STAT 210 Applied Statistics and Data Analysis

COURSE PROJECT REPORT

Towards a cleaner future: an investigation of global electricity production

Group	# 6
Members:	Xiao Shao, Mojtaba AlShams, Sarah Mutahar
Instructor:	Joaquin Ortega Sanchez
Date:	December 2, 2021

Contents

Abstract

Several aspects of global electricity generation are investigated. A simple linear regression model is proposed to reveal the relationship between electricity access/consumption and gross domestic product (GDP) per catipa. Rich countries tend to consume more electricity. Analysis of variance (ANOVA) is done for the low-carbon energy source shares among different groups of country according to GDP per capita. No significant difference is found, meaning rich countries do not necessarily have a higher share of low-carbon energy source for electricity production. Annual electricity production from low-carbon source in Europe, United States, and China is compared with total electricity consumption of electric vehicles (EVs) from these regions. It is found that the growth of electricity consumption of EVs outpaces that of electricity generation from low-carbon source.

1. Introduction

The availability of energy has changed the course of civilization during the previous few centuries since the Industrial Revolution. Not only have new energy sources been discovered – first fossil fuels, then nuclear, and now various renewable technologies such as hydropower, solar, wind – but also the amount of energy we can generate and use. Human energy consumption using natural resources consists mainly of three parts: electricity, heating and transportation. Therefore, electricity is only one dimension of energy consumption and this dimension will be in focus in the report.

In the first part of this project, how people's access to electricity and consumption vary across the world is studied. Gross domestic product (GDP) per capita will be used as the indication of the economic development level. The income level of a country can be estimated by GDP per capita as well since GDP is a standard measure of monetary value of all goods and services made in a country during a certain period. A positive correlation between the GDP per capita and electricity consumption per capita is expected.

In the second part, the relationship between the shares of low-carbon energy source and countries' economic development level is investigated. Energy source structure of electricity production varies across the world because of natural resources limitations, technology, capital, etc. Hence, the variance of the ratio of low-carbon electricity consumption to the overall (both low-carbon and high-carbon) electrical consumption is analyzed among different groups of countries divided by GDP per capita. Note that electricity is considered from low-carbon source if the generation process produces low greenhouse gas emissions compared to conventional fossil fuel power generation. Nuclear, wind, solar, hydropower, and other renewable sources like geothermal are all considered as low-carbon energy source.

The third part aims to evaluate if electric vehicles (EVs) are really crabon-free. Electrical vehicles gained tremendous momentum during the past few years. Many people believe that electric vehicles will help resolve the climate change crisis by reducing CO_2 emission to a great extent. The key problem is where does the electricity consumed by EVs come from. Annual electricity production from low-carbon source from 2010 to 2019 in Europe, United States, and China will be compared with total electricity consumption of EVs from these regions.

More than one data set will be used in this project. Some of the data sets describe people's behavior towards energy consumption and others describe the electricity generation sources themselves. Also, a data set that describes electric cars sales was used. The following table summarizes this data.

DATA/VARIABLE NAME	DATA/VARIABLE DESCRIPTION
Access To Electricity (% of Population)	This data set represents a percentage of the population with access to electricity. Electrification data are collected from industry, national surveys and international sources around 1990 to 2019. It is published by World Development Indicators - World Bank. [1]
GDP Per Capita, PPP (Constant 2017 International $)	This data set represents the GDP per capita from 1990 – 2020 based on purchasing power parity (PPP). PPP GDP is gross domestic product converted to international dollars using purchasing power parity rates. It is published by World Development Indicators - World Bank. [1]
Per capita electricity (KWH)	This data set represents the electricity consumption data for the following countries: Australia, United States, Japan, China, South Africa , World, United Kingdom, France, India, Brazil, Sweden from 1985 – 2020. The data is compiled from two resources BP Statistical Review of World Energy & Ember. [2]
Percentage of Electricity from Coal/Gas/Hydro/Solar/Wind/Oil/Nuclear/Other Renewables	This data set provides the percentage of energy from different resources (not just electricity) consumption data from 1985 – 2020. The data is compiled from two resources: BP Statistical Review of World Energy & Ember. [2]

Fossil Fuel/Nuclear/ Renewable Electricity Per Capita (KWH)	This data set represents the electricity per capita generated from different resources from 1985 – 2020. The data is compiled from two resources: BP Statistical Review of World Energy & Ember. [2]
Year	The year of each data used

2. Electricity consumption and GDP per capita

To understand how people's availability and consumption of electricity vary across the world, we use gross domestic product (GDP) per capita to indicate economic level and citizen income for different countries. Fig. 1 and Fig. 2 show the relationship between access to electricity, and electricity consumption per capita measured against average income (GDP per capita), respectively. Note that the *x*-axis in Fig. 1 is in log scale.

