New high-resolution gridded dataset of daily mean, minimum, and maximum temperature and relative humidity for Central Europe (HYRAS)

Christène Razafimaharo¹ · Stefan Krähenmann¹ · Simona Höpp¹ · Monika Rauthe¹ · Thomas Deutschländer¹

Received: 5 December 2019 / Accepted: 15 September 2020 / Published online: 26 September 2020

© The Author(s) 2020

Abstract

This study presents daily high-resolution (5 km × 5 km) grids of mean, minimum, and maximum temperature and relative humidity for Germany and its catchment areas, from 1951 to 2015. These observational datasets (HYRAS) are based upon measurements gathered for Germany and its neighbouring countries, in total more than 1300 stations, gridded in two steps: first, the generation of a background field, using non-linear vertical temperature profiles, and then an inverse distance weighting scheme to interpolate the residuals, subsequently added onto the background field. The modified Euclidian distances used integrate elevation, distance to the coast, and urban heat island (UHI) effect. A direct station-grid comparison and cross-validation yield low errors for the temperature grids over most of the domain and greater deviations in more complex terrain. The interpolation of relative humidity is more uncertain due to its inherent spatial inhomogeneity and indirect derivation using dew point temperature. Compared with other gridded observational datasets, HYRAS benefits from its high resolution and captures complex topographic effects. HYRAS improves upon its predecessor by providing datasets for additional variables (minimum and maximum temperature), integrating temperature inversions, maritime influence and UHI effect, and representing a larger area. With a long-term observational dataset of multiple meteorological variables also including precipitation, various climatological analyses are possible. We present long-term historical climate trends and relevant indices of climate extremes, pointing towards a significantly warming climate over Germany, with no significant change in total precipitation. We also evaluate extreme events, specifically the summer heat waves of 2003 and 2015.

Keywords Temperature measurement · Relative humidity measurement · Daily interpolation · High-resolution grid · Non-linear vertical temperature gradient · Climate trends · Climate extremes

1 Introduction

High-resolution observation grids of climate variables have become important components of climatological studies. These grids consist of station measurements interpolated at a desired resolution using one of the various interpolation methods available. They are used as inputs for various models (Heininger and Cullmann (2015), Gusyev et al. (2016), Stahl et al. (2017), Höllering et al. (2018)), and they can also be used for the bias correction of climate models (Imbery et al.

(2013), Willkofer et al. (2018), Meyer et al. (2019), Krähenmann et al. (2020)). High-resolution observation grids furthermore allow for spatial analyses based on continuous surfaces rather than on irregularly spaced point measurements, and long-term daily grids allow for analyses on different spatial scales and for the analysis of trends.

Table 1 presents examples of different daily observational datasets of temperature and relative humidity developed in the last decades, covering the area of interest, namely, Germany and Central Europe. At a global scale, Piper and Stewart (1996) used a nearest-neighbour approximation to develop a daily $1^\circ \times 1^\circ$ grid of mean temperature and daily temperature range for the year 1987. Caesar et al. (2006) also developed a global, albeit coarser (2.5° latitude $\times$ 3.75° longitude) daily grid of minimum and maximum temperature, over half a century (1946–2000), using angular-distance weighting interpolation. Haylock et al.

✉ Christène Razafimaharo
christene.verosoaa@outlook.com

¹ Deutscher Wetterdienst, Offenbach am Main, Germany

Table 1 Overview of existing grids for temperature (mean, minimum, and maximum) and relative humidity

Citation & Name of Dataset (if available)	Variable	Interpolation method	Time period	Temporal resolution	Region	Spatial resolution
Piper and Stewart (1996)	Mean temperature and daily temperature range	Nearest-neighbour	1987	Daily	Global	1° x 1°
Caesar et al. (2006) HadGHCND	Minimum and maximum temperature	Angular distance weighting interpolation of daily anomalies with respect to reference period	1946–2000	Daily	Global	2.5° latitude x 3.75° longitude
Haylock et al. (2008) E-OBS Version 17.0	Mean, minimum and maximum temperature	Monthly background fields from thin-plate splines and daily anomalies from external drift kriging	1950–2017	Daily	Europe	0.25° x 0.25°
Cornes et al. (2018) E-OBS Version 20.0e	Mean, minimum and maximum temperature	Monthly background fields from thin-plate splines and daily fields from reduced-rank thin-plate splines	1950–2019	Daily	Europe	0.25° x 0.25°
Krähenmann et al. (2011)	Minimum and maximum temperature	Monthly background fields from block kriging and interpolation of daily anomalies from block kriging	2005–2008	Daily	Europe (WMO Region 6)	5 km x 5 km
Ntegeka et al. (2013) EFAS-Meteo	Mean, minimum and maximum temperature	Inverse distance weighting	1990–2011	Daily	Europe	5 km x 5 km
DWD Climate Data Center (CDC) (2018) DWD Monthly	Mean, minimum and maximum temperature	Background fields from multiple linear regression, followed by inverse distance weighting	1881–last completed month	Monthly	Germany	1 km x 1 km
Frick et al. (2014) HYRAS-2006	Mean temperature, relative humidity	Optimal interpolation; relative humidity derived from temperature and vapour pressure.	1951–2006	Daily	Germany and neighboring countries (catchments)	5 km x 5 km
Krähenmann et al. (2018) TRY	Mean temperature, relative humidity	Daily background field from non-linear temperature gradient, multi-dimensional inverse distance weighting of anomalies. Relative humidity derived from temperature and dew point	1995–2012	Hourly	Germany	1 km x 1 km
Frei (2014)	Mean temperature	Daily background field from non-linear temperature gradient, multi-dimensional inverse distance weighting of anomalies.	1961–2010	Daily	Switzerland	1 km x 1 km
Hiebl and Frei (2016) SPARTACUS	Minimum and maximum temperature	Daily background field from non-linear temperature gradient, multi-dimensional inverse distance weighting of anomalies.	1961–2013	Daily	Austria	1 km x 1 km
Delvaux et al. (2015) FORBIO Climate dataset	Minimum and maximum temperature	External drift kriging	1980–2013	Daily	Belgium	4 km x 4 km
Berezowski et al. (2016) CPLFD-GDPT5	Minimum and maximum temperature	External drift kriging	1951–2013	Daily	Poland (Vistula and Oder catchments)	5 km x 5 km

This table lists the variables interpolated, the selected interpolation method, the spatial and temporal coverage, and resolutions. Datasets compared with HYRAS are italicized over a grey background

(2008) used a combination of thin-plate splines and kriging to interpolate daily $0.25^\circ \times 0.25^\circ$ grids of mean, minimum, and maximum temperature over Europe; known as the E-OBS dataset, these grids are available from 1950 to 2017 (as of the Version 17.0, used in this study for comparison). These grids were later on extended to 2019 (as of the latest Version 20.0e), based on an ensemble of daily fields (Comes et al. 2018). Higher resolution grids have also been developed, which cover the area of interest. The Joint Research Commission (Ntegeka et al. 2013) applied an inverse distance weighting (IDW) scheme to develop $5 \text{ km} \times 5 \text{ km}$ daily (1990–2011) grids of mean temperature over Europe. Krähenmann et al. (2011) generated daily grids of minimum and maximum temperature from 2005 to 2008 at a resolution of $5 \text{ km} \times 5 \text{ km}$ over Europe (WMO Region 6), using block kriging. Looking specifically at Germany, Frick et al. (2014) developed long-term (1951–2006) daily grids of mean temperature and relative humidity for Germany and its catchments ($5 \text{ km} \times 5 \text{ km}$), applying the Optimal Interpolation (Gandin 1966) and converting mean temperature to relative humidity using vapour pressure. These grids (HYRAS-2006) constitute the previous version of the dataset presented in this paper. Krähenmann et al. (2018) developed hourly grids of mean temperature and relative humidity over the period from 1995 to 2012 over Germany at a resolution of $1 \text{ km} \times 1 \text{ km}$, adapting the method developed by Frei (2014). The method of Frei (2014) was indeed first developed and applied to a $1 \text{ km} \times 1 \text{ km}$ daily (1961–2010) dataset of mean temperature for Switzerland, using the concept of non-linear vertical temperature gradients and IDW based on non-Euclidean distance metrics. Hiebl and Frei (2016) implemented a similar method for high-resolution grids of minimum and maximum temperature over Austria. Other daily minimum and maximum grids for neighbouring countries include the FORBIO dataset for Belgium (Delvaux et al. 2015) for the period 1980 to 2013 at a spatial resolution of $4 \text{ km} \times 4 \text{ km}$ and the CPLFD-GDPT5 dataset covering the period from 1951 to 2013 for the Oder and Vistula catchments in Poland (Berezowski et al. 2016) at a spatial resolution of $5 \text{ km} \times 5 \text{ km}$, both interpolated using external drift kriging.

The German Federal Ministry of Transport and Digital Infrastructure's project "Network of Experts" aims at understanding the effects of climate change on the German federal transport infrastructure, in order to adapt to climate change. Within this context, the high-resolution gridded daily dataset HYRAS for temperature and relative humidity, first developed within the project KLIWAS (Frick et al. 2014), was updated and expanded. The HYRAS dataset covers Germany and its catchment areas in neighbouring countries. This work presents the updated HYRAS dataset (HYRAS-2015) for mean, minimum, and maximum temperature and relative humidity (HYRAS-TAS v3.0, HYRAS-TMIN v3.0, HYRAS-TMAX v3.0, and HYRAS-HURS v3.0). This new version builds on its predecessor (HYRAS-TAS v1.01 and

HYRAS-HURS v1.01, also known as HYRAS-2006; Frick et al. (2014)), with the implementation of a new interpolation method based on Krähenmann et al. (2018) and Frei (2014), the temporal expansion of the dataset from 1951–2006 to 1951–2015, and the addition of new datasets for minimum and maximum temperature.

The objectives of this paper are (a) to present the new methodology used for the updated HYRAS dataset (HYRAS-2015) for temperature and relative humidity; (b) to assess the quality of the dataset as it compares with station observations, other observational datasets, and its predecessor (HYRAS-2006); and (c) to present applications of the HYRAS dataset in analysing long-term historical climate in Germany.

The database used for the interpolation is presented in Section 2. The interpolation methodology is then described in Section 3. The quality of the dataset is assessed through different validation methods presented in Section 4. In Section 5, climatological analyses based on the HYRAS dataset are presented. Finally, Section 6 presents the conclusions drawn from this study.

2 Database

The HYRAS-gridded dataset is based on daily measurements gathered for Germany and 7 neighbouring countries: Austria, Belgium, Czechia, France, Luxembourg, the Netherlands, and Switzerland, thus covering the major river basins of Germany (Rhine, Danube, Elbe, Ems, and Weser). The different data providers are listed in detail in Table 2. Station data from the Institute of Meteorology and Water Management in Poland (IMGW) were acquired at a later stage and were integrated in a new interpolation of HYRAS to include the Oder river basin. However, this new gridded dataset has not yet undergone a validation as thorough as the version presented here. Station data from Czechia were indeed used in deriving the HYRAS grids presented here. Upon an initial agreement with the data provider (the Czech Hydrometeorological Institute, CHMI), we hereby do not publish any information pertaining to these stations or results derived for that country¹.

The station measurements collected are not all temporally continuous, such that the number of stations used in the interpolation varies with time. Figure 1 shows the evolution of the number of stations available per month between January 1951 and December 2015, for daily mean temperature and relative humidity. The numbers of stations for minimum and for maximum temperature are not shown here, as they usually correspond to stations where the daily mean temperature is provided. By 1977, the total number of stations for daily mean temperature exceeds 1000 stations and increases until the year 1995, where it reaches a maximum of 1207 stations. From this

¹ The CHMI authorized the display of the data after this paper was prepared.

