



# How to correctly apply Gaussian statistics in a non-stationary climate?

Reinhold Steinacker<sup>1,2</sup>

Received: 17 October 2019 / Accepted: 17 March 2021 / Published online: 5 April 2021  
© The Author(s) 2021

## Abstract

Time series with a significant trend, as is now being the case for the temperature in the course of climate change, need a careful approach for statistical evaluations. Climatological means and moments are usually taken from past data which means that the statistics does not fit to actual data anymore. Therefore, we need to determine the long-term trend before comparing actual data with the actual climate. This is not an easy task, because the determination of the signal—a climatic trend—is influenced by the random scatter of observed data. Different filter methods are tested upon their quality to obtain realistic smoothed trends of observed time series. A new method is proposed, which is based on a variational principle. It outperforms other conventional methods of smoothing, especially if periodic time series are processed. This new methodology is used to test, how extreme the temperature of 2018 in Vienna actually was. It is shown that the new annual temperature record of 2018 is not too extreme, if we consider the positive trend of the last decades. Also, the daily mean temperatures of 2018 are not found to be really extreme according to the present climate. The real extreme of the temperature record of Vienna—and many other places around the world—is the strongly increased positive temperature trend over the last years.

## 1 Introduction

The average temperature 2018 in Vienna resulted in 13.0 °C, an unprecedented value, nearly 1 degree higher than the warmest year in the whole 244-year time series since 1775 (Hiebl et al. 2019). If we compare this value with the mean temperature of the last climate normal period 1961–1990 ( $\hat{T} = 9.7^\circ\text{C}$ ), it yields a positive anomaly of +3.3 °C. This corresponds to more than 4 times the standard deviation (SDEV) of that period. Even if we compare the 13.0°C with the average of the last completed three decades 1981–2010 ( $\hat{T} = 10.5^\circ\text{C}$ ), the deviation is +2.5 °C, corresponding to more than three times the SDEV of that period. Taking longer periods, e.g., the twentieth century or the whole 244-year series (including the year 2018), the deviation was even larger, roughly 5 SDEV and 4 SDEV respectively. Hence, the year 2018 is commonly be seen as extraordinary and an extreme

outlier temperature-wise. But is this simple statistical evaluation adequate?

The World Meteorological Organization (WMO) has defined a climate normal period long ago as a 30-year period (WMO 1989). Starting in 1900, the periods 1901–1930, 1931–1960, and 1961–1990 have been defined and the next normal period 1991–2020 will be used from 2021 onwards. Due to the rapid global temperature increase now also the use of 30-year overlapping periods (e.g., 1981–2010) are recommended (WMO 2007). A thirty year period has been neatly chosen for temperature, because on one hand it is long enough to give a statistically reasonable mean value and statistical moments, on the other hand it is short enough, that climate trends do not have a strong impact on the statistical moments (Angel et al. 1993). This was certainly true until the eighties of the twentieth century, when the climatological global temperature trends were never larger than a few tenths of a degree within two consecutive 30-year periods. Hence, the climatological time series were close enough to stationarity. Since around 1980 however, the temperature trends, locally, regionally as well as globally have significantly increased in the course of global warming. Hence, for the statistical treatment and interpretation of recent temperature time series, e.g., the determination of statistical moments, we have to consider trends.

✉ Reinhold Steinacker  
reinhold.steinacker@gmail.com

<sup>1</sup> Department of Meteorology and Geophysics, University of Vienna, Althanstrasse 14, 1090 Vienna, Austria

<sup>2</sup> Karl-Innerebnerstrasse 75 b, 6020, Innsbruck, Austria

In Section 2, different methods to smooth (filter) time series are discussed and a new method is introduced, which allows a reasonable determination of trends from the very beginning until the very end of a time series. In Section 3, the application of the filter is used to compute the trend of the temperature time series of Vienna and to determine time dependent mean values and the corresponding statistical moments. The concluding Section 4 gives suggestions, how to treat time series and to receive adequate statistical measures in times of significant climate trends.

## 2 Filtering time series

There are many ways how to filter time series (Hamming 1989; Schönwiese 2013; Shumway and Stoffer 2006). The simplest—and mathematically cheapest—is the application of MOVing Average Smoothing, furthermore denoted as MOVAS. This method, however, has some unwelcome properties. Part of high frequency variations may still be present after smoothing and the spectral response may even invert the phase of some wavelike variations. A further problem is the fact, that moving averages do not allow to determine values at both ends of a time series, as long as we stay with centered averaging. If we continue the averaging until the very end of the series, we have to take one-sided averages, which usually lead to a significant underestimation of the magnitude of a trend. The problem of determining reasonable trends at the end of time series has been extensively investigated (e.g., Arguez et al. 2008; Mann 2004, 2008). Besides the determination of trends at the end of time series also trend extrapolation is applied in many scientific disciplines (Box and Jenkins 1970; Livezey et al. 2007).

