Nowcasting and Forecasting the Spreading of Novel Coronavirus 2019-nCoV and its Association With Weather Variables in 30 Chinese Provinces: A Case Study

Preprint · February 2020

DOI: 10.13140/RG.2.2.11932.85124

2 authors:

N. Al-Rousan

27 PUBLICATIONS **95** CITATIONS

SEE PROFILE

Hazem Al-Najjar

27 PUBLICATIONS **68** CITATIONS

SEE PROFILE

Nowcasting and Forecasting the Spreading of Novel Coronavirus 2019-nCoV and its Association With Weather Variables in 30 Chinese Provinces: A Case Study

Nadia Al-Rousan*, Hazem Al-Najjar
Department of Computer Engineering
Faculty of Engineering and Architecture
Istanbul Gelisim university, Istanbul, Turkey
*Corresponding Author: nadia.rousan@yahoo.com

Summary

Background On 1st of February, 2020, 2019-noV coronavirus outbreak was announced to the public and it was classified as epidemic. Although the disease was discovered in Hubei province, China, but it was exported to all other Chinese provinces and spread globally. The novel coronavirus disease couldn't contained and it continued its spreading to all the world to exceed 34000 cases in main land China. Finding the environment effect on the epidemic spreading would help in understanding the growth of the disease. Besides, forecasting the size of infected, death, and recovered cases in China provinces would help to understand the domestic of this disease on global health.

Methods

2019-nCoV dataset from John Hopkins University that collected from World Health Organization [3], Chinese Center for Disease Control and Prevention (CDC)[22], and European Centre for Disease Prevention and Control [4]. The data covers the period from 22nd of January, 2020 to 4th of February, 2020. The published data consists of numeration to the number of confirmed cases, death, and recovered cases in all the infected regions globally. Data was filtered to select the infected provinces in China that considered as the first highest infected country in the world. To study the effect of environment and metrological variables on coronavirus disease spreading in all Chinese provinces, metrological data from the Global Forecast System (GFS) Web service that produced by the National Centers for Environmental Prediction (NCEP). The website delivers time series data for several metrological data from one month ago up to three days in future. Data on all Chinese provinces were obtained. Time series prediction using several forecasting models namely, Brown, Holt, Simple, Auto Regressive Integrated Moving Average (ARIMA) were used to forecast the expected confirmed, death, and recovered cases in china.

Findings

Studying the effect of weather conditions on coronavirus spreading found strong effect of weather on most of the Chinese provinces. None of the weather variables effect in all provinces where the effected variables vary from one province to another. It was found that the short wave radiation and temperature are the most effected variables, and the confirmed cases are the most effected cases of weather variables. On the other hand, this study estimated the number of confirmed, death, and recovered cases until 1st of September, 2020. It was found that the number of confirmed and death cases are growing faster than the recovered cases. The growth rate varies from one province to another based on weather conditions. The growth in some provinces is linear, while some of provinces have exponential growth.

Interpretation

China country and its provinces are the highest infected area by Coronavirus 2019-nCoV epidemic. The spreading of the disease in China was faster than other countries globally. Although the spreading of disease is centralized on Chinese cities, but the size of infected cases vary from one Chinese city to another which have different environmental variables and geographical nature. The effect of environment on the disease spreading probably would either infect other places have same environment or would increase the size of recovered cases. Studying different factors and attributes would help quickly to mitigate the disease.

Research in context

Evidence before this study

The causative of the novel coronavirus 2019-nCoV was identified as a zoonotic disease. The disease was centralized in Hubei province and other Chinese provinces. By 22nd of January, 2020, emergency outbreak was announced from World Health Organization (Who) to mitigate spreading this disease to other places outside China. It was announced that 444 cases were confirmed on 22nd of January 2020 besides hundreds of cases outside Hubei and spreading overseas. The disease had uncertain and unexpected behavior before that date. The studies which published in this issue were to discuss the effect of transmission in China and to outside China. None of the published works discussed the effect of environmental variables on spreading the disease on Chinese provinces. In addition, it is found very few research to forecast and estimate the number of probable confirmed, death, and recovered cases. While such researches would contribute to the field to understand the behavior of this epidemic.

Added value of this study

The relationship between fast spreading of novel coronavirus 2019-nCoV inside and outside China and other variables, factors, and attributes (i.e. environment, demography, geography, and transport, etc.) is not clear until this moment. The fears of uncontrollable coronavirus size is increased. The effect of environment on spreading the disease is discussed based on metrological variables inside China. Time series data that represent the infected cases are used to forecast the expected size in the future.

Implications of all the available evidence

The effect of environmental factors and attributes should be considered and adopted to recover the infected cases and mitigate the breakout of coronavirus disease. Besides, expected the probable infected cases in China should be considered to faster contain the spreaded disease.

Introduction

Recently, scientists have announced to the world their detecting to a novel decease that is improved edition of coronavirus. The novel coronavirus 2019-nCoV was identified as an outbreak of respiratory illness. On 31st of December 2019, Wuhan city in Hubei Chinese province has reported 27 pneumonia of unknown etiology [1]. The Wuhan Municipal Health Commission has announced about seven severe cases were detected have same respiratory illness [2]. These cases had several sypotoms (i.e. very dry cough, extremely high fever, and difficult dyspnea). Besides, examining their bodies tested positive for radiological findings of bilateral lung infiltrates [3-5]. On 9th of January 2020, Chinese Centre of Disease Control and Prevention (CCDC) has reported to the world the news of detecting the novel coronavirus 2019-nCoV in those sever cases [6]. In addition, from 15 to around 59 cases have been suspected to be infected by 2019-nCoV in Hubei [7]. On 22nd of January 2019, 2019-nCoV coronavirus outbreak was announced to the public and it was classified as epidemic, and all data about the illness travelling was shared [8]. However, the announcement from CCDC about detecting the novel 2019-nCov coronavirus caused announcing the emergency cases from World Health Organization (WHO) [9]. This is due to confirmed several cases globally besides the concentration of this novel virus in Chinese provinces. China or as commonly named People's Republic of China (PRC) is a country located 35.8617° North and 104.1954° East in continent of Asia. It consists of more than 32 provinces including the special administrative regions (i.e. Hong Kong, Inner Magnolia, and Macau). By time, the number of confirmed infected cases has been rapidly increased especially in Hubei and other Chinese provinces. By the end of 5th of February 2020, 24303 cases were confirmed in China country while several other cases were confirmed globally as shown in Figure 1 that represents the distribution of confirmed cases in Chinese provinces [10].

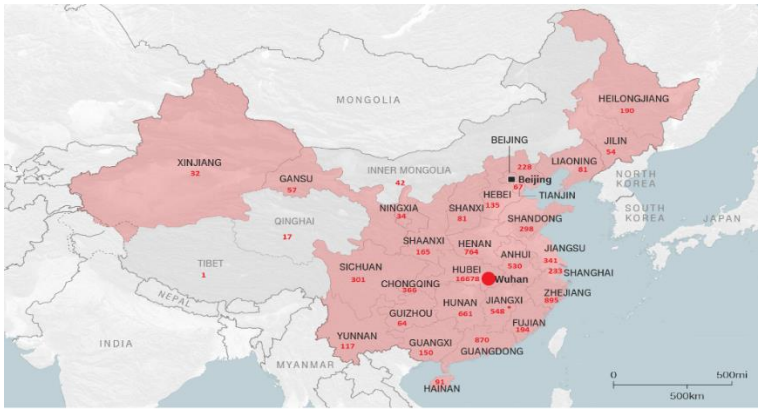

Figure 1. Distribution of confirmed coronavirus cases in China.

As shown in China map, Hubei province was reported as the highest infected area. Figure 2 shows the number of confirmed cases in Hubei from 22/1/2020 to 5/2/2020.

