Environmental Research 219 (2023) 115062

Contents lists available at ScienceDirect

Environmental Research

journal homepage: www.elsevier.com/locate/envres

High-resolution spatiotemporal modeling of daily near-surface air temperature in Germany over the period 2000–2020

Nikolaos Nikolaou [a,b,*], Marco Dallavalle [a,b], Massimo Stafoggia [c], Laurens M. Bouwer [d], Annette Peters [a,b], Kai Chen [e,f], Kathrin Wolf [a,1], Alexandra Schneider [a,1]

[a] Institute of Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany
[b] Institute for Medical Information Processing, Biometry, and Epidemiology, Pettenkofer School of Public Health, LMU Munich, Munich, Germany
[c] Department of Epidemiology, Lazio Regional Health Service, Rome, Italy
[d] Climate Service Center Germany (GERICS), Helmholtz-Zentrum Hereon, Hamburg, Germany
[e] Department of Environmental Health Sciences, Yale School of Public Health, New Haven, CT, USA
[f] Yale Center on Climate Change and Health, Yale School of Public Health, New Haven, CT, USA

ARTICLE INFO

Keywords:
Near-surface air temperature
Land surface temperature
Spatiotemporal modeling
Validation
Exposure assessment
Environmental epidemiology

ABSTRACT

The commonly used weather stations cannot fully capture the spatiotemporal variability of near-surface air temperature (T_{air}), leading to exposure misclassification and biased health effect estimates. We aimed to improve the spatiotemporal coverage of T_{air} data in Germany by using multi-stage modeling to estimate daily 1 × 1 km minimum (T_{min}), mean (T_{mean}), maximum (T_{max}) T_{air} and diurnal T_{air} range during 2000–2020. We used weather station T_{air} observations, satellite-based land surface temperature (LST), elevation, vegetation and various land use predictors. In the first stage, we built a linear mixed model with daily random intercepts and slopes for LST adjusted for several spatial predictors to estimate T_{air} from cells with both T_{air} and LST available. In the second stage, we used this model to predict T_{air} for cells with only LST available. In the third stage, we regressed the second stage predictions against interpolated T_{air} values to obtain T_{air} countrywide. All models achieved high accuracy ($0.91 \leq R^2 \leq 0.98$) and low errors (1.03 °C ≤ Root Mean Square Error (RMSE) ≤ 2.02 °C). Validation with external data confirmed the good performance, locally, i.e., in Augsburg for all models ($0.74 \leq R^2 \leq 0.99$, 0.87 °C ≤ RMSE ≤ 2.05 °C) and countrywide, for the T_{mean} model ($0.71 \leq R^2 \leq 0.99$, 0.79 °C ≤ RMSE ≤ 1.19 °C). Annual T_{mean} averages ranged from 8.56 °C to 10.42 °C with the years beyond 2016 being constantly hotter than the 21-year average. The spatial variability within Germany exceeded 15 °C annually on average following patterns including mountains, rivers and urbanization. Using a case study, we showed that modeling leads to broader T_{air} variability representation for exposure assessment of participants in health cohorts. Our results indicate the proposed models as suitable for estimating nationwide T_{air} at high resolution. Our product is critical for temperature-based epidemiological studies and is also available for other research purposes.

1. Introduction

Climate change is one of the greatest global challenges for humans and their entire living environment in the 21st century. It has been at the center of various social and research disciplines, from economics (Hertel and Rosch, 2010) and animal welfare (Lacetera, 2019) to land management and food security (Shukla et al., 2019), with a particular attention to the human health domain (Peters and Schneider, 2021; Vicedo-Cabrera et al., 2021; Watts et al., 2019). Near surface air temperature (T_{air}) is one of the most important meteorological parameters and a key indicator of climate change. T_{air} is observed to be steadily increasing globally since pre-industrial times, with the 10 warmest years on record to have occurred after 2000 (Lindsey and Dahlman, 2021). Further increases from 3 °C to 6.2 °C are expected by the end of 2100, if no action is taken (IPCC, 2022). In Germany, 32 of the last 34 years are characterized by annual T_{air} above the 1961–1990 average (DWD, 2022).

Many epidemiological studies have documented the adverse impact of T_{air} on mortality (Guo et al., 2016; Zanobetti and Schwartz, 2008) and morbidity (Ye et al., 2012), especially when exposure to extreme T_{air}

* Corresponding author. Ingolstädter Landstr. 1, D-85764, Neuherberg, Germany.
E-mail address: nikolaos.nikolaou@helmholtz-muenchen.de (N. Nikolaou).
[1] Shared last authorship.

https://doi.org/10.1016/j.envres.2022.115062
Received 3 July 2022; Received in revised form 9 December 2022; Accepted 12 December 2022
Available online 17 December 2022

values occurs (Gronlund et al., 2018; Kovats and Kristie, 2006). Besides heat waves and cold spells, increases or decreases in more moderate T_{air} ranges to which people are exposed most of the time during their life, also contribute to the observed temperature-related mortality burden (Gasparrini et al., 2015). Human health can be adversely affected by T_{air} either after short (Breitner et al., 2014) or long-term (Zafeiratou et al., 2021) exposure. Therefore, T_{air} extremes and variations pose a major threat for public health, especially with continuing global warming and the higher frequency and intensity of extreme events (Meehl and Tebaldi, 2004). In this regard, high spatiotemporally-resolved T_{air} exposure datasets are needed for improved exposure assessment in epidemiological studies.

The vast majority of environmental epidemiological studies that investigate health effects of T_{air} or implement T_{air} in their analyses as a confounder or an effect modifier, use observational data from meteorological stations, often provided by a national monitoring network. These datasets are generally highly accurate, quality controlled and publicly available and consist of various meteorological parameters, including T_{air} 2 m above the ground. However, the monitoring locations are irregularly scattered, often placed in rural or park-like environments, and their number is too limited to fully capture spatial temperature variations across urban and rural landscapes. Furthermore, in most cases airport stations are used, which, by definition, are located out of the cities. Therefore, the commonly used weather station observations are not capable to represent the full variability of T_{air} in space and in time, leading to exposure misclassification and bias of health effect estimates towards the null hypothesis of no association (Armstrong, 1998). Over the last years, researchers have developed methods to provide high-resolution spatiotemporally T_{air} exposure outputs on local, countrywide or even global scales. Several interpolation techniques have been suggested such as regression-kriging (Kilibarda et al., 2014; Li et al., 2020; Sekulić et al., 2020), modified inverse distance weighting (IDW) and thin plate spline (TPS) interpolation. For example, Sekulić et al. (2020) predicted daily mean T_{air} (T_{mean}) in 1 × 1 km across Croatia for 2008 using a regression kriging model with T_{mean} observations and geometrical temperature trend, digital elevation model (DEM) and topographic wetness index as covariates (R^2 = 0.98, Root Mean Square Error (RMSE) = 1.2 °C). Jobst et al. (2017) introduced a multi-layer approach, including TPS and lapse rate models, to estimate daily maximum T_{air} (T_{max}) and minimum T_{air} (T_{min}) in 1 × 1 km from 1990 to 2014, in the alpine Clutha catchment, New Zealand (T_{max} RMSE = 2.38 °C, T_{min} RMSE = 2.93 °C). Other studies compared multiple interpolation approaches in the same region. In middle Ebro Valley, Spain, Vicente-Serrano et al. (2003) compared the results of annual T_{air} models of global or local interpolators as well as geo-statistical and mixed methods. R^2 ranged from 0.39 (co-kriging) to 0.75 (regression--based) and RMSE from 0.80 °C (co-kriging) to 0.56 °C (IDW, r = 2). However, the traditional interpolation methods are subject to specific limitations. For instance, they are highly affected by the weather stations locations, without fully accounting for between-station variability. This issue is more profound in complex geo-climatic areas and landscapes characterized by high spatial heterogeneity. Interpolation also leads to neighbouring effects and cannot capture the T_{air} variations in city-level analysis and consequently the urban heat island (UHI) effects are not well represented. Finally, the weather stations are often poorly scattered across a country and fail to provide complete T_{air} time series.

To improve the interpolation between locations, several studies have used satellite data for their main predictors (Benali et al., 2012; Fluckiger et al., 2022; Vancutsem et al., 2010; Xu et al., 2014; Zhu et al., 2013). For instance, Xu et al. (2014) applied a linear regression and a random forest (RF) model to predict T_{max} across British-Columbia, Canada. The RF model achieved higher model's accuracy (R^2 = 0.74, mean absolute error (MAE) of 2.02 °C) in comparison with the linear regression model (R^2 = 0.64, MAE = 2.41 °C). Recently, Jin et al. (2022) estimated high-resolution spatiotemporal T_{mean} from land surface temperature (LST) and a variety of spatial predictors using a three-stage ensemble model in Sweden over a long period. The ensemble model consisted of a generalized additive model, a generalized additive mixed model, a RF model and an extreme gradient boosting model (R^2 = 0.98, RMSE = 1.38 °C). In recent years, several studies applied a multi-stage regression-based approach introduced by Kloog et al. (2014), and making use of the moderate resolution imaging spectroradiometer (MODIS) LST products to predict daily T_{air} in 1 × 1 km (Kloog et al., 2017; Rosenfeld et al., 2017; Shi et al., 2016). This approach is straightforward to model, with high accuracy and small errors.

T_{mean} is the most frequently modeled T_{air} measure and the most commonly used in studies of environmental epidemiology. However, we also need to focus on T_{min} and T_{max}. Climate change strongly affects T_{min} and T_{max} (Modala et al., 2017) and there is evidence that T_{min}, which corresponds to the nighttime temperatures, has been increased more than T_{max} during the 20th century (Gil-Alana, 2018). Due to this substantial T_{min} increase, especially the urban areas face extensive heat stress nights, a phenomenon that will be strengthened in the future (Chapman et al., 2017). Modeling T_{min} and T_{max} also facilitates the estimates of the diurnal T_{air} range (DTR). There is already evidence that DTR affects, independently from T_{mean}, the human health (Cheng et al., 2014; Davis et al., 2020). These effects are critically important for future policies implementation, but lack of broad epidemiological investigation, especially at national scale due to the scarcity of fully spatially covered, high resolution, daily DTR data.

In this study, we aimed to extend and improve the spatiotemporal coverage of T_{air} data in the complex terrain of Germany, using remote sensing and regression-based modeling. More specifically, we aimed to map daily T_{min}, T_{mean}, T_{max} and DTR in 1 × 1 km across Germany during the period 2000–2020 to provide harmonized T_{air} data for epidemiological research like the German National Cohort (NAKO) with more than 200,000 participants spread around the country.

