Received: 14 July 2021 | Revised: 21 February 2022 | Accepted: 27 February 2022 | Published on: 15 March 2022

DOI: 10.1002/joc.7597

International Journal of Climatology RMetS

RESEARCH ARTICLE

Modelling daily air temperature at a fine spatial resolution dealing with challenging meteorological phenomena and topography in Switzerland

Benjamin Flückiger[1,2] | **Itai Kloog**[3] | **Martina S. Ragettli**[1,2] | **Marloes Eeftens**[1,2] | **Martin Röösli**[1,2] | **Kees de Hoogh**[1,2]

[1]Department of Epidemiology and Public Health, Swiss Tropical and Public Health Institute, Basel, Switzerland

[2]University of Basel, Basel, Switzerland

[3]Department of Geography and Environmental Development, Ben-Gurion University of the Negev, Beer-Sheva, Israel

Correspondence
Kees de Hoogh, Department of Epidemiology and Public Health, Swiss Tropical and Public Health Institute, Basel, Switzerland.
Email: c.dehoogh@swisstph.ch

Funding information
Federal Office of Public Health, Grant/Award Number: 3727; Swiss Federal Office for the Environment

Abstract

Epidemiological studies investigating the relationship between air temperature or heat and health, still, by and large, rely on either information from the nearest weather station or on coarse gridded temperature predictions, thereby ignoring small-scale intra-urban variations. Recent methodological advances show promise in achieving high spatiotemporal temperature predictions, thus improving the characterization of spatial variations in temperature and decreasing bias in health studies. Here, we applied a two-stage approach using random forest to (a) impute missing moderate resolution imaging spectroradiometer (MODIS) land surface temperature at a 1 × 1 km resolution and (b) to use the gap-filled MODIS data to explain spatiotemporal variation in the measured ground-based air temperature data at a 100 × 100 m resolution across Switzerland using a range of predictor variables, including meteorological parameters, normalized difference vegetation index, impervious surface and altitude. Models presented here managed to capture temporal and spatial variations in air temperature in Switzerland from 2003 to 2018 at a fine spatial resolution of 100 × 100 m. Stage 1 models achieved an overall R^2 of 0.98 and a root mean squared error (RMSE) of 1.49°C (independent validation), and the stage 2 model performed well for all years with R^2 and RMSE ranging from 0.94 to 0.99 and 1.05 to 1.86°C, respectively. We were also able to capture the urban heat island effect and some typical weather phenomena caused by Switzerland's complex topography, like the foehn effect and inversion conditions. The resulting daily temperature surfaces for 2003–2018 will facilitate ongoing epidemiological research investigating the health effects of heat.

KEYWORDS

air temperature, MODIS, random forest, remote sensing, urban heat island

1 | INTRODUCTION

Heat stress has been identified as an urgent health threat at the global scale (Lee *et al*., 2019; Vicedo-Cabrera *et al*., 2021). A body of evidence links heat stress to increased morbidity and mortality especially in the vulnerable population including the elderly, people with chronic diseases, small children, the poor and outdoor workers (Åström *et al*., 2011; Ye *et al*., 2012; Benmarhnia *et al*., 2015; Song *et al*., 2017; Levi *et al*., 2018). Populations in cities are potentially more affected by high ambient temperatures, especially during the night, than those in the surroundings, because built environments absorb and store heat, resulting in a so-called urban heat island (UHI) effect. UHIs can be differentiated into four types, depending on scale and location: subsurface, surface, canopy layer and boundary layer UHIs (Christen *et al*., 2017). In the context of our study, we are looking at the canopy UHI, which is the most influential one for air temperature near the ground. In our study, we define the canopy UHI as the mean daily temperature difference between city and countryside.

The observed intensity of this phenomenon is usually in a range of 2–4°C of increased air temperature, in extreme cases (during the night in large cities) the intensity can reach 5–10°C (Heaviside *et al*., 2017). The number of vulnerable people exposed to extreme temperatures and heatwaves is increasing in every region of the world (Watts *et al*., 2021). The reasons for this increase are climate change and the associated rise in temperature and frequency of heatwaves as well as ageing population and progressing urbanization (Coffel *et al*., 2017; Wouters *et al*., 2017; Ilango *et al*., 2020).

Epidemiological studies investigating the relationship between temperature and health outcomes based on historic data have mostly relied on crude measures of exposure. Typically, exposures of populations in cities and regions were based on measurements of ambient temperature at a single meteorological station (Gasparrini *et al*., 2015), thereby ignoring intra-urban variation at a finer spatial scale (Kloog, 2019). Two studies in Switzerland, for example, investigated heat-related mortality in eight cities and only used one representative monitoring station for each city to estimate exposure (Ragettli *et al*., 2017; Lee *et al*., 2021). While this approach may be suitable for health assessments based on aggregated health data on city or area level (de Schrijver *et al*., 2021), exposure misclassification may occur when individual-level health records with information on home addresses are available. In addition, for studies investigating temperature-related health effects of a whole country including both rural and urban areas, station-based weather data of sufficient quality may not be available to capture the spatiotemporal exposure variability. This especially applies for Switzerland with large topographic heterogeneity and associated specific meteorological conditions. For example, micro-climates in valleys or plateaus result in unique inversion conditions, and the occurrence of foehn winds (warm, dry winds on the lee side of mountains), are especially challenging. Besides topography also UHI effects can lead to a high spatial variability within cities. For Switzerland, this effect was estimated to increase the maximum night-time temperature during extreme heatwaves up to 6–7°C compared to rural sites (Gehrig and Scherrer, 2018). Not accounting for small scale variability in exposure due to topography and the built environment may lead to a bias in health effect estimates (Zeger *et al*., 2000). Other factors influencing health effect estimates as a result of heat include socioeconomic position, building characteristics and the availability of air conditioning (Heaviside *et al*., 2017).

Recent developments in statistical modelling techniques aimed at improving spatial and temporal temperature patterns have taken advantage of satellite based land surface temperature (LST) data (Kloog, 2019). Kloog *et al*. (2014) pioneered this approach in Northeastern United States, using mixed regression models to predict daily near-surface air temperature combining satellite (moderate resolution imaging spectroradiometer [MODIS]), ground-based weather station data and a range of spatiotemporal predictors achieving a high precision. Shi *et al*. (2016) build on this method to estimate daily air temperature data for Southeastern United States using the same high-resolution satellite data demonstrating reliable prediction models across an area with very different climatic conditions compared to Northeastern United States. In France, the same approach was successfully applied to predict daily air temperature at a 1×1 km resolution (Kloog *et al*., 2017). Hough *et al*. (2020) further developed this method in France using data from the Aqua, Terra and Landsat satellites. A combination of mixed effect, random forest and gradient boosting models were used to downscale 1×1 km MODIS LSTs (daily T_{mean}, T_{max} and T_{min}) to a finer spatial resolution of 200×200 m and to convert to ambient temperature using a number of predictor variables including Landsat normalized difference vegetation index (NDVI) and elevation. A similar methodology was used in Israel to estimate daily T_{mean}, T_{max} and T_{min} near surface temperature at a slightly coarser spatial scale of 1×1 km (Rosenfeld *et al*., 2017). Although not taking advantage of LST data, Dirksen *et al*. (2020) combined linear regression and multiple adaptive regression splines to downscale daily air temperature to a 1×1 km resolution in the Netherlands using a range of meteorological and land use predictor variables.

To facilitate future studies in Switzerland investigating the relationship between temperature and health we developed models characterizing small scale intra-urban variation in temperature. Here we present the development of a spatiotemporal framework, applying random forest in estimating daily minimum, maximum and mean temperature across Switzerland from 2003 to 2018 at a fine spatial scale of 100 × 100 m, downscaling from existing 2 × 2 km temperature surfaces (Frei, 2014). This framework consists of 2 stages; (a) imputation of missing data in satellite-based daily surface temperature (T_s) at 1 × 1 km; and (b) converting T_s to surface air temperature (T_a) at 100 × 100 m using a combination of weather re-analysis data, land use variables and ground monitoring stations, thereby improving on previous methods reducing from three to two stages. With these models, we aim to capture both the UHI effect within Swiss cities and temperature differences due to topography-related weather phenomena, like the foehn effect and inversion conditions. Our motivation for the spatiotemporal resolution of our model output is driven by the resolution of the health data in future epidemiological studies. The health data are generally available at a daily resolution and we generally do not know at what time of the day the event occurred, justifying the choice for T_{mean}, T_{min} and T_{max}. We typically know the geocoded residential address location of the participants in the health studies and therefore can leverage this information by linking it to the fine scale 100 × 100 m temperature surfaces. The method presented here can be used as a blueprint for national or continental models.

2 | DATA AND METHODS

2.1 | Study area

The study area consists of the entire territory of Switzerland with an area of 41,285 km^2. Despite the relatively small size, Switzerland contains a variety of topographic features. The main features are the Jura Mountains in the northwest (typical altitude 500–1,600 m above sea level), the central plateau in the middle (300–800 m) and the Alps in the south and east (300–2000 m for inhabited areas). The climatic conditions reflect the topographic variation with arctic temperatures in the Alps (e.g. the average of daily mean temperatures in Davos from 1998 to 2018 was 4.1°C, with a minimum daily mean of −22.2°C and a maximum daily mean of 20.8°C), mild conditions in the central plateau (e.g. the average of daily mean temperatures in Kloten from 1998 to 2018 was 10.0°C, with a minimum daily mean of −13.0°C and a maximum daily mean of 27.7°C) and hot and humid summers in the Mediterranean influenced southern side of the Alps (e.g. the average of daily mean temperatures in Lugano from 1998 to 2018 was 13.1°C, with a minimum daily mean of −5.3°C and a maximum daily mean of 28.5°C). In addition, inner alpine valleys have special climatic conditions, as they are shielded from precipitation from north and south. Downslope winds on the lee side of the mountains (foehn) can strongly influence the temperature and humidity in the valleys and occasionally stretch far into the central plateau due to adiabatic warming.

