Analysis of COVID-19: Model-Fitting and Correlation with Commodities and Pollutants

Dhruv Gupta

Department of Physics, Indian Institute of Technology Hyderabad, EP17BTECH11006. Email: ep17btech11006@iith.ac.in

Abstract

This report concerns the COVID-19 pandemic and its effects on prices of commodities such as crude oil, petrol, diesel, and also checks the correlation between the cases of COVID-19 in India with the air quality/pollution levels.

Keywords: COVID-19, model-fitting, correlation, Hill Equation, Spearman, Perason, Kendall Tau

1 Introduction

COVID-19 is the official name given to the novel coronavirus that is currently known to have originated in Wuhan, China in December 2019, according to the World Health Organisation [1]. Since the outbreak in China, the virus has spread rapidly across the world, primarily affecting the first world countries, now up-to a total count of over one million cases as of writing this report. The scale of the pandemic is vast, and its effects are far-reaching, seeing lock-downs across the world in response and economies slowing down. In this report, my objective would be to reproduce the results of a paper [2] that attempted to model the cases of COVID-19 in countries using the Hill Equation. Additionally, this report will look at three different correlations - one, between crude oil prices and the total infected cases worldwide; two, the fuel prices across India with cumulative cases in the country; three, the air quality index and pollution levels across cities in response to the recent lock-down and rise of cases, recently also analyzed by the Centre for Pollution Control Board (CPCB) in a report recently [3].

Data Sources

My data sources are linked in the individual subsections wherever mentioned and referenced in the references at the end of the paper. All data has been taken for the date range Jan 22-March 29, 2020.

- COVID-19 Global Cases - Johns Hopkins CSSE [4]
- COVID-19 State-wise India Cases - COVID-19 Corona Virus India Dataset [5]
- Pollutants Concentration - Automatic Monitoring Data [6]
- Air Quality Index - Air Now DoS [7]
- Petrol Prices (India) - PPAC [8]
- Crude Oil Prices - WTI Historical Prices [9]

Code

All code for generating graphs and computing various statistics has been uploaded to my GitHub repository [10].

2 Model-Fitting

Growth-Model: The Hill Function

The cumulative data of cases is fit using a growth model called the Hill function. The Hill function is used to model population-dependent growth or decline [11], in this case modelling the infected population.

$$I(T) = \frac{AT^m}{K + T^m} \tag{1}$$

Here, T is the time variable and the parameters are described; A is a scaling factor for the population, K captures the time after which growth is half-complete before saturation, and m is the growth-factor which defines how quickly the population rises. Shown below is an example:

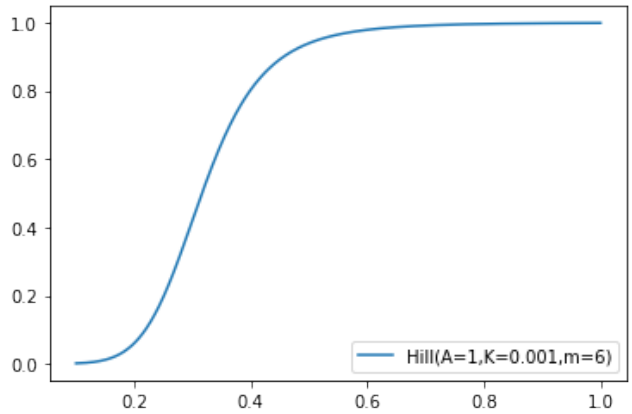


Figure 1. Sample Growth Model: Hill Equation

The countries for which data is being fitted are: China, France, India, Iran, Italy, South Korea, United Kingdom, US. The cumulative cases for the countries are plotted below through the timeline:

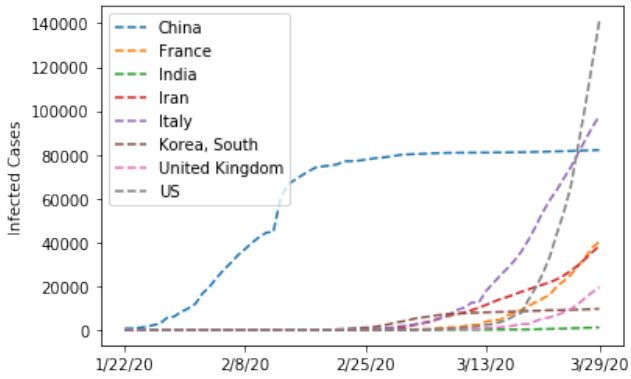


Figure 2. Global COVID-19 Cases [4]

2.1 Fitting Curves for 8 Infected Countries*

Using the *curvefit* module from scipy-optimize, the datasets for 8 countries fitted to corresponding Hill functions are shown.

