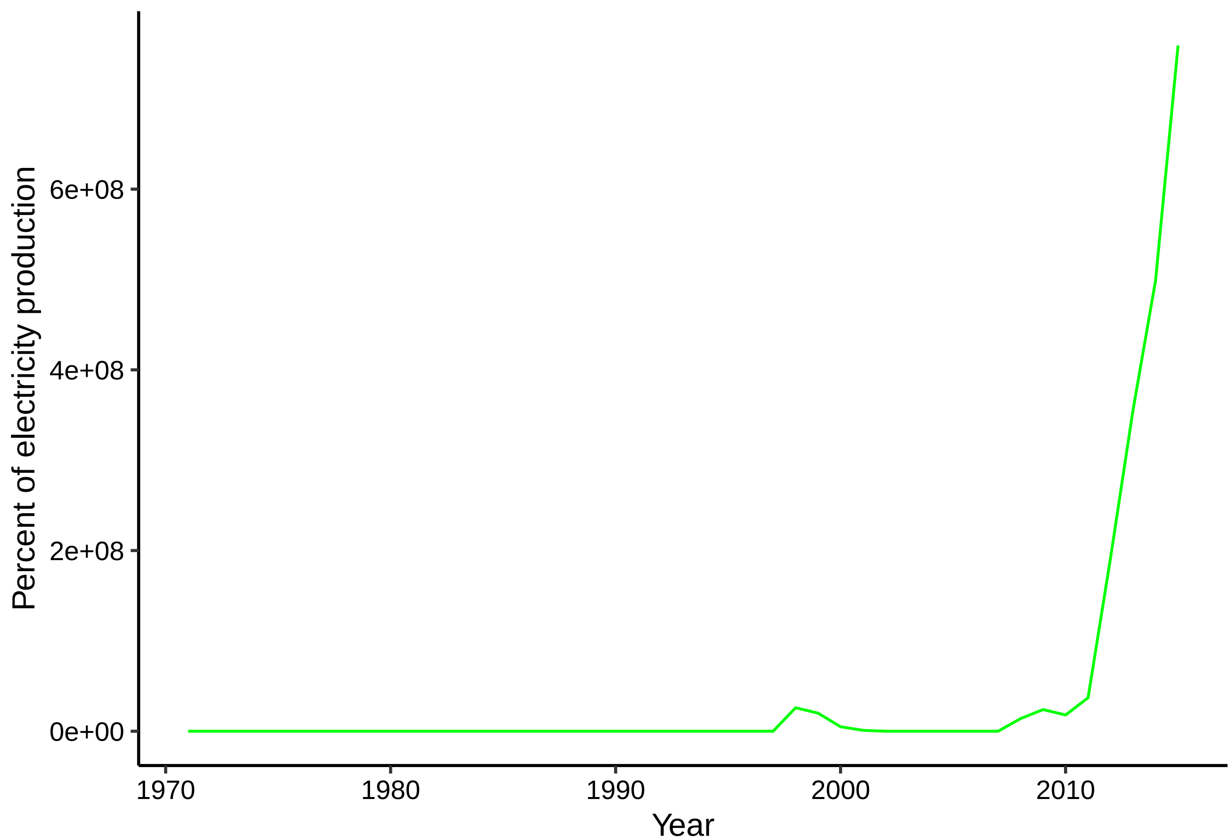
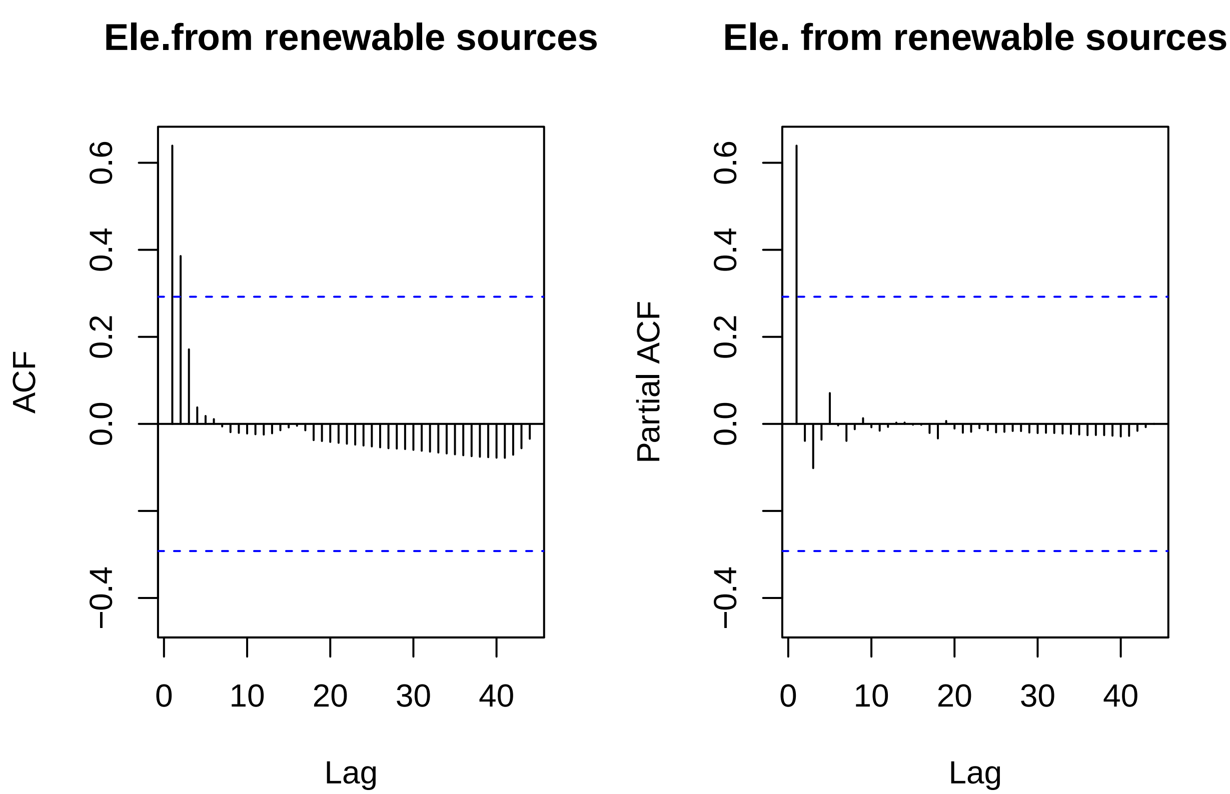


Figure 2: Trend of electricity production from other renewable sources excluding hydropower in Ethiopia from 1971-2015

The plot show, electricity production from other renewable sources other than hydropower started in early 2000, but remain in its initiation stage until the begining of 2010. When the hydro dominated systems have been severely affected by drought, the Government of Ethiopia started to take action to diversify the generation mix with other sources such as solar, wind, and geothermal that will result in a more climate-resilient power system (US ITA, 2022). As a result, in 2010 and onwards following the intorudction of government intiatives and incentives for the prmotion of clean energy soultions in the country, a significantly dramatic increase of electricity production from renewable energy sources observed. In addition, as it can be seen from the plot, the electricity production from renewable sources other than hydro has no seasonal pattern or component in it.

```
# Transforming the data into a time series object
ts_ele_pro_ren <- ts(ele_pro_source$ele_pro_ren, start = c(1971,1), frequency = 1)
# Plotting the Acf and Pacf plots
par(mfrow=c(1,2))
Acf(ts_ele_pro_ren, lag.max = 60, main="Ele.from renewable sources")
Pacf(ts_ele_pro_ren, lag.max = 60, main="Ele. from renewable sources")
```



The ACF and PACF plots of electricity production from renewable sources other than hydropower show an AR process with order p=1, q=0

Running stationarity tests

```
# Conducting Mann Kendall and ADF tests
MKtest_ele_ren <- MannKendall(ts_ele_pro_ren)
print(MKtest_ele_ren)
```

```
## tau = 0.577, 2-sided pvalue =9.5367e-07
```

The p-value is less than the significance level, thus, we reject the null hypothesis and say the series has a trend. We conduct the ADF test and check whether the trend is stochastic or not.

```
ADFtest_ele_ren <- adf.test(ts_ele_pro_ren)
print(ADFtest_ele_ren)
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  ts_ele_pro_ren
## Dickey-Fuller = 1.731, Lag order = 3, p-value = 0.99
## alternative hypothesis: stationary
```

The p-value is greater than the significance level, thus, we fail to reject the null hypothesis thus the series has a unit root. And conclude that the series has a stochastic trend. Therefore, we use differencing to achieve stationarity.

