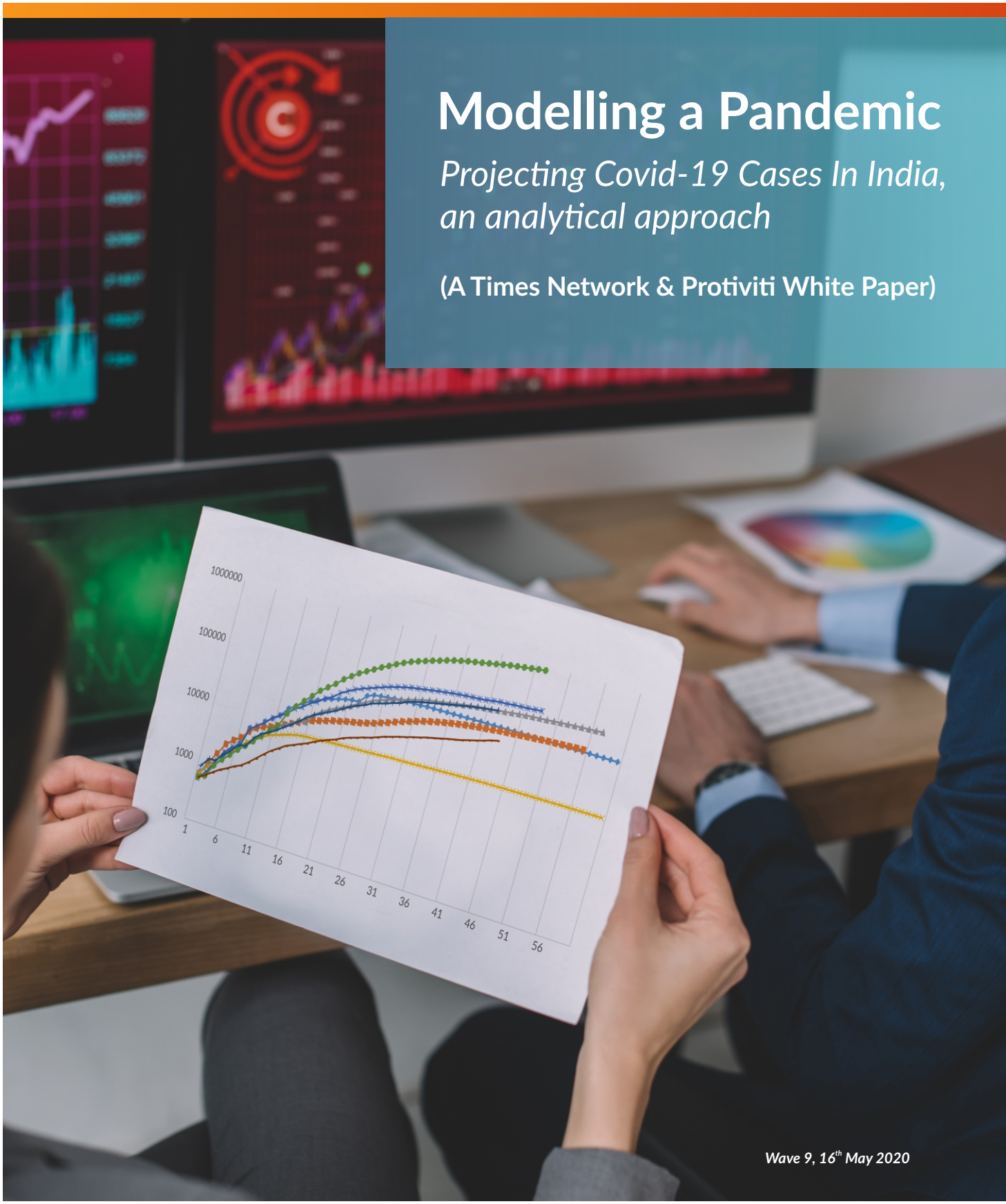
TIMES NETWORK
NOW OR NOTHING

protiviti®
Face the Future with Confidence

Modelling a Pandemic

Projecting Covid-19 Cases In India, an analytical approach

(A Times Network & Protiviti White Paper)



Wave 9, 16th May 2020

Internal Audit, Risk, Business & Technology Consulting

Summary

Covid-19 is changing the world around us. Almost every country is feeling the unprecedented impact of the pandemic and is devising ways of dealing with the situation. In an attempt to understand how this will impact India as a country, Protiviti, a global consulting firm and Times Network, a premium broadcast network and part of India's largest media conglomerate, The Times Group, came together to project the path of the pandemic in India.

The paper prepared by Protiviti's data analytics team tries to look at the trajectory of the COVID-19 outbreak as an epidemiological event and attempts to model the same using different modeling techniques. This paper focusses on time series based quantitative analysis and SEIR modeling already developed for prior epidemics but tailored and advanced for COVID-19.

Covid-19 outbreak

The Coronavirus caused the COVID-19 pandemic presents an unprecedented and severe threat to the lives of entire countrymen, the Indian healthcare system, and the economy. The overall idea of this paper is to propose our view on this pandemic spread and how Indian leaders can use this study to limit the spread of COVID-19.

