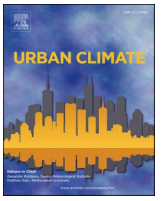
Urban Climate 58 (2024) 102226

Contents lists available at ScienceDirect

Urban Climate

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



# Filling gaps in urban temperature observations by debiasing ERA5 reanalysis data

Amber Jacobs [a,*], Sara Top [a], Thomas Vergauwen [a,b], Juuso Suomi [c], Jukka Käyhkö [c], Steven Caluwaerts [a,b]

[a] Department of Physics and Astronomy, Ghent University, Krijgslaan 281, 9000 Gent, Belgium
[b] Royal Meteorological Institute, Brussels, Belgium
[c] Department of Geography and Geology, University of Turku, FI-20014 Turun yliopisto, Finland

ARTICLE INFO

*Keywords:*
Urban climate networks
Urban heat island
Gap filling
Bias-adjusted time series
ERA5 reanalysis

ABSTRACT

Gaps in urban meteorological time series complicate the analysis and usage of datasets. Various gap-filling techniques exist, including the debiasing of ERA5 reanalysis data. Unfortunately, an extensive evaluation of these debiasing techniques is lacking for urban datasets. This research compares five gap-filling techniques for urban temperature time series, including three debiasing techniques that employ a learning period and time window to take into account the seasonal and diurnal ERA5 temperature bias. The evaluation, performed by filling manually constructed gaps, reveals that short gaps are best filled by linear interpolation, while longer gaps benefit from ERA5 debiasing. The bias correction is crucial for urban locations, with all debiasing techniques performing similarly. The exact length and placement of the learning period and time window have limited impact on the performance, however a symmetrical placement of the learning period with a minimum length of 10 days and a small time window provide the best outcome. Based on these results, a gap-filling algorithm is designed which efficiently fills all gaps in temperature time series by selecting the most optimal technique for each gap. The algorithm can reproduce the urban heat island effect, although a small over- or underestimation might occur.

## 1. Introduction

A common phenomenon in meteorology is the occurrence of gaps in observational time series (Diouf and Dème, 2022). A gap in a time series is defined as a sequence of missing values, for which different origins exist (Henn et al., 2013; Yozgatligil et al., 2013), e.g. due to communication problems, the malfunctioning of the measuring device or the removal of values by quality control. The length of a gap is described by the number of missing values or the length of the missing time period, and can vary from one missing data point to several months or even years. Gaps occur in a wide variety of meteorological datasets, although the susceptibility of a dataset to contain gaps depends on its origin. For instance, crowdsourced data holds a greater possibility of gaps compared to synoptic datasets, mainly due to their difference in data quality (Coney et al., 2022). The occurrence of gaps complicates the analysis and further use of

*Abbreviations:* GF, gap-filling; LI, linear interpolation; LR, linear regression; MB, mean bias; WMB, weighted mean bias; DTR, daily temperature range; MSE, mean squared error; UHI, urban heat island; LCZ, local climate zone.

* Corresponding author at: Krijgslaan 281/S9, 9000 Gent, Belgium.
*E-mail addresses:* amber.jacobs@ugent.be (A. Jacobs), sara.top@ugent.be (S. Top), thomas.vergauwen@meteo.be (T. Vergauwen), juuso.suomi@utu.fi (J. Suomi), jukka.kayhko@utu.fi (J. Käyhkö), steven.caluwaerts@ugent.be (S. Caluwaerts).

https://doi.org/10.1016/j.uclim.2024.102226
Received 17 June 2024; Received in revised form 19 August 2024; Accepted 24 November 2024
Available online 5 December 2024

the data. For example, excluding data due to gaps during analysis might lead to bias if these gaps are present in a systematic way (Aieb et al., 2019; Kang, 2013). In addition, a gap can prevent the calculation of certain meteorological values e.g. annual heat stress (Top et al., 2020). Furthermore, a complete observational dataset may be required for climatology studies (Afrifa-Yamoah et al., 2020; Yozgatligil et al., 2013) or as input for another application e.g. forcing data for a numerical model (Garen, 2013).

In urban climate research, the identification and analysis of data gaps is particularly important. For instance, the examination of the intensity of the urban heat island (UHI) phenomenon (Arnfield, 2003; Oke, 1982) necessitates the combination of multiple observational time series that span the same time period. First of all, only temperature measurements that are simultaneously present for both urban and rural location can contribute to the examination of the UHI, since the UHI intensity is defined as their difference. Secondly, when comparing the UHI intensity of different urban locations, it is recommended to only include measurements spanning the same time period to avoid a biased outcome and incorrect interpretations. Faulty conclusions might for example be drawn if there is a mismatch in (anti)cyclonic weather conditions between the measurements of two urban locations, since the UHI intensity depends on weather type (Ivajnšič and Žiberna, 2019). The condition to only utilize simultaneous data leads to the ignorance of useful data during time periods when a gap is present in another time series.

Since gaps induce problems, a variety of gap-filling (GF) techniques have been proposed to complete gapped time series. These techniques typically utilize available observations or model datasets to estimate the missing values, and are clustered into three categories based on their reconstruction technique: temporal, spatial and spatiotemporal GF techniques (Henn et al., 2013). Temporal GF techniques rely on the autocorrelation of the time series and are built on observations before and/or after the gap. Common examples are linear and polynomial interpolation (Claridge and Chen, 2006), or the utilization of statistical forecasts (Liston and Elder, 2006; Walton, 1996). On the other hand, spatial GF techniques employ observations of nearby locations of the same time period as the gap. Multiple spatial interpolation techniques exist (Hartkamp et al., 1999), including inverse-distance weighting (Dhevi, 2014), thin-plate splines (Hutchinson, 1991) and kriging (Garen et al., 1994). Lastly, if a GF technique is performed in both time and spatial dimension, such as the application of empirical orthogonal functions (Beckers and Rixen, 2003), it is classified as a spatiotemporal method. In all three categories, machine learning approaches can be implemented (Bellido-Jiménez et al., 2021; Faramarzzadeh et al., 2023; Sarafanov et al., 2020). The machine learning techniques for GF vary in their implemented algorithm (e.g. LASSO regression, random forest, support vector method, neural network) and input data (e.g. temporal and/or spatial, external datasets).

Besides previous GF techniques, missing values can also be reconstructed through a complete and homogeneous external dataset, such as the ERA5 reanalysis data (Hersbach et al., 2020). It is well known that reanalysis datasets are prone to biases with respect to observations (Cucchi et al., 2020; Simmons et al., 2020). In particular for the near-surface temperature, the ERA5 bias typically shows a diurnal and seasonal variance (Betts et al., 2019; Haiden et al., 2018). The ERA5 bias is even more pronounced for urban locations, since urban observations are not taken into account and urban effects are not modeled due to the low resolution of ERA5 compared to the urban sub-grid scale (Lee and Dessler, 2024). In addition, the bias may be further increased due to a spatial mismatch between the ERA5 grid point and the observational site, introducing e.g. discrepancies in land-sea characteristics for coastal locations or altitude differences. The ERA5 dataset can be included in the spatial and spatiotemporal reconstruction techniques by considering these data as additional stations (Cerlini et al., 2020). On the other hand, a more straight-forward GF method is to replace the missing values with the corresponding values of the bias-adjusted ERA5 dataset (Lipson et al., 2022; Lompar et al., 2019). Regardless of the applied GF technique, it is important to account for the bias of the ERA5 dataset by debiasing the data before its utilization for GF.

To the authors' knowledge, there is hardly any research about filling gaps in urban meteorological time series. The existing GF techniques cannot be directly applied to urban time series, since they do not necessarily account for the specific urban microclimate characteristics like e.g. the UHI. Therefore, comprehensive studies on GF procedures for urban datasets are needed. Filling gaps in meteorological datasets with a technique fit for the urban environment will facilitate the use and thus potential of such observational datasets. In this paper we investigate multiple ERA5-based GF techniques on an urban dataset. Different debiasing techniques for the ERA5 reanalysis data have been suggested and tested, e.g. performing linear regression (Lompar et al., 2019) and calculating mean bias (Lipson et al., 2022). Here we present for the first time an extensive comparison of different ERA5 debiasing GF techniques specifically focused on completing urban temperature time series. Their performances are evaluated by filling artificially constructed gaps in the originally complete hourly temperature time series of the TURCLIM urban climate network (TURCLIM, 2024). The gap construction is based on the gap distribution of the incomplete urban climate network MOCCA (Caluwaerts et al., 2020). By comparing the performances of the GF as a function of the gap length, conclusions are drawn about the most optimal GF technique for each situation. Based on these conclusions, a general GF algorithm is designed to effectively complete a series of gaps in hourly temperature time series. To evaluate the performance of this newly created GF algorithm, specifically for urban temperature series, the impact of the GF on the UHI intensity is studied.

The next section provides more details about the utilized urban climate datasets and the applied methods. This includes a detailed explanation on the different GF techniques and the GF algorithm, along with their evaluation metrics. The results are presented in a subsequent section, followed by a discussion. Finally, the last section provides a general conclusion on GF techniques for urban temperature time series.

## 2. Materials and methods

### 2.1. Utilized datasets

#### 2.1.1. Urban climate network TURCLIM

The TURCLIM urban climate network in the city Turku (Finland), managed by the Geography Section of the University of Turku

(TURCLIM, 2024), is utilized for the evaluation of the GF techniques and algorithm. The TURCLIM network currently consists of 83 observation sites in which air temperature and relative humidity are measured on a half-hour interval (converted to 1-h interval for this study). The measurement height is 3 m above the ground. Turku is a middle-size coastal city (195,000 inhabitants in 2021) with a hemiboreal, humid Dfb climate (Peel et al., 2007). The annual average temperature at Turku Airport was 5.8 °C (1991–2020; Jokinen et al., 2021). The coldest month was February with average temperature of −4.5 °C, average daily minimum temperature of −7.1 °C and average daily maximum temperature of −1.2 °C. The warmest month was July, when the respective temperatures were 17.5 °C, 12.5 °C and 22.6 °C. The UHI of the city center is on average 2 °C, but occasionally it can reach 10 °C. The observations of the TURCLIM network have been utilized in several UHI modeling studies (Hjort et al., 2011; Suomi and Käyhkö, 2012; Suomi et al., 2012; Suomi, 2014; Suomi and Meretoja, 2021; Kivimäki et al., 2023; Suomi et al., 2024).

