IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

Statistical Correlation Analysis between Environmental cum Demographic Parameters and EMF Radiation with COVID-19 Infection

Kalmekar Rajkumar V.

Scientist-F
Environmental Science Department
Naval Material Research Laboratory (DRDO), Ambernath (E) - 421506, Maharashtra, India

Abstract:
Objective: Number of infection per unit population of country and related deaths due to COVID-19 is being reported different in all countries, seems not related with population density of country or hygiene level of country as anticipated. Factors contributing for different infection rates is not completely understood. This study is aimed at analyzing the correlation of COVID-19 infection by statistical method using open source data available with demographic and environmental parameters including EMF radiation of mobile and broadband with wi-fi router.
Statistical analysis design: Open source data on target parameter COVID-19 infections per unit population and related mortality were analysed for correlation with open source data on 14 various predictor parameters like Population Density, Net Migrants, Median Age, Vegetarian diet (%), % Urban Population, % Mobile Penetration in country, Mobile Connection Speed, Number of Fixed Broad Band Subscriptions in the country with wi-fi router, Fixed Broad Band Speed, Average February – March Temperature, % BCG Immunization Coverage, % Total & Urban Population using at least basic sanitation services and Per Capita CO2 Emissions. For statistical analysis, both actual data and average of data within sets of ranging infections per unit population have been considered. Also a mathematical model is generated for prediction of infections, using regression technique on best correlated parameter and also using combination of highly correlated parameters. This regression model could be used to calculate expected infections in the country under influence of set of parameters. Additionally, difference between actual infection and predicted infection calculated by this model, could be useful for finding out effectiveness of implementing other measures by the country to control infection below expected value and vice versa.
Results: with respect to target parameter total infection cases per unit population, top 7 predictor parameters in descending order of correlation coefficient values (ρ) are: Net Migrant / 1M Population (0.784), Fixed Broad Band Speed (0.749), Number of Fixed Broad Band with wi-fi router Subscriptions per unit population (0.694), Per Capita CO2 Emission (0.689), Median Age (0.637), % Urban Population (0.598) and % Total Population using at least basic sanitation services (0.597). Also with respect to another target parameter i.e. death per unit cases in the country, relatively lower ρ values are obtained. Top 5 parameters in descending order of correlation coefficient values (ρ) are: Number of Fixed Broad Band with wi-fi router Subscriptions per unit population (0.678), Tot cases/ 1M population (0.623), % BCG Immunization Coverage (-0.576; negative directional), Median Age (0.552) and Fixed Broad Band Speed (0.503). Additionally regression analysis helped in generating mathematical model of finding a combined parameters with higher correlation ρ value 0.846. The regression equation helped to calculate expected infection numbers and difference with actual could be utilized to access effectiveness of country in controlling infection by other means.
Conclusion: Statistical analysis showed highest correlation on COVID-19 infection and related deaths with use of high speed broadband services with wi-fi router, may be due to ill effects on health and immune power of electromagnetic radiation. There is need for further research on mechanism by which electromagnetic radiation influences COVID-19 infection. Also need to find out whether the effect of EMF radiation is temporary or permanent i.e. if radiation is reduced; infections and deaths are reduces or not.
Keywords: COVID-19, Corona Virus, Statistical correlation analysis, statistical regression analysis, Environmental effect, mobile radiation effect, broadband with wi-fi router radiation effect, electromagnetic (EMF) effect.

1. Introduction

Since December 2019 entire world is distressed with a new corona virus infection originated from Wuhan city of China. On 11 March 2020, WHO declared this illness caused by SARS-CoV-2 Corona virus as COVID-19 a global pandemic. Till first week of May 2020, it has reached to almost entire world ranging from developed to least developed countries. However, rate of growth of infection and related death is varying in all countries.

All though, infection is known to be through one to one public contact and due to contact with aerosol and droplets, the exact parameters deciding different rate of infection is not yet completely understood. Because of this; social distancing, isolating infected person, tracing contact history and curtailing all contact persons are the only preventive measures being utilized to curtain COVID-19 spread. Lot of speculations are being made on relation between rate of spread and effect of various parameters like population density, temperature, humidity, BCG vaccinations, sanitization standard, food habits, immunity, 5G mobile and EMF radiation etc.

Since the COVID-19 outbreak is very recent, there is paucity of peer revived research papers reporting correlation between COVID-19 and influencing parameters. However, some reports available, reporting association of COVID-19 with few individual parameters like air pollution and EMF radiation. Further, comparative correlation studies and detailed methodology for regression analysis on assorted possible combination parameters not have been reported.

Lippi et al. on 27 April 2020 [17] reported association between environmental pollution and prevalence of Coronavirus disease 2019 (COVID-19) in Italy using linear regression analysis and Pearson's correlation. Paper reported correlation coefficient ρ=0.66 with overall pollution levels (PM10 and ozone) justifying such polluted environment may be fertile for biological settings for high correlation. Similar study confirming correlation was reported by Marco et al. on 16 April 2020 [18] analyzing a link between air pollution and COVID-19 in England using some markers of poor air quality, nitrogen oxides and ozone. Aminoff et al. on 14 April 2020 [20] published "Reflections and Recommendations on COVID-19, 5G and Wireless Radiation". It is reported that wireless radiation is linked to oxidative stress and adverse impact on immune function. It is further quoted that spreading of virus is accelerated by electromagnetic radiation.

This paper is aimed at finding statistical correlation between COVID-19 infection with total 14 parameters including demographic, environmental and electromagnetic radiation. Here Fixed broadband connections are also considered as a source of EMF radiation since now a days mostly wi-fi routers are used to convert wired data in to wireless data for multiple use and ease of connection. Also it is aimed at providing detailed methodology for generating regression model using combination of certain highly correlated parameters for increasing accuracy of prediction.

This study may help in identifying most influencing parameters for COVID-19 infection and effectiveness of implemented infection control measures so that some additional measures can be initiated for further controlling the spread of the decease. In the present study; only data available through open internet sources has been used and no controlled experimental work has been done to find the correlation coefficient between parameters.