Next we describe the data used in our two-stage modelling framework. In stage 1, elevation, azimuth, NDVI and day length were used in a random forest model to explain spatiotemporal variation in surface temperature with the aim to impute missing surface temperature data. In stage 2 the imputed surface temperature was combined with meteorological variables, elevation, NDVI, impervious surface, land cover, radiation and sky-view in a random forest model to predict the measured air temperature.

2.2 | Air temperature data

We obtained measured daily near surface air temperature (T_a) data from the Federal Office of Meteorology and Climatology (MeteoSwiss) for the years 2003–2018. The network consists of measurement stations operated by MeteoSwiss and partner institutions. The stations cover all regions of Switzerland and different altitudes and the coverage of the network, which was expanded substantially between 2003 and 2018 from 294 to 534 stations (Figure 1). We checked the data on consistency and removed values which were implausible, affected by logging errors, impossible, or affected by anomalies, specifically:

- We considered values outside of the known range to be implausible: temperatures lower than −50°C, higher than 50°C, %Relative Humidity (RH) below 0 or above 120.
- We removed sequences of measurements where during 3 days or more in a row (up to 23 days), the same exact temperature, RH, T_{max} and T_{min} were recorded at a single site, likely due to logging errors.
- Removed cases where T_{mean} was higher than T_{max}, T_{mean} was lower than T_{min}, T_{max} was lower than T_{min}, or T_{min} was higher than T_{max}.
- We removed values which were more than 20°C lower or higher than the 'expected temperature' at a particular site i for a particular day j considering the station ID$_i$ and the Swiss median temperature$_j$ (median

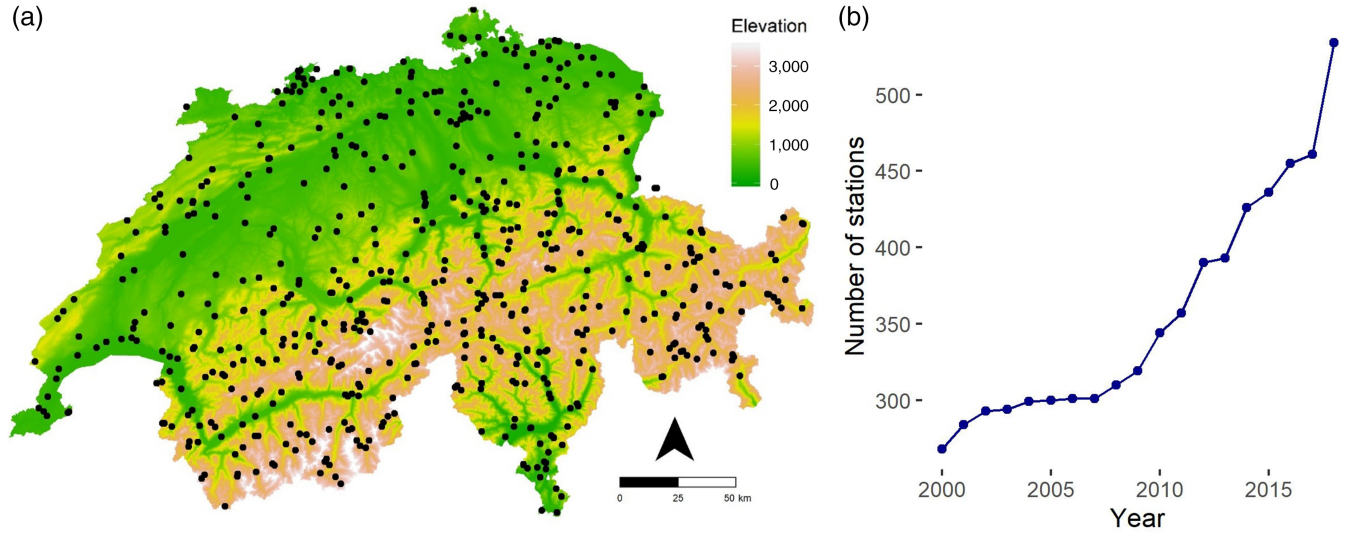

FIGURE 1 Location of all 534 weather stations (dots) from the MeteoSwiss network on a digital elevation map of Switzerland (a) and the number of available meteorological stations for each year from 2000 to 2018 in Switzerland (b) [Colour figure can be viewed at wileyonlinelibrary.com]

temperature across all Swiss sites on that day) using a mixed model with a random intercept for Station ID. This flagged values which were untypical for the location (e.g. a mountain station with typical range −20 to 5°C reporting 32°C) or untypical for the time of year (e.g. if all Swiss stations are measuring temperatures 10–25°C and one station is measuring −15°C.

- During the cleaning process, we only removed a very small percentage of data, for example, for T_{mean} we removed 0.094% of the data (2,673 values from a total of 2,846,109 values).

Table S1 shows the distribution (mean, minimum, maximum, first and 99th percentile) of T_{mean}, T_{min} and T_{max} by calendar year and the mean altitude of the meteorological stations for each year.

2.3 | Surface temperature data

The MODIS version 6 from the satellites Aqua (MYD11A1) and Terra (MOD11A1) was used to provide daily surface temperature data (T_s). MODIS T_s is derived from the thermal infrared channels 31 (10.78–11.28 μm) and 32 (11.77–12.27 μm). The MODIS T_s version 6 data are corrected for atmospheric and emissivity effects and have made improvements in removing cloud contamination. The spatial resolution is approximately 1 × 1 km, each satellite orbiting Switzerland twice per day. 2003 is the first full year with data from both satellites. This limits our model to the years from 2003 onwards. Overpass times are 10:30 a.m. and 10:30 p.m. for Terra and 1:30 p.m. and 1:30 a.m. for Aqua (local solar time). This pattern of overpass times allows us to capture diurnal differences, which can be helpful to detect UHI effects. Data from the corresponding MODIS tile h18v04 was obtained for 2003–2018 from EarthData (https://ladsweb.modaps.eosdis.nasa.gov/, accessed 19 November 2019). As proposed by Hough et al. we used the quality assessment band to exclude pixels with an LST error >2 K and pixels over snow and water (NDVI < 0) (Hough et al., 2020). To assess the quality of MODIS data, we compared measured temperature at Payerne weather station (daily mean) with the four MODIS variables for the year 2017. We found high correlations (∼0.9) of MODIS surface temperature data with ground based measurements.

2.4 | Normalized difference vegetation index

Vegetation or greenness affects temperature specifically through shade, evapotranspiration and the effect of reduced natural ventilation. Greenness can also imply less built up areas and therefore a higher skyview factor that helps the emission of longwave radiation. We did not distinguish between high and low vegetation. To account for the spatial and temporal variation of vegetation the NDVI was obtained as a proxy for greenness. The NDVI is measured daily by the MODIS instrument on board of the Terra satellite at a 1 × 1 km resolution. Monthly NDVI values, used in stage 1, were obtained from the MOD13A3 product (https://lpdaac.usgs.gov/products/mod13a3v006/, accessed 19 November 2019) for the years 2003–2018.

2.5 | Elevation plus elevation-derived variables

Two products from the digital elevation model (DHM25) from Swisstopo were obtained. The 200 × 200 m DHM25 product was aggregated to 1 × 1 km to match the MODIS resolution for the first stage. For the second stage, the 25 × 25 m DHM25 product was aggregated to match the 100 × 100 m resolution of the final temperature rasters. Air pressure is usually lower at higher altitudes; therefore, elevation can be a good predictor to model temperature. Slope and aspect were calculated for stage 2 at a 100 × 100 m resolution using the SLOPE and ASPECT tools available in ArcGIS 10.6. Slope identifies the steepest slope (in degrees) between the cell and its neighbouring cells. Aspect depicts the downslope direction of the steepest slope (in degrees from 0 to 359.9). Slope and aspect have a major effect on solar radiation with steep south-facing slopes receiving more radiation than north-facing slopes. Therefore aspect, slope and elevation were considered as important predictor variables. The skyview factor, (stage 2) a measure of the visible sky based on the digital terrain model, was calculated using the SAGA tool Sky View Factor in QGIS 3.4.4. The skyview factor can affect the emission of longwave radiation and can be of importance in the formation of UHIs.

2.6 | Solar radiation, day length, azimuth

Diffuse and direct solar radiation and day length were estimated for each grid cell using the 'solrad' package in R (Seyednasrollah *et al*., 2013). The package uses day of the year, coordinates, slope, aspect and elevation to estimate the potential diffuse and direct solar radiation in Watt per square metre as well as the day length in hours and the solar azimuth angle in degrees. Solar radiation is directly related to temperature and therefore potentially an important predictor. Day length and solar azimuth angle were chosen as seasonal indicators.

2.7 | Meteorological data

Daily meteorological data were extracted from the ERA-interim (global atmospheric reanalysis) data set from the European Centre for Medium-Range Weather Forecasts (ECMWF) (Dee *et al*., 2011). The data set has an approximate resolution of 80 × 80 km (T255 spectral) from which we obtained the provided downscaled resolution of 10 × 10 km. Parameters used were daily boundary layer height, 2 m temperature, 10 m U wind component, 10 m V wind component, total cloud cover and total precipitation modelled for 3 p.m. (UTC) in the afternoon.

