nature portfolio

https://doi.org/10.1038/s43588-021-00136-6

# Supplementary information

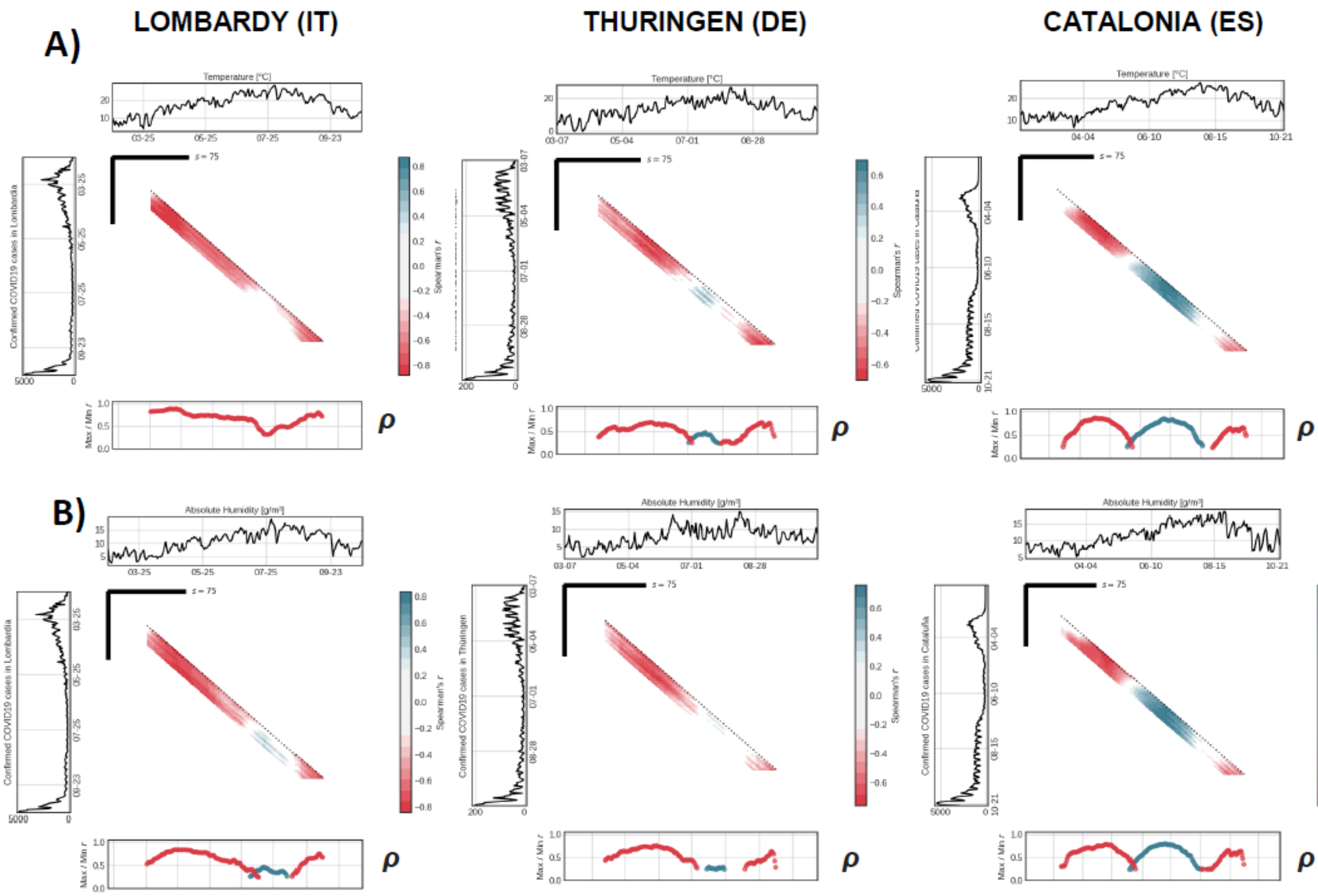
# Climatic signatures in the different COVID-19 pandemic waves across both hemispheres