This study focuses on temperature, given its widespread availability and its prominent role in urban research. Based on the TURCLIM data quality, near-surface air temperature time series from 1 January 2012 till 31 December 2021 (expressed in UTC + 2) are selected for six stations of this quality-controlled network. After a time transformation to UTC and a coarsening to hourly data, the time series are nearly complete, with the exception of a few short gaps which are no longer than 2 h. To be able to perform the evaluation procedure (described in Section 2.3), the dataset is completed by performing a linear interpolation on these short gaps, based on the approach used by Richard et al. (2021).

Fig. 1 shows the locations of the selected stations on a local climate zone (LCZ) map (Stewart and Oke, 2012). The selection consists of three urban and three rural stations. The urban stations Betel and Puutori are located in LCZ 2, while the urban station Virastotalo corresponds to LCZ 5. All three rural stations, named Kurala, Tuorla and Ylijoki, are located in LCZ D.

### 2.1.2. ERA5 reanalysis dataset

The external dataset employed for GF needs to contain complete temperature time series, i.e. no gaps can be present. Several external datasets meet this requirement, e.g. ERA5Land (Muñoz-Sabater et al., 2021) or UERRA (CDS Climate Data Store, 2019), but we opted for ERA5 as it is a global dataset that is widely used within the climate community. ERA5 is the latest version of the reanalysis product of the European Centre for Medium-Range Weather Forecast (Hersbach et al., 2020). This homogeneously distributed dataset represents the past weather conditions all over the globe. The reanalysis product is available from 1940 till present at an hourly temporal resolution. ERA5 covers the entire Earth on a grid with a spatial resolution of 0.25° x 0.25°.

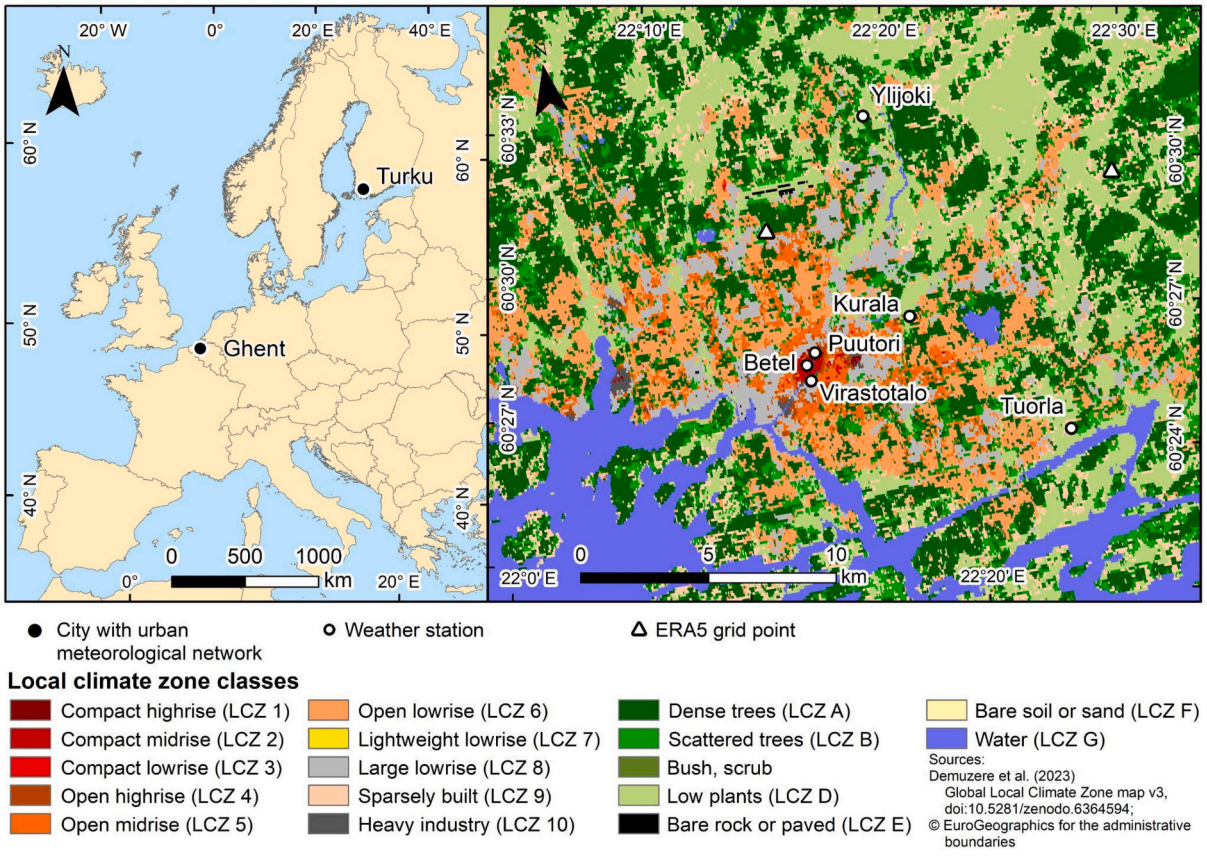



**Fig. 1.** Locations of the utilized urban climate networks on a European map (left), and the locations of the selected TURCLIM stations with nearest ERA5 grid points on the background of a LCZ map (right).

For each selected station of the TURCLIM network, the corresponding ERA5 near-surface temperature data is retrieved by selecting the data of the nearest grid point (Fig. 1). When selecting ERA5 data, it is important to minimize the bias from spatial mismatches between ERA5 grid points and observational sites. For coastal cities, applying a land-sea mask can prevent including ERA5 data with sea characteristics, which could skew near-surface meteorological fields. For the TURCLIM stations used in this research, the nearest land-based grid point was chosen instead of interpolating between grid points that are partially over water. To address altitude mismatches, an altitude correction using a lapse rate coefficient could be applied. However, due to minimal orographic variation in the Turku region, such a correction was not applied for the TURCLIM data.

### 2.1.3. Urban climate network MOCCA

To construct pertinent evaluation metrics, it is important to work with a realistic distribution of gap lengths for urban climate networks. The gap distribution may vary between urban climate networks, depending on e.g. the quality of the communication technology and maintenance, but a general picture is obtained by looking at the distribution of the urban climate network MOCCA (Caluwaerts et al., 2020). This network is located in Ghent, a typical mid-size European city (268,000 inhabitants in 2023 (Statbel, 2024)) located in the North of Belgium at the confluence of the rivers Lys and Scheldt. The region has a flat topography and experiences a mild maritime climate with an average minimum and maximum temperature of 1.1 °C and 6.7 °C in January and 13.4 °C and 23.4 °C in July. The average annual precipitation is 875.6 mm (Royal Meteorological Institute of Belgium RMI, 2024).

The MOCCA network consists of six weather stations located in different LCZs. The temperature data runs from 1 June 2016 till 31 December 2023 (expressed in UTC), with the exception of the Sint-Bavo station which was only operational until 3 May 2022. The original time frequency is 5 min, but the data is coarsened to match the hourly frequency of the TURCLIM data. In Fig. 2, the occurrence of gaps is visualized for the hourly temperature observations of the MOCCA network, after performing quality control with the Python package Metobs-toolkit (Vergauwen et al., 2024). In addition, the distribution of the gap lengths is shown, with respect to both the number of gaps and the number of missing values that correspond to a gap with the corresponding gap length. The gap lengths show a wide variety, ranging from one missing datapoint to 112 consecutive days of missing observations. Although short gaps are more abundant, long gaps are not negligible due to their large number of missing values.