This study has been performed by Protiviti and Times Network during ***April 2020***, to estimate the extent to which COVID-19 is expected to influence India as a country using aggregated statistical measures computed based on three types of forecasting models: **% Based Models, Time Series model & SEIR Model.**

We analyzed India as a whole, top **10 Indian states, and 3 top hot-spots** along with the entire country for a period of close to 50 days to project the evolution of the COVID-19 pandemic using the 'Number of daily new cases' along with the total 'active' patients count.

The projected development of infected cases assumes an inherited effect of countrywide lockdown applied from **24th March 2020**, and the numbers proposed for the future also consider that this period of lockdown continues for the next **45 days or so.**

Further, Scenario Analysis is performed to understand the effect of dates of lifting the lockdown on the number of infected cases in India.

About COVID-19

As defined by World Health Organization (WHO), Coronavirus disease COVID-19 is a highly contagious disease that is primarily said to spread through saliva droplets or through the discharge of a sneezing infected person.

Some of the prominent symptoms experienced by the COVID-19 virus infected people include mild to moderate respiratory illness while recovering with no specific treatment. However, older people and people with existing medical conditions of diabetics, cancer or some other chronic diseases are more prone to developing severe illness with this virus.

With no specific vaccines or treatment available for COVID-19, the best way to prevent and slower the rate of transmission is by being well informed about the virus, the symptoms and how its spreads. Washing your hands several times in a day or using a alcohol based sanitizer and avoiding to touch one's face multiple times are some of the options to protect oneself and others.

The World Health Organization (WHO) declared the 2019–20 coronavirus outbreak a Public Health Emergency of International Concern (PHEIC) on 30th January 2020, and since then, 4,805,229* new cases have been registered, of which approx. 36% recovered, and 7% were fatal. A total of 184 new countries have since reported an outbreak in a short period of 2 months.

India's response to COVID-19 has been pre-emptive, pro-active and graded with high-level political commitment and a 'whole government' approach to respond to the COVID-19 pandemic. Ministry of Health and Family Welfare is working to strengthen surveillance, build capacity of health system and optimize 'window of opportunity' created by mandatory physical distancing in India.

In India the total cases registered as on 17th May, 2020 was 95,698 with 55 days of mandatory social distancing within the 404 districts. The most affected states / union territories are Maharashtra, Tamil Nadu and Delhi. Out of 736 districts, 520 districts have reported confirmed numbers. The silver lining is that the poorest of districts have the least number of cases. A total of 36 districts have seen no fresh cases in the last 14 days, and 46 districts have seen no fresh cases in the last 7 days. Fewer cases also mean, quarantine and track & trace procedures are easier to apply, and form an effective first line of defense.

Coverage & Assumptions

01 This paper presents a glimpse of the study performed by Protiviti for Times Network as of 20th April 2020, to estimate the extent to which COVID-19 is expected to influence India as a country using aggregated statistical measures computed based on three types of forecasting models: % Based Models, Time Series model & SEIR Model.

02 We analyzed the top 10 Indian states and 3 top hot-spots along with the entire country for a period of close to 50 days to project the evolution of the COVID-19 pandemic using the 'Number of daily new cases' along with the total 'active' patients count.

03 The projected evolution of infected cases assumes an inherited effect of countrywide lockdown applied from 24th March 2020, and the numbers proposed for the future also consider that this period of lockdown continues for the next 45 days or so.

04 The projections in the study are primarily based on time series modeling. Therefore, we understand that this complex and dynamic nature of the influence of COVID-19 shall evolve daily and is a continuous work in progress.

05 The models developed making use of the existing standard epidemiological modeling techniques like SIR and SEIR are with the limitations of limited data, dynamic environmental factors, and external government based interventions.

06 The proposed SEIR model does not take into account the asymptomatic transmission explicitly; however, it cannot also be discarded entirely in the actual scenario.

** - as of 18th May, 06:09 GMT.*

07 Additionally, no stratified cohorts of the population are considered for studying this transmission-based COVID-19 pandemic. A later version of the model might include age-based or other stratifying features that may account for differing demographics in India.

08 We have taken the utmost care while collecting, preparing, and processing the required infected count data. However, the overall quality of data will have a material impact on the quality and reliability of outputs.

09 This work is ultimately a scenario-based modeling framework wherein Protiviti's analysts analyzed this pandemic situation and have come up with a robust statistical way to model it under several key assumptions. Therefore, the outputs of the study are not an indication of the future but rather probable situations for consideration and further deliberation.

10 This study does not take into account COVID-19 testing data in India and all the states, timing of testing, external factors like immunity, BCG vaccination, and international travel history data for a specific period, specific cluster density, demographics.

11 This predictive modeling considers the learning of progression rate from other countries like the US, Italy, China, Malaysia, South Korea, and incorporate in the models to create scenarios with certain assumptions.

12 The study depends on several assumptions that may vary at a granular level.

13 Lastly, the modeling here is regulated, assuming that the cases listed on trustworthy government sourced websites are an average over time. We completely understand the public health care system in India. We can reasonably assume that the number of cases considered in our study is on a lower side as many cases are not getting reported, and the testing rate of India is towards the lower side when compared globally.