Model fitting - electricity production from other renewable sources other than hydropower

```
# Since the series has a stochastic trend it needs to be differenced
# to achieve stationarity.
# Checking for the number of differences needed
n_diff_ele_ren <- ndiffs(ts_ele_pro_ren, alpha = 0.1, test = c("kpss", "adf", "pp"),
max.d = 2)
cat("Number of differencing needed: ",n_diff_ele_ren)
```

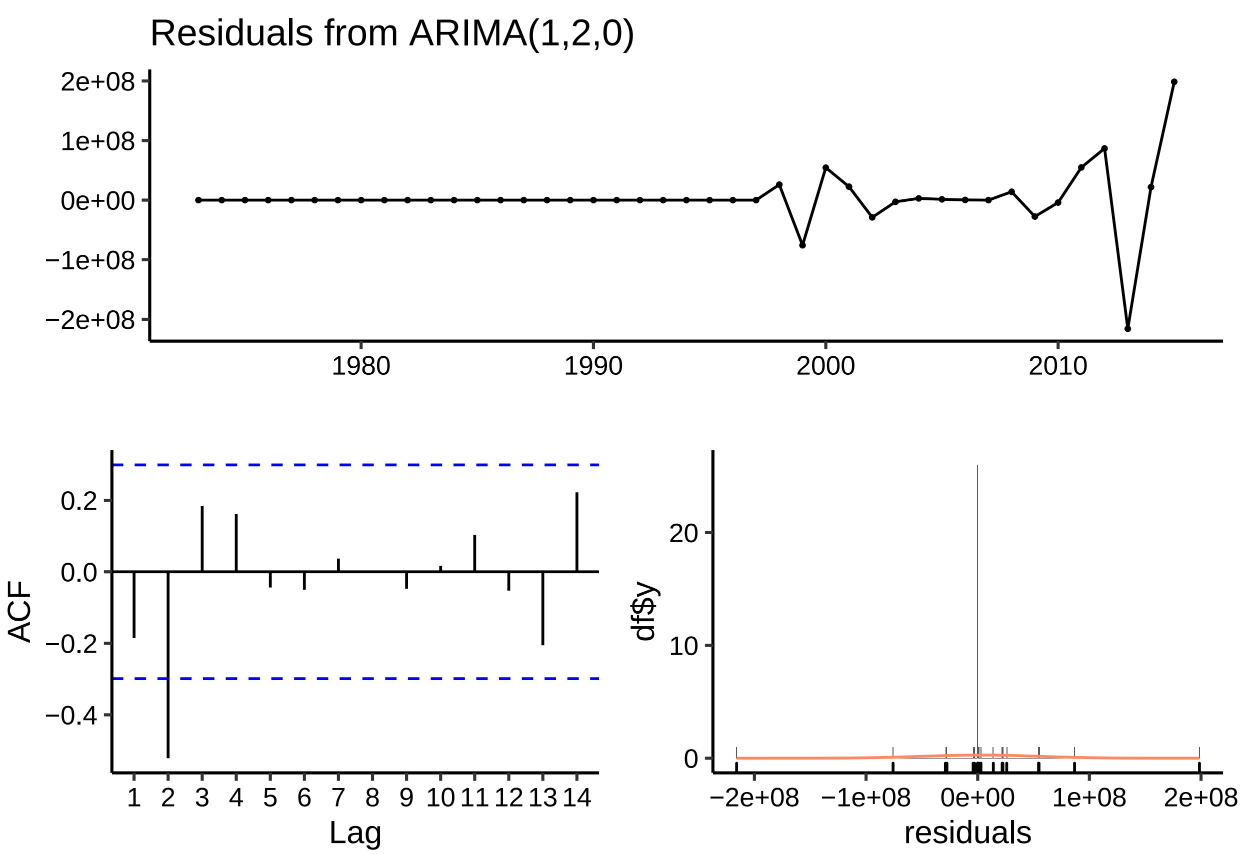
```
## Number of differencing needed:  2
```

Therefore, the ARIMA model for the electricity generation from other renewable sources excluding hydropower is ARIMA(1,2,0)

```
# Differencing the series to remove the trend.
diff_ts_ele_pro_ren <- diff(ts_ele_pro_ren, differences=2,lag=1)
# Fitting an arima model to the differenced series
arima_model_ren <- Arima(diff_ts_ele_pro_ren, order = c(1,2,0),
include.mean = FALSE, include.drift = FALSE)
print(arima_model_ren)
```

```
## Series: diff_ts_ele_pro_ren
## ARIMA(1,2,0)
##
## Coefficients:
##           ar1
##       -0.3155
## s.e.   0.1683
##
## sigma^2 = 2.72e+15:  log likelihood = -786.28
## AIC=1576.56   AICc=1576.88   BIC=1579.99
```

```
# Plotting the residual series of the ARIMA fit
checkresiduals(arima_model_ren, lag=12)
```




```
##
##  Ljung-Box test
##
```

```
## data:  Residuals from ARIMA(1,2,0)
## Q* = 18.598, df = 11, p-value = 0.06869
##
## Model df: 1.   Total lags used: 12
```

The residual series looks like a white noise series with no trend. As shown in the ACF plot as well there is no trend that can be seen.

5.1.3. Electricity production from oil sources

The other major source of electricity that the country depend on is oil sources. The electricty production from oil sources mainly relay on diesel fuel which is the main refined oil product in the country that is used for thermal power plants, and for private and public diesel generators in parts of the country where electrical power from the national grid can not be accessed.

```
# plotting the trend of electricity production from oil over time
ggplot(ele_pro_source, aes(x=year, y=ele_pro_oil))+
  geom_line(color="red")+
  xlab("Year")+
  ylab("Percent of electricity production")
```