2. Materials & Methods

2.1. Study Design

Pearson Correlation coefficient (ρ) is considered for analysis and finding relation between two parameters. Pearson Correlation coefficient (ρ) measures linear correlation between two variables X and Y. Correlation coefficient is defined as the covariance of the variables divided by the product of their standard deviations. MS Excel program has been used to calculate the correlation coefficient (ρ)

$$\rho = Correl\ (X,Y) = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2 \sum(y-\bar{y})^2}} \qquad \text{(Eq. 1)}$$

where, x and y are the sample mean Average of parameter-1 and parameter-2

It has a value between +1 and -1, where +1 indicates a perfect positive linear correlation, 0 indicates no linear correlation, and -1 indicates perfect negative correlation. Therefore negative or positive values only signify its direction of correlation i.e. ρ is positive if parameter 1 increases with increase of parameter 2, or vice versa. Values above 0.7 are considered good, near 0.5 are considered moderate and values below 0.3 are considered having weak relation.

2.2. Sourcing data for statistical analysis

Data taken for present study on COVID-19 is from internet website [1] www.worldometers.info/coronavirus/, as retrieved on 11 May 2020 and only countries having total COVID-19 cases above 4700 have been selected. Therefore total 58 countries were found above this cutoff number of 4700 as on 11 May 2020 and considered in the study as listed in Data Sheet Table 1.

Total 14 predictor parameters are used for correlation analysis all taken from internet websites as listed under [references]. Out of which 5 are demographic parameters viz. P1: Population Density (Nos/Km²) [2], P2: Net Migrants / 1M Population [2], P3: Median Age (Yrs) [2], P4: Vegetarian diet (%) [13], P5: % Urban Population [2]. Remaining 9 are environmental parameters viz P6: % Mobile Penetration in the country [3], P7: Mobile connection speed (Mb/s) [4], P8: No of Fixed Broad Band Subscriptions/ 100 [5], P9: Fixed Broad Band Speed (Mbps) [6], P10: Average Feb-Mar Temperature (Deg. C) [7], P11: % BCG immunization coverage [8], P12: % Total Population using at least basic sanitation services [12], P13: % Urban Population using at least basic sanitation services [12] and P14: Per Capita CO2 Emissions (Tons) [9]. In this study special attention is given to the parameters which are easily controllable for managing infection like P4 and P6 to P11 and discussed in details. Two sets of target parameter of COVID-19 Infection is considered, these are P15: Total Cases/ 1M population and P16: Death / 1000 cases. For ease of interpretation, data of countries have been sorted out in descending order of Total Cases/ 1M population. Non available data of country have been kept blank, which will not have any effect on analysis result.

Table 1: Data Sheet

Parameter Type →		(a) Demographic Predictor Parameters					(b) Environmental Predictor Parameters									(c) Target Parameters	
Parameter No →		P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16
Set No	Country Name	Population density (P/Km²)	Net migrant / 1M Pop.	Median age (Yrs)	Vegetarian diet (%)	Urban pop %	% Mobile penetration in country	Mobile connection speed (Mb/s)	No of Fixed Broad Band subscriptions/ 100	Fixed Broad Band speed (Mbps)	Avg Feb-Mar temp. (Deg. C)	% BCG immunization coverage	% Total Population using at least basic sanitation services	% Urban Population using at least basic sanitation services	Per capita CO2 emissions (Tons)	Tot cases/ 1M pop	Death / 1000 cases
Reference		[2]			[13]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[12]		[9]	[1]	
S1	Qatar	248	13.88	32		96	82.7	24.6	9.70	94.94	19.85	99	100	100	37.29	7,817	0.64
	Spain	94	0.86	45	1.5	80	75.4	24.8	31.22	127.76	11.15	0	100	100	5.40	5,661	100.51
	Ireland	72	4.78	38	4.3	63	49.8	16.2	29.43	78.68	6.20	86	91	89	8.32	4,657	63.35
	Belgium	383	4.14	42	7	98	81.8	34.2	38.31	84.67	3.55	0	99	99	8.34	4,612	162.84
S2	USA	36	2.88	38	8	83	88.1	21.3	33.85	132.55	3.95	0	100	100	15.52	4,133	59.04
	Singapore	8,358	4.62	42		100	85.9	39.3	25.76	197.26	26.85	99	100	100	8.56	4,072	0.74
	Italy	206	2.46	47	10	69	61.8	19.9	27.94	55.19	7.55	0	99	99	5.90	3,623	139.39
	Switzerland	219	6.01	43	14	74	82.0	35.2	45.42	148.24	-0.05	0	100	100	4.73	3,502	60.54
S3	UK	281	3.84	40	7	83	67.6	21.7	39.31	66.99	5.05	0	99	99	5.55	3,229	145.25
	Bahrain	2,239	28.09	32		89	76.5	13.9	14.31	27.00	19.00		100	100	17.15	2,904	1.72
	France	119	0.56	42	5	82	58.7	25.2	43.75	136.45	6.10	82	99	99	5.13	2,711	149.02
	Portugal	111	-0.59	46	1.2	66	70.8	21.6	34.6	109.99	12.15	74	100	100	4.86	2,705	41.04
S4	Sweden	25	3.96	41	10	88	86.5	30.8	37.7	134.56	-3.15	4	99	99	4.54	2,606	122.41
	Belarus	47	0.92	40		79	57.0		33.41	58.12	-3.35	95	98	98	6.63	2,530	5.53
	Netherlands	508	0.93	43	5	92	88.8	42.4	42.33	112.68	3.55	0	98	98	9.62	2,488	127.41
	Germany	240	6.49	46	10	76	61.3	22.6	40.45	94.73	1.85	0	99	99	9.44	2,051	43.88
S5	Peru	26	3.00	31		79	72.8	11.7	7.18	27.57	17.55	65	74	80	1.87	2,041	27.93
	Kuwait	240	9.25	37		100	86.8	16.2	2.74	62.51	16.95	84	100	100	25.65	2,034	6.88
	Panama	58	2.60	30		68	67.1	7.2	10.88	80.11	26.35	97	83	92	2.87	1,958	29.11
	Israel	400	1.16	30	13	93	61.3	13.6	28.14	82.01	13.00	75	100	100	8.04	1,905	15.22
S6	UAE	118	4.04	33		86	79.9	19.9	-6.55	100.95	20.45	15	99	99	23.37	1,840	10.87
	Canada	4	6.41	41	9.4	81	82.1	42.5	38.01	120.98	-6.55		99	99	18.58	1,824	70.72
	Denmark	137	2.62	42	5	88	80.0	34.6	43.17	134.13	1.05	0	100	100	6.65	1,815	50.14
	Austria	109	7.22	43	9	57	70.8	27.5	28.75	54.53	0.60	90	100	100	8.43	1,762	39.16
S7	Ecuador	71	2.06	28		63	44.3	10.5	10.13	23.03	20.60	82	88	91	2.43	1,675	72.24
	Turkey	110	3.37	32		76	60.5	17.1	14.77	24.15	5.15	42.0	97.0	100	4.61	1,644	27.37
	Russia	9	1.25	40	4	74	57.0	12	21.44	64.57	-13.10	0	90	95	11.44	1,517	9.23
	Chile	26	5.84	35	6	85	63.6	12	16.94	91.36	15.40	96	100	100	4.46	1,510	10.60
S8	Norway	15	5.16	40	4	83	89.9	48.2	40.23	125.85	-2.30	0	98	98	8.28	1,495	26.76
	Iran	52	-0.65	32		76	58.6	16.7	12.39	10.10	7.75	7.0	88.0	92.0	8.08	1,301	61.49
	Moldova	123	-0.34	38		43			14.42	65.66	0.80		76	86	2.03	1,221	34.40
	Serbia	100	0.46	42		56	75.2	21.5	21.21	55.26	2.85	89	98	100	4.65	1,158	21.59
S9	Saudi Arabia	16	3.88	32		84	67.8	13.6	7.59	61.03	19.20	33	100	100	15.94	1,122	6.24
	Finland	18	2.53	43	6	86	82.8	27	30.95	87.59	-6.40	90	99	99	9.31	1,080	44.44
	Dominican Republic	225	-2.77	28		85	56.3	8.5	7.3	23.80	23.80	86	84	86	2.26	954	37.74
	Romania	84	-3.85	43		55	65.4	13.2	24.29	151.55	1.00	99	84	95	3.98	810	62.96