Our Approach to Forecasting

Our complete modeling exercise is based on Time-Series Modeling, and SEIR model, wherein none of the models provide for any causal factors leading to the spread of COVID-19 in India. The reason for considering time as the independent variable for this study is two-fold:

01 *Firstly, the data on the count of daily infected cases is readily available in comparison to country/state wise data available for any other relevant factors*

02 *Secondly, modeling external causal factors like testing rate, demographic variables or climatic conditions requires extensive time and resource investments which of course wouldn't have helped us in coming out with these findings in time*

Some details on the models used for projecting the influence of Covid-19 in India and specific states/cities are presented below:

Percentage Based Model

- Apply the growth progression rate of the infected cases for USA and Italy on India data and project the cases using the same trends overtime to get the forecasted figures and no mathematical/statistical model has been applied
- Considering the infected cases as a function of the total population and use the percentage of Hubei to project the trend in India based on early data points available for India

Hybrid Model

- Segregate the data on two major segments – before lockdown and after lockdown, and two distinct models have been built for two sections considering the trends only and no other external factors have been taken into account
- Within lockdown, there are three stages- Progression rate, Flattening rate and Decay Rate : 3 different time series models have been fit separately based on the learning from China, Malaysia, and South Korea
- All the coefficients are based on Ordinary Least Square methodology, and adjusted R-Square is the primary metric considered for accuracy; R-Square is more than 90% for all the models

Polynomial Regression Models

- To overcome the problem of Bias-Variance tradeoff while selecting an appropriate and effective statistical model, we made use of Polynomial Regression models to model the COVID-19 infected Indians data.
- Polynomial Regression model assumes that the relationship between the dependent (Daily Count of Infected Patients) and independent variable (Time - Days) is curvilinear (specifically Polynomial)
- The coefficients of the proposed models are estimated using Ordinary Least Square (OLS) method while the model accuracy and performance is assessed using adjusted R-square values along with observing the Mean Squared Error (MSE) values
- Polynomial Regression of degree 2 and 3 provided for relatively a better fit along with an average forecasting accuracy of over 94% & 98% respectively

The Epidemic Modeling – SEIR MODEL

Based on the nuances and the complexity associated with limited pandemic data, we adopted a methodology of forecasting infected cases using the Susceptible Exposed Infected Recovered (SEIR) Model, mainly used for such pandemics both by researchers and academicians across the globe.

Susceptible Exposed Infected Recovered (SEIR) Model

In order to estimate the extent of the Covid-19 pandemic in India, we made use of the SEIR model proposed by Betten Court and Ribeiro[1] & Gani et al.[2]. This standard deterministic SEIR epidemic model classifies each individual into susceptible S(t), latent/exposed E(t), infective I(t) and recovered R(t) individual. Considering time as the independent variable, we made use of the 'desolve' package in R for solving the simultaneous differential equations representing the four probabilistic states of existence as mentioned above (S, E, I, R).

1. *Bettencourt, L. M. A. and Ribeiro, R. M., Real time Bayesian estimation of the epidemic potential of emerging infectious diseases. PLoSOne, 2008, 3(5), e2185.*
2. *S. R. Gani, Sk. TaslimAli and A. S. Kadi, The transmission dynamics of pandemic influenza A/H1N1 2009–2010 in India, CURRENT SCIENCE, 101(8) 25 Oct. 2011.*

SEIR model assumes that the infectious period, latent period, transmission probability and infection rate per person per unit time is dynamic and varies based on individuals and their respective location. The pandemic spread curve presents a transmission rate/infection rate and its evolution over a period of time and we estimated these parameters based on the actual available day wise infected count data from Ministry of Health & Family Welfare (MoHFW) website of India.

SEIR model is a class of Compartmental Models, *Figure 1*, which are primarily used for modeling spread of infectious disease. In such models, population is divided into homogeneous compartments with individuals in a cohort possessing similar characteristics. Such models are usually investigated using ordinary differential equations (ODEs) but in some cases are viewed as in stochastic framework too in order to model a more realistic scenario.

SEIR model used in our study, helps predict properties of how COVID-19 may spread in India, especially for the top 12 states, highlighting the total count of infected people and the duration of this pandemic.

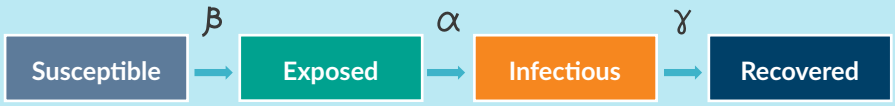


Figure 1: Susceptible Exposed Infectious Recovered (SEIR) Model Depiction

β – Infectious Rate, controls the rate of spread by representing the probability of transmitting disease between a susceptible person and an infectious individual

σ – Incubation Rate, suggests the rate of latent individuals becoming infectious

γ – Recovery Rate, provides the average duration (D = 1/ γ) of the infection

Dynamic Nature of Reproduction Number R_0

One of the most common measures of estimating the strength of an epidemic is the basic Reproduction Number (R0), which in simple words, suggests the number of secondary cases generated by one primary infected case in a susceptible population.