Table 2 Overview of observation data providers for the interpolation of HYRAS-TAS, HYRAS-TMIN, HYRAS-TMAX, and HYRAS-HURS

Country	Provider
Austria	Zentralanstalt für Meteorologie und Geodynamik (ZAMG)
Belgium	Royal Meteorological Institute of Belgium (RMIB)
Czechia	Czech Hydrometeorological Institute (CHMI)
France	Météo France (MFRA)
Germany	Deutscher Wetterdienst (DWD)
Luxembourg	Luxembourg Institute of Science and Technology/MeteoLux (LUXE)
The Netherlands	Royal Netherlands Meteorological Institute (KNMI)
Switzerland	MeteoSwiss (MSWI)

point to December 2015, the number of stations fluctuates around 1100 stations. For relative humidity, the total number of stations exceeds 1000 stations in 1992, reaches its peak of 1065 stations in 1995, and decreases to 800 stations by the end of December 2015. Furthermore, the spatial coverage is not completely consistent across the area of interest (see also Fig. 2 in Frick et al. (2014), now with data for all of Switzerland and Austria). The station density is highest in southwestern Germany, the Alps (Switzerland and Austria), Luxembourg, and Eastern France. In comparison, the stations in Northern Germany and the Netherlands are more dispersed.

Additionally, not all data providers use similar methods for aggregating their hourly measurements into daily data. Possible inconsistencies in daily aggregation methods are already presented in detail in Frick et al. (2014). These differences in daily aggregation of hourly data may lead to non-negligible discontinuities in the resulting interpolated datasets and should be considered when validating or utilizing them.

The station measurements collected were already quality controlled by the providing institutes. Nevertheless, some additional quality controls were applied to the station measurements, to discard outliers that may generate unrealistic interpolation results. For daily mean, minimum, and maximum temperature, values below $-50\text{ }^{\circ}\text{C}$ and above $50\text{ }^{\circ}\text{C}$ were discarded, as they

Fig. 2 Interpolation regions, defined according to climate characteristics

were deemed too extreme and thus unrealistic for the area studied. Based on these criteria, no outlier daily mean temperature measurements were found. Over the whole investigation period, 4 outlier minimum temperature measurements were found, and 19 outlier maximum temperature measurements were identified. For relative humidity, values below 0 % and above 100 % were discarded, as these are unrealistic. Over the entire investigation period, this resulted in a total of 410 outliers identified. Errors were usually limited to a few stations, over consecutive days, indicating a possible malfunctioning of the sensors during that period of time.

For trend analyses, the interpolated gridded dataset should be based on homogeneous measurements, in order to avoid producing spurious trends. Before producing the previous version of the HYRAS grids (HYRAS-2006), the station time series were tested for homogeneity, as detailed in Frick et al. (2014) and Rauthe et al. (2013). However, homogenizing the time series at all stations provided is computationally demanding, especially for daily data (Della-Marta and Wanner (2006), Mestre et al. (2011)). Additionally, this procedure requires detailed data on station history and the measurement devices used, which may not be feasible as the station data were collected from a variety of providers. As described by Guijarro

Fig. 1 Evolution of the monthly number of (a) temperature and (b) relative humidity observation stations used for the interpolation of HYRAS from 1951 to 2015; the different colours indicate different data providers

et al. (2017) and Venema et al. (2012), there exist several homogenization techniques, and they are still undergoing development and improvements. Therefore, most available gridded daily datasets are entirely or partially based on non-homogenized time series (see Table 1). Some institutions (e.g. MeteoSwiss or the Royal Netherlands Meteorological Institute) provide homogenized times series, but only for a small number of stations and only for monthly data, which cannot be used to calculate a high-resolution daily dataset. Thus, testing the station measurements for homogeneity and homogenizing them was not undertaken in this project and should also be considered when utilizing the gridded dataset.

3 Methodology

Several interpolation methods have been used in the past to generate regularly spaced grids from point measurements. According to Vicente-Serrano et al. (2003), these methods can generally be classified as global methods (trend surfaces, multiple linear regression), local interpolation methods (Thiessen polygons, IDW, splines), geostatistical methods (varieties of kriging), and mixed methods. The choice of interpolation method is highly dependent on the variable interpolated, the quality and density of observational data, the orography of the area, climate forcing factors (Daly 2006), and the target spatial and temporal resolution. Some studies delving closely into the comparison of interpolation methods for temperature were performed by Vicente-Serrano et al. (2003), Stahl et al. (2006), Hofstra et al. (2008), and Eiselt et al. (2017).

The previous HYRAS-TAS and HYRAS-HURS (Frick et al. 2014) datasets were interpolated using a method based on the Optimal Interpolation (Gandin 1966). However, improvements were required to better represent temperature and relative humidity in areas with complex topography, where temperature inversions may occur and consider additional factors such as the urban heat island (UHI) effect and cold air ponding. A different interpolation method integrating non-linear elevation temperature profiles (Frei 2014) and IDW using multi-dimensional distances was therefore implemented. Adapted at Germany's National Meteorological Service (Deutscher Wetterdienst, DWD) for the generation of high-resolution hourly grids of meteorological variables for Germany (Krähenmann et al. 2018), this method furthermore allows us to generate datasets for minimum and maximum temperature (HYRAS-TMIN and HYRAS-TMAX).

The general interpolation procedure for all four variables consists of first interpolating a background field, based on non-linear elevation temperature profiles. Then, the residuals between the observations and the interpolated background field are gridded using IDW with multi-dimensional distances weighted for elevation, distance to coast, and urban heat island effect. The resulting residual field is then added onto the

background field, to obtain the final grid. The interpolation procedure is applied for each day within the investigation period (1951 to 2015) and is based on subregions that are then assembled together. These subregions are defined subjectively, based on the different geographical features of Germany and its catchment areas. In total, 13 regions were defined, which overlap each other (see Fig. 2).

The HYRAS grids are first interpolated at a $1 \text{ km} \times 1 \text{ km}$ resolution and are subsequently aggregated onto a $5 \text{ km} \times 5 \text{ km}$ grid. This spatial resolution conforms to the requirements for the diverse applications of the dataset (e.g. hydrological modelling, bias correction of climate models) and to the resolution of the other HYRAS-2015 datasets of precipitation and global radiation (the latter being unpublished). The projection of the final grids is the Lambert conformal conic (LCC) projection for Europe, based on the European Terrestrial Reference System 1989 (ETRS89), which is the common projection of regional climate models in Europe. As specified in Frick et al. (2014), the ETRS89-LCC Europe projection uses the standard parallels 65°N and 35°N , with the reference coordinates 52°N and 10°E . The grid is nearly equidistant for our region of interest, as required for the selected interpolation algorithm.

3.1 Interpolation of daily mean temperature

The interpolation method is based on non-linear temperature gradients for different subregions in the domain of interest, fitted to the station observations. Therefore, only representative stations should be used in deriving the non-linear temperature profiles. To ensure this, stations in locations prone to cold air ponding and in locations with a highly pronounced urban heat island (UHI) effect were first identified and excluded from the vertical temperature profiles.

3.1.1 Removal of locations of cold air ponding

Locations prone to cold air ponding contain anomalies that are not representative of a subregion's vertical temperature profile. We decided to implement a method to identify stations prone to cold air ponding based on their elevation and minimum temperature measurements. This method is applied for each month of the investigation period. For each subregion, the average minimum temperature for a station is compared with its elevation. Each subregion is divided into 6 elevation layers between 200 and 1200 m.a.s.l., and the 25th percentile (lower quartile) value of the average minimum temperatures at the stations in each layer is selected. A linear piecewise interpolation function is then fitted between the 25th percentile of each layer and the layer's mean elevation. For a given subregion in the studied month, should a station's average minimum temperature be inferior to the function's approximation at that same elevation, the station is eliminated and not

considered in determining the subregion's non-linear vertical temperature profile. Depending on the subregion and the month evaluated, between 11 and 32 % of the stations have an average minimum temperature lower than the fitted function's value at their corresponding elevation. Therefore, these stations are disregarded when fitting the non-linear vertical temperature profile. At this point, additional outlier verification is implemented. Values exceeding 6 standard deviations for temperature and dew point (3 standard deviations for minimum and maximum temperature) were thus also excluded from the vertical temperature profile. The stations excluded in this step are nonetheless further used in the subsequent IDW interpolation of residuals.

3.1.2 Determining UHI effect

As the HYRAS grids are interpolated at a high spatial and temporal resolution, UHI can have a pronounced effect on the measured daily minimum temperatures and result in high temperatures due to the land surface rather than atmospheric influences. The potential UHI effect at the stations therefore needs to be estimated, as described in Krähenmann et al. (2018), based on the method of Wienert et al. (2013).

A UHI mask was created by using the CORINE land use gridded data (Keil et al. 2011). For different land uses, a coefficient for the UHI potential between 0 and 1 was assigned to each different land use, with 1 describing highly urbanized areas and coefficients near 0 describing green spaces (parks, forests) and water bodies. To determine the maximum UHI intensity, the UHI potential obtained was converted to an approximation of the population. Based on approximations in three large agglomerations (Berlin, Munich, and Frankfurt), the following relationship between the UHI potential and the population was determined:

$$E_w = 5 \cdot 10^6 \cdot UHI_{\text{potential}}^{2.1}$$

where E_w is the population and $UHI_{\text{potential}}$ the estimated UHI potential based on land use. The maximum UHI intensity is then the product of a function of the population and of a function of the time of the year, further factored by a ratio accounting for local effects.

The UHI potential at an observation station consists of the average UHI potential at all grid cells within a radius of 3 km. Stations exhibiting a maximal UHI potential of 2 °C were excluded prior to fitting a function to the non-linear vertical temperature gradient.

3.1.3 Fitting a non-linear vertical gradient for the daily background field

For a given subregion, the following non-linear function representing the vertical temperature gradient is fitted to the

measured daily temperatures (mean, dew point, and differences between minimum/maximum and mean) at the stations in this subregion (Frei 2014):

$$T(z) = \begin{cases} T_0 - \gamma z, & z > h_1 \\ T_0 - \gamma z - \frac{a}{2} \left(1 + \cos \left(\pi \cdot \frac{z - h_0}{h_1 - h_0} \right) \right), & h_0 \leq z \leq h_1 \\ T_0 - \gamma z - a, & z < h_0 \end{cases}$$

where h_0 is the lower and h_1 is the upper limit of the inversion layer, T_0 is the x -intercept (at an elevation $z = 0$) of the gradient, γ is the linear gradient, a the temperature offset between the layers above and below the inversion, and z the station elevation in metres. With this configuration, the temperature profile consists of 3 parts: first, a bottom layer ($z < h_0$) with linear gradient γ , then the inversion layer (between h_0 and h_1), and finally an upper layer ($z > h_1$), with the same linear gradient γ , offset from the bottom layer by a temperature difference a . Depending on the subregion, weather conditions, and season, inversions can vary in steepness (a) and in depth (h_0 and h_1). Figure 4 in Frei (2014) provides a schematic of this non-linear function. As the subregions in the domain overlap each other, some stations can be considered for the derivation of the non-linear temperature profiles in multiple subregions. Illustrations of the resulting non-linear temperature profiles for January 31, 2006, are provided in Fig. 3, for the subregions of the Thuringian Forest and the Bavarian Forest. These temperature profile functions are then applied to the elevation model (GTOPO30, USGS (1993)), to obtain the background field for temperature on the selected day. For grid cells where subregions overlap each other, the estimated background temperature is a weighted average of the temperature estimates using the different regions' non-linear gradients.