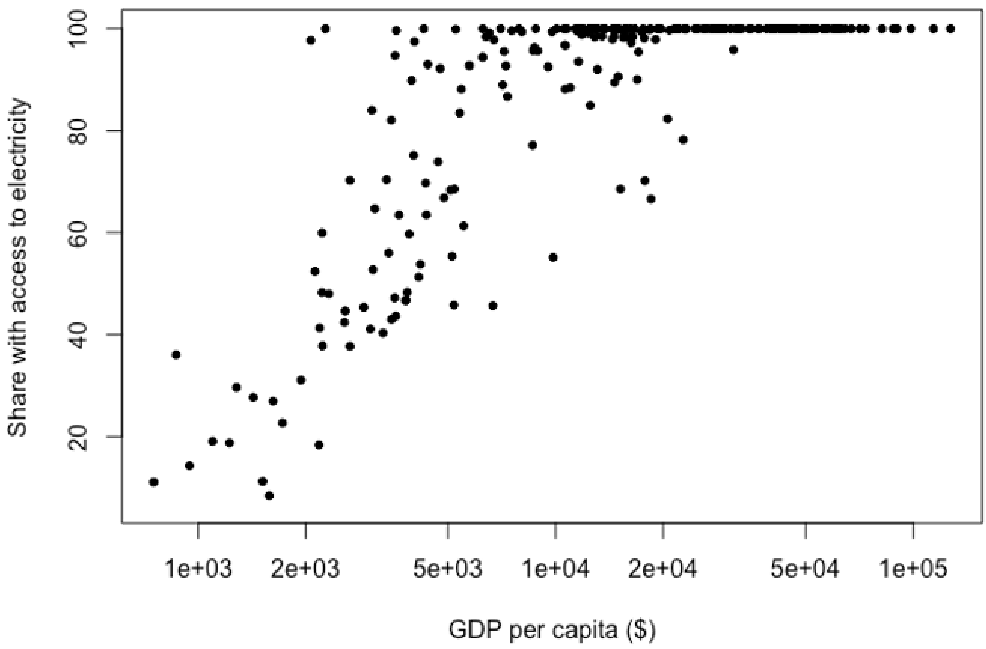

Fig. 1 Access to electricity vs. GDP per capita, 2019.

