Supplementary Material

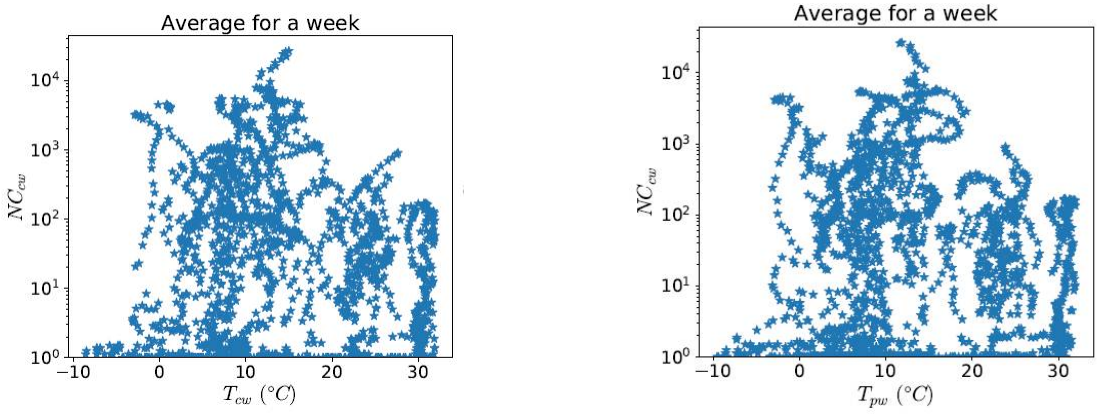
Investigating the Impact of Regional Temperature on COVID-19 Pandemic During 2020

Murk Marvi, Asad Arfeen, Mohammed Raza Mehdi, and Zahid Rauf

May 12, 2021

In order to further support the research presented in the paper and for verification of the results, we performed a number of experiments by making slightly different assumptions.

In order to analyze the relationship between average temperature and new cases (NC) reported per day for the current week, the scatter plots are reported in Figure S1 by considering data from January to March 2020. Figure S1a depicts the average NC reported in the current week (NC_{cw}) against average temperature of the current week (T_{cw}), whereas Figure S1b depicts NC_{cw} against the average temperature of the past week (T_{pw}). Here, by the current week we mean the past 7 days, and by the past week we mean the past 8 to 15 days. The similar patterns are observed in both the plots reported in Figure S1, that is, the locations bearing temperature in a moderate range has reported on average a higher number of cases.



(**a**) (**b**)

Figure S1. Scatter plot of average NC reported in the current week (NC_{cw}) in each location against: (**a**) average temperature of the current week (T_{cw}), (**b**) average temperature of the past week (T_{pw}).

We fit simple linear models by assuming NC_{cw} as a dependent and T_{cw} or T_{pw} as an independent variable, and the results are reported in Table S1. A significant inverse relationship has been observed between NC_{cw} and T_{cw}, and it is in accordance with the results reported in Section 5.2 where NC reported per day is considered as a dependent

variable instead of NC_{cw}. Moreover, a significant relationship has not been observed between NC_{cw} and T_{pw}. This implies that average temperature of the past week has not significantly affected the average NC reported in the current week.

Table S1. Analysis of simple linear models obtained by assuming NC_{cw} or ND_{cw} as a dependent and T_{cw} or T_{pw} as an independent variable.