Table 1: Data Sheet (contd..)

Parameter No →		P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16
Set No	Country Name	Population density (P/Km^2)	Net migrant / 1M Pop.	Median age (Yrs)	Vegetarian diet (%)	Urban pop %	% Mobile penetration in country	Mobile connection speed (Mb/s)	No of Fixed Broad Band subscriptions/ 100	Fixed Broad Band speed (Mbps)	Avg Feb-Mar temp. (Deg. C)	% BCG immunization coverage	% Total Population using at least basic sanitation services	% Urban Population using at least basic sanitation services	Per capita CO2 emissions (Tons)	Tot cases/ 1M pop	Death / 1000 cases
S10	Brazil	25	0.10	33	14	88	57.6	13	13.7	53.31	25.40	56	88	93	2.25	765	67.97
	Czechia	139	2.06	43	2	74	83.7	31.5	28.82	57.81	0.45		99	99	10.53	759	34.26
	Poland	124	-0.78	42	8	60	64.9	17.3	18.48	96.88	-0.30	93	99	99	7.81	428	49.07
	Ukraine	75	0.23	41	5.2	69		11.2	12.55	51.83	-0.45	93	96	97	5.22	358	25.14
S11	Kazakhstan	7	-0.96	31		58	61.6	11.4	14.14	41.15	-7.25	90	98	97	13.01	274	7.30
	Mexico	66	-0.47	29	8	84	70.5	14.9	13.26	37.23	18.40	48	91	93	3.58	272	99.26
	Australia	3	6.21	38	12	86	85.3	37.4	32.4	43.40	21.85		100	100	17.10	272	14.71
	Colombia	46	4.02	31	12	80	62.8	10	12.88	28.63	24.00	99	90	93	1.61	217	41.47
S12	S. Korea	527	0.23	44	3	82	97.0	52.4	41.58	120.11	3.05	52	100	100	11.85	213	23.47
	Malaysia	99	1.54	30		78	68.9	11.5	8.5	79.86	26.30	99	100	100	8.68	208	14.42
	South Africa	49	2.45	28		67	65.7	15	2.99	28.70	20.85	92	76	76	6.95	169	17.75
	Morocco	83	-1.39	30		64	64.7	11.2	3.86	15.50	13.05	99	89	94	1.64	169	29.59
S13	Pakistan	287	-1.06	23		35	61.5	6.2	0.93	8.49	15.45	6	60	77	0.87	140	21.43
	Argentina	17	0.11	32	5	93	68.1	12.8	17.78	38.02	21.85	78	96	96	4.61	134	52.24
	Algeria	18	-0.23	29		73	40.9	3.1	7.66	3.63	12.45	86	88	90	3.85	131	83.97
	Japan	347	0.57	48	4.7	92	94.3	33	31.68	100.56	5.70	85	100	100	9.70	125	40.00
S14	Philippines	368	-0.61	26	5	47	54.4	7	3.24	23.80	25.85	91	74	78	1.22	101	69.31
	Bangladesh	1,265	-2.24	28		39	40.5	5.7	4.43	25.78	23.35	1	48	51	0.47	95	10.53
	Egypt	103	-0.37	25		43	45.4	8.6	5.35	30.91	16.00	98	94	98	2.32	92	54.35
S15	China	153	-0.24	38	5	61	74.0	24.0	26.86	111.96	4.35	64	85	91	7.38	58	51.72
	Indonesia	151	-0.36	30		56	65.7	6.9	2.29	20.13	26.30	78	73	80	2.03	52	76.92
	India	464	-0.39	28	31	35	83.1	6.8	1.33	35.98	21.50	4	60	72	1.91	49	40.82

2.3. Averaging Data in to set of countries

Infection data is taken from actual social field without any scientific experiments under controlled condition. Data of demographic and environmental parameters may have actual measurement limitations of accuracy and time lag and some of the data was not available. Further, there are chances that certain countries infection rates may deviate from expected values due to effect of other parameters like strict social distancing measures implemented inside country, effect of use of other medications and vaccinations etc which are not considered in this study. This may have chance of certain data working as noise in the statistical correlation analysis and may pollute the overall results. To avoid such possible influence, after shorting countries according to total number of infections per unit population, countries are grouped in to 15 sets with 4 countries in each set except last two sets having only 3 countries. All data is averaged within each n=15 number of sets.