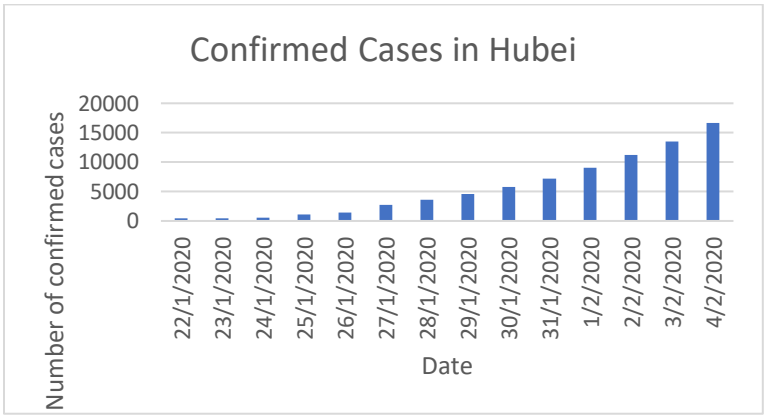

Figure 2. Daily change in the number of coronavirus confirmed cases in Hubei

As shown in Figure 1, the number of confirmed cases has been increased exponentially starting from around 444 cases in 22/1/2020 to reach 16678 in 4/2/2020. On the other hand, it is clear that the number of infected and confirmed cases vary from one province to another as shown in Figure 3.

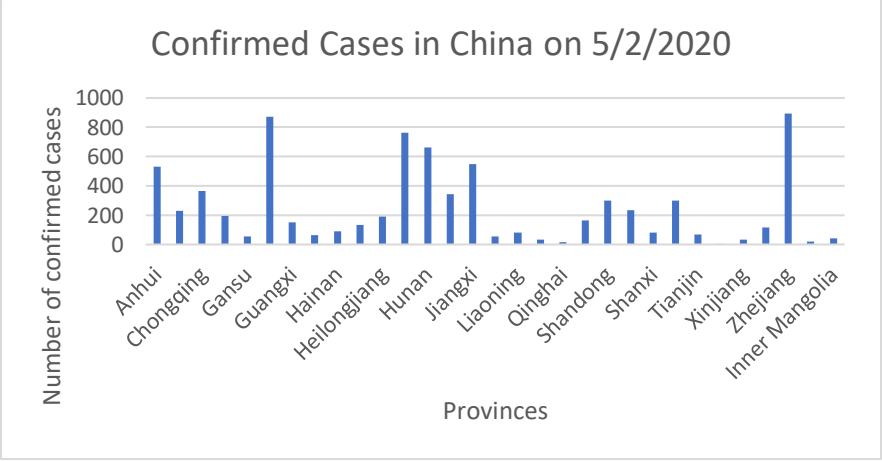

Figure 3. Variation of number of confirmed cases in China provinces

Figure 3 shows that the second high infected province is Zhejiang then Guangdong provinces with 895 and 870 cases respectively. While Tibet and Qinghai the special administrative provinces has reported one and 17 confirmed case respectively, which represent the lowest number of cases compared to other provinces.

However, Several researches have focused on analyzing the medical conditions of coronavirus 2019-nCoV illness, and the causes to spread the disease very fast globally and in China provinces as a case study [11-12]. It is questionable that some provinces effected by coronavirus more than other provinces. In addition, it is clear from China map that the variation in the number of confirmed cases does not related with distance between these provinces that some far provinces were infected faster than other close provinces. On the other hand, studying the most populous Chinese provinces that shown in Figure 4 revealed that Guangdong, which is the second infected province, is the first top most populous province as well. While it is not the case for some other provinces (i.e. Zhejiang, Shandong, and Hunan). It shown that Hubei is the ninth populous province while it is the first infected province. These results revealed that the infection is not related to the people density.

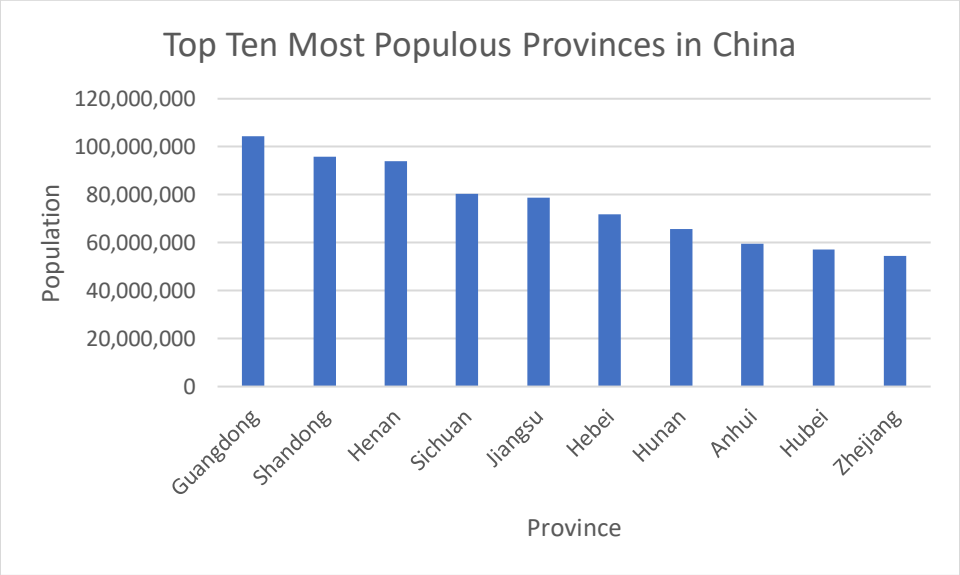

Figure 4. Top Ten Most Populous Provinces in China

Several researches were conducted to discover the reason to cause such disease. In the late of 2019, a Wuhan's Huanan Seafood Wholesale Market was reported as a first causative to the novel coronavirus. The main reason to report such claim was selling several strange live animals (i.e. fishes, bats, and snakes, etc.)[13]. Unfortunately, samples from the market tested positive for 2019-nCoV. Thus, the market was closed on 1st of January 2020 [14]. However, although the market was closed to the public, and the aetiological cause was identified, but the disease could not contained in Hubei and also continuing its spreading through China and outside China globally [15]. Therefore, other factors and attributes may are the main causable to the disease spreading (i.e. Geographical nature, environment, air and train transport, population, etc.).

By the end of January, several researches are published to analyze some probably reasons to cause 2019-nCoV coronavirus, or to estimate and predict the probably number of infections in the future [16-18]. These researches aimed to mitigate the spreading of the dangerous non stopped virus.

Joseph et al, [19] have provided a nowcast and forecast to the probable size of coronavirus. Besides to the probable course of spread coronavirus globally. On the other hand, this study have provided the effect of the social and personal potential impact that have been progressively and quickly implemented in January, 2020 on the disease spreading. The global flight bookings data from the Official Aviation Guide (OAG) was purchased to study the sequence of disease spreading outside Hubei. The results revealed how the disease would be very dangerous and killer globally

unless several potential impacts are implemented (i.e. limit the transportation to and from the infected areas, and reduce within-population contact rates, etc.).

Several other researches have used different Artificial Intelligence (AI) learning techniques to estimate the probable spreading size of coronavirus [20]. While other researches have focused on the main Genomic characterisation and epidemiology of the infected cases by coronavirus disease or the characterization of the 2019-nCoV itself [21]. However, the fast spreading inside and outside china is interesting and important case to study. Evedintly, some research questions should be answered as well to mitigate the spreading of coronavirus 2019-nCoV. These questions mainly about the main causable attributes to spread 2019-nCoV epidemic. In addition, it is important to know if the variation in geographical nature and weather variables related to widely spreading the novel coronavirus in some provinces rather than others. Besides, to study the relationship between temperatures, relative humidity, pressure, wind speed, wind direction, rainfall rate, snowfall rate, snow depth, and shortwave irradiation variables and the spreading of 2019-nCoV coronavirus. The aims of this research is to contribute to the current available sources and information about several causative of novel coronavirus 2019-nCoV spreading. The main target is to study the effect of each metrological variable in China country on the size of coronavirus disease, and to find their effects on confirmed, death, and recovery cases in each Chinese provinces. Then, to forecast the probable confirmed, death, and recovery cases in china provinces in the future based on different forecasting methods. To the best of authors' knowledge, no research was published to discuss the effect of the variation of environmental variables and geographical nature on the spreading of the novel 2019-nCoV coronavirus epidemic. This study would help other researchers to mitigate the spreading of coronavirus disease, or in increasing the number of recovered cases by preparing the appropriate environment variables for their recovery. In addition, to select the optimum forecasting models to estimate the worst and best scenarios in coronavirus spreading.

