STEM Fellowship Big Data Challenge 2017-2018

A Model of Renewable Energy Generation in the World with Focus on Canada

Yingyi Liang, Zhamilya Bilyalova, Haohao Fu, Cheng Guo

Princeton International School of Mathematics and Science

April 9, 2018

Abstract

Renewable energy is becoming a more and more important topic as our natural resources drain. Many social factors are influencing and influenced by the use of renewable energy, such as economy, education and environment. To explore this, a model is built with selected social factors that might influence the use of renewables: education expenditure, GDP, population density, surface area, electricity access and research expenditure for each country, and renewable is presented by the ratio of electricity generated by renewables to total electricity generation for the countries. Data from World Bank, Government of Canada and Wikipedia are analyzed with R in RStudio, and a statistically significant model was built. Our model shows that GDP and education level have a positive relationship, and population density, electricity access have a negative relationship with the electricity generation percentage by renewable energy. Change of this percentage over time for provinces in Canada is then looked at more closely and we see that there's a gap between Canada provinces with high generation percentage by renewable energy and provinces with low percentage. Future researches could be done on adding more factors to the model, then building and comparing Canada's model with model for the world to see if the same factors are influential.

Keywords
Renewable energy, Linear model, Canada, Data analysis

1 Introduction

Renewable energy is the energy derived from renewable resources that can be continuously replenished. Due to Canada's vast landscapes and diversified geography, it has an access to many renewable resources that include wind, solar, biomass, geothermal, moving water and ocean energy and can be used to produce energy[1]. Our exploration is based on energy generation. The generation of renewable energy can be affected by such parameters as education, GDP, population density, surface area, electricity access, and research expenditure.

In developing a model that analyzes the different social parameters that affects the percentage generation of renewable energy, our project aims to understand the world situation of renewable energy better by using big data, and possibly aid the government decision on fields to invest in in order to to increase the production of renewable energy. Canada, being the world leader of renewable energy, is also of our interest, so we took a closer look at Canada's place in the world, and the development of renewables in Canada by provinces. Studies have been done on analyzing the factors that influence renewable energy usage of a country as well as influence of renewable energy on other social factors. Renewable energy generation is found to reduces imports to a significant extent [2], and renewable energy consumption can increase economic growth[3][4].

2 Materials & Methods

Our data about world renewable energy generation is from Wikipedia(https://en.wikipedia.org/wiki/List_of_countries_by_electricity_production_from_renewable_sources) and the different parameters is from World bank (https://data.worldbank.org/indicator). Data specifically on Canada is from Canada's Renewable Power Landscape – Energy Market Analysis 2016 report. This data include provin-

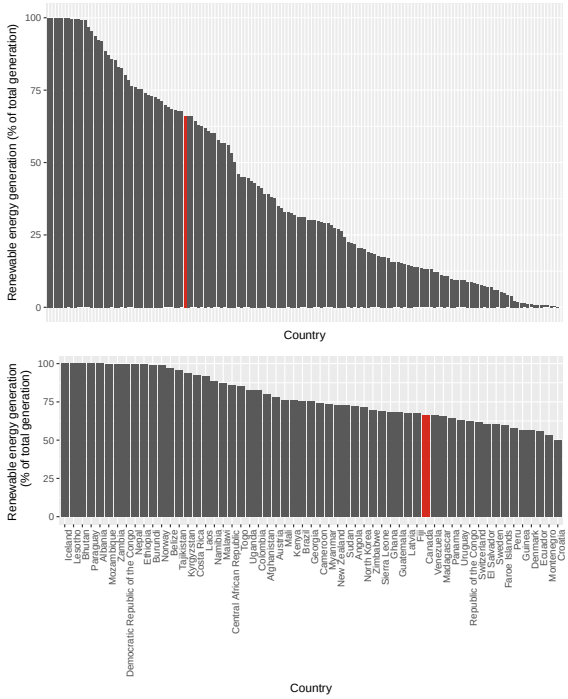

Figure 1: Renewable energy generation percentage of Canada (marked in red) comparing to other countries in the world.