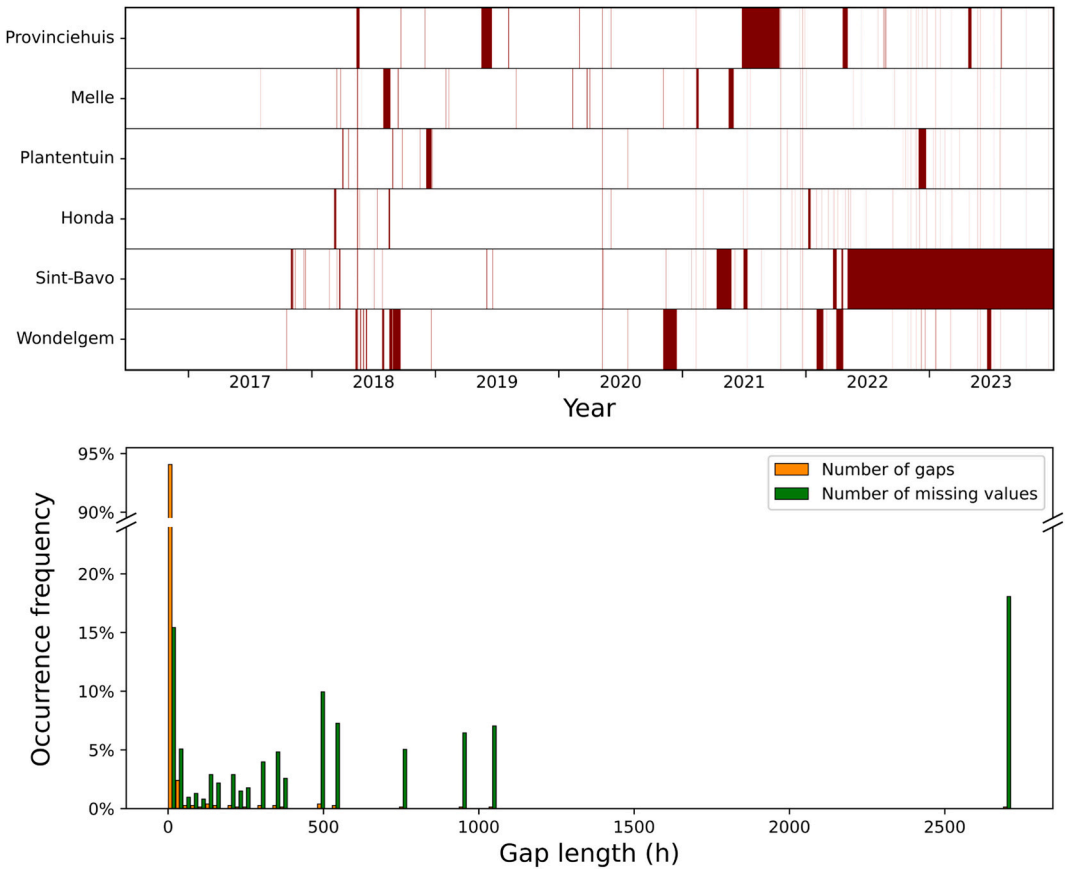



**Fig. 2.** Overview of the gaps for each station (top) and the distribution in function of gap length (bottom) of the hourly temperature observations of MOCCA. The distribution of the gap lengths, both in terms of the number of gaps (orange) and the number of missing values (green), is presented with a bin width of 24 h. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

### 2.2. Gap-filling methodologies

This section first introduces the GF techniques, which are able to fill a single gap in an hourly temperature time series. The second part of this section explains the working mechanism of the GF algorithm, which is designed to handle a series of gaps.

#### 2.2.1. Gap-filling techniques to fill a single gap

Five different GF techniques are implemented, whereof two can be considered as basic while the remaining three are based on debiasing ERA5. The basic GF methods (Fig. 3) serve as a benchmark for the performance evaluation of the debiasing techniques. The first basic one does not rely on an external dataset, and simply estimates the missing values by performing a linear interpolation (LI) in time (Fig. 3a). The second basic method replaces the missing values with original ERA5 data of the nearest grid point to the observation site, without performing any debiasing procedure (Fig. 3b).

The three remaining GF techniques employ the ERA5 data of the nearest grid point, albeit they correct for the ERA5 bias (Fig. 4). The debiasing procedures benefit from the simultaneous availability of observations and ERA5 data in the time period preceding and/or following the gap to assess an eventual bias of ERA5. All three debiasing GF techniques follow the same main steps, starting with the selection of a time period around the gap. After determining the eventual bias of the ERA5 temperature corresponding to the selected time period, this information is applied to debias the ERA5 data during the gap. Finally, the gap is filled by replacing the missing values with the debiased ERA5 temperature data.

When selecting the time period for the bias determination, it is important to take into account the diurnal and seasonal dependency of the ERA5 temperature bias. To account for a potential seasonal bias dependency, a learning period with a limited length is placed either before, after or symmetrically around the gap. For the symmetrical positioning, the possibility exists to consider the part before and after the gap as two separate learning periods, performing the debiasing procedure twice and taking a linear combination of the two results based on the location of the missing value in the gap. This option is specifically designed for very long gaps for which the time period before and after the gap may occur in different seasons. A potential diurnal dependency of the ERA5 bias is addressed by including only data points from the same time of the day, via a time window positioned around the hour of the missing value that is filled. The learning period and time window are defined by three selection parameters: positioning, seasonal span and time variation, for which the definitions are visualized in Fig. 4a. Positioning defines the location of the learning period with respect to the gap (left, right, both or separate), seasonal span ($s$) determines the length of the learning period (expressed in number of days) and time variation ($tv$) is the deviation from the hour of the missing value (expressed in hours, such that the total number of values in the time window is $2tv + 1$). Note that for a certain gap, the learning period is the same for each missing value of this gap, while the placing of the time window depends on the timestamp of the missing value and therefore differs between missing values of the same gap.

While the data selection is identical for all three debiasing GF techniques, they differ in the calculation of the bias and their debiasing procedure. The first debiasing GF technique is based on the method proposed by Lompar et al. (2019), and corrects the ERA5 temperature data by computing a linear regression (LR) relation (Fig. 4b). The ERA5 temperature corresponding to the missing value is debiased by inserting its value into the LR formula. The second debiasing GF technique is based on the work of Lipson et al. (2022), and calculates a mean bias (MB) value which represents the mean difference between the ERA5 and observational temperature of the selected data (Fig. 4c). By subtracting the MB value from the ERA5 temperature of the missing value, a corrected ERA5 value is obtained. The last debiasing GF technique builds on the previous one but assumes a weighted mean bias (WMB). The weights are based on the daily temperature range (DTR) in such a way that days with a DTR similar to the one of the day of the missing value contribute more to the calculation of the mean bias (Fig. 4d). The ERA5 data corresponding to the missing value is corrected by subtracting the WMB. The decision to include the DTR in the bias determination stems from the hypothesis of a DTR dependence of the ERA5 bias, which is plausible since the weather type, which impacts the DTR, has an influence on the bias of ERA5.

#### 2.2.2. Gap-filling algorithm to fill an incomplete time series

Previously explained GF techniques are only able to fill a single gap in a temperature time series, while in practice a temperature dataset may contain several gaps. The designed GF algorithm fills every gap in the temperature time series by selecting the most optimal GF technique, which mainly depends on the gap length as will be shown in Section 3.1. First, the algorithm completes short gaps through linear interpolation, and afterwards longer gaps are one by one filled by the MB debiasing technique. A distinction is

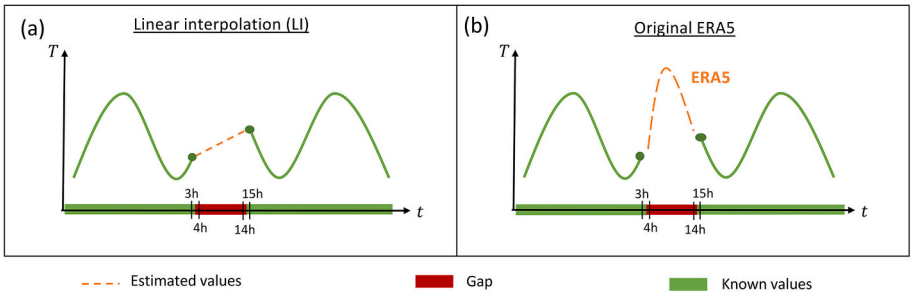



**Fig. 3.** Schematic illustration of the basic GF techniques to fill a single gap: linear interpolation (a) and original ERA5 (b).

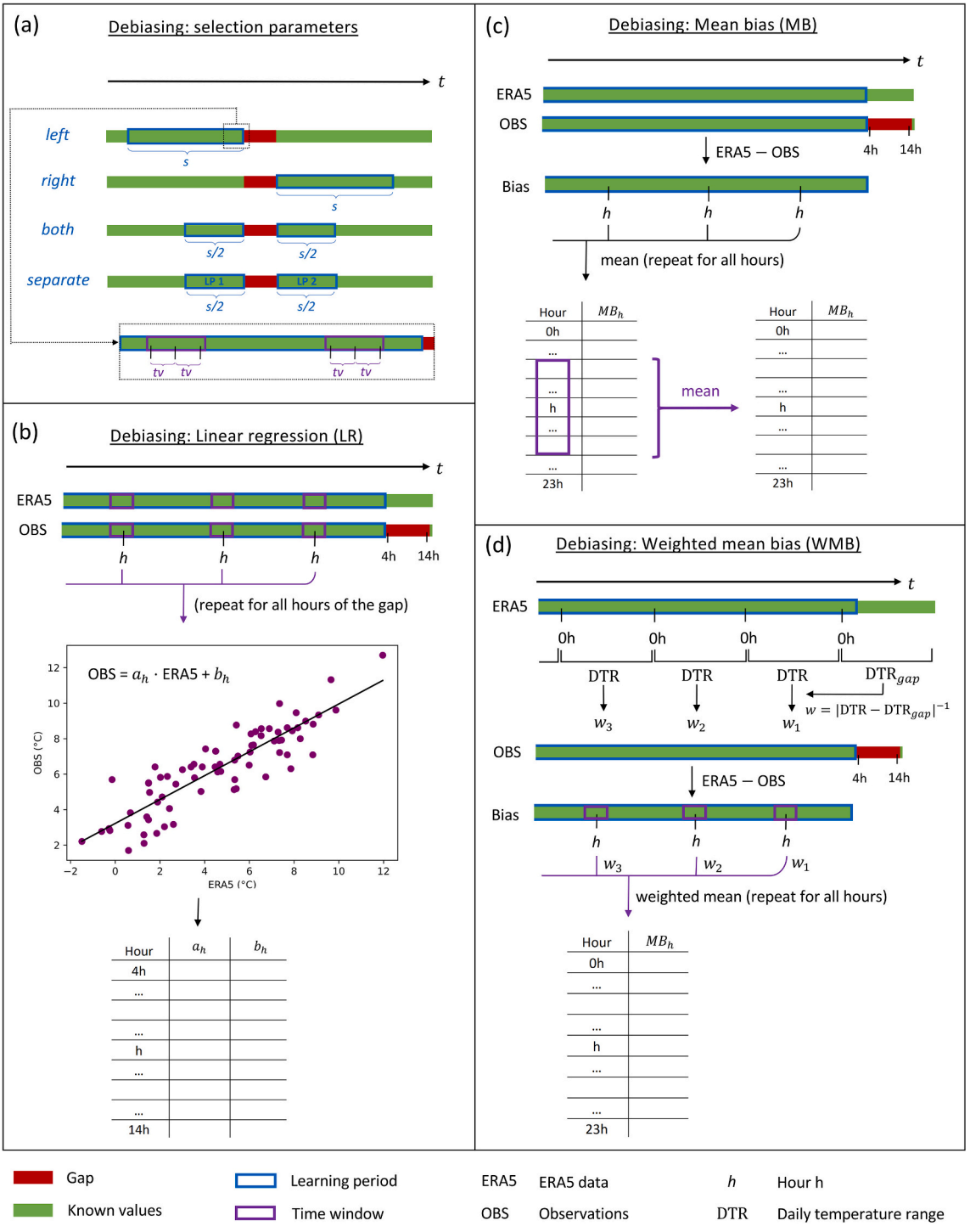



**Fig. 4.** Schematic illustration of the debiasing GF techniques to fill a single gap: the definition of the three selection parameters (a) and the bias calculation procedure of the three debiasing techniques: linear regression (b), mean bias (c) and weighted mean bias (d).