Our estimate of the reproduction number for India is 1.98 with 95% CI [1.95, 2.01], whereas our evaluation of the top 12 highly infected states of India shows a reproduction number that varies between 1.2 to 2.6. For our modeling, we made use of an R package by the name R0 wherein out of the multiple ways to estimate R0, the Maximum Likelihood method provided us with R0 estimate both for state/city wise and India as a whole.

We performed numerous simulations for predicting the influence of changing R_0 based on the following assumptions:

- Exposed population % varies from 30-65% of the total population of a state
- The influence period varies between 3-5 days
- Values of gamma and beta for estimating R_0 were also simulated to generate various scenarios
- The pandemic spread curve presents a transmission rate/infection rate and its evolution over a period of time

Some Insights into the Indian States

According to our study, some of the Indian states highly infected by COVID-19 include Maharashtra, Gujarat, Delhi, Rajasthan, Tamil Nadu, Madhya Pradesh seems that they would easily fall in the close range of 20K – 100K as a worst-case scenario, thereby raising some serious concerns in front of the health ministry to handle shortages of healthcare infrastructure for some of these states.

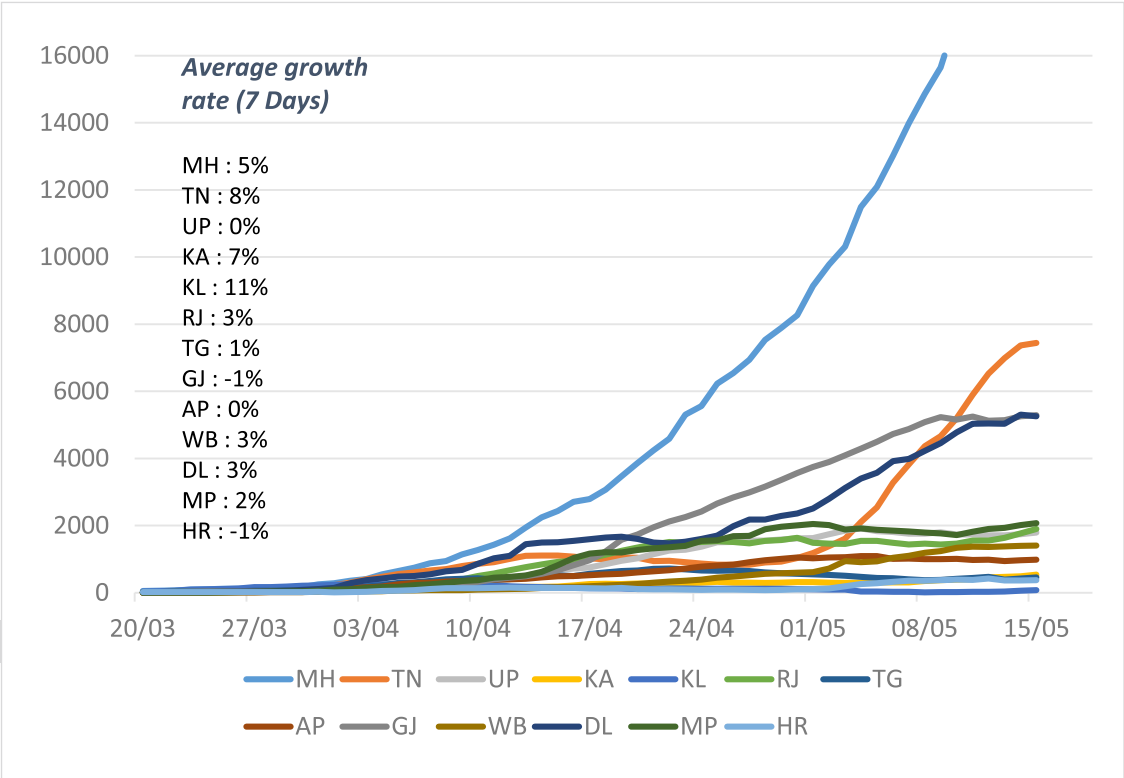
We expect that during the peaking time of this infectious disease in India, an expected 2,000 to 8,000 may be recorded daily in Maharashtra, Uttar Pradesh, Delhi, and Gujarat, which will also put stress on the existing resources in these states.

Our model suggests that all these highly probable infected states would hit their peak during the first fortnight of May and provides for the time in advance to prepare for such a crisis beforehand.

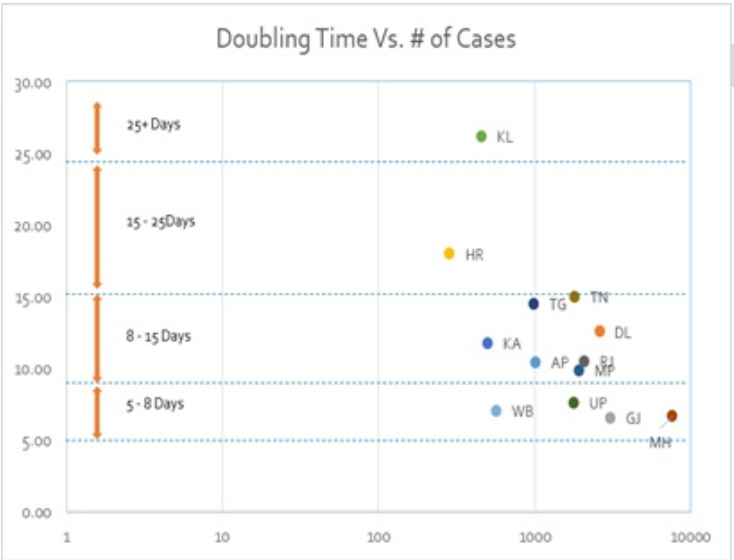
Places like Gujarat needs to pay special attention to creating arrangements for handling the massive shortfall of medical facilities.