3.1.4 IDW for the daily residual field

Next, the residuals between the temperatures measured at the stations and the background field values at the nearest grid cells are calculated and interpolated. The interpolation consists of an IDW, implemented in 4 different regions from the north to the south of the domain, defined according to climate features and based on the aggregation of the 13 subregions previously described: a coastal region, an alpine region, and in between a northern and a southern region. Unlike in Frei (2014), where this interpolation is done using non-Euclidian distances, the distances are based on multi-dimensional Euclidian distances. In addition to the two-dimensional Euclidian distance between a grid cell and its neighbouring stations, this variant integrates the station elevation, the distance to the coast, and the UHI potential. The station elevation and the distance to the coast are additionally weighted to consider topographic obstructions (Frei 2014), and the potential UHI is weighted to be converted into an equivalent distance. The weights are obtained for each day and for each station,

Fig. 3 Examples of non-linear temperature profiles on January 31, 2006, and for different regions as defined in Fig. 2: (a) Thuringian Forest and (b) Bavarian forest. Each dot represents a station in the region. Grey dots are stations excluded from fitting the non-linear temperature profile, such

as stations prone to cold air ponding, outliers, and stations with a maximum UHI potential > 2 °C. The size of the dots indicates the weights given to the stations in fitting the non-linear temperature profile

based on a selection of possible values. The combination of weights is optimized by minimizing the root mean square error from a leave-one-out cross-validation scheme. The results for the 4 regions are subsequently merged using a weighted average. The coastal region's grid cells are weighted based on their distance to the coast. The other 3 regions are weighted by a factor of 1. Should 2 regions overlap, partial weights are used in each region over these areas. The sum of the partial weights from 2 overlapping regions is equal to 1. The resulting interpolated residuals are then added onto the background field, thus yielding the final daily mean temperature field.

3.2 Interpolation of relative humidity

The interpolation of relative humidity is achieved indirectly, as it depends on parameters representing the moisture content of the atmosphere (such as vapour pressure or specific humidity) and on the air temperature. Furthermore, Peixoto and Oort (1996) demonstrate that dew point temperature yields equivalent information as vapour pressure and can therefore be used in the derivation of relative humidity. This approach is adopted here, with first the conversion of observed temperature and relative humidity into observed dew point, followed by its interpolation onto a grid, and finally the conversion into a relative humidity grid using the previously interpolated mean temperature grids.

To obtain equivalent dew point temperature observations, the Magnus formula is used, which relates dew point to the observed mean temperature and relative humidity (Magnus 1844):

$$T_{\text{dew}} = K_3 \times \frac{\frac{K_2 T_{\text{air}}}{K_3 T_{\text{air}}} + \log(RH/100)}{\frac{K_2 K_3}{K_3 T_{\text{air}}} - \log(RH/100)}$$

where K_2 and K_3 are coefficients for calculating saturated vapour pressure, T_{air} is the observed daily mean temperature in °C, and RH the observed daily mean relative humidity in %. The coefficients for calculating vapour pressure only over water were used, as the daily mean temperatures were not very low, and the temperature threshold for condensation or freezing does not necessarily exactly correspond to 0 °C.

The dew point temperature observations obtained were then interpolated using the same method as for the mean temperature interpolation. Dew point temperatures exceeding the mean temperature were set to a ceiling value equal to the mean temperature. Finally, the interpolated dew point grid was converted into a relative humidity grid, following the relationship:

$$RH = \frac{e^{T_{\text{dew}}}}{e^{T_{\text{air}}}} \times 100$$

3.3 Interpolation of minimum and maximum temperature

Prior to interpolating the minimum and maximum temperature fields, the input station measurements were verified to ensure that the daily mean temperature did not exceed maximum temperature and that minimum temperature did not exceed daily mean temperature. Typically, the absolute minimum and maximum temperature measurements are directly interpolated onto grids (e.g. Caesar et al. (2006), Stahl et al. (2006), Haylock et al. (2008), Krähenmann et al. (2011), Hiebl and Frei (2016), Aalto et al. (2016), Werner et al. (2019)). However, here, we interpolate their deviation from the mean temperature observations, applying the method described previously. Rather than the absolute values, the deviations are indeed interpolated, in an attempt to minimize the range of values and facilitate fitting a non-linear vertical gradient to the values. Thus, the fields of residuals are obtained, which are subsequently added to the interpolated

daily mean temperature field to obtain the final minimum and maximum temperature grids.

4 Validation

In order to assess their quality, different methods were implemented to validate the HYRAS-2015 temperature and relative humidity datasets. These include the direct comparison between the station observations and the grids (Section 4.1), cross-validation (Sections 4.2 and 4.3), a comparison with other observational datasets (Section 4.4), and finally a comparison with previous versions of HYRAS (HYRAS-2006, Section 4.5).

4.1 Comparison with station observations

The quality of the interpolation can first be evaluated by comparing the station measurements to the gridded values at the nearest grid cell, selected by minimizing the two-dimensional Euclidian distance. Different error measures were calculated based on the difference between the station measurements and the grid cell values, such as the average bias and the mean average error (MAE). The formulas used to calculate these error measures are presented in the [Appendix](#). To facilitate the upcoming comparison with the results of Frick et al. (2014), we calculate the differences by subtracting the grid cell value from the station measurement. The error measures were calculated at each station and for each day from 1951 to 2015 and aggregated based on different temporal scales (annual cycle or complete time period) and spatial criteria (over the whole interpolation domain or according to elevation).

4.1.1 Annual cycles of error measures

Figure 4 shows the long-term annual cycle of the bias and of the MAE, averaged over all stations and grouped (averaged) by month, for the mean daily temperature, the relative humidity, the minimum, and the maximum temperature. The bias is always positive, meaning the measured temperature is warmer than at the grid cell and varies on average between 0.17 in December and January and 0.32 °C in April. On the one hand, this can be due to the actual station elevation being lower than the corresponding DEM elevation (see analysis in Section 4.1.2) but can also point to temperature inversions not being fully captured. It is still possible that some areas prone to cold air ponding were not correctly eliminated prior to fitting the non-linear vertical profile, leading to (unrealistic) cooler temperatures at higher elevations. On average for the mean temperature, the MAE varies between 0.62 in December and January and 0.73 °C in April. These low errors over all stations and their weak annual cycle are satisfying indicators of the consistency of the interpolation method throughout the year and different seasons.

The error measures for the minimum temperature have the least pronounced annual cycle. Indeed, the bias is always positive and varies on average between 0.15 in January and 0.2 °C in April, while the MAE varies on average between 0.74 in January and 0.67 °C in November, the range of values being wider in winter (80 % of the values between 0.64 and 0.82 °C in January) than in the summer (80 % of the values between 0.64 and 0.75 °C in June). Meanwhile, the error measures of the maximum temperature have a highly pronounced annual cycle. The bias, while remaining positive throughout the year, varies on average between 0.16 in January and above 0.35 °C in April, whereas on average, the MAE will be around 0.67 °C in January and increases above 0.92 °C in May. Since the horizontal temperature gradient is reduced in winter, this could explain the more pronounced cycle for the maximum temperature as well as the greater errors in summer than in winter. A reduced horizontal temperature gradient leads to a smoother temperature field and therefore to an improved representation of station data in the gridded dataset.

For the relative humidity, the bias is positive in autumn and winter (up to 0.5 % in January) and sinks under 0 % in the summer (− 0.2 % in June). The MAE varies on average between 2.7 in December and January and nearly 2.2 % in August. Unlike for the temperature error measures, the range of values of the error measures for relative humidity has a visible seasonality. Based on the percentiles of the calculated bias, values are greater in winter (80 % of the values between − 0.2 and 0.75 % in January) than in summer (80 % of the values between − 0.25 and 0 % in June). For the MAE, the range of values is also larger in winter (80 % of the values lying between 2.3 and 3.3 % in January) than in summer (80 % of the values between 2 and 2.5 % in June). The seasonality in these errors could be due to the variety of factors that determine the relative humidity (e.g. elevation of the inversion line). Additionally, the indirect interpolation using dew point temperature also adds uncertainty to the results.

All in all, when comparing the station measurements to their corresponding nearest grid cell value, the error measures are low for all variables, such that the interpolation generally works well over most of the domain.

4.1.2 Spatial distribution of error measures

Besides the annual cycle of the error measures, it is also necessary to understand how the errors vary in space. For example, when the long-term average errors at each station are evaluated (not illustrated here), the bias in mean temperature remains within ± 1 °C at most stations, with greater values occurring at stations in mountainous areas.

Elevation is indeed an important factor in deciding the quality of the interpolation. Therefore, the distribution of the error measures as a function of elevation was also analysed. Figure 5 shows the long-term average daily errors (bias and

Fig. 4 Long-term (1951–2015) annual cycle of the error measures from the comparison of the station and grid values, as averages over all stations and as percentiles. **(a)** Daily mean temperature, **(b)** relative humidity, **(c)** daily minimum temperature, **(d)** daily maximum temperature. Red lines and circles are for the bias; lines and triangles are for the MAE

MAE) over all stations for the mean daily temperature, grouped by elevation difference between the station and its nearest grid cell. Elevation intervals of 100 m were selected. Furthermore, to account for the elevation difference between the station and its nearest-grid cell, the grid cell temperature was adjusted accordingly. This was achieved as described in Frick et al. (2014), by multiplying the elevation difference by a constant moist adiabatic temperature lapse rate ($\partial T/\partial h = -6.5^\circ \text{C}/\text{km}$) and adding the resulting temperature increment to the grid cell temperature. The error measures between this adjusted grid cell value and the station value were also aggregated and displayed in Fig. 5. As can be seen, both the bias and MAE are around 0°C when averaged over stations where the elevation difference with the grid cell is ± 100 m. The majority of station-grid cell pairs used for calculating the long-term daily error measures also belong to this interval, which contributes to reducing the average over all pairs. As the difference in elevation increases however, so do the errors, the number of station-grid cell pairs decreases. The bias is positive (up to 4.8°C) at locations where the station is 1000 m lower than the grid cell, i.e. the grid cell underestimates the station temperature as it is located higher than the station. Conversely, the bias is negative (up to -4°C) at locations where the station is 1000 m higher than the grid cell, such that the grid cell overestimates the station temperature. The MAE

increases up to 5°C at 1000-m elevation difference. By considering the elevation difference and adjusting the grid cell temperature accordingly, the range of MAE values will decrease to 2.2°C for elevation differences of 1000 m. The bias decreases but also reverses, adopting negative values up to -1.3°C when the station is 1000 m lower than the grid cell and up to 2°C when the station lies 1000 m above the grid cell. Although the linear height correction reduces the differences between the station and its nearest grid cell, this adjustment still remains approximate, as it neglects varying humidity (by utilizing a constant rate) and the occurrence of inversions (by utilizing a linear relationship).

This analysis yields similar results for the minimum and maximum temperature; hence, the results are not shown here. For relative humidity, no clear pattern with respect to elevation could be identified. This nevertheless highlights the importance of considering the underlying DEM used as basis for the interpolation and the propagation of differences due to the differences between the grid cell elevation and the actual station elevation.

4.2 Leave-one-out cross-validation

A leave-one-out cross-validation is applied to optimize the selection of factors for weighting the elevation, the distance

Fig. 5 Long-term average daily errors over all stations for the mean daily temperature, grouped by elevation difference between the station and its nearest grid cell. Elevation intervals of 100 m were selected. The results are shown for the direct differences in values (solid lines) as well as for

the differences, linearly adjusted based on elevation (dashed lines). Red lines and circles are for the bias; lines and triangles are for the MAE. The black dots indicate the number of pairs in the elevation difference interval considered

to the coast, and the UHI potential, used in the IDW of residuals (as described in Section 3.1.4). This procedure is used to evaluate the performance of the interpolation method when stations are eliminated from the interpolation. This is achieved in multiple runs, each time removing a different station from the ensemble and estimating the value at the removed station's location. The estimation is then verified against the observation.