In the format provided by the authors and unedited

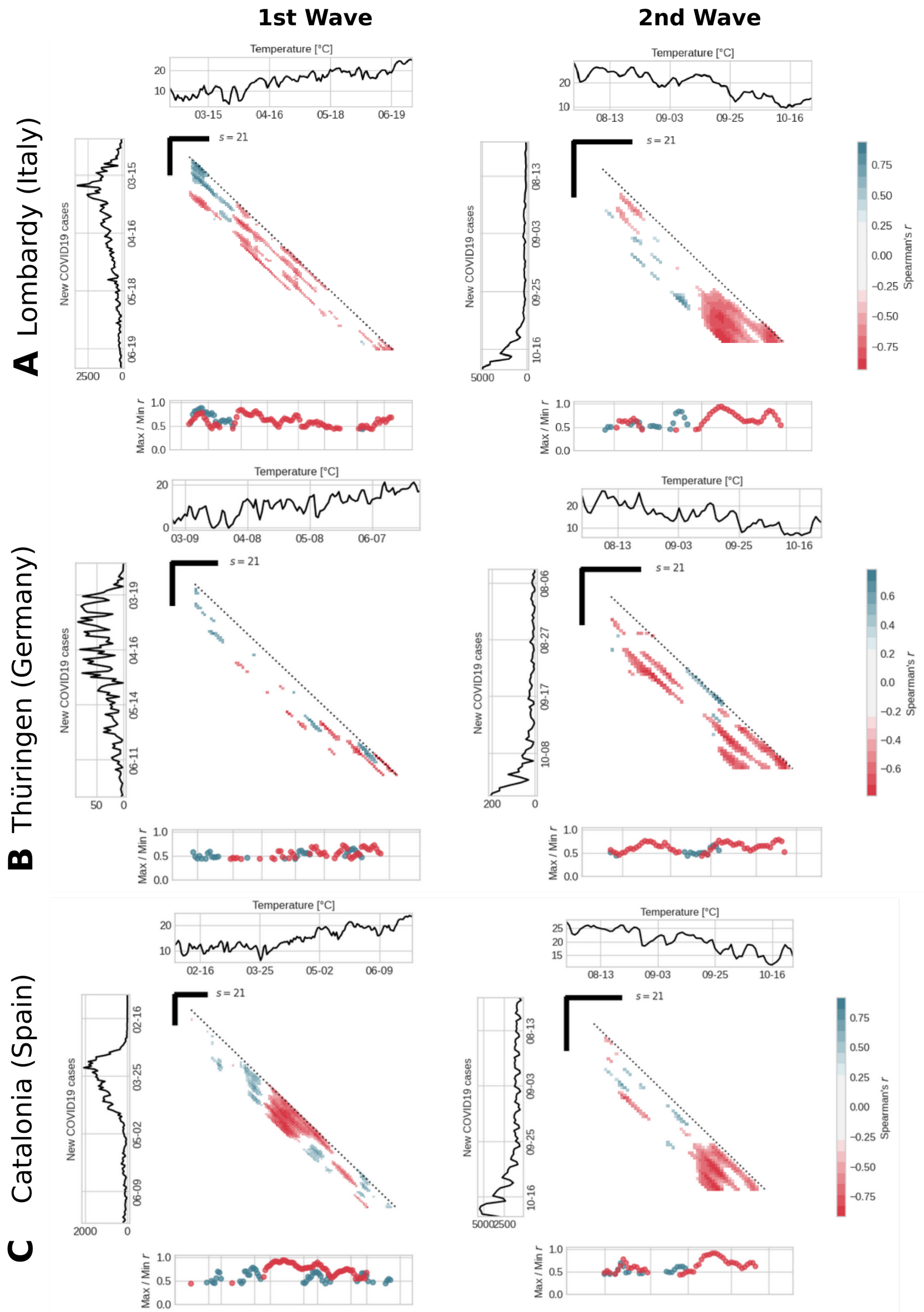
# Supplementary Information: Climatic signatures in the different COVID-19 pandemic waves across both hemispheres

**Alejandro Fontal**[1], **Menno J. Bouma**[1], **Adrià San José**[1], **Leonardo Lopez**[1], **Mercedes Pascual**[2], **and Xavier Rodó**[1,3,✉]