2.8 | Land use

Different land uses can have a significant effect on the local climate. Compared to natural surfaces built-up areas have a lower permeability, a lower albedo, a higher heat capacity and decreased turbulent heat transport. The reduced vegetation results in a reduced evaporative cooling. However, a number of factors increase the turbulence over built-up areas. The warmer surface in a city can increase the instability of the atmosphere and therefore enhance buoyancy, especially during heatwaves with low wind speeds. Also, the roughness of a city surface is relatively large, which promotes the development of atmospheric turbulence, which explains why the atmospheric boundary layer over a city is usually deeper than over the nearby countryside. Buildings can cause a reduction of wind speed and are a source of anthropogenic heat. The canyon structure of streets results in a reduced albedo compared to the countryside and a reduced skyview factor resulting in a reduced longwave radiation. These effects are major factors in the creation of UHIs (Christen *et al*., 2017).

To control for this effect road and building density are included into the model as predictors. From the SwissTLM3D dataset from Swisstopo the road density was calculated for each 100 m grid cell. To account for the wider surroundings, a focal mean function with a circular moving window of 500 m was applied. The same approach was used for the building density. Both were combined by summing the values to a combined index variable to account for sealed surfaces. Land use data were extracted from the Corine Land Cover dataset (2012). The different land use categories were recoded into five main categories (artificial, open space, forest, agriculture/wetlands and glaciers) (see Table S2).

The predictor variables are summarized in Table S3 (stage 1) and Table S4 (stage 2).

2.9 | Statistical methods

A two-stage modelling approach was used to estimate daily near surface temperature (T_{mean}, T_{min}, T_{max}) at a fine spatial resolution. We chose random forest from the R 'ranger' package (Wright and Ziegler, 2017) as our method in both the stages as previous studies have shown random forest to be a good and reliable method for spatiotemporal temperature modelling (Noi *et al*., 2017; Li and Zha, 2018; Zhou *et al*., 2020). In the first stage, the

imputation of missing satellite data, we impute missing values in four T_s variables (MODIS Aqua and Terra, both for day and night at 1 × 1 km) with the equation:

$$T_{S_{\mathrm{MODIS}_{ij}}} \sim \mathrm{RF}\begin{pmatrix}\mathrm{Day\_length}_{jy}, & \mathrm{Day\_of\_year}_{j}, & \mathrm{NDVI}_{im} \\ \mathrm{Elevation}_{i,} & \mathrm{Azimuth}_{jxy} & \end{pmatrix}, \quad (1)$$

where $T_{S_{\mathrm{MODIS}_{ij}}}$ is the MODIS surface temperature at grid cell i on day j; $\mathrm{Day\_length}_{jy}$ is length of day in hours for day i by latitude y; $\mathrm{Day\_of\_year}_{j}$ is day of year for day j; NDVI_{im} is the Landsat NDVI of cell i for calendar month m; $\mathrm{Elevation}_{i}$ is the elevation in cell i; and $\mathrm{Azimuth}_{jxy}$ is the azimuth solar angle for day j by longitude x and latitude y. For each year, four models were built to impute missing data in existing MODIS Aqua (1:30 a.m. and 1:30 p.m.) and Terra (10:30 a.m. and 10:30 p.m.) surfaces (mtry parameter and minimal node size were both set to 4 after hyperparameter tuning using the R package 'caret' [Kuhn, 2008], the number of trees was set to 300).

The validation process is illustrated in Figure 2. In stage 1, a randomly selected 20% of the grid cells with T_s measurements was set aside as a truly independent validation dataset to test the performance of the models by year. The remaining 80% of data were randomly split into five equally sized subsets and a five-fold cross-validation resampling was performed where five models were iteratively developed on four training subsets and validated on one validation subset. Out of the five models, the best performing model was then used to predict the independent validation set calculating the 'out-of-sample' R^2 and root mean squared error (RMSE). This best model was then used to impute T_s in the grid cells with missing data. The resulting 1 × 1 km daily 'gap filled' estimates of the four T_s variables (different satellites and overpasses) were then averaged to form one surface temperature variable and linked to measured ground level temperature measurements (T_a).

In stage 2, we then applied another random forest model where we explained the spatiotemporal variation in the measured data at the ground monitoring stations at a 100 × 100 m resolution using the following equation:

where Ta_{ij} is the temperature at 2 m height at grid cell i on day j; $\mathrm{Ts\_modis}_{ij}$ is the gapfilled mean of MODIS

$$T_{aij} = \widetilde{\mathrm{RF}}\begin{pmatrix}\mathrm{Ts\_modis}_{ij}, \mathrm{Ts\_modis\_lag\_lead}_{ij}, \mathrm{Ta\_era}_{ij}, \mathrm{Diff\_radiation}_{jxy}, \mathrm{Dir\_radiation}_{jxy}, \\ \mathrm{Elevation}_{i}, \mathrm{NDVI}_{ia}, \mathrm{Day\_of\_year}_{j}, \mathrm{Latitude}_{i}, \mathrm{Impervious\_surface}_{i}, \mathrm{Skyview}_{i}, \\ \mathrm{Boundary\_layer\_era}_{ij}, \mathrm{Wind\_v10\_era}_{ij}, \mathrm{Wind\_u10\_era}_{ij}, \mathrm{Cloud\_cover\_era}_{ij}, \\ \mathrm{Precipitation\_era}_{ij}, \mathrm{Landcover\_clc}_{i}\end{pmatrix}, \quad (2)$$

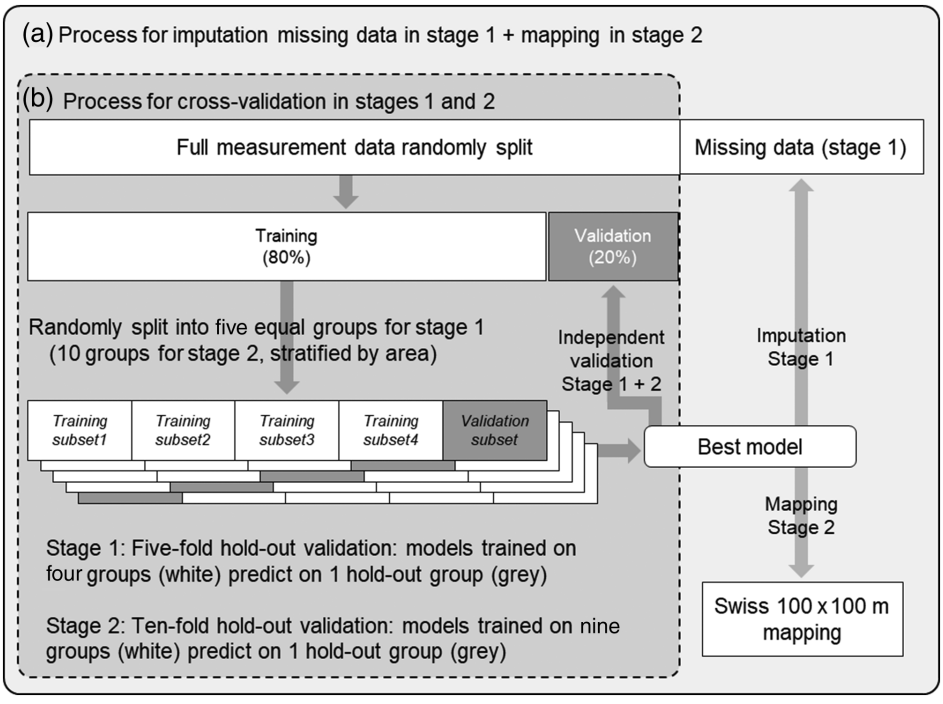

FIGURE 2 Cross-validation strategy for the models in stages 1 and 2. For both models, 20% of the data was kept aside for a final evaluation. The models were trained with 80% of the data, using hold-out validation

TABLE 1 Performance indicators for annual models tested on the hold-out dataset showing R^2, RMSE, slope and intercept by year, satellite and overpass