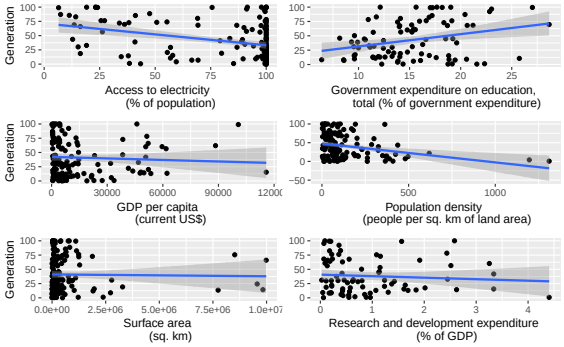

Figure 2: Renewable energy generation percentage of countries graphed with each parameter in the model.

```
Residuals:
    Min      1Q  Median      3Q     Max
-57.21 -15.22   0.40  12.16  69.80

Coefficients:
                     Estimate Std. Error t value Pr(>|t|)
(Intercept)         4.310e+01  2.103e+01   2.050  0.04556 *
Education           2.649e+00  9.265e-01   2.860  0.00613 **
GDP                 5.259e-04  2.495e-04   2.108  0.04000 *
Population_density -4.406e-02  1.791e-02  -2.460  0.01734 *
Surface_area       -9.703e-07  2.030e-06  -0.478  0.63472
Electricity_access -4.213e-01  1.477e-01  -2.852  0.00626 **
Research_expenditure -6.237e+00  5.359e+00  -1.164  0.24991
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 25.8 on 51 degrees of freedom
  (223 observations deleted due to missingness)
Multiple R-squared:  0.3917,    Adjusted R-squared:  0.3202
F-statistic: 5.474 on 6 and 51 DF,  p-value: 0.0001963
```

Table 1: Result from model of generation with education expenditure, GDP, population density, surface area, electricity access and research expenditure in R.

cial and Canadian total renewable electricity capacities from wind, solar, biomass, hydro and other renewable from 2005-2015(http://open.canada.ca/data/en/dataset/7e50d70e-f462-48ea-97f5-6c0bc48cfd14). Data was cleaned, combined, and analyzed with R in RStudio (source code can be provided via github for reproducing the result if needed). The packages used are readr, dplyr, tidyr, ggplot2, plotly, knitr, kableExtra, gridExtra, stringr, syuzhet, tidyverse, RColorBrewer, rvest, and ggrepel.

3 Results

Our data analysis focuses on the ratio between generation of electricity from renewable sources to generation of total electricity (in %). We first looked at the data for all countries in the world with focus on Canada. The distribution is shown in Figure 1, with Iceland and Lesotho leading the way, their generation percentage approaching a hundred percent. Canada is marked red, and we see that it is in the first one third of all countries for the ratio of renewable generation. A zoomed in graph shows us that with generation of 66% in 2015, Canada is in between Fiji and Venezuela.

Then, six parameters: education expenditure, GDP, population density, surface area, electricity access and research expenditure were introduced and graphed with the generation of renewables to see their correlations (Figure 2). Most correlations seem insignificant, with population density and electricity access negatively related to renewable generation ratio, and education expenditure positively related to the ratio. The negative correlation of electricity access was unexpected, so we graphed it independently with country names labeled (Figure 3) to examine the relationship. The six parameters are then put together to make a linear model with the generation ratio, results are shown in Table 1. The linear model yielded an adjusted r-square of about 0.32, and a p-value of 0.0002. In all the factors, education and the access of electricity are the most significant, then the gross domestic product, and population density. We failed to find any correlation between surface area and research expenditure of a country with its renewable energy production.

After that, we focused specifically on the renewable energy production of Canada. Change

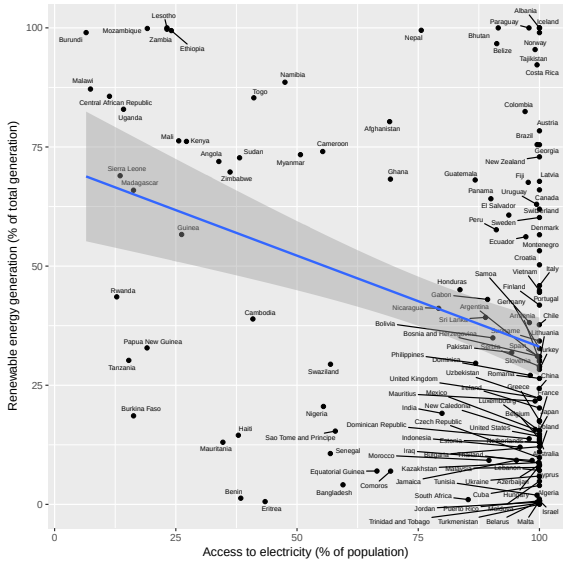

Figure 3: Renewable energy generation percentage graphed with access of electricity for all countries. Country names labeled.