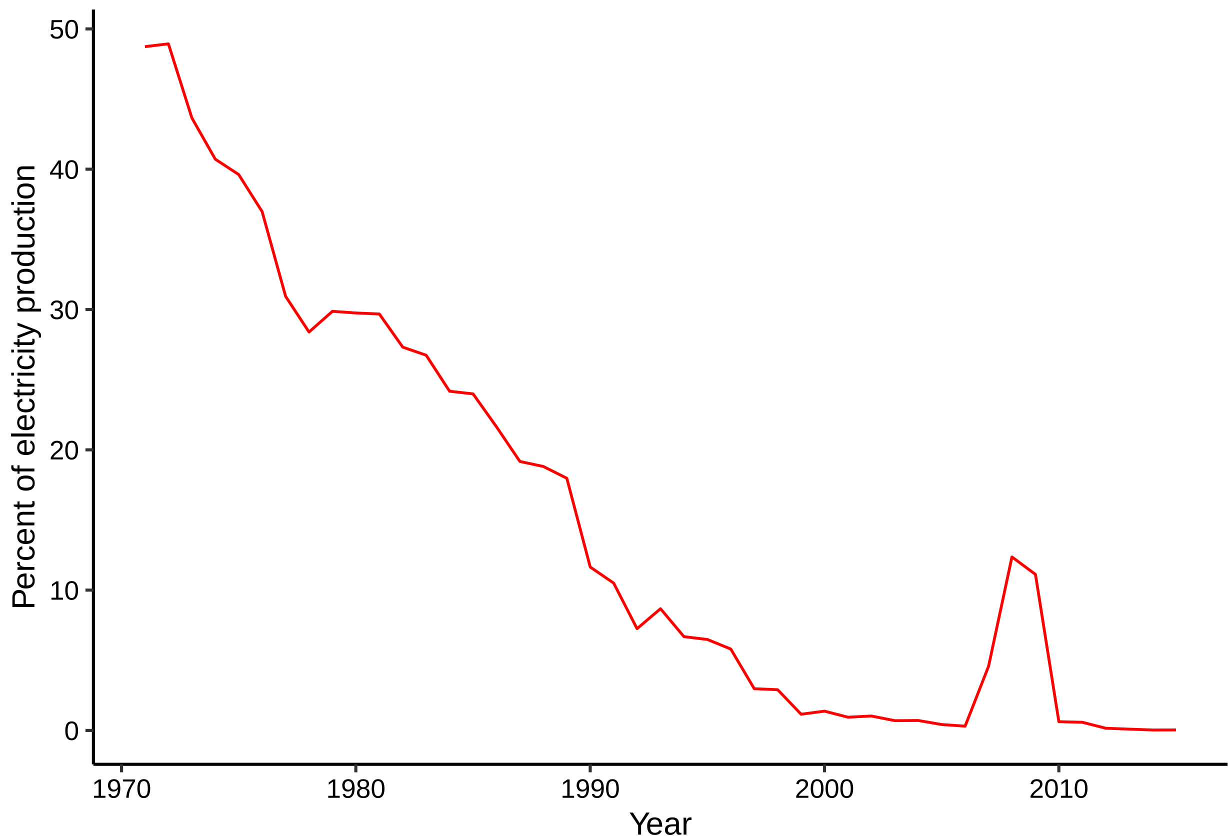
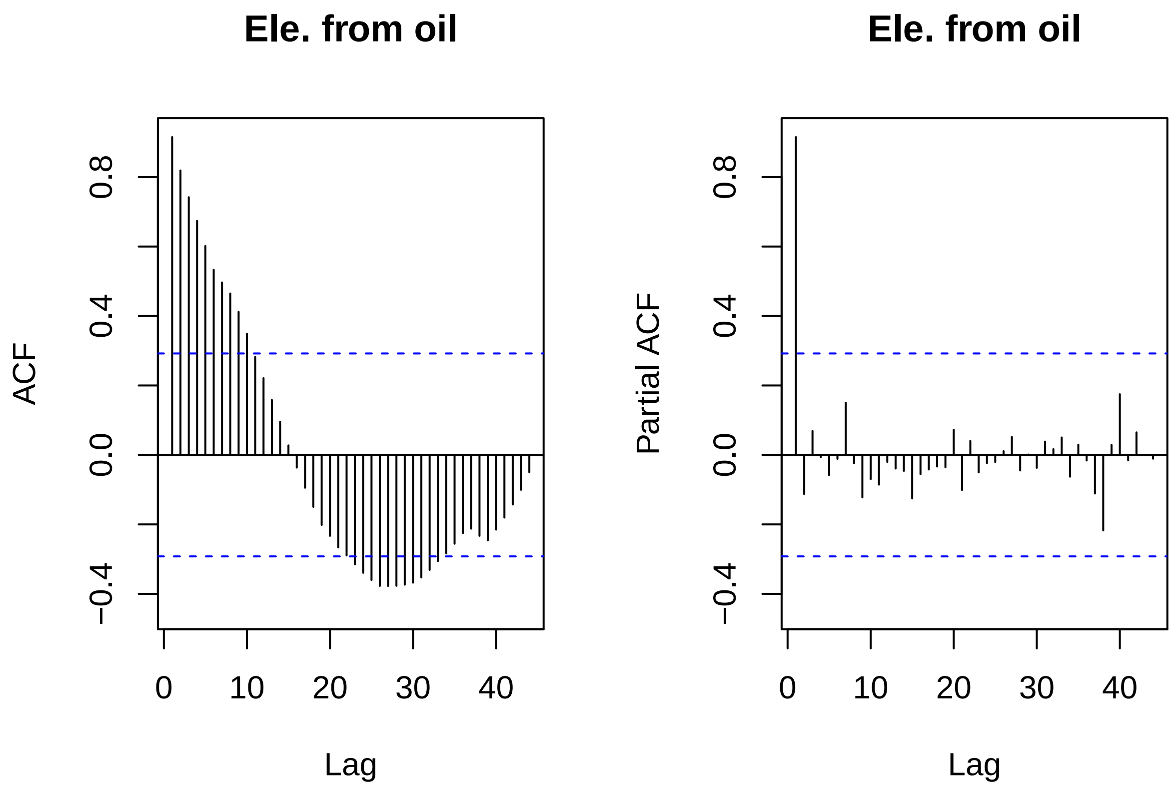


Figure 3: Trend of electricity production from oil in Ethiopia from 1971-2015

The plot shows an overall significant drop of percent of electricity production from oil sources over time. This indicate that the country is decreasing its dependency on oil for the production of electricity over time. As shown in the plot, there was a sudden increase in the percent of electricity production from oil sources that shifted the decreasing trend for sometime in 2008 until 2010 where it came back to a decreasing trend. As described above, due to the sever drought events in those years, following the reduction in the generation capacity of the hydropower plants, the governemnt took a measure to increase the use of oil as a source of electricity in an effrot to continually fulfill the electricity demand in the country.

```
# Transforming the data into a time series object
ts_ele_pro_oil <- ts(ele_pro_source$ele_pro_oil, start = c(1971,1), frequency = 1)
# Plotting the Acf and Pacf plots
par(mfrow=c(1,2))
Acf(ts_ele_pro_oil, lag.max = 60, main="Ele. from oil")
Pacf(ts_ele_pro_oil, lag.max = 60, main="Ele. from oil")
```



The ACF and PACF plots of electricity production from oil sources also show an AR process with order p=1, q=0

Running stationarity tests

```
# Conducting Mann Kendall and ADF tests

MKtest_ele_oil <- MannKendall(ts_ele_pro_oil)
print(MKtest_ele_oil)
```

```
## tau = -0.883, 2-sided pvalue =< 2.22e-16
```

The p-value is less than the significance level, thus, we reject the null hypothesis and say the series has a trend. We conduct the ADF test and check whether the trend is stochastic or not.

```
ADFtest_ele_oil <- adf.test(ts_ele_pro_oil)
print(ADFtest_ele_oil)
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  ts_ele_pro_oil
## Dickey-Fuller = -1.3983, Lag order = 3, p-value = 0.8126
## alternative hypothesis: stationary
```

The p-value is greater than the significance level, thus, we fail to reject the null hypothesis thus the series has a unit root. And conclude that the series has a stochastic trend. Therefore, we use differencing to achieve stationarity.

Model fitting - electricity production from oil sources

```
# Since the series has a stochastic trend it needs to be differenced
# to achieve stationarity.
# Checking for the number of differences needed
n_diff_ele_oil <- ndiffs(ts_ele_pro_oil, alpha = 0.1, test = c("kpss", "adf", "pp"),
max.d = 2)
cat("Number of differencing needed: ",n_diff_ele_oil)
```