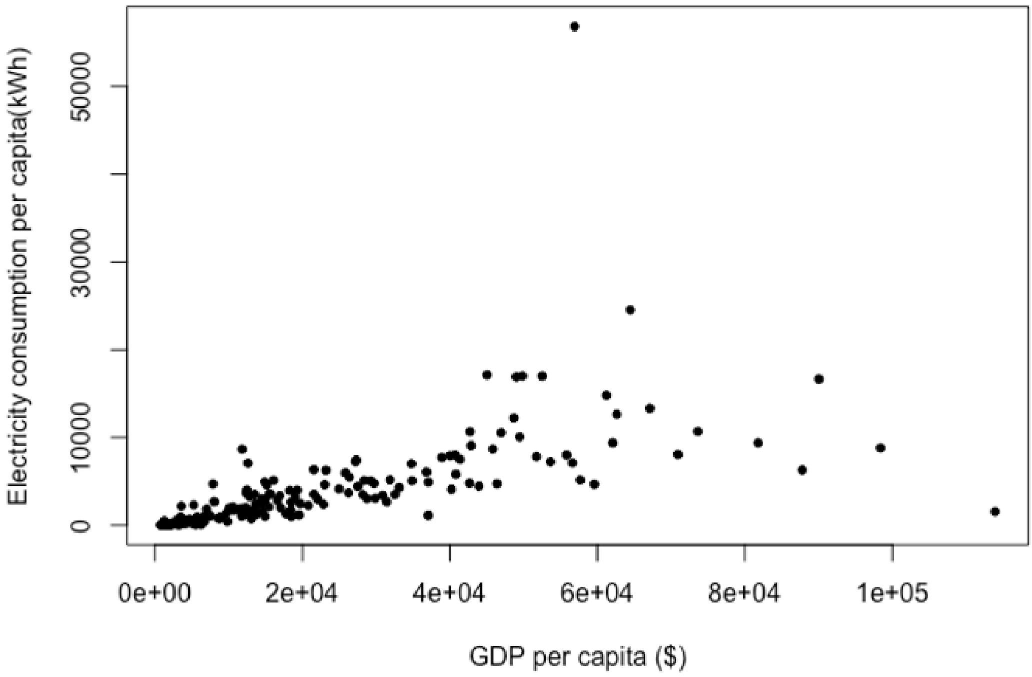

Fig. 2 Electricity consumption vs. GDP per capita, 2019

In both scatterplots, we see a strong positive correlation: electricity access and consumption are low in poorer countries and increase as incomes increase. In Fig. 1 when GDP per capita reaches a certain value, the share of electricity access reaches the upper limit of 100%. That's why many data points congregate at 100% as GDP per capita increases. On the other hand, such kind of percentage limit is not a problem for electricity consumption plot. Therefore, it is more convenient to do regression analysis for the correlation between electricity consumption per capita and GDP per capita.

We fit a simple linear regression for electricity consumption per capita as a function of GDP per capita and draw the diagnostic plots, as shown in Fig. 3 and Fig. 4. From the plots we can see point 96 (Luxembourg) and 73 (Iceland) are outliers and marked in Fig. 3.

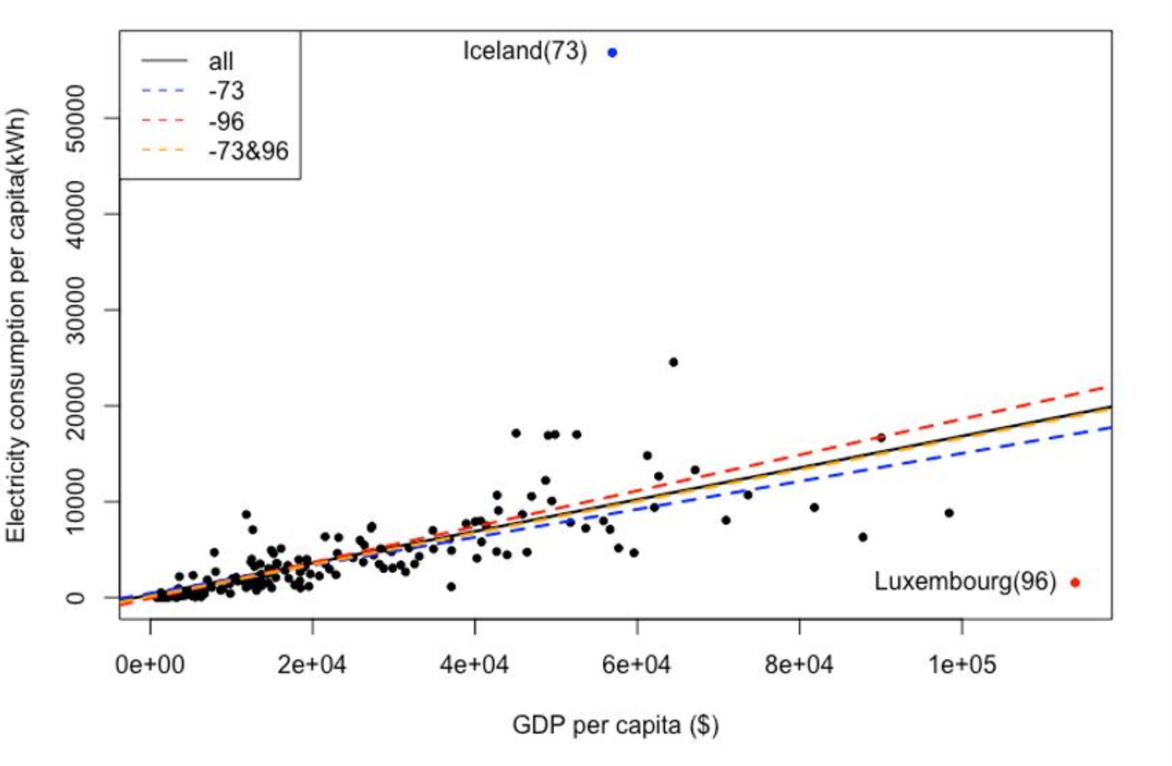

Fig. 3 Linear models with and without outlier(s).