2. Methods

2.1. Study area

Germany is located in central Europe, covering 357,021 km^2, with a population of 83.2 million people (Statistisches Bundesamt, 2022). The country consists of a diverse landscape, starting from the Alps in the south to the northern coast lines of the North and Baltic Seas, including big cities, small towns, mountains, various water bodies, forests and arable land. Elevation ranges from 3.54 m below sea level near Neuendorf-Sachsenbande to 2,962 m in the Alpine mountain Zugspitze. Climate is temperate to continental according to the Köppen climate classification. There is a warm summer humid continental climate in south-eastern regions and a temperate oceanic climate in north-western regions (Beck et al., 2018b). The lowest T_{air} ever recorded in Germany was −37.8 °C measured on February 12th, 1929 in Wolznach-Hüll (DWD, 2017), while the highest was 41.2 °C measured on July 25th, 2019 in Duisburg and in Tönisvorst of North Rhine-Westphalia (DWD, 2020). We divided Germany's mainland into 366,536 grid cells of 1 × 1 km based on the European INSPIRE (Infrastructure for Spatial Information in the European Community) standard using the Lambert Azimuthal Equal-Area projection, EPSG: 3035 (©GeoBasis-DE/BKG (2021)).

2.2. Materials

We collected a large number of publicly available earth- and satellite-based data derived from multiple sources for the period 2000–2020 across Germany, with the best fitting temporal and spatial resolution for our analysis. LST based on satellite data was the main predictor for T_{air} as they were strongly correlated. Moreover, we implemented in the modeling process various spatial predictors such as remote sensing elevation, vegetation, urban fabric, arable land, pastures, forests and inland waters to increase the percentage of explained

variance and enhance the model performance overall.

2.2.1. T_{air} data

We downloaded daily data of T_{min}, T_{mean} and T_{max} observations 2 m above the ground from 1080 sites, which are publicly available from the Deutscher Wetterdienst (DWD) - German Meteorological Service online database (DWD, 2021). All data were quality controlled by DWD and their metadata (e.g., station relocation or time zones) were provided as well. We did not find any unusual values such as temperatures lower or higher than the observed temperature extreme records in the country or any seasonal outliers. We excluded stations that stopped measuring before 2000 (N = 358), stations that weren't operating continuously over the entire study period (N = 309, > 70% NAs) and stations located outside the German mainland (N = 7). Thus, we included T_{air} data from 406 weather stations scattered across the country (Fig. 1), with each station located in a single grid cell of our gridded dataset. The DWD T_{air} dataset had only 1.4% of station-days with missing values during our study period.

2.2.2. Remote sensing data

2.2.2.1. TERRA MODIS data. We downloaded and preprocessed TERRA MODIS LST and normalized difference vegetation index (NDVI) data from the server (NASA, 2021) through the R package MODIStsp (Busetto and Ranghetti, 2016). Since TERRA MODIS started measuring on February 24th, 2000, our analysis also starts on that date.

2.2.2.1.1. LST data. LST defines the radiative temperature of the earth's surface, as derived from infrared radiation and measured in the direction of the remote sensor. We used the product MOD11A1v006 that provides LST (using the generalized split-window algorithm) data in a daily temporal resolution and a spatial resolution of 1 × 1 km, corrected for emissivity (Wan et al., 2015). We used daytime LST to model T_{max} and nighttime LST to model T_{min} and T_{mean}, as suggested by previous studies (Rosenfeld et al., 2017), regardless their quality assurance flag to avoid reducing the input sample size since for the most problematic cells affected by cloud effects, instrumental problems or other reasons, LST was not produced. For more insight into MODIS LST and its retrieval, we refer to the existing literature (Wan, 2014).

2.2.2.1.2. NDVI data. NDVI is used as a proxy for greenness and quantifies the amount of vegetation by calculating the near-infrared and the red light difference. We used the product MOD13A3v006 that provides monthly NDVI data, as greenness does not change considerably within a month, in a spatial resolution of 1 × 1 km (Didan, 2015). We also tested the enhanced vegetation index (EVI) as alternative. Since we observed strong positive correlations between NDVI and EVI (21-year average r = 0.81), negligible differences in the models' validation results (after the 3rd or 4th decimal point) and extremely correlated T_{air} predictions (r = 0.999988), we kept NDVI as model predictor.

2.2.2.2. DEM data. We used the DEM (GTOPO30) developed by the US Geological Survey's Earth Resources Observation Systems Data Center. Its spatial resolution was 30-arc-second and we aggregated it to 1 × 1 km grid cells over mainland Germany, borders and shorelines included (Fig. S1).

2.2.2.3. Land use data. From Copernicus CORINE Land Cover 2012 (CLC2012, 250 m resolution) https://land.copernicus.eu/pan-european/corine-land-cover/clc-2012 and 2018 (CLC2018, 100 m resolution) https://land.copernicus.eu/pan-european/corine-land-cover/clc2018, we extracted the variables urban fabric (classes: "continuous urban

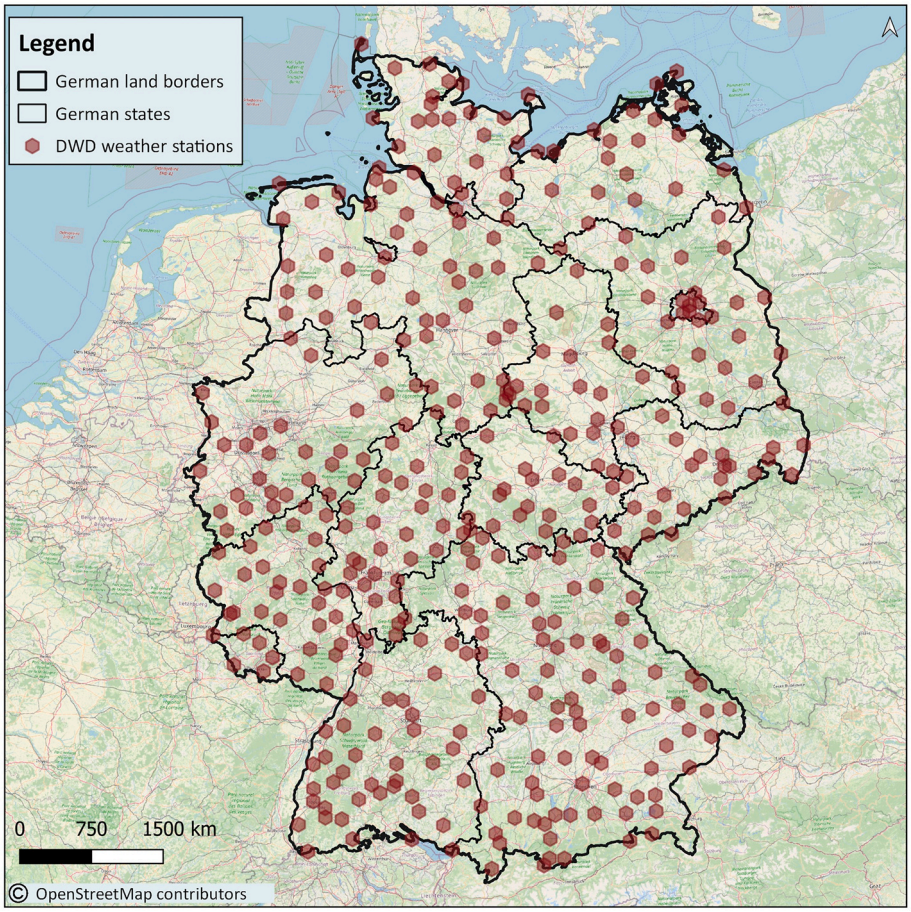

Fig. 1. Map of Germany showing the spatial distribution of the 406 T_{air} weather stations included in our analysis (2000–2020).

fabric" and "discontinuous urban fabric" from the subcategory urban fabric of the artificial surfaces category), arable land (classes: "non-irrigated arable land", "permanently irrigated arable land" and "rice fields" from the subcategory arable land of the agricultural areas category), pastures (class: "pastures" from the subcategory arable land of the agricultural areas category), forests (classes: "broad-leaved forest", the "coniferous forest" and the "mixed forest" from the subcategory forest of the forest and seminatural areas category) and inland waters (classes: "water courses" and the "water bodies" from the subcategory inland waters of the water bodies category). We then combined the classes of each variable and calculated each variable's proportion of CORINE pixels to the INSPIRE 1 × 1 km grid cells over mainland Germany. CLC2012 was used for modeling until 2016 and CLC2018 from 2017 on.

2.3. Statistical analysis and validation

We applied a three-stage regression-based model, following the approach from Kloog et al. (2014). In previous studies following the same approach basis, the results were promising. Kloog et al. (2014) mapped T_{mean} in the northeast and mid-Atlantic USA (RMSE = 2.16 °C). Shi et al. (2016) predicted T_{mean} in southeastern USA ($R^2 = 0.97$, RMSE = 1.38 °C). Rosenfeld et al. (2017) estimated T_{min}, T_{mean} and T_{max} in Israel using LST data from both TERRA and AQUA satellites (RMSE < 1.7 °C). Finally, Kloog et al. (2017) estimated T_{mean} in France ($R^2 > 0.93$, RMSE < 1.7 °C). Given the method's high accuracy, the small errors, its straightforward way of modeling and its successful application in different and geographically complex areas around the world, we also followed the basic concept of this approach, adjusting it to our needs with regard to Germany's unique spatial and geo-climate features, differentiating the process where it was necessary and introducing a TPS technique implementation on the third stage of the model. More specifically, due to Germany's unique surface and altitude fluctuation, we first adjusted our model for elevation. We also added information on water bodies, forests and arable land which corresponded to a large proportion of the country, i.e., arable land covers around 34% of Germany (and 28.32% in stage 1 - calibration stage of our analysis). Additionally, we implemented a TPS interpolation given the number and the distribution of the DWD weather stations and used grid-cell specific intercepts and slopes in the third stage to capture spatial differences in the relationship of interpolated observed T_{air} with predicted T_{air}. Our code was developed in R software, version 4.0.2 (R Core Team, 2020). The linear mixed models analyses were conducted with the R package "LM4" (Bates et al., 2014), while figures were produced either using the R package "ggplot2" (Wickham, 2009) or QGIS, version 3.10.5-A Coruña (QGIS Development Team, 2020). We applied the three-stage modeling process separately for each year and each T_{air} measure.

2.3.1. First stage: T_{air} and LST available

The first stage included grid cells where both T_{air} observations from DWD weather stations and satellite-derived LST values were available. We regressed T_{air} on LST and additional spatial information to understand and describe the T_{air}-LST relationship in the best way possible for Germany over the last two decades. A daily random slope for LST was implemented in the model to account for daily variations in the T_{air}-LST relationship.

The general mathematical formula of the first stage linear mixed effects model was the following:

$$T_{airij} = b_o + u_j + (b_1 + v_j) * LST_{ij} + b_2 * DEM_i + b_3 * NDVI_{ij} + b_4 * UrbanFabric_i + b_5 * ArableLand_i + b_6 * Pastures_i + b_7 * Forests_i + b_8 * InlandWaters_i + \varepsilon_{ij} \quad (1)$$

where,

- T_{airij} stands for the T_{air} observation at the grid cell i on day j;
- b_o and u_j stand for the fixed and the random intercepts, respectively;
- LST_{ij} stands for the daytime or nighttime LST measurement at the grid cell i on day j;
- b_1 and v_j stand for the fixed and the random slopes, respectively;
- DEM_i stands for the elevation at grid cell i;
- $NDVI_{ij}$ stands for the monthly NDVI measurement at the grid cell i on the month that day j falls in;
- $UrbanFabric_i$, $ArableLand_i$, $Pastures_i$, $Forests_i$ and $InlandWaters_i$ stand for the percentages of the urbanism, the land under temporary agricultural crops, the pastures, the forests and the water bodies at the grid cell i.
- ε_{ij} is the error term at grid cell i on day j.

