

Supplementary material

COVID-19 pandemic increased the magnitude of mortality risks associated with cold temperature in Italy: A nationwide time-stratified case-crossover study

Table of Contents

1. **The calibration of ERA5 temperature data**
2. **Fig S1.** Distribution of weather stations that used to calibrate the ERA5 daily temperature and dew point temperature.
3. **Fig S2.** 10-fold cross-validation of the random forest model that using ERA5 daily mean temperature and dew point temperature and other covariates to predict the daily mean temperature and dew point temperature observed by weather stations.
4. **Fig S3.** Cumulative temperature-mortality associations along 21 lag days in 2020 COVID-19 pandemic and the same months in the single year of 2015-2019.
5. **Fig S4.** Cumulative temperature-mortality associations with different lag days in 2020 COVID-19 pandemic and the same months during 2015-2019.
6. **Table S1.** The distribution of temperature from March to May in Italy, 2015–2020.

1. The calibration of ERA5 temperature data

Data collection

We collected the ERA5 hourly surface (at 2 meters above the land surface) ambient temperature and ambient dew point temperature at $0.1^\circ \times 0.1^\circ$ spatial resolution from January 1st, 2015 to May 31st, 2020 for Italy and its surrounding areas. This data was downloaded from the ERA5-Land hourly data (<https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-land?tab=form>).

To valid and calibrate the ERA5 data, we downloaded hourly ambient temperature and ambient dew point temperature observed by weather stations from the Integrated Surface Database (<https://www.ncdc.noaa.gov/isd>) hosted by National Oceanic and Atmospheric Administration of US. A total of 519 weather stations with hourly weather observations during 2015-2020 for Italy and its surrounding areas were used to validate the ERA5 hourly dataset (**Fig S1**). Hourly observations were transformed into daily observations by averaging all hourly observations within each day. Validation and calibration were based on a total of 692,002 daily observations. We did not use the daily temperature observations from weather stations for our analyses mainly for two reasons: 1) the spatial coverage of those stations is not satisfactory, with many provinces not covered; 2) the daily temperature observations from many weather stations were often incomplete. In contrast, the ERA5 data has perfect spatial and temporal coverage, despite its less accuracy.

To improve the performance of the calibration model, we also collected data on elevation and leaf area index (LAI) that may affect temperatures. LAI is an indicator for the density of vegetation cover. The monthly LAIs for both high and low vegetation at $0.1^\circ \times 0.1^\circ$ spatial resolution were downloaded from ERA5-Land monthly averaged data (<https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-land-monthly-means?tab=form>). The global elevation data at 7.5 arc second (i.e., $7.5/3600$ degree \times $7.5/3600$ degree, about $250\text{m} \times 250\text{m}$) spatial resolution was downloaded from the NASA Shuttle Radar Topographic Mission (SRTM <http://srtm.csi.cgiar.org/srtmdata/>). We aggregated the elevation data to $0.1^\circ \times 0.1^\circ$ spatial resolution by averaging elevation values of the 48×48 blocks of $7.5/3600$ degree \times $7.5/3600$ degree pixels. ERA5 daily temperature, ERA5 LAI, and elevation data were linked to all weather stations according to date, longitude, and latitude.

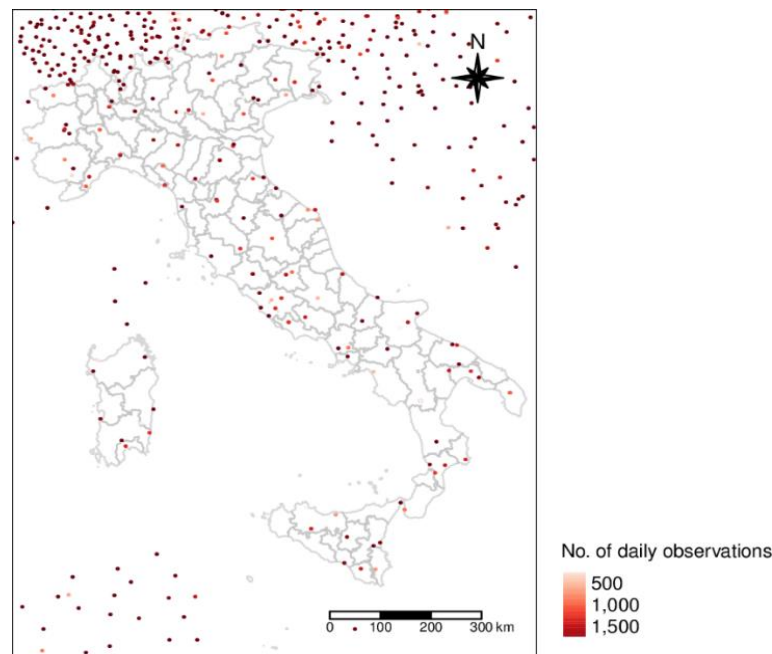


Fig S1. Distribution of weather stations that used to calibrate the ERA5 daily temperature and dew point temperature.