Methods

In this study, the first step is to study the effect of several environmental variables namely, temperatures, relative humidity, pressure, wind speed, wind direction, rainfall rate, snowfall rate, snow depth, and shortwave irradiation on coronavirus 2019-nCoV confirmed, death, and recovered cases China. The effects of these metrological attributes on coronavirus disease in each province were studied to highlight the causative of spreading the disease in some provinces rather, faster, and more prevalent than other provinces. The second step is to forecast the expected number of confirmed, death, and recovered cases in all Chinese provinces after 210 days for both best and worst cases. Several forecasting methods were used to accurate the expected results.

Data sources and assumptions

The daily metrological data for all Chinese provinces excluding Inner Mangolia and Hong Knog was obtained. Time series metrological data from the Global Forecast System (GFS) Web service since 22nd of January, 2020 to 4th of February, 2020 were obtained. Different nine environmental variables were adopted to complete this study including delivers time series temperature at two meters above the ground in (Kelvin) unit, relative humidity at two meters above the ground in percent (%), pressure at ground level in Hectopascal (hPa), wind speed at 10 meters above the ground in (m/s), wind direction at 10 meters above ground in degrees (i.e. 0 means from North, and 90 from East), rainfall rate in (kg/m^2), snowfall rate in (kg/m^2), snow depth in meter, and Surface downward short-wave irradiation in (watt hour/m^2).

While a 2019-nCoV dataset from John Hopkins University was used to forecast the infected, death, and recovered cases. The data also covered the period from 22nd of January, 2020 to 4th of February, 2020. The numeration of confirmed cases, death, and recovered cases in all Chinese provinces excluding both Inner Mangolia and Hong kong was used to forecast the size of the disease after 210 days.

Finding environment effect on 2019-nCoV in China

Correlation analysis wass used to find the effect of the environmental variables on the enumeration of infected or suspected cases including confirmed, death, and recovered cases. Several environmental variables including temperatures, relative humidity, pressure, wind speed, wind direction, rainfall rate, snowfall rate, snow depth, and shortwave irradiation along with day were adopted as input variables, while the output is the number of confirmed, death, and recovered cases separately. Pearson correlation coefficient was adopted to find the relationship between environmental variables and the size of confirmed, death, and recovered cases in all Chinese provinces separately. The magnitude of the Pearson correlation coefficient indicates the strength of the relationship between considered variables depends on how the coefficient is close to -1 or 1, while the sign of the correlation coefficient indicates direction of the relationship between these variables.

Pearson correlation coefficient can be calculated by using Equation (1) [23].

$$R = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} \qquad (1)$$

where, n is the number of samples, x_i, y_i are the single samples indexed with i, and $\bar{x}, \bar{y}$ are the means of samples.

Forecasting the spread of 2019-nCoV in China within 210 days

Day variable along with the current available dataset for enumerated number of confirmed, death, and recovered cases were adopted to forecast the expected size of confirmed, death, and recovered cases during the coming seven months (210 days) starting from 5th of February, 2020. Different four forecasting methods that represent the mostly used models were employed to estimate the results for each province separately, and the results for the optimum model in each province is considered. Brown, Holt linear trend model, Simple, and Autoregressive Moving Average (ARIMA) models are employed in this study to forecast the infected cases.
Brown model is a linear exponential smoothing model that depends on two different smoothed series that are centered at different points in time. Holt linear trend model is another form of linear exponential smoothing models as well. This model is able to allow forecasting data with a trend based on three equations one fore forecasting and two for smoothing. On the other hand, simple exponential smoothing (SES) method is the simplest model for forecasting and suitable for forecasting data that have no clear trend or seasonal pattern. While ARIMA model is another exponential smoothing model to provide complementary approaches and describe the autocorrelations in the data [24].

Results

Date and weather correlation analysis

To study the impact of weather and date variables on Deaths, confirmed, and recovered variables, a correlation analysis is adopted as shown in Table 1. The results showed that date variable (day, month, and year) is significant with strong correlation against strong correlation all studied variables except Deaths variables (in case of Fujian, Hebei, and Shandong), and recovered variable (in case of Ningxia and Shaanxi). Temperature is significant with Auhui, Hubei, Jilin (confirmed), Hunan (confirmed), Liaoning(recovered), Ningxia, Shanxi(confirmed), Tianjin(recovered), and Tibet. Humidity showed a negative correlation with Auhui, Hubei, Ningxia, Shanghai, Shanxi, where pressure showed a positive movement with Fujian, Guangdong, Guangaxi, Hainan. Wind speed has negative correlation and strong movement with Jiangsu only, whilw wind direction indicated a negative sign with Hebei (Deaths cases) and positive sign with Sichuan (recovered cases). Rainfall showed strong positive movement with Fujian (Deaths cases) and negative movement with Shangai (confirmed and recovered). Snowfall did not show any significant correlations, where snow depth showed negative correlation with Jilin (recovered) and Ningxia (confirmed). Short wave irradiation

has positive correlation with Auhui, Beijing, Gansu, Henan, Hubei, Jiangsu, Shanghai, Sichuan and Tibet. The results indicated that the weather variables have a small effect on spreading corona virus and no prove can be extracted between the impact of weather and Deaths, Recovered, and Confirmed case in all provinces. In addition, weather variables have variation effects on each province as shown in Table 2. The results indicated that weather variables have effect on Auhui, Fujian, Gansu, Guangdong, Guangxi, Hainan, Hebei, Henan, Hubei, Hunan, Jiangsu, Jilin, Liaoning., Ningxia, Shanghai, Sichuan, Shanxi, Tianjin and Tibet only, where the rest of studied provinces do not have impacts on weather variables.

Besides, the results revealed that the date variable is the major variable that effected strongly on the movement of Corona virus in all Chinese's provinces which gives an indicator that the growth of deaths and confirmed cases will be increased rapidly within the couple of months. To estimate the growth speed of each case in different provinces, different forecasting models are used as discussed in the following section.

Table 1: Correlation Analysis between independent and dependent variables.