For each T_{air} measure, we applied a separate regression.

2.3.2. Second stage: T_{air} not available and LST available

In the second stage, T_{air} was predicted for the combinations of grid cells and days without available T_{air} observations, but with available LST data by applying the regression coefficients derived from EQ. (1).

2.3.3. Third stage: neither T_{air} nor LST available

In the third stage, we predicted T_{air} for grid cells and days with neither T_{air} observations nor LST data. We regressed second stage T_{air} predictions against daily interpolated T_{air} values in 1 × 1 km across Germany via a linear mixed model with random grid-cell-specific intercepts and slopes. Since previous studies suggested that TPS outperformed alternative interpolation techniques such as kriging or IDW for T_{air} modeling (Wu et al., 2015), we applied it to interpolate the DWD T_{air} data. The smoothing parameter was chosen by a generalized cross validation (CV) method. We used the R package "fields" (Nychka et al., 2017) for the TPS interpolation.

The general mathematical formula of the third stage linear mixed effects model was the following:

$$\text{Second stage } T_{airij} = a_i + b_i * intT_{airij} + \varepsilon_{ij} \quad (2)$$

where,

- Second stage T_{airij} stands for the T_{air} predictions given by the second stage at the grid cell i, on the day j;
- a_i and b_i stand for the i grid-cell-specific intercepts and slopes;
- $intT_{airij}$ stands for the interpolated T_{air} values at the grid cell i on day j;
- ε_{ij} is the error term at grid cell i on day j.

For each T_{air} measure, we applied a separate regression.

After predicting T_{min}, T_{mean} and T_{max}, we also calculated the DTR by taking the difference of T_{min} from T_{max}. All dates are represented in the standard time zone for Germany, i.e., UTC+1, without adjusting for daylight-saving time.

2.3.4. Internal validation

The models' performance was evaluated through 10-fold CV separately for the first and third stage by randomly dividing the respective datasets into testing and training sets (10:90) ten times. The models were then re-fitted in each of the ten training sets and T_{air} predicted in the respective testing sets.

For the first stage, we calculated the corresponding percent of explained T_{air} variability R^2 and the RMSE between observed and predicted T_{air} for each run for the cell-days with both T_{air} and LST available. The temporal and spatial performance (R^2 and RMSE) were also computed (Shi et al., 2016). Briefly, the temporal statistics derived by regressing (a) against (b), where: (a) is the difference of the DWD observed T_{airij} of each j day with the annual DWD observed T_{airi} in weather station location i, and (b) is the difference of the predicted T_{airij}

of each j day with the annual predicted T_{air} at the same weather station location i. Spatial statistics derived by regressing (c) against (d), where: (c) is the annual DWD observed T_{airi} and (d) is the annual predicted T_{airi} in weather station location i.

For the third stage, R^2 and RMSE derived from linearly regressing the predicted T_{air} against DWD observed T_{air} for each run. Before the beginning of the modeling process, this specific sub-sample of DWD weather stations T_{air} observations in cell-days without LST available, was completely held-out from the modeling process. Thus, we used these cell-days to validate our third stage predictions and to quantify the respective errors also under conditions such as of cloudy days. We also compared the predicted T_{air} of the third stage against the second stage T_{air}, i.e., the dependent variable of the third stage model, for all grid cells across Germany.

We additionally quantified the bias through measuring the mean signed error between the DWD observed T_{air} and our models' T_{air} as predicted from all the modeling stages.

2.3.5. Validation with external data

We carried out two validations using external datasets and compared our predictions i) on the small scale with a monitoring network (HOBO-Logger) set up in the region of the city of Augsburg in 2012 in a cooperative work of Helmholtz Munich (Institute of Epidemiology, Environmental Risks) and the University of Augsburg (Institute of Geography, Physical Geography and Quantitative Methods) (Beck et al., 2018a) and ii) on the large scale with predictions from another German-wide model dataset developed by DWD within the "Testreferenzjahre" (TRY) project by using a completely different methodological approach (Krähenmann et al., 2018).

For the small-scale validation, we considered T_{air} measurements of 4 min resolution from 82 HOBO-Logger devices (ONSET, Type Pro v2) (2013–2018) with most of them to be located in the city of Augsburg where we did not have prior information from DWD on the first - calibration model's stage (Fig. S2). For a detailed description of the monitoring network and the measurements' quality assurance we refer to the corresponding paper (Beck et al., 2018a). To proceed with our comparison, we aggregated the 4-min data to daily T_{mean} values and we also considered the daily T_{min} and T_{max} values. We additionally investigated the intraseasonal models' performance.

For the large scale validation, we downloaded openly available daily T_{mean} predictions from the DWD TRY project on a 1 × 1 km spatial resolution (Krähenmann et al., 2016), generated by a 3-step interpolation method. A daily background field was constructed from a non-linear temperature gradient and it was estimated seven times a day. Then, two hourly background fields were calculated by weighting the three temporally closest background fields and they also conducted an hourly residual interpolation. For a detailed description of the modeling process, we refer to the corresponding paper (Krähenmann et al., 2018). Our comparison was restricted to mainland Germany and the overlapping time period between the two datasets from 2001 to 2012. In addition to an overall comparison of both T_{mean} models' predictions (all our model predictions against all DWD TRY model predictions across the country for every year), we also conducted several sensitivity analyses subsetting the predictions by season, without their extreme values, to their extreme values, and comparing the urban versus the rural Augsburg area.

For both 2.3.4 and 2.3.5, all stated R^2 values correspond to the fraction of variance explained by the respective models.

2.4. Case study - Augsburg

We used Augsburg as a case study to examine the spatiotemporal variability as well as the distribution of the modeled daily T_{mean} in comparison with the observed daily T_{mean} at the DWD sites. Augsburg is the third largest city in Bavaria, Germany. The overall population of its urban district and its surrounding districts (Landkreis Augsburg in the west and Aichach-Friedberg in the east), is around 900,000 people, of which approx. 300,000 live in the city center. We chose Augsburg as it is the study region of the Cooperative Health Research in the Region of Augsburg (KORA) cohort (Holle et al., 2005) and one of the 18 study centers of the NAKO study (German National Cohort Consortium, 2014).

2.5. Spatial and temporal patterns in Germany during 2000–2020

We calculated the main descriptive statistics of the three T_{air} measures and investigated the spatiotemporal patterns of T_{air} across entire Germany, but also for the 18 study centers of the NAKO study, which are scattered across the country and represent both rural and urban areas including the biggest German cities (Fig. S3), focusing on study regions that have more than 2000 inhabitants/km^2 and cover a large population percentage (Mannheim, Leipzig, Kiel, Hannover, Hamburg, Essen, Düsseldorf, Berlin and Augsburg).

3. Results

3.1. Models' accuracy

T_{air} and LST were highly correlated, with an average $R^2 = 0.91$, an intercept of 4.79 °C and a slope of 0.88 over the period 2000–2020, after regressing T_{air} against LST. Table 1 shows the prediction accuracy results for the first and the third stage model of each T_{air} measure for 2000–2020 on average. The detailed results per year are in the Supplementary material (Tables S1 and S2). For the first stage, the 21-year average R^2 equalled 0.91 (yearly range: 0.86–0.93), 0.96 (yearly range: 0.95–0.97) and 0.96 (yearly range: 0.95–0.97) for the T_{min}, T_{mean} and T_{max} model in an average of 45,432, 48,925 and 42,155 cell-days, respectively. We additionally observed low values of RMSE for all the models. For the T_{min} model, the 21-year average RMSE equalled 2.02 °C (yearly range: 1.91 °C–2.13 °C), while for the T_{mean} and T_{max} equalled 1.41 °C (yearly range: 1.32 °C–1.54 °C) and 1.77 °C (yearly range: 1.67 °C–1.85 °C), respectively. The spatial and temporal R^2 remained quite high, while the corresponding errors stayed low. In the case of T_{mean}, overall spatial $R^2 = 0.88$ (yearly range: 0.84–0.93) and temporal $R^2 = 0.97$ (yearly range: 0.95–0.98), while $RMSE_{spatial} = 0.49$ °C (yearly range: 0.42 °C–0.59 °C) and $RMSE_{temporal} = 1.32$ °C (yearly range: 1.25 °C–1.45 °C).

For the third stage model (Table 1), the 21-year average $R^2 = 0.97$ (yearly range: 0.95–0.98), 0.98 (yearly range: 0.97–0.99) and 0.97 (yearly range: 0.95–0.98), while the RMSE = 1.25 °C (yearly range: 1.17 °C–1.38 °C), 1.03 °C (yearly range: 0.88 °C–1.12 °C) and 1.41 °C (yearly range: 1.22 °C–1.49 °C) for the T_{min}, T_{mean} and T_{max} model in a number of 81,166, 87,040 and 84,330 cell-days, respectively. Table S3 shows the comparison results between the third and second stage T_{air} predictions.

21-year average mean signed error was found to be 0.10 °C, 0.05 °C and −0.16 °C, for T_{min}, T_{mean} and T_{max} model, respectively. We report it in a yearly basis, together with the intercepts and slopes of the linear regressions between our predictions and the DWD observations in Table S4.

The percentage of T_{air} predictions that was provided by each stage of the modeling procedure can be found in Table S5 of the Supplementary material. On average, the first stage resulted in 0.04%, 0.04%, 0.03% of the final T_{min}, T_{mean} and T_{max} predictions, respectively, the second stage a 36.1%, 35.6% and 32.7%, while the third stage filled in the remaining approximately 62%, 62.5% and 65.4%. The missing values of our output T_{air} predictions' dataset over the period 2000–2020 were close to 1% for all the models.

3.2. Validation with external data

The small-scale validation in the Augsburg area showed that all models achieved high correspondence ($0.95 \leq R^2 \leq 0.99$) and low

Table 1
Prediction accuracy for the first and the third stage predictions: 10-fold CV results for daily T_{min}, T_{mean} and T_{max} in Germany, averaged for 2000–2020.