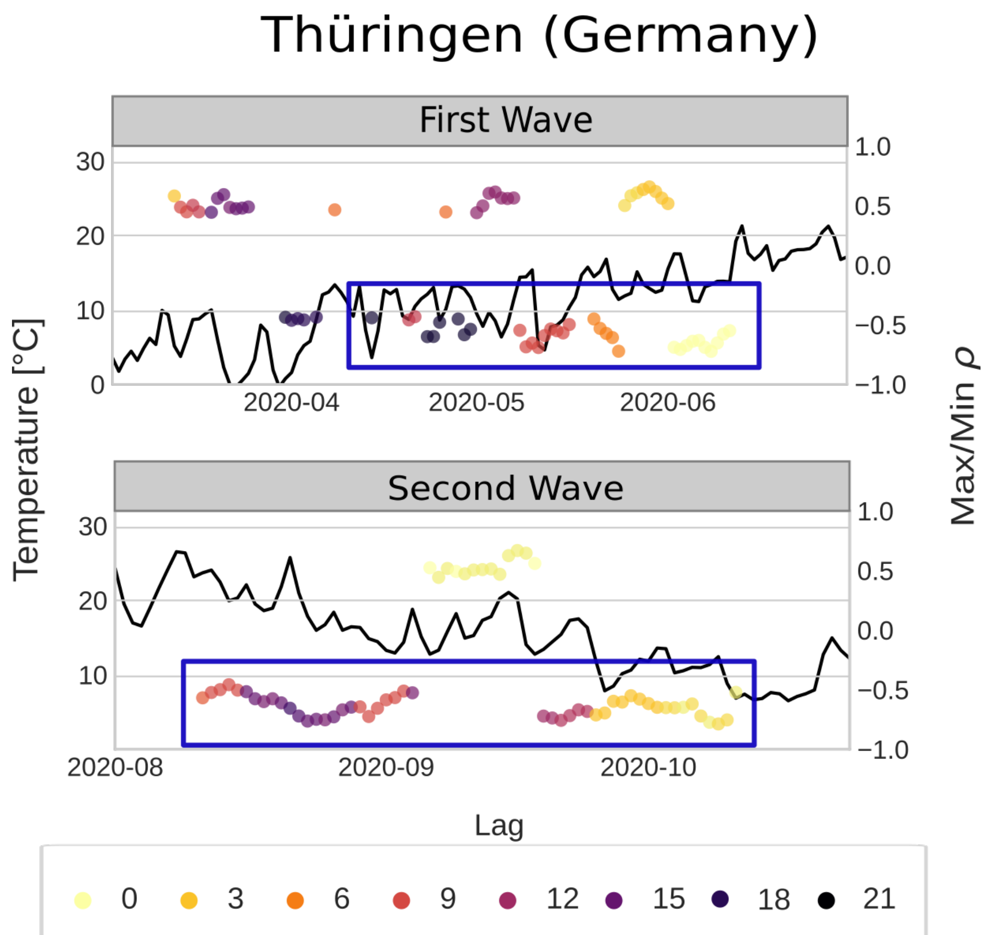
[1]Climate and Health (CLIMA) Program, Barcelona Institute for Global Health (ISGlobal), Barcelona, Catalonia, Spain
[2]Department of Ecology and Evolution. University of Chicago, Chicago, Illinois, USA
[3]ICREA, Barcelona, Catalonia, Spain