$\widehat{NC}_{cw} = \beta_0^c + \beta_1^c T_{cw}$		$\widehat{NC}_{cw} = \beta_0^c + \beta_1^c T_{pw}$		$\widehat{ND}_{cw} = \beta_0^d + \beta_1^d T_{cw}$		$\widehat{ND}_{cw} = \beta_0^d + \beta_1^d T_{pw}$	
(β_0^c, β_1^c)	p-value	(β_0^c, β_1^c)	p-value	(β_0^d, β_1^d)	p-value	(β_0^d, β_1^d)	p-value
(6879, −51.58)	0.006	(6067, −9.64)	0.61	(126.4, −0.25)	0.387	(118,0.186)	0.517

We also fit simple linear models by assuming average of new deaths (ND) reported in the current week (ND_{cw}) as a dependent and T_{cw} or T_{pw} as an independent variable, and the results are reported in Table S1. A significant relationship has not been observed between ND_{cw} and T_{cw}, and it is in accordance with the results reported in Section 5.2 where ND reported per day is considered as a dependent variable instead of ND_{cw}. Similarly, a significant relationship has not been observed between ND_{cw} and T_{pw}. This implies that average temperature of the past week has not significantly affected the average ND reported in the current week.

Although the main focus of the manuscript is about investigating the impact of regional temperature on COVID-19 pandemic, for additional insights we also fit simple linear models by assuming average relative humidity (RH) of the current week (H_{cw}) or the past week (H_{pw}) as an independent variable. From the results reported in Table S2, a significant positive relationship has been observed between NC_{cw} and T_{cw} or T_{pw}. Similarly, a significant inverse relationship has been observed between ND_{cw} and T_{cw} or T_{pw}.

Table S2. Analysis of simple linear models obtained by assuming NC_{cw} or ND_{cw} as a dependent and H_{cw} or H_{pw} as an independent variable.

$\widehat{NC}_{cw} = \beta_0^c + \beta_1^c H_{cw}$		$\widehat{NC}_{cw} = \beta_0^c + \beta_1^c H_{pw}$		$\widehat{ND}_{cw} = \beta_0^d + \beta_1^d H_{cw}$		$\widehat{ND}_{cw} = \beta_0^d + \beta_1^d H_{pw}$	
(β_0^c, β_1^c)	p-value	(β_0^c, β_1^c)	p-value	(β_0^d, β_1^d)	p-value	(β_0^d, β_1^d)	p-value
(2856,45)	0.00	(3384, 37.9)	0.00	(167.2,−0.68)	0.00	(174, −0.78)	0.00

We perform multiple LR analysis by assuming nineteen independent variables defined in Table 2 and NC_{cw} as a dependent variable. The results of multiple LR analysis performed by assuming average temperature and RH of the current week (T_{cw}, H_{cw}) are reported in Figure S2a. Moreover, the results of multiple LR analysis performed by assuming average temperature and RH of the past week (T_{pw}, H_{pw}) are reported in Figure S2b. A significant inverse relationship has been observed between NC_{cw} and T_{cw} or T_{pw}. This is in accordance with the results reported in Section 5.3 wherein the number of NC reported per day is considered as a dependent variable. Moreover, a significant relationship between NC_{cw} and T_{cw} or T_{pw} has not been observed. This is also in accordance with the results reported in Section 5.3.

We also perform multiple LR analysis by assuming nineteen independent variables defined in Table 2 and ND_{cw} as a dependent variable. The results of multiple LR analysis performed by assuming average temperature and RH of the current week (T_{cw}, H_{cw}) are reported in Figure S3a. Moreover, the results of multiple LR analysis performed by assuming average temperature and RH of the past week (T_{pw}, H_{pw}) are reported in Figure S3b. A significant inverse relationship has been observed between ND_{cw} and T_{cw} or T_{pw}. This is in accordance with the results reported in Section 5.3 wherein the number of ND reported per day is considered as a dependent variable. Moreover, a significant relationship between ND_{cw} and T_{cw} or T_{pw} has not been observed. This is also in accordance with the results reported in Section 5.3.

In a nutshell, the following insights are extracted from the above mentioned analysis.

- Under simple LR analysis, a significant inverse relationship has been observed between NC_{cw} and T_{cw}. However, a significant relationship has not been observed between NC_{cw} and T_{pw}. This implies that the variations in average temperature of the past week have not significantly affected the average NC reported in the current week. Similarly, a significant relationship between ND_{cw} and T_{cw} or T_{pw} has not been observed. This implies that the variations in average temperature of the current or past week have not significantly affected the average ND reported in the current week.