	Aqua day				Aqua night				Terra day				Terra night			
Year	R^2	RMSE (°C)	Slope	Intercept (°C)	R^2	RMSE (°C)	Slope	Intercept (°C)	R^2	RMSE (°C)	Slope	Intercept (°C)	R^2	RMSE (°C)	Slope	Intercept (°C)
2003	0.98	1.72	1.00	−0.48	0.98	1.45	1.01	−1.53	0.98	1.64	1.00	−0.26	0.99	1.11	1.00	−0.95
2004	0.97	1.75	1.00	−0.60	0.97	1.52	1.01	−1.89	0.98	1.66	1.00	−0.73	0.98	1.19	1.00	−1.33
2005	0.98	1.73	1.00	−0.48	0.97	1.49	1.01	−1.62	0.98	1.67	1.00	−0.37	0.98	1.17	1.00	−1.04
2006	0.98	1.74	1.00	−0.82	0.97	1.53	1.01	−1.90	0.98	1.67	1.00	−0.58	0.98	1.23	1.00	−1.03
2007	0.97	1.71	1.00	−0.73	0.96	1.44	1.01	−2.08	0.97	1.64	1.00	−0.33	0.98	1.13	1.01	−1.42
2008	0.97	1.68	1.00	−0.76	0.97	1.44	1.01	−2.02	0.98	1.62	1.00	−0.54	0.98	1.18	1.01	−1.39
2009	0.98	1.70	1.00	−0.57	0.97	1.48	1.01	−1.58	0.98	1.66	1.00	−0.56	0.98	1.21	1.00	−1.19
2010	0.98	1.67	1.00	−0.54	0.98	1.48	1.01	−1.47	0.98	1.64	1.00	−0.44	0.99	1.17	1.00	−1.22
2011	0.97	1.72	1.00	−0.42	0.97	1.41	1.01	−1.53	0.97	1.69	1.00	−1.00	0.98	1.12	1.00	−1.22
2012	0.98	1.63	1.00	−0.80	0.97	1.39	1.01	−1.66	0.98	1.55	1.00	−0.65	0.98	1.14	1.00	−1.13
2013	0.98	1.65	1.00	−0.33	0.97	1.48	1.01	−1.47	0.98	1.60	1.00	−0.43	0.98	1.20	1.00	−1.38
2014	0.96	1.67	1.00	−1.09	0.96	1.47	1.01	−2.76	0.97	1.62	1.00	−0.97	0.97	1.23	1.01	−2.19
2015	0.97	1.75	1.00	−0.81	0.97	1.41	1.01	−1.65	0.97	1.68	1.00	−0.17	0.98	1.12	1.00	−1.25
2016	0.97	1.71	1.00	−0.78	0.97	1.42	1.01	−1.67	0.98	1.63	1.00	−0.73	0.98	1.14	1.00	−1.15
2017	0.98	1.69	1.00	−0.79	0.97	1.48	1.00	−1.38	0.98	1.64	1.00	−0.30	0.98	1.20	1.00	−1.31
2018	0.97	1.77	1.00	−0.84	0.97	1.40	1.01	−1.95	0.97	1.71	1.00	−1.00	0.98	1.19	1.00	−1.16

surface temperature of Aqua and Terra day and night at grid cell i on day j; Ts_modis_lag_lead$_{ij}$ is the average of lag and lead value of Ts_modis at grid cell i on day j; Ta_era$_{ij}$ is the temperature at 2 m height from the ERA-Interim reanalysis data at grid cell i on day j; Diff_radiation$_{jxy}$ is the diffuse radiation on day j by longitude x and latitude y; Dir_radiation$_{jxy}$ is the potential direct radiation on day j by longitude x and latitude y; Elevation$_i$ is the elevation in cell i; NDVI$_{ia}$ is the Landsat NDVI of cell i for the year a; Day_of_year$_j$ is day of year for day j; Latitude is the latitude at y; Impervious_surface is the area covered by buildings or roads at grid cell i; Skyview is the calculated skyview factor at grid cell i; Boundary_layer_era$_{ij}$ is the boundary layer height at grid cell i on day j; Wind_v10_era$_{ij}$ is the 10-m V wind component at grid cell i on day j; Wind_u10_era$_{ij}$ is the 10-m U wind component at grid cell i on day j; Cloud_cover_era$_{ij}$ is the total cloud cover at grid cell i on day j; Precipitation_era$_{ij}$ is the total precipitation at grid cell i on day j; Landcover_clc$_i$ is the Corine Land Cover class for grid cell i. Land use was handled as an unordered factor covariate.

Models for each year predicting daily T_a were separately developed for T_{mean}, T_{min} and T_{max}. Like in the first stage, the initial ground monitoring data were split randomly into a training data set (80%) and an independent validation set (20%). Different from stage 1, the training set was split into 10 folds, stratified by area, so that each area contained approximately 10% of measurement stations. The second stage cross-validation was conducted by leaving out each time a full area (Figure 2). As every year has a different number of measurement stations available, the area boundaries used to split the data into 10 equal groups changed slightly year by year. The random forest mtry parameter (the number of predictors randomly sampled as candidates at each split) was set to 6 and minimal node size was set to 4 after hyperparameter tuning using the R package 'caret'. The

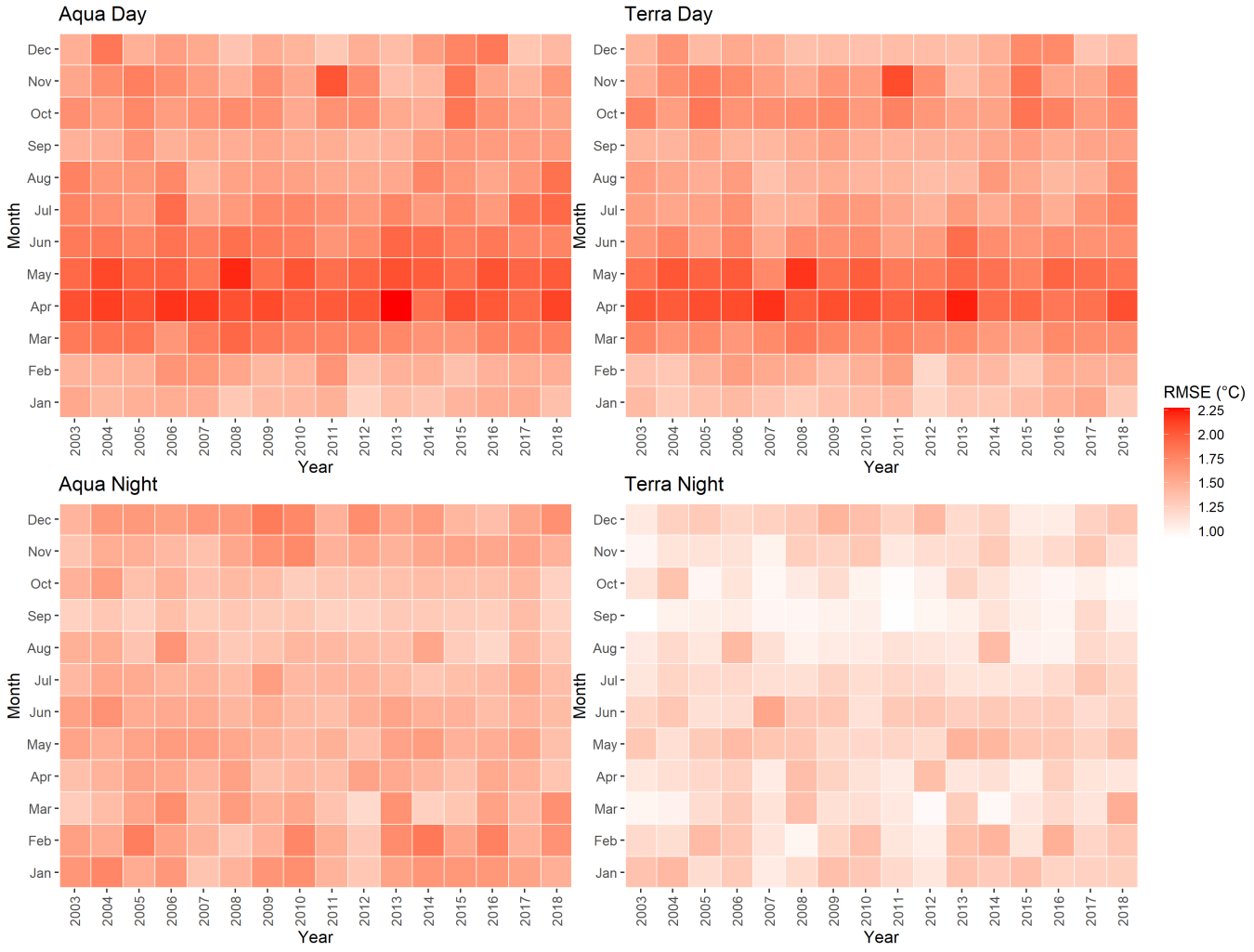

FIGURE 3 Model performance (RMSE on the hold-out dataset) of stage 1 averaged over the whole area by month, year, satellite and overpass time [Colour figure can be viewed at wileyonlinelibrary.com]

number of trees was set to 400 after manual model parameter optimization. The best model was applied to predict T_a for the independent validation set and R^2 and RMSE. The best model was also used to predict daily T_a (T_{mean}, T_{min} and T_{max}) across Switzerland at a 100 × 100 m resolution.

We compared our temperature predictions with forecasts from an existing temperature model from MeteoSwiss. The Federal Office of Meteorology and Climatology, MeteoSwiss produces the daily mean temperature (TabsD) at an approximate 2 × 2 km resolution across Switzerland for 2003–2018 as part of their MeteoSwiss Grid-Data Products (https://www.meteoschweiz.admin.ch/home/klima/schweizer-klima-im-detail/raeumliche-klimaanalysen.html, accessed 22 June 2021). The model uses a deterministic method to spatially interpolate daily temperature station data. The underlying statistical analysis of this product integrates information from weather stations, radar and satellite. Firstly, vertical temperature dependencies were calculated using station data constructing a smooth surface across Switzerland and secondly, residuals between the station measurements and the vertical profiles were interpolated with non-Euclidean distances, instead taking into account the complex topography (Frei, 2014). We aggregated our 100 × 100 m grid cells to match the 2 × 2 km grid resolution of the MeteoSwiss product and calculated for both products monthly means.

In addition, we assessed the ability of the model to capture the UHI effect and temperature differences due to topography-related weather phenomena typical for Switzerland, like the foehn effect and inversion conditions.

3 | RESULTS

3.1 | Stage 1

Figure S1 shows the percentage of missing satellite T_s data, mainly caused by cloud cover, by year for the different satellites and overpasses. All years have between 50 and 70% of missing data for each overpass. Aqua day generally has most missing T_s data (up to 68% in 2010), followed by Terra day, Terra night and Aqua night. Within-year variation (not shown) is substantial with missing data varying from 32.8% (Aqua night April 2007) to 83.4% (Aqua day May 2006). Figure S2 shows a map of the overall data availability for Terra and Aqua over Switzerland, clearly showing the impact of elevation with a higher percentage of missing MODIS data at lower altitudes. The percentage of missing MODIS data for 2003–2018 was calculated by summing up available data for all satellites and overpasses for each grid cell divided by the total amount of potential data points at the specific grid cell.