Calibration model

We used a random forest model to calibrate the ERA5 daily mean temperature and dew point temperature against the observations from weather stations. In the random forest models, the dependent variable was either the daily mean temperature or daily mean dew point temperature, and the independent variables (predictors) included ERA5 daily mean temperature, ERA5 daily mean dew point temperature, monthly LAI, elevation, year, day of year (1 to 366), latitude, and longitude. The random forest models were fitted by the “ranger” (version 0.10.1) R package with default settings. The models built based on all the 692,002 daily observations from 519 weather stations were then used to calibrate the ERA5 daily temperature observations. The calibrated ERA5 daily mean temperature and ERA5 daily mean dew point temperature were then used to calculate daily mean relative humidity, using the algorithm provided by the “humidity” R package.

To evaluate the accuracy of the calibration model, we performed a 10-fold cross-validation by randomly splitting the 519 stations into 10 folds. As showed in **Fig S2**, the calibrated ERA5 temperatures (i.e., predicted values by the calibration model) were highly correlated with the observed values from weather stations, suggesting a high accuracy of the calibration model.

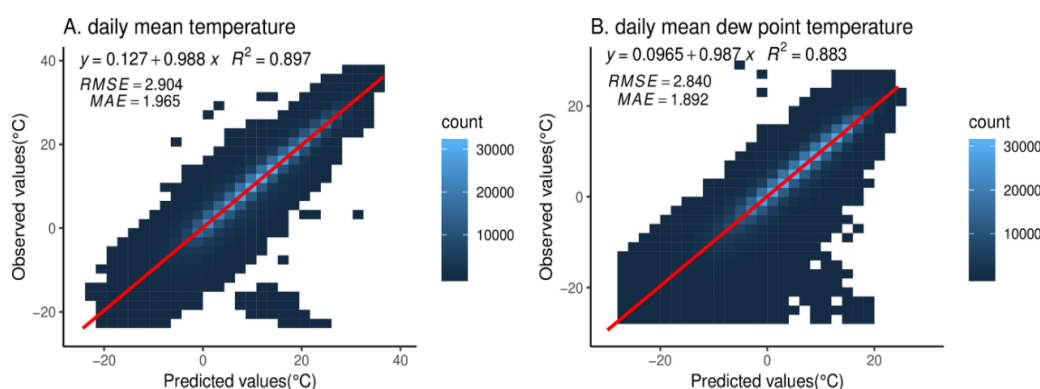


Fig S2. 10-fold cross-validation of the random forest model that using ERA5 daily mean temperature and dew point temperature and other covariates to predict the daily mean temperature and dew point temperature observed by weather stations.

Notes: other covariates included elevation, year, day of year (1 to 366), latitude, longitude, and leaf area index. We performed a 10-fold cross-validation by randomly splitting the 519 stations into 10 folds. RMSE, root mean square error. MAE, mean absolute error.

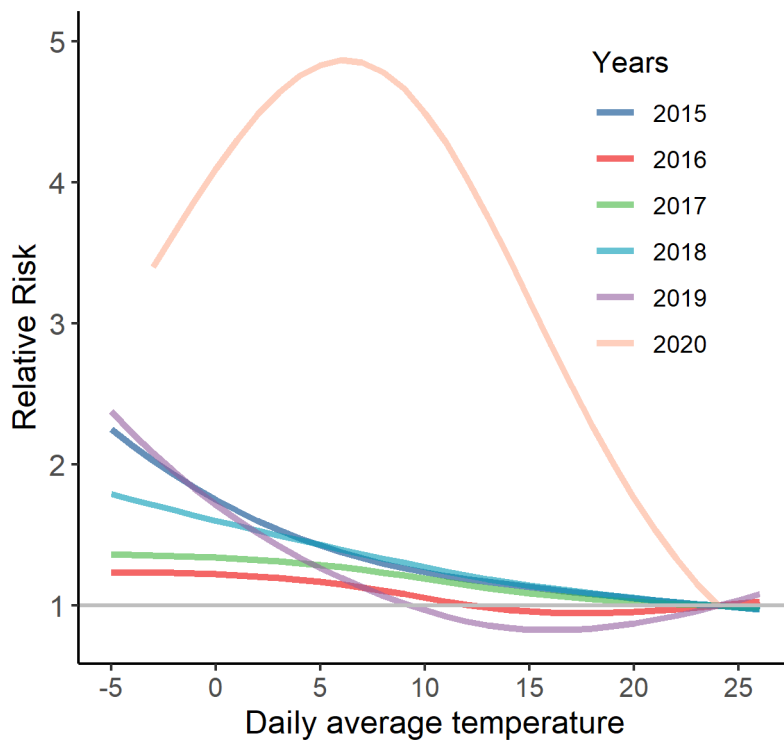


Fig S3. Cumulative temperature–mortality associations along 21 lag days in 2020 COVID-19 pandemic and the same months in the single year of 2015-2019.