Additionally, states like UP and Delhi may also drift quite quickly into a flagged zone and hence needs to be prepared in advance.

This section represents how some of the states has been doing in the last one month. We can see from the graphs below that some of the states have started showing of decline while some of the states show growth.



Graph 1



Graph 2

The first graphs talks about the growth of the number of active cases and 7 day moving average.

The second graph talks about the Doubling Rate or the number of days it took to double.

If we look at these information we can clearly see that there are 3 buckets of states

1. Where the growth has gone done considerably
2. Where the growth has stabilized and
3. Where there are still very high rate of growth

We can categorize Haryana (HR) and Kerala (KL) in the first bucket.

Karnataka (KA), Telangana (TG), Delhi (DL), Andhra Pradesh (AP), Rajasthan (RJ), Tamilnadu (TN) and Madhya Pradesh (MP) falls under the second bucket with a 7 day moving average ranging between 4%-8% and a Doubling Rate of 8-15 Days.

Finally, there are the high growth states, West Bengal (WB), Gujarat (GJ), Uttar Pradesh (UP) and Maharashtra where the growth rate is in 8% - 10% and the Doubling rate hovering around 6/7 days.

These states are further analyzed in details in the following sections.

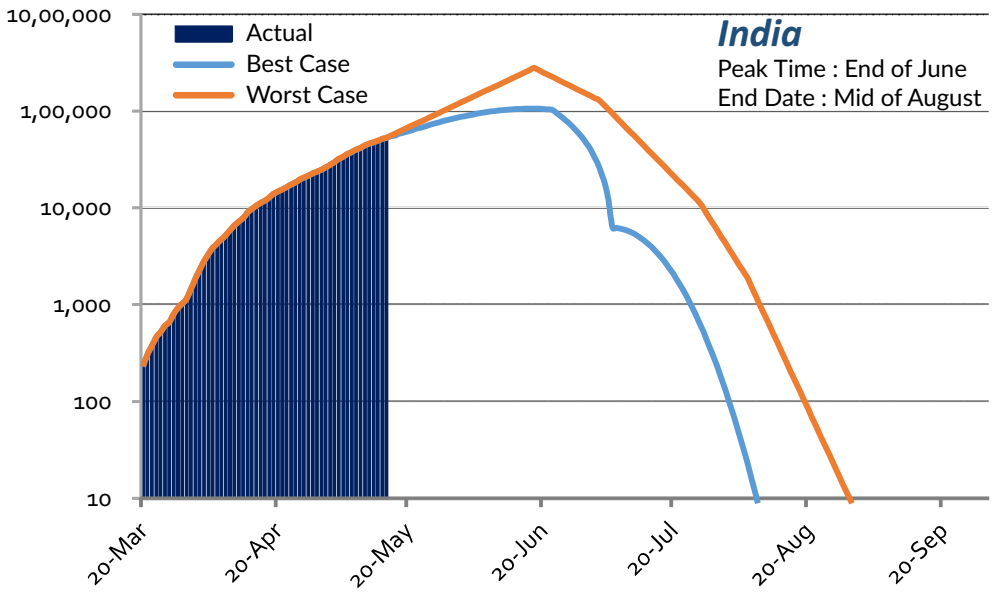
Sensitivity Analysis for Accurate Projections

Considering the dynamic nature of the COVID-19 pandemic and understanding that modeling the influence of this infection over a large population would always be quite complex and ever-evolving, we considered running a scenario analysis based on the simulated values of infected cases as per the changing R0 values. The objective of performing this analysis is to help understand how the number of infected cases could vary based on the period of lockdown.

Our present study, along with the literature, strongly suggests that with an increasing value of R0 based on time, the number of infected cases rises, whereas a decline in its value below 1.5 indicates that the situation typically may be getting relatively better. In the scenario analysis, we consider only a selected range of R0 values, and it changes daily for India.

Results

What will be the impact of the novel Coronavirus (COVID-19) in India? Answering this question requires an accurate forecast of the spread of confirmed cases as well as analysis of the number of deaths and recoveries. Forecasting, however, requires ample historical data; often, the correct data useful in forecasting is not available. At the same time, no prediction is precise as the future rarely repeats itself in the same way as the past. The graph is an aggregation of 3 types of forecasting model: % Based Models, Time Series model & SEIR Model.