made between moderately long gaps, for which the positioning *both* is selected, and very long gaps, for which the positioning *separate* is applied. Additionally, the algorithm regulates the location of the learning period for the MB debiasing technique, such that the learning period only contains known values and no missing values. The values obtained through linear interpolation of short gaps can be part of a learning period. However, it is decided not to include those acquired from the MB debiasing technique of previous gaps to prevent potential errors from propagating to subsequent gaps. Therefore, the presence of nearby gaps may force the learning period to be shifted or even altered in length to avoid the selection of missing values.

Two problems can arise when altering the length and positioning of the learning period, but for each a solution is provided. First of all, to ensure a sufficiently large data selection, the shortening of the learning period is limited. If the number of available observations from the previous till the next gap is insufficient, an exception is made to include gaps previously filled by the MB debiasing technique to extend the learning period forward. Secondly, extreme shifting of the learning period, due to the appearance of another gap nearby, may result in an asymmetric learning period that is located almost completely before or after the gap. This raises a problem if the *separate* positioning is selected, since the extreme shifting drastically downsizes the length of one of the separate learning periods, making it impossible to perform the debiasing in a statistically correct way. When this situation occurs, the selected positioning is switched to *both*. In order to address various complex scenarios, the algorithm's operational framework is determined by multiple parameters. Table 1 lists all the parameters of the algorithm, including their definition and default values.

### 2.3. Performance evaluation

To evaluate the performance of the GF techniques and GF algorithm, a reference is required. Therefore, this study starts from the complete TURCLIM dataset, in which artificial gaps are subsequently created by removing known values, based on the method used by Lompar et al. (2019) and Henn et al. (2013).

#### 2.3.1. Evaluation of gap-filling techniques

For the examination of the GF techniques, a single gap is created in the temperature time series. To ensure an unhindered placement of the learning period, the gap is placed far enough from the edges of the time series. The performance of the GF techniques is evaluated by calculating the mean squared error (*MSE*) between the estimated values obtained by performing the GF technique (*EST*) and the original observations (*OBS*) of all the missing values of the gap:

$$MSE = \sum_{t=1}^{N_{gap}} \frac{(EST_t - OBS_t)^2}{N_{gap}} \tag{1}$$

with $N_{gap}$ the number of missing values in the gap. Each of the five GF techniques is applied to the same gap, such that a MSE-value is attained for each technique.

To ensure the independence of the performance evaluation with respect to the position of the single gap, this process is repeated multiple times, including a variation in gap length. The general error ($E$) is obtained by calculating the average over all MSE-values, for each GF technique ($gf$) and gap length ($l$) separately:

$$E_{gf,l} = \sum_{i=1}^{n} \frac{MSE_{gf,l,i}}{n} \tag{2}$$

with $n$ the number of repetitions for each gap length.

The uncertainty of the error $E$ is visualized by calculating the standard error (*SE*),

$$SE_{gf,l} = \frac{std\left(MSE_{gf,l,i}\right)}{\sqrt{n}} \tag{3}$$

and determining the 95 % confidence interval (*CI*):

$$CI_{gf,l} = [E - 1.96\ SE, E + 1.96\ SE] \tag{4}$$

The evaluation of the GF techniques is performed in two separate parts. In the first part, the performances of the five different GF techniques are compared, while keeping the three selection parameters (positioning, seasonal span and time variation) constant. Eqs. (2) and (4) are used to calculate $E$ and *CI* for each GF technique separately in function of the gap length. In the second part of the evaluation, the most optimal settings for the selection parameters are evaluated for the MB debiasing technique. Each parameter is varied separately, while keeping the other two parameters constant. Eqs. (2) and (4) are used to calculate $E$ and *CI* for each set of

**Table 1**
Parameters of the GF algorithm, including their description and default value.

| Name | Description | Default value |
|---|---|---|
| Threshold LI | Threshold for the switch between LI and MB debiasing technique. If the gap length (expressed in hours) is smaller than this threshold, LI is applied. | 5 |
| Threshold *both-separate* | Threshold for the switch between positioning *both* and *separate*. If the gap length (expressed in days) equals or exceeds this threshold, the positioning *separate* is applied. | 15 |
| Default seasonal span | Most optimal value for the parameter seasonal span (expressed in days) when applying the MB debiasing technique. | 60 |
| Minimum seasonal span | Minimum value for the parameter seasonal span (expressed in days) when applying the MB debiasing technique. | 30 |
| Minimum learning period one side | Minimum length of each learning period (expressed in days) when applying the MB debiasing technique with positioning *separate*. | 5 |
| Time variation | Value for the parameter time variation (expressed in hours) when applying the MB debiasing technique. | 1 |

selection parameter values in function of the gap length. For both parts of the evaluation, the following choices were made for the values of the constant parameters, based on empirical testing and the values used by Lompar et al. (2019) and Lipson et al. (2022):

- positioning = *both*;
- seasonal span = 60 days;
- time variation = 1 h.

The process of creating a gap and calculating the MSE is repeated 1000 times for each gap length to ensure a robust evaluation. The evaluation is performed for a range of gap lengths from 1 h till 2 weeks, mirroring the distribution of gap lengths found for the MOCCA network with exclusion of the extremely rare gap lengths.

### 2.3.2. *Evaluation of gap-filling algorithm*

The performance of the algorithm for an urban dataset is evaluated by comparing the UHI intensity of the urban TURCLIM stations before and after execution of the GF algorithm. The UHI intensity for each urban station is defined as the mean difference between its temperature (*Turban*) and a rural reference temperature (*Trural*), according to the following formula:

$$UHI(station, season, hour) = \sum_{t=1}^{N} \frac{Turban_t - Trural_t}{N} \tag{5}$$

As a rural reference station, the Ylijoki station is selected. Since the UHI is influenced by both seasonality and diurnal variations (Oke, 1982), the UHI is calculated for each meteorological season (summer JJA, autumn SON, winter DJF, spring MAM) and each hour of the day separately, meaning that $t$ runs over timestamps corresponding to the same season and hour of the day. Using formula (5), the UHI of the originally complete temperature time series, referred to as original UHI, is calculated for each urban station, season, and hour of the day.

To calculate the UHI after GF, first a series of gaps is constructed in the temperature time series of the urban stations and the Ylijoki station (Fig. 5). The gaps are identical for all stations. To guarantee a realistic gap pattern, the gaps are randomly constructed based on the gap frequency and the distribution of the gap lengths of the MOCCA data as visualized in Fig. 2. Next, the gaps are filled by performing the GF algorithm on each time series, with the default values as described in Table 1. Finally, the UHI intensity of the reconstructed time series is calculated for each station, season and hour of the day using formula (5), but note that solely the gap-filled data are utilized. By excluding the non-gap-filled portion in the reconstructed time series, which by definition is identical to the corresponding values in the original observed time series, the score unambiguously reflects the performance of the gap-filling procedure. Since the timing of the gaps is identical for the urban locations and the reference rural station, the UHI intensity can always be calculated based solely on estimated urban and estimated rural temperature values.

The procedure of constructing a series of gaps and calculating the UHI after GF is repeated multiple times, as visualized in Fig. 5. The reason for this is twofold: on the one hand it ensures the independence of the evaluation with respect to the exact location of the gaps, and on the other hand it also ensures the usage of enough data points. After 100 repetitions, an estimated UHI is determined for each station, season and hour of the day by calculating the mean over the 100 UHI-values after GF. In addition, the standard error of this mean is calculated, to determine the 95 % confidence interval. In the end, the performance of the GF algorithm is evaluated by comparing the original UHI with the estimated UHI and the corresponding confidence interval.

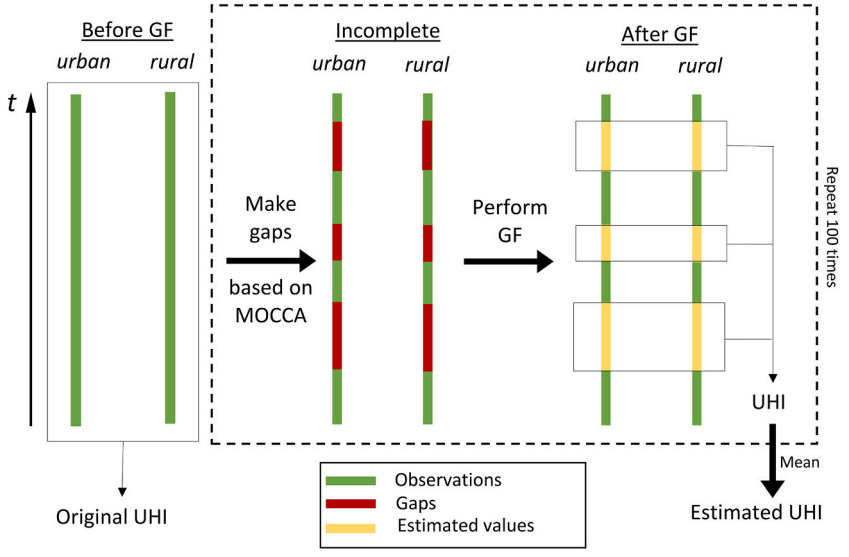



**Fig. 5.** Schematic visualization of the evaluation procedure of the GF algorithm for one urban station.The UHI is calculated two times, once based on the original values (original UHI), and once based on gap-filled values (estimated UHI).