Table 2 shows average data within 15 numbers of sets of 58 countries under study. For sake of comparison, correlation coefficient values for both average data set and actual data have been calculated and plotted in the result. However, further regression analysis has been done based on average data set only.

2.4. Statistical Regression analysis and Goodness of Fit

Statistical regression analysis was done using average data set with two variable combinations: firstly with a single most correlated and controllable parameter and secondly with combination of several most correlated parameters.

For single parameter regression model, as shown in Fig. 3, a scatter plot was plotted between single predictor (Y) and target (X) parameter. Using MS Excel program, a best fit trendline is plotted. Thereafter, regression equation is established and coefficient of determination (R^2) is obtained. The obtained regression equation is used to find P17: Expected cases / 1M population, P18: Per unit difference; which is difference between actual cases and expected cases per unit actual cases as shown in Table 4. Obtained P18 indicates effectiveness of countries in controlling the infection; may be using some different measures and parameters not considered in this study.

For combination parameter regression model, individual countries combination parameter values (Y_c) are calculated using several (n) number of respective countries predictor parameters as given below (Eq. 2) and the parameter is termed as P19: Combination parameter. It is expected that the obtained combination parameter should show highest correlation ρ value more than considered individual parameters. Further modeling processing is done similar to single parameter regression model. P20: Expected infection and P21: per unit difference based on combination parameter regression model is calculated as shown in Table 4.

$$(Y_c) = \frac{\sum_{i=1}^{i=n}\left(\frac{y_i}{\sum y_i}\right) r_i}{n} \qquad \text{(Eq. 2)}$$

Where, Y_c= combined parameter value
y= individual parameter value
i= parameter No
n= total parameters
ρ= correlation value of individual parameter

Table 2: Average Data within the Set

Set No	(a) Demographic Predictor Parameters					(b) Environmental Predictor Parameters									(c) Target Infection Parameters		(d)
	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	P19
	Population density (P/Km^2)	Net migrant / 1M Pop.	Median age (Yrs)	Vegetarian diet (%)	Urban pop %	% Mobile penetration in country	Mobile connection speed (Mb/s)	No of Fixed Broad Band subscriptions/ 100	Fixed Broad Band speed (Mbps)	Avg Feb-Mar temp. (Deg. C)	% BCG immunization coverage	% Total Population using at least basic sanitation services	% Urban Population using at least basic sanitation services	Per capita CO2 emissions (Tons)	Tot cases/ 1M pop	Death / 1000 cases	Combination Parameter
S1	199.3	5.92	39.25	4.27	84.25	72.44	24.95	27.17	96.51	10.19	46.3	97.5	97.0	14.84	5686.8	81.8	19.9
S2	2204.8	3.99	42.50	10.67	81.50	79.45	28.93	33.24	133.31	9.58	24.8	99.8	99.8	8.68	3832.5	64.9	18.1
S3	687.5	7.98	40.00	4.40	80.00	68.40	20.60	32.99	85.11	10.58	52.0	99.5	99.5	8.17	2887.3	84.3	20.4
S4	205.0	3.08	42.50	8.33	83.75	73.40	31.93	38.47	100.02	-0.28	24.8	98.5	98.5	7.56	2418.8	74.8	16.2
S5	181.0	4.00	32.00	13.00	85.00	72.01	12.18	12.24	63.05	18.46	80.3	89.3	93.0	9.61	1984.5	19.8	12.8
S6	92.0	5.07	39.75	7.80	78.00	78.20	31.13	25.85	102.65	3.89	35.0	99.5	99.5	14.26	1810.3	42.7	18.9
S7	54.0	3.13	33.75	5.00	74.50	56.37	12.90	15.82	50.78	7.01	55.0	93.8	96.5	5.74	1586.5	29.9	10.8
S8	72.5	1.16	38.00	4.00	64.50	74.57	28.81	22.06	64.22	2.28	32.0	90.0	94.0	5.76	1293.8	36.1	10.1
S9	85.8	-0.05	36.50	6.00	77.50	68.08	15.58	17.53	80.99	9.40	77.0	91.8	95.0	7.87	991.5	37.8	9.5
S10	90.8	0.40	39.75	7.30	72.75	68.74	18.25	18.39	64.96	6.28	80.7	95.5	97.0	6.45	577.5	44.1	9.3
S11	30.5	2.20	32.25	10.67	77.00	70.02	18.43	18.17	37.60	14.25	79.0	94.8	95.8	8.83	258.8	40.7	10.4
S12	189.5	0.71	33.00	3.00	72.75	74.06	22.53	14.23	61.04	15.81	85.5	91.3	92.5	7.28	189.8	21.3	8.8
S13	167.3	-0.15	33.00	4.85	73.25	66.21	13.78	14.51	37.68	13.86	63.8	86.0	90.8	4.76	132.5	49.4	6.3
S14	578.7	-1.08	26.33	5.00	43.00	46.77	7.10	4.34	26.83	21.73	63.3	72.0	75.7	1.34	96.0	44.7	2.3
S15	256.0	-0.33	32.00	18.00	50.67	74.25	12.57	10.16	56.02	17.38	48.7	72.7	81.0	3.77	53.0	56.5	5.9

3. Results and discussion

Table 3 shows Pearson Correlation coefficient (ρ) for both: (a) average data set and (b) actual data. As expected overall correlation coefficients of actual data is less than that of set average due to the noise influence possibilities as discussed previously. Correlation patters remains more or less same for total cases / 1M population, thus ρ values suggest conclusive results. However, some deviation in ρ values observed for death/ 100 cases between average data and actual data for some of the parameters viz. P1: Population density, P2: Net migrant / 1M Population and P14: Per capita CO2 emissions (Tons) indicating non-conclusive results or no correlation wrt death/ 100 cases for these parameters.