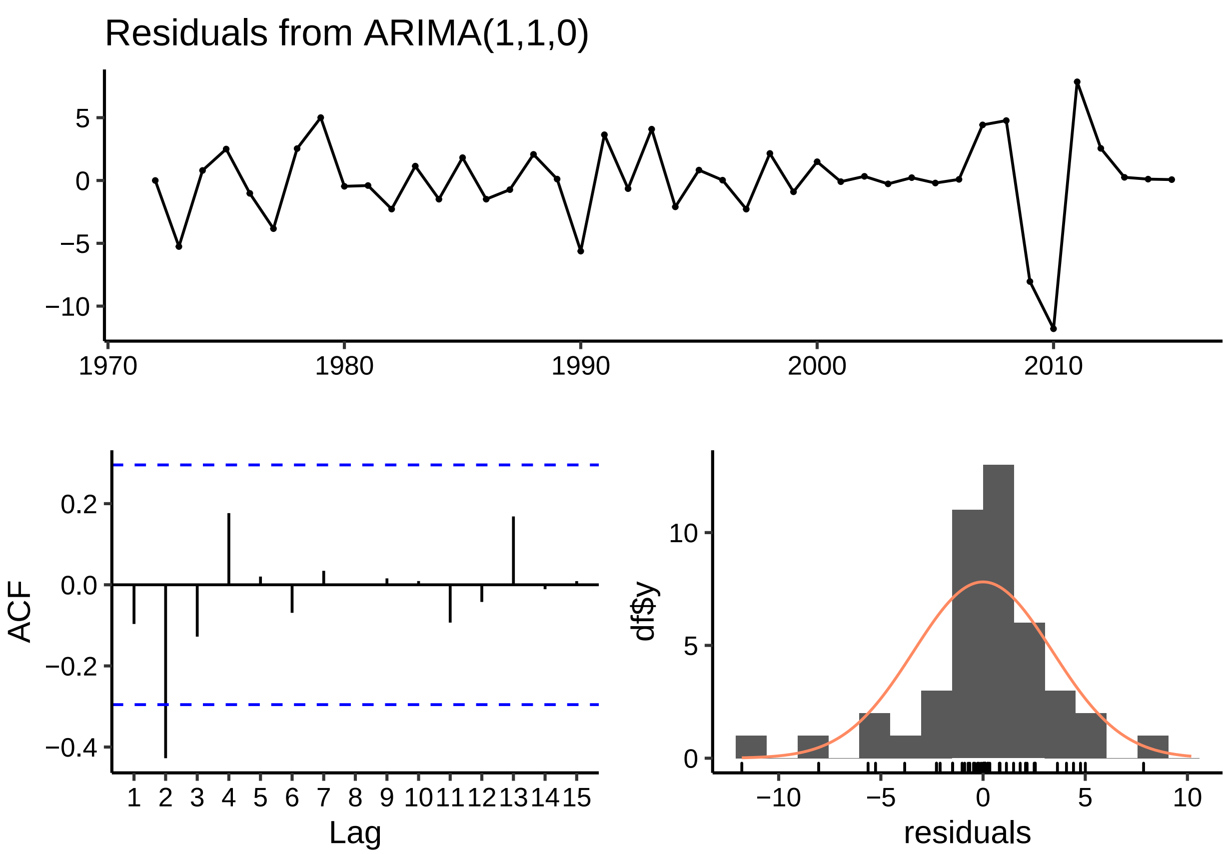
```
## Number of differencing needed:  1
```

Therefore, the ARIMA model for the electricity generation from oil sources is ARIMA(1,1,0)

```r
# Differencing the series to remove the trend.
diff_ts_ele_pro_oil <- diff(ts_ele_pro_oil, differences=1,lag=1)
# Fitting an arima model to the differenced series
arima_model_oil <- Arima(diff_ts_ele_pro_oil, order = c(1,1,0),
include.mean = FALSE, include.drift = FALSE)
print(arima_model_oil)
```

```
## Series: diff_ts_ele_pro_oil
## ARIMA(1,1,0)
##
## Coefficients:
##           ar1
##       -0.2816
## s.e.   0.1486
##
## sigma^2 = 11.81:  log likelihood = -113.64
## AIC=231.28   AICc=231.58   BIC=234.8
```

```r
# Plotting the residual series of the ARIMA fit
checkresiduals(arima_model_oil, lag=12)
```




```
##
##  Ljung-Box test
##
```

```
## data:  Residuals from ARIMA(1,1,0)
## Q* = 12.64, df = 11, p-value = 0.3175
##
## Model df: 1.   Total lags used: 12
```

The residual series looks like a white noise series with no trend. As shown in the ACF plot as well there is no trend that can be seen.

5.2. Summary of trend of electricity production and model fitting by major sources

```r
# Plotting the three major sources together
ts_ele_pro_source <- ts(ele_pro_source[,2:4], start = c(1971,1), frequency = 1)
plot(ts_ele_pro_source, main="Trends of electricity production by source")
```

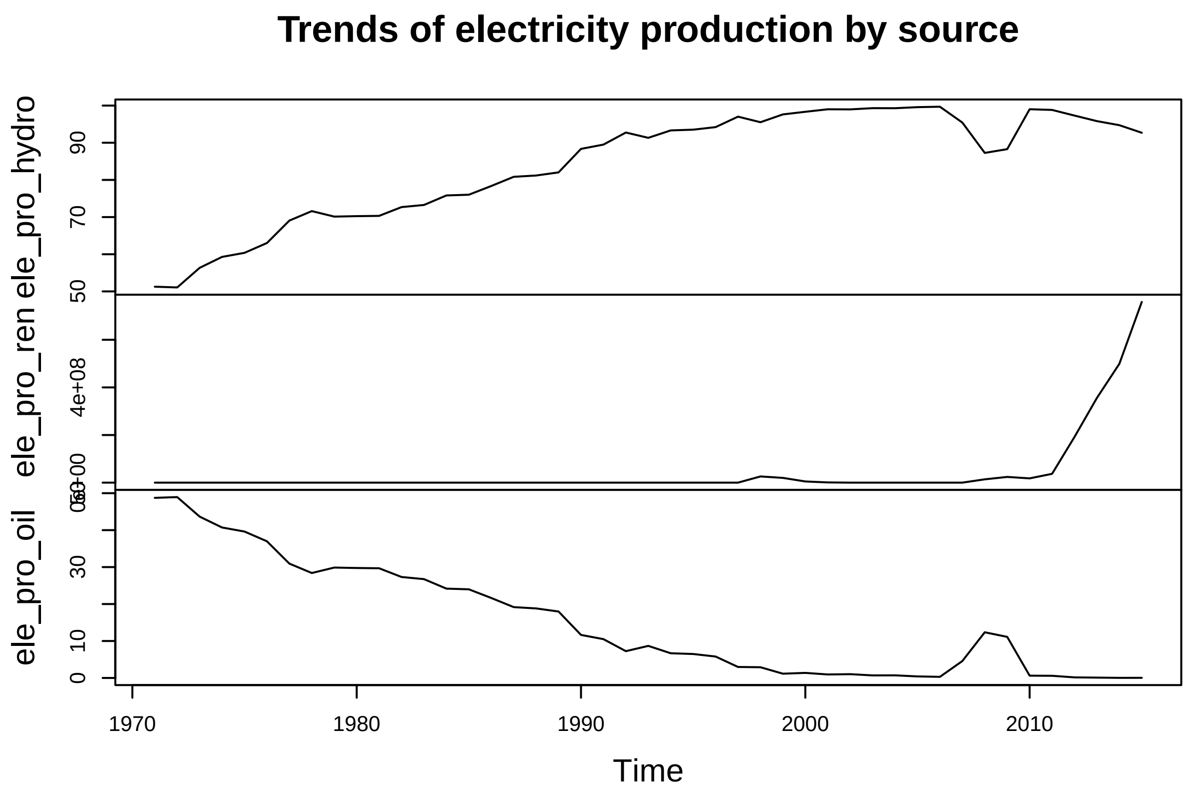


Figure 4: Trends of electricty production from three major sources (1971-2014)

```r
ggplot(ele_pro_source, aes(x=year))+
  geom_line(aes(y= ele_pro_hydro, color="ele_pro_hydro"))+
  geom_line(aes(y= ele_pro_ren, color="ele_pro_ren"))+
  geom_line(aes(y= ele_pro_oil, color= "ele_pro_oil"))+
  labs(color="")+
  scale_color_manual(values = c("ele_pro_hydro"="blue","ele_pro_ren"="green","ele_pro_oil"="black"),
  labels=c("Hydropower", "Other renewables", "Oil")) +
  ylim(c(0,100)) +
  xlab("Year") +
  ylab("% of electricty production")
```