TABLE 2 Performance indicators (R^2, RMSE, slope and intercept) of the stage 2 models for T_{mean}, T_{min} and T_{max} by year

	T_{min}				T_{mean}				T_{max}			
Year	R^2	RMSE (°C)	Slope	Intercept (°C)	R^2	RMSE (°C)	Slope	Intercept (°C)	R^2	RMSE (°C)	Slope	Intercept (°C)
2003	0.97	1.58	1.02	0.03	0.98	1.22	1.01	−0.02	0.97	1.70	1.01	−0.13
2004	0.95	1.86	1.01	−0.04	0.98	1.26	1.01	−0.05	0.96	1.79	1.01	−0.08
2005	0.97	1.52	1.01	0.02	0.98	1.29	1.01	0.01	0.97	1.78	1.01	−0.04
2006	0.95	1.77	1.01	−0.01	0.98	1.32	1.01	−0.03	0.96	1.80	1.01	−0.09
2007	0.96	1.55	1.02	0.02	0.97	1.25	1.01	−0.03	0.96	1.64	1.01	−0.05
2008	0.95	1.80	1.02	−0.03	0.97	1.34	1.01	−0.01	0.97	1.55	1.01	−0.09
2009	0.97	1.49	1.01	0.05	0.98	1.22	1.01	−0.01	0.97	1.70	1.01	−0.08
2010	0.97	1.61	1.01	0.02	0.99	1.16	1.01	−0.01	0.97	1.68	1.01	−0.04
2011	0.95	1.66	1.02	−0.02	0.98	1.17	1.01	−0.04	0.96	1.74	1.01	−0.06
2012	0.97	1.51	1.01	0.01	0.98	1.24	1.01	−0.06	0.97	1.73	1.01	−0.07
2013	0.96	1.67	1.01	−0.02	0.99	1.05	1.01	−0.02	0.97	1.63	1.01	−0.08
2014	0.96	1.48	1.02	−0.02	0.98	1.12	1.01	−0.05	0.97	1.60	1.02	−0.11
2015	0.96	1.55	1.02	−0.02	0.98	1.14	1.01	−0.05	0.97	1.63	1.01	−0.10
2016	0.96	1.51	1.02	−0.04	0.98	1.15	1.01	−0.06	0.97	1.59	1.01	−0.14
2017	0.97	1.60	1.02	−0.04	0.98	1.17	1.01	−0.05	0.97	1.64	1.01	−0.10
2018	0.97	1.50	1.01	−0.03	0.98	1.29	1.01	−0.06	0.97	1.77	1.01	−0.11

The stage 1 random forest models explained large parts of the spatiotemporal variation in the T_s data (Table 1). The annual stage 1 models achieved an overall R^2 of 0.98 and an RMSE of 1.49°C (hold-out validation). The slope was close to 1 (from 1.00 to 1.01) and the intercept was closer to 0 for the Aqua and Terra day (from −1.09 to −0.17) than the Aqua and Terra night (from −2.76 to −0.95). Within-year variability of predictive power was substantial as illustrated in Figure 3 showing RMSE for each satellite, daytime, year and month for the hold-out data (R^2 showed in Figure S3). Day models generally performed better from October to March (Figure S3) opposed to April to June, which is matched with higher RMSEs observed for those months (Figure 3). Differences between satellites were most apparent in the night with Terra models generally outperforming Aqua models.

Figure S4 shows the overall variable importance plot for all the year models (stage 1) as box plots. The variable importance was calculated with the mean decrease in impurity, which gives an indication of how much a feature contributes relatively to the outcome estimation. The 'length of day' variable is the strongest variable in almost all models followed by 'day of the year'. NDVI and elevation are similar with azimuth only having a small influence in the random forest models.

Figure S5 illustrates the results of 'gapfilling' showing T_s data before and after stage 1 for the Aqua daytime overpass on 16 February 2015. Large parts of north-west and south-east Switzerland were covered by clouds around 10 a.m. on 16 February 2015 (first map). The stage 1 model imputes the missing clear-sky T_s data, resulting in a complete 'gapfilled' Aqua day (second map).

3.2 | Stage 2

The performance (R^2 and RMSE) of the stage 2 models on the hold-out validation set predicting daily T_{min}, T_{mean} and T_{max} by year is shown in Table 2. Slope and intercept were calculated with a linear regression between

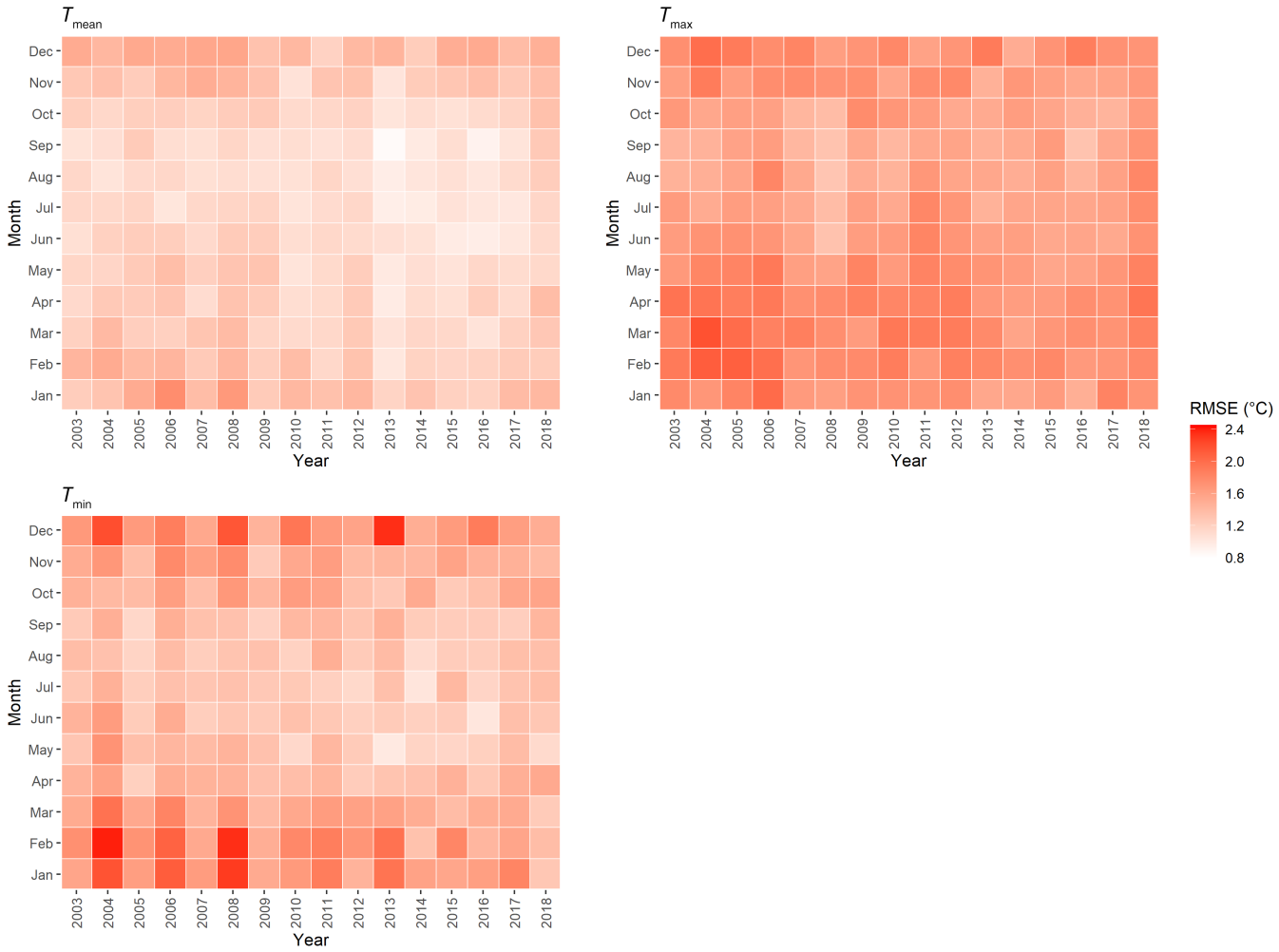

FIGURE 4 Mean RMSE by month, year, T_{mean}, T_{max} and T_{min} of the stage 2 model on the hold-out dataset [Colour figure can be viewed at wileyonlinelibrary.com]

measured and predicted data on all datasets (training data and hold-out validation). The stage 2 model performed well for all years with R^2 and RSME ranging from 0.94 to 0.99 and 1.05 to 1.86°C, respectively. The within-year variability of the models is shown in Figure 4 with monthly RMSE scores for T_{min}, T_{mean} and T_{max} (and R^2 in Figure S6). As before in the stage 1 models the best performing months for T_{mean} and to a lesser extent for T_{min} and T_{max} are April to September also reflected in similar patterns for RMSE. Stage 2 models showed the best performance when using all satellites data averaged (daily mean of aqua day, aqua night, terra day and terra night) as one single predictor.