First stage predictions								
Measure	R^2	$R^2_{spatial}$	$R^2_{temporal}$	RMSE (°C)	SD (°C)	$RMSE_{spatial}$ (°C)	$RMSE_{temporal}$ (°C)	Sample size (cell-days number)
T_{min}	0.91	0.68	0.92	2.02	6.76	0.87	1.83	45,432
T_{mean}	0.96	0.88	0.97	1.41	7.56	0.49	1.32	48,925
T_{max}	0.96	0.84	0.97	1.77	9.12	0.77	1.60	42,155

Third stage predictions				
Measure	R^2	RMSE (°C)	SD (°C)	Sample size (cell-days number)
T_{min}	0.97	1.25	6.44	81,166
T_{mean}	0.98	1.03	7.22	87,040
T_{max}	0.97	1.41	8.02	84,330

*SD: standard deviation of the DWD observed T_{air}.

errors, even below 1 °C (Table 2). By season comparison led to similar findings of high models' performance. All T_{air} models, and especially the T_{min} model, achieved slightly better performance during the winter rather than the summer period (Table 2). Detailed results are provided in Tables S6 and S7. The linear regressions between the predicted T_{air} from our daily T_{min}, T_{mean} and T_{max} models and the HOBO-Logger observed daily T_{min}, T_{mean} and T_{max} during the comparison period, gave an intercept of 0.61, 0.28 and 0.41, and a slope of 1.04, 1.02 and 0.98, respectively (Table S8). Fig. 2 and S4 also indicates the strong correspondence of our model predictions and the HOBO-Logger network observations.

The large-scale comparison with the TRY dataset suggested a good correlation between the two models' outputs ($0.71 \leq R^2 \leq 0.99$) while most RMSE were below 1 °C (Table 3). For year by year results, please see Table S9-**S13** in the Supplementary material. In Fig. 3, we visualized the model correspondence for a randomly selected example year (2010). Most of the two models' predictions met in the slope of 1.

Our model predictions captured a wider T_{air} distribution and representation of spatial T_{air} variations in small scale analysis as we observe in the example Fig. S5 for the years 2003, 2006, 2008 and 2012 (randomly chosen) in the Augsburg area.

3.3. Case study - Augsburg

We first calculated the distances between the geocoded residential addresses of the KORA study participants and the two available DWD stations across the Augsburg area (Fig. S6). Most of the participants lived 5–15 km far from a station. Additionally, we found that T_{mean} varied substantially over space (Fig. 4). The city centre was way hotter that the surrounding rural areas (variation close to 2 °C) and in these rural regions there was also substantial variation even in neighbouring tiles. The actual difference in exposure assessment from the DWD observations and our model predictions could also be seen in the long-term assignment (Fig. 5). The DWD T_{mean} was only representative for the left-hand side distribution queue (cooler Augsburg areas). The higher temperatures are not captured by the DWD stations.

Fig. 6 is a short-term T_{mean} exposure distribution example, using an average of 7 days, which is often used for exposure assessment in epidemiological studies. Our model's predictions were close to the DWD observations at both stations. But, both stations values were below the distribution's mean and especially the T_{mean} of DWD station 1 was lower than the first quartile (Q1) of the distribution. This was mainly affected by their location, which was outside the city centre as seen in Fig. 4.

3.4. Descriptive statistics and spatiotemporal T_{air} patterns

Table 4 shows a selection of descriptive statistics (mean, standard deviation (SD), Q1, median and third quartile (Q3)) regarding the T_{min}, T_{mean} and T_{max} in Germany for the period 2000–2020 resulting from the DWD weather stations observations and our model predictions. The observed 21-year average T_{min}, T_{mean} and T_{max} from the DWD stations were 5.15 °C (SD = 6.59 °C), 9.44 °C (SD = 7.39 °C) and 13.85 °C (SD = 8.77 °C), respectively, while our models gave predicted 21-year average T_{min}, T_{mean} and T_{max} of 5.24 °C (SD = 5.89 °C), 9.57 °C (SD = 7.36 °C) and 14 °C (SD = 8.75 °C), respectively.

We also present the 21-year averaged predicted T_{min}, T_{mean} and T_{max} maps of Germany (Fig. 7, plot 1). The T_{air} spatial variability exceeded 15 °C annually on average, depending on the measure. We saw specific spatial patterns for T_{air}, including mountainous regions, rivers, lakes, forests and coastlines. For instance, the Alps and the Harz highland area were characterized by the lowest T_{air} values nationwide, while the dense urban cores (e.g., from Stuttgart to Frankfurt) or big individual cities as Munich and Berlin had much higher values of T_{air}, especially for T_{min} and T_{mean}, than the surrounding rural areas. We also observed the high contrasts our output provided even for small areas, due to its high resolution of 1 × 1 km. We additionally present in Fig. S7, the German-wide

Table 2
Accuracy results of the small-scale external validation with HOBO-Logger T_{min}, T_{mean} and T_{max} observations in the Augsburg area during 2013–2018, overall and by season.

Overall						
Measure	R^2	RMSE (°C)	SD (°C)	7-day average R^2	7-day average RMSE (°C)	SD (°C)
T_{min}	0.95	1.80	6.84	0.97	1.44	6.44
T_{mean}	0.99	1.07	7.72	0.99	0.90	7.37
T_{max}	0.98	1.37	9.11	0.98	1.08	8.50

By season												
Measure	Winter			Spring			Summer			Fall		
	R^2	RMSE (°C)	SD (°C)	R^2	RMSE (°C)	SD (°C)	R^2	RMSE (°C)	SD (°C)	R^2	RMSE (°C)	SD (°C)
T_{min}	0.83	1.53	4.10	0.89	1.78	4.74	0.74	2.05	3.13	0.87	1.72	4.51
T_{mean}	0.93	1.00	3.94	0.97	1.06	5.15	0.92	1.24	3.49	0.97	1.02	5.10
T_{max}	0.92	1.30	4.68	0.96	1.29	6.26	0.91	1.43	4.81	0.96	1.36	6.59

*SD: standard deviation of the dependent variable (HOBO-Logger T_{air}).

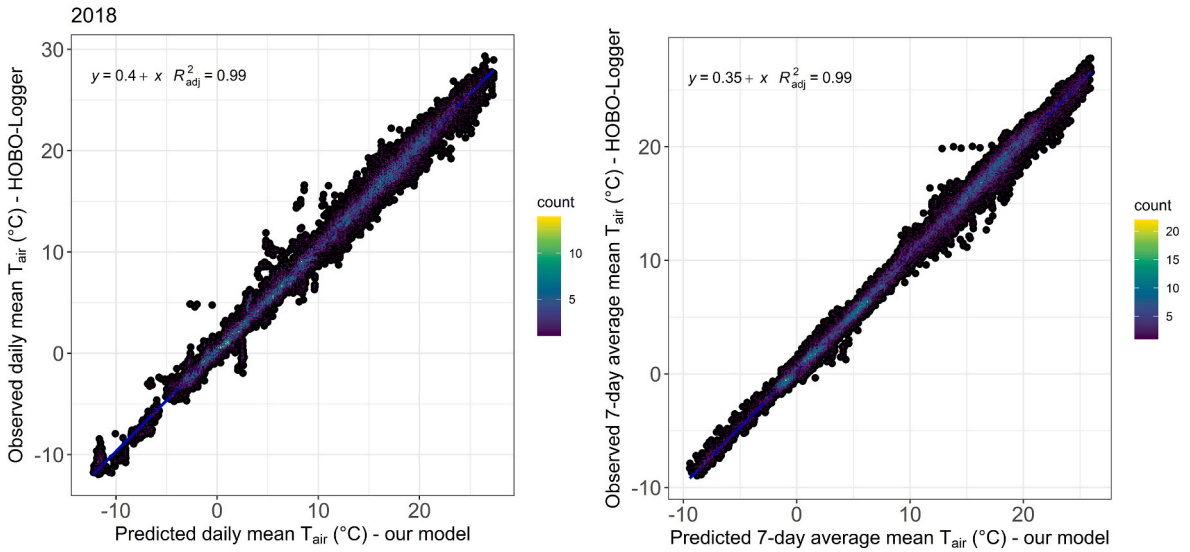

Fig. 2. Density scatterplots between the model daily T_{mean} predictions and the HOBO-Logger daily T_{mean} observations for 2018, daily average and 7-day average.

Table 3
Comparison between our model daily T_{mean} predictions and the DWD TRY model daily T_{mean} predictions in Germany during 2001–2012.

	R^2	RMSE (°C)
Overall	0.99	0.90
Season		
Winter	0.94	0.98
Spring	0.97	0.90
Summer	0.94	0.88
Fall	0.97	0.86
Without extremes		
5th pctl. < T_{mean} < 95th pctl.	0.99	0.79
To extremes		
T_{mean} < 5th pctl.	0.76	1.19
T_{mean} > 95th pctl.	0.71	0.93
District		
Augsburg Landkreis (rural)	0.99	0.84
Stadt Augsburg (urban)	0.99	0.85

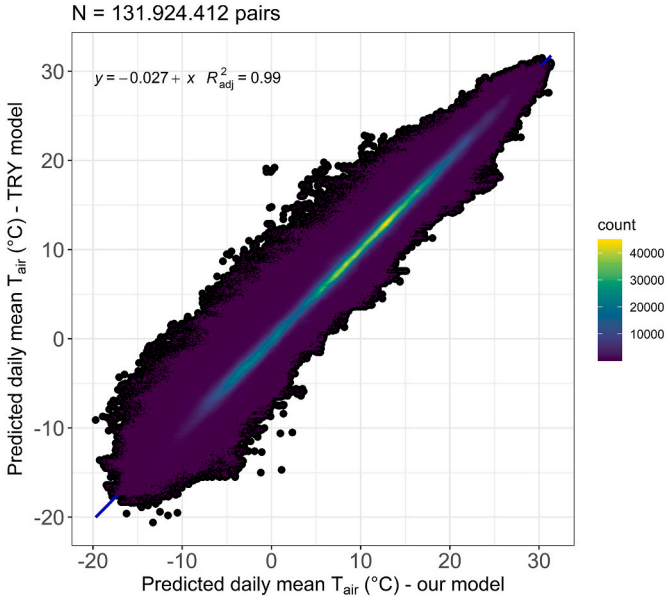

Fig. 3. Density scatterplot between our model daily T_{mean} predictions and the project TRY model daily T_{mean} predictions for 2010.

DTR maps for 2016–2019, i.e., the NAKO study baseline years. Eastern Germany and more intensely the North-Eastern part of the country experienced higher DTR variations, while cities or mountains and large water bodies were characterized by smaller DTR.

Regarding the temporal T_{air} variability in Germany (and in NAKO study centers) over the first two decades of the 21st century, we report the differences between the predicted T_{mean} yearly averages and the 20-year average (Fig. 7, plot 2 and Fig. S8). The year 2000 was excluded as the model predictions started from late February. There was an obvious tendency of increased averaged T_{mean} for the last 5–7 years (continuously beyond 2016). The hottest years recorded during the studied period were 2018 (T_{mean} = 10.45 °C) and 2020 (T_{mean} = 10.42 °C). Additionally, we mapped the number of heat (T_{max} > 30 °C) and cold (T_{min} < 0 °C) days by 3-digit zip code through the years. Fig. 8 presents an example comparing the years 2001 (as a reference) and 2015, where the observed difference was pronounced. For 2015 the number of heat days increased dramatically since 2001 mainly in eastern and south-eastern Germany, whereas the cold days dropped, even if 2015 was not among the hottest years of the study period (Fig. 7, plot 2).