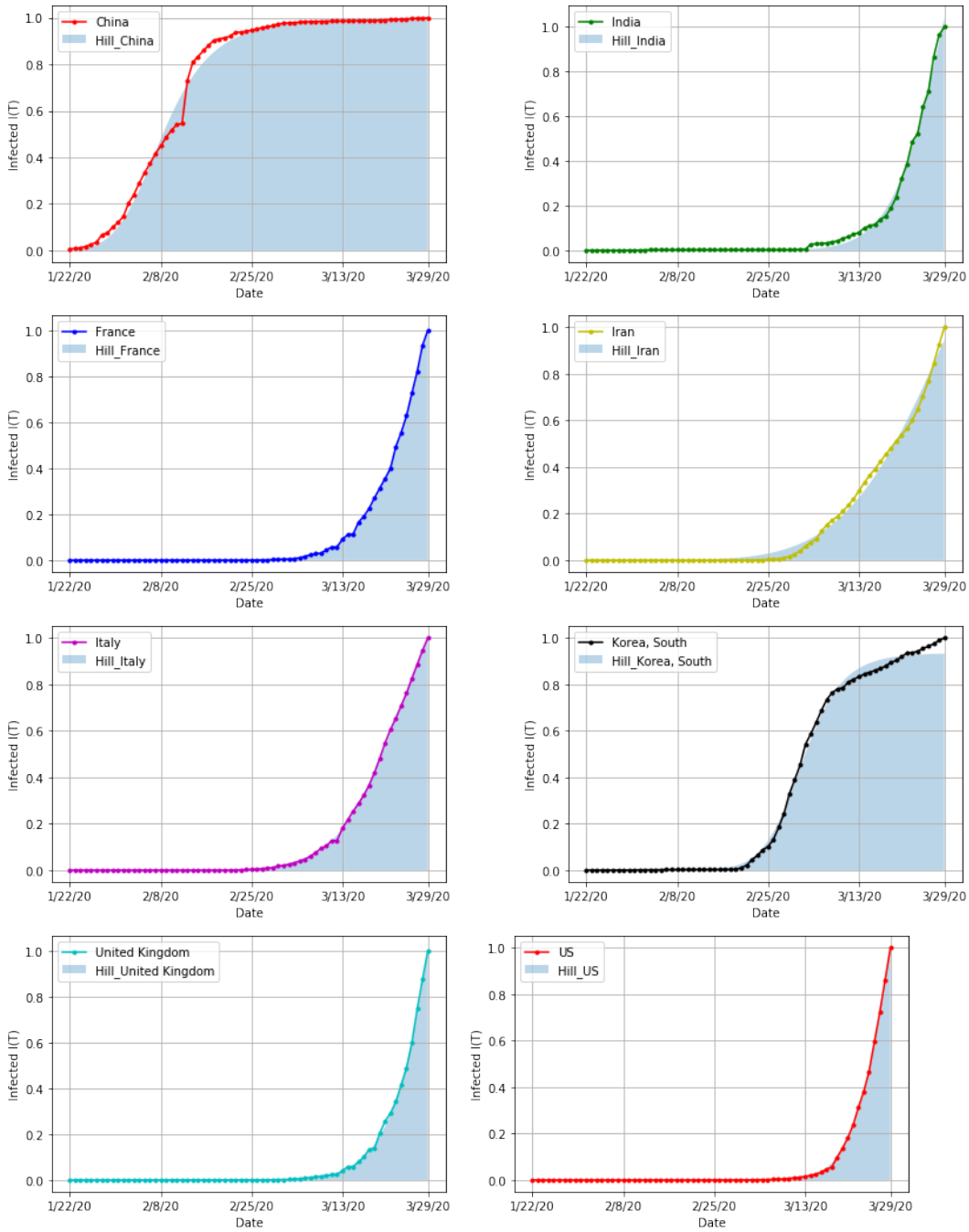


Figure 3. Hill-Equations Fitted to COVID-19 Infected Data

3 Correlation Coefficients

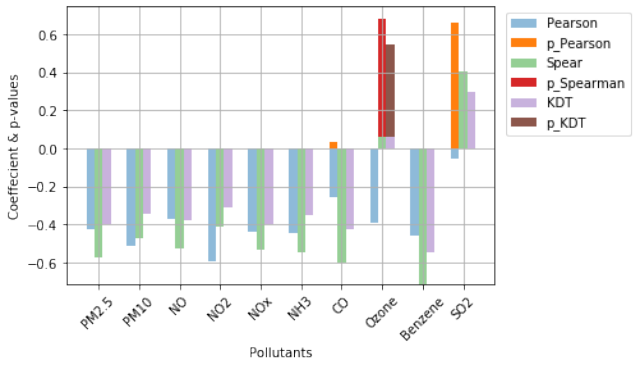
Method

Data for infected cases in India is taken from the JH CSSE [4]. The pollutants concentration is taken from CPCB [6]. The petrol prices are taken from the PPAC [8]. The correlations are drawn across multiple pollutants and the air-quality-index(AQI) for three cities, and for the petrol-diesel prices

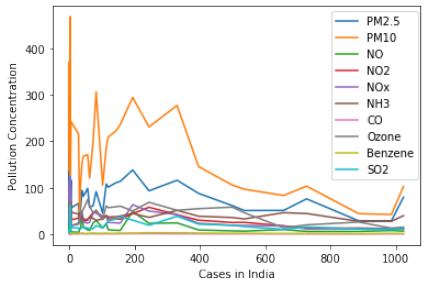
for four cities.
The correlation coefficients used are the Pearson, Spearman, and Kendall-Tau coefficients. The corresponding p-values to test the null hypothesis are also shown in the tables.

3.1 Pollutants-India

The concentration of pollutants in the Mundka district of Delhi are tested for correlation with the cumulative cases in India. The below stacked bar plots show the correlation coefficients and the corresponding p-values.



(a) Correlation Coefficients



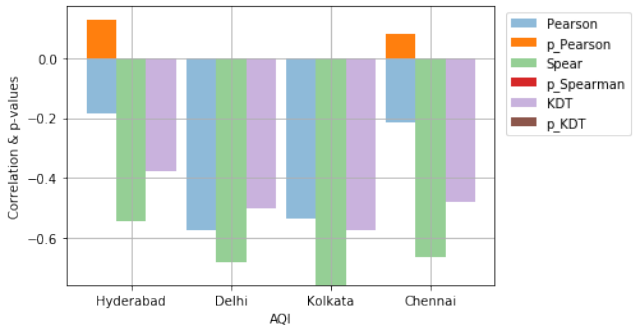
(b) Raw Data

Figure 4. Pollution Correlation With COVID-19 Indian Cases

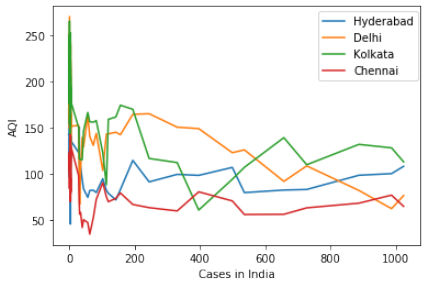
There seems to be an anti-correlation for all the pollutants,except for Ozone and SO2. Since their p-values are quite significant we will have to reject the null-hypothesis that the data-sets are anti-correlated.

3.2 Air Quality Index-India

The Air Quality Indices [7], indicative of the overall air quality of four cities (Hyderabad, Delhi, Chennai, Kolkata), are tested for correlation with the cumulative cases in the corresponding states of the cities. The below stacked bar plots show the correlation coefficients and the corresponding p-values. The higher the AQI is, the more hazardous it is.



(a) Correlation Coefficients



(b) Raw Data

Figure 5. AQI Correlation With COVID-19 Indian Cases

There seems to be an anti-correlation for all the pollutants, with the concentrations decreasing with increase in the infected population.

3.3 Petrol Prices-India

The petrol prices in four cities [8](Mumbai, Delhi, Kolkata, Chennai) are tested for correlation with the cumulative cases in the corresponding states of the cities. The below stacked bar plots show the correlation coefficients and the corresponding p-values.