Table 3 shows the obtained long-term (1951–2015) cross-validation scores (bias and MAE), averaged over all stations, also separated according to seasons (winter and summer). These scores correspond to the results obtained with the optimized weights for the IDW. This cross-validation was performed for all variables. In the case of relative humidity, this was indirectly achieved by cross-validating the dew point temperature and converting it to relative humidity using the mean temperature (as described in Section 3.2).

For all variables, the overall bias is close to 0 °C (or 0 %), meaning that the interpolation method does not systematically under- or overestimate the values when stations are absent from the interpolation. This is consistent with the findings of Hiebl and Frei (2016) for minimum and maximum temperature, where the method of Frei (2014) was also applied to Austria. Here, the application of the interpolation method to dew point temperature to eventually obtain relative humidity is also verified. The bias in mean temperature is greater in the winter than in the summer, likely due to

maritime effects or UHI effect not being captured when stations are neglected from the interpolation.

The MAE for the temperature variables varies between 0.6 and 0.8 °C, depending on the variable and the season considered. These values are generally comparable with those obtained in Frei (2014) and in Hiebl and Frei (2016). The error scores are greater for minimum temperature than for maximum temperature, possibly due to unresolved UHI effect that could not be completely reduced. The MAE for the relative humidity is around 4.5 %, which is influenced by uncertainties in the conversions of measurements into dew point, the uncertainties in the interpolation of dew point, and finally the uncertainties in the final conversion into relative humidity.

The bias values for the temperature variables are lower than the values obtained in the station-grid cell comparison presented previously, further emphasizing the role of the grid cell location and elevation in the resulting errors. Besides, it should be noted that this leave-one-out cross-validation procedure does not require recalculating the non-linear gradients for the background fields. Therefore, it does not consider changes in temperature profiles and their ability to capture temperature inversions. As only one station is removed in each cross-validation run, this requires increased computing effort, although a single station may have little influence on the temperature profile. This is nonetheless addressed in the following fivefold cross-validation (Section 4.3).

Table 3 Summary of long-term (1951–2015) leave-one-out cross-validation scores (bias and MAE) over all stations, for mean, minimum, and maximum temperature and relative humidity

	TAS		TMIN		TMAX		HURS	
	Bias	MAE	Bias	MAE	Bias	MAE	Bias	MAE
Annual	0.01 °C	0.67 °C	0.03 °C	0.75 °C	− 0.03 °C	0.72 °C	0.2 %	4.42 %
Winter	0.02 °C	0.73 °C	0.03 °C	0.72 °C	− 0.03 °C	0.69 °C	0.21 %	4.59 %
Summer	0 °C	0.64 °C	0.03 °C	0.8 °C	− 0.02 °C	0.73 °C	0.18 %	4.36 %

The differences stem from the comparison of the station observation and the estimation at the same location. The overall averages and the seasonal averages are listed

4.3 Fivefold cross-validation

Another cross-validation was also performed to evaluate the quality of the grid over a reduced dataset and validate it against an independent dataset (not used as input for the interpolation). For the HYRAS-TAS and the HYRAS-HURS datasets, fivefold cross-validation was implemented. In this case, the non-linear gradients are also recalculated, such that we can evaluate the effect of the reduced station dataset on reproducing non-linear temperature profiles. The cross-validations of HYRAS-TMIN and HYRAS-TMAX were not performed, as they would yield similar results as the cross-validation of HYRAS-TAS, and already underwent an intensive leave-one-out cross-validation.

For each cross-validation run, the previously mentioned difference measures (bias, MAE) were calculated, this time comparing the station and its nearest grid cell. Figure 6 shows the long-term (1951–2015) annual cycle of each error measure, aggregated over all 5 runs of the cross-validation of the temperature and relative humidity datasets, respectively. Besides the long-term average, the different percentiles (median, 10th, 25th, 75th, and 90th) could also be calculated, to evaluate the range of the error measures.

For the mean daily temperature, the bias is mostly positive throughout the year, with the 90th percentile of bias values at 0 °C in January or more in other months. On average, the bias is 0.2 °C in December and increases to 0.4 °C in April. The bias remains close to that of the original interpolation in winter, such that the removal of stations from the interpolation does not generate greater positive or negative errors. In the summer however, the bias from the cross-validation exceeds that of the original interpolation by up to + 0.1 °C, indicating that the removal of stations from the interpolation will generate greater positive errors, such as when inversions at higher elevations are not captured. Indeed, the absence of stations will modify the non-linear profiles of the background fields. A comparison of the long-term spatial distribution of bias based on elevation (not illustrated here) shows that the cross-validation bias deviates from the original interpolation's bias by less than ± 0.1 °C at 53 % of the stations in the domain,

the majority of those (90 %) being located at elevations of 500 m.a.s.l. or less. Conversely, the cross-validation bias deviates from the original interpolation's bias by more than ± 1 °C at only 13 out of 2425 stations evaluated, all located at 2100 m.a.s.l. or above, highlighting the increase in error due to missing stations at high elevations. The MAE lies between 1 in October and 1.2 °C in April. The range of possible values (based on the 10th and 90th percentiles) is greater in the winter than in the summer (0.85 to 1.3 °C in January, 1 to 1.25 °C in July). The greater range of errors in winter can be attributed to the occurrences of inversions, which may not be properly captured when certain stations at high elevations are missing from the interpolation (see previous explanation for the bias). The MAE resulting from the fivefold cross-validation is also greater than the MAE from the original interpolation with all stations (on average and in terms of range of values), also indicating that stations missing from the interpolation will generate greater errors.

For the relative humidity, the annual cycles of the error measures are mostly positive, with bias values up to the 75th percentile at 0 % or more in December/January. Nevertheless, from April to August, the average and the median of the bias are slightly below 0 %. As was the case for the mean temperature, the obtained average bias is greater than the average bias from the original interpolation, with values throughout the year between -0.2 in April and 0.75 % in January. The range of values based on the percentiles also increases, which is due to the greater uncertainty generated by the reduced station ensemble. Accordingly, the average MAE increases by + 2 % in comparison with the original interpolation. With the reduction of the number of stations used in the cross-validation interpolation, the inter-decile range also widens to a range of values between 3.5 (10th percentile) and 7 % (90th percentile) in December/January, versus 3.5 % (10th percentile) and 5 % (90th percentile) in the summer months.

4.4 Comparison with other datasets

In this section, the HYRAS-dataset is compared with other observational datasets, available for a similar time period

Fig. 6 Results of the cross-validation of (a) the mean daily temperature and (b) the relative humidity, averaged over each month and all stations in the HYRAS domain. The error scores result from the comparison of the station value against its nearest grid cell. The annual cycles of the bias (red lines and circles) and the MAE (blue lines and triangles) are shown as averages and as percentiles

and geographical region. The datasets used for comparison and their characteristics are italicized and shaded in grey in Table 1.

4.4.1 Comparison with E-OBS

First, HYRAS is compared with the gridded European Climate Assessment & Dataset E-OBS dataset (Version 17.0, Haylock et al. 2008). E-OBS consists of daily grids from 1950 to 2017 for Europe and parts of the Middle East, at a resolution of $0.25^\circ \times 0.25^\circ$, i.e. approximately $25 \text{ km} \times 25 \text{ km}$, for daily mean, minimum, and maximum temperature. The E-OBS dataset is interpolated from up to 1100 stations over Germany (Cornes et al. 2018) using a combination of thin-plate splines for the interpolation of monthly mean temperature and external drift kriging for the interpolation of daily anomalies.

Figure 7 a and b show a direct comparison of the long-term daily fields of mean temperature for HYRAS and E-OBS, over the complete time period from 1951 to 2015. To achieve this, HYRAS was aggregated onto the coarser E-OBS grid, using a bilinear interpolation implemented through the Climate Data Operators (CDO, Schulzweida et al. 2018). To calculate the differences between the datasets, the long-term bias and MAE were estimated for each grid cell, subtracting E-OBS from HYRAS, such that positive bias differences imply a warmer HYRAS value. Generally, there is good agreement between the two datasets over the flatter regions of Germany (bias of $\pm 0.8^\circ \text{C}$, MAE under 0.4°C). Some mountainous regions in Germany stand out with a negative bias of -0.8 to -2.4°C , such as the Harz Mountains, the Thuringian Forest, and the Black Forest: as the topographic detail may not be captured by the coarser resolution of E-OBS, it will yield warmer temperatures than HYRAS. The differences between the two datasets in the Alps are also highly variable, up to 3.6°C of absolute difference and a bias exceeding $\pm 2^\circ \text{C}$. When averaged over the domain, the bias -0.01°C and the MAE is 0.4°C . Similar results are obtained when comparing the fields of daily minimum and maximum temperatures (not shown here). As seen in Table 4, the bias for mean temperature deviates most from 0°C in the winter (-0.03°C), when inversions are not captured by E-OBS. This is also the case for the maximum temperature (-0.01°C). The minimum temperature datasets deviate most from each other, with a bias of -0.07 to -0.08°C throughout the year. This is mainly due to greater deviations in minimum temperatures in the Alps, such as when local cold air ponding in valleys are not properly captured by the coarser E-OBS grid.

4.4.2 Comparison with DWD monthly dataset

Next, HYRAS was also compared with the DWD monthly datasets of mean, minimum, and maximum temperature. The

DWD monthly datasets are routine products made available from 1881 to the last completed month, at a spatial resolution of $1 \text{ km} \times 1 \text{ km}$ over Germany (DWD Climate Data Center (CDC) 2018). They are interpolated using the method of Müller-Westermeier (1995), which consists of a multiple linear regression (using elevation, slope, and vegetation as predictors) to construct a long-term monthly climatology, followed by an IDW of monthly averages measured at stations, where temperatures are linearly adjusted to sea level to use a two-dimensional Euclidian distance.

As was done for the comparison with E-OBS, Figs. 7 c and d show a direct comparison of the long-term (1951–2015) monthly fields of mean temperature for HYRAS and the DWD monthly datasets, by showing the spatial distribution of the bias and the MAE. To achieve this, the finer DWD monthly grids were aggregated onto the HYRAS $5 \text{ km} \times 5 \text{ km}$ grid, using a bilinear interpolation implemented through CDO (Schulzweida et al. 2018), and the HYRAS daily datasets were averaged monthly. The differences were obtained by subtracting the DWD monthly grids from the HYRAS grid, such that positive differences imply a warmer HYRAS grid. As was already the case in the comparison with E-OBS, here the DWD monthly datasets and HYRAS agree well with each other over Germany. The bias over Germany is mostly within $\pm 0.8^\circ \text{C}$, without any evident pattern. Again, a positive bias of 1.6°C or more is noticeable in the Black Forest and the Alpine foothills. In these areas of greater elevation, this positive bias is likely the result of differences in the consideration of elevation in the interpolation methods. Nevertheless, the average over Germany remains low at -0.02°C . The MAE is under 0.8°C almost everywhere, with mountain ranges such as the Black Forest and the Alpine foothills showing MAE exceeding 1.2°C but less than 2.0°C . Averaged over Germany, the MAE is 0.16°C . Whereas the DWD monthly datasets only consider a linear relationship between temperature and elevation (by linearly adjusting the measured temperatures to sea level), a complex non-linear vertical gradient is used in HYRAS. Thus, inversions will be represented by HYRAS, yielding warmer temperatures at higher elevations.