4. Discussion

In this paper, we developed reliable high spatiotemporally-resolved T_{min}, T_{mean}, T_{max} and DTR datasets for Germany during 2000–2020, following a regression-based method which consists of three stages. We combined meteorological and remote sensing data as well as multiple land cover predictors. All models attained very good performance, and consequently their predictive ability appears to have a strong foundation, with overall high explained variance ($0.91 \leq R^2 \leq 0.98$) and low errors (1.03 °C ≤ RMSE ≤ 2.02 °C), calculated through CV. In addition, bias was found to be close to 0 (−0.16 °C ≤ mean signed error ≤ 0.10 °C). The external small ($0.74 \leq R^2 \leq 0.99$, 0.87 °C ≤ RMSE ≤ 2.05 °C) and large-scale validation ($0.71 \leq R^2 \leq 0.99$, 0.79 °C ≤ RMSE ≤ 1.19 °C) confirmed the high performance of the models. We additionally showed the benefits of our spatiotemporal T_{air} modeling in terms of exposure assessment for participants of epidemiological studies, conducting a case study in the Augsburg area.

For Germany, except the datasets of a coarser resolution of 5 km or more (Brinckmann and Bissolli, 2015; Frick et al., 2014), there is the 1 × 1 km hourly T_{air} dataset for 1995–2012, generated by Krähenmann et al. (2018), who applied a 3-step interpolation method (monthly RMSE ≈ 1 °C). We used this product to externally compare our model findings with another model across the country. There was a good overall

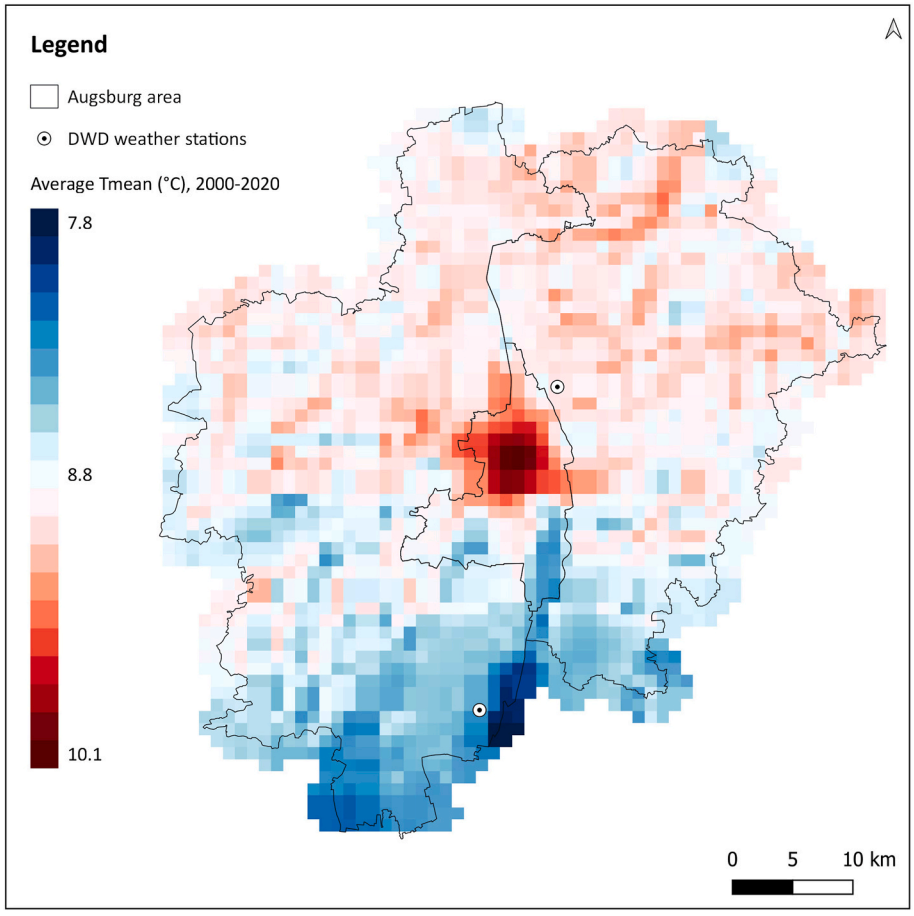

Fig. 4. Spatial pattern of the averaged predicted T_{mean} in the Augsburg area during 2000–2020.

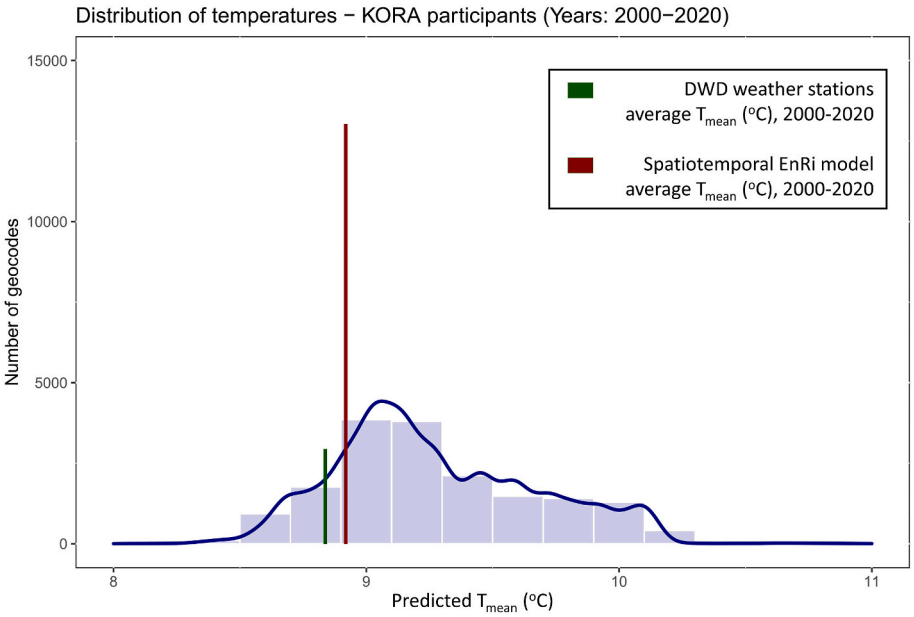

Fig. 5. Distribution of the predicted T_{mean}, assigned to KORA participants for 2000–2020 (in blue). The green and red lines show the exposure assignment based on the nearest monitoring station location. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

correspondence. However, our model predictions captured broader T_{air} variations, especially in city level (Fig. S5), as the interpolation methods are highly affected by the weather stations locations with limitations to fully represent between-station variability. Especially in Germany, a complex geo-climate study domain with high spatial heterogeneity, interpolation leads to neighbouring motives, thus closer regions are assigned rather similar values, and cannot capture sufficiently either the small-scale T_{air} variability or its extreme values. Therefore, we also

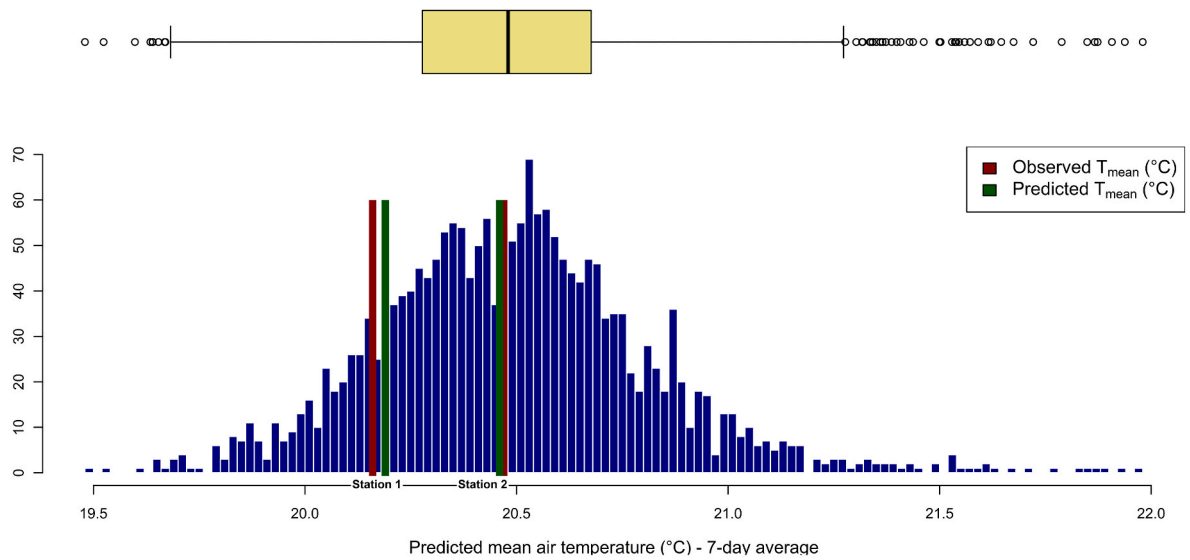

Fig. 6. Predicted 7-day average T_{mean} distribution (with blue) vs the 2 available DWD stations T_{mean} observations (with red) across the Augsburg area for 31.08.2019. The green lines represent our model's predictions at the stations locations. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 4
Observed and predicted T_{min}, T_{mean} and T_{max} in Germany during 2000–2020.

Source	Measure	Mean (°C)	SD (°C)	Q1 (°C)	Median (°C)	Q3 (°C)
DWD observations (n = 406 stations)	T_{min}	5.15	6.59	0.34	5.33	10.33
	T_{mean}	9.44	7.39	3.71	9.59	15.27
	T_{max}	13.85	8.77	6.81	13.91	20.68
Models predictions (n = 366,536 grid cells)	T_{min}	5.24	5.89	0.24	5.24	10.29
	T_{mean}	9.57	7.36	3.76	9.71	15.19
	T_{max}	14.00	8.75	6.90	14.05	20.67

observed worse correspondence of our and TRY models to extremes, i.e., 5th and 95th percentiles (Table 3). Krähenmann et al. (2018) also lack validation with completely independent T_{air} datasets while validating the interpolation-based predictions with input data could be biased due to their strong dependence. Finally, our output had a longer and more recent temporal extent and can be used in recent German cohorts.

Aiming to produce a helpful dataset for scientists working in the field of environmental epidemiology and especially those who investigate the T_{air} health effects or implement T_{air} in their analysis as a confounder or an effect modifier, the case study example we demonstrate for Augsburg is of great importance. The prevailing way for studies exploring the T_{air} health effects is to collect their exposure data from monitoring networks consisting of a limited number of ground-based weather stations, unevenly distributed across the country and insufficient to capture the spatial T_{air} variability, especially in city centers. Taking into consideration the Augsburg area, an epidemiological study would usually assign to all participants the T_{air} observations of the station which has the shortest distance from their residential address or a weighted average of the two available stations, again depending on distances or a simple interpolation technique. Hence, the participants would not be assigned with the exposure value of their actual location, but of the station's location even 10 km away (Fig. S6), implying the need of finer resolutions. People living in the city centre, where there is no available station, would be assigned with a way lower T_{air} exposure than their representative one, as we observed in Figs. 5 and 6. All the aforementioned issues lead to exposure error and consequently the health effects are biased towards the null (Zeger et al., 2000). On the other hand, our output captured the T_{air} variability and trends and reduced the exposure misclassification. Hence, we achieved a better representation of T_{air} variability and fulfilled one of our primary goals that was to provide more accurate T_{air} exposure assessment to German epidemiological studies.