3.1. **Correlation of (a) average data set with respect to P15:Total Cases/ 1M population**

Based on Table 3, correlation coefficient values of all predictor parameters with target parameter P15:Total Cases/ 1M population is graphically plotted as shown in Fig. 1.

A good correlation with highest values among all parameters was observed (with ρ>0.7) for P2: Net Migrant / 1M Population and P9: Fixed Broad Band Speed, having ρ values 0.784 and 0.749, respectively. This suggests primary initiation or nucleation of infection may be due to net international migrants in the country. However, propagation of infection may be aggravated by ill health effects due to higher electromagnetic radiations of Fixed broadband connections with wi-fi routers. Fixed Broadband transmission is mostly used with wi-fi wireless routers, which have high speed data transmission (@ 3 to 4 times of mobile data) and have high frequency than wireless mobile transmissions. Though, identifying reason for correlation was not in the scope of this study, several research papers elsewhere have reported reasons for effect of electromagnetic radiations on health especially on immune system of human body. Jahansson et al. in 2009 [10] reported loss of immune power due to even non ionizing electromagnetic radiation and termed electromagnetic radiations as immune-disrupting. Bonhomme-Faivre et al. in 2003 [11] found statistically significantly lower immune system; total lymphocyte, CD4, and CD3 counts, and significantly increased natural killer (NK) cells, in subjects exposed to electromagnetic field vs. controls. Simko et al. in 2019 [19] reviewed 94 relevant paper on 5G wireless communication and health effect using 6 to 100 GHz frequency range and reported that 80% in vivo studies showed response to exposures while 58% in vitro studies demonstrated effect. The study reported various health effects like release of free radicals, cell effects similar to normal heating, cell biological and morphological changes, increase in motility of human spermatozoa and immune system related parameter change etc. Aminoff et al. on 14 April 2020 [20] reported spreading of virus is accelerated by electromagnetic radiations due to reduction in overall immune power of human body. However, there is requirement of more research on finding exact mechanism how electromagnetic radiation of fixed broadband communication influences COVID-19 infection. If the effect of EMF radiation is permanent and non reversible, there may not be immediate results observed with reducing EMF radiation but will benefit in long run. However, if effect of EMF radiation is temporary and reversible, immediate results of reducing Broadband and other EMF radiation on reduction of infection could be expected.

Above moderate to good correlation was observed (with $0.6<\rho<0.7$) for P3: Median Age, P8: No of Fixed Broad Band Subscriptions/ 100 and P14: Per Capita CO2 Emissions, having ρ values 0.637, 0.694 and 0.689, respectively. Higher aged people are having less immunity and other health complications and are therefore prone to infections. Effect of large number of population using high speed Broadband services related to more infection is as discussed above. Correlation of higher infection with higher per capita CO2 emission in the country may be due to zinc and other nutrients deficiency and reduction in immunity power as reported by Myers et al. in 2015 [14] and Smith et al. in 2018 [15]. Also, COVID-19 mainly attacks respiratory system and lung thus supply of adequate pure oxygen is very important against this infection and therefore higher CO_2 emission showed positive correlation with infection numbers. As, per capita CO2 emission is a controllable parameter, all measures of reducing CO2 emission may be taken like closing high CO2 emitting industries and activities in the infected areas for controlling infection.

Table 3: Pearson Correlation Coefficient (ρ) of (a) Average Data Set and (b) Actual data

Type of Parameter		(a) Demographic Predictor Parameters					(b) Environmental Predictor Parameters									(c)	(e)
Parameter No →		P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P19
	(a) Average Data Set and (b) Actual data	Population density (P/Km²)	Net migrant / 1M Pop.	Median age (Yrs)	Vegetarian diet (%)	Urban pop %	% Mobile penetration in country	Mobile connection speed (Mb/s)	No of Fixed Broad Band subscriptions/ 100	Fixed Broad Band speed (Mbps)	Avg Feb-Mar temp. (Deg. C)	% BCG immunization coverage	% Total Population using at least basic sanitation services	% Urban Population using at least basic sanitation services	Per capita CO2 emissions (Tons)	Tot cases/ 1M pop	Combination Parameter
P12:Tot cases/ 1M pop	a	0.401	**0.784**	0.637	-0.125	0.598	0.338	0.514	**0.694**	**0.749**	-0.338	-0.522	0.597	0.550	**0.689**		**0.846**
	b	0.213	**0.469**	0.367	-0.184	0.421	0.234	0.320	0.388	**0.467**	-0.124	-0.244	0.399	0.360	**0.441**		**0.628**
Death / 1000 cases	a	0.374	0.499	0.552	-0.008	0.180	0.167	0.368	**0.678**	0.503	-0.258	-0.576	0.273	0.219	0.232	**0.623**	
	b	-0.184	-0.195	0.276	-0.050	0.087	-0.036	0.174	**0.414**	0.149	-0.178	-0.359	0.100	0.093	-0.292	0.245	

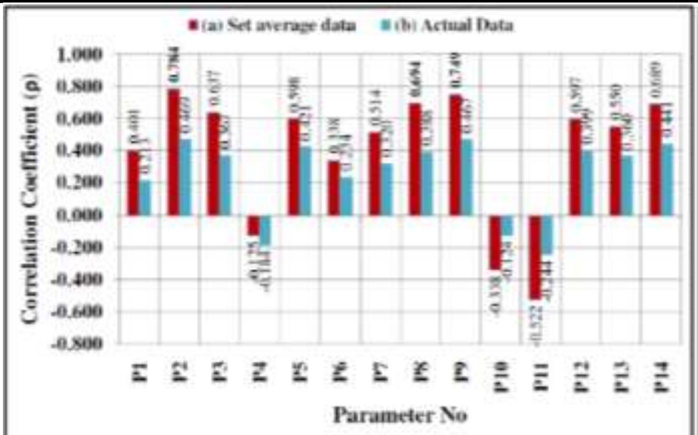

Fig. 1: Correlation Coefficient of (a) set average data and (b) actual data wrt P15: Total cases/ 1M Pop.