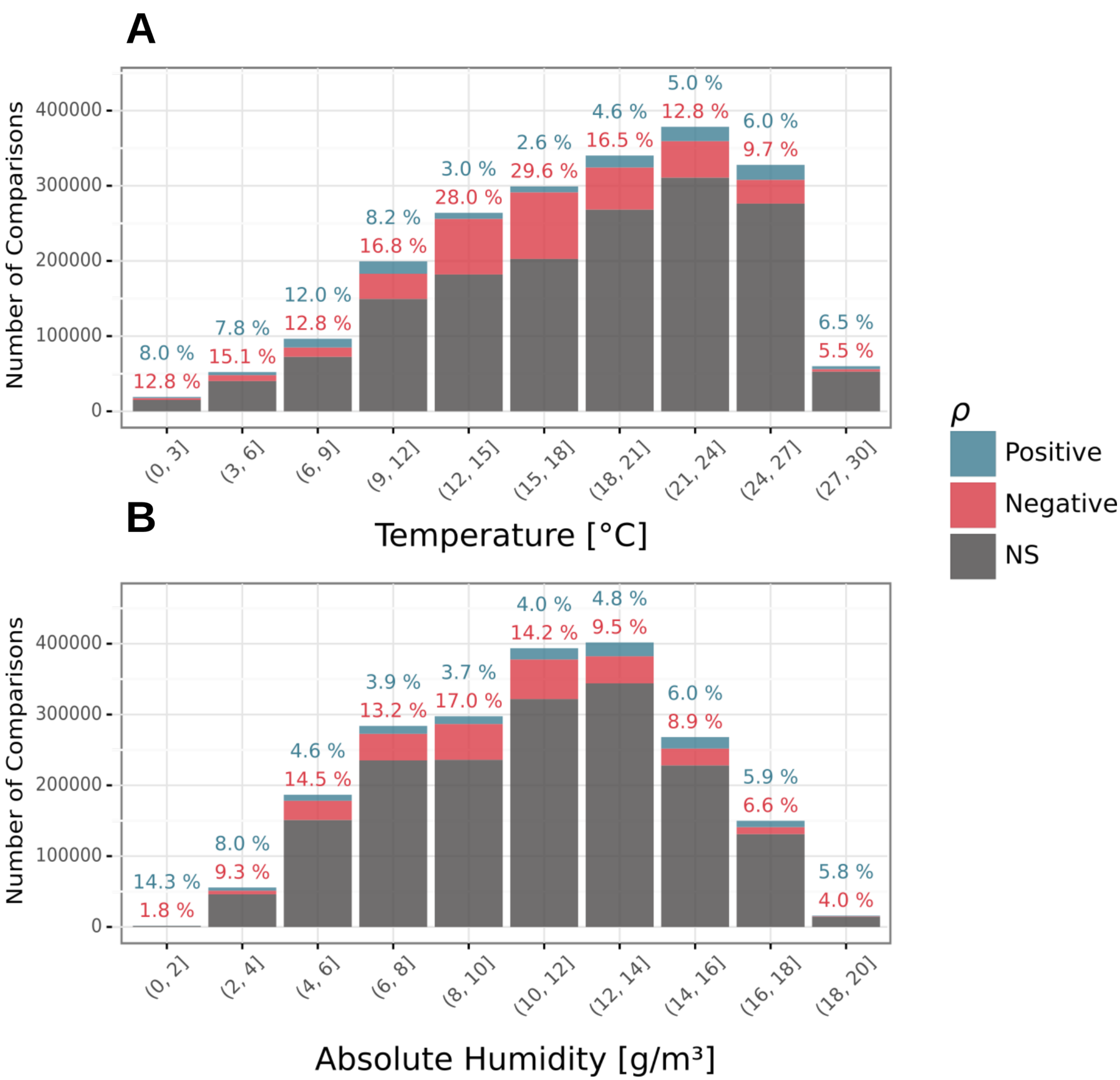

**Supplementary Figure 1. Long-window SDC analyses evaluating the correlation of Temperature and Absolute Humidity with COVID-19 cases for three European regions.** Two-way SDC plots (s=75 and with lags from 0 to 21 days) for the aggregated COVID19 daily new cases time series of three regions in highly affected countries in Europe, namely Italy (Lombardia); Germany (Thüringen) and Spain (Catalonia), and their daily population density-weighted averages of (A) T and (B) AH from March to October 2020. In the figures, each cell of the grid is coloured according to the Spearman correlation coefficient for two s-day centred windows for each of the time series, with distances from the diagonal indicating the temporal lag across the two time series compared. Panels above and left display the time series and panels below show the maximum correlation coefficient obtained for each data point in time. Only those comparisons with a significant p-value at an $\alpha$=0.05 from a non-parametric randomized test are shown.



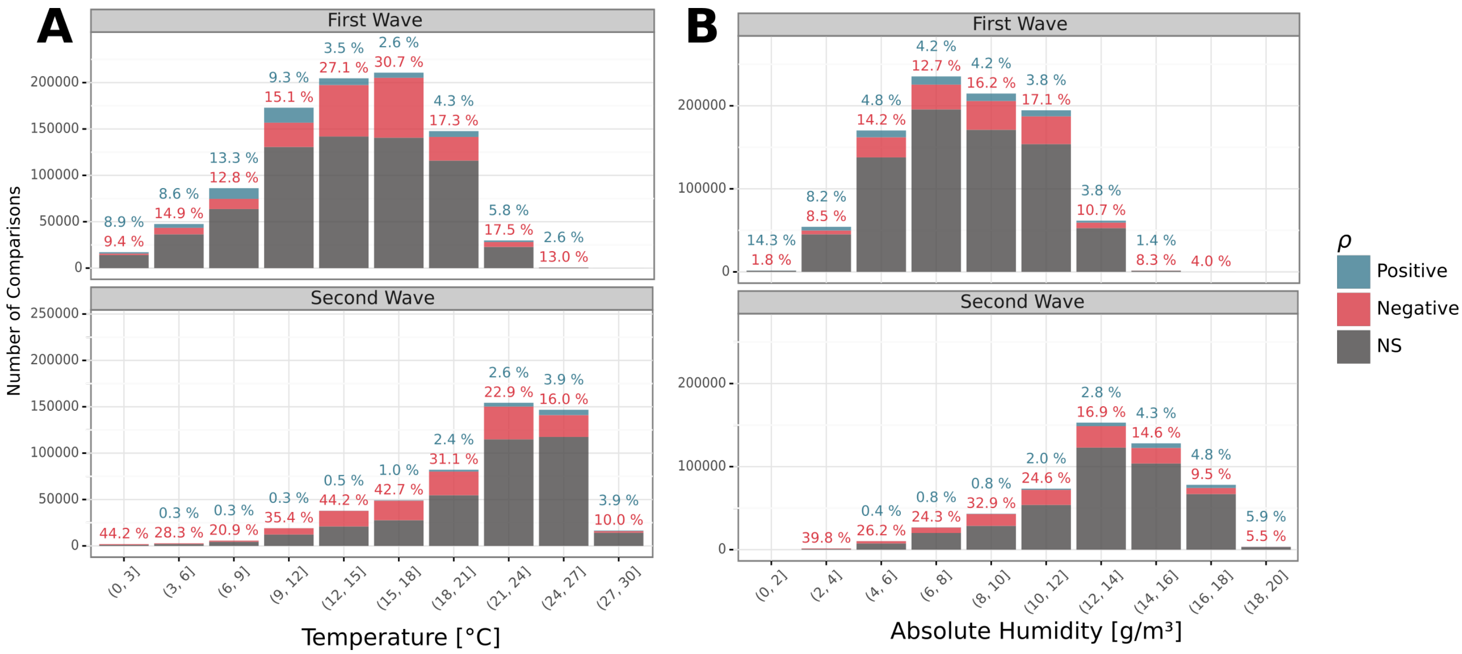

**Supplementary Figure 2. Short-window SDC analyses evaluating transitory correlations between Temperature and COVID-19 cases in three European regions.** These SDC analyses were, performed at a high temporospatial resolution, with the time series are broken down in two intervals: from March to June 2020 to represent the first wave (left), and from August to October 2020 to represent the second wave (right). Fragment sizes used in these analyses are s=21 days, which allows to identify multiple short-term transient couplings between the COVID19 incidence and meteorological factors. Three subnational regions are shown (A-Lombardy in Italy, B-Thüringen in Germany and, C- Catalonia in Spain).