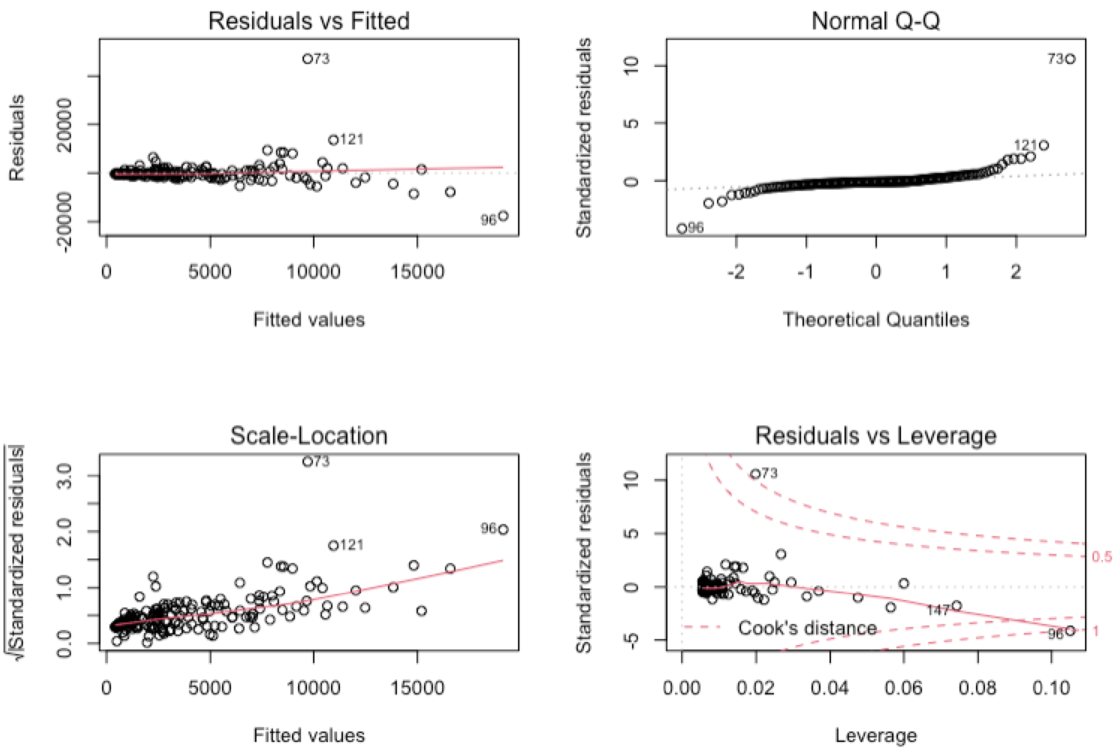

Fig. 4 Diagnostic plots of the linear model including all data points.

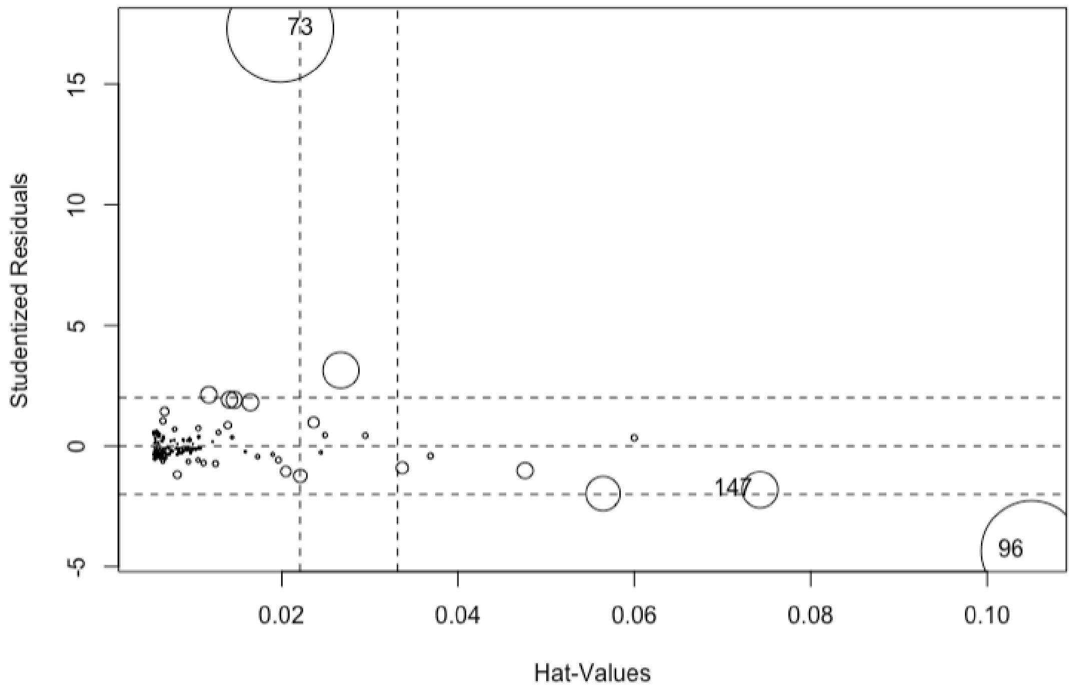

Fig. 5 Regression influences plot of full data model.