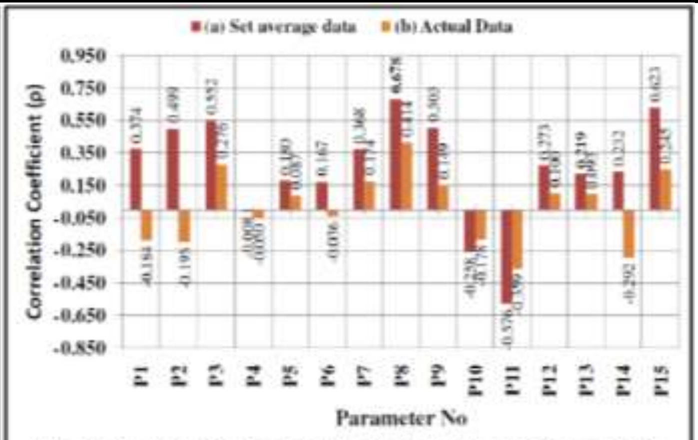

Fig. 2: Correlation Coefficient of (a) set average data and (b) actual data wrt P16: Death/ 1000 cases

A moderate correlation was observed (with $0.5 < \rho < 0.6$) for P5: Urban Population %, P7: Mobile Connection Speed, P11: % BCG immunization coverage, P12: % Total Population using at least basic sanitation services and P13: % Urban Population using at least basic sanitation services with ρ values 0.598, 0.514, -0.522, 0.597 and 0.550, respectively. This indicates a moderate correlation with more % of urban population, higher mobile connection speed, less BCG immunization and higher sanitation levels in country. This may be due to chances of urban population coming in contact with international travelers more. Human body living in higher sanitation level may have less immunity to fight with unknown new infection. It is already reported by Stanwell-Smith in 2001 [16] that too much cleanliness may be harmful to the development and regulation of the immune system. An indicative outcome from this result is, managing controllable parameters for reducing infection by reducing mobile data speed or frequency of radiation and increasing BCG immunization.

Weak correlation was observed (with $\rho < 0.5$) for P1: Population Density, P4: Vegetarian diet %, P6: % Mobile Penetration, and P10: Avg Feb-Mar Temperature having ρ values 0.401, -0.125, 0.338 and -0.338, respectively. This gives only indication about direction of correlation with very weak correlation values which are suggestive but not necessary for reducing infections by reducing population density, practicing vegetarian diet and maintaining higher surrounding temperature.

3.2. Correlation of (a) average data set with respect to P16:Death/ 1000 cases

Based on Table 3, correlation coefficient values of all predictor parameters with target parameter P16: Death / 1000 cases is graphically plotted as shown in Fig. 2.

Overall correlation values of parameters wrt Death/ 1000 cases is marginally less as compared to wrt Total Cases/ 1M population (0.678 vs 0.784). This indicates correlation of considered parameters on death due to COVID-19 is less due to effect of local healthcare facilities which is not considered in this study.

However, again above moderate but highest correlation among all parameter was observed (with $0.6 < \rho < 0.7$) for P8: No of Fixed Broad Band Subscriptions/ 100 with ρ values 0.678. This suggests effect of higher EMF radiations even for higher deaths similar to effect on higher infection numbers. This finding supports requirement of reducing the controllable Fixed Broadband and all other EMF radiations for reducing both: COVID-19 infections and related deaths. Correlation coefficient with respect to P15: Total cases/ 1M Population was 0.623, indicating death are related to number of infections in the country which is obvious.

Moderate correlation was observed (with $0.5<\rho<0.6$) for P3: Median age, P9: Fixed Broad Band Speed, and P11: % BCG immunization coverage having ρ values 0.552, 0.503 and -0.576, respectively. This further supports effect of reducing EMF radiation and increasing BCG immunization for reducing COVID-19 infections and related deaths.

Rest other parameters showed weak correlation with ρ values < 0.5, indicating COVID-19 infections may be more influenced than related death due to further treatment under controlled conditions and health care facilities and other support system in the country.

3.3. Regression mathematical model

(A) Single parameter regression model: As per correlation of (a) average data set with respect to P15: Total Cases/ 1M population, maximum correlation coefficient was obtained for P2: Net migrations/ 1M population and P9: Fixed Broad Band speed. P9 being a controllable parameter, single parameter regression model was generated for same. Fig. 3 shows scatter plot of Fixed Broad Band Speed on X axis and Total Cases / 1M Population on Y axis. Best fit power regression trendline equation obtained with $R^2 = 0.600$, as given below-

$$(Y) = 0.014X^{2.615} \quad \text{(Eq. 3)}$$

where, Y=sets average (Fixed Broad Band Speed)