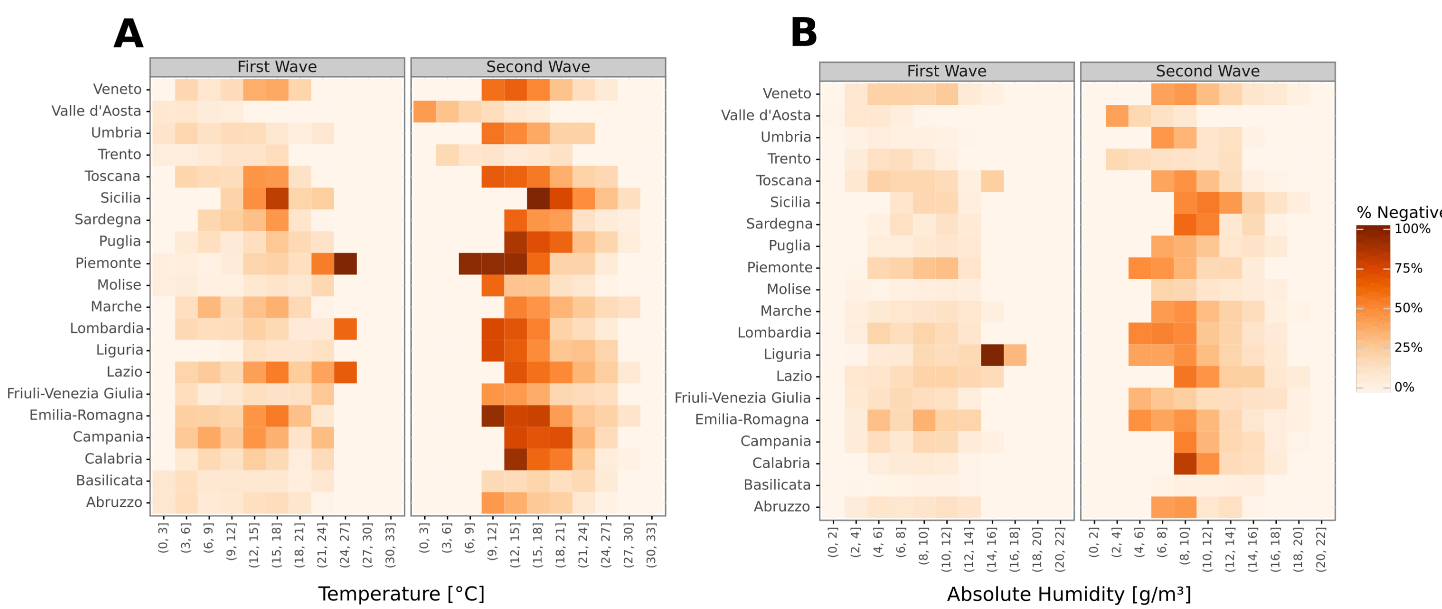

**Supplementary Figure 3. Evaluation of the dominant lags at which temperature correlates with COVID-19 cases for the initial waves of the pandemic in Thüringen.** As in Figure 3B,C but for the Thüringen region (Germany).