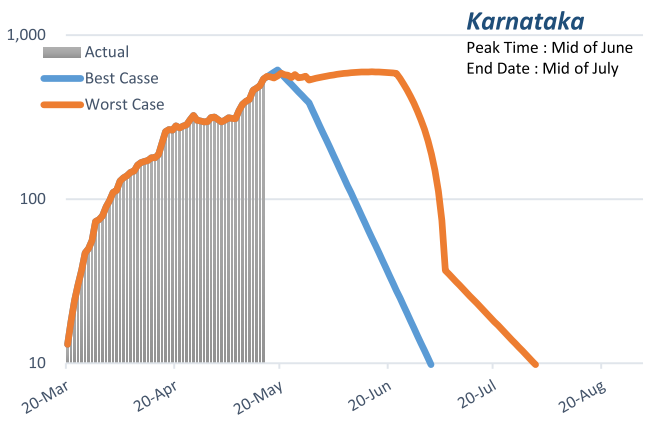


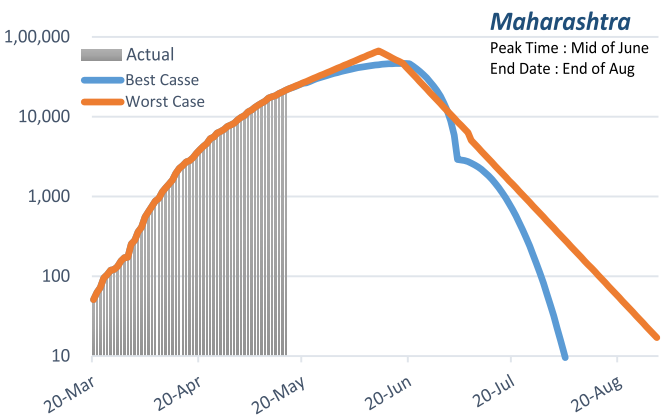
At its probable worst, COVID-19 might reach a peak of 2.8 lakh patients, of which approximately 30% may require ICU facilities. If the range is greater than 30% in total cases, this number soon becomes a devastating scenario.

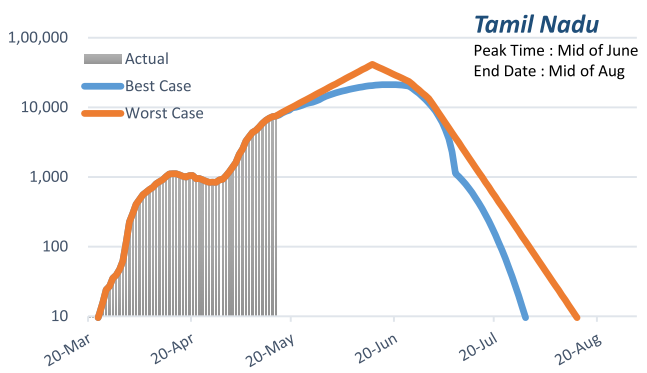
We expect that in the best-case scenario, India will reach about 1,05,000 cases, approximately 1,90,000 cases in the most likely scenario and 280,000 cases in the worst-case scenario. In the most likely scenario, we expect that India will be 95% free around the 1st week of August.

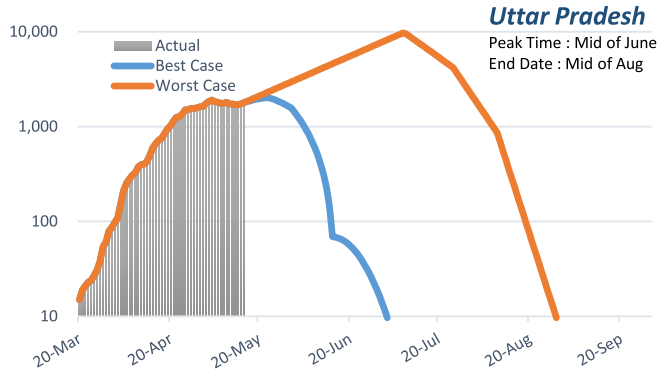
The above model is broken down into the below numbers for each of the top 10 affected states. The values identify the peaks in worst-case scenarios.