The variable importance for the T_{mean}, T_{min} and T_{max} models is shown in Figure S7 as box plots over the 2003–2018 period. LST from MODIS (T_s_modis and T_s_modis_Iag_lead) consistently explained most of the spatiotemporal variation in daily air temperature, followed by ERA-interim temperature (T_a_era), diffuse radiation (diff_radiation) and elevation as the most influential predictor variables. This confirms the crucial contribution of satellite-derived data in our temperature models.Figure 5 shows as an example a map of the predicted minimum, mean and maximum temperature on 15 July 2015 at a 100 × 100 m resolution.

A complete overview over all annual models is given in Figure S8. It shows the annual temperature difference of T_{mean} with the overall mean from 2003 to 2018. A warming trend is observed with the last 5 years on average warmer than the 2003–2018 average.

3.3 | Comparison with MeteoSwiss model

Monthly aggregated predictions from both models across Switzerland for 2003–2018 were identical, strengthening the robustness of the predictions of both models (Figure S9). We conducted a more detailed comparison, plotting the mean, first and 99th percentile of daily mean temperature for 2017 from both models (Figure S10), which showed a very good temporal agreement with a correlation coefficient of $R = 0.997$. The complex topography in Switzerland, with narrow valleys and steep mountain slopes in the Alps, can lead to significant temperature changes over short distances. When using coarse

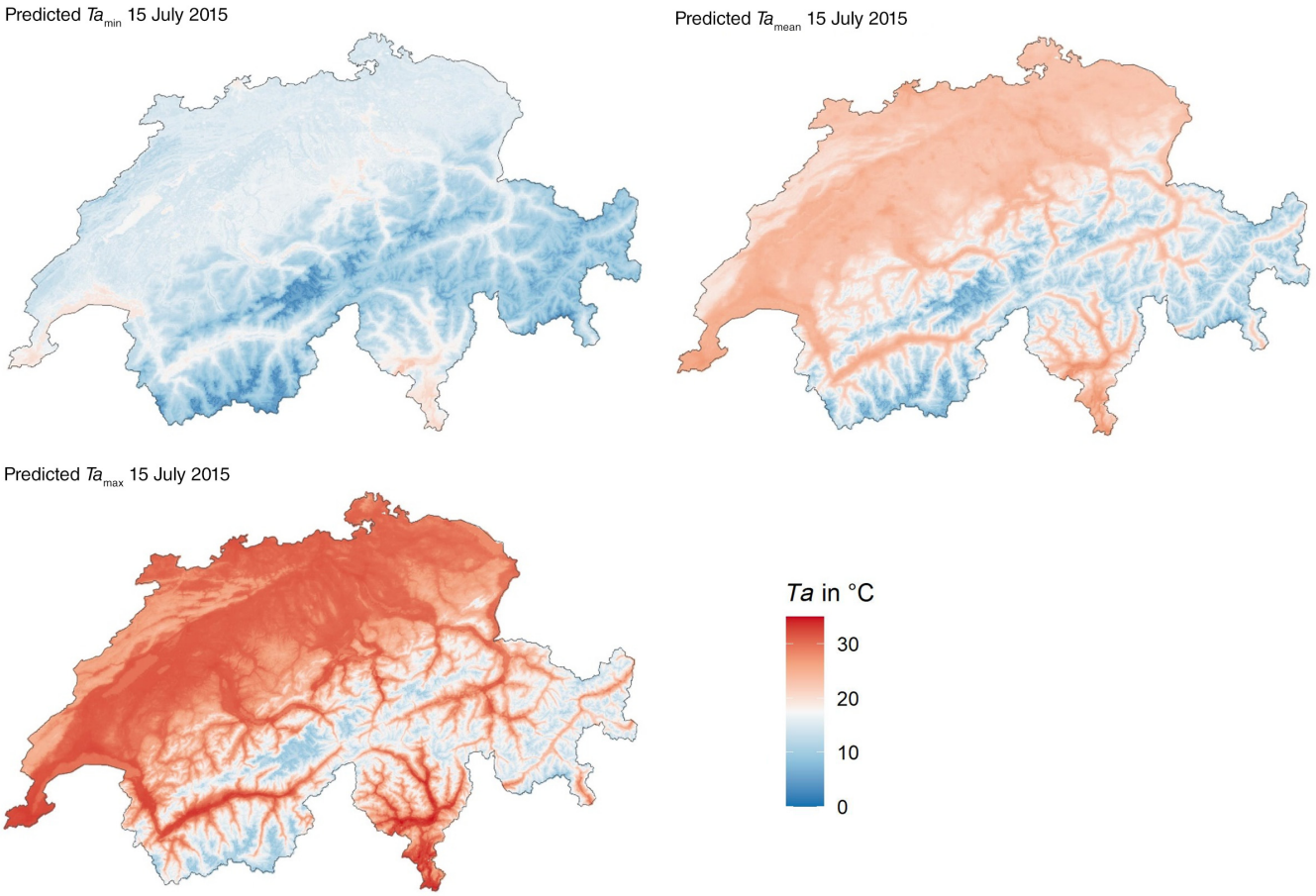

FIGURE 5 Ta_{min}, Ta_{mean} and Ta_{max} predictions for 15 July 2015 [Colour figure can be viewed at wileyonlinelibrary.com]

temperature grids, edge effects can potentially lead to significant exposure misclassification. To illustrate this, two maps in Figure S11 show an alpine valley (with the city of Bellinzona in the centre) with on the left the daily mean temperature at a 2 × 2 km resolution predicted by MeteoSwiss and on the right predictions at the 100 × 100 m resolution from our model. In absolute terms, the temperature range of two model predictions are similar (from 7 to 37°C), but the high temperatures in our model are restricted to the valley, and do not appear at higher altitudes, something unavoidable with the coarser 2 × 2 km resolution. We also compared our estimates with the MeteoSwiss model and a downscaled version of the MeteoSwiss model at a 100 × 100 m resolution (using bilinear interpolation) during a heatwave in Zurich (Figure S12). It shows a profile of predicted daily mean temperature along a transect through Zurich and surrounding area on a day during a heatwave on 5 July 2015. Our model (Swiss TPH) predicts more localized differences than both the coarser and downscaled MeteoSwiss model. By modelling at a finer spatial resolution we therefore managed to minimize potential exposure misclassification, especially in these types of areas with complex terrain.

4 | DISCUSSION

This paper describes the development of daily air temperature (daily mean, min and max) at a fine spatial resolution of 100 × 100 m across Switzerland leveraging satellite, atmospheric reanalyses data and other spatiotemporal predictor variables in a machine learning framework. Our modelling framework is a two-stage approach. Firstly, we imputed missing values caused by cloud-cover in satellite-derived surface temperature with length of day the strongest predictor variable, followed by day of the year, NDVI and elevation. Secondly, we developed random forest models predicting fine-resolution air temperature based on various spatiotemporal predictors, including surface temperature from ground monitoring, temperature predictions from ECMWF reanalysis and potential solar radiation as the most important predictors. The improved spatial resolution represents a substantial improvement upon exposure assessment methods commonly used in epidemiological studies, which typically use central monitors or coarser resolution exposure maps.

4.1 | Comparison with other modelling studies

We are the first study estimating Swiss nationwide daily ambient temperature at a 100 × 100 m resolution with excellent results (mean 2003–2018: R^2 = 0.98, RMSE = 1.21°C). Our findings are similar to recent studies carried out in Europe. Hough *et al.* (2020) modelled daily ambient temperature (mean, max and min) across France from 2000 to 2016 at a 1 × 1 km resolution across France and at a 200 × 200 m resolution in urban areas with R^2 ranging from 0.92 to 0.97 (RMSE 1.3–1.9°C). The findings in this paper also are comparable with other studies modelling at a coarser resolution of 1 × 1 km conducted in France (Kloog *et al.*, 2017) (R^2 = 0.95–0.96, RMSE = 2.16°C) and Israel (Rosenfeld *et al.*, 2017) (Aqua R^2 = 0.99, RMSE 0.70°C, Terra R^2 = 0.99, RMSE 0.67°C). A study in the Netherlands (Dirksen *et al.*, 2020), also modelling at the coarser 1 × 1 km, reported RMSE's below 1°C, which could be explained by the mostly flat terrain, and therefore, less contrast in temperature, compared countries with complex terrain like Switzerland.

4.2 | Urban heat island effect

One aim of this study was to capture the UHI effect with our model. Without independent air temperature measurement data, we were not able to conduct a quantitative assessment of characterizing the UHI effect. A qualitative assessment is presented in Figure S12 which shows a profile of predicted daily mean Temperature along a transect through Zurich and surrounding area on a day during a heatwave on 5 July 2015. The profile graph shows the elevated temperatures modelled in the city of Zurich compared to temperatures in the surrounding area. Additionally, it shows variability within the city of Zurich. By predicting temperature levels at a fine resolution of 100 × 100 m, we are getting closer at capturing the UHI effect in Swiss cities, compared to using a central weather station, the coarser gridded (2 × 2 km) MeteoSwiss product or indeed a 100 × 100 m downscaled version of the MeteoSwiss product. However, further improvements capturing the UHI effect can be achieved by addressing some data limitations in the current study. In 2018, only 15 stations out of 576 were located in inner cities. Weather stations from routine monitoring networks are typically positioned to avoid local influences and are thus located in suburban areas, in rural locations, or, in Switzerland, also at mountain tops. For epidemiological studies investigating the relationship between heat and health, temperature exposure surfaces are needed which better capture temperature gradients in areas where people live and work, which are typically not at airports, rural areas, or, at least in Switzerland, at mountain tops. To better capture the UHI effect, weather station networks need to expand into urban areas. Another venue to explore is to develop fusion models

combining routine weather station data with weather data collected using cheap sensors in citizen's science initiatives or to use non-official weather data from for example weather underground (Muller *et al*., 2015; Chapman *et al*., 2017). Another improvement could be made by including some of the variables that potentially explain microclimatic differences like wind speed and direction at lower heights, and shadows cast by buildings and street canyons. This was recently illustrated by Koopmans *et al*. (2020) who build a fine-scale heatmap for Wageningen in the Netherlands by including amongst other variables building configuration, wind speed and tree register data.