The diagnostic plot does not look good, and adjusted R^2 is 0.3884. The regression influence plot of Fig. 5 clearly shows point 73 and 96 are highly influential. To evaluate how influential these two points are on the model, Fig. 3 also shows the regression lines without point 73, without point 96, and without both. We observe that after removing both points, the regression line almost matches the original full data model. However, adjusted increases to 0.6527 without these two influential points. The new model excluding data points of Luxembourg and Iceland is

$$Electricity\_consumption\_per\_capita = 134.1 + 0.165 \times \ GDP\_per\_capita$$

The diagnostic plots of the new model shown in Fig. 6 are still not satisfactory, especially the quantile plot. The residuals increase and become more uncertain with increasing fitted values.

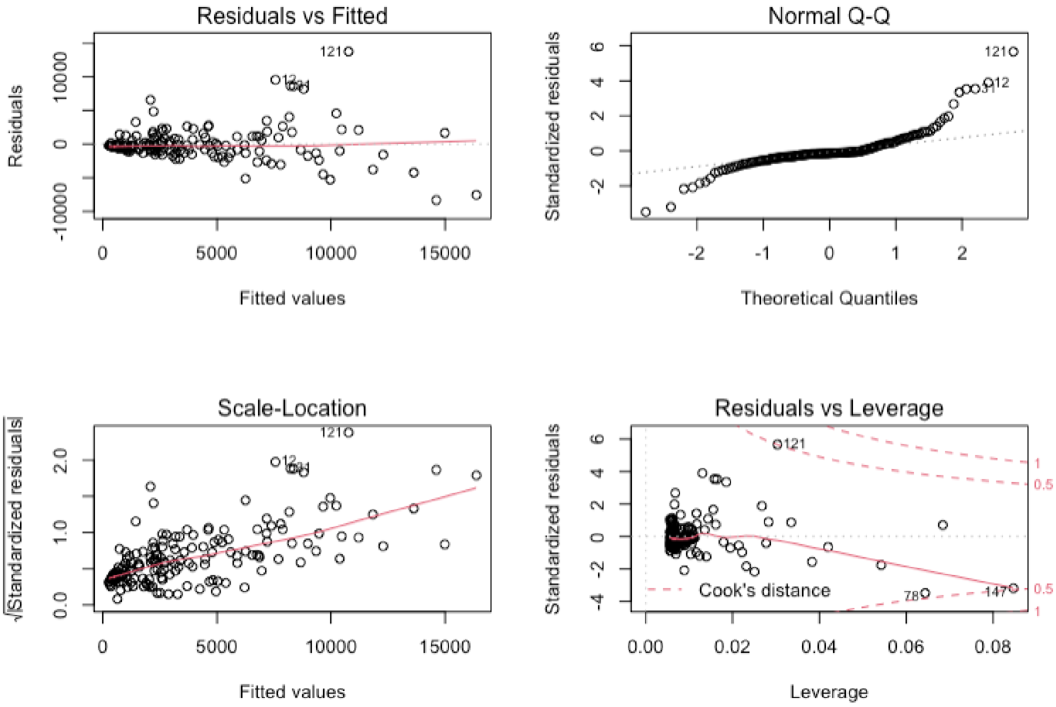

Fig. 6 Diagnostic plots excluding data point 96 (Luxembourg) and 73 (Iceland).

3. Electricity consumption from low-carbon source and GDP per capita:

Electricity generation produces carbon emissions while using the natural resource as a fuel. Depending on the amount of carbon emission, these sources are classified into low-carbon or high carbon resources. In this section of the report, the focus will be on people's consumption of electricity which was generated using low-carbon sources. We are trying to provoke the question whether replacing conventional vehicles with electrical ones will help reduce carbon emissions and help solve the climate change crisis. It will give an insight if the world is ready to transform to an elegant solution such as using electrical vehicles as primer transportation.