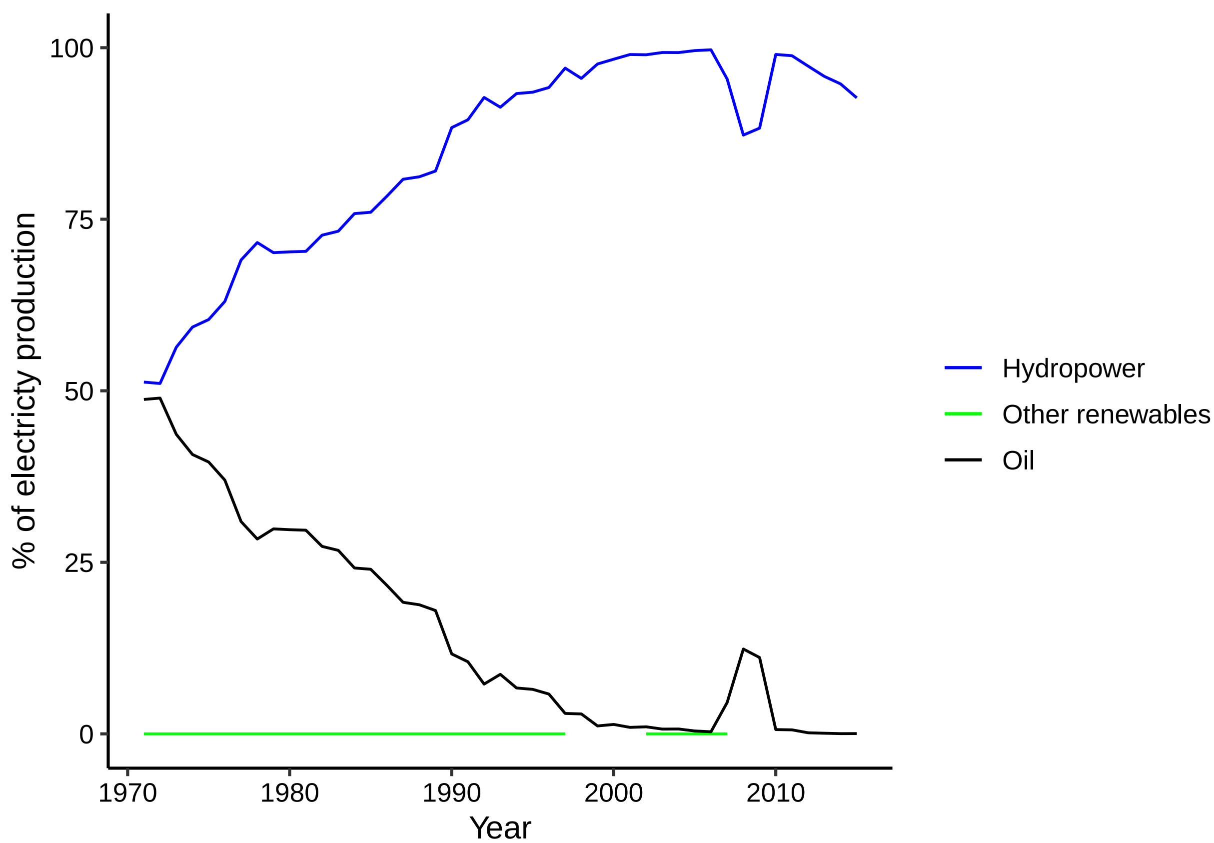


Figure 5: Trends of electricty production from three major sources (1971-2014)

The trends of perecent of electricity production from hydropower, other renewables and oil sources clealry indicate that there is an overall significant increase for production of electricity from hydropower and other renewable sources and an overall decrease for production of electricity from oil sources.

In the years 2008-2010, where there was sever drought events that significantly affected the electricity generation capacity of almost all hydropower dams, there was a significant decrease of perecent of electricity production from hydropower which then get back to its pre 2008 levels in 2010. In the same period where there were a significant dent, the generation of electricity from oil sources was significantly increased and stayed high for the same duration of a decrease in hydropower electricity production. And then following the increase in production from hydropower, the production of electricity from oil sources decreased signifcantly showing an inverse relationship between the production of electricity from hydropower and the production of electricity from oil sources.

This shows that the country is making an effort to transition to renewable energy sources over a period of time which aligns with the ultimate goals of the Climate Resilient Green Economy Strategy and the National Electrification Program of the Government. In addition, this is also a strong indication of the importance of energy generation mix in order to reduce the negative impacts of climate change in the supply of electricity.

In all the three sources of electricity production, there is a highly significant trend that can be observed but there is no seasonal pattern that can be seen.

Based on the trend analysis the models identified are ARIMA(1,2,0) for both electricity production from hydropower and electricity production from other renewable sources other than hydropower, and ARIMA (1,1,0) for percent of electricity production from oil sources.

6. Trend analysis of electricity access and consumption

6.1. Access to electiricty

The national total access to electricity stands at 48.27% in 2019 more than 95% of which is concentrated in urban and peri-urban areas of the country. Although those areas are better to have access, the frequent interruption of electricity services for various reasons makes the supply unreliabile. This problem of unreliability worsens as we move far away from urban areas. In almost all of rural areas, access to electricity is so limited that people heavily depend on forest resources, agricultural residues and animal wastes for their energy need. This contributes to the social, economic and environmental problems of the country.

The governemnt of Ethiopia has developed and launched a National Electrification Program in 2017 which aims to address the energy need of the nations by expanding its grid coverage as well as increasing its off-grid solutions to easily inaccessible areas by 2025.

```
# Extracting data for the access to electricity from 1971 to 2015
ele_access <- ETH_ele_pro_cons_clean %>%
  select(year, Total_Access) %>%
  filter(year>='2000-01-01' & year<='2019-01-01')
summary(ele_access)
```

```
##       year             Total_Access  
##  Min.   :2000-01-01   Min.   :12.70  
##  1st Qu.:2004-10-01   1st Qu.:26.79  
##  Median :2009-07-02   Median :29.88  
##  Mean   :2009-07-02   Mean   :31.18  
##  3rd Qu.:2014-04-02   3rd Qu.:35.68  
##  Max.   :2019-01-01   Max.   :48.27  
##                       NA's   :1
```

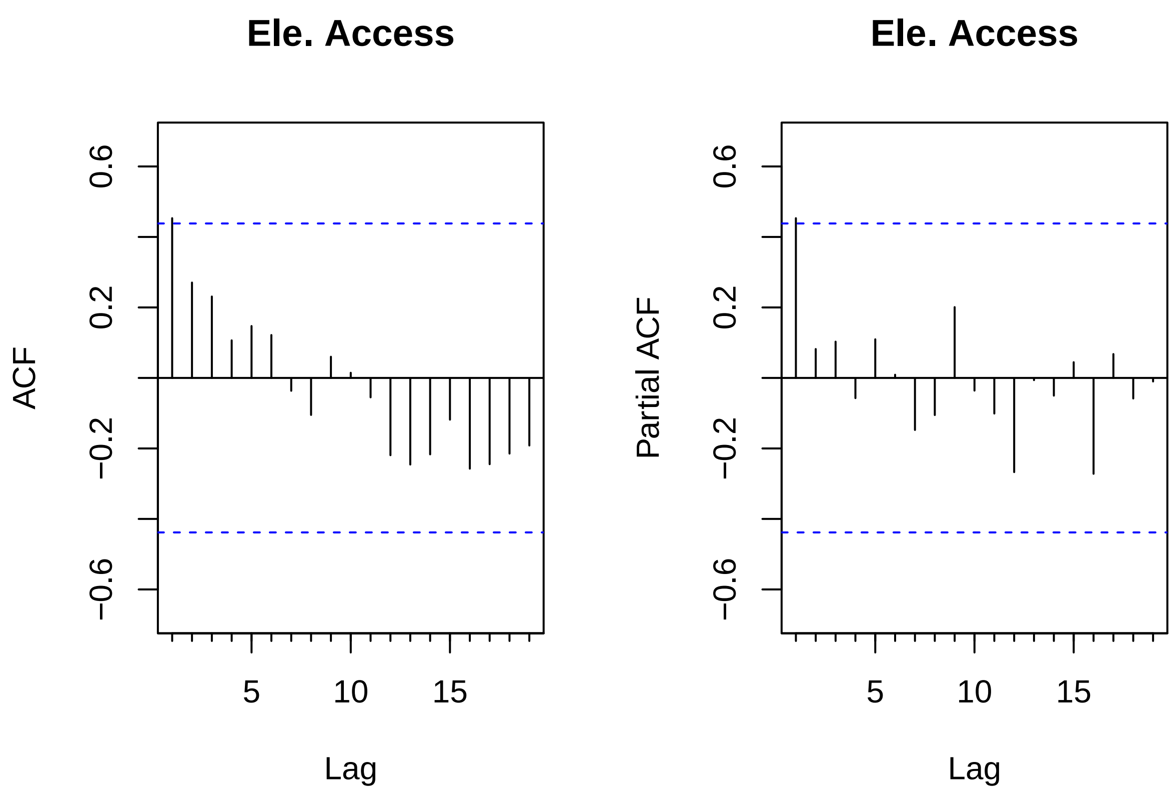
```
# Filling the missing value
Tot_access_full <- na.approx(ele_access$Total_Access)
Tot_access_full.df <- as.data.frame(Tot_access_full)
ele_access <- cbind(ele_access$year,Tot_access_full.df)
colnames(ele_access) <-c("year", "Total_Access")
```

```r
# Trend of access to electricity
ggplot(ele_access, aes(x=year, y=Total_Access))+
  geom_line(color="green")+
  xlab("Year")+
  ylab("% of population")
```