4.4.3 Comparison as yearly cycles over Germany

Figure 8 expands the comparison between HYRAS and E-OBS and between HYRAS and the DWD monthly datasets by showing the long-term (1951–2015) annual cycles of difference measures, for the monthly mean, minimum, and maximum temperatures, as averages over Germany only (domain of the DWD monthly dataset).

The absolute differences between HYRAS and E-OBS are greater than the absolute differences between HYRAS and the DWD monthly grids. This can first occur due to the differences in grid resolution, as both HYRAS and the DWD monthly grids are interpolated on a $1 \text{ km} \times 1 \text{ km}$ grid, such

Table 4 Comparison of HYRAS-2015 and E-OBS and HYRAS-2015 and HYRAS-2006 as long-term (1951–2015 and 1951–2006, respectively) seasonal and annual biases

	HYRAS-2015 vs E-OBS			HYRAS-2015 vs HYRAS-2006	
	TAS	TMIN	TMAX	TAS	HURS
Spring	− 0.01 °C	− 0.08 °C	0 °C	0.03 °C	− 0.24 %
Summer	− 0.01 °C	− 0.08 °C	0 °C	0.02 °C	− 0.09 %
Fall	0 °C	− 0.07 °C	0 °C	0.04 °C	− 0.26 %
Winter	− 0.03 °C	− 0.07 °C	− 0.01 °C	− 0.01 °C	− 0.09 %
Annual	− 0.01 °C	− 0.08 °C	0 °C	0.02 °C	− 0.17 %

The spatial averages over a common domain are listed. A positive bias corresponds to a warmer HYRAS-2015 dataset

that the same topographic detail will be captured by both datasets, whereas E-OBS is interpolated on a $0.1^\circ \times 0.1^\circ$ grid (approximately $10 \text{ km} \times 10 \text{ km}$). The bias is least pronounced in October (-0.02°C) and most pronounced in May (-0.05°C), being negative throughout the year. This means that the temperatures across the HYRAS grid are on average lower than across the E-OBS grid, possibly as a result of the different original grids used for the interpolation (see previous explanation). Conversely, the absolute differences between HYRAS and E-OBS are at a minimum in October (0.02°C) and reach their peak in May (0.05°C). Similar results were previously obtained by Frick et al. (2014). Between HYRAS

and the DWD monthly grids, the bias is negative throughout the year, being less pronounced in July (-0.01°C) than in January (-0.04°C), and the absolute differences are at their maximum in January (0.04°C) and their minimum in June (0.01°C). The greater differences in winter can be attributed to the differences in interpolation methods, and more specifically in the integration of elevation, as previously explained: as temperature inversions occur more frequently in winter but are not modelled in the DWD monthly datasets, this will lead to differences in winter. Another major difference in the interpolation methods is the consideration of UHI for HYRAS, which is not the case for the DWD monthly grids. As a result,

Fig. 7 Comparison of the long-term fields of daily mean temperature between HYRAS and E-OBS (a + b) and monthly mean temperature between HYRAS and DWD Monthly (c + d), over the complete time period from 1951 to 2015. The long-term (a + c) MAE and (b + d) bias were estimated for each grid cell. The HYRAS grid was scaled to the $25 \text{ km} \times 25 \text{ km}$ E-OBS grid before calculating the differences. The DWD Monthly grid was scaled to the $5 \text{ km} \times 5 \text{ km}$ HYRAS grid before calculating the differences

Fig. 8 Long-term (1951–2015) annual cycles of difference measures (red lines and circles, bias; blue lines and triangles, MAE) between HYRAS and E-OBS (solid lines) and between HYRAS and DWD Monthly (dashed lines), for the monthly (a) mean, (b) minimum, and (c) maximum temperatures, as averages over Germany

higher temperatures in densely populated regions can spread to the surroundings in the DWD monthly grids and generate warmer temperatures.

In comparison with the mean temperature, the differences in minimum temperature have a less pronounced annual cycle. Between HYRAS and E-OBS, the differences are minimized in September (bias: -0.02 °C; MAE: 0.03 °C). The bias is worst in May (-0.04 °C), while the absolute differences are the largest in January (0.04 °C). Between HYRAS and the DWD monthly grids, the differences are at their highest in August (bias: -0.05 °C; MAE: 0.05 °C) and at their lowest in December (bias: -0.04 °C; MAE: 0.04 °C). Overall, with little variation in the annual cycle of errors, HYRAS-TMIN agrees well with the corresponding E-OBS and with the DWD monthly grids.

The differences in maximum temperature have the most pronounced annual cycle. Between HYRAS and E-OBS, the differences are at their minimum in December (bias: -0.02 °C; MAE: 0.03 °C), whereas the differences greatly increase in spring and in summer, reaching their peak values in May (bias: -0.07 °C; MAE: 0.07 °C). The differences in grid resolution and in the underlying topography (see previous explanation) can also contribute to lower temperatures in HYRAS-TMAX. The greater absolute values of the maximum temperature, especially in the summer, can also in part explain the larger differences between the datasets. Between HYRAS and the DWD monthly grids however, the

differences are again lower than with E-OBS. The bias is negative from September to April, sinking to -0.03 °C in December and rising above 0 °C up to 0.01 °C in September. The MAE fluctuates between 0.01 from March to October and 0.03 °C in December. Overall, the datasets agree well with each other over Germany, with minor differences arising due to different grid resolutions and integration of elevation in the interpolation.

4.5 Comparison with previous HYRAS version (HYRAS-2006)

Finally, in order to assess the enhancement in the interpolation method, the HYRAS dataset was compared with its previous version (Version 1.01, Frick et al. (2014)), for mean temperature and relative humidity. As no previous dataset was available for minimum and maximum temperature, such a comparison is not possible. HYRAS-2006 covered a similar domain as HYRAS-2015, at the same spatial and temporal resolution over the period from 1951 to 2006.

When comparing the long-term (1951–2006) spatial differences between HYRAS-2015 and HYRAS-2006 (not shown here), the two datasets agree well with each other over most of the common domain. The differences were obtained by subtracting the previous version from the newer version, such that positive differences imply warmer temperatures or more humidity in the new version. Bias in mean temperature mainly

remains within ± 0.2 °C (on average 0.01 °C over Germany), with a few areas standing out with a bias up to ± 0.4 °C. A clear pattern is visible in the Swiss and Austrian Alps, where the temperatures from HYRAS-2015 are up to +1.0 °C warmer as in HYRAS-2006. This can be attributed to the interpolation method of HYRAS-2015 accounting for temperature inversions, which would yield warmer temperatures than HYRAS-2006 at higher elevations but also to the expansion of the HYRAS area in Version 3.0. Notably, more data in Austria was acquired such that the whole country is included in the interpolation of HYRAS-2015, resulting in boundary effects in that region. When averaged over a common domain (see Table 4), the bias is mostly positive, on average throughout the year 0.02 °C, only sinking under 0 °C in the winter (-0.01 °C).

Bias in relative humidity is on average -0.17 % across the common domain, being most pronounced in the spring (-0.24 %) and the fall (-0.26 %) (see Table 4). Across Germany, the bias generally remains within ± 1 %, varying between ± 1.5 % in some areas. As with the temperature, the larger differences can be found in the Alps. Along the alpine crest, HYRAS-2015 yields less relative humidity than HYRAS-2006, with deviations below -5 %. Deviations can notably stem from the differing derivations of relative humidity, using vapour pressure in HYRAS-2006 and dew point temperature in HYRAS-2015.

Figure 9a shows the long-term (1951–2006) annual cycles of the bias and the MAE, both for HYRAS-2015 and for HYRAS-2006, as averages over all stations. It should be noted that the ensemble of stations used for their respective interpolations and for this comparison is not completely identical. The bias generally increases throughout the year. As for the MAE, it does not vary greatly between HYRAS-2006 and HYRAS-2015. Whereas in HYRAS-2006, the MAE in December was 0.6 °C, it slightly increases to 0.63 °C in HYRAS-2015. From April to October however, the MAE improves for HYRAS-2015, with values decreasing from 0.64 to 0.58 °C in October, for example. As the error measures from HYRAS-2015 are generally very close to those of HYRAS-2006, it can in part be attributed to the expansion

of the HYRAS dataset, e.g. in the Austrian Alps, where greater deviations are more likely to occur.

Figure 9b is similar to Fig. 9a but for relative humidity. The bias on the other hand does not present a clear picture. From May to August, the bias does not change between HYRAS-2006 and HYRAS-2015. In all other months, the bias increases by 0.25 to 0.5 %. It is difficult to perform a comparison for the relative humidity, as relative humidity is spatially more heterogeneous than temperature. Nonetheless, clear improvements in the MAE are noticeable from HYRAS-2006 to HYRAS-2015. The MAE decreases by -0.75 to -1 % in all months.

Furthermore, the spatial analysis of the error measures in daily mean temperature according to the elevation difference between the station and the grid cell (see Section 4.1.2) was also performed by Frick et al. (2014) for HYRAS-2006. It can therefore be compared with the analysis for HYRAS-2015 (results for 1951–2015 shown in Fig. 5, results for 1951–2006 not shown here). Despite the inherent errors due to the elevation differences between stations and their nearest grid cell, the unadjusted bias (-0.1 °C for 0 to 100 m difference, 4.8 °C for -1000 to -900 m difference) and MAE (0.21 °C for 0 to 100 m difference, 4.96 °C for -1000 to -900 m difference) for HYRAS-2015 are both lower than for HYRAS-2006 (bias: -0.1 °C for 0 to 100 m difference, 5.2 °C for -1000 to -900 m difference; MAE: 0.3 °C for 0 to 100 m difference, 5.3 °C for -1000 to -900 m difference). This improvement can be attributed to the better representation of vertical structures in the interpolation of HYRAS-2015.

Frick et al. (2014) also conducted a fivefold cross-validation of the HYRAS-2006 mean temperature dataset, whose results can be compared with those of the cross-validation of HYRAS-2015 (results for 1951–2015 shown in Fig. 6a, results for 1951–2006 not illustrated here). As was the case for HYRAS-2006, the monthly bias of the cross-validation of HYRAS-2015 remains close to the bias of the original interpolation, albeit with a deviation of up to +0.1 °C, as previously described and explained in Section 4.3. The monthly MAE increases in comparison with the original interpolation, as was the case for HYRAS-2006. The increase in

Fig. 9 Comparison of the long-term (1951–2006) annual cycles of the respective bias (red lines and circles) and MAE (blue lines and triangles) between HYRAS-2015 (solid lines) and HYRAS-2006 (dashed lines), for (a) HYRAS-TAS and (b) HYRAS-HURS. The error scores are calculated as averages over all stations present in the respective domains of each HYRAS version

MAE is also larger, at approximately $+0.5\text{ }^{\circ}\text{C}$ each month, as opposed to $+0.3\text{ }^{\circ}\text{C}$ each month for HYRAS-2006. Conclusions regarding the quality of the cross-validation results and the datasets validated cannot be directly drawn from this comparison, as the stations used in the interpolation and the spatial coverages of the grids differ from each other. As was already discussed in Section 4.3, differences in station ensembles, and especially at higher elevations, can lead to non-negligible deviations. For example, the bias from the cross-validation of HYRAS-2015 decreases by up to $-0.2\text{ }^{\circ}\text{C}$ when excluding all stations located in Austria from the spatial average. This could explain differences with HYRAS-2006, where only the northwestern part of Austria was included in the HYRAS domain.

Overall, the HYRAS-2015 dataset provides improved representation of mean temperature and relative humidity in comparison with HYRAS-2006, especially in areas of complex topography. Although the spatial expansion of the HYRAS domain leads to some considerable differences between HYRAS-2015 and HYRAS-2006 (particularly at the edges of the domain), HYRAS-2015 provides a more complete representation of temperature and relative humidity across Germany and its relevant catchment areas.