State	Variables		Days	Temp	Humidity	Pressure	Wind Speed	Wind direction	Rainfall	Snowfall	Snow depth	Short Wave irradiation
Auhui	Confirmed	Corr	.974**	.615*	-.721**	-.111	-.455	-.188	-.534*	-.219	-.216	.835**
		Sig.	.000	.019	.004	.705	.103	.520	.049	.452	.459	.000
	Recovered	Corr	0.811	0.653	-.486	-.014	-.401	-.277	-.383	-.186	-.182	0.622
		Sig.	.000	.011	.078	.961	.155	.337	.177	.525	.533	.018
Beijing	Confirmed	Corr	.991**	-.371	-.541*	-.118	.439	.401	.405	.364	-.131	.543*
		Sig.	.000	.191	.046	.689	.116	.156	.151	.201	.655	.045
	Deaths	Corr	.832**	.094	-.157	-.536*	.281	.355	.259	.252	.113	.267
		Sig.	.000	.749	.592	.048	.331	.213	.372	.385	.700	.356
	Recovered	Corr	.858**	-.650*	-.460	.130	.343	.172	.208	.195	-.186	.599*
		Sig.	.000	.012	.098	.659	.229	.557	.475	.505	.525	.024
Chongqing	Confirmed	Corr	.996**	-.099	-.044	.246	-.594*	.136	.056			.016
		Sig.	.000	.737	.882	.397	.025	.642	.850			.956
	Deaths	Corr	.842**	-.185	.287	.216	-.450	-.206	.358			-.268
		Sig.	.000	.527	.320	.459	.106	.481	.209			.355
	Recovered	Corr	.796**	-.078	.112	.227	-.454	-.208	.198			.005
		Sig.	.001	.792	.702	.436	.103	.476	.498			.987
Fujian	Confirmed	Corr	.990**	-.493	.018	.851**	-.575*	-.168	-.292			-.021
		Sig.	.000	.073	.950	.000	.031	.567	.310			.942
	Deaths	Corr	-.241	.077	.328	-.299	.346	.293	.775**			-.333
		Sig.	.407	.793	.253	.299	.226	.309	.001			.244
	Recovered	Corr	.557*	.008	.436	.257	-.213	-.414	-.034			-.373
		Sig.	.039	.977	.119	.374	.464	.141	.909			.189
Gansu	Confirmed	Corr	.992**	.283	-.612*	-.155	.342	.328	.211			.750**
		Sig.	.000	.350	.026	.613	.253	.274	.489			.003
	Recovered	Corr	.743**	.560*	-.571*	-.162	.206	.174	-.156			.807**
		Sig.	.004	.046	.042	.597	.499	.569	.611			.001
Guangdong	Confirmed	Corr	.972**	-.267	.095	.723**	-.392	-.495	.129			-.339
		Sig.	.000	.355	.746	.003	.165	.072	.660			.236
	Recovered	Corr	.902**	-.201	.206	.600*	-.364	-.458	.116			-.393
		Sig.	.000	.490	.480	.023	.200	.100	.693			.165
Guangxi	Confirmed	Corr	.994**	-.324	-.130	.722**	-.552*	-.387	.010			-.054
		Sig.	.000	.258	.658	.004	.041	.171	.973			.854
	Recovered	Corr	.795**	-.111	.080	.487	-.365	-.279	-.087			-.062
		Sig.	.001	.706	.787	.077	.199	.335	.768			.832
Guizhou	Confirmed	Corr	.924**	.269	.241	.238	-.459	.213	.030	-.333	-.287	-.276
		Sig.	.000	.352	.407	.413	.099	.465	.919	.245	.320	.340
	Recovered	Corr	.850**	.447	.014	.159	-.336	.258	-.080	-.304	-.274	-.082
		Sig.	.000	.109	.962	.587	.240	.372	.785	.291	.343	.779
Hainan	Confirmed	Corr	.994**	-.110	-.063	.642*	.373	-.044	-.152			.164
		Sig.	.000	.708	.830	.013	.189	.881	.603			.576
	Deaths	Corr	.832**	-.462	.058	.771**	.612*	.004	-.012			-.042
		Sig.	.000	.096	.843	.001	.020	.990	.967			.888
	Recovered	Corr	.825**	.325	-.177	.363	-.132	.028	-.295			.365
		Sig.	.000	.256	.546	.203	.653	.925	.307			.200
Hebei	Confirmed	Corr	.988**	-.400	-.329	-.213	.376	-.034	.215	.258	.438	.544*

		Sig.	.000	.156	.250	.464	.185	.908	.460	.374	.117	.044
	Deaths	Corr	.447	-.038	.084	.286	.355	-.662**	.090	.077	.077	.003
		Sig.	.109	.897	.775	.321	.213	.010	.759	.794	.794	.993
	Recovered	Corr	.793**	-.735**	-.148	.195	.225	-.004	.356	.387	.573*	.457
		Sig.	.001	.003	.613	.504	.438	.989	.212	.172	.032	.100
Heilongjiang	Confirmed	Corr	.923**	-.530	.161	-.504	-.122	.304	-.199	.135	.116	.278
		Sig.	.000	.051	.583	.066	.678	.290	.495	.646	.693	.336
	Deaths	Corr	.906**	-.159	-.118	-.547*	-.201	-.007	-.342	.210	-.296	.442
		Sig.	.000	.586	.689	.043	.491	.982	.231	.472	.305	.114
	Recovered	Corr	.769**	-.651*	.400	-.434	.001	.519	-.119	.259	.413	.000
		Sig.	.001	.012	.157	.121	.996	.057	.686	.372	.142	1.000
Henan	Confirmed	Corr	.977**	.448	-.415	-.316	.274	.214	-.422	-.403		.644*
		Sig.	.000	.108	.140	.271	.343	.462	.133	.154		.013
	Deaths	Corr	.885**	.400	-.556*	-.503	-.025	.522	-.511	-.478		.753**
		Sig.	.000	.157	.039	.067	.934	.056	.062	.084		.002
	Recovered	Corr	.736**	.223	-.210	-.044	.482	-.087	-.237	-.236		.493
		Sig.	.003	.443	.470	.880	.081	.767	.416	.417		.073
Hubei	Confirmed	Corr	.951**	.799**	-.798**	-.365	-.510	-.107	-.453	-.247	-.240	.701**
		Sig.	.000	.001	.001	.199	.063	.717	.104	.394	.409	.005
	Deaths	Corr	.976**	.796**	-.827**	-.354	-.512	-.126	-.493	-.241	-.220	.743**
		Sig.	.000	.001	.000	.214	.061	.667	.073	.406	.449	.002
	Recovered	Corr	.896**	.748**	-.697**	-.283	-.414	-.117	-.364	-.195	-.292	.612*
		Sig.	.000	.002	.006	.327	.141	.691	.200	.504	.312	.020
Hunan	Confirmed	Corr	.980**	.693**	-.311	.122	-.618*	-.474	-.453	-.251		.335
		Sig.	.000	.006	.279	.679	.019	.087	.104	.387		.242
	Recovered	Corr	.782**	.562*	.020	.005	-.306	-.303	-.251	-.165		.114
		Sig.	.001	.037	.945	.986	.288	.292	.386	.572		.698
Jiangsu	Confirmed	Corr	.980**	-.170	-.619*	.209	-.718**	.101	-.269			.775**
		Sig.	.000	.562	.018	.474	.004	.732	.353			.001
	Recovered	Corr	.884**	.086	-.526	.324	-.680**	-.130	-.206			.638*
		Sig.	.000	.770	.053	.258	.007	.659	.479			.014
Jiangxi	Confirmed	Corr	.961**	.485	-.361	.429	-.529	-.252	-.381			.361
		Sig.	.000	.079	.205	.126	.052	.385	.180			.205
	Recovered	Corr	.898**	.500	-.200	.347	-.417	-.286	-.320			.215
		Sig.	.000	.069	.492	.224	.138	.322	.265			.460
Jilin	Confirmed	Corr	.907**	-.678**	.246	-.141	.273	.297	.042	-.165	-.558*	.491
		Sig.	.000	.008	.396	.630	.346	.302	.887	.572	.038	.074
	Recovered	Corr	.868**	-.580*	.477	-.463	.213	.178	.005	-.209	-.696**	.412
		Sig.	.000	.030	.084	.095	.465	.542	.988	.474	.006	.143
Liaoning	Confirmed	Corr	0.993	-0.623	-0.211	-0.16	0.048	0.447	0.391	0.347	.	0.248
		Sig.	.000	0.017	0.468	0.586	0.87	0.109	0.166	0.224	.	0.393
	Recovered	Corr	0.882	-0.786	-0.216	-0.038	0.231	0.533	0.219	0.243	.	0.332
		Sig.		0.001	0.458	0.898	0.427	0.05	0.451	0.402	.	0.246
Ningxia	Confirmed	Corr	.975**	.810**	-.805**	-.367	-.062	-.065	-.228	-.292	-.773**	.544*
		Sig.	.000	.000	.001	.197	.833	.825	.433	.312	.001	.045
	Recovered	Corr	.628*	.930**	-.664**	-.626*	-.222	-.162	-.189	-.190	-.464	.514
		Sig.	.016	.000	.010	.017	.445	.579	.518	.514	.095	.060
Qinghai	Confirmed	Corr	.967**	.370	-.636*	.221	.238	-.062	-.093	-.012	-.566*	.658*
		Sig.	.000	.193	.015	.448	.412	.833	.752	.968	.035	.011
Shaanxi	Confirmed	Corr	.982**	.653*	-.513	-.431	-.093	.127	-.004	.236	.027	.569*