ANOVA will be applied on the refined data to find if the percentage of electrical consumption from low-carbon sources to the overall consumption is different among countries. The share of low-carbon energy souce is defined as:

$$r_{low} = \frac{EC_{low}}{EC_{low} + EC_{high}},$$

where EC_{low} is electrical consumption using low carbon resources and EC_{high} is the consumption using high carbon resources. Therefore, the null hypothesis to the ANOVA test would be r_{low} is the same for all countries. Fig.7 shows the boxplot of r_{low} for five income classes of countries based on their GDP per capita.

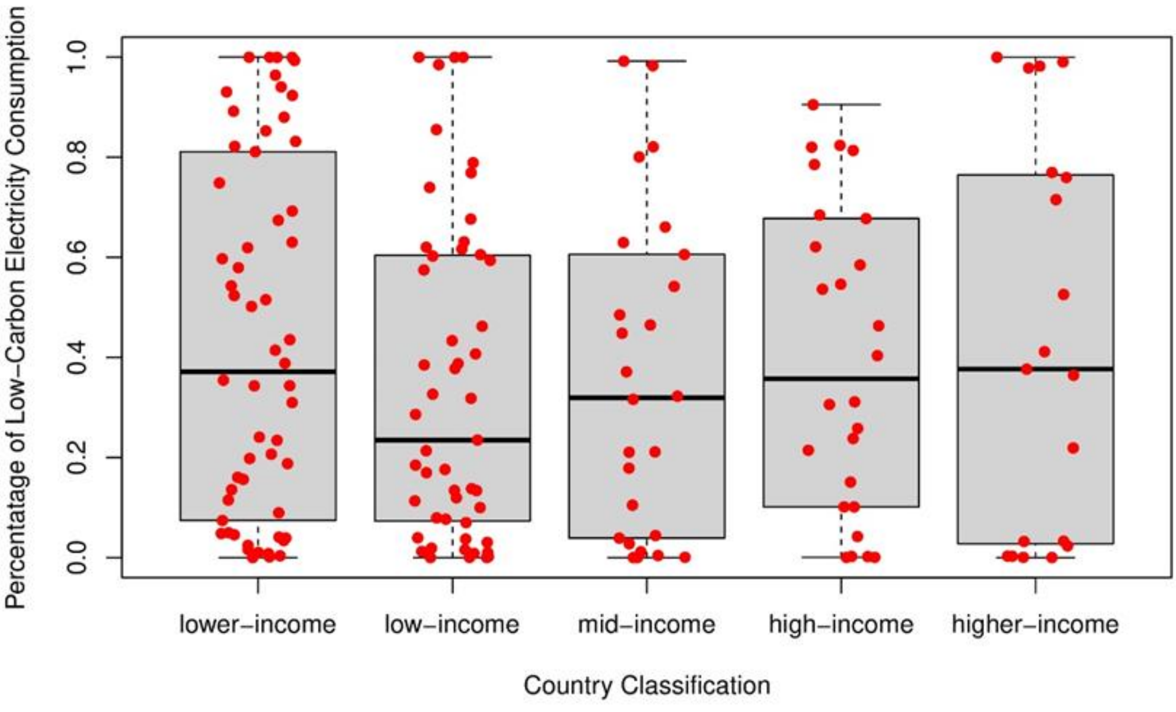

Fig. 7 Electricity consumption using low-carbon resources to the overall consumption ratio by GDP class.

ANOVA test returned a p-value of 0.652, hence, we fail to reject the null hypothesis. Therefore, from the boxplot and the ANOVA test, it is possible to conclude that all countries' electric consumption share from low-carbon resources to the overall consumption is the same. Diagnostic plots will be used to confirm if these results are valid. Please refer to Fig.8.