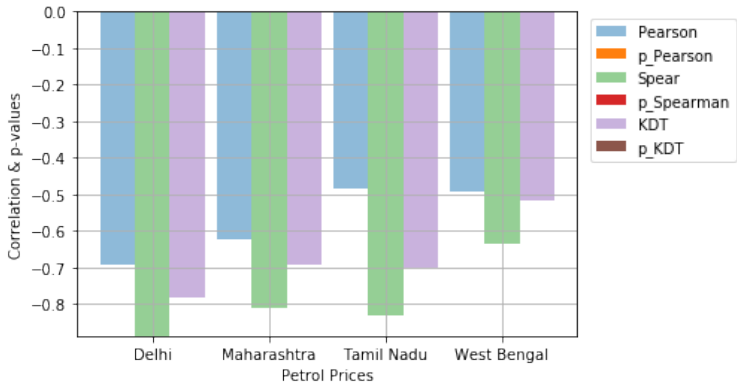
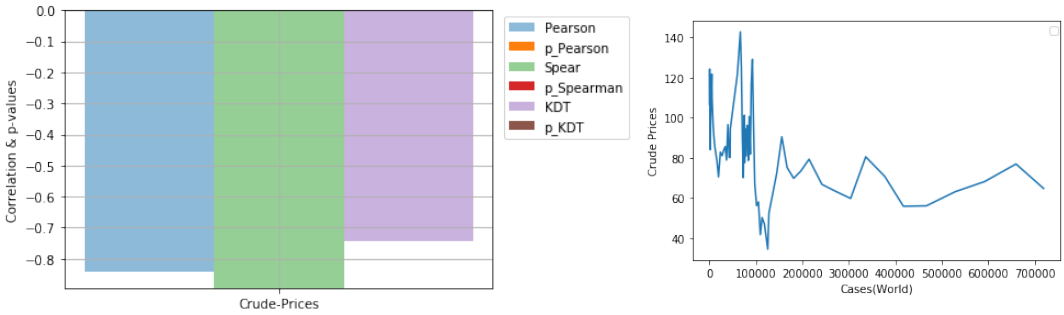


Figure 6. Petrol Prices Correlations

There seems to be a strong anti-correlation.

3.4 Crude Prices-World

The global standard crude oil prices [9] are tested for correlation with the cumulative cases of COVID-19 in the world for the given timeline. The below stacked bar plots show the correlation coefficients and the corresponding p-values.



(a) Correlation Coefficients

(b) Raw Data

Figure 7. Crude Price Correlation With COVID-19 Global Cases

There seems to be a strong anti-correlation.

4 Results

The below tables show the parameter-fits for the model-fitting of cumulative COVID-19 cases in the various countries and the correlation coefficients and the p-values for the correlation tests performed.

Table 1. Parameters for Hill Equation Fits

Country	A	K	Growth Factor(m)
China	1.004	0.0029	5.34
France	2.64	1.61	12.07
India	2.43	1.33	13.90
Iran	2.35	1.48	6.83
Italy	1.52	0.53	10.97
South Korea	0.94	0.0034	12.88
United Kindom	7.68	6.58	13.83
US	2.09	1.09	20.32

Table 2. Pollution Correlation Statistics

Coeff|p-value	PM2.5	PM10	NO	NO2	NOx	NH3	CO	Ozone	Benzene	SO2
Pearson	-0.43	-0.51	-0.37	-0.6	-0.44	-0.45	-0.26	-0.39	-0.46	-0.05
p-Pearson	2.9E-04	9.4E-06	1.8E-03	8.7E-08	1.8E-04	1.4E-04	3.5E-02	9.4E-04	8.2E-05	6.5E-01
Spearman	-0.57	-0.47	-0.53	-0.41	-0.53	-0.55	-0.6	0.06	-0.72	0.41
p-Spearman	3.32-07	5.2E-05	3.7E-06	4.7E-04	3.5E-06	1.4E-06	7.6E-08	6.1E-01	7.3E-12	6.0E-04
Kendall-Tau	-0.4	-0.34	-0.38	-0.31	-0.4	-0.35	-0.43	0.06	-0.55	0.3
p-Kendall-Tau	5.8E-06	1.0E-04	1.7E-05	4.8E-04	6.8E-06	6.2E-05	1.3E-06	4.8E-01	6.5E-10	8.0E-04

Table 3. AQI Correlation Statistics

Coeff|p-value	Hyderabad	Delhi	Kolkata	Chennai
Pearson	-0.28	-0.57	-0.54	-0.36
p-Pearson	2.03E-02	3.17E-07	2.35E-06	2.45E-03
Spearman	-0.55	-0.68	-0.76	-0.69
p-Spearman	1.49E-06	1.35E-10	7.12E-14	6.11E-11
Kendall-Tau	-0.39	-0.5	-0.58	-0.5
p-Kandall-Tau	1.20E-05	1.33E-08	6.88E-11	1.38E-08