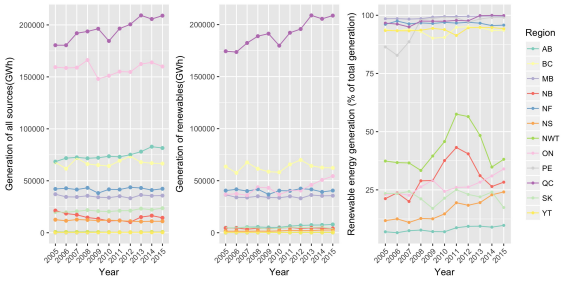

Figure 4: Linear model of change of the ratio between total energy and renewable energy generation during 2005-2015 for Canada provinces (Nunavut excluded).

in the total energy generation, renewable energy generation and their ratio in twelve provinces (Nunavut was excluded because it has almost no renewable generation and a lack of data) of Canada between the years 2005-2015 are shown in Figure 4. We can see that despite Saskatchewan, ratio for all the provinces are increasing. And it can be clearly seen that there are two groups, with Prince Edward Island, Manitoba, Quebec, Newfoundland and Labrador, Yukon, and British Columbia on the top of the graph, their ratio approaching 100%; Northwest Territories, New Brunswick, Ontario, Saskatchewan, Nova Scotia and Alberta on the bottom, their ratio below 50%. Ratios are fitted in linear model in Figure 5, and these results are visualized on map in Figure 6. We took a closer look at the energy generation from different sources of Prince Edward Island (Figure 7) since it has the highest rate of increase, ap-

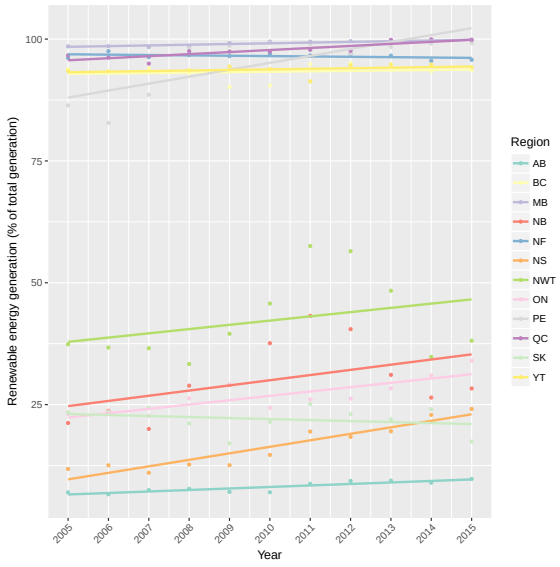

Figure 5: Change of the generation of total energy, renewable energy, and the ratio between them during 2005-2015 for Canada provinces (Nunavut excluded).

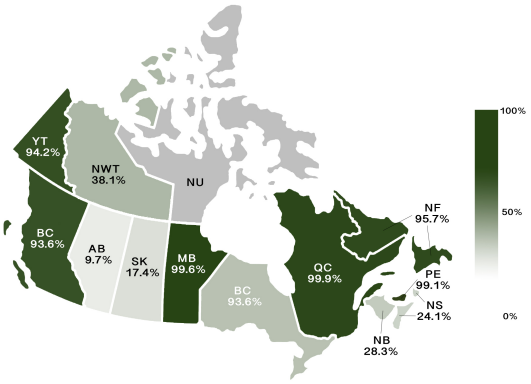

Figure 6: Ratio between generation of renewable energy and total energy in 2015 for Canada provinces visualized on map (Nunavut represented in gray).

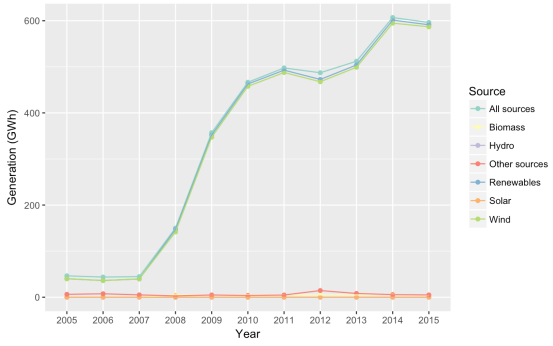

Figure 7: Energy generation from different sources for Prince Edward Island.

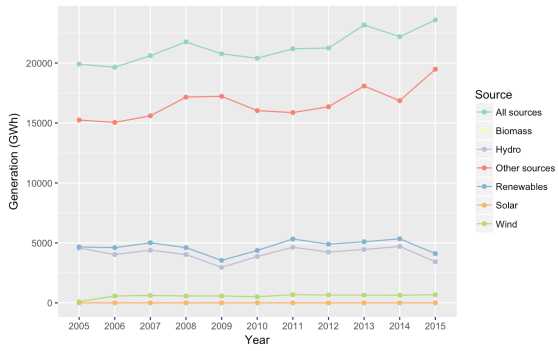

Figure 8: Energy generation from different sources for Saskatchewan.

proaching 100%. We also looked at the energy generation by source of Saskatchewan (Figure 8) because it appears to be the only one with decreasing ratio.