## 3. Results

### 3.1. Evaluation of the gap-filling techniques

The evaluation of the five GF techniques is performed based on the time series of each TURCLIM station separately. Fig. 6 presents the results for one rural station (Ylijoki) and one urban station (Betel), yet similar outcomes are observed for the other weather stations.

For both the rural and urban stations, a distinct difference in error pattern is evident between the LI and ERA5-based GF techniques. The error of the LI technique increases with gap length, with a steep increase for gap lengths below one day. Since the LI method is not able to reproduce the diurnal and seasonal variations in temperature, its poor performance for long gaps does not come as a surprise. In contrast, the errors of the GF techniques based on ERA5 stay fairly constant and still give good results for long gaps. This independence of the error toward the gap length is expected, since the variations in temperature are already accounted for in the external dataset. To summarize, for short gaps the best accuracy is achieved by the LI method, while for longer gaps the (debiased) ERA5 techniques obtain better performance. The transition lies around a gap length of 5 h (rural) to 6 h (urban).

Whether debiasing has an impact on the scores largely depends on the location of the station. For rural stations, debiasing yields negligible improvement, likely indicating that ERA5 does not exhibit significant temperature biases for the rural locations in the region surrounding Turku. On the other hand, for urban stations the results show a substantial difference in performance between the original ERA5 and the debiased ERA5 techniques. Applying a debiasing method instead of directly using ERA5 data results in a significant decrease in the error, reducing it by approximately 2 $°C^2$ across the entire range of gap lengths. Debiasing is indeed expected to yield a notable influence at locations characterized by microclimates not resolved by the ERA5 dataset (e.g. the UHI of cities), as well as in regions marked by complex topography.

The different debiasing techniques exhibit similar performance, with overlapping confidence intervals across the majority of gap lengths. This demonstrates that the specifics of the applied bias procedure exert only a limited influence on the performance. For urban locations, the techniques can be ranked in the following order, starting with the best performance: LR, WMB and MB technique. The difference between WMB and MB suggests that a small improvement is gained by including the DTR variable into the debiasing process. For rural locations, the differences in performance between the debiasing techniques are more limited, with a slightly better performance for the LR and MB technique compared to the WMB technique. The almost non-existing differences between the debiasing techniques for rural stations could be expected, since the debiasing does not yield a large impact on the performance, as stated in the previous paragraph.

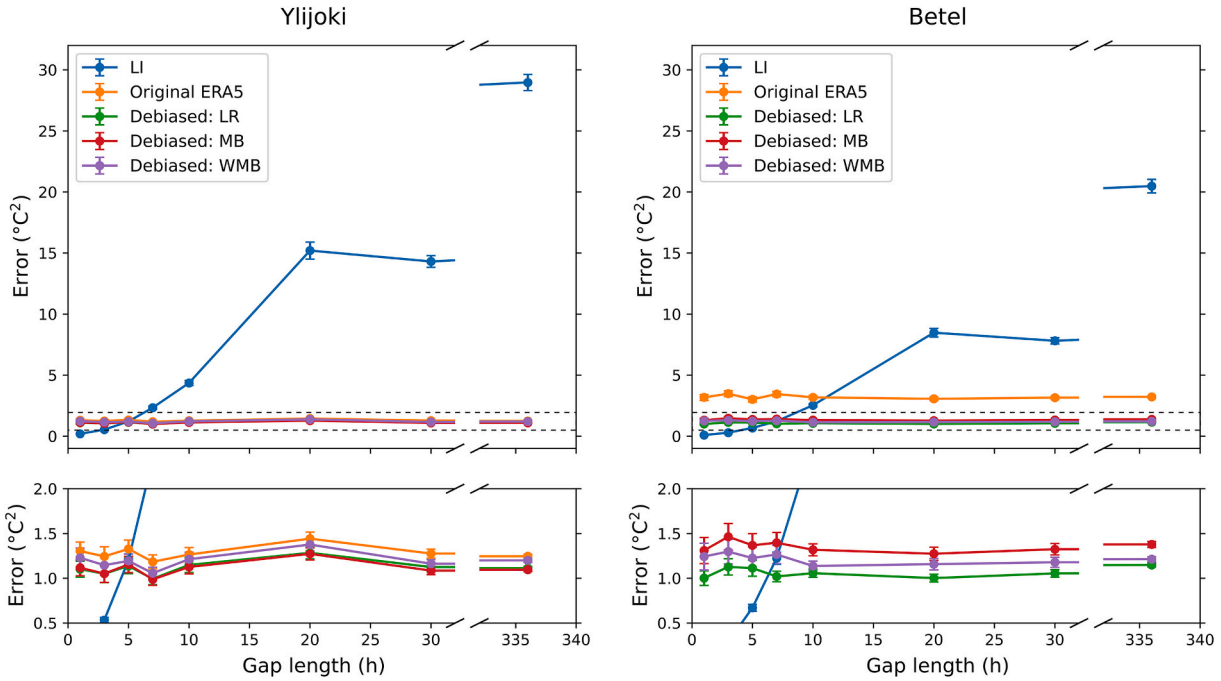



**Fig. 6.** Evaluation of the five GF techniques: linear interpolation (LI), filling in original ERA5, debiasing by calculating linear regression (LR), debiasing by calculating mean bias (MB) and debiasing by calculating a weighted mean bias (WMB). The evaluation is visualized for one rural station (Ylijoki, left) and one urban station (Betel, right). At the bottom a close-up is shown of the error region indicated by the dotted lines. The error bars represent the 95 % confidence intervals.

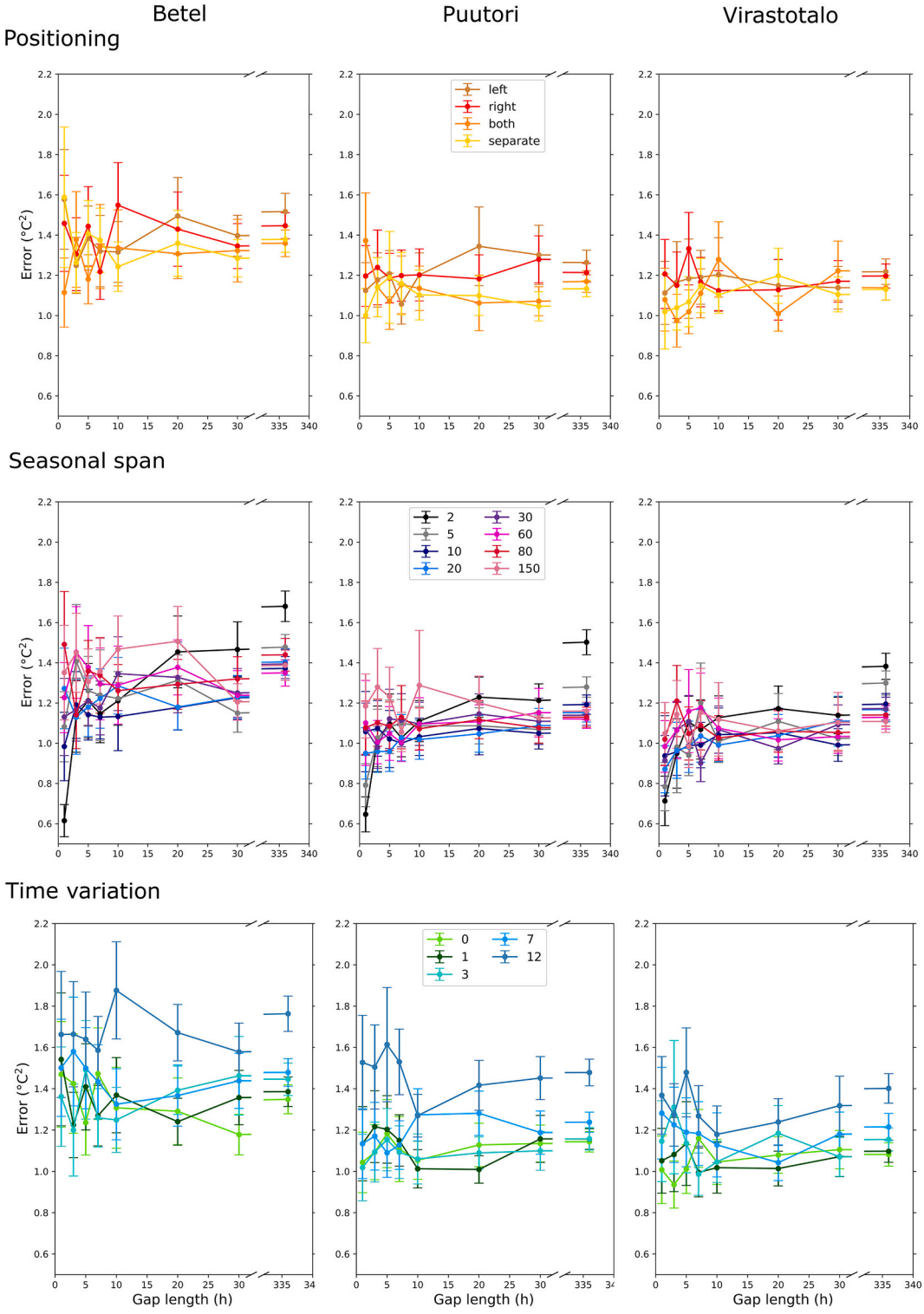



**Fig. 7.** Evaluation of the selection parameters: positioning (top), seasonal span (middle) and time variation (bottom). The evaluation is performed for three urban stations: Betel (left), Puutori (middle) and Virastotalo (right). The error bars represent the 95 % confidence intervals.