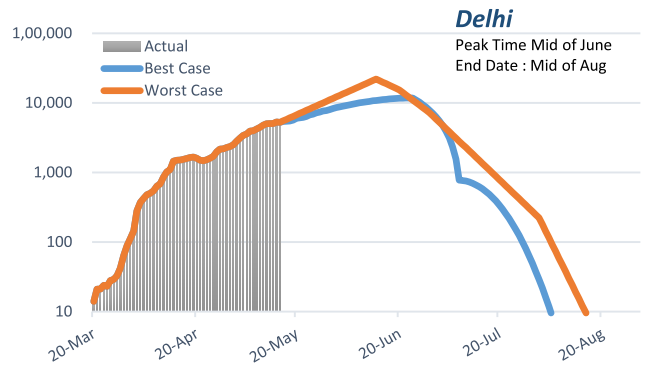
State and City Projections

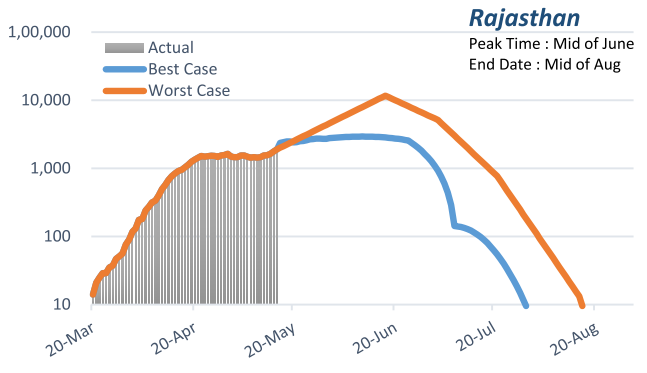




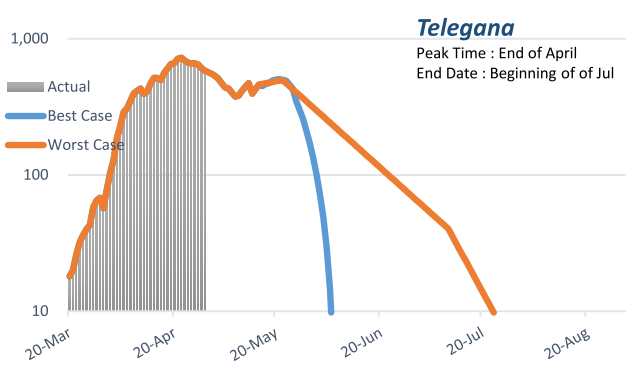


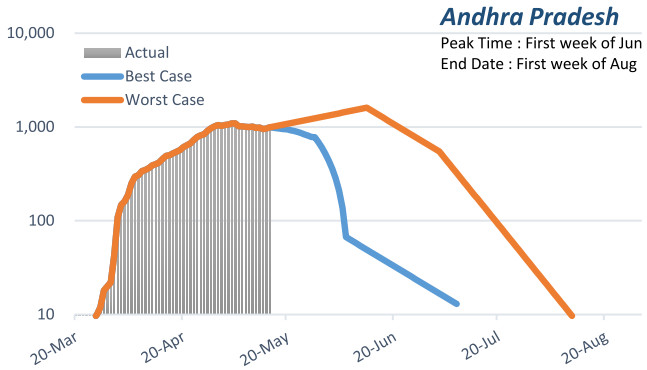


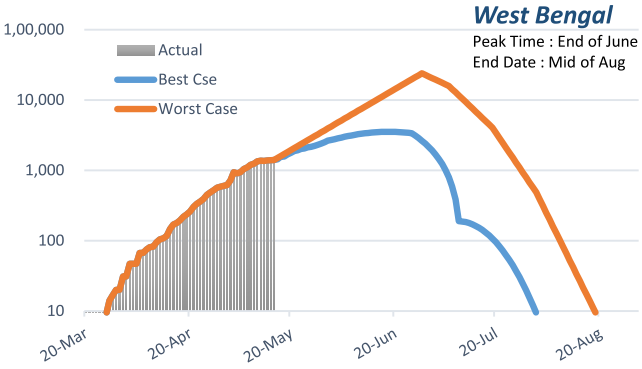


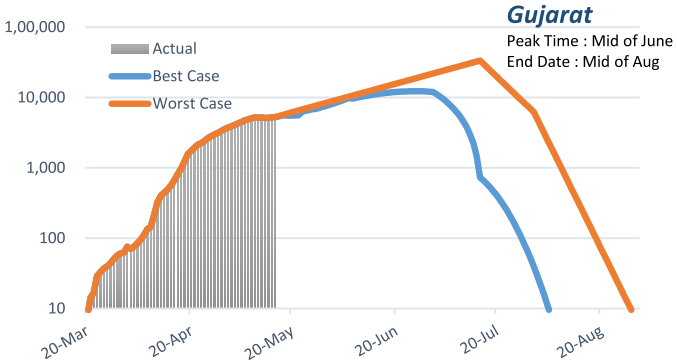


State and City Projections



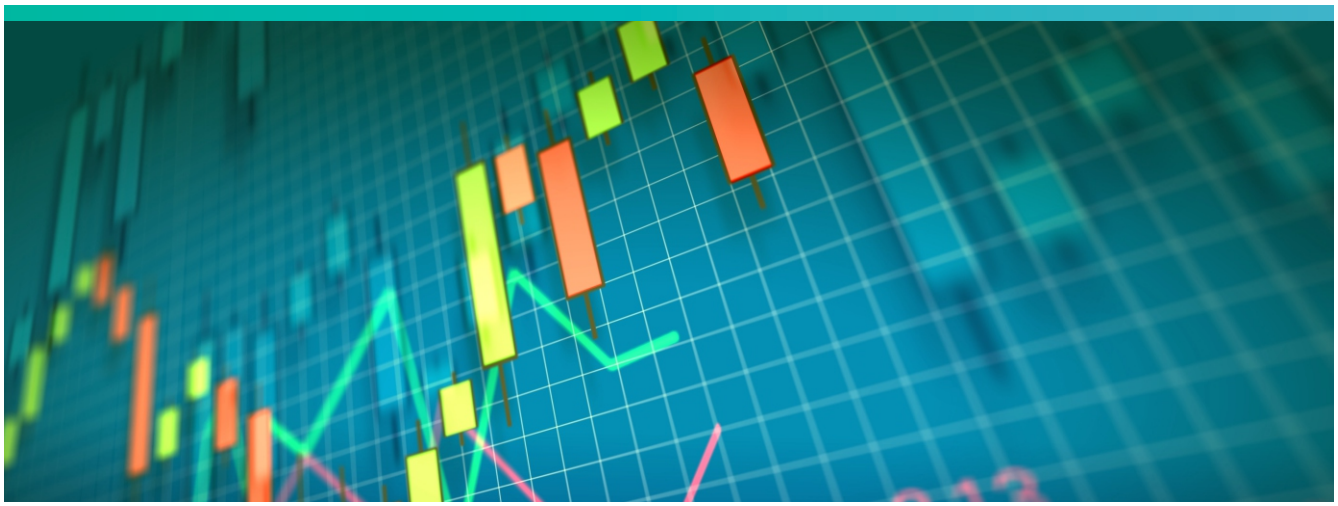






COVID-19 DISCLAIMER

The pandemic novel coronavirus 2019, is rapidly evolving, with new findings and insights being discovered daily. This document is a series of observations that represents Protiviti's viewpoint, at a certain point in time. This document does not constitute or claims to serve as an advisory for any medical, safety or regulatory action.



About TIMES NETWORK

TIMES NETWORK, a premium broadcast network is part of India's largest media conglomerate, The Times Group. Times Network with its credo **Now or Nothing** takes the centre stage of influencing the influencer, by delivering the most compelling and irreplaceable content to over 100 million urban affluent viewers in India and is available in over 100 countries across the globe. It houses best in class News, Hollywood and Bollywood brands including; **TIMES NOW** – India's No. 1 English News channel, **ET NOW** - India's leading English Business News channel, **MIRROR NOW** - the channel that Fights for 'You' and channelizes its focus on issues that impact people's lives; **MOVIES NOW** and **MOVIES NOW HD** – Home to Superheroes and India's destination for Hollywood Blockbusters; **MN+** – The Gold class of Hollywood; **Romedy NOW** and **Romedy NOW HD** – a unique destination for Love. Laugh and Live; **MNX** and **MNX HD** – Hollywood's wild child, **Zoom** - India's leading entertainment destination and **The Zoom Studios** – original content destination/creator that brings stories to life with relatable and compelling narratives.Timesnownews.com, the digital entity of the Network, has seamlessly adapted itself to serve the diverse nature and interests of the content-savvy internet audience.

About the Authors

Dhrubabrata Ghosh

Dhrubabrata Ghosh Dastidar, BE, Masters of Technology, PGDBM, leads the Data Analytics and Digital Transformation practice for Protiviti in India. He is a data evangelist with a focus on developing and architecting end to end data science solutions.

Dr. Nitin Sachdeva