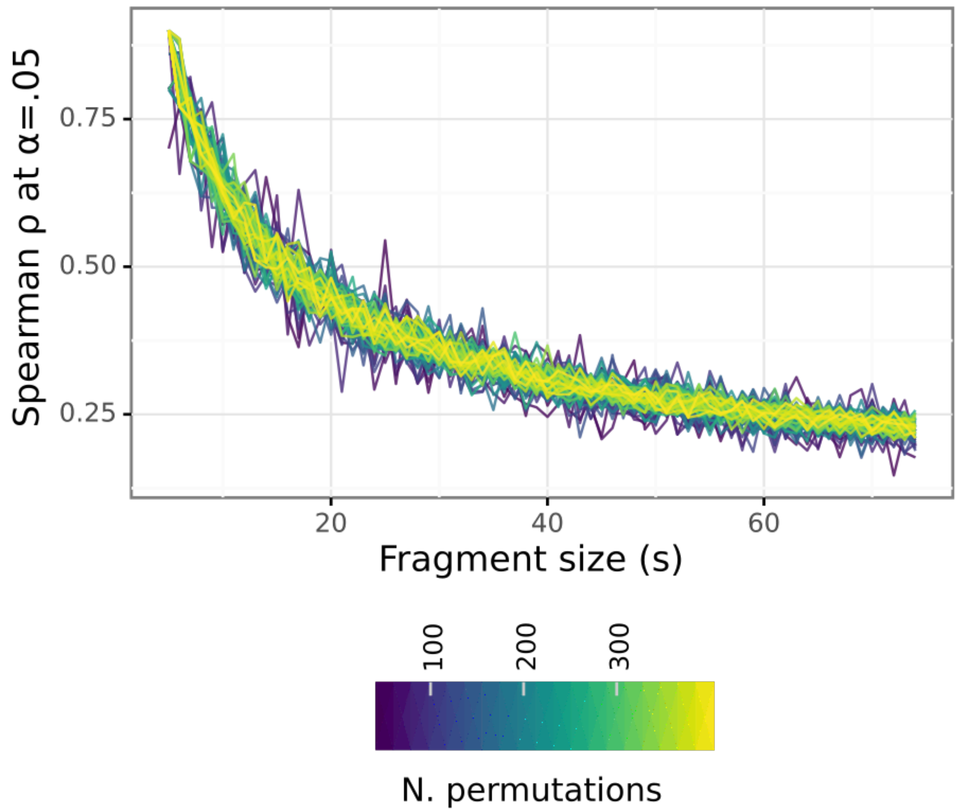

**Supplementary Figure 4. Distribution of the temperature and absolute humidity ranges studied and the proportion of significant correlations with COVID-19 cases in Italian regions.** Proportion of significant correlations of (A) T and (B) AH with COVID19 cases. SDC analysis (s=21 days) is computed with respect to all the comparisons made containing each range of values of the forcing variable, for time series covering all italian regions (n=20). Colours denote negative (red), positive (blue) or non-significant (dark gray) Spearman rank correlations (see Methods).



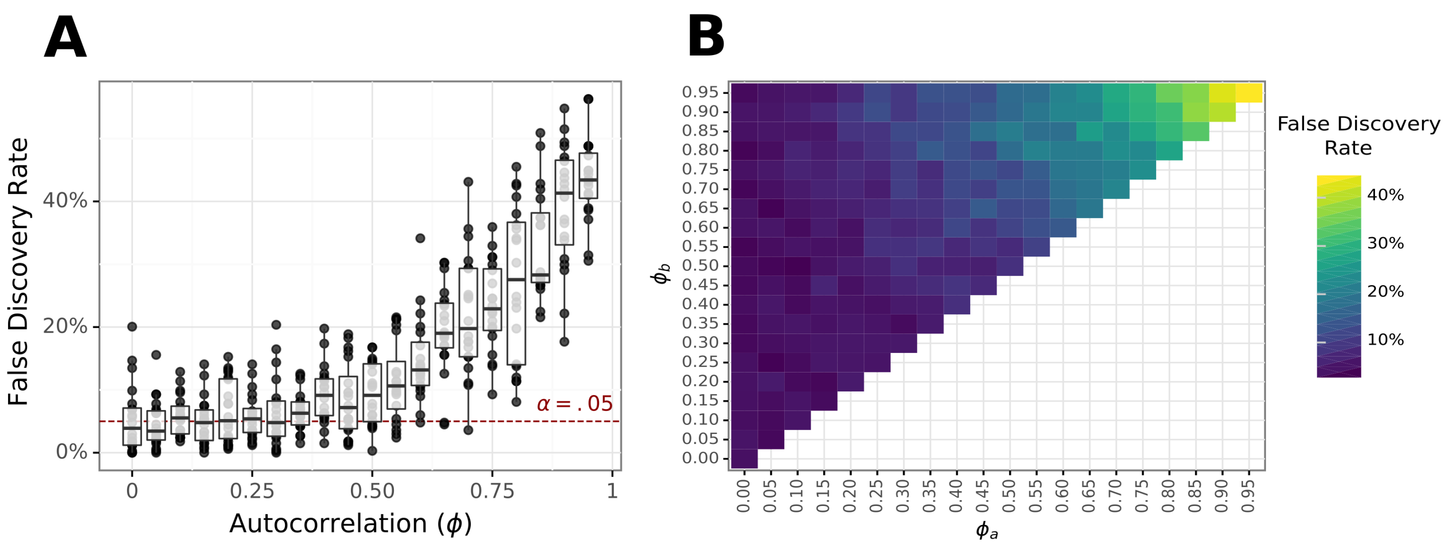
**Supplementary Figure 5. Distribution of the temperature and absolute humidity ranges studied and the proportion of significant correlations with COVID-19 cases in Italian regions for the two initial waves.** Same as Supplementary Figure 4 but separating between the (A) T and; (B) AH, for the first (top) and second (bottom) pandemic waves.



**Supplementary Figure 6. Comparison of the proportion of negative correlations between temperature/humidity and COVID-19 at different variable ranges for all Italian regions.** Heatmap representation for all separate regions in Italy, of the proportion of significant negative relationships obtained with a SDC analysis for s=21. Results are shown for the range of values covered by each meteorological variable analyzed (T, AH) and for each of the Italian regions studied. (A) T; (B) AH. Color scale denotes the percentage of significant negative monotonic relationships in temperate conditions. A and B are split between the first and second waves, respectively.