- Under multiple LR analysis, a significant inverse relationship has been observed between NC_{cw} and T_{cw} or T_{pw}. Similarly, a significant inverse relationship has been observed between ND_{cw} and T_{cw} or T_{pw}. This implies that the variations in average temperature of the current or past week have significantly affected the average NC or ND reported in the current week.

- Under simple LR analysis, a significant positive relationship has been observed between NC_{cw} and H_{cw} or H_{pw}. Similarly, a significant inverse relationship has been observed between ND_{cw} and H_{cw} or H_{pw}. This implies that the variations in average RH of the current or past week have significantly affected the average NC or ND reported in the current week.

- However, under multiple LR analysis, a significant relationship between NC_{cw} and H_{cw} or H_{pw} has not been observed. Similarly, a significant relationship between ND_{cw} and H_{cw} or H_{pw} has not been observed.

From above insights, it is evident that NC_{cw} has shown a significant inverse relationship with average temperature under all the conducted analysis except one. Similarly, ND_{cw} has shown either a significant inverse or an insignificant relationship with average temperature depending on the conducted analysis. Moreover, NC_{cw} has shown either a significant positive or an insignificant relationship with average RH depending on the conducted analysis. Similarly, ND_{cw} has shown either a significant inverse or an insignificant relationship with average RH depending on the conducted analysis.

OLS Regression Results

Dep. Variable:	avg_NCCW	R-squared:	0.763
Model:	OLS	F-statistic:	1715.
Method:	Least Squares	Prob (F-statistic):	0.00
No. Observations:	10122		
Df Residuals:	10102		
Df Model:	19		

	coef	t	P>|t|
Intercept	9222.5548	2.767	0.006
total_cases	0.0086	52.602	0.000
total_deaths	0.0279	3.602	0.000
total_cases_per_million	0.0869	6.058	0.000
total_deaths_per_million	-3.8951	-5.394	0.000
population	1.869e-06	5.251	0.000
population_density	-0.0224	-0.276	0.782
median_age	-216.1237	-4.537	0.000
aged_65_older	103.4816	0.680	0.496
aged_70_older	344.3180	1.676	0.094
gdp_per_capita	-0.0098	-1.236	0.217
cardiovasc_death_rate	-4.5922	-2.558	0.011
diabetes_prevalence	141.2690	3.235	0.001
female_smokers	2.7390	0.172	0.864
male_smokers	35.4353	3.321	0.001
hospital_beds_per_thousand	-193.1478	-4.436	0.000
life_expectancy	-214.0369	-4.112	0.000
human_development_index	1.43e+04	4.433	0.000
avg_TCW	-29.5960	-2.398	0.017
avg_HCW	-0.1932	-0.030	0.976

(**a**)

OLS Regression Results

Dep. Variable:	avg_NCCW	R-squared:	0.763
Model:	OLS	F-statistic:	1715.
Method:	Least Squares	Prob (F-statistic):	0.00
No. Observations:	10122		
Df Residuals:	10102		
Df Model:	19		

	coef	t	P>|t|
Intercept	9481.7515	2.847	0.004
total_cases	0.0086	52.679	0.000
total_deaths	0.0282	3.654	0.000
total_cases_per_million	0.0879	6.143	0.000
total_deaths_per_million	-3.9376	-5.449	0.000
population	1.85e-06	5.197	0.000
population_density	-0.0176	-0.218	0.828
median_age	-215.3716	-4.521	0.000
aged_65_older	96.3051	0.633	0.526
aged_70_older	355.0971	1.729	0.084
gdp_per_capita	-0.0097	-1.230	0.219
cardiovasc_death_rate	-4.7774	-2.658	0.008
diabetes_prevalence	143.2411	3.282	0.001
female_smokers	1.1760	0.074	0.941
male_smokers	36.8398	3.450	0.001
hospital_beds_per_thousand	-196.3724	-4.511	0.000
life_expectancy	-214.5241	-4.123	0.000
human_development_index	1.421e+04	4.425	0.000
avg_TPW	-35.7133	-2.947	0.003
avg_HPW	-1.1398	-0.175	0.861