4.3 | Weather phenomena

We were able to model temperature differences during specific meteorological phenomena for Switzerland. The 'foehn' wind in Switzerland occurs when there is a difference in pressure between the south- and the north-side of the Alps. There are two foehns, a south foehn, the most frequent one, occurring when pressure on the north-side is lower than on the south-side, and a north foehn, when the opposite occurs. The effect of the foehn wind is, apart from high wind speeds, an increase in temperature in the Alpine valleys. On 23 November 2003, a south foehn occurred and Figure S13 shows temperatures of around 15°C in the Alpine valleys, more than expected on a 'normal' day in November compared with temperatures in the Espace Mittelland (i.e. Swiss plateau).

Another meteorological phenomenon in Switzerland is inversion. Inversions occur mainly in winter or autumn for example with fog or low stratus clouds, resulting in higher temperatures above the inversion layer compared to below. The 22nd of December 2007 illustrates an inversion in Switzerland with low temperatures at ground level, rising with altitude and then dropping again (Figure S14). Because our model only predicts temperature at surface level, we mimicked the balloon measurements by a transect line starting at the same location into the Alps. The temperature on this line was then aggregated by height. The result is also a temperature gradient by height, comparable with the balloon measurements. As we can see in Figure S14, both show a similar trend, lower temperature at ground level, higher temperatures in average altitudes and then again cold temperatures in high altitudes. The balloon measurements were started at noon, this likely explains the higher temperatures than shown on the aggregated transect (representing daily mean temperatures).

4.4 | Predictor variables

The absence of satellite data during cloudy conditions is a limitation in developing temperature models based on satellite derived temperature estimates (Kloog, 2019). Switzerland has prolonged periods of cloud cover, especially in the Espace Mittelland, leading to a high percentage of missing data for satellite-observed surface temperature. Both the Aqua and Terra satellite data have between 50 and 70% missing data for each overpass. This potentially can lead to less precise temperature estimates in areas and/or during periods affected. However, with our imputation step (stage 1) we were able to build random forest models allowing prediction of air temperature on days without satellite-observed surface temperature data. Due to the higher cloud cover during the cold seasons, more surface temperature data had to be imputed during these months. This resulted in the models of the summer months (May–September, 2003–2018) showing a better performance (RMSE of T_{mean} = 1.10°C) than the models of the winter months (October–April 2003–2018, RMSE of T_{mean} = 1.28°C). This difference is however not extremely large, and we consider the great value of having a high-resolution temperature model for Switzerland which covers the entire year, also in light of possible future applications. Another limitation is that we used meteorological predictor variables in our models from the slightly older ERA-interim product compared to the now available ERA5 product. ERA5 has an improved spatial and temporal resolution and representation of geophysical processes and, therefore, meteorological variables (e.g. wind speed, wind direction, precipitation etc.) will be more accurate in future studies using ERA5 (Hersbach *et al*., 2020).

5 | CONCLUSION

This study presents a novel simplified modelling framework, to predict daily ambient temperature for a 16-year time series in Switzerland. By reducing the spatial resolution to 100 × 100 m compared to the coarser 2 × 2 km of existing models we were able to capture the UHI effect and some typical weather phenomena caused by Switzerland's complex topography, like the foehn effect and inversion conditions. The temperature models presented here were able to accurately capture temporal and spatial variations in air temperature in Switzerland using random forest from 2003 to 2018, with stage 1 models achieving an overall R^2 of 0.98 and an RMSE of 1.49°C (hold-out validation), and the stage 2 model performing well for all years with R^2 and RMSE ranging from 0.94 to 0.99 and 1.05 to 1.86°C, respectively. Independent hold-

out validation further showed a high performance of the model. The resulting daily temperature surfaces for 2003–2018 will facilitate ongoing epidemiological research investigating the health effects of heat.

AUTHOR CONTRIBUTIONS

Benjamin Flückiger: Data curation; formal analysis; investigation; methodology; visualization; writing – original draft; writing – review and editing. **Itai Kloog:** Formal analysis; methodology; writing – review and editing. **Martina S. Ragettli:** Project administration; writing – original draft; writing – review and editing. **Marloes Eeftens:** Data curation; formal analysis; writing – original draft; writing – review and editing. **Martin Röösli:** Conceptualization; writing – review and editing. **Kees de Hoogh:** Conceptualization; formal analysis; investigation; methodology; supervision; writing – original draft; writing – review and editing.

ACKNOWLEDGEMENTS

This work was supported by the Swiss Federal Office for the Environment and Federal Office of Public Health (grant number 3727). Open Access Funding provided by Universitat Basel. [Correction added on 20 May 2022, after first online publication: CSAL funding statement has been added.].

CONFLICT OF INTEREST

All authors declare that they have no conflicts of interest.

ORCID

Kees de Hoogh https://orcid.org/0000-0001-5974-2007

REFERENCES

Benmarhnia, T., Deguen, S., Kaufman, J.S. and Smargiassi, A. (2015) Review article: vulnerability to heat-related mortality: a systematic review, meta-analysis, and meta-regression analysis. *Epidemiology*, 26(6), 781–793.

Chapman, L., Bell, C. and Bell, S. (2017) Can the crowdsourcing data paradigm take atmospheric science to a new level? A case study of the urban heat island of London quantified using Netatmo weather stations. *International Journal of Climatology*, 37(9), 3597–3605.

Christen, A., Mills, G., Voogt, J.A. and Oke, T.R. (Eds.). (2017) Urban heat island. In: *Urban Climates*. Cambridge: Cambridge University Press, pp. 197–237.

Coffel, E.D., Horton, R.M. and de Sherbinin, A. (2017) Temperature and humidity based projections of a rapid rise in global heat stress exposure during the 21st century. *Environmental Research Letters*, 13(1), 014001.

Corine Land Cover (CLC 2012) Raster database [Internet]. Available at: https://land.copernicus.eu/pan-european/corine-land-cover/clc-2012/view.

de Schrijver, E., Folly, C.L., Schneider, R., Franco Duran, O.H., Gasparrini, A., Vicedo Cabrera, A.M, de Schrijver Evan, Folly Christophe L., Schneider Rochelle, Royé Dominic, Franco Oscar H., Gasparrini Antonio and Vicedo-Cabrera Ana M. (2021) A comparative analysis of the temperature-mortality risks using different weather datasets across heterogeneous regions. GeoHealth, 5(5). https://doi.org/10.1029/2020gh000363

Dee, D., Uppala, S., Simmons, A., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C.M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J., Haimberger, L., Healy, S.B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N. and Vitart, F. (2011) The ERA-interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137(656), 553–597. https://doi.org/10.1002/qj.828

Dirksen, M., Knap, W.H., Steeneveld, G.-J., Holtslag, A.A.M. and Tank, A.M.G.K. (2020) Downscaling daily air-temperature measurements in The Netherlands. *Theoretical and Applied Climatology*, 142(1), 751–767. https://doi.org/10.1007/s00704-020-03313-1

Frei, C. (2014) Interpolation of temperature in a mountainous region using nonlinear profiles and non-Euclidean distances. *International Journal of Climatology*, 34(5), 1585–1605.

Gasparrini, A., Guo, Y., Hashizume, M., Lavigne, E., Zanobetti, A., Schwartz, J., Tobias, A., Tong, S., Rocklöv, J., Forsberg, B., Leone, M., De Sario, M., Bell, M.L., Guo, Y.-L.L., Wu, C.-f., Kan, H., Yi, S.-M., de Sousa Zanotti Stagliorio Coelho, M., Saldiva, P.H.N., Honda, Y., Kim, H. and Armstrong, B. (2015) Mortality risk attributable to high and low ambient temperature: a multicountry observational study. *The Lancet*, 386(9991), 369–375.

Gehrig, R.K.N. and Scherrer, S. Städtische Wärmeinseln in der Schweiz – klimatologische Studie mit Messdaten in fünf Städten. 2018.

Heaviside, C., Macintyre, H. and Vardoulakis, S. (2017) The urban Heat island: implications for health in a changing environment. *Current Environmental Health Reports*, 4(3), 296–305.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R.J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S. and Thépaut, J.-N. (2020) The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730), 1999–2049. https://doi.org/10.1002/qj.3803

Hough, I., Just, A.C., Zhou, B., Dorman, M., Lepeule, J. and Kloog, I. (2020) A multi-resolution air temperature model for France from MODIS and Landsat thermal data. *Environmental Research*, 183, 109244.

Ilango, S.D., McElroy, S. and Schwarz, L. (2020) Recommendations for epidemiologic studies of aging populations in a changing climate. *International Journal of Public Health*, 65(8), 1431–1432.

Kloog, I. (2019) Use of earth observations for temperature exposure assessment in epidemiological studies. *Current Opinion in Pediatrics*, 31(2), 244–250.

Kloog, I., Brent, A.C., Francesco, N. and Joel, S. (2014) Predicting spatiotemporal mean air temperature using MODIS satellite surface temperature measurements across the Northeastern USA. *Remote Sensing of Environment*, 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. and Schwartz, J. (2017) Modelling spatio-temporally resolved air temperature across the complex geo-climate area of France using satellite-derived land surface temperature data. *International Journal of Climatology*, 37(1), 296–304.