Notes: Relative Risk: overall cumulative relative risk of all-cause mortality in previous 21 days, compared with the minimum mortality temperature of 24°C.

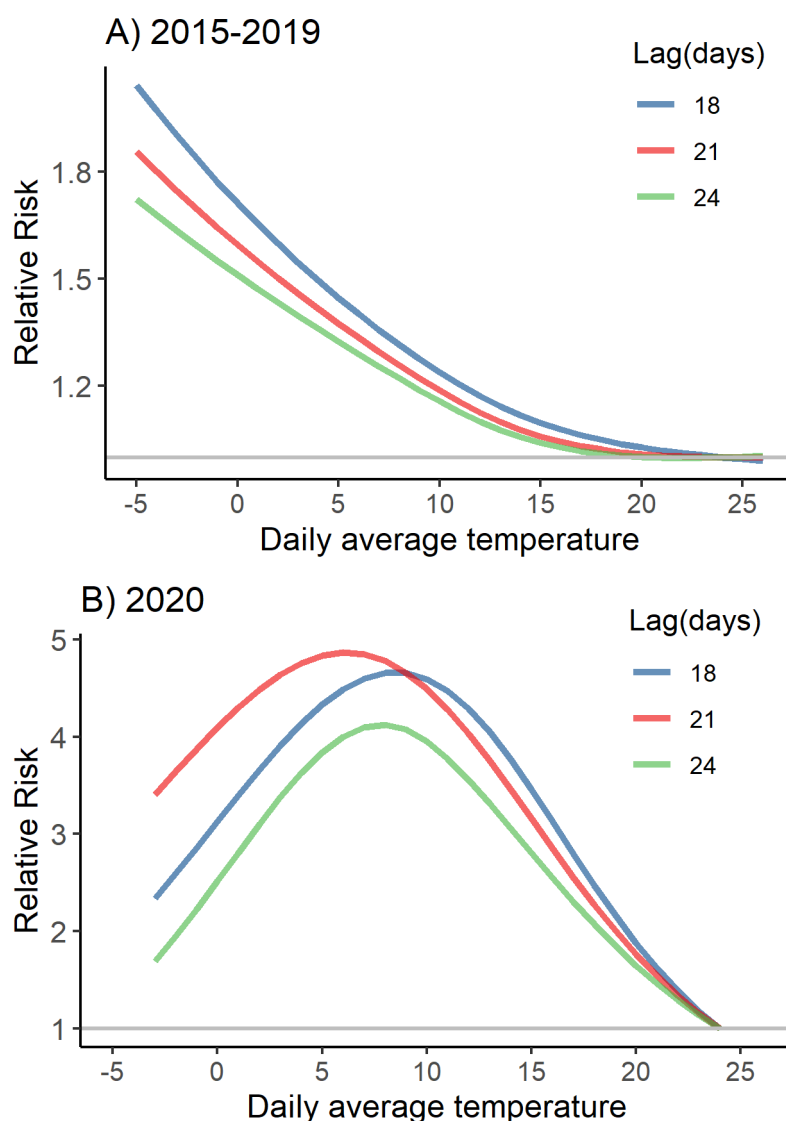


Fig S4. Cumulative temperature–mortality associations with different lag days in 2020 COVID-19 pandemic and the same months during 2015-2019.

Notes: A) 2015-2019: the normal period from March 1st to May 31st in 2015-2019; B) 2020: the pandemic period from March 1st to May 31st, 2020; We used a natural cubic spline with 3 degrees of freedom for the temperature distribution, and a natural cubic spline with 3 degrees of freedom for lags days; Relative Risk: overall cumulative relative risk of all-cause mortality in previous 21 days, compared with the minimum mortality temperature of 24°C.

Table S1. The distribution of daily temperature (°C) from March to May in Italy, 2015–2020.

year	Mean	SD	MIN	P25	P50	P75	MAX
2015	12.32	4.51	-3.11	8.77	12.36	15.69	23.71
2016	12.20	4.14	-3.89	9.10	12.74	15.11	24.36
2017	13.01	4.07	-0.62	10.23	12.77	15.49	25.97
2018	12.75	5.18	-6.79	9.07	13.19	16.76	24.68
2019	11.48	3.10	-0.37	9.47	11.51	13.62	19.80
2020	12.52	4.40	-3.13	9.46	12.73	15.83	24.50

Notes: SD: Standard deviation; MIN: minimum daily average temperature; P25: 25th percentile; P50: 50th percentile; P75: 75th percentile; MAX: maximum daily average temperature.