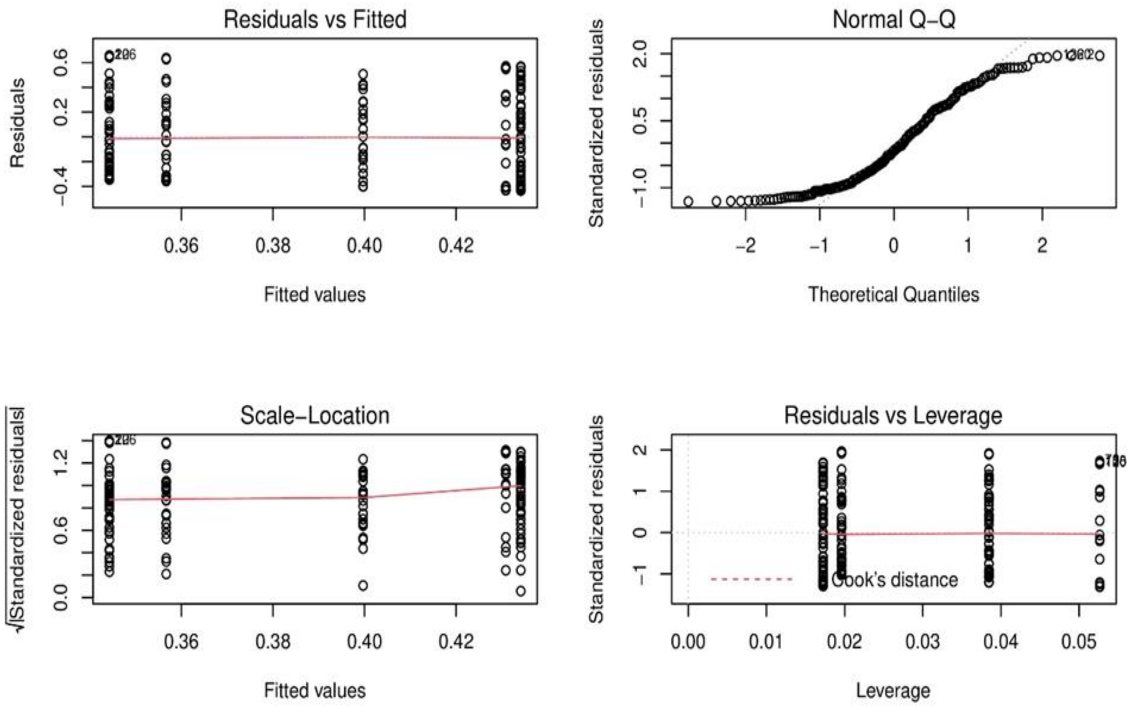

Fig. 8 Diagnostic plots of the ANOVA test

All plots look good, except for the normal qq-plot. We find from the three plots that symmetry and homoscedasticity both exist and cook's distance seems to be reasonable. QQ-plot shows the middle points follow a normal distribution, however, the tails do not. Nevertheless, combining the diagnostic plots results with the obtained boxplot give us a reasonable justification to accept ANOVA results.

4. Low-carbon energy source and electric vehicles (EVs)

Electricity is easier to decarbonize as compared to other two major components that make up total energy production, i.e., transport and heating. Shifting to electric vehicles (EVs) is broadly seen as a solution for a low-carbon future. The International Energy Agency, for example, projects that by 2030, global electricity demand for electric vehicles will increase five-to eleven-fold from levels in 2019. If we are to reap the climate benefits of electric vehicles, this electricity needs to be as low-carbon as possible.

A typical EV consumes around 200 Wh per kilometer. Assuming driving averagely 100 km each day, an EV cosumes 7.3 MWh a year. From sale data of EVs from 2010 to 2019 and assuming all these EVs have been working well since then, Fig. 9 shows the total electricity consumption of EVs in Europe, China, and USA. A near-exponential growth is observed, especially for China.