		Sig.	.000	.011	.061	.124	.752	.665	.990	.417	.926	.034
	Recovered	Corr	.588*	.537*	-.389	-.342	.160	-.094	-.202	-.123	-.242	.433
		Sig.	.027	.048	.169	.232	.585	.748	.489	.675	.404	.122
Shandong	Confirmed	Corr	.995**	-.379	-.084	-.131	-.012	.351	.251	.141		.452
		Sig.	.000	.181	.776	.654	.967	.219	.386	.631		.104
	Deaths	Corr	-.103	-.065	-.176	-.369	-.027	.350	-.183	-.147		-.042
		Sig.	.726	.827	.548	.195	.928	.220	.531	.615		.886
	Recovered	Corr	.835**	-.320	.057	.213	.036	-.064	.332	.256		.276
		Sig.	.000	.265	.848	.464	.902	.828	.247	.377		.340
Shanghai	Confirmed	Corr	.990**	.078	-.783**	.459	-.597*	.000	-.685**			.854**
		Sig.	.000	.790	.001	.099	.024	1.000	.007			.000
	Deaths	Corr	.713**	-.443	-.549*	-.097	-.033	.171	-.505			.534*
		Sig.	.004	.113	.042	.741	.911	.559	.066			.049
	Recovered	Corr	.969**	-.021	-.871**	.475	-.607*	.123	-.751**			.903**
		Sig.	.000	.944	.000	.086	.021	.676	.002			.000
Sichuan	Confirmed	Corr	.992**	.187	-.413	.289	-.035	.480	-.334	-.450	-.286	.757**
		Sig.	.000	.522	.142	.316	.905	.082	.243	.107	.322	.002
	Deaths	Corr	.868**	.249	-.206	.330	.124	.308	-.261	-.349	-.294	.714**
		Sig.	.000	.390	.479	.249	.673	.284	.367	.221	.307	.004
	Recovered	Corr	.760**	.248	-.411	.433	-.202	.731**	-.256	-.145	.108	.560*
		Sig.	.002	.393	.144	.122	.489	.003	.378	.621	.714	.037
Shanxi	Confirmed	Corr	.988**	.738**	-.819**	-.397	.486	.375	.090	-.100	-.723**	.641*
		Sig.	.000	.003	.000	.160	.078	.186	.759	.735	.003	.014
	Recovered	Corr	.848**	.475	-.721**	-.135	.381	.115	-.114	-.194	-.476	.657*
		Sig.	.000	.086	.004	.646	.179	.695	.699	.505	.085	.011
Tianjin	Confirmed	Corr	.977**	-.637*	-.229	-.088	.310	-.058	.237	.234	.669**	.397
		Sig.	.000	.014	.430	.765	.281	.843	.415	.421	.009	.160
	Recovered	Corr	.692**	-.880**	-.106	.268	.359	-.044	-.122	-.132	.539*	.364
		Sig.	.006	.000	.719	.354	.207	.882	.678	.652	.047	.200
Tibet	Confirmed	Corr	0.912	0.799	-0.731	-0.081	-0.411	-0.563	0.288	0.189	-0.972	0.835
		Sig.	0.001	0.01	0.025	0.835	0.272	0.114	0.452	0.626	0	0.005
Xinjiang	Confirmed	Corr	.987**	.310	-.015	.364	-.121	.218	-.426			-.111
		Sig.	.000	.280	.959	.200	.679	.454	.128			.705
	Recovered	Corr	.814**	.454	-.409	.328	-.026	.262	-.333			.309
		Sig.	.000	.103	.146	.252	.930	.366	.245			.282
Yunnan	Confirmed	Corr	.987**	.310	-.015	.364	-.121	.218	-.426			-.111
		Sig.	.000	.280	.959	.200	.679	.454	.128			.705
	Recovered	Corr	.814**	.454	-.409	.328	-.026	.262	-.333			.309
		Sig.	.000	.103	.146	.252	.930	.366	.245			.282
Zhejiang	Confirmed	Corr	.985**	-.138	-.500	.650*	-.348	-.275	-.541*			.376
		Sig.	.000	.638	.068	.012	.223	.341	.046			.186
	Recovered	Corr	.857**	.104	-.150	.579*	-.474	-.386	-.335			.034
		Sig.	.000	.723	.608	.030	.087	.172	.241			.909

*Correlation is significant at the 0.01 level (2-tailed).

**Correlation is significant at the 0.05 level (2-tailed).

Table 2: The most effected weather variables with movement direction on each province

Province	Variable	Effected variables and directions
Auhui	Confirmed	+Short wave irradiation (IR), -Humidity
Fujian	Confirmed	+Pressure
	Deaths	+Rainfall
Gansu	Confirmed	+IR
	Recovered	+IR
Guangdong	Confirmed	+Pressure
Guangxi	Confirmed	+Pressure
Hainan	Deaths	+Pressure
Hebei	Deaths	+Wind direction
Henan	Deaths	+IR
Hubei	Confirmed	+ Temperature, -Humidity, +IR
	Deaths	+ Temperature, -Humidity, +IR
Hunan	Confirmed	+Temperature
Jiangsu	Confirmed	-Wind speed, +IR
	Recovered	-Wind speed
Jilin	Confirmed	-Temperature
	Recovered	-Snow depth
Liaoning	Recovered	-Temperature
Ningxia	Confirmed	+Temperature, -Humidity, +Snow depth
	Recovered	+Temperature, -Humidity
Shanghai	Confirmed	-Humidity, -Rainfall, +IR
	Recovered	-Humidity, -Rainfall, +IR
Sichuan	Confirmed	+IR
	Deaths	+IR
	Recovered	+Wind direction
Shanxi	Confirmed	+Temperature, -Humidity, -Snow depth
	Recovered	-Humidity
Tianjin	Confirmed	+Snow depth,
	Recovered	-Temperature
Tibet	Confirmed	+Temperature, -snow depth, +IR

Forecasting Deaths, Confirmed, and Recovered of Corona virus in China

Figures 5 to 8 showed forecasting of confirmed, deaths, recovered variables of different Chinese's provinces up to 210 days until 1st of September 2020. For brevity, only the forecasting results are presented and model performance metrics including R2 and erro functions are eliminated and will be available in the separated files with data. The results are built based on the best forecasting models as shown in Table 3. The results showed that forecasting deaths variable for Fujian, Gansu, Guangdong, Guangxi, Guizhou, Hebei, Hunan, Jiangsu, Jiangxi, Jilin, Liaoning, Ningxia, Qinghai, Shaanxi, Shandong, Tianjin, Tibet, Xinjiang, Yunnan, and Zhejiang are not applicable, since no cases are recorded. For confirmed ad recovered variables different forecasting model are considered to reach the optimal results. To consider the worst-case forecasting scenarios and the best-case forecasting scenarios, an upper control limit (UCL, i.e., 5%) and lower control limit (LCL i.e., 5%) are considered, respectively. To sum up, forecast and worst scenarios for confirmed, deaths and recovered variables, calculations for all studied provinces are considered. The overall expected results after 210 days of all Chinese provinces for confirmed, deaths, and recovered cases are 5878678, 90571, and 69137, respectively, where in the worst scenario, the results are 7585212, 11948 and 205443, for confirmed, deaths, and recovered cases, respectively. Moreover, if the rate of growth of the infected cases did not solve then after 210 days, China will not be able to control the spreading of Corona virus.

Therefore, if no antidot is found for the new Corona virus, the number of deaths will be increased, since the currant capability of Chinese's hospitals are not able to treat that huge number of patients, which makes number of deaths patients to increase rapidly in very short time. After considering the total number of beds in all over China, which is 6.12 million beds, in the worst scenario the number of beds will be not enough for this huge number, which will increase the number of deaths. As a result, the number of deaths many rapidly increased to reach 1661952 (119468+87271+(7585212-6129999(beds))) in case all the Chinese hospitals worked together. Unfortunately, Hubei has the majority of confirmed cases and the ability of Hubei hospitals are very limited which add extra demands to

stop Corona virus spreading. Based on current status of Hubei's hospitals, the number of deaths after 210 days will exceed 5 million, since the capacity of Hubei's hospitals are very limited.

The outcome of this study suggested that extra efforts from international governments should be made to stop the virus from spreading to other countries and inside China. This can be done by increasing the number of medical staffs and hospitals in Hubei and other provinces to contain the future cases, isolating all the suspected cases, stopping any direct contact with the patients' families, returning to the history of each patient to understand the way of infection, combining the medical history of the patients with current diagnosis to extract information about the virus.

Table 3: The best forecasting models for each variable.