5 Climatology

In addition to temperature and relative humidity, HYRAS-2015 also comprises a precipitation dataset (HYRAS-PR). This is an extension of the HYRAS-2006 version, presented in Rauthe et al. (2013). The same interpolation method used for HYRAS-2006 was applied to HYRAS-2015, with minor modifications in the derivation of the background field. Along with the temperature and relative humidity grids, these datasets allow us to perform different climatological analyses. Section 5.1 describes long-term observed trends based on the dataset, and Section 5.2 illustrates the use of HYRAS in analysing past extreme events.

5.1 Trends

With a dataset available over a period of 65 years, long-term climate trends can be assessed. Temperature trends in Germany have usually been assessed based on station observations (e.g. Łupikasza et al. (2011), Brien et al. (2013)), requiring stations that are spatially representative in order to derive meaningful trends valid for the study region. In this study, this is overcome by the use of the interpolated grids. It should however be noted that all available station observations were used in the regionalization of the different HYRAS variables. As a result, HYRAS is not solely based on homogeneous (or homogenized) station observations, as is also the case for other observational gridded datasets (e.g. E-OBS,

Haylock et al. (2008), and the other datasets listed in Table 1). Furthermore, the number of stations varies with time (see Fig. 1). Both can influence the trends; however, with averaging in time and space, the trends and their directions are very robust. As shown by Begert et al. (2005), the general conclusions regarding the direction of the trends in temperature and precipitation do not change with homogenization.

The long-term reference (1971–2000) annual temperature averaged over Germany is $8.6\text{ }^{\circ}\text{C}$. Figure 10a shows the yearly anomalies with respect to the reference mean over the period from 1951 to 2015, and the linear trend over that period is presented in Table 5. A clear trend in the average temperature can be seen, increasing significantly (at the 95 % confidence level) by $+1.6\text{ }^{\circ}\text{C}$ from 1951 to 2015. The moving average nevertheless shows a flatter increase in the first half of the period (1951 to around 1982) than in the second half. Similar trends based on stations were found in Klein Tank et al. (2002), for approximately the same period. During most years until 1990, the average annual temperature anomalies are negative and last between 3 and 6 consecutive years, occasionally being followed by 1 or 2 years with positive temperature anomalies under $+0.5\text{ }^{\circ}\text{C}$. However, starting in 1990, very few years with negative temperature anomalies (only 4) can be identified. The positive anomalies usually exceed $+0.5\text{ }^{\circ}\text{C}$. Eight of the 10 warmest years between 1951 and 2015 were recorded in the twenty-first century. On a seasonal basis, the increase in average temperature is most pronounced in the spring ($+1.7\text{ }^{\circ}\text{C}$) and least pronounced in the fall ($+0.95\text{ }^{\circ}\text{C}$), nevertheless being significant in all seasons.

Conversely, the average annual relative humidity tends to decrease from 1951 to 2015. As indicated in Table 5, this decrease occurs at a rate of $-0.2\text{ } \%/ \text{decade}$ (significant at the 95 % level). This is notably the case in the spring ($-0.3\text{ } \%/ \text{decade}$) and in the summer ($-0.5\text{ } \%/ \text{decade}$), while no significant trend can be detected either in winter or in the fall. This can be seen in Fig. 10b, which also shows the yearly anomalies with respect to the reference mean of $78.6\text{ } \%$. Most years from 1951 to 1988 are more humid than the reference average, with individual years (in total, 6 years) standing out for being less humid than average. From 1989 to 1994, a series of relatively dry years occurred, with relative humidity of -1.5 to $-2\text{ } \%$ under the reference average. Since 1995, most of the years have been more humid than average, with the positive anomalies generally increasing. Nevertheless, the year 2003 stands out for being the driest year in the period, with the relative humidity sinking to $-3.5\text{ } \%$ under the reference average. This extreme year and more specifically its extreme summer and heat wave are examined more closely in Section 5.2.1.

Unlike the temperature and the relative humidity, the deviation in total annual precipitation with respect to the reference average of 774 mm is highly variable in time, with relatively wetter and drier years alternating with each other (see Fig. 10c). The longest succession of relatively wet years was from 1998 to

Fig. 10 Yearly anomalies with respect to the long-term reference (1971–2000) mean, averaged over Germany for (a) temperature, (b) relative humidity, and (c) precipitation, based on their corresponding HYRAS-2015 dataset. A linear trend was fit to the time series (dashed lines), and an 11-year moving window average was calculated and is shown as well (solid black line)

2002, which was immediately followed by one of the driest years recorded in 2003 (less than -20% annual precipitation compared with the reference average). The longest succession of relatively dry years consisted of 3 years between 1962 and 1964 and between 1971 and 1973. Overall, trends in precipitation are less pronounced than for temperature and relative humidity, and not significant (Table 5). The largest increase over the whole time series is observed in winter ($+26\%$). The lack of significance in contrast to other analyses (such as in Rauthe et al. (2013)) can be explained due to the large regional variability of the precipitation and therefore the large regional variability of the trends. Furthermore, precipitation is highly variable in time, which can significantly influence the results of the linear trend calculation when the length of the times series slightly differs. It should nonetheless be mentioned that some significant trends can be identified for indices of extreme precipitation, based on other analyses within the project “Network of Experts”. Extreme precipitation is however not the focus of this study.

An advantage of using daily high-resolution grids is the possibility of assessing long-term climate indices. Such indices are defined, for example, by the WMO or within the CLIMDEX project (CLIMDEX 2018). In the “Network of

Experts”, over 30 indices were calculated based on the HYRAS temperature and relative humidity datasets. A selection of these indices is presented here, and their trends are evaluated. For example, summer days (days when the maximum temperature is equal to or exceeds $25\text{ }^{\circ}\text{C}$) are good indicators of hot extremes. From 1951 to 2015, the average of summer days across Germany has varied between 15 to 63 days, reaching its maximum during the unprecedented summer of 2003. Overall, the number of summer days per year has increased significantly by $+18$ days from 1951 to 2015 (or $+2.7$ days per decade), as indicated in Table 5. Similarly, hot days (days when the maximum temperature is equal to or exceeds $30\text{ }^{\circ}\text{C}$) tend to increase during that same period, at a rate of almost $+1$ day per decade. Conversely, frost days (days when the minimum temperature is below $0\text{ }^{\circ}\text{C}$) are good indicators of cold extremes. In the same way that summer days tend to increase, frost days tend to significantly decrease from 1951 to 2015, at a rate of -3.2 days per decade (on average in Germany).

Moreover, the anomalies in mean temperature, relative humidity, and precipitation can be analysed together. For example, Fig. 11 considers the anomalies in the summer season, as

Table 5 Reference period (1971–2000) average over Germany and absolute trends for different variables and indices

Variable	Time period	1971–2000 average over Germany	Average decadal trend	Absolute trend from 1951 to 2015
tasMean	Winter	0.8	<i><u>0.27</u></i>	<i><u>1.77</u></i>
Daily mean temperature [°C]	Spring	8.1	<i><u>0.29</u></i>	<i><u>1.89</u></i>
	Summer	16.6	<i><u>0.26</u></i>	<i><u>1.68</u></i>
	Fall	8.7	<i><u>0.15</u></i>	<i><u>0.95</u></i>
	Annual	8.6	<i><u>0.24</u></i>	<i><u>1.57</u></i>
tasminMean	Winter	− 1.87	<i><u>0.28</u></i>	<i><u>1.82</u></i>
Daily minimum temperature [°C]	Spring	3.61	<i><u>0.21</u></i>	<i><u>1.34</u></i>
	Summer	11.6	<i><u>0.17</u></i>	<i><u>1.11</u></i>
	Fall	5.27	<i><u>0.13</u></i>	<i><u>0.88</u></i>
	Annual	4.7	<i><u>0.02</u></i>	<i><u>0.2</u></i>
tasmaxMean	Winter	3.53	<i><u>0.03</u></i>	<i><u>0.28</u></i>
Daily maximum temperature [°C]	Spring	12.9	<i><u>0.03</u></i>	<i><u>0.34</u></i>
	Summer	21.9	<i><u>0.03</u></i>	<i><u>0.32</u></i>
	Fall	12.7	<i><u>0.01</u></i>	<i><u>0.12</u></i>
	Annual	12.8	<i><u>0.03</u></i>	<i><u>0.27</u></i>
hurs	Winter	84.7	− 0.04	− 0.28
Daily mean relative humidity [%]	Spring	74.1	− <i><u>0.32</u></i>	− <i><u>2.09</u></i>
	Summer	73.1	− <i><u>0.52</u></i>	− <i><u>3.38</u></i>
	Fall	82.7	0.13	0.83
	Annual	78.6	− <i><u>0.19</u></i>	− <i><u>1.26</u></i>
rMean	Winter	181	4.06	26.4 (16 %)
Average precipitation total [mm]	Spring	176	0.86	5.57 (3.3 %)
	Summer	231	− 2.55	− 16.6 (− 6.6 %)
	Fall	188	3.66	23.8 (13.8 %)
	Annual	774	5.25	34.1 (4.5 %)
su	Annual	29.5	<i><u>2.69</u></i>	<i><u>17.5</u></i>
Summer days				
Days where maximum temperature ≥ 25 °C				
[Number of days]				
hd	Annual	4.6	<i><u>0.93</u></i>	<i><u>6.07</u></i>
Hot days				
Days where maximum temperature ≥ 30 °C				
[Number of days]				
fd	Annual	86.0	− <i><u>3.16</u></i>	− <i><u>20.5</u></i>
Frost days				
Days where minimum temperature < 0 °C				
[Number of days]				

Trends significant at the 95 % level are italicized and underlined; trends significant at the 90 % level are italicized. For the precipitation totals, the relative trends (percentages) are also indicated in parentheses

averages over Germany. First, as was the case for the annual temperatures, there is a tendency for warmer temperatures in the later decades. Considering the relative humidity, these warmer years also correspond to drier years. The total annual precipitation however does not tend to increase or decrease as a function of the temperature, as both positive and negative anomalies are found in later decades. Extreme years stand out clearly, such as the summer of 2003 and the summer of 2015, which are both located in the lower-right quadrant, signifying higher than average temperatures and lower than average

precipitation, along with dry relative humidity conditions compared with the average. This combination of high temperatures and dryness contributed to extreme conditions, which are evaluated closely in Section 5.2.

5.2 Extreme events

Daily high-resolution grids of meteorological variables are useful for a variety of applications, one of them being the possibility of reproducing and analysing the characteristics

Fig. 11 Summer (June, July, August) temperature anomalies with respect to the long-term reference (1971–2000) mean, averaged over Germany for temperature, precipitation (plotted against each other) and relative humidity (size of the dots), based on their corresponding HYRAS-2015 dataset. The colours indicate different decades

of historical extreme events. Here, we focus on the heat wave of the summer 2003 across Europe and the heat wave of the summer 2015 in south Germany. The former was previously assessed in Frick et al. (2014) and will be re-evaluated based on the new version of the HYRAS dataset and different climate indices based on new variables available. The latter provides an example of an extreme event occurring in the extended period of the HYRAS dataset. Ionita et al. (2017) performed a similar Europe-wide analysis based on the E-OBS dataset (Haylock et al. 2008); however, the results presented here are restricted to Germany.

5.2.1 Heat wave of 2003

In the summer of 2003, a heat wave struck large parts of Europe, leading to record temperatures being measured. This heat wave corresponded to a large-scale anomaly, resulting from a northwards shift of the Atlantic subtropical high-pressure system (Ionita et al. 2017), remaining stationary above Europe.