Figure 6: Trend of access to electricity in Ethiopia from 2000-2019

The plot clearly indicates that there is a significantly increasing trend of perecent of population who have access to electricity over time. Although the total national level access rate is below 50%, the increasing trend is attributed to a continued increase in the generation of electricity and the increase in the economic growth of the country in the same period (WB data). The decrease in perecent of the population who have access to electricity in some years that is observed in the plot is mainly the result of the decrease in electricity generation following lower water levels of hydropower dams as a result of drought events. The plot also shows that there is no seasonal pattern in the percentage of population who have access to electricity overtime.

```r
# Transforming the data into a time series object for access to electricity
ts_ele_access <- ts(ele_access$Total_Access, start = c(2000,1), frequency = 1)
# Plotting the Acf and Pacf plots
par(mfrow=c(1,2))
Acf(ts_ele_access, lag.max = 60, main="Ele. Access")
Pacf(ts_ele_access, lag.max = 60, main="Ele. Access")
```



The time series plot also confirms that the time series for access to electricity has no seasonal component in it.

The ACF and PACF plots show that the access to electricity series is a white noise.

Running stationarity tests

```
# Conducting Mann Kendall and ADF tests

MKtest_ele_access <- MannKendall(ts_ele_access)
print(MKtest_ele_access)
```

```
## tau = 0.695, 2-sided pvalue =2.1338e-05
```

The p-value is less than the significance level, thus, we reject the null hypothesis and say the series has a trend. We conduct the ADF test and check whether the trend is stochastic or not.

```
ADFtest_ele_access <- adf.test(ts_ele_access)
print(ADFtest_ele_access)
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  ts_ele_access
## Dickey-Fuller = -1.9603, Lag order = 2, p-value = 0.5875
## alternative hypothesis: stationary
```

The p-value is greater than the significance level, thus, we fail to reject the null hypothesis thus the series has a unit root. And conclude that the series has a stochastic trend. Therefore, we use differencing to achieve stationarity.

Model fitting - access to electricity

```
# Since the series has a stochastic trend it needs to be differenced
# to achieve stationarity.
# Checking for the number of differences needed
n_diff_ele_access <- ndiffs(ts_ele_access, alpha = 0.1, test = c("kpss", "adf", "pp"),
max.d = 2)
cat("Number of differencing needed: ",n_diff_ele_access)
```

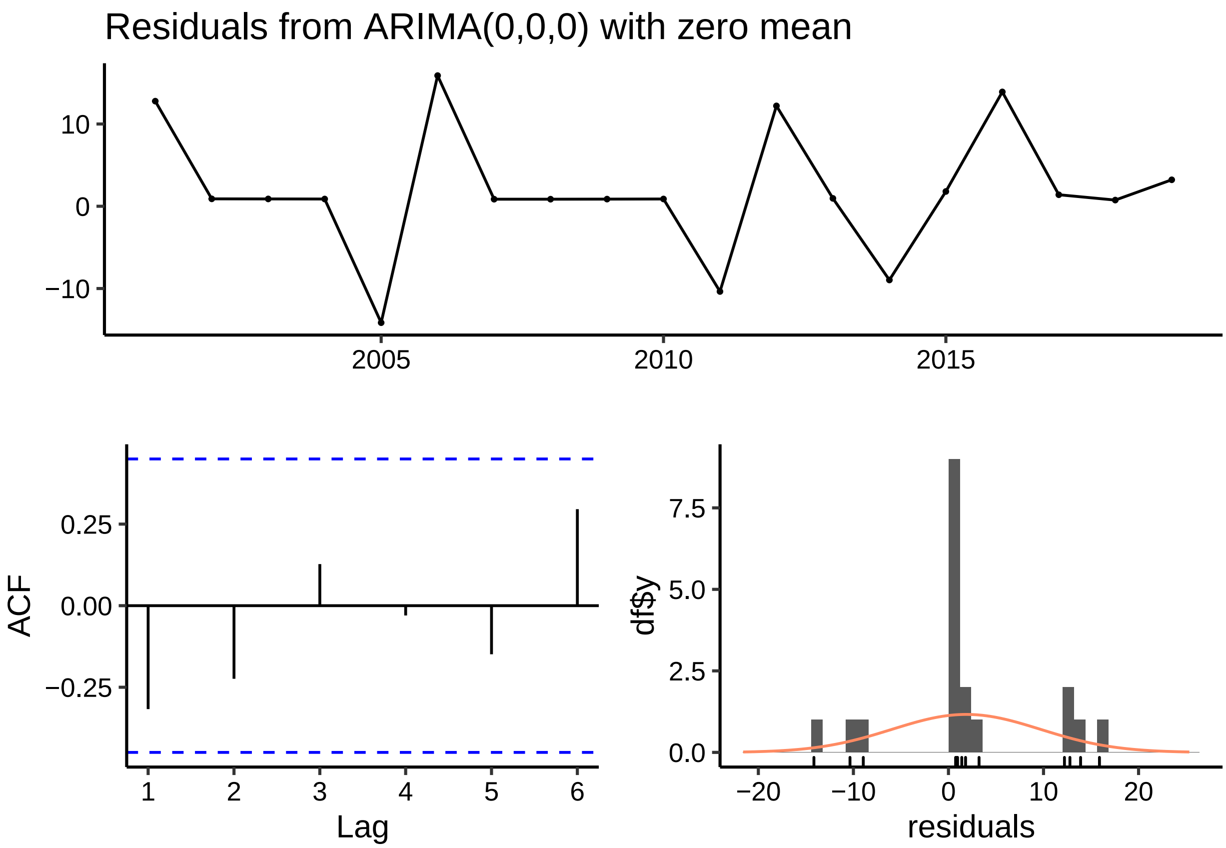
```
## Number of differencing needed:  1
```

```
# Differencing the series to remove the trend.
diff_ts_ele_access <- diff(ts_ele_access, differences=1,lag=1)
# Fitting an arima model to the differenced series
arima_model_access <- Arima(diff_ts_ele_access,
include.mean = FALSE, include.drift = FALSE)
print(arima_model_oil)
```

```
## Series: diff_ts_ele_pro_oil
## ARIMA(1,1,0)
##
```

```
## Coefficients:
##           ar1
##       -0.2816
## s.e.   0.1486
##
## sigma^2 = 11.81:  log likelihood = -113.64
## AIC=231.28   AICc=231.58   BIC=234.8
```

```
# Plotting the residual series of the ARIMA fit
checkresiduals(arima_model_access, lag=12)
```




```
##
##  Ljung-Box test
##
## data:  Residuals from ARIMA(0,0,0) with zero mean
## Q* = 9.7135, df = 12, p-value = 0.6411
##
## Model df: 0.   Total lags used: 12
```

Checking the residual series also confirms that the series is a white noise series with no trend or seasonality. As shown in the histogram also, the residuals follow a normal distribution, and in the ACF plot as well there is no trend or seasonality that can be traced.