A key finding of our analysis were the observed changes in T_{air}, which are mainly attributed to climate change that is already noticeable in Germany (Rüth et al., 2019). We showed that the four hottest years, based on an area-weighted averaging of the temperature during last two decades across the country, all occurred after 2014, while the last three consecutive years found to be the hottest overall (Fig. 7, plot 2). This finding was consistent and even more pronounced for the big constantly growing German cities (Fig. S7). Additionally, due to the high spatio-temporal resolution of the models, we detected climate change effects that cannot be captured by crude German-wide T_{air} averages. For instance, for 2015, we observed a substantial increase in hot days since 2001 even if this year's average T_{air} was lower than the 20-year average. The results of this analysis showing the impact of climate change on T_{air} locally and countrywide, are large, even over this short temporal period. With this tool, impacts on human health could be detected which then might contribute to climate change adaptation and risk reduction policies that German authorities need to enact in the following years.

4.1. Strengths

Best of our knowledge, this is the first study of T_{air} modeling which validates the models' prediction so extensively using external data. First locally, via a ground-based dense monitoring network for 6 years and then nationwide with another model based on a different approach for 12 years. The results indicated good performance and low errors in both cases, boosting our confidence in the quality of our product. An additional strength is the models' spatial resolution and spatial and temporal extent. They are German-wide and have a temporal extent of 21 years. Their national scale combined with the fine resolution of 1 × 1 km and the daily temporal resolution, provided us with the opportunity to study the spatiotemporal patterns of T_{air} all across Germany but also in specific places around the country, containing both urban and rural settings.

Our output product is an excellent fit for many individual-level epidemiological studies in Germany, without limitations on the study area(s). Mapping four different T_{air} measures was also important for environmental epidemiology as many recent studies report health effects of different temperature measures (Cheng et al., 2014; Guo et al., 2016; Oberheim et al., 2020; Wong et al., 2020), and there is a special

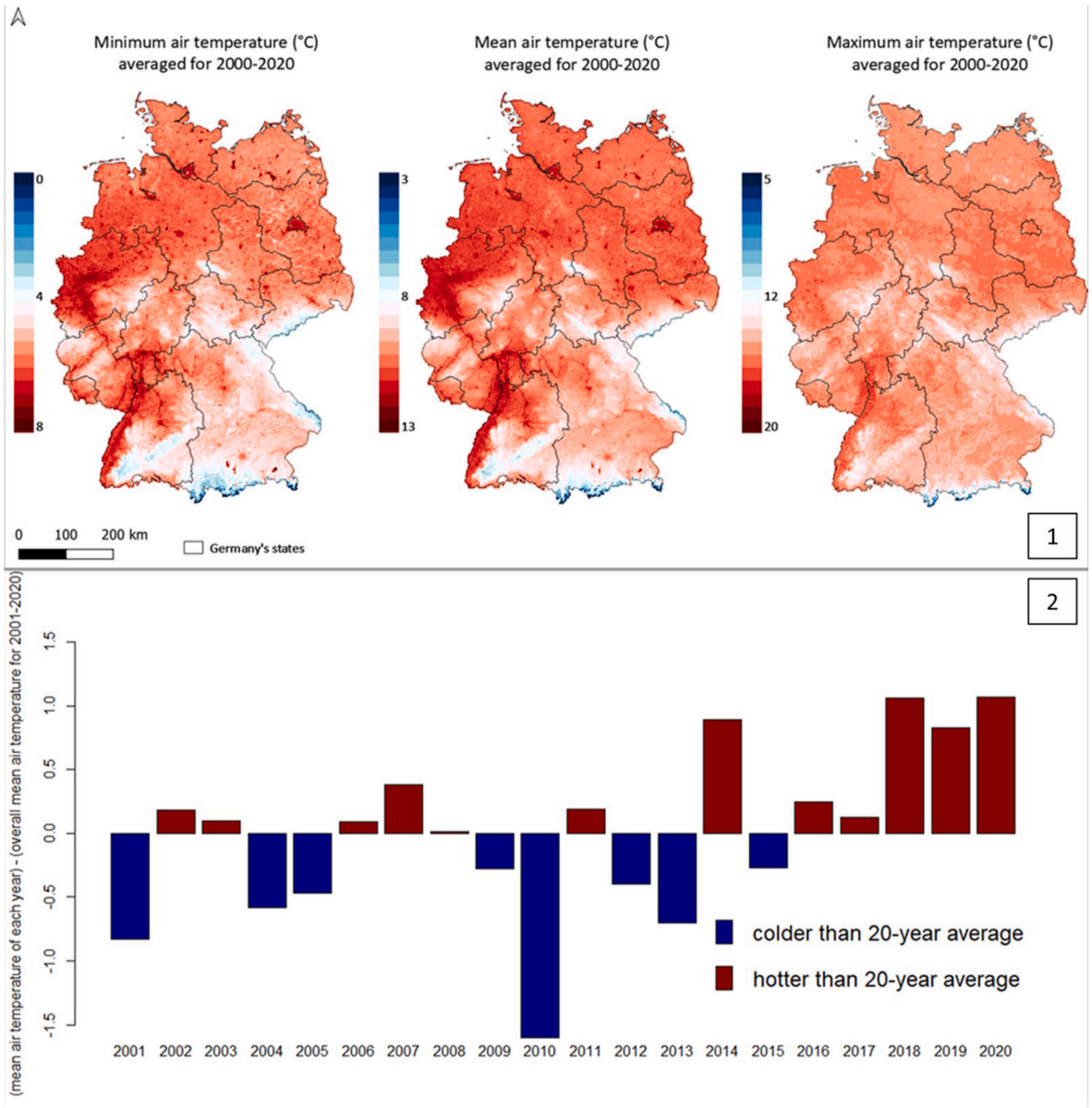

Fig. 7. Spatiotemporal T_{air} patterns in Germany for 2000–2020. Plot 1: Spatial patterns of the predicted T_{min}, T_{mean} and T_{max} in Germany, averaged for 2000-2020. Plot 2: Difference between the predicted T_{mean} yearly averages and the predicted T_{mean} 20-year average (2001-2020), German-wide.

need for further research with high spatiotemporally-resolved T_{min}, T_{mean}, T_{max} and DTR. Our final output can also be used for other research purposes outside the health field. For example, we are currently developing a high-resolution hybrid spatiotemporal RF model in order to predict daily mean relative humidity (RH) in Germany. To accomplish it, we use our daily T_{mean} predictions dataset as the main predictor in the RF model, due to its strong association (negative) with RH (Nikolaou et al., 2022).

4.2. Limitations

On the other hand, there are some limitations. First of all, the main predictor for estimating T_{air} is LST either at its daily or nightly value. It is well known that LST datasets include a high percentage of missing values because of cloud coverage, atmosphere dust, snow or sensor failure (Ghafarian Malamiri et al., 2018). However, using the TPS interpolated T_{air} data in the third stage model, we observed high accuracy for third stage predictions with quite low errors even when we compared them with independent observations from ground-based weather station networks. Moreover, we could not estimate the models performance, internally, in locations that have not been trained on. We tackled this issue by externally validating our T_{air} predictions with the HOBO-Logger T_{air} measurements for cells where the models were not trained on, and we observed high performance. Therefore, we are confident that even in the cell-days without available DWD T_{air} or LST data, our predictions are equally reliable. An additional factor that limits our product is its spatial resolution of 1 × 1 km, which is sufficient for country-wide analysis, but it might be a bit coarse for small-area/local analyses, especially for studies exploring the UHI effect. However, even in small scale analysis, our 1 × 1 km resolution dataset provides a better representation of T_{air} variability in comparison with the existing weather stations, as we showed in the case study of Augsburg's area. Higher spatiotemporal resolution, at least for the cities, might be a very good future upgrade in the framework of T_{air} modeling in Germany, given the example of previous studies in neighbouring countries (Hough et al., 2020). Finally, the 21-year extent we used to understand the spatiotemporal T_{air} patterns over Germany might be short to investigate climate change (usually a 30-year period). It is

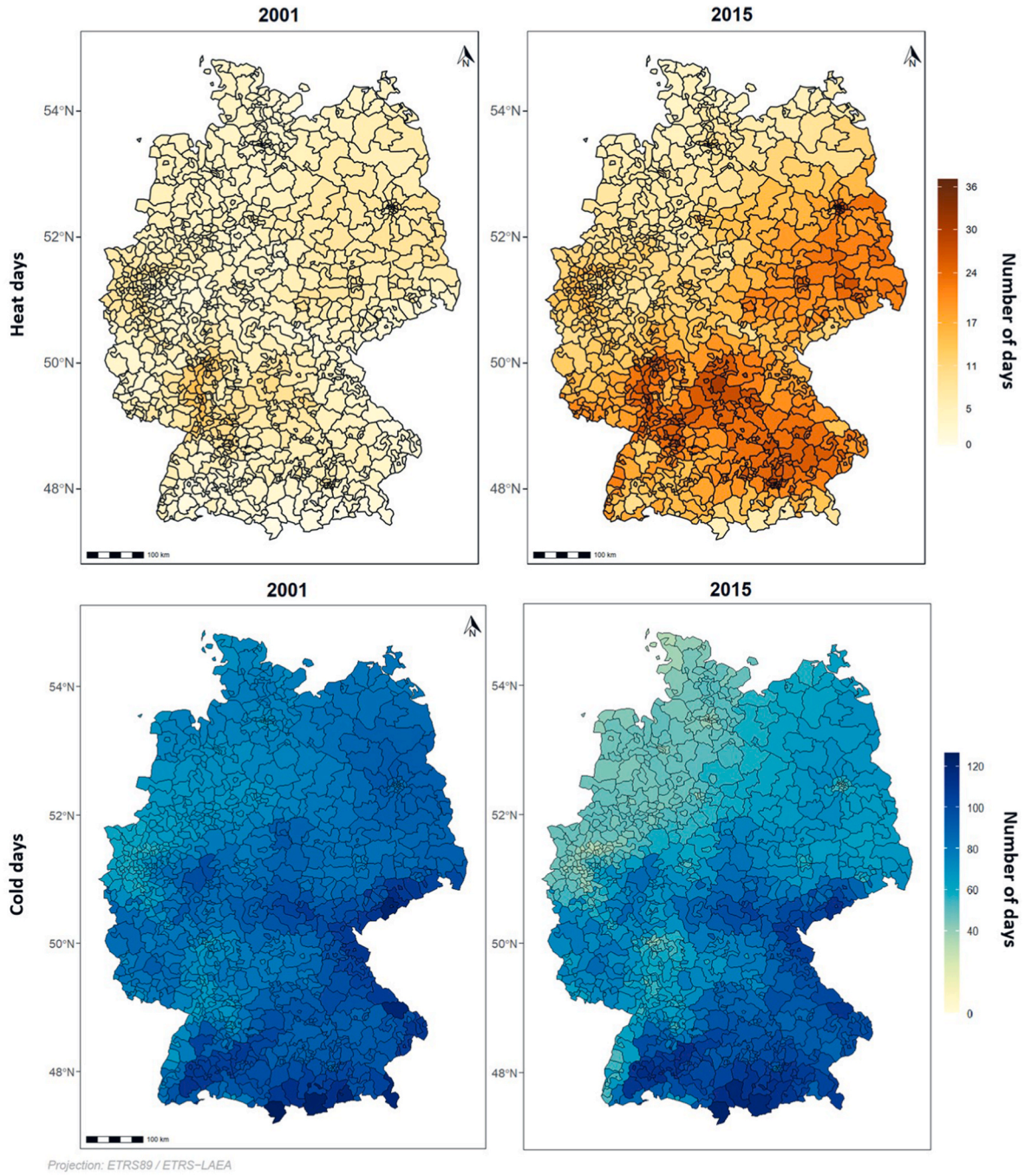

Fig. 8. Number of heat ($T_{max} > 30$ °C) and cold ($T_{min} < 0$ °C) days by 3-digit zip code in Germany, 2001) versus 2015.

nevertheless a good starting point to understand the T_{air} patterns and find useful climate change indications in Germany.