The climate characteristics of this extreme event can also be reproduced by the HYRAS dataset. During that summer, the average daily mean temperature highly deviated by + 3 °C from its reference average (1971–2000 average over Germany: 16.6 °C). Figure 12a shows that the deviations varied spatially between + 2 and + 5 °C. Larger deviations are mainly noticeable in southern Germany as was already concluded by Schönwiese et al. (2004) based on station data. The number of summer days (not shown here) was above average, with up to + 40 summer days in southern Germany. Furthermore, an increase in the number of hot days compared with the reference average is observable across the

whole country, exceeding + 20 days in southern Germany (Fig. 12b). In that summer, higher than average radiation between + 4 at the coast and + 20 % in southern Germany was recorded (not shown here), based on the HYRAS dataset for global radiation (unpublished). While positive anomalies for average and maximum temperature were observed everywhere, tropical nights (on days where the minimum temperature exceeds 20 °C) were only recorded in specific areas, namely, the Rhine-Main area and the Saar-Nahe Hills. To further exasperate this heat wave, the relative humidity (Fig. 12c) also highly decreased by – 6.5 % on average in comparison with its reference value (average across Germany in the summer: 73.1 %), with values reaching – 10 to – 12 % across most of Germany, leading to very dry conditions. During that summer, the total precipitation was lower than average across most of Germany, up to – 50 % in some areas (Fig. 12d). As a result of this heat wave and the resulting dry conditions, navigable waterways had extremely low water levels, rendering them not practicable (Beck et al. 2004; BfG 2006; Jonkeren et al. 2007).

5.2.2 Heat wave of 2015

As the time period covered by the HYRAS datasets was extended, more recent heat waves such as in the summer of 2015 can also be analysed (Hoy et al. 2017). Indeed, that summer was the third warmest summer in Germany since 1881 (Becker et al. 2015) and the second warmest in the HYRAS period of 1951 to 2015. Although this heat wave was not as intense and widespread as in the summer of 2003 (Ionita et al. 2017), important deviations from reference averages could again be observed (Muthers et al. 2017).

Fig. 12 Deviations from reference (1971–2000) average during the summer of 2003 (a–d) and 2015 (e–h). (a + e) Average temperature, (b + f) number of hot days, (c + g) relative humidity, and (d + h) precipitation

The summer of 2015 was a highly variable summer. Hot air inflows from Africa were alternating with cooler air inflows from the North Atlantic (Becker et al. 2015), but a stagnant high-pressure system also contributed to increased radiative forcing (Ionita et al. 2017). Due to the northwards inflow of hot air, southern Germany was particularly affected by that heat wave. This can be clearly seen in Fig. 12e, where the average temperature exceeded its reference average by + 2 to + 4 °C. Accordingly, the number of summer days (not shown here) and the number of hot days (Fig. 12f) exceed their reference average by up to + 30 days in some regions during that summer. The relative humidity anomaly was heterogeneous during that summer, as shown in Fig. 12g. The Saar-Nahe Hills and parts of Bavaria were particularly dry, with – 10 % relative humidity in comparison with the reference average. While the precipitation in the northern lowlands tended to exceed its reference average, very dry conditions occurred again in southern Germany, with precipitation deficits reaching – 50 % (Fig. 12h). Although not as intense as the heat wave in 2003, this extreme event still had negative consequences on transportation systems, leading again to low water levels (Belz et al. 2015a; Belz et al. 2015b; Van Lanen et al. 2016) and also damaging road infrastructure (UNECE 2019).

6 Conclusion

This study presents daily high-resolution (5 km × 5 km) grids of mean, minimum, and maximum temperature and relative

humidity for Germany and its catchment areas over the period from 1951 to 2015. These observational datasets were developed and used within the German Federal Ministry of Transport and Digital Infrastructure's project "Network of Experts", with the greater aim at understanding the effects of climate change on transport infrastructure and adapting to it. Along with other meteorological variables (precipitation, global radiation), the HYRAS observational datasets are used within the "Network of Experts" project for hydrological modelling and for the bias correction of climate models, to cite some examples of possible applications.

The HYRAS-TAS and HYRAS-HURS datasets are updated and expanded versions of the datasets presented in Frick et al. (2014), interpolated using a different regionalization method considering temperature inversions, UHI effect, locations prone to cold air ponding, and maritime effects. This new method furthermore enabled the interpolation of minimum and maximum temperature datasets. These observational datasets are based upon measurements gathered from the Germany's National Meteorological Service and neighbouring countries' (hydro-)meteorological authorities. The station measurements undergo quality control, after which they are gridded using the method presented by Krähenmann et al. (2018), based on Frei (2014).

In order to assess the quality of the gridded datasets, the interpolated grids are directly compared with the nearest station measurements. The results are generally in good agreement with the measurements, with difficulties arising in areas of more complex topography. A leave-one-out cross-validation indicates that

the regionalization method is unbiased when stations are removed from the interpolation. Moreover, fivefold cross-validation of HYRAS-TAS and HYRAS-HURS are performed, allowing an evaluation of the interpolation method against an independent dataset, along with an evaluation of the impact of a reduced station dataset. The resulting bias in temperature is close to the bias of the original interpolation, indicating that the interpolation method is generally unbiased and that its quality is not exclusively dependent on the original station database used for the regionalization, especially in areas of low topographic complexity. The bias in relative humidity however increases, due to the greater uncertainty generated by the removal of stations. This also leads to larger absolute errors. The HYRAS datasets are also compared with other observational datasets of daily mean, minimum, and maximum temperatures, namely, the E-OBS-gridded dataset (Haylock et al. 2008) and the DWD monthly dataset. The former shows the advantages of interpolating at a higher spatial resolution, while the latter highlights the variations resulting from different approaches in integrating elevation and effects such as UHI into the interpolation. Finally, the HYRAS-2015 temperature and relative humidity datasets were compared with their respective previous versions (HYRAS-2006), to estimate the improvement in the interpolation methods. The two datasets agree well with each other over most of the common domain, with boundary effects occurring due to the expansion of the HYRAS domain in HYRAS-2015. The error measures from the station-grid comparison and from the cross-validation tend to be higher for HYRAS-2015, with non-negligible effects due to deviations, e.g. in Austria (which was only partially covered by HYRAS-2006).

Long-term observational gridded datasets are useful for analyses of past climate. Along with the presented temperature and relative humidity grids, the HYRAS precipitation dataset is used to quantify long-term climate trends. These reveal significant increases in temperatures in all seasons and decreases in relative humidity, mainly in the spring and in the summer. However, no significant trend in mean precipitation was identified. High-resolution daily-gridded datasets can also serve as the basis for calculating climate indices and study the occurrence of climate extremes. As examples for extreme events, the heat waves of the summer of 2003 and 2015 are evaluated in this study, whereby the HYRAS datasets show important deviations from reference conditions during these events.

The datasets presented here satisfy some of the required improvements named by Frick et al. (2014): the HYRAS-TAS and HYRAS-HURS datasets were prolonged and expanded, and daily high-resolution datasets of minimum and maximum temperature were produced. Further developments of the HYRAS datasets include the finalization of the expanded dataset covering the Oder river catchment area, based on recently acquired station data for Poland. Additionally, the improvement of the (yet unpublished) solar radiation daily-gridded dataset is desired. Eventually, the development of a HYRAS dataset at a subdaily

resolution could be considered. Finally, the temporal prolongation of all HYRAS datasets is planned, in order to expand the observation period to include recent years.

Acknowledgements The work presented here was conducted within the “Network of Experts” (*Expertennetzwerk*) project, funded by the German Federal Ministry of Transport and Digital Infrastructure (*Bundesministerium für Verkehr und digitale Infrastruktur*, BMVI). The authors thank all project partners and other users of HYRAS for their feedback on the dataset. Additionally, the authors acknowledge and thank all data providers: the Zentralanstalt für Meteorologie und Geodynamik, the Royal Meteorological Institute of Belgium, Météo France, the Luxembourg Institute of Science and Technology/MeteoLux, the Royal Netherlands Meteorological Institute, and MeteoSwiss. The authors furthermore acknowledge the use of the E-OBS dataset from the EU-FP6 project UERRA (<http://www.uerra.eu>) and the Copernicus Climate Change Service and the data providers in the ECA&D project (<https://www.ecad.eu>). The authors acknowledge that the interpolation procedure described here was implemented in R (R Development Core Team 2017). Finally, the authors are also grateful to the anonymous reviewer, whose comments helped to improve the manuscript.

Funding Open Access funding enabled and organized by Projekt DEAL. The authors received funding from the German Federal Ministry of Transport and Digital Infrastructure (*Bundesministerium für Verkehr und digitale Infrastruktur*, BMVI) as part of the “Network of Experts” (*Expertennetzwerk*) project.

Appendix

The following formulas were applied in this study to evaluate the HYRAS datasets.

Bias:

$$\text{Bias} = \frac{1}{n} \sum_{i=1}^n x_i - y_i$$

Mean absolute error, MAE:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |x_i - y_i|$$

x_i : station measurement (or HYRAS-2015 grid cell value)

y_i : grid value at the nearest grid cell (or grid cell value of other dataset)