Koopmans, S., Heusinkveld, B.G. and Steeneveld, G.J. (2020) A standardized physical equivalent temperature urban heat map at 1-m spatial resolution to facilitate climate stress tests in the Netherlands. *Building and Environment*, 181, 106984.

Kuhn, M. (2008) Building predictive models in R using the caret package. *Journal of Statistical Software*, 28(5), 2008. https://doi.org/10.18637/jss.v028.i05

Lee, J.Y., Kim, H., Gasparrini, A., Armstrong, B., Bell, M.L., Sera, F., Lavigne, E., Abrutzky, R., Tong, S., de Sousa Zanotti Stagliorio Coelho, M., Saldiva, P.H.N., Correa, P.M., Ortega, N. V., Kan, H., Garcia, S.O., Kyselý, J., Urban, A., Orru, H., Indermitte, E., Jaakkola, J.J.K., Ryti, N.R.I., Pascal, M., Goodman, P.G., Zeka, A., Michelozzi, P., Scortichini, M., Hashizume, M., Honda, Y., Hurtado, M., Cruz, J., Seposo, X., Nunes, B., Teixeira, J.P., Tobias, A., Íñiguez, C., Forsberg, B., Åström, C., Vicedo-Cabrera, A.M., Ragettli, M.S., Guo, Y.-L.L., Chen, B.-Y., Zanobetti, A., Schwartz, J., Dang, T.N., Do Van, D., Mayvaneh, F., Overcenco, A., Li, S. and Guo, Y. (2019) Predicted temperature-increase-induced global health burden and its regional variability. *Environment International*, 131, 105027.

Lee, J.Y., Röösli, M. and Ragettli, M.S. (2021) Estimation of heat-attributable mortality using the cross-validated best temperature metric in Switzerland and South Korea. *International Journal of Environmental Research and Public Health*, 18(12), 6413.

Levi, M., Kjellstrom, T. and Baldasseroni, A. (2018) Impact of climate change on occupational health and productivity: a systematic literature review focusing on workplace heat. *La Medicina del lavoro*, 109(3), 163.

Li, L. and Zha, Y. (2018) Mapping relative humidity, average and extreme temperature in hot summer over China. *Science of the Total Environment*, 615, 875–881.

Muller, C.L., Chapman, L., Johnston, S., Kidd, C., Illingworth, S., Foody, G., Overeem, A. and Leigh, R.R. (2015) Crowdsourcing for climate and atmospheric sciences: current status and future potential. *International Journal of Climatology*, 35(11), 3185–3203.

Noi, P.T., Degener, J. and Kappas, M. (2017) Comparison of multiple linear regression, cubist regression, and random Forest algorithms to estimate daily air surface temperature from dynamic combinations of MODIS LST data. *Remote Sensing*, 9(5), 398. https://doi.org/10.3390/rs9050398

Oudin Åström, D., Bertil, F. and Joacim, R. (2011) Heat wave impact on morbidity and mortality in the elderly population: a review of recent studies. *Maturitas*, 69(2), 99–105.

Ragettli, M.S., Vicedo-Cabrera, A.M., Schindler, C. and Röösli, M. (2017) Exploring the association between heat and mortality in Switzerland between 1995 and 2013. *Environmental Research*, 158C, 703–709.

Rosenfeld, A., Dorman, M., Schwartz, J., Novack, V., Just, A.C. and Kloog, I. (2017) Estimating daily minimum, maximum, and mean near surface air temperature using hybrid satellite models across Israel. *Environmental Research*, 159, 297–312.

Seyednasrollah, B., Kumar, M. and Link, T.E. (2013) On the role of vegetation density on net snow cover radiation at the forest floor. *Journal of Geophysical Research: Atmospheres*, 118(15), 8359–8374. https://doi.org/10.1002/jgrd.50575

Shi, L., Liu, P., Kloog, I., Lee, M., Kosheleva, A. and Schwartz, J. (2016) Estimating daily air temperature across the southeastern United States using high-resolution satellite data: a statistical modeling study. *Environmental Research*, 146, 51–58.

Song, X., Wang, S., Hu, Y., Yue, M., Zhang, T., Liu, Y., Tian, J. and Shang, K. (2017) Impact of ambient temperature on morbidity and mortality: an overview of reviews. *Science of the Total Environment*, 586, 241–254.

Vicedo-Cabrera, A.M., Scovronick, N., Sera, F., Royé, D., Schneider, R., Tobias, A., Astrom, C., Guo, Y., Honda, Y., Hondula, D.M., Abrutzky, R., Tong, S., de Sousa Zanotti Stagliorio Coelho, M., Nascimento Saldiva, P.H., Lavigne, E., Matus Correa, P., Valdes Ortega, N., Kan, H., Osorio, S., Kyselý, J., Urban, A., Orru, H., Indermitte, E., Jaakkola, J.J.K., Ryti, N., Pascal, M., Schneider, A., Katsouyanni, K., Samoli, E., Mayvaneh, F., Entezari, A., Goodman, P., Zeka, A., Michelozzi, P., De'Donato, F., Hashizume, M., Alahmad, B., Hurtado Diaz, M., De La Cruz Valencia, C., Overcenco, A., Houthuijs, D., Ameling, C., Rao, S., Di Ruscio, F., Carrasco-Escobar, G., Seposo, X., Silva, S., Madureira, J., Holobaca, I.H., Fratianni, S., Acquaotta, F., Kim, H., Lee, W., Iniguez, C., Forsberg, B., Ragettli, M.S., Guo, Y.L.L., Chen, B.Y., Li, S., Armstrong, B., Aleman, A., Zanobetti, A., Schwartz, J., Dang, T.N., Dung, D.V., Gillett, N., Haines, A., Mengel, M., Huber, V. and Gasparrini, A. (2021) The burden of heat-related mortality attributable to recent human-induced climate change. *Nature Climate Change*, 11(6), 492–500. https://doi.org/10.1038/s41558-021-01058-x

Watts, N., Amann, M., Arnell, N., Ayeb-Karlsson, S., Beagley, J., Belesova, K., Boykoff, M., Byass, P., Cai, W., Campbell-Lendrum, D., Capstick, S., Chambers, J., Coleman, S., Dalin, C., Daly, M., Dasandi, N., Dasgupta, S., Davies, M., Di Napoli, C., Dominguez-Salas, P., Drummond, P., Dubrow, R., Ebi, K.L., Eckelman, M., Ekins, P., Escobar, L.E., Georgeson, L., Golder, S., Grace, D., Graham, H., Haggar, P., Hamilton, I., Hartinger, S., Hess, J., Hsu, S.-C., Hughes, N., Mikhaylov, S.J., Jimenez, M.P., Kelman, I., Kennard, H., Kiesewetter, G., Kinney, P.L., Kjellstrom, T., Kniveton, D., Lampard, P., Lemke, B., Liu, Y., Liu, Z., Lott, M., Lowe, R., Martinez-Urtaza, J., Maslin, M., McAllister, L., McGushin, A., McMichael, C., Milner, J., Moradi-Lakeh, M., Morrissey, K., Munzert, S., Murray, K.A., Neville, T., Nilsson, M., Sewe, M.O., Oreszczyn, T., Otto, M., Owfi, F., Pearman, O., Pencheon, D., Quinn, R., Rabbaniha, M., Robinson, E., Rocklöv, J., Romanello, M., Semenza, J.C., Sherman, J., Shi, L., Springmann, M., Tabatabaei, M., Taylor, J., Triñanes, J., Shumake-Guillemot, J., Vu, B., Wilkinson, P., Winning, M., Gong, P., Montgomery, H. and Costello, A. (2021) The 2020 report of The Lancet Countdown on health and climate change: responding to converging crises. *The Lancet*, 397(10269), 129–170.

Wouters, H., De Ridder, K., Poelmans, L., Willems, P., Brouwers, J., Hosseinzadehtalaei, P., Tabari, H., Vanden Broucke, S., van Lipzig, N.P.M. and Demuzere, M. (2017) Heat stress increase under climate change twice as large in cities as in rural areas: a study for a densely populated midlatitude maritime region. *Geophysical Research Letters*, 44(17), 8997–9007. https://doi.org/10.1002/2017gl074889

Wright MN, Ziegler A. ranger: A fast implementation of random forests for high dimensional data in C++ and R. *Journal of Statistical Software*; Vol 1, Issue 1 (2017). 2017.

Ye, X., Wolff, R., Yu, W., Vaneckova, P., Pan, X. and Tong, S. (2012) Ambient temperature and morbidity: a review of epidemiological evidence. *Environmental Health Perspectives*, 120(1), 19–28.

Zeger, S.L., Thomas, D., Dominici, F., Samet, J.M., Schwartz, J., Dockery, D. and Cohen, A. (2000) Exposure measurement error in time-series studies of air pollution: concepts and consequences. *Environmental Health Perspectives*, 108(5), 419–426.

Zhou, B., Erell, E., Hough, I., Rosenblatt, J., Just, A.C., Novack, V. and Kloog, I. (2020) Estimating near-surface air temperature across Israel using a machine learning based hybrid approach. *International Journal of Climatology*, 40(14), 6106–6121. https://doi.org/10.1002/joc.6570

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Flückiger, B., Kloog, I., Ragettli, M. S., Eeftens, M., Röösli, M., K. de Hoogh (2022). Modelling daily air temperature at a fine spatial resolution dealing with challenging meteorological phenomena and topography in Switzerland. *International Journal of Climatology*, *42*(12), 6413–6428. https://doi.org/10.1002/joc.7597