6.2 Per capita electric power consumption

Elecric power consumption refers to consumption of electricity produced from different sources by individual households, private and public businesses as well as organizations throughout the country. Of course, the level/ amount of consumption various significantly among those users, as well as by location, urban and peri-urban areas predominantly take the lion shre in consumption.

```
# Extracting data for electric power consumption from 1971 to 2014
ele_power_cons <- ETH_ele_pro_cons_clean %>%
  select(year, Power_Consumption) %>%
  filter(year>='1971-01-01' & year<='2014-01-01')

# Trend of per capita consumption of electricity
ggplot(ele_power_cons, aes(x=year, y=Power_Consumption))+
  geom_line(color="green")+
  xlab("Year")+
  ylab("Consumption in kwh")
```

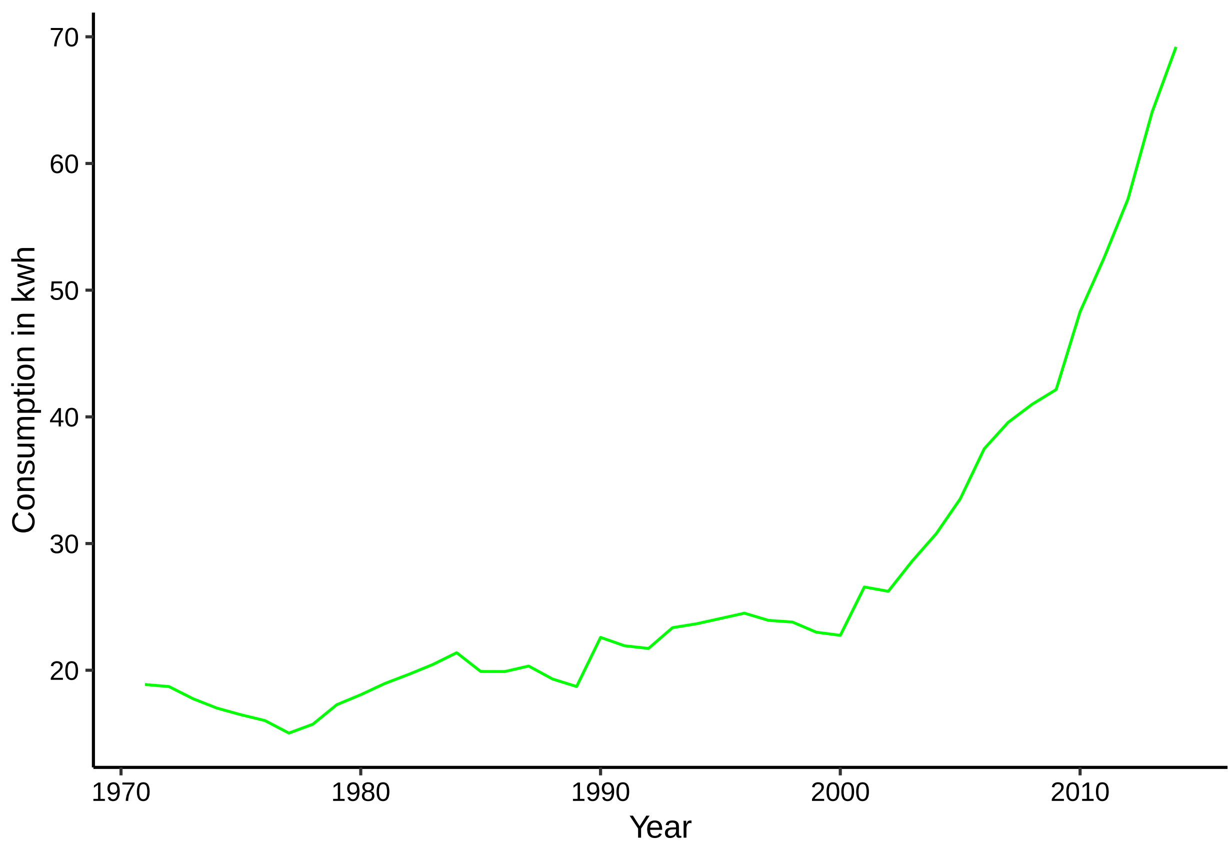
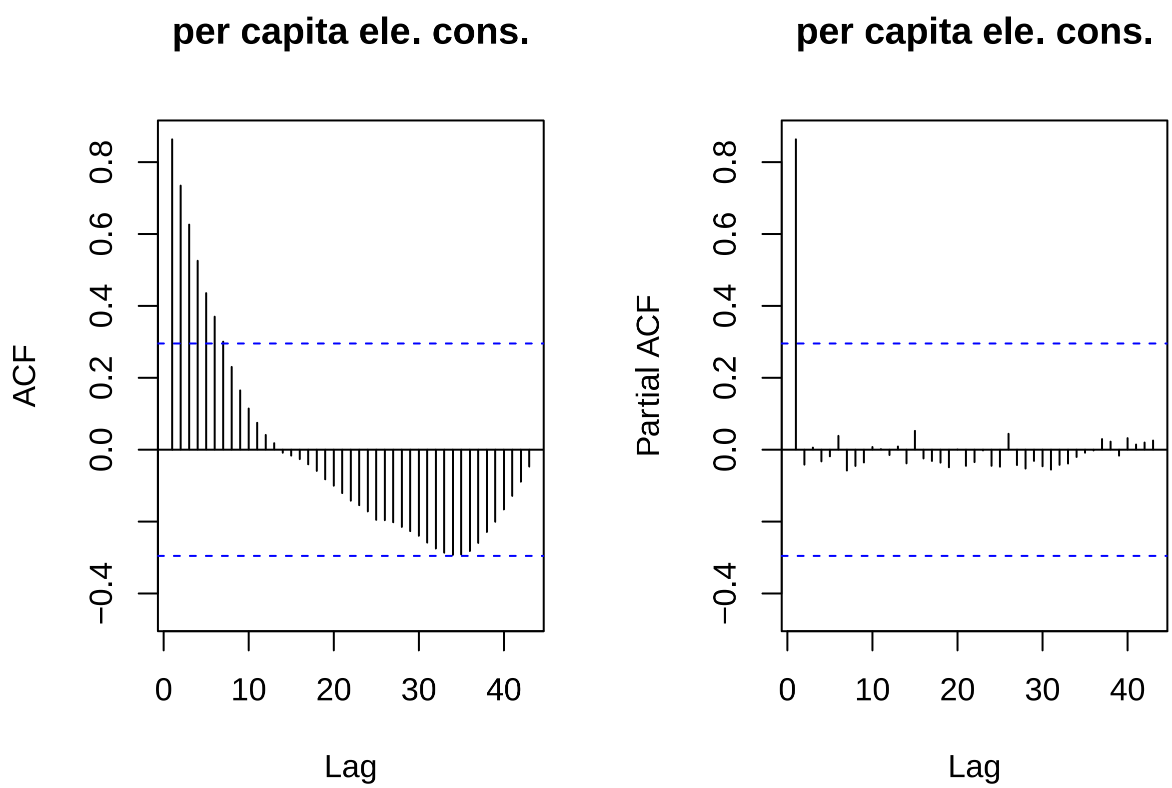


Figure 7: Trend of per capita electricity consumption (kwh) in Ethiopia from 1971-2014

The plot clearly indicate that there is an overall significantly increasing trend of per capita consumption of electricity over time. However, the per capita consumption of electricity was well below 30kwh until 2000. It is from 2000 onwards that the per capita consumption of electricity show a significant increasing trend reaching 70kwh in 2014. Similar to the access to electiricity, the increasing trend of electric power consumption is also directly related to the increase in the real annual GDP per capita growth as shown by the World Bank development indicators for annual GDP and GDP per capita for Ethiopia. From the plot it can be seen that there is no seasonal pattern in the per capita consumption of electricity.

```
# Transforming the data into a time series object
ts_ele_cons <- ts(ele_power_cons$Power_Consumption, start = c(1971,1), frequency = 1)
# Plotting the Acf and Pacf plots
par(mfrow=c(1,2))
Acf(ts_ele_cons, lag.max = 60, main="per capita ele. cons.")
Pacf(ts_ele_cons, lag.max = 60, main="per capita ele. cons.")
```



The ACF and PACF plots also confirms that the time series for per capita electricity consumption has no seasonal component in it. These plots also show that the series follow an AR process with order p=1, q=0

Running stationarity tests

```
# Conducting Mann Kendall and ADF tests
MKtest_ele_cons <- MannKendall(ts_ele_cons)
print(MKtest_ele_cons)
```

```
## tau = 0.839, 2-sided pvalue =< 2.22e-16
```

The p-value is less than the significance level, thus, we reject the null hypothesis and say the series has a trend. We conduct the ADF test and check whether the trend is stochastic or not.

```
ADFtest_ele_cons <- adf.test(ts_ele_cons)
```

```
## Warning in adf.test(ts_ele_cons): p-value greater than printed p-value
```

```
print(ADFtest_ele_cons)
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  ts_ele_cons
## Dickey-Fuller = 1.8572, Lag order = 3, p-value = 0.99
## alternative hypothesis: stationary
```

The p-value is greater than the significance level, thus, we fail to reject the null hypothesis thus the series has a unit root. And conclude that the series has a stochastic trend. Therefore, we use differencing to achieve stationarity.