### *3.2. Evaluation of the selection parameters*

The evaluation of the selection parameters is performed for the urban stations of TURCLIM. In Fig. 7 the results are shown for each

parameter.

The positioning of the learning period for the debiasing has a limited impact on the error values. The sequence of the positioning errors is not consistent across gap lengths and urban stations, with often overlapping confidence intervals. However, for the longest gaps the error is consequently larger if the learning period is situated completely before (*left*) or after (*right*) the gap, compared to a more symmetric positioning (*both* and *separate*). Although it is expected that *both* and *separate* would perform better for all gap lengths, their advantage is only visible for long gaps. For the longest gaps the *separate* method is expected to perform better, since it assigns more weight to the learning period closest to the filled value. However, this is not the case for all tested urban stations, probably due to the limited length of the tested gaps with a maximum of only 2 weeks.

A large overlap in confidence intervals is also present for the different parameter values of the seasonal span, which expresses the total number of days included in the learning period(s) of the debiasing. Nevertheless, for the longest gap of 2 weeks, the seasonal span of 2 and 5 days shows a significantly larger error compared to longer learning periods. To calculate the bias in a statistically meaningful way, the learning period needs to be large enough, explaining the preference for a long learning period. However, the results of the longest gap indicate that the length of the learning period has no significant impact once a sufficient length is exceeded. On the other hand, for the shortest gap of 1 h, the seasonal span of 2 and 5 days correspond to the lowest error values over all urban locations. For short gaps it seems sufficient to only take into account the day before and after the gap. This excellent result could originate from the good autocorrelation of the temperature time series, which agrees with the good results of the LI method for short gaps.

Similar to the previous two discussed parameters, there is a large overlap in confidence intervals for the majority of the time variation values, which determine the width of the time window. However, the time variation of 12 h consistently corresponds to the highest error values for all stations and gap lengths, with sometimes a non-overlapping confidence interval. Since for this parameter value no distinction is made between the hours of the day, this result could be expected because of the hourly character of the bias. Also in agreement with the expectations, a time variation of 7 h has generally higher error values compared to the parameter values 0, 1 and 3 h.

### 3.3. Evaluation of the gap-filling algorithm

The UHI of each season, both before and after performing the GF algorithm, is visualized in Fig. 8 for the urban station Betel. Other urban stations display very similar results, which are shown in Appendix A.

In general, the GF algorithm is capable of reproducing the UHI trend, with a strong resemblance between the curves of the original and estimated UHI for all seasons. For Betel, the maximal absolute difference between the original and estimated UHI is 0.2° reached in spring at 3 UTC. The GF algorithm not only resolves the general UHI pattern, characterized by high intensity at night and lower intensity during the day, but it also accurately reproduces the local peak around noon during summer and spring. After GF, the seasonal dependence of the UHI is characterized very well, resulting in distinguishable curves between the seasons.

However, the original UHI value does not always lie within the confidence interval of the estimated UHI. The largest deviations between the original and estimated UHI occur during time periods that correspond to the largest seasonal differences in UHI intensity and during time periods of the maximum and minimum UHI intensity. During the night, the UHI intensity of the summer and spring season are systematically underestimated, while for winter and autumn an overestimation occurs. During the day, the opposite occurs.

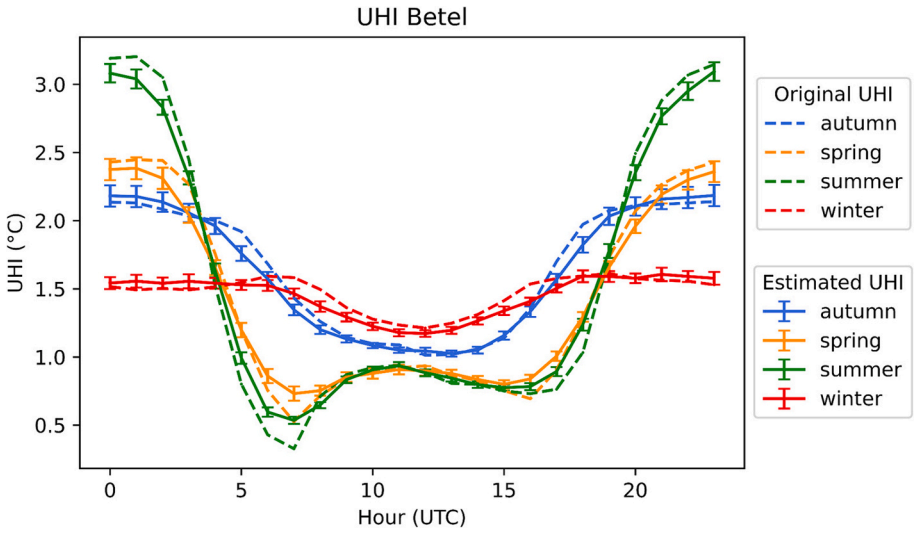



**Fig. 8.** UHI of the urban station Betel, both before and after performing the GF algorithm. The error bars represent the 95 % confidence intervals.

## 4. Discussion

### 4.1. Evaluation of the gap-filling techniques and selection parameters

The performances of the GF techniques justify the designed working mechanism of the GF algorithm, which includes a distinction between small and large gaps. The results state that small gaps are preferably filled with linear interpolation, which is in agreement with results from literature (Cerlini et al., 2020; Henn et al., 2013). On the other hand, large gaps benefit from ERA5-based GF techniques, preferably including a debiasing method. The results indicate that debiasing could be inessential for a rural location, but since results in the literature suggest that debiasing is relevant for rural location, such as complex orographic regions and islands (Lompar et al., 2019), the choice is made to include debiasing for each location. Given the nearly negligible differences in performance among the debiasing techniques, it is justifiable to employ the simplest and computationally least demanding MB technique in the algorithm.

The performance evaluation of the selection parameters results in some interesting key findings. First of all, the choice for the parameter values has a limited impact on the performance. Nevertheless, the evaluation provides some guidance on the optimal choices for the parameter values of the debiasing method, which in case of long gaps corresponds to a symmetrical placement of the learning period with a minimal length of 10 days and a small time window. These preferred parameter values for long gaps correspond well with the empirically chosen parameter values in Section 2.2.2 and 2.3.1, justifying the usage of the latter. Since these empirically chosen values are partially based on values from literature (Lipson et al., 2022; Lompar et al., 2019), it can be concluded that the results obtained through the parameter evaluation agree with these earlier studies.

As the evaluation of the selection parameters is performed for each parameter separately, it is not possible to draw a solid conclusion about the most optimal combination of the three selection parameter values. A better approach would be to simultaneously vary multiple parameters and visualize the results by e.g. a heatmap, but given the many potential parameter combinations, this would result in a high computational cost. However, since the values of the constant parameters lie well in the range of best values according to their single evaluation, the combination of the best values of each single evaluation can be considered to correspond to the best performance.

The results of these evaluations demonstrate that increasing the complexity of the GF technique is not always preferred. For example, implementing the DTR in the MB technique causes worse results for rural locations, and separating the learning period does not yield a significant improvement for the MB debiasing technique compared to the normal symmetrical placement. In addition, more complex techniques will have a higher computational cost. Therefore, when filling a large series of gaps, one has to wonder if the slightly higher accuracy received by increasing the complexity is worth raising the computational cost.

The insights gained by these results have served as a basis for the creation of a GF module of the python package MetObs-toolkit, which is specifically designed to process and analyze raw measurements of urban meteorological networks (Vergauwen et al., 2024).

### 4.2. Evaluation of the gap-filling algorithm

The excellent representation of the UHI by the GF algorithm can be explained by the selective usage of the time series' autocorrelation. Since the UHI intensity is highly dependent on the time during the year and day, as visualized by the curves of the results, it is crucial to only exploit data corresponding to the same time periods during the year and day for the GF. Since this seasonal and hourly selection is (indirectly) embedded in both the LI technique and the MB debiasing technique of the GF algorithm, the excellent UHI resemblance is expected.

The observed deviations between the original and the estimated UHI (Fig. 8) can also be understood by this UHI time dependence. Due to the substantial length of the learning period, it may partially lie in one of the adjoining seasons. This seasonal mixing shifts the UHI intensity toward the one of the adjoining seasons, especially during time periods with large seasonal differences in UHI intensity. As a result, a slight overestimation of the UHI is present for seasons with a limited UHI compared to other seasons, and an underestimation occurs for high UHI-values. Due to a different order of UHI-intensities of the seasons between day and night, the over- or underestimation depends on the time during the day. This may indicate that the most ideal value of the default seasonal span (Table 1) should be smaller. But since the estimated UHI of the different seasons are still distinguishable, it is concluded that overall the length of the learning period is not too extensive and only a minor adjustment may be needed. A similar reasoning can be conducted for the mixing of hours within the time window, leading to a flattening of the UHI curve, as clearly seen in the results.

It should be noted that the presented evaluation is an artificial situation in which only estimated values are included in the calculation of the estimated UHI intensity. In a realistic scenario, with a temperature time series only partially consisting of gaps, the UHI intensity calculated based on the entire completed time series will resemble the UHI effect even better.