**Supplementary Figure 7. Evaluation of the relationship between Fragment size and significance threshold for Spearman correlation between two white noise time series.** Line plot showing the smallest value of Spearman ρ which yields a significant estimate (at α=.05) as a function of the fragment size s when running a non-parametric permutation test. Here, the permutation test computes the Spearman coefficient ρ between two white noise time series, with each permutation randomly shifting the indices of the series, breaking their temporal shape. This permutation test enables to estimate the probability of finding spurious significant correlations, and can thus be used as a non-parametric test of significance for pairs of time series of interest. The threshold decays rapidly as the fragment size grows, with high ρ values being rarer the longer the time series compared. As indicated by the colour of each line, estimation of the significance threshold becomes more stable as the number of permutations used to estimate it increases.




**Supplementary Figure 8. Evaluation of the effect that autocorrelation in time series fragments has on the likelihood of finding spurious couplings.** To estimate this, we generated pairs of autocorrelated (red) noise series of 365 time steps, with $\mu$=0, $\sigma$=1 and varying degrees of autocorrelation (parameter $\phi$ in the figure) from 0 (equivalent to white noise) to .95 in steps of .05. We did this 20 times for every unique combination of parameters to get a robust estimate. We then looked for significant couplings in either direction in each of these synthetic time series pairs with an SDC analysis (s=30) and estimated, for each of them, the False Discovery Rate (FDR, also Type I Error), as the rate of couplings deemed as significant at an $\alpha = .05$. In A, we can see a how the FDR changes as a function of the autocorrelation of the two fragment pairs (here we only show those cases where $\phi_a = \phi_b$). In B, we show the average FDR of the tests made with each pair of autocorrelation values. In both cases, it is quite notable that the chance of finding spurious couplings increases monotonically with the autocorrelation of the time series evaluated.

| Location and wave | Model | $\chi^2$ | AIC | BIC |
|---|---|---|---|---|
| **Barcelona (first wave)** | $\beta(T)$ | 1.10E-05 | -6.31E-03 | -6.26E-03 |
| | Constant $\beta$ | 2.23E-04 | -7.10E-02 | -7.06E-02 |
| | Seasonal $\beta$ | 2.13E-04 | -7.09E-02 | -7.00E-02 |
| **Barcelona (second wave)** | $\beta(T)$ | 5.10E-04 | -5.43E-03 | -5.38E-03 |
| | Constant $\beta$ | 4.10E-03 | -3.33E-02 | -3.28E-02 |
| | Seasonal $\beta$ | 4.13E-03 | -3.53E-02 | -3.47E-02 |
| **Barcelona (third wave)** | $\beta(T)$ | 1.66E-04 | -5.36E-03 | -5.32E-03 |
| | Constant $\beta$ | 2.65E-03 | -3.06E-02 | -3.02E-02 |
| | Seasonal $\beta$ | 2.45E-03 | -3.03E-02 | -3.00E-02 |
| **Catalunya (first wave)** | $\beta(T)$ | 4.84E-06 | -3.89E-03 | -3.85E-03 |
| | Constant $\beta$ | 3.54E-04 | -2.98E-02 | -2.76E-02 |
| | Seasonal $\beta$ | 3.51E-04 | -2.96E-02 | -2.71E-02 |
| **Catalunya (second wave)** | $\beta(T)$ | 2.98E-05 | -6.78E-03 | -6.73E-03 |
| | Constant $\beta$ | 2.50E-04 | -6.66E-02 | -6.60E-02 |
| | Seasonal $\beta$ | 2.45E-04 | -6.61E-02 | -6.54E-02 |
| **Catalunya (third wave)** | $\beta(T)$ | 1.97E-06 | -4.78E-03 | -4.75E-03 |
| | Constant $\beta$ | 1.78E-05 | -3.78E-02 | -3.66E-02 |
| | Seasonal $\beta$ | 1.77E-05 | -3.77E-02 | -3.65E-02 |
| **Thüringen (first wave)** | $\beta(T)$ | 3.48E-06 | -2.76E-03 | -2.71E-03 |
| | Constant $\beta$ | 2.55E-05 | -3.45E-02 | -3.39E-02 |
| | Seasonal $\beta$ | 2.54E-05 | -3.44E-02 | -3.37E-02 |
| **Thüringen (second wave)** | $\beta(T)$ | 1.66E-06 | -4.52E-03 | -4.46E-03 |
| | Constant $\beta$ | 3.55E-05 | -2.96E-02 | -2.66E-02 |
| | Seasonal $\beta$ | 3.51E-05 | -2.93E-02 | -2.57E-02 |
| **Lombardia (first wave)** | $\beta(T)$ | 3.83E-06 | -2.78E-03 | -2.75E-03 |
| | Constant $\beta$ | 2.44E-05 | -3.41E-02 | -3.27E-02 |
| | Seasonal $\beta$ | 2.43E-05 | -3.40E-02 | -3.26E-02 |
| **Lombardia (second wave)** | $\beta(T)$ | 3.69E-04 | -4.76E-03 | -4.71E-03 |
| | Constant $\beta$ | 2.34E-05 | -3.40E-02 | -3.27E-02 |
| | Seasonal $\beta$ | 2.31E-05 | -3.39E-02 | -3.26E-02 |

**Supplementary Table 1.** Residual statistical measures. The indices measure the goodness of fit for the different models in the different locations and waves.