n : number of pairs of values

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain

permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Aalto J, Pirinen P, Jylhä K (2016) New gridded daily climatology of Finland: permutation-based uncertainty estimates and temporal trends in climate. *J Geophys Res Atmos* 121:3807–3823. <https://doi.org/10.1002/2015JD024651>
- Beck C, Grieser J, Trömel S (2004) Die Trockenperiode des Jahres 2003 in Deutschland im Kontext langzeitlicher Niederschlagsvariabilität. Deutscher Wetterdienst (DWD), Offenbach am Main, Germany
- Becker P, Imbery F, Friedrich K, Rauthe M, Matzarakis A, Grätz A, Janssen W (2015) Klimatologische Einschätzung des Sommer 2015. Deutscher Wetterdienst (DWD), Offenbach am Main, Germany
- Begert M, Schlegel T, Kirchhofer W (2005) Homogeneous temperature and precipitation series of Switzerland from 1864 to 2000. *Int J Climatol* 25:65–80
- Belz JU, Mürlebach M, Krahe P, Rademacher S (2015a) Weithin fortgesetzt fallende Wasserstände. Bundesanstalt für Gewässerkunde (BfG), Koblenz, Germany
- Belz JU, Mürlebach M, Wiechmann W, Krahe P, Rademacher S (2015b) Ausweitung der Niedrigwassersituation. Bundesanstalt für Gewässerkunde (BfG), Koblenz, Germany
- Berezowski T, Szcześniak M, Kardel I, Michałowski R, Okruszko T, Mezghani A, Piniewski M (2016) CPLFD-GDPT5: high-resolution gridded daily precipitation and temperature data set for two largest Polish river basins *Earth Syst Sci Data* 8:127–139. <https://doi.org/10.5194/essd-8-127-2016>
- BfG (2006) Niedrigwasserperiode 2003 in Deutschland. Ursache - Wirkungen - Folgen. Bundesanstalt für Gewässerkunde (BfG), Koblenz, Germany
- Brienen S, Kapala A, Mächel H, Simmer C (2013) Regional centennial precipitation variability over Germany from extended observation records. *Int J Climatol* 33:2167–2184. <https://doi.org/10.1002/joc.3581>
- Caesar J, Alexander L, Vose R (2006) Large-scale changes in observed daily maximum and minimum temperatures: creation and analysis of a new gridded data set *J Geophys Res Atmos* 111 doi:<https://doi.org/10.1029/2005JD006280>
- CLIMDEX (2018) Indices. <https://www.climdex.org/learn/indices/>. Accessed 16.09.2019
- Cornes RC, van der Schrier G, van den Besselaar EJM, Jones PD (2018) An ensemble version of the E-OBS temperature and precipitation data sets. *J Geophys Res Atmos* 123:9391–9409. <https://doi.org/10.1029/2017JD028200>
- Daly C (2006) Guidelines for assessing the suitability of spatial climate data sets. *Int J Climatol* 26:707–721. <https://doi.org/10.1002/joc.1322>
- Della-Marta PM, Wanner H (2006) A method for homogenizing the extremes and mean of daily temperature measurements. *J Climate* 19:4179–4197
- Delvaux C, Journée M, Bertrand C (2015) The FORBIO Climate data set for climate analyses. *Adv Sci Res* 12:103–109. <https://doi.org/10.5194/asr-12-103-2015>
- DWD Climate Data Center (CDC) (2018) Raster der Monatsmittel der Lufttemperatur (2 m) für Deutschland, Version 1.0 edn. DWD Climate Data Center (CDC), Offenbach am Main, Germany
- Eiselt K-U, Kaspar F, Mölg T, Krähenmann S, Posada R, Riede J (2017) Evaluation of gridding procedures for air temperature over Southern Africa. *Adv Sci Res* 14:163–173. <https://doi.org/10.5194/asr-14-163-2017>
- Frei C (2014) Interpolation of temperature in a mountainous region using nonlinear profiles and non-Euclidean distances. *Int J Climatol* 34:1585–1605. <https://doi.org/10.1002/joc.3786>
- 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* 23:15–32. <https://doi.org/10.1127/0941-2948/2014/0560>
- Gandin LS (1966) Objective analysis of meteorological fields. By L. S. Gandin. Translated from the Russian. Jerusalem (Israel Program for Scientific Translations), 1965. Pp. vi, 242: 53 Figures; 28 Tables. £4 1 s. *QJR Meteorol Soc* 92:447–447. <https://doi.org/10.1002/qj.49709239320>
- Guijarro JA, López JA, Aguilar E, Domonkos P, Venema VKC, Sigró J, Brunet M (2017) Comparison of homogenization packages applied to monthly series of temperature and precipitation: the MULTITEST project. Ninth Seminar for Homogenization and Quality Control in Climatological Databases and Fourth Conference on Spatial Interpolation Techniques in Climatology and Meteorology (Budapest, Hungary, 03 – 07 April 2017), Geneva: WMO, Climate Data and Monitoring WCDMP-No. 85: 46–62 pp https://repositorio.aemet.es/bitstream/20.500.11765/8392/1/Comparison_GuijarroWCDMP_85.pdf
- Gusyeve M, Gädeke A, Cullmann J, Magome J, Sugiura A, Sawano H, Takeuchi K (2016) Connecting global- and local-scale flood risk assessment: a case study of the Rhine River basin flood hazard. *J Flood Risk Manag* 9:343–354. <https://doi.org/10.1111/jfr3.12243>
- Haylock M, Hofstra N, Klein Tank A, Klok L, Jones P, New M (2008) A European daily high-resolution gridded dataset of surface temperature and precipitation for 1950–2006. *J Geophys Res Atmos* 113: 113. <https://doi.org/10.1029/2008JD010201>
- Heininger P, Cullmann J (2015) Sediment matters. Springer International Publishing, Switzerland. <https://doi.org/10.1007/978-3-319-14696-6>
- Hiebl J, Frei C (2016) Daily temperature grids for Austria since 1961—concept, creation and applicability. *Theor Appl Climatol* 124:161–178. <https://doi.org/10.1007/s00704-015-1411-4>
- Hofstra N, Haylock M, New M, Jones P, Frei C (2008) Comparison of six methods for the interpolation of daily. *European Climate Data J Geophys Res Atmos* 113. <https://doi.org/10.1029/2008jd010100>
- Höllering S, Wienhöfer J, Ihringer J, Samaniego L, Zehe E (2018) Regional analysis of parameter sensitivity for simulation of streamflow and hydrological fingerprints. *Hydrol Earth Syst Sci* 22:203–220. <https://doi.org/10.5194/hess-22-203-2018>
- Hoy A, Hänsel S, Skalak P, Ustrnul Z, Bochníček O (2017) The extreme European summer of 2015 in a long-term perspective. *Int J Climatol* 37:943–962. <https://doi.org/10.1002/joc.4751>
- Imbery F, Plagemann S, Namyslo J (2013) Processing and analysing an ensemble of climate projections for the joint research project KLIWAS. *Adv Sci Res* 10:91–98. <https://doi.org/10.5194/asr-10-91-2013>
- Ionita M et al (2017) The European 2015 drought from a climatological perspective. *Hydrol Earth Syst Sci* 21:1397–1419. <https://doi.org/10.5194/hess-21-1397-2017>
- Jonkeren O, Rietveld P, Ommeren J (2007) Climate change and inland waterway transport: welfare effects of low water levels on the river. *Rhine J Transport Econ Pol* 41:387–411
- Keil M, Bock M, Esch T, Metz A, Nieland S, Pfizner A (2011) CORINE Land Cover 2006 - Europaweit harmonisierte Aktualisierung der Landbedeckungsdaten für Deutschland. Deutsches Zentrum für Luft- und Raumfahrt e.V., Deutsches Fernerkundungsdatenzentrum, Weßling, Germany
- Klein Tank AMG et al (2002) Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate

- Assessment. *Int J Climatol* 22:1441–1453. <https://doi.org/10.1002/joc.773>
- Krähenmann S, Bissolli P, Rapp J, Ahrens B (2011) Spatial gridding of daily maximum and minimum temperatures in Europe. *Meteorog Atmos Phys* 114:151–161. <https://doi.org/10.1007/s00703-011-0160-x>
- Krähenmann S, Haller M, Walter A (2020) Bias adjustment and statistical downscaling making use of PCA-derived predictors. *Meteorol Z*. Manuscript submitted for publication.
- Krähenmann S, Walter A, Brien S, Imbery F, Matzarakis A (2018) High-resolution grids of hourly meteorological variables for Germany. *Theor Appl Climatol* 131:899–926. <https://doi.org/10.1007/s00704-016-2003-7>
- Lupikasza EB, Hänsel S, Matschullat J (2011) Regional and seasonal variability of extreme precipitation trends in southern Poland and central-eastern Germany 1951–2006. *Int J Climatol* 31:2249–2271. <https://doi.org/10.1002/joc.2229>
- Magnus G (1844) Versuche über die Spannkraft des Wasserdampfes. *Ann Phys Chem*
- Mestre O, Gruber C, Prieur C, Caussinus H, Jourdain S (2011) SPLIDHOM: a method for homogenization of daily temperature observations. *J Appl Meteorol Climatol* 50:2343–2358. <https://doi.org/10.1175/2011JAMC2641.1>
- Meyer J, Kohn I, Stahl K, Hakala K, Seibert J, Cannon AJ (2019) Effects of univariate and multivariate bias correction on hydrological impact projections in alpine catchments. *Hydrol Earth Syst Sci* 23:1339–1354. <https://doi.org/10.5194/hess-23-1339-2019>
- Müller-Westermeier G (1995) Numerische Verfahren zur Erstellung klimatologischer Karten. Deutscher Wetterdienst (DWD), Offenbach am Main, Germany
- Muthers S, Laschewski G, Matzarakis A (2017) The summers 2003 and 2015 in South-West Germany: heat waves and heat-related mortality in the context of climate change. *Atmos* 8. <https://doi.org/10.3390/atmos8110224>
- Ntegeka V, Salamon P, Gomes G, Sint H, Lorini V, Zambrano-Bigiarini M, Thielen J (2013) EFAS-Meteo A European daily high-resolution gridded meteorological data set for 1990–2011. <https://doi.org/10.2788/51262>
- Peixoto JP, Oort AH (1996) The climatology of relative humidity in the atmosphere. *J Clim* 9:3443–3463. [https://doi.org/10.1175/1520-0442\(1996\)009<3443:Tcorhi>2.0.Co;2](https://doi.org/10.1175/1520-0442(1996)009<3443:Tcorhi>2.0.Co;2)
- Piper SC, Stewart EF (1996) A gridded global data set of daily temperature and precipitation for terrestrial biospheric modeling. *Global Biogeochem Cycles* 10:757–782. <https://doi.org/10.1029/96gb01894>
- R Development Core Team (2017) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria
- Rauthe M, Steiner H, Riediger U, Mazurkiewicz A, Gratzki A (2013) A Central European precipitation climatology - Part I: generation and validation of a high-resolution gridded daily data set (HYRAS). *Meteorol Z* 22:235–256. <https://doi.org/10.1127/0941-2948/2013/0436>
- Schönwiese CD, Staeger T, Trömel S (2004) The hot summer 2003 in Germany. Some preliminary results of a statistical time series analysis. *Meteorol Z* 13:323–327. <https://doi.org/10.1127/0941-2948/2004/0013-0323>
- Schulzweida U, Kornblüeh L, Quast R (2018) CDO user's guide. Climate data operators, Version 1.9.3.
- Stahl K, Dan Moore RD, Floyer JA, Asplin M, McKendry I (2006) Comparison of approaches for spatial interpolation of daily air temperature in a large region with complex topography and highly variable station density. *Agric For Meteorol* 139:139–236. <https://doi.org/10.1016/j.agrformet.2006.07.004>
- Stahl K et al. (2017) The snow and glacier melt components of the streamflow of the River Rhine and its tributaries considering the influence of climate change. Final report to the International Commission for the Hydrology of the Rhine basin (CHR). English version, March 2017.
- UNECE (2019) Report of the group of experts on climate change impacts and adaptation for transport networks and nodes. Informal document WP.5 (2019) No. 6. United Nation Economic Commission for Europe (UNECE), Working Party on Transport Trends and Economics (WP.5), Geneva, Switzerland
- USGS (1993) Digital elevation models: data user guides 5. - Technical Report. U. S. Geological Survey (USGS), Reston, Virginia, USA
- Van Lanen HAJ et al (2016) Hydrology needed to manage droughts: the 2015 European case. *Hydrol Process* 30:3097–3104. <https://doi.org/10.1002/hyp.10838>
- Venema VKC, Mestre O, Aguilar E, Auer I, Guijarro JA, Domonkos P, Vertacnik G, Szentimrey T, Stepanek P, Zahradnické P, Viarre J, Müller-Westermeier G, Lakatos M, Williams CN, Menne MJ, Lindau R, Rasol D, Rustemeier E, Kolokythas K, Marinova T, Andresen L, Acquafotta F, Fratianni S, Cheval S, Klancar M, Brunetti M, Gruber C, Prohom Duran M, Likso T, Esteban B, Brandsma T (2012) Benchmarking homogenization algorithms for monthly data. *Clim Past* 8:89–115
- Vicente-Serrano S, Saz M, Cuadrat J (2003) Comparative analysis of interpolation methods in the middle Ebro Valley (Spain): application to annual precipitation and temperature. *Clim Res* 24:161–180. <https://doi.org/10.3354/cr024161>
- Werner AT, Schnorbus MA, Shrestha RR, Cannon AJ, Zwiers FW, Dayon G, Anslow F (2019) A long-term, temporally consistent, gridded daily meteorological dataset for northwestern North America. *Sci Data* 6:180299. <https://doi.org/10.1038/sdata.2018.299>
- Wienert U, Kreienkamp F, Spekat A, Enke W (2013) A simple method to estimate the urban heat island intensity in data sets used for the simulation of the thermal behaviour of buildings. *Meteorol Z* 22: 179–185. <https://doi.org/10.1127/0941-2948/2013/0397>
- Willkofer F, Schmid F-J, Komischke H, Korck J, Braun M, Ludwig R (2018) The impact of bias correcting regional climate model results on hydrological indicators for Bavarian catchments. *J Hydrol Reg Stud* 19:25–41. <https://doi.org/10.1016/j.ejrh.2018.06.010>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.