### 4.3. Limitations and future work

This research only focuses on the GF of hourly urban temperature time series, but the debiasing GF techniques and GF algorithm could be extended to other types of datasets. First of all, the GF techniques can be applied to other meteorological variables. However, since other variables do not necessarily behave in the same way as temperature, it may be necessary to alter the debiasing procedure by e.g. working with a multiplicative bias correction instead of an additive correction. Secondly, by including a frequency alteration of the ERA5 data, the debiasing techniques could be applied to datasets with other time frequencies. Lastly, the presented techniques and algorithm could also be extended to perform GF in other micro-environments such as forests, lakes, mountains,.... The designed

debiasing techniques will however only employ their full potential if the bias has diurnal and seasonal characteristics. Besides the extension to other gapped datasets, the GF algorithm could also be extended by including other debiasing methods, i.e. machine learning algorithms, or other external datasets, i.e. ERA5Land.

Extending the results of this research to other datasets has to be handled with caution, since the performance evaluation is only performed for hourly temperature time series of one urban climate network. The most optimal GF technique and selection parameter values probably depend on the characteristics of the dataset (urban characteristics, time frequency, meteorological variable, climate zone, …), e.g. for meteorological time series of tropical climates the seasonal selection seems less important. For rural locations, Lompar et al. (2019) already indicate the dependency of the GF performance on the rural characteristics of their tested dataset. To get the best GF results, the GF algorithm could be tuned with respect to the incomplete dataset. It could be considered to first execute a GF evaluation on a complete section of the incomplete time series to select optimal parameters before performing the actual GF.

The designed GF algorithm has some fundamental restrictions. First of all, the algorithm is unable to fill gaps in real-time since the ERA5 data is released with a five-day lag. In addition, the algorithm cannot fill long gaps located at the very beginning of the time series that are followed quickly by another long gap. In this situation it is impossible to place a learning period large enough to perform debiasing. A possible workaround for this problem, which is not yet implemented in the algorithm, is to start the GF at another time location with a lower density of gaps.

## 5. Conclusions

Currently, limited effort is being made to fill gaps in urban meteorological time series. This study demonstrates a variety of simple, yet highly effective, GF methodologies, including different ERA5 debiasing techniques. The results indicate the importance of a distinction between short gaps (<5 h), for which LI yields the best outcomes, and long gaps (≥5 h), for which the ERA5-based techniques are most feasible. For urban locations it is crucial to implement a bias correction on the ERA5 data, although the precise method and selection parameters have a minimal impact on the performance and the most optimal settings will probably depend on the characteristics of the (urban) dataset. Furthermore, based on the results of the evaluation of the individual GF techniques, a newly designed GF algorithm is presented to handle a series of gaps in hourly temperature time series. This algorithm is able to reproduce the hourly and seasonal temperatures in urban environments while conserving the urban characteristics, which is not guaranteed for other existing GF techniques. The utilization of such an urban-aware GF algorithm can ensure a broader utilization of urban observational datasets, without the loss of reliable detection of UHI patterns. This study offers simple but effective methods to fill gaps in hourly urban temperature time series, which are also of broader importance due to their expandability to other meteorological variables, frequencies and micro-climates supporting e.g. climate adaptation measures in city planning.

## CRediT authorship contribution statement

**Amber Jacobs:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation. **Sara Top:** Writing – review & editing, Visualization, Validation, Methodology, Conceptualization. **Thomas Vergauwen:** Writing – review & editing, Data curation. **Juuso Suomi:** Writing – review & editing, Resources. **Jukka Käyhkö:** Writing – review & editing, Resources. **Steven Caluwaerts:** Writing – review & editing, Validation, Supervision, Methodology, Funding acquisition, Conceptualization.

## Declaration of competing interest

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

## Data availability

The code is available in a public github repository https://github.com/amberJ99/Gapfilling_debiasingERA5. Data will be made available on request.

## Acknowledgements

A. Jacobs and S. Top are funded by the FWO (Fund for Scientific Research) of the Flemish regional government, fellowships 11PBN24N and 1270723N respectively. T. Vergauwen en S. Caluwaerts are respectively supported by BELSPO projects B2/202/P1/CS-MASK and FED-tWIN 2020-018_AURA. The Geography Section of the University of Turku and the Urban Environment Division of the City of Turku have by their financial support enabled maintaining the TURCLIM urban climate network. Contacts in the context of the FAIRNESS COST Action CA20108 have contributed to the research development and data exchange.

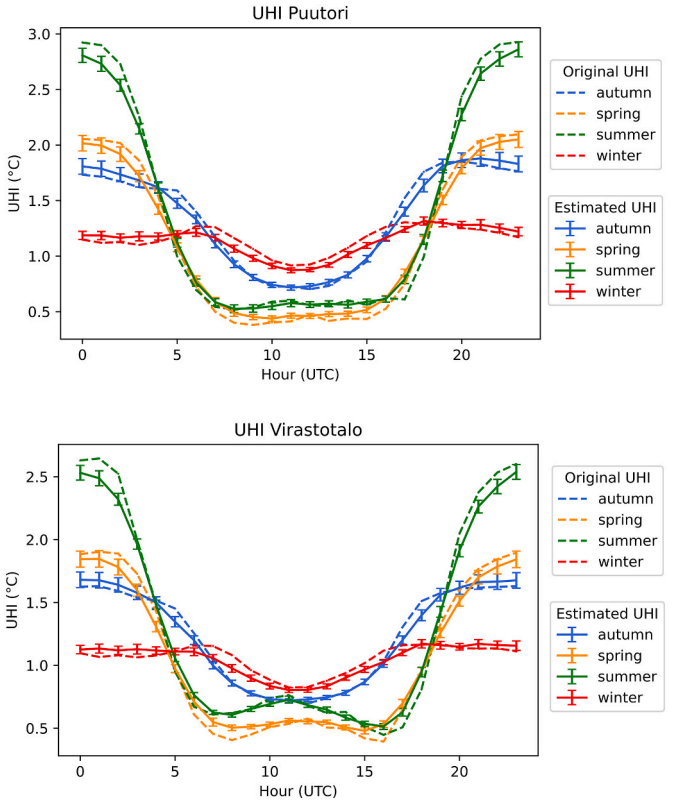



**Fig. A.1.** UHI of the urban station Puutori (top) and Virastotalo (bottom), both before and after performing the GF algorithm. The error bars represent the 95 % confidence intervals.

## References


Afrifa-Yamoah, E., Mueller, U.A., Taylor, S.M., Fisher, A.J., 2020. Missing data imputation of high-resolution temporal climate time series data. Meteorol. Appl. 27 (1), e1873. https://doi.org/10.1002/met.1873.

Aieb, A., Madani, K., Scarpa, M., Bonaccorso, B., Lefsih, K., 2019. A new approach for processing climate missing databases applied to daily rainfall data in Soummam watershed, Algeria. Heliyon 5 (2). https://doi.org/10.1016/j.heliyon.2019.e01247.

Arnfield, A.J., 2003. Two decades of urban climate research: a review of turbulence, exchanges of energy and water, and the urban heat island. Int. J. Climatol. 23 (1), 1–26. https://doi.org/10.1002/joc.859.

Beckers, J.M., Rixen, M., 2003. EOF calculations and data filling from incomplete oceanographic datasets. J. Atmos. Oceanic Tech. 20 (12), 1839–1856. https://doi.org/10.1175/1520-0426(2003)020%3C1839:ECADFF%3E2.0.CO;2.

Bellido-Jiménez, J.A., Gualda, J.E., García-Marín, A.P., 2021. Assessing machine learning models for gap filling daily rainfall series in a semiarid region of Spain. Atmosphere 12 (9), 1158. https://doi.org/10.3390/atmos12091158.

Betts, A.K., Chan, D.Z., Desjardins, R.L., 2019. Near-surface biases in ERA5 over the Canadian prairies. Front. Environ. Sci. 7, 129. https://doi.org/10.3389/fenvs.2019.00129.

Caluwaerts, S., Hamdi, R., Top, S., Lauwaet, D., Berckmans, J., Degrauwe, D., Dejonghe, H., De Ridder, K., De Troch, R., Duchêne, F., Maiheu, B., Van Ginderachter, M., Verdonck, M., Vergauwen, T., Wauters, G., Termonia, P., 2020. The urban climate of Ghent, Belgium: a case study combining a high-accuracy monitoring network with numerical simulations. Urban Clim. 31, 100565. https://doi.org/10.1016/j.uclim.2019.100565.

CDS (Climate Data Store), 2019. Complete UERRA Regional Reanalysis for Europe from 1961 to 2019. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). https://cds.climate.copernicus.eu/datasets/reanalysis-uerra-europe-complete?tab=overview (accessed 13 August 2024).

Cerlini, P.B., Silvestri, L., Saraceni, M., 2020. Quality control and gap-filling methods applied to hourly temperature observations over Central Italy. Meteorol. Appl. 27 (3), e1913. https://doi.org/10.1002/met.1913.

Claridge, D.E., Chen, H., 2006. Missing data estimation for 1–6 h gaps in energy use and weather data using different statistical methods. Int. J. Energy Res. 30 (13), 1075–1091. https://doi.org/10.1002/er.1207.

Coney, J., Pickering, B., Dufton, D., Lukach, M., Brooks, B., Neely III, R.R., 2022. How useful are crowdsourced air temperature observations? An assessment of Netatmo stations and quality control schemes over the United Kingdom. Meteorol. Appl. 29 (3), e2075. https://doi.org/10.1002/met.2075.

Cucchi, M., Weedon, G.P., Amici, A., Bellouin, N., Lange, S., Müller Schmied, H., Herbach, H., Buontempo, C., 2020. WFDE5: bias-adjusted ERA5 reanalysis data for impact studies. Earth Syst. Sci. Data 12 (3), 2097–2120. https://doi.org/10.5194/essd-12-2097-2020.

Dhevi, A.S., 2014. Imputing missing values using inverse distance weighted interpolation for time series data. In: 2014 Sixth International Conference on Advanced Computing (ICoAC), pp. 255–259. https://doi.org/10.1109/ICoAC.2014.7229721.