Province	Confirmed	Deaths	Recovered
Auhui	Brown	ARIMA(0,0,0)	Brown
Beijing	ARIMA(0,1,0)	Simple	Brown
Chongqing	ARIMA(0,1,0)	Brown	Brown
Fujian	ARIMA(0,1,0)	ARIMA(0,0,0)	Brown
Gansu	ARIMA(0,1,0)	ARIMA(0,0,0)	Brown
Guangdong	Holt	ARIMA(0,0,0)	Brown
Guangxi	Holt	ARIMA(0,0,0)	Brown
Guizhou	Brown	ARIMA(0,0,0)	Holt
Hainan	Holt	Simple	Brown
Hebei	Brown	ARIMA(0,0,0)	Brown
Heilongjiang	Holt	Holt	Brown
Henan	Holt	ARIMA(0,1,0)	Brown
Hubei	ARIMA(0,2,0)	ARIMA(0,2,0)	Brown
Hunan	Holt	ARIMA(0,0,0)	Brown
Jiangsu	Holt	ARIMA(0,0,0)	Brown
Jiangxi	Holt	ARIMA(0,0,0)	Holt
Jilin	Brown	ARIMA(0,0,0)	Simple
Liaoning	Holt	ARIMA(0,0,0)	Holt
Ningxia	ARIMA(1,1,0)	ARIMA(0,0,0)	Holt
Qinghai	ARIMA(0,1,0)	ARIMA(0,0,0)	ARIMA(0,0,0)
Shaanxi	Holt	ARIMA(0,0,0)	Brown
Shandong	ARIMA(0,1,0)	ARIMA(0,0,0)	Holt
Shanghai	Brown	Simple	ARIMA(0,1,0)
Shanxi	Brown	Simple	ARIMA(0,1,0)
Sichuan	Holt	Simple	Brown
Tianjin	Brown	ARIMA(0,0,0)	Holt
Tibet	Simple	ARIMA(0,0,0)	ARIMA(0,0,0)
Xinjiang	ARIMA(0,1,0)	ARIMA(0,0,0)	ARIMA(0,0,0)
Yunnan	Brown	ARIMA(0,0,0)	Brown
Zhejiang	ARIMA(0,1,0)	ARIMA(0,0,0)	ARIMA(0,2,0)

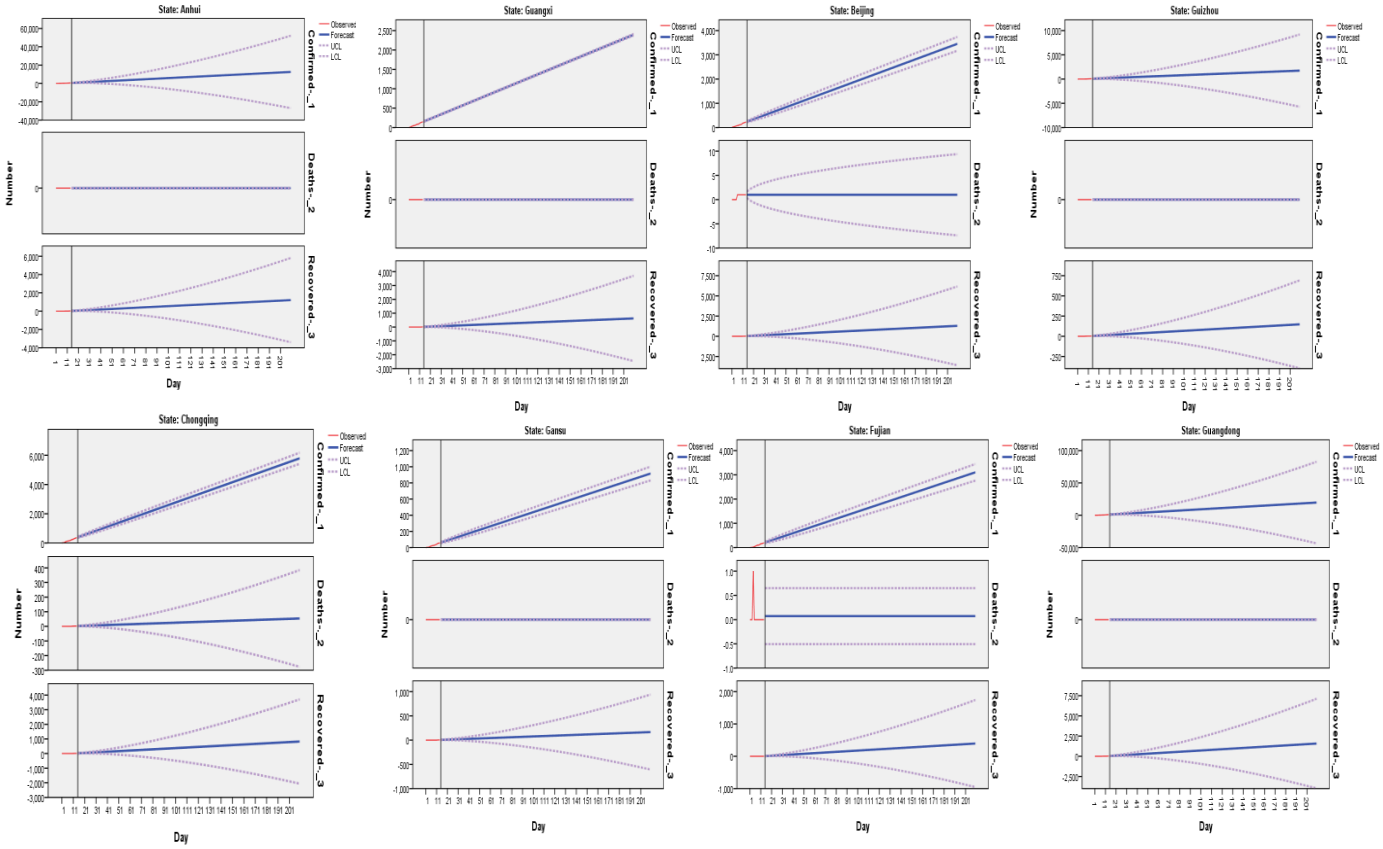

Figure 5: Forecasting Deaths, Confirmed, and Recovered for Auhui, Guangxi, Beijing, Guizhou (Upper), Chongqing, Gansu, Fujian, and Guangdong (Down)