(**b**)

Figure S2. Results of multiple LR analysis by assuming NC_{cw} as a dependent variable. (**a**) By considering average temperature and RH of the current week. (**b**) By considering average temperature and RH of the past week. Here, the considered variables are said to be significantly affecting the dependent variable NC_{cw} if the corresponding p-value ($P > |t|$) is smaller than $\alpha = 0.05$.

OLS Regression Results

Dep. Variable:	avg_NDCW	R-squared:	0.665
Model:	OLS	F-statistic:	1057.
Method:	Least Squares	Prob (F-statistic):	0.00
No. Observations:	10122		
Df Residuals:	10102		
Df Model:	19		

	coef	t	P>|t|
Intercept	425.9804	7.078	0.000
total_cases	-1.257e-05	-4.282	0.000
total_deaths	0.0060	42.803	0.000
total_cases_per_million	0.0021	8.294	0.000
total_deaths_per_million	-0.1351	-10.358	0.000
population	3.826e-08	5.953	0.000
population_density	0.0053	3.592	0.000
median_age	1.0783	1.254	0.210
aged_65_older	-5.7438	-2.091	0.037
aged_70_older	6.9150	1.864	0.062
gdp_per_capita	-0.0011	-7.603	0.000
cardiovasc_death_rate	0.0488	1.504	0.133
diabetes_prevalence	3.1889	4.043	0.000
female_smokers	1.7311	6.008	0.000
male_smokers	-2.1962	-11.398	0.000
hospital_beds_per_thousand	0.8374	1.065	0.287
life_expectancy	-4.8995	-5.213	0.000
human_development_index	32.8602	0.564	0.573
avg_TCW	-1.4608	-6.554	0.000
avg_HCW	0.1562	1.340	0.180

(**a**)

OLS Regression Results

Dep. Variable:	avg_NDCW	R-squared:	0.665
Model:	OLS	F-statistic:	1057.
Method:	Least Squares	Prob (F-statistic):	0.00
No. Observations:	10122		
Df Residuals:	10102		
Df Model:	19		

	coef	t	P>|t|
Intercept	421.7026	7.013	0.000
total_cases	-1.242e-05	-4.237	0.000
total_deaths	0.0060	42.869	0.000
total_cases_per_million	0.0022	8.524	0.000
total_deaths_per_million	-0.1362	-10.436	0.000
population	3.796e-08	5.906	0.000
population_density	0.0052	3.571	0.000
median_age	1.1188	1.301	0.193
aged_65_older	-5.8824	-2.143	0.032
aged_70_older	7.0876	1.911	0.056
gdp_per_capita	-0.0011	-7.731	0.000
cardiovasc_death_rate	0.0490	1.511	0.131
diabetes_prevalence	3.1930	4.052	0.000
female_smokers	1.7160	5.952	0.000
male_smokers	-2.1907	-11.362	0.000
hospital_beds_per_thousand	0.8119	1.033	0.302
life_expectancy	-4.8828	-5.197	0.000
human_development_index	35.0698	0.605	0.545
avg_TPW	-1.4976	-6.844	0.000
avg_HPW	0.1704	1.453	0.146

(**b**)

Figure S3. Results of multiple LR analysis by assuming ND_{cw} as a dependent variable. (**a**) By considering average temperature and RH of the current week. (**b**) By considering average of temperature and RH for the past week. Here, the considered variables are said to be significantly affecting the dependent variable ND_{cw} if the corresponding p-value ($P > |t|$) is smaller than $\alpha = 0.05$.