4 Discussion

Canada has a relatively high renewable energy generation percentage among other countries in the world, it is in the first one third of the distribution graph(Figure 1). Then we looked at possible social factors that might have caused the difference in this generation of renewable. For the linear model, six factors are chosen to represent different aspects of a country: use of electricity, importance of education, economic condition, demographics, area, and focus on research. All of them were expected to be positively related to renewable energy generation despite population density. Although according to Figure 2, the linear model of relationships between each factor with generation ratio is vague, when we combine them together, we end up with a statistically significant model (Table 3) which has an adjusted r-squared interesting enough to consider. In reality, the access of electricity is also found to be negatively proportional to the generation ratio. We assumed it was because of some countries in Africa that has low accessibility to electricity but uses mostly solar energy. Figure 3 proves our idea, with African countries such as Burundi, Malawi, and Mozambique on the top left corner of the graph. Similarly, for population density, European countries like Iceland and Norway, also African countries like Paraguay and Zambia with low population density and high renewable energy generation percentage appears on the upper left corner, making the parameter negatively related. Government expenditure on education was positively related as expected, we think highly educated people are more likely to pay attention on using sustainable energy. Also, it shows that the government pays high attention on future developments of the country, which includes education and energy use. Finally, a country's GDP is also positively related to renewable energy use. Countries with good financial situation are able to put money into investigating new energy resources and building the power plants needed - European countries like Switzerland and Denmark on the upper right of the graph explains the trend.

Then we examined closely for the generation of renewable in different provinces of Canada. The data for electricity generation was chosen instead of electricity capacity, to stay consistent with analysis of the world data. According to Figure 4, although provinces like Ontario and Alberta have a high net generation of electricity, their low renewable generation causes a low percentage. While provinces like Yukon and British Columbia have low generation but high percentages. In Figure 5, we can see that there are two clusters of data. And according to the map in Figure 6, the bigger coast provinces generally have higher renewable energy generation ratio. For Figure 7, Prince Edward Island is analyzed for its high increase rate of renewable generation ratio. We can see that most of its electricity are generated by renewables, specifically wind. And its generation started to increase sharply from 2007. Saskatchewan is analyzed in Figure 8 for its decreasing ratio of electricity generation from renewables. We can see that its total generation of electricity has been slowly increasing over time, while the generation by renewable energy (mostly hydro) is dropping, making the overall ratio drop. Another interesting thing we noticed in Figure 4 (right) is that although the ratios for all provinces are generally increasing, many of them dropped around 2007, peaks around 2011 and then reached another low point around 2014.

In all these analysis, errors might have been

introduced by a lack of data for the parameters for some countries. Also, although we tried to used all the world data from 2012, some lack of data are compensated by using data from 2013 or 2011, which could also introduce error. This disagreement of dates also exist in the wikipedia data for world energy generation.

Conclusions

In this paper, we successfully created a model with the parameters about education, GDP, population density, surface area, electricity access and research expenditure of the countries that predicts their generation percentage of renewable energy. The model has an adjusted R-squared 0.3202 and p value 0.0002, which is less than 0.05, showing that this model is statistically significant. The parameters for education (government expenditure on education, total, % of government expenditure), access to electricity (access to electricity, % of population) are the most significant. With education expenditure directly proportional and access to electricity inversely proportional to the generation percentage. GDP(GDP per capita, US$) and population density (people per sq. km of land area) are also two parameters that are influential to some extent. So investigating in education can potentially increase the generation of renewable energy. In future research, more parameters could be included to have a better understanding of which social parameters influences the renewable energy generation percentage.

We looked more closely at Canada, on its change of renewable generation percentage over time by different provinces. We see that the values range from 9.7%(Alberta) to 99.9%(Quebec), some provinces are far ahead of others, with Prince Edward Island having the highest rate of increase. For next steps, we can make a similar model to predict renewable energy generation percentage with social parameters for provinces of Canada, then compare it with the world model to see the similarities and differences.

References

[1] About Renewable Energy. (2017, December 13). Retrieved January 19, 2018, from http://www.nrcan.gc.ca/energy/renewable-electricity/7295

[2] Vaona, A. (2016). The effect of renewable energy generation on import demand. Renewable Energy, 86, pp.354-359.

[3] Rafindadi, A. and Ozturk, I. (2017). Impacts of renewable energy consumption on the German economic growth: Evidence from combined cointegration test. Renewable and Sustainable Energy Reviews, 75, pp.1130-1141.

[4] Inglesi-Lotz, R. (2016). The impact of renewable energy consumption to economic growth: A panel data application. Energy Economics, 53, pp.58-63.