X= Total Cases / 1M Population

Table 4: Regression Analysis data

Parameter Type →		(d) Single Regression Model		(e) Combination Regression Model			Parameter Type →		(d) Single Regression Model		(e) Combination Regression Model		
Parameter No →		P17	P18	P19	P20	P21	Parameter No →		P17	P18	P19	P20	P21
Set No	Country Name	Expected cases / 1M pop	Per unit difference	Combination Parameter	Expected cases / 1M pop	Per unit difference	Set No	Country Name	Expected cases / 1M pop	Per unit difference	Combination Parameter	Expected cases / 1M pop	Per unit difference
S1	Qatar	2076	0.73	32.75	8930	-0.14	S9	Saudi Arabia	654	0.42	14.02	1431	-0.28
	Spain	4512	0.20	13.48	1314	0.77		Finland	1681	-0.56	14.87	1624	-0.50
	Ireland	1270	0.73	16.20	1953	0.58		Dominican Republic	56	0.94	1.10	6	0.99
	Belgium	1539	0.67	17.00	2167	0.53		Romania	7052	-7.71	7.92	416	0.49
S2	USA	4968	-0.20	18.79	2690	0.35	S10	Brazil	459	0.40	6.29	253	0.67
	Singapore	14049	-2.45	20.19	3142	0.23		Czechia	567	0.25	13.41	1300	-0.71
	Italy	502	0.86	12.50	1115	0.69		Poland	2188	-4.11	9.69	644	-0.51
	Switzerland	6656	-0.90	21.08	3450	0.01		Ukraine	426	-0.19	7.62	384	-0.07
S3	UK	834	0.74	15.18	1698	0.47	S11	Kazakhstan	233	0.15	7.90	414	-0.51
	Bahrain	77	0.97	40.04	13782	-3.75		Mexico	179	0.34	5.21	169	0.38
	France	5359	-0.98	14.65	1573	0.42		Australia	268	0.01	19.47	2906	**-9.68**
	Portugal	3050	-0.13	11.57	944	0.65		Colombia	90	0.58	9.22	578	-1.67
S4	Sweden	5167	-0.98	17.31	2255	0.13	S12	S. Korea	3839	**-17.02**	15.62	1805	**-7.47**
	Belarus	575	0.77	11.37	909	0.64		Malaysia	1320	-5.35	9.97	685	-2.29
	Netherlands	3249	-0.31	15.46	1766	0.29		South Africa	91	0.46	7.82	406	-1.40
	Germany	2064	-0.01	20.71	3321	-0.62		Morocco	18	0.89	1.82	17	0.90
S5	Peru	82	0.96	7.50	371	0.82	S13	Pakistan	4	0.97	0.93	4	0.97
	Kuwait	696	0.66	22.56	3995	-0.96		Argentina	190	-0.41	6.87	306	-1.29
	Panama	1331	0.32	9.62	634	0.68		Algeria	0	1.00	3.69	80	0.39
	Israel	1415	0.26	11.69	966	0.49		Japan	2413	**-18.30**	13.73	1366	**-9.93**
S6	UAE	2437	-0.32	16.33	1987	-0.08	S14	Philippines	56	0.45	2.51	35	0.65
	Canada	3912	-1.14	23.76	4468	-1.45		Bangladesh	69	0.28	0.86	3	0.96
	Denmark	5124	-1.82	17.19	2220	-0.22		Egypt	110	-0.20	3.55	74	0.20
	Austria	487	0.72	18.21	2516	-0.43	S15	China	3195	**-54.08**	11.40	915	**-14.78**
S7	Ecuador	51	0.97	6.65	286	0.83		Indonesia	36	0.31	3.03	52	-0.01
	Turkey	58	0.96	9.54	622	0.62		India	164	-2.35	3.31	63	-0.29
	Russia	757	0.50	12.03	1028	0.32	**Average**			-1.74			-0.72
	Chile	1877	-0.24	15.01	1657	-0.10	**Median**			0.26			0.21
S8	Norway	4338	-1.90	19.66	2968	-0.99	**Standard Dev.**			7.85			2.90
	Iran	6	1.00	5.47	188	0.86							
	Moldova	791	0.35	6.57	278	0.77							
	Serbia	504	0.56	8.88	533	0.54							

Based on above (Eq. 3), P17: Expected cases / 1M population and P18: Per unit difference between actual and expected cases is calculated as shown in Table 4. Median deviation 0.267, average -1.746 and standard deviation 7.854 is obtained. In the table, negative values indicate actual number of cases reported less than mathematically expected and vice versa. Country with maximum deviation is found to be China having per unit deviation value -54.08 as against average of -1.746.

(A) Combination parameter regression model: As per combination parameter regression model (Eq. 2), calculated P19: Combination parameter value is shown in Table 4, using n=5 for multiple most correlating parameters i.e. P2, P3, P8, P9 and P14. A scatter plot is generated as shown in Fig. 4 with P19: Combination parameter on X axis and P15: Total cases / 1M population on Y axis. The best fit power regression trendline equation is obtained With $R^2 = 0.726$, as given below-

$$(Y) = \mathbf{4.781}X^{2.159} \quad \text{(Eq. 4)}$$

where, Y=sets average (P16: combination parameter)

X= Total Cases / 1M Population

Based on combination parameter, mathematical estimation of P20: Expected cases / 1M population and P21: Per unit difference between expected and actual cases is calculated as shown Table 4. Obtained P21 indicates more accurate effectiveness of countries in controlling the infection by adopting other measures which are not considered in this study. Again country with maximum per unit deviation is found to be China having value -14.78 against average value -0.721. It has been observed that above average value becomes closer to zero to thee decimal places i.e. 0.0008 after ignoring values of 4 most deviating countries. Thus proposed combination parameter regression model indicates accurate identification of noise data in the population and accuracy of predictive regression mathematical model.

With respect to single parameter regression model, there is no significant change in median deviation value, which is 0.215 compared to earlier 0.267. However significant change in average value and standard deviation which are now closer to zero viz. -0.721 vs earlier -1.746 and 2.90 vs earlier 7.854, respectively. Further average value becomes almost zero after removal of noise data, this suggests a better correlation and estimation than single parameter.

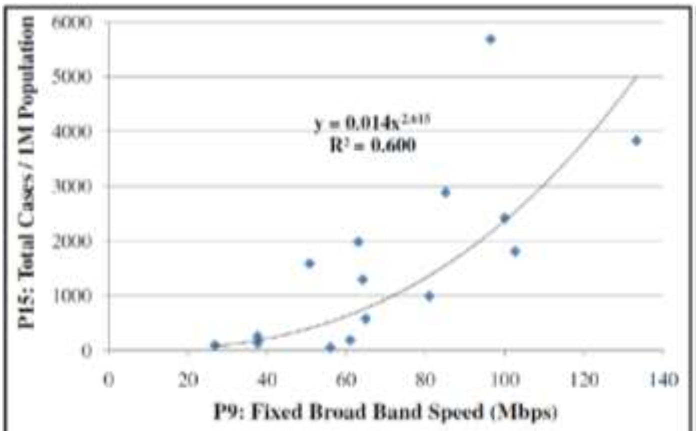

Fig. 3: Scatter plot
(Single parameter regression model)