Diouf, S., Dème, A., 2022. Imputation methods for missing values: the case of Senegalese meteorological data. Afr. J. Appl. Stat. 9 (1), 1245–1278. https://doi.org/10.16929/ajas/2022.1245.267.

Faramarzzadeh, M., Ehsani, M.R., Akbari, M., Rahimi, R., Moghaddam, M., Behrangi, A., Klöve, B., Haghighi, A.T., Oussalah, M., 2023. Application of machine learning and remote sensing for gap-filling daily precipitation data of a sparsely gauged basin in East Africa. Environ. Process. 10 (1), 8. https://doi.org/10.1007/s40710-023-00625-y.

Garen, D.C., 2013. Choosing and assimilating forcing data for hydrological prediction. In: Pomeroy, J.W., Whitfield, P.H., Spence, C. (Eds.), Putting Prediction in Ungauged Basins into Practice. Canadian Water Resources Association, pp. 89–100.

Garen, D.C., Johnson, G.L., Hanson, C.L., 1994. Mean areal precipitation for daily hydrologic modeling in mountainous regions. JAWRA J. Am. Water Resourc. Assoc. 30 (3), 481–491. https://doi.org/10.1111/j.1752-1688.1994.tb03307.x.
Haiden, T., Sandu, I., Balsamo, G., Arduini, G., Beljaars, A., 2018. Addressing biases in near-surface forecasts. ECMWF Newsletter 157, 20–25. https://doi.org/10.21957/eng71d53th.
Hartkamp, A.D., De Beurs, K., Stein, A., White, J.W., 1999. Interpolation techniques for climate variables. NRG-GIS Series 90-01. Cimmyt.
Henn, B., Raleigh, M.S., Fisher, A., Lundquist, J.D., 2013. A comparison of methods for filling gaps in hourly near-surface air temperature data. J. Hydrometeorol. 14 (3), 929–945. https://doi.org/10.1175/JHM-D-12-027.1.
Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., et al., 2020. The ERA5 global reanalysis. Q. J. Roy. Meteorol. Soc. 146 (730), 1999–2049. https://doi.org/10.1002/qj.3803.
Hjort, J., Suomi, J., Käyhkö, J., 2011. Spatial prediction of urban–rural temperatures using statistical methods. Theor. Appl. Climatol. 106, 139–152. https://doi.org/10.1007/s00704-011-0425-9.
Hutchinson, M.F., 1991. The application of thin plate smoothing splines to continent-wide data assimilation. Data Assimilation Syst. 104–113.
Ivajnšič, D., Žiberna, I., 2019. The effect of weather patterns on winter small city urban heat islands. Meteorol. Appl. 26 (2), 195–203. https://doi.org/10.1002/met.1752.
Jokinen, P., Pirinen, P., Kaukoranta, J.P., Kangas, A., Alenius, P., Eriksson, P., Johansson, M., Wilkman, S., 2021. Tilastoja Suomen Ilmastosta ja Merestä 1991-2020. https://doi.org/10.35614/isbn.9789523361485.
Kang, H., 2013. The prevention and handling of the missing data. Korean J. Anesthesiol. 64 (5), 402. https://doi.org/10.4097/kjae.2013.64.5.402.
Kivimäki, M., Batty, G.D., Pentti, J., Suomi, J., Nyberg, S.T., Merikanto, J., Nordling, K., Ervasti, J., Suominen, S.B., Partanen, A.I., Stenholm, S., Käyhkö, J., Vahtera, J., 2023. Climate change, summer temperature, and heat-related mortality in Finland: multicohort study with projections for a sustainable vs. fossil-fueled future to 2050. Environ. Health Perspect. 131 (12), 127020. https://doi.org/10.1289/EHP12080.
Lee, J., Dessler, A.E., 2024. Improved surface urban heat impact assessment using GOES satellite data: a comparative study with ERA-5. Geophys. Res. Lett. 51 (1), e2023GL107364. https://doi.org/10.1029/2023GL107364.
Lipson, M., Grimmond, S., Best, M., Chow, W., Christen, A., Chrysoulakis, N., Coutts, A., Crawford, B., Earl, S., Evans, J., Fortuniak, K., Heusinkveld, B.G., Hong, J., Hong, J., Järvi, L., Jo, S., Kim, Y., Kotthaus, S., Lee, K., et al., 2022. Harmonized gap-filled datasets from 20 urban flux tower sites. Earth Syst. Sci. Data Discuss. 2022, 1–29. https://doi.org/10.5194/essd-14-5157-2022.
Liston, G.E., Elder, K., 2006. A meteorological distribution system for high-resolution terrestrial modeling (MicroMet). J. Hydrometeorol. 7 (2), 217–234. https://doi.org/10.1175/JHM486.1.
Lompar, M., Lalić, B., Dekić, L., Petrić, M., 2019. Filling gaps in hourly air temperature data using debiased ERA5 data. Atmosphere 10 (1), 13. https://doi.org/10.3390/atmos10010013.
Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D., Piles, M., Rodríguez-Fernández, N., Zsoter, E., Buontempo, C., Thépaut, J.N., 2021. ERA5-land: a state-of-the-art global reanalysis dataset for land applications. Earth Syst. Sci. Data 13 (9), 4349–4383. https://doi.org/10.5194/essd-13-4349-2021.
Oke, T.R., 1982. The energetic basis of the urban heat island. Q. J. Roy. Meteorol. Soc. 108 (455), 1–24. https://doi.org/10.1002/qj.49710845502.
Peel, M.C., Finlayson, B.L., McMahon, T.A., 2007. Updated world map of the Köppen-Geiger climate classification. Hydrol. Earth Syst. Sci. 11 (5), 1633–1644. https://doi.org/10.5194/hess-11-1633-2007.
Richard, A., Fine, L., Rozenstein, O., Tanny, J., Geist, M., Pradalier, C., 2021. Filling Gaps in Micro-Meteorological Data. https://doi.org/10.1007/978-3-030-67670-4_7.
Royal Meteorological Institute of Belgium (RMI), 2024. Climate in Your Community. https://www.meteo.be/nl/klimaat/klimaat-van-belgie/klimaat-in-uw-gemeente (accessed 5 June 2024).
Sarafanov, M., Kazakov, E., Nikitin, N.O., Kalyuzhnaya, A.V., 2020. A machine learning approach for remote sensing data gap-filling with open-source implementation: an example regarding land surface temperature, surface albedo and NDVI. Remote Sens. (Basel) 12 (23), 3865. https://doi.org/10.3390/rs12233865.
Simmons, A., Soci, C., Nicolas, J., Bell, B., Berrisford, P., Dragani, R., Flemming, J., Haimberger, L., Healy, S., Hersbach, H., Horányi, A., Inness, A., Muñoz-Sabater, J., Radu, R., Schepers, D., 2020. Global Stratospheric Temperature Bias and other Stratospheric Aspects of ERA5 and ERA5, p. 1. https://doi.org/10.21957/rcxqfmg0.
Statbel, 2024. Population by Place of Residence, Nationality (Belgian/non-Belgian), Marital Status, Age and Gender. https://bestat.statbel.fgov.be/bestat/crosstable.xhtml?view=fc14c1ce-7361-4d42-a892-fce8e81a1b79 (accessed 3 June 2024).
Stewart, I.D., Oke, T.R., 2012. Local climate zones for urban temperature studies. Bull. Am. Meteorol. Soc. 93 (12), 1879–1900. https://doi.org/10.1175/BAMS-D-11-00019.1.
Suomi, J., 2014. Characteristics of Urban Heat Island (UHI) in a High Latitude Coastal City - a Case Study of Turku, SW Finland.
Suomi, J., Käyhkö, J., 2012. The impact of environmental factors on urban temperature variability in the coastal city of Turku, SW Finland. Int. J. Climatol. 32 (3), 451–463. https://doi.org/10.1002/joc.2277.
Suomi, J., Meretoja, M., 2021. Trends and irregular variation of spatial temperature differences in the high-latitude coastal city of Turku, Finland. Climate Res. 84, 41–57. https://doi.org/10.3354/cr01649.
Suomi, J., Hjort, J., Käyhkö, J., 2012. Effects of scale on modelling the urban heat island in Turku, SW Finland. Clim. Res. 55 (2), 105–118. https://doi.org/10.3354/cr01123.
Suomi, J., Saranko, O., Partanen, A.I., Fortelius, C., Gonzales-Inca, C., Käyhkö, J., 2024. Evaluation of surface air temperature in the HARMONIE-AROME weather model during a heatwave in the coastal city of Turku, Finland. Urban Clim. 53, 101811. https://doi.org/10.1016/j.uclim.2024.101811.
Top, S., Milošević, D., Caluwaerts, S., Hamdi, R., Savić, S., 2020. Intra-urban differences of outdoor thermal comfort in Ghent on seasonal level and during record-breaking 2019 heat wave. Build. Environ. 185, 107103. https://doi.org/10.1016/j.buildenv.2020.107103.
TURCLIM, 2024. Turku Urban Climate Research Group. https://sites.utu.fi/turclim/ (accessed 15 May 2024).
Vergauwen, T., Vieijra, M., Covaci, A., Jacobs, A., Top, S., Dewettinck, W., Vandelanotte, K., Hellebosch, I., Caluwaerts, S., 2024. MetObs-a Python toolkit for using non-traditional meteorological observations. J. Open Source Software 9 (95), 5916. https://doi.org/10.5281/ZENODO.10794417.
Walton, T.L., 1996. Fill-in of missing data in univariate coastal data. J. Appl. Stat. 23 (1), 31–40. https://doi.org/10.1080/02664769624332.
Yozgatligil, C., Aslan, S., Iyigun, C., Batmaz, I., 2013. Comparison of missing value imputation methods in time series: the case of Turkish meteorological data. Theor. Appl. Climatol. 112, 143–167. https://doi.org/10.1007/s00704-012-0723-x.