Dr. Nin Sachdeva, Ph.D., Operational Research, is a seasoned advanced analytics professional with 14+ years of consulting, research, and academic experience in designing, developing and deploying data science and forecasting solutions.

Kallol Kumar

Kallol Kumar, M.Sc. (Statistics), MBA, is an analytics professional with 17+ years of experience in advanced analytics, data science and providing advisory to multiple clients across the globe.

About Protiviti

Protiviti (www.protiviti.in) is a global consulting firm that delivers deep expertise, objective insights, a tailored approach and unparalleled collaboration to help leaders confidently face the future. Through its network of more than 85 offices in over 25 countries, Protiviti and its independently owned Member Firms provide clients with consulting solutions in finance, technology, operations, data analytics, governance, risk and internal audit.

Named to the 2019 Fortune 100 Best Companies to Work For® list, Protiviti has served more than 60 percent of Fortune 1000® and 35 percent of Fortune Global 500® companies. The firm also works with smaller, growing companies, including those looking to go public, as well as with government agencies. In India, Protiviti's member firm (Protiviti India Member Pvt. Ltd.) is a leading provider of business consulting, internal audit, risk management, technology, forensic and fraud investigation, data analytics, digital transformation, and transaction services. Protiviti member firms are separate and independent legal entities, are not agents of other firms in the Protiviti network, and have no authority to obligate or bind other firms in the Protiviti network.

Protiviti is a wholly owned subsidiary of Robert Half (NYSE: RHI). Founded in 1948, Robert Half is a member of the S&P 500 index.

Contact Us:

Protiviti Member Firm

Amit Ray
Managing Director
Phone: +971 52896 4143
Email: amit.ray@protivitiglobal.me

Dhrubabrata Ghosh
Director
Phone: +91 9739 546 661
Email: dhrubabrata.ghosh@protivitiglobal.in

Times Network

Vivek Srivastava
President & Business Head English Entertainment Cluster
Times Network, Mumbai
Phone: +91 98922 20165
Email: srivastava.vivek@timesgroup.com

Jignesh Kenia
EVP and Head, Corporate Strategy and Business Development
Times Network, Mumbai
Phone: +91 98198 07552
Email: jignesh.kenia@timesgroup.com

protiviti®