5. Conclusion

In this study, we applied a high-resolution hybrid spatiotemporal modeling approach to estimate daily T_{min}, T_{mean} and T_{max} as well as to calculate DTR across Germany over the period 2000–2020. We achieved excellent models' performances, validated extensively both locally and nationwide. Our product contributes substantially to exposure misclassification decrease accomplishing a better representation of T_{air} variability, and helps towards understanding the spatiotemporal T_{air} patterns and observing the impact of climate change during the last decades in Germany. Finally, our dataset is a great fit for recent German health cohorts and environmental epidemiology studies overall, but could also be used for other research purposes.

Author contributions

Nikolaos Nikolaou: Conceptualization, Data curation, Methodology, Formal analysis, Visualization, Writing – original draft, Writing – review & editing. **Marco Dallavalle**: Data curation, Methodology, Visualization, Writing – review & editing. **Massimo Stafoggia**: Methodology, Writing – review & editing. **Laurens M. Bouwer**: Data curation, Validation, Writing – review & editing. **Annette Peters**: Conceptualization, Writing – review & editing, Supervision. **Kai Chen**: Conceptualization, Data curation, Methodology, Formal analysis, Writing – review & editing. **Kathrin Wolf**: Conceptualization, Data

curation, Methodology, Writing – review & editing, Supervision. **Alexandra Schneider:** Conceptualization, Methodology, Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

This work was supported by the Helmholtz Climate Initiative (HI-CAM) project, which is funded by the Helmholtz Association's Initiative and Networking Fund. The authors are responsible for the content of this publication.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envres.2022.115062.

References

Armstrong, B.G., 1998. Effect of measurement error on epidemiological studies of environmental and occupational exposures. Occup. Environ. Med. 55 (10), 651–656. https://doi.org/10.1136/oem.55.10.651.

Bates, D., Mächler, M., Bolker, B., Walker, S., 2014. Fitting linear mixed-effects models using lme4. arXiv preprint arXiv: 1406.5823. https://doi.org/10.48550/arXiv.1406.5823.

Beck, C., Straub, A., Breitner, S., Cyrys, J., Philipp, A., Jacobeit, J., 2018a. Air temperature characteristics of local climate zones in the Augsburg urban area (Bavaria, southern Germany) under varying synoptic conditions. Urban Clim. 25, 152–166. https://doi.org/10.1016/j.uclim.2018.04.007.

Beck, H.E., Zimmermann, N.E., McVicar, T.R., Vergopolan, N., Berg, A., Wood, E.F., 2018b. Present and future Koppen-Geiger climate classification maps at 1-km resolution. Sci. Data 5 (1), 1–12. https://doi.org/10.1038/sdata.2018.214.

Benali, A., Carvalho, A., Nunes, J., Carvalhais, N., Santos, A., 2012. Estimating air surface temperature in Portugal using MODIS LST data. Remote Sens. Environ. 124, 108–121. https://doi.org/10.1016/j.rse.2012.04.024.

Breitner, S., Wolf, K., Devlin, R.B., Diaz-Sanchez, D., Peters, A., Schneider, A., 2014. Short-term effects of air temperature on mortality and effect modification by air pollution in three cities of Bavaria, Germany: a time-series analysis. Sci. Total Environ. 485, 49–61. https://doi.org/10.1016/j.scitotenv.2014.03.048.

Brinckmann, S., Bissolli, P., 2015. Gridded Daily Mean Near-Surface (2 M) Air Temperature for Europe (Project DecReg/MiKlip). Version V001, 2015. DWD Climate Data Center (CDC). https://doi.org/10.5676/DWD_CDC/DECREG0110v1.

Bundesamt, Statistisches, 2022. Current population - population by nationality and sex [Online]. Available from: https://www.destatis.de/EN/Themes/Society-Environment/Population/Current-Population/Tables/census-sex-and-citizenship-2021.html. (Accessed 19 January 2022). Accessed.

Busetto, L., Ranghetti, L., 2016. MODIStsp: an R package for automatic preprocessing of MODIS Land Products time series. Comput. Geosci. 97, 40–48. https://doi.org/10.1016/j.cageo.2016.08.020.

Chapman, S., Watson, J.E., Salazar, A., Thatcher, M., McAlpine, C.A., 2017. The impact of urbanization and climate change on urban temperatures: a systematic review. Landsc. Ecol. 32 (10), 1921–1935. https://doi.org/10.1007/s10980-017-0561-4.

Cheng, J., Xu, Z., Zhu, R., Wang, X., Jin, L., Song, J., Su, H., 2014. Impact of diurnal temperature range on human health: a systematic review. Int. J. Biometeorol. 58 (9), 2011–2024. https://doi.org/10.1007/s00484-014-0797-5.

Davis, R.E., Hondula, D.M., Sharif, H., 2020. Examining the diurnal temperature range enigma: why is human health related to the daily change in temperature? Int. J. Biometeorol. 64 (3), 397–407. https://doi.org/10.1007/s00484-019-01825-8.

Didan, K., 2015. MOD13A3 MODIS/terra vegetation indices monthly L3 global 1km SIN grid V006 [data set]. NASA EOSDIS Land Processes DAAC. https://doi.org/10.5067/MODIS/MOD13A3.006.

DWD, 2017. Wetterextreme in Deutschland und weltweit [Online]. Available from: https://www.dwd.de/DE/wetter/thema_des_tages/2017/2/16.html. (Accessed 19 January 2022). Accessed.

DWD, 2020. DWD-Stationen Duisburg-Baerl und Tönisvorst jetzt Spitzenreiter mit 41, 2 Grad Celsius [Online]. Available from: https://www.dwd.de/DE/presse/pressemitteilungen/DE/2020/20201217_annulierung_lingen_news.html. (Accessed 19 January 2022). Accessed.

DWD, 2021. DWD Climate Data Center (CDC): Historical Daily Station Observations (Temperature, Pressure, Precipitation, Sunshine Duration, etc.) for Germany version v21.3, 2021.

DWD, 2022. Zeitreihen und Trends [Online]. Available from: https://www.dwd.de/DE/leistungen/zeitreihen/zeitreihen.html. (Accessed 30 January 2022). Accessed.

Fluckiger, B., Kloog, I., Ragettli, M.S., Eeftens, M., Roosli, M., de Hoogh, K., 2022. Modelling daily air temperature at a fine spatial resolution dealing with challenging meteorological phenomena and topography in Switzerland. Int. J. Climatol. https://doi.org/10.1002/joc.7597.

Frick, C., Steiner, H., Mazurkiewicz, A., Riediger, U., Rauthe, M., Reich, T., Gratzki, A., 2014. Central European high-resolution gridded daily data sets (HYRAS): mean temperature and relative humidity. Meteorol. Z. 23 (1), 15–32. https://doi.org/10.1127/0941-2948/2014/0560.

Gasparrini, A., Guo, Y., Hashizume, M., Lavigne, E., Zanobetti, A., Schwartz, J., Forsberg, B., 2015. Mortality risk attributable to high and low ambient temperature: a multicountry observational study. Lancet 386 (9991), 369–375. https://doi.org/10.1016/S0140-6736(14)62114-0.

German National Cohort (GNC) Consortium, 2014. The German National Cohort: aims, study design and organization. Eur. J. Epidemiol. 29 (5), 371–382. https://doi.org/10.1007/s10654-014-9890-7.

Ghafarian Malamiri, H.R., Rousta, I., Olafsson, H., Zare, H., Zhang, H., 2018. Gap-filling of MODIS time series land surface temperature (LST) products using singular spectrum analysis (SSA). Atmosphere 9 (9), 334. https://doi.org/10.3390/atmos9090334.

Gil-Alana, L.A., 2018. Maximum and minimum temperatures in the United States: time trends and persistence. Atmos. Sci. Lett. 19 (4), e810 https://doi.org/10.1002/asl.810.

Global 30 Arc-Second Elevation (GTOPO30). DOI:/10.5066/F7DF6PQS.

Gronlund, C.J., Sullivan, K.P., Kefelegn, Y., Cameron, L., O'Neill, M.S., 2018. Climate change and temperature extremes: a review of heat-and cold-related morbidity and mortality concerns of municipalities. Maturitas 114, 54–59. https://doi.org/10.1016/j.maturitas.2018.06.002.

Guo, Y., Gasparrini, A., Armstrong, B.G., Tawatsupa, B., Tobias, A., Lavigne, E., Tong, S., 2016. Temperature variability and mortality: a multi-country study. Environ. Health Perspect. 124 (10), 1554–1559. https://doi.org/10.1289/EHP149.

Hertel, T.W., Rosch, S.D., 2010. Climate change, agriculture, and poverty. Appl. Econ. Perspect. Pol. 32 (3), 355–385. https://doi.org/10.1093/aepp/ppq016.

Holle, R., Happich, M., Löwel, H., Wichmann, H.E., null for the MONICA/KORA Study Group, 2005. KORA-a research platform for population based health research. Gesundheitswesen 67 (S 01), 19–25. https://doi.org/10.1055/s-2005-858235.

Hough, I., Just, A.C., Zhou, B., Dorman, M., Lepeule, J., Kloog, I., 2020. A multi-resolution air temperature model for France from MODIS and Landsat thermal data. Environ. Res. 183, 109244 https://doi.org/10.1016/j.envres.2020.109244.

IPCC, 2022. In: Pörtner, H.-O., Roberts, D.C., Tignor, M., Poloczanska, E.S., Mintenbeck, K., Alegría, A., Craig, M., Langsdorf, S., Löschke, S., Möller, V., Okem, A., Rama, B. (Eds.), Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press. Cambridge University Press, Cambridge, UK and New York, NY, USA, p. 3056. https://doi.org/10.1017/9781009325844.