## Supplementary Discussion

We observed consistent negative correlations between COVID-19 and weather conditions, T and AH, for regions and individual countries. This consistency across scales is observed in 2020 for the rise of the first wave, its decline towards the advancing summer in both hemispheres, and the second wave after the summer. These associations remain robust and become stronger on finer temporal scales within countries during the seasonal rise and fall of cases, reflecting similar patterns of temporal variation between disease incidence and climate factors. The strong correlations become more frequent where and when seasonal mean temperatures fall roughly between 12-18°C, and AH levels, between 4-12 g/m³, conditions that diverge from those observed for the mid-winter global emergence of SARS-CoV-2 (S1), suggesting permissive temperatures above 5°C. Our results are also consistent with earlier studies more limited in scope (S2–S4) and close to the ranges estimated here for both flu A and flu B. We then separated the three different epidemic waves up to January 2021 to compare the frequency of significant negative local correlations found for different ranges of AH, again by means of SDC analyses (s=21). This was done for all Italian and Spanish regions across the three defined periods of the COVID19 pandemic (Extended Data Figure 6 and see Methods). The aim was to enlarge the temporal span of potential correlations, given that in this last wave the NH was entering its coldest months. Results reinforce the range of AH and T values previously described, while also serve to clarify the lower edges of the environmental relation to COVID-19 (Extended Data Figure 6).

The global correlation between $R_0$ and climate covariates across countries specifically considered the timing of initiation of the first epidemic in each location to avoid the confounding effects of later interventions. Then, by addressing correlation patterns locally in time and zooming in for finer scales, the SDC analysis detected time intervals with similar patterns of variation that are not simply the result of ups and downs over the seasons or the typical durations of intervention policies. Identification of significant correlations between temporal patterns at different and embedded window lengths is considered an efficient way to infer plausible causality with this multiscale approach (S5, S6). The new MSDCis particularly useful for this purpose, as shown for different countries in continents at the height of their respective winter seasons. The implemented randomizations indicated that the spurious occurrence of the described correlations from either noisy fluctuations or red-noise autocorrelated structures, is highly unlikely (e.g. see Supplementary Figures 7 and 8).

A higher plausible role of aerosols than fomites on surfaces, indicates the need to examine climate forcing in mathematical models formulated to represent these routes of transmission explicitly (e.g.(S7)). The open question of whether and how this representation matters goes beyond modeling of COVID-19 and applies to other respiratory viruses.

## Supplementary References


[S1] M. M. Sajadi, P. Habibzadeh, A. Vintzileos, S. Shokouhi, F. Miralles-Wilhelm, and A. Amoroso. Temperature, Humidity and Latitude Analysis to Predict Potential Spread and Seasonality for COVID-19. *SSRN*, page 3550308, March 2020. ISSN 1556-5068. doi: 10.2139/ssrn.3550308.

[S2] M. Holtmann, M. Jones, A. Shah, and G. Holtmann. Low ambient temperatures are associated with more rapid spread of COVID-19 in the early phase of the endemic. *Environmental Research*, 186:109625, July 2020. ISSN 0013-9351. doi: 10.1016/j.envres.2020.109625.

[S3] S. Mallapaty. Why does the coronavirus spread so easily between people? *Nature*, 579(7798):183, 2020. ISSN 1476-4687. doi: 10.1038/d41586-020-00660-x.

[S4] J. Wang, K. Tang, K. Feng, X. Lin, W. Lv, K. Chen, and F. Wang. High Temperature and High Humidity Reduce the Transmission of COVID-19. SSRN Scholarly Paper ID 3551767, Social Science Research Network, Rochester, NY, March 2020.

[S5] X. Rodo and M. A. Rodriguez Arias. A new method to detect transitory signatures and local time/space variability structures in the climate system: The scale-dependent correlation analysis. *Climate Dynamics*, 27:441–458, October 2006. doi: 10.1007/s00382-005-0106-4.

[S6] J. REMAIS, B. ZHONG, E. J. CARLTON, and R. C. SPEAR. Model approaches for estimating the influence of time-varying socio-environmental factors on macroparasite transmission in two endemic regions. *Epidemics*, 1(4):213–220, December 2009. ISSN 1755-4365. doi: 10.1016/j.epidem.2009.10.001.

[S7] R. O. J. H. Stutt, R. Retkute, M. Bradley, C. A. Gilligan, and J. Colvin. A modelling framework to assess the likely effectiveness of facemasks in combination with 'lock-down' in managing the COVID-19 pandemic. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 476(2238):20200376, June 2020. doi: 10.1098/rspa.2020.0376. Publisher: Royal Society.