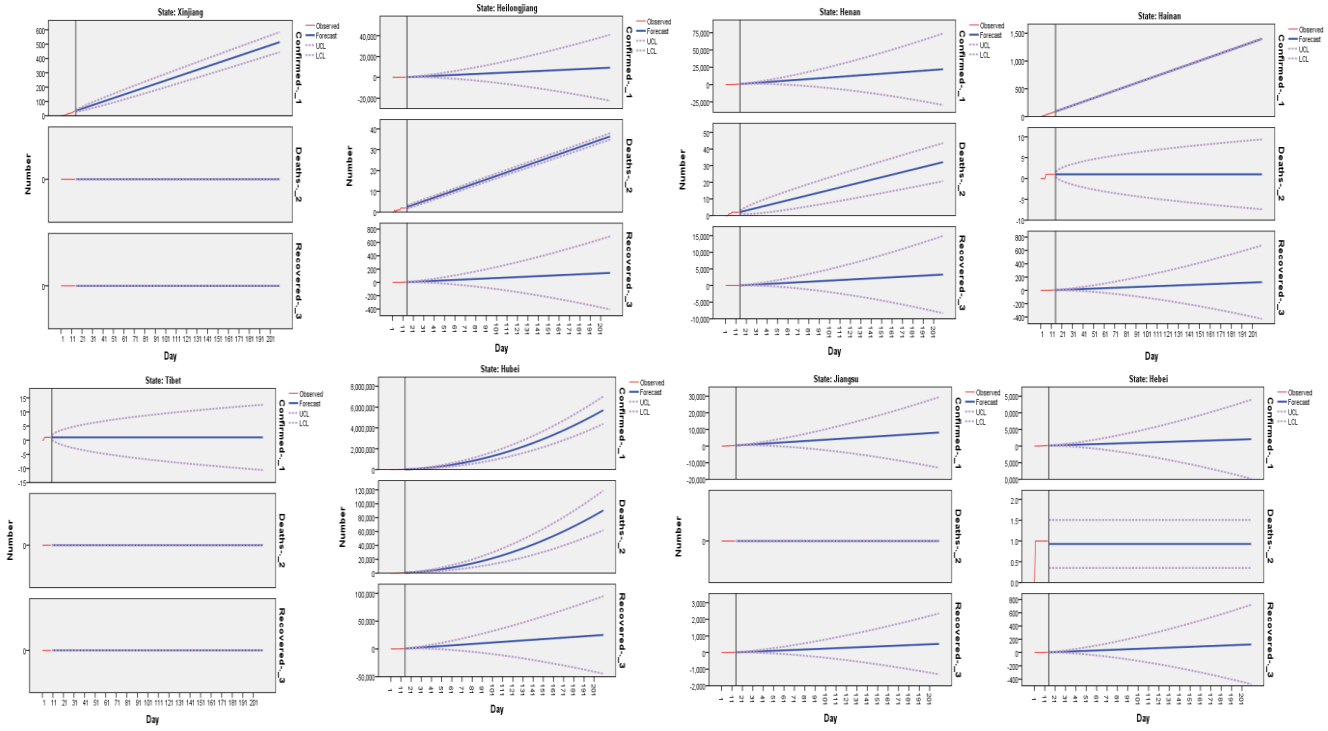

Figure 6: Forecasting Deaths, Confirmed, and Recovered for Xinjiang, Heilongjiang, Henan, Hainan (Upper), Tibet , Hubei, Jiangsu and Hebei (Down)

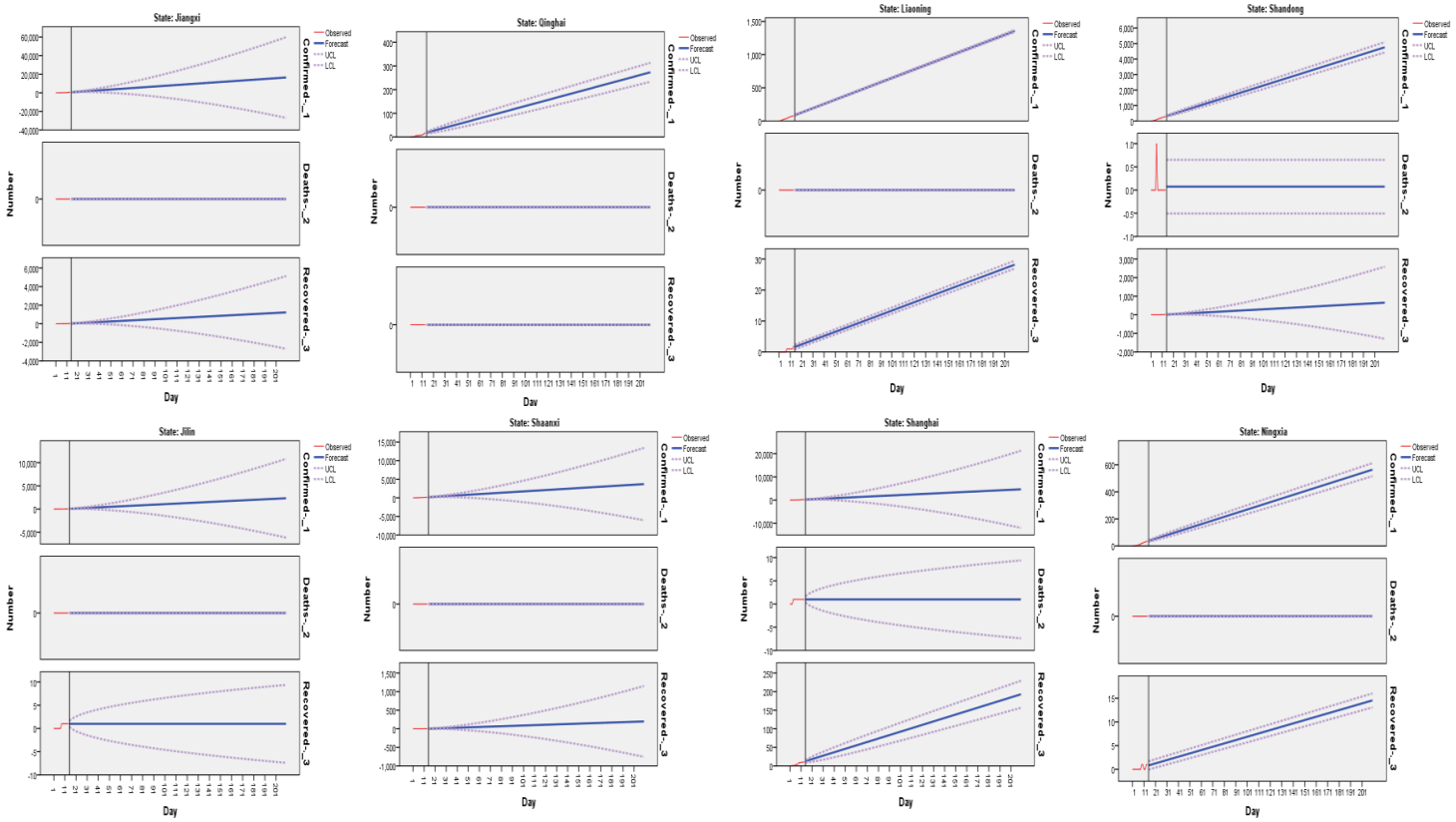

Figure 7: Forecasting Deaths, Confirmed, and Recovered for Jiangxi, Qinghai, Liaoning, Shandong(Upper), Jilin, Shaanxi, Shanghai and Ningxia (Down)

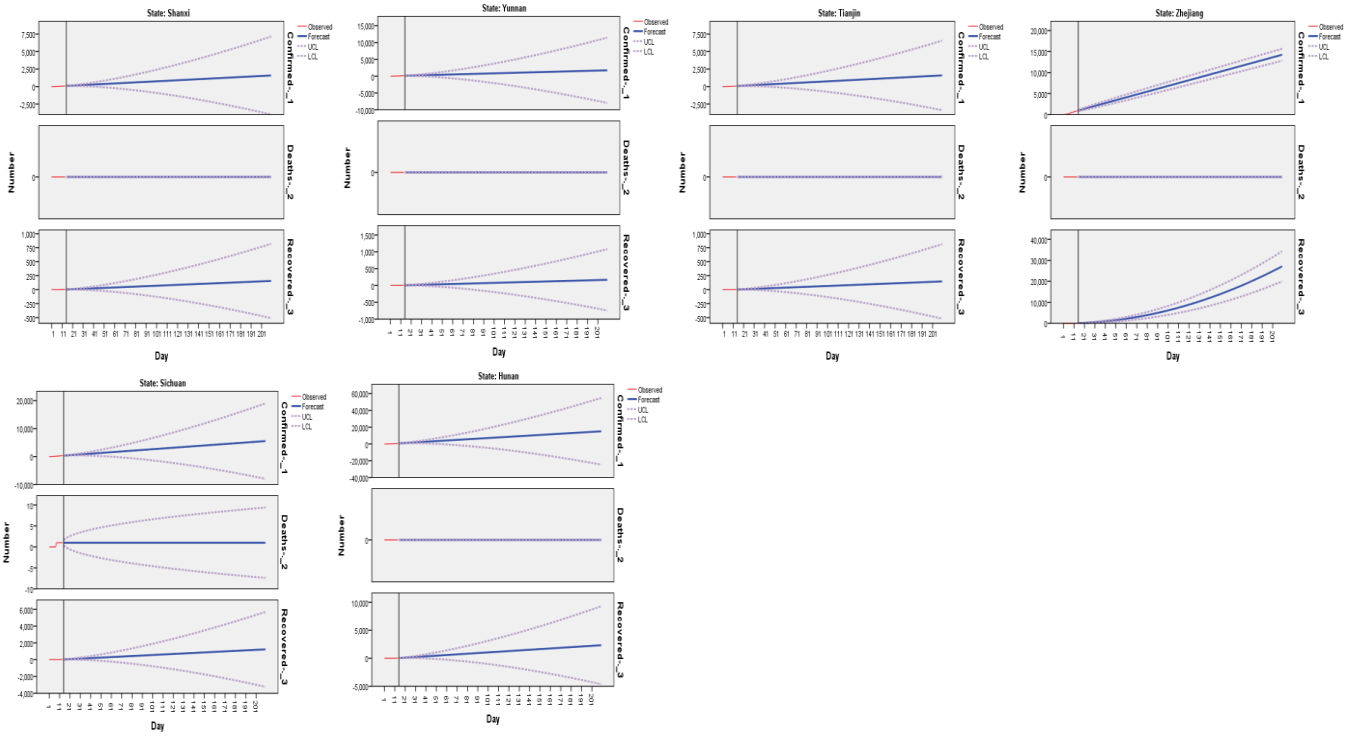

Figure 8: Forecasting Deaths, Confirmed, and Recovered for Shanxi, Yunnan, Tianjin, Zhejiang (Upper), Sichuan, and Hunan (Down)