Jin, Z., Ma, Y., Chu, L., Liu, Y., Dubrow, R., Chen, K., 2022. Predicting spatiotemporally-resolved mean air temperature over Sweden from satellite data using an ensemble model. Environ. Res. 204, 111960 https://doi.org/10.1016/j.envres.2021.111960.

Jobst, A.M., Kingston, D.G., Cullen, N.J., Sirguey, P., 2017. Combining thin-plate spline interpolation with a lapse rate model to produce daily air temperature estimates in a data-sparse alpine catchment. Int. J. Climatol. 37 (1), 214–229. https://doi.org/10.1002/joc.4699.

Kilibarda, M., Hengl, T., Heuvelink, G.B., Gräler, B., Pebesma, E., Perčec Tadić, M., Bajat, B., 2014. Spatio-temporal interpolation of daily temperatures for global land areas at 1 km resolution. J. Geophys. Res. Atmos. 119 (5), 2294–2313. https://doi.org/10.1002/2013JD020803.

Kloog, I., Nordio, F., Coull, B.A., Schwartz, J., 2014. Predicting spatiotemporal mean air temperature using MODIS satellite surface temperature measurements across the Northeastern USA. Remote Sens. Environ. 150, 132–139. https://doi.org/10.1016/j.rse.2014.04.024.

Kloog, I., Nordio, F., Lepeule, J., Padoan, A., Lee, M., Auffray, A., Schwartz, J., 2017. Modelling spatio-temporally resolved air temperature across the complex geo-climate area of France using satellite-derived land surface temperature data. Int. J. Climatol. 37 (1), 296–304. https://doi.org/10.1002/joc.4705.

Kovats, R.S., Kristie, L.E., 2006. Heatwaves and public health in Europe. Eur. J. Publ. Health 16 (6), 592–599. https://doi.org/10.1093/eurpub/ckl049.

Krähenmann, S., Walter, A., Brienen, S., Imbery, F., Matzarakis, A., 2016. Daily means of hourly grids of air temperature for Germany (project TRY Advancement), Version V001, DWD Climate Data Center (CDC). https://doi.org/10.5676/DWD_CDC/TRY_Basis_v001, 2016.

Krähenmann, S., Walter, A., Brienen, S., Imbery, F., Matzarakis, A., 2018. High-resolution grids of hourly meteorological variables for Germany. Theor. Appl. Climatol. 131 (3), 899–926. https://doi.org/10.1007/s00704-016-2003-7.

Lacetera, N., 2019. Impact of climate change on animal health and welfare. Animal Frontiers 9 (1), 26–31. https://doi.org/10.1093/af/vfy030.

Li, S., Griffith, D.A., Shu, H., 2020. Temperature prediction based on a space–time regression-kriging model. J. Appl. Stat. 47 (7), 1168–1190. https://doi.org/10.1080/02664763.2019.1671962.

Lindsey, R., Dahlman, L., 2021. Climate Change: Global Temperature [Online]. Available from: https://www.climate.gov/news-features/understanding-climate/climate-change-global-temperature. (Accessed 29 January 2022). Accessed.

Meehl, G.A., Tebaldi, C., 2004. More intense, more frequent, and longer lasting heat waves in the 21st century. Science 305 (5686), 994–997. https://doi.org/10.1126/science.1098704.

Modala, N.R., Ale, S., Goldberg, D.W., Olivares, M., Munster, C.L., Rajan, N., Feagin, R. A., 2017. Climate change projections for the Texas high plains and rolling plains. Theor. Appl. Climatol. 129 (1), 263–280. https://doi.org/10.1007/s00704-016-1773-2.

NASA, 2021. LP DAAC [Online]. Available: https://lpdaac.usgs.gov/. (Accessed 17 October 2022). Accessed.

Nikolaou, N., Bouwer, L., Valizadeh, M., Dallavalle, M., Wolf, K., Stafoggia, M., Peters, A., Schneider, A., 2022. High-resolution Hybrid Spatiotemporal Modeling of Daily Relative Humidity across Germany for Epidemiological Research: a Random Forest Approach. EGU General Assembly 2022, Vienna, Austria. https://doi.org/10.5194/egusphere-egu22-6543, 23–27 May 2022, EGU22-6543.

Nychka, D., Furrer, R., Paige, J., Sain, S., 2017. fields: tools for spatial data. R package version 9 (10.5065), D6W957CT. https://doi.org/10.5065/D6W957CT.

Oberheim, J., Höser, C., Lüchters, G., Kistemann, T., 2020. Small-scaled association between ambient temperature and campylobacteriosis incidence in Germany. Sci. Rep. 10 (1), 1–12. https://doi.org/10.1038/s41598-020-73865-9.

Peters, A., Schneider, A., 2021. Cardiovascular risks of climate change. Nat. Rev. Cardiol. 18 (1), 1–2. https://doi.org/10.1038/s41569-020-00473-5.

QGIS Development Team, 2020. QGIS geographic information system. Open source geospatial foundation project. Available from: http://qgis.osgeo.org.

R Core Team, 2020. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. Available from: https://www.R-project.org/.

Rosenfeld, A., Dorman, M., Schwartz, J., Novack, V., Just, A.C., Kloog, I., 2017. Estimating daily minimum, maximum, and mean near surface air temperature using hybrid satellite models across Israel. Environ. Res. 159, 297–312. https://doi.org/10.1016/j.envres.2017.08.017.

Rüth, P.V., Schönthaler, K., Andrian-Werburg, S.V., Buth, M., 2019. Monitoringbericht 2019 zur Deutschen Anpassungsstrategie an den Klimawandel Bericht der Interministeriellen Arbeitsgruppe Anpassungsstrategie der Bundesregierung. Umweltbundesamt (UBA), Dessau-Roßlau. Available from: https://www.umweltbundesamt.de/publikationen/umweltbundesamt-2019-monitoringbericht-2019-zur.

Sekulić, A., Kilibarda, M., Protić, D., Tadić, M.P., Bajat, B., 2020. Spatio-temporal regression kriging model of mean daily temperature for Croatia. Theor. Appl. Climatol. 140 (1), 101–114. https://doi.org/10.1007/s00704-019-03077-3.

Shi, L., Liu, P., Kloog, I., Lee, M., Kosheleva, A., Schwartz, J., 2016. Estimating daily air temperature across the Southeastern United States using high-resolution satellite data: a statistical modeling study. Environ. Res. 146, 51–58. https://doi.org/10.1016/j.envres.2015.12.006.

Shukla, P., Skea, J., Calvo Buendia, E., Masson-Delmotte, V., Pörtner, H., et al., 2019. IPCC, 2019: Climate Change and Land: an IPCC Special Report on Climate Change, Desertification, Land Degradation, Sustainable Land Management, Food Security, and Greenhouse Gas Fluxes in Terrestrial Ecosystems.

Vancutsem, C., Ceccato, P., Dinku, T., Connor, S.J., 2010. Evaluation of MODIS land surface temperature data to estimate air temperature in different ecosystems over Africa. Remote Sens. Environ. 114 (2), 449–465. https://doi.org/10.1016/j.rse.2009.10.002.

Vicedo-Cabrera, A.M., Scovronick, N., Sera, F., Royé, D., Schneider, R., Tobias, A., et al., 2021. The burden of heat-related mortality attributable to recent human-induced climate change. Nat. Clim. Change 11 (6), 492–500. https://doi.org/10.1038/s41558-021-01058-x.

Vicente-Serrano, S.M., Saz-Sánchez, M.A., Cuadrat, J.M., 2003. Comparative analysis of interpolation methods in the middle Ebro Valley (Spain): application to annual precipitation and temperature. Clim. Past 24 (2), 161–180. https://doi.org/10.3354/cr024161.

Wan, Z., 2014. New refinements and validation of the collection-6 MODIS land-surface temperature/emissivity product. Remote Sens. Environ. 140, 36–45. https://doi.org/10.1016/j.rse.2013.08.027.

Wan, Z., Hook, S., Hulley, G., 2015. MOD11A1 MODIS/terra land surface temperature/emissivity daily L3 global 1km SIN grid V006 [data set]. NASA EOSDIS Land Processes DAAC. https://doi.org/10.5067/MODIS/MOD11A1.006.

Watts, N., Amann, M., Arnell, N., Ayeb-Karlsson, S., Belesova, K., Boykoff, M., et al., 2019. The 2019 report of the Lancet Countdown on health and climate change: ensuring that the health of a child born today is not defined by a changing climate. Lancet 394 (10211), 1836–1878. https://doi.org/10.1016/S0140-6736(19)32596-6.

Wickham, H., 2009. Ggplot2: Elegant Graphics for Data Analysis, second ed. Springer, New York, p. 35. https://doi.org/10.1007/978-0-387-98141-3. 10.1007.

Wong, S., Cantoral, A., Téllez-Rojo, M.M., Pantic, I., Oken, E., Svensson, K., et al., 2020. Associations between daily ambient temperature and sedentary time among children 4–6 years old in Mexico City. PLoS One 15 (10), e0241446. https://doi.org/10.1371/journal.pone.0241446.

Wu, W., Xu, A.-D., Liu, H.B., 2015. High-resolution spatial databases of monthly climate variables (1961–2010) over a complex terrain region in southwestern China 119 (1), 353–362. https://doi.org/10.1007/s00704-014-1123-1.

Xu, Y., Knudby, A., Ho, H.C., 2014. Estimating daily maximum air temperature from MODIS in British Columbia, Canada. Int. J. Rem. Sens. 35 (24), 8108–8121. https://doi.org/10.1080/01431161.2014.978957.

Ye, X., Wolff, R., Yu, W., Vaneckova, P., Pan, X., Tong, S., 2012. Ambient temperature and morbidity: a review of epidemiological evidence. Environ. Health Perspect. 120 (1), 19–28. https://doi.org/10.1289/ehp.1003198.

Zafeiratou, S., Samoli, E., Dimakopoulou, K., Rodopoulou, S., Analitis, A., Gasparrini, A., et al., 2021. A systematic review on the association between total and cardiopulmonary mortality/morbidity or cardiovascular risk factors with long-term exposure to increased or decreased ambient temperature. Sci. Total Environ. 772, 145383 https://doi.org/10.1016/j.scitotenv.2021.145383.

Zanobetti, A., Schwartz, J., 2008. Temperature and mortality in nine US cities. Epidemiology 19 (4), 563. https://doi.org/10.1097/EDE.0b013e31816d652d.

Zeger, S.L., Thomas, D., Dominici, F., Samet, J.M., Schwartz, J., Dockery, D., Cohen, A., 2000. Exposure measurement error in time-series studies of air pollution: concepts and consequences. Environ. Health Perspect. 108 (5), 419–426. https://doi.org/10.1289/ehp.00108419.

Zhu, W., Lű, A., Jia, S., 2013. Estimation of daily maximum and minimum air temperature using MODIS land surface temperature products. Remote Sens. Environ. 130, 62–73. https://doi.org/10.1016/j.rse.2012.10.034.