Model fitting - per capita electricity consumption

```
# Since the series has a stochastic trend it needs to be differenced
# to achieve stationarity.
# Checking for the number of differences needed
n_diff_ele_cons <- ndiffs(ts_ele_cons, alpha = 0.1, test = c("kpss", "adf", "pp"),
max.d = 2)
cat("Number of differencing needed: ",n_diff_ele_cons)
```

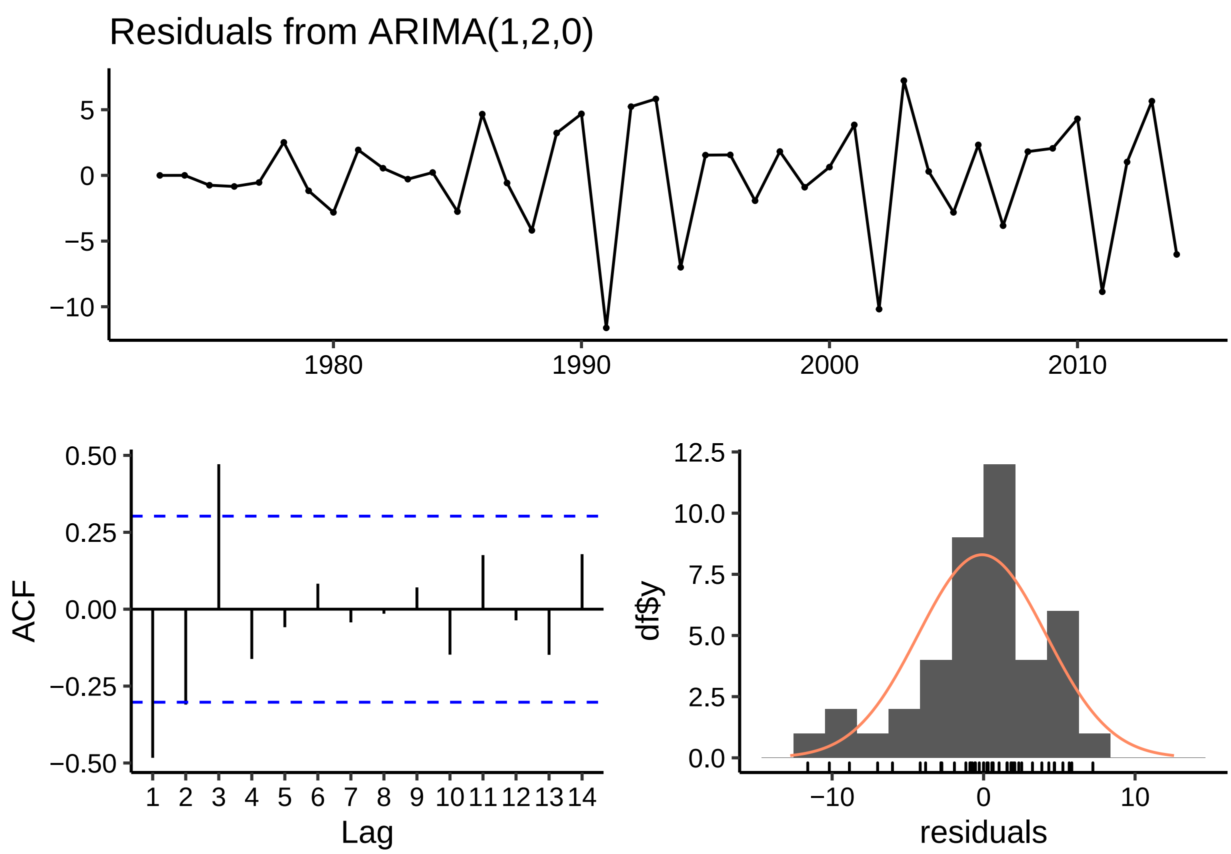
```
## Number of differencing needed:  2
```

```
# Differencing the series to remove the trend.
diff_ts_ele_cons <- diff(ts_ele_cons, differences=2,lag=1)
```

```
# Fitting an arima model to the differenced series
arima_model_cons <- Arima(diff_ts_ele_cons, order = c(1,2,0),
include.mean = FALSE, include.drift = FALSE)
print(arima_model_cons)
```

```
## Series: diff_ts_ele_cons
## ARIMA(1,2,0)
##
## Coefficients:
##           ar1
##       -0.6728
## s.e.   0.1136
##
## sigma^2 = 18.75:  log likelihood = -115.18
## AIC=234.36   AICc=234.68   BIC=237.74
```

```
# Plotting the residual series of the ARIMA fit
checkresiduals(arima_model_cons, lag=12)
```




```
##
##  Ljung-Box test
##
## data:  Residuals from ARIMA(1,2,0)
## Q* = 30.852, df = 11, p-value = 0.001163
```

```
##
## Model df: 1.   Total lags used: 12
```

The residual series looks like a white noise series with no trend. This can also be seen from the ACF plot where there is not trend or seasonality that can be traced and from the histogram where the residuals follow a normal distribution.

6.3. Summary of trend analysis and model fitting of access to electricity and per capita electricity consumption

Both the percentage of population who have access to electricity and the per capita consumption of electricity show an increasing trend overtime. In both cases the increase was highly significant starting from 2010 which are attributed by the introduction of an intensive Growth and Transformation Plan in 2005 with the aim of increasing the economic activity of the country in many dimensions. In both cases, there is no seasonal pattern that can be traced.

The trend analysis of the percent of population who have access to to electricity indicate that the series is a white noise, and thus, there is no model identified for it. Whereas, the trend analysis of the per capita consumption of electricity indicate that the model identified is ARIMA(1,2,0).

7. The way forward

Continue to gather a relatively complete historical data for Electricity generation from different sources, access to electricity and the Blue Nile river flow to continue to investigate the trend of electricity generation, access and consumption, and the pros and cons of using Blue Nile river for the generation of electricity. Moreover, gathering datasets and additional information on precipitation and temperature to look at their impacts on electricity generation along with the flow of the Blue Nile river.

Modelling the trends of electricity generation and consumption and forecasting how they will look like in the coming future

Generating scenarios taking into consideration of the rainfall, temperature, drought events and electricity generation patterns around the Blue Nile river, and investigating the impacts of GERD to the downstream countries.

8. References

- Ministry of Finance (MoF). https://www.mofed.gov.et/programmes-projects/crge-facility/. accessed in April 2022.

- Ministry of Water, Irrigation and Energy (MoWIE), 2019. Lighting to All:National Electrification Program 2.0. Integrated Planning for Universal Access. Addis Ababa, Ethiopia.

- National Planning Commission, 2016.Growth and Transformation Plan II (GTP II)(2015/16-2019/20).Volume I: Main Text.Addis Ababa, Ethiopia

- US International Trade Administration (US ITA, 2022). https://www.trade.gov/country-commercial-guides/ethiopia-energy#:~:text=Ethiopia%20has%20abundant%20renewable%20energy,wind%2C%20solar%20and%20 in April 2022.

- World Bank, World Development Indicators. https://data.worldbank.org/indicator/EG.ELC.ACCS.ZS?locations=ET). access in April 2022.

- World Food Programme (WFP), 2022. An Analysis of the Impacts of Ongoing Drought across the Eastern Horn of Africa, 2020 - 2022 Period. Food Security Analysis.Research, Analysis and Monitoring (RAM) Unit World Food Programme, Regional Bureau Nairobi.