Discussion

At the beginning of a new year, a novel Corona virus is (nCOV) raised in the horizon to threat all over the world about a new epidemic version of viruses that can kill many people when human-to- human transmission is made. The first case is reported from patient that visited Wuhan fish market, afterward hundreds of cases are reported from different places all over the world that either visited or have direct/indirect contact from Wuhan residence. In this study, we have estimated the outbreak spreading size of 2019-nCoV thus far in different Chinese's provinces to estimate confirmed, recovered and deaths cases and the relationship between weather variables and the previous cases in 30 provinces. Our findings show that different Chinese's provinces have different weather conditions, which makes confirmed, deaths and recovered cases of each province to be affected from different metrological weather variables. Therefore, in the absence of researches' articles that discussing the methods of stopping the virus. This research comes to highlight the future estimation number of deaths, recovered and confirmed cases in case no antidot is found. The research found that the growth of confirmed, and deaths cases are increasing more faster than the recovered cases. This gives an initial alarm to different governments all over the world, about the complexity and the expectation in the near future, if no plans are considered. Additionally, estimating the future infected number of patients as well as number of deaths will give the decision makers a vision about how to manage the problem and what is the needed plans that must be followed to contain nCov cases. Besides, the relationship between weather variables and provinces can give a hint to the future researchers on the most important variables that need to be considered to understand the virus's behavior. The used forecasting techniques in this study are simple and many researchers can regenerate the results and validate the estimated cases.

Lack of studies tried to consider using different forecasting models to estimate the expected number of infected cases in the near future by using 30 Chinese's provinces. Therefore, this study considers as a first article that suggested the optimal forecasting model for each province and tried to analysis the estimated number of deaths until 1st of September 2020. The study finds that in the worst-case scenario the total number of deaths will reach 5 million and in the normal growth the number will reach 1.6 million.

Contributors

Nadia AL-Rousan wrote a part of the manuscript, analyze the collected data, and designed the experiments, and Hazem AL-Najjar designed the experiments, interpreted the results, and wrote a part of the manuscript.

Declaration of interests

We declare no competing interests. Public epidemiological data from John Hopkins University and WHO websites were used in this research. Metrological data were purchased from from the Global Forecast System (GFS) Web service. All these data are available free and public, and any researcher can purchase the data from these websites.

References

1. Wuhan City Health Committee (WCHC). Wuhan Municipal Health and Health Commission's briefing on the current pneumonia epidemic situation in our city 2019 [updated 31 December 201914 January 2020]. Available from: http://wjw.wuhan.gov.cn/front/web/showDetail/2019123108989.
2. Chan, J. F. W., Yuan, S., Kok, K. H., To, K. K. W., Chu, H., Yang, J., ... & Tsoi, H. W. (2020). A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. The Lancet.
3. World Health Organization. (2020). Home care for patients with suspected novel coronavirus (nCoV) infection presenting with mild symptoms and management of contacts: interim guidance.
4. European Centre for Disease Prevention and Control (ECDC). Risk assessment: Outbreak of acute respiratory syndrome associated with a novel coronavirus, Wuhan, China; first update 2020 [updated 22 January 2020]. Available from: https://www.ecdc.europa.eu/en/publications-data/risk-assessment-outbrea

5. Chen, N., Zhou, M., Dong, X., Qu, J., Gong, F., Han, Y., ... & Yu, T. (2020). Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. The Lancet.
6. Corman, V. M., Landt, O., Kaiser, M., Molenkamp, R., Meijer, A., Chu, D. K., ... & Mulders, D. G. (2020). Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. Eurosurveillance, 25(3), 2000045.
7. News X. Experts claim that a new coronavirus is identified in Wuhan 2020 [14 January 2020]. Available from: http://www.xinhuanet.com/2020-01/09/c_1125438971.htm
8. Holmes E. Initial genome release of novel coronavirus 2020 [14 January 2020]. Available from: http://virological.org/t/initial-genome-release-of-novel-coronavirus/319.
9. Chan, J. F. W., Kok, K. H., Zhu, Z., Chu, H., To, K. K. W., Yuan, S., & Yuen, K. Y. (2020). Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan. Emerging Microbes & Infections, 9(1), 221-236.
10. Wang, H., Kanmangne, D., Li, R., Qian, Z., Xia, X., Wang, X., & Wang, T. (2020). miR-30a-3p suppresses the proliferation and migration of lung adenocarcinoma cells by downregulating CNPY2. Oncology Reports, 43(2), 646-654.
11. Wang, W., Tang, J., & Wei, F. (2020). Updated understanding of the outbreak of 2019 novel coronavirus (2019-nCoV) in Wuhan, China. Journal of Medical Virology.
12. Phan, L. T., Nguyen, T. V., Luong, Q. C., Nguyen, T. V., Nguyen, H. T., Le, H. Q., ... & Pham, Q. D. (2020). Importation and human-to-human transmission of a novel coronavirus in Vietnam. New England Journal of Medicine.
13. Hui, D. S., I Azhar, E., Madani, T. A., Ntoumi, F., Kock, R., Dar, O., ... & Zumla, A. (2020). The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health—The latest 2019 novel coronavirus outbreak in Wuhan, China. International Journal of Infectious Diseases, 91, 264-266.
14. Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., ... & Cheng, Z. (2020). Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. The Lancet.
15. Li, X., Zai, J., Wang, X., & Li, Y. (2020). Potential of large 'first generation'human-to-human transmission of 2019-nCoV. Journal of Medical Virology.
16. Bassetti, M., Vena, A., & Roberto Giacobbe, D. (2020). The Novel Chinese Coronavirus (2019-nCoV) Infections: challenges for fighting the storm. European Journal of Clinical Investigation, e13209.
17. Benvenuto, D., Giovannetti, M., Ciccozzi, A., Spoto, S., Angeletti, S., & Ciccozzi, M. (2020). The 2019-new coronavirus epidemic: evidence for virus evolution. Journal of Medical Virology.
18. Zhao, S., Lin, Q., Ran, J., Musa, S. S., Yang, G., Wang, W., ... & Wang, M. H. (2020). Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: A data-driven analysis in the early phase of the outbreak. International Journal of Infectious Diseases.
19. Wu, J. T., Leung, K., & Leung, G. M. (2020). Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. The Lancet.
20. Jung, S. M., Akhmetzhanov, A. R., Hayashi, K., Linton, N. M., Yang, Y., Yuan, B., ... & Nishiura, H. (2020). Real time estimation of the risk of death from novel coronavirus (2019-nCoV) infection: Inference using exported cases. medRxiv.
21. Lu, R., Zhao, X., Li, J., Niu, P., Yang, B., Wu, H., ... & Bi, Y. (2020). Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. The Lancet.
22. Chinese Center for Disease Control and Prevention (CDC), 2019 Novel Coronavirus, 2020, Beijing, China
23. Benesty, J., Chen, J., Huang, Y., & Cohen, I. (2009). Pearson correlation coefficient. In Noise reduction in speech processing (pp. 1-4). Springer, Berlin, Heidelberg.
24. Hyndman, R. J., & Athanasopoulos, G. (2018). Forecasting: principles and practice. OTexts.