Table 4. Petrol Prices(India): Correlation Statistics

Coeff|p-value	Delhi	Mumbai	Chennai	Kolkata
Pearson	-0.69	-0.62	-0.48	-0.49
p-Pearson	6.22E-11	1.26E-08	3.03E-05	1.88E-05
Spearman	-0.89	-0.81	-0.83	-0.64
p-Spearman	7.39E-24	4.32E-17	1.98E-18	5.45E-09
Kendall-Tau	-0.78	-0.69	-0.7	-0.52
p-Kandall-Tau	7.47E-17	6.42E-13	5.63E-13	1.92E-07

Table 5. Crude Prices(Global): Correlation Statistics

Coeff|p-value	Crude
Pearson	-0.84
p-Pearson	1.7E-19
Spearman	-0.90
p-Spearman	7.54E-25
Kendall-Tau	-0.74
p-Kandall-Tau	4.71E-19

5 Conclusion

The objective of this report was to replicate the results of the paper [2] that attempted to model the COVID-19 cases against a growth function - here, the Hill function. As is seen for the eight countries, there is good agreement between the data and the Hill function. Here it was essential to normalize the cases first and define an appropriate domain to work with to avoid complications with finding the optimized parameters.

After that came the correlations. It is to be noted here that these are not well-fortified correlations as there are other factors at play. The concentration of pollutants and the general AQI seems to have been slightly decreasing with the increase in Indian COVID-19 cases, however this is also largely due to the lock-downs in the recent month. Petrol prices can be seen to be dipping with the COVID-19 cases in India as well, with the most in Delhi, followed by Maharashtra, Tamil Nadu, and West Bengal. This could be explained as follows; the initial cases of COVID-19 were in Delhi [5], soon-after which cases in Maharashtra started piling up, while West Bengal only recently reported cases.

The strongest of correlations seemed to have been shown by the crude oil prices which are much more closely linked to the international scene, where COVID-19 had already spread from China and through to the Middle-East and the US. The growth in cases can be seen to be strongly anti-correlated to the COVID-19 pandemic.

Concluding, there is a lot more scope in the given topic, using more complicated models such as the SIR model [12], and taking into account various factors such as a country's transportation, social structure, etc. to model the growth/decline of COVID-19, better.

References

[1] Who-covid19 coronavirus disease (covid-19) pandemic. https://www.who.int/emergencies/diseases/novel-coronavirus-2019. Accessed: Apr 3, 2020.

[2] Jagadish Kumar and K. P. S. S. Hembram. Epidemiological study of novel coronavirus (covid-19). 2020.

[3] Impact of janta curfew lockdown on air quality. https://www.cpcb.nic.in/air/NCR/jantacurfew.pdf. Accessed: March 31, 2020.

[4] 2019 novel coronavirus covid-19 (2019-ncov) data repository by johns hopkins csse. https://github.com/CSSEGISandData/COVID-19. Accessed: March 29, 2020.

[5] Covid-19 corona virus india dataset. https://www.kaggle.com/imdevskp/covid19-corona-virus-india-dataset/. Accessed: March 30, 2020.

[6] Central control room for air quality management - all india. https://app.cpcbccr.com/ccr/#/caaqm-dashboard-all/caaqm-landing. Accessed: March 30, 2020.

[7] Airnow department of state. https://www.airnow.gov/index.cfm?action=airnow.global_summary#India. Accessed: March 31, 2020.

[8] Retail selling price (rsp) of petrol and diesel in metro cities since 16.6.2017. https://www.ppac.gov.in/content/149_1_PricesPetroleum.aspx. Accessed: March 31, 2020.

[9] Wti crude oil prices - 10 year daily chart. https://www.macrotrends.net/2516/wti-crude-oil-prices-10-year-daily-chart. Accessed: March 31, 2020.

[10] Covid-19 dsa project: Code. https://github.com/DhruvG99/dsa/.

[11] Hill equation. http://systems-sciences.uni-graz.at/etextbook/sw2/hill_function.html. Accessed: March 31, 2020.

[12] Wikipedia contributors. Compartmental models in epidemiology — Wikipedia, the free encyclopedia. https://en.wikipedia.org/w/index.php?title=Compartmental_models_in_epidemiology&oldid=948887492, 2020. [Online; accessed 3-April-2020].