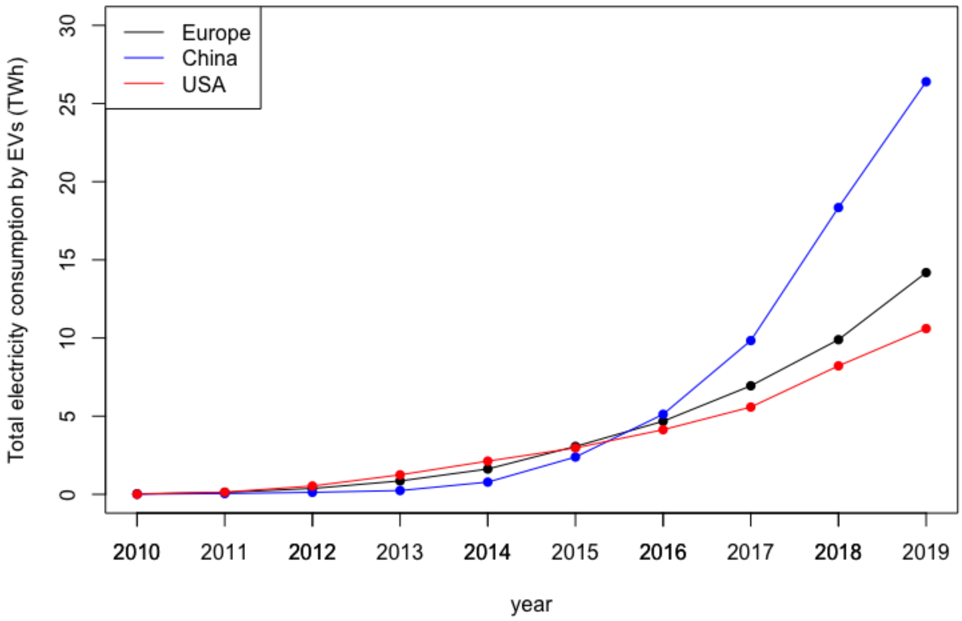

Fig. 9 Annual electricity production from low-carbon source in Eurpe, China, and USA.

Fig.10 shows the growth of electricity production from low-carbon source in Europe, China, and USA from 2010 to 2019. We observe a faster growth rate of production in China. However, none of these countries/region shows a growth rate as fast as that of EV consumption.

Although the total electricity consumption by EVs is still small as compared to total electricity production, the fact that the maket share of EVs is still below 10% and growing fast prompts us to question whether EVs are really carbon-free if the growth of electricity production from low-carbon source cannot catch up.

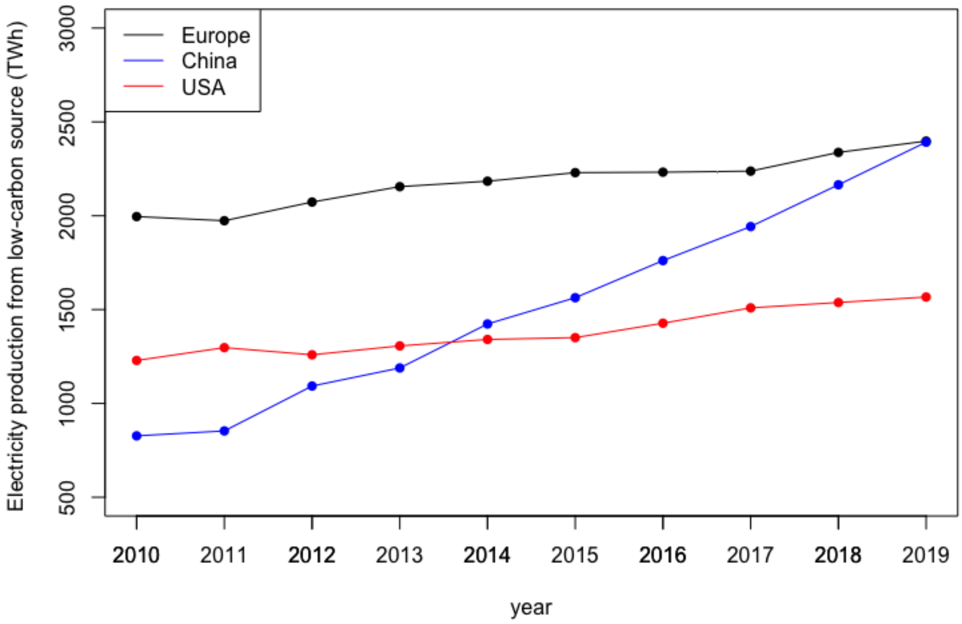

Fig. 10 Annual electricity production from low-carbon source in Eurpe, China, and USA.

5. Conclusion:

A simple linear model to describe the relationship between electricity consumption per capita and GDP per capita is proposed. Diagnostic plots is used to evaluate the model.

An ANOVA is used to test whether all countries had the same ratio of electricity consumption from low-carbon sources to the overall electrical consumption by capita. A very large p-value rejects the null hypothesis. Plotting the boxplot and analyzing the diagnostic plots justified our conclusion and accepted ANOVA test results that there is no obvious difference of low-carbon source ratio between different groups of country divided by GDP per capita.

Rich countries do not necessarily perform better than poor countries on using low-carbon energy source.

EVs are thought as carbon-free at the first glance. The the growth of electricity consumption of EVs outpaces that of electricity generation from low-carbon source. It is questionable that EVs are really carbon-free if the electricity they need does not come from low-carbon souce.

References

[1] http://data.worldbank.org/data-catalog/world-development-indicators

[2] BP Statistical Review of World Energy:

https://www.bp.com/en/global/corporate/energy-economics/statistical-review-of-world-energy.html

Ember: https://ember-climate.org/data/