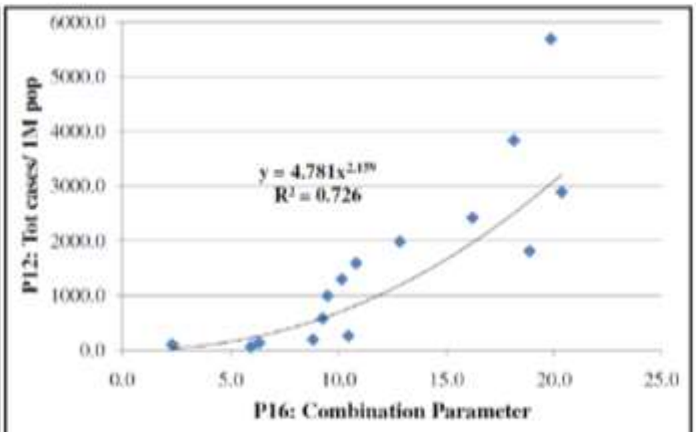

Fig. 4: Scatter plot
(Combination parameter regression model)

The combination regression model using number of parameters having high correlation values is having around 8% higher correlation value ρ_c=0.846 than maximum of a single parameters ρ_s=0.784. Also, a better R^2 value is obtained for combination parameter R_c^2 = 0.726 than for single parameter R_s^2=0.600. Therefore, above technique of finding combination parameter can be used to find effective correlating combination parameter. Also could be used more accurately to calculate expected infection cases in the country with given set of affecting parameters. Per unit difference between actual and expected cases could be used as a better indicator of effectiveness of controlling the infection by using other measures by the country.

4. Conclusion

COVID-19 infections and related death per unit population of country analyzed by statistical method with respect to various demographic and environmental parameters exhibits very good correlation with respect to net migrants per unit population and fixed broad band's both speed and number of subscriptions in the country, used with wireless technology as compared to other analyzed parameters. The study shows correlation of ill health effects of high speed electromagnetic radiation on COVID-19 infection and related death.

More experiments and research may be carried out to find exact mechanism how an EMF radiation influences COVID-19 infection and death. If the effect of EMF emitted radiation is permanent and non reversible, reducing EMF radiation may not show immediate benefit however will help in long term. But if effect is temporary and reversible, immediate benefits can be expected. Additionally before launching of any new technology with higher data transfer rate and higher frequency like 5G, detailed clinical trials must be done since existing levels only are showing ill health effects.

Declaration

Analysis and results are based on data available from open sources as referred appropriately and statistical techniques. Author declares no relevant conflicts of interest to the content of manuscript.

References:

[1] Worldometers.info. 11 May 2020. Coronavirus. retrieved from https://www.worldometers.info/coronavirus/

[2] Worldometers.info. 11 May 2020. Population. retrieved from https://www.worldometers.info/world-population/population-by-country/

[3] Wikipedia. 15 May 2020. List of countries by 4G LTE penetration (2016-19). retrieved from https://en.wikipedia.org/wiki/List_of_countries_by_4G_LTE_penetration

[4] Wikipedia, 15 May 2020, List of countries by Internet connection speeds (May 2019), retrieved from https://en.wikipedia.org/wiki/List_of_countries_by_Internet_connection_speeds

[5] Wikipedia. 15 May 2020. List of countries by number of broadband Internet (2019), retrieved from https://en.wikipedia.org/wiki/List_of_countries_by_number_of_broadband_Internet_subscriptions

[6] Speedtest Global Index, Mar 2020. Mobile and Fixed Broadband speed, retrieved from https://www.speedtest.net/global-index

[7] Weatherbase. 15 May 2020. Weather Averages - All Countries. retrieved from https://www.weatherbase.com/weather/countryall.php3

[8] WHO. 12 April 2020. BCG Vaccine. retrieved from https://www.who.int/

[9] Worldometers.info. 15 May 2020. Environment. CO2 Emissions. retrieved from https://www.worldometers.info/co2-emissions/

[10] Jahansson Olle and ELSEVIER. 2009. Disturbance of immune system by electromagnetic field – A potentially underlying cause for cellular damage and tissue repair reduction which could lead to disease and impairment. Journal of Pathophysiology. 16:157-177.

[11] Bonhomme-Faivre Laurence, Marion Sylvie, Forestier François, Santini Roger, Auclair Huguette. Nov 2003. Effects of Electromagnetic Fields on the Immune Systems of Occupationally Exposed Humans and Mice. Arch Environ Health, 58(11): 712-717.

[12] WHO. 1 Jun 2020. Basic and safely managed sanitation services Data by country. retrieved from https://apps.who.int/gho/data/node.main.WSHSANITATION?lang=en

[13] Wikipedia. 1 Jun 2020. Vegetarianism by country. retrieved from https://en.wikipedia.org/wiki/Vegetarianism_by_country

[14] Myers Samuel S, Ryan Wessells K, Kloog Itai, Zanobetti Antonella and Schwartz Joel. 16 July 2015. Effect of increased concentrations of atmospheric carbon dioxide on the global threat of zinc defi ciency: a modelling study.. www.lancet.com/lancetgh.

[15] Smith Matthew R and Myers Samuel S. SEPTEMBER 2018. Impact of anthropogenic CO2 emissions on global human nutrition. Nature Climate Change, 8: 834–839.

[16] Stanwell-Smith R. 2001. Hygiene and the Immune System. Journal of Infection, 43: 61–64.

[17] Lippi Giuseppe, Sanchis-Gomar Fabian, Henry M. Brandon. 27 April 2020. Association between Environmental Pollution and Prevalence of Coronavirus disease 2019 (COVID-19) in Italy, medRxiv

[18] Travaglio Marco, Popovic Rebeka, Yu Yizhou, Leal Nuno Santos, Miguel Luis. 16 April 2020. Links between air pollution and COVID-19 in England

[19] Simko Myrtill, Mattsson Mats-Olof. 2019. International Journal of Environmental Research and Public Health, 16:3406

[20] Aminoff Sarah, Foster Susan, Kheel Kate. 14 April 2020. Retrieved from http://www.avaate.org/spip.php?article2871