A better method than MOVAS, which is frequently used to filter climatological time series (e.g., ZAMG 2019) is the Gaussian Smoothing filter, furthermore denoted by GAUSS. Here, the averaging is performed by weighting the data within an averaging window according to the Gaussian frequency distribution. Such a Gaussian filter smooths out the high frequency parts of a time series much better than MOVAS but still does not allow to get reasonable trends at the ends of a time series. Other frequently used time filters in many fields of science are the closely related Locally Estimated Scatterplot Smoothing (LOESS) and the Locally Weighted Scatterplot Smoothing (LOWESS) (Cleveland 1981; Savitzky and Golay 1964). Within an arbitrary time window usually a quadratic polynomial is used for LOESS and a linear polynomial is fitted to weighted data for LOWESS. In addition a LOWESS2 has been investigated here, which takes a weighted quadratic polynomial fit. The suggested weighting function is given by

$$w(\tau) = \begin{cases} \left(1 - \left|\frac{\tau}{\Delta\tau}\right|\right)^3 & \text{if } \left|\frac{\tau}{\Delta\tau}\right| \leq 1, \\ 0 & \text{otherwise} \end{cases}, \quad (1)$$

where  $\tau$  is the time difference with respect to the center of the particular time window and  $\Delta\tau$  represents the half-width of the smoothing time window. This function has some similarity to the Gaussian function. Another smoothing filter is based on a selective cutoff frequency of harmonic waves. Only the lowest  $b$  harmonics of a time series are considered, whereas the higher harmonics are completely disregarded. For a periodic temperature time series, this can be expressed as:

$$\begin{aligned} \tilde{T}(t) = & \frac{1}{n} \sum_{t=1}^n T(t) + \sum_{k=1}^b T_{k,s} \sin\left(\frac{2\pi kt}{n}\right) \\ & + \sum_{k=1}^b T_{k,c} \cos\left(\frac{2\pi kt}{n}\right), \end{aligned} \quad (2)$$

where  $\tilde{T}(t)$  denotes the smoothed value of the variable at time  $t$  and  $T_{k,s}$  and  $T_{k,c}$  denote the coefficients (amplitudes) of the  $k$ th sine and cosine harmonics, which are determined by a least square method. For a non-periodic time series, the first term on the right hand side of Eq. 2 can be replaced by the intercept and slope of the linear regression of  $T(t)$  with regard to time. This type of filter is denoted by SPECTral Smoothing filter (SPECS) in the following. A new filter method proposed here is the mean value smoothing spline, furthermore denoted by MEVSS. For that, discrete values  $\tilde{T}(t)$  of the smoothest possible function are searched, which equal the mean value of  $T(t)$  of the original time series within arbitrary, non-overlapping smoothing time windows. Mathematically this can be expressed for a time series with  $n$  data points by the condition:

$$\begin{aligned} \sum_{t=2}^{n-1} [\tilde{T}(t-1) - 2\tilde{T}(t) + \tilde{T}(t+1)]^2 \rightarrow \text{Min} & \quad \text{for non-periodic time series and} \\ \sum_{t=1}^n [\tilde{T}(t-1) - 2\tilde{T}(t) + \tilde{T}(t+1)]^2 \rightarrow \text{Min} & \quad \text{for periodic time series,} \end{aligned} \quad (3)$$

where we set for the latter  $\tilde{T}(0) = \tilde{T}(n)$  and  $\tilde{T}(n+1) = \tilde{T}(1)$ .

The constraint for Eq. 3 for all  $i$  non-overlapping smoothing time intervals is

$$\lambda_i \sum_{t=m_i-\Delta t_i}^{m_i+\Delta t_i} \tilde{T}(t) = \lambda_i \sum_{t=m_i-\Delta t_i}^{m_i+\Delta t_i} T(t) \quad (4)$$

where  $m_i$  denotes the center data point and  $\Delta t_i$  is the half-width of the  $i$ th smoothing interval of the time series and  $\lambda_i$  are Lagrangian multipliers. For open time series, the first/last smoothing interval must start/end with the time series, whereas for periodic series no such limitation exists. The smoothed values for MEVSS are yielded by deriving Eqs. 3 and 4 by all  $\tilde{T}(t)$  and  $\lambda_i$  and solving the resulting set of  $n + i$  linear

equations. Even if the number of data points is large, the numerical solution is very effective and fast, because the matrix of coefficients is an extremely sparse matrix. It should be noted that MEVSS is not equal to the conventional smoothing spline procedure, also known as variational interpolation (Sasaki 1955), furthermore denoted in this paper by SPLIS. There, a function is searched for by the condition

$$\sum_{t=2}^{n-1} (\tilde{T}(t-1) - 2\tilde{T}(t) + \tilde{T}(t+1))^2 + \gamma \sum_{t=1}^n (\tilde{T}(t) - T(t))^2 \rightarrow \text{Min}. \tag{5}$$

$\gamma$  determines the degree of smoothness of the filtered time series. The smaller  $\gamma$ , the smoother is the filtered time series.

To study the effect of different smoothing procedures, we take two synthetic temperature time series with 360 discrete points of  $T(t)$ . The first is based on an analytic sinusoidal function  $T_s(t)$  superposed by a random Gaussian noise  $T_n(t)$ . This is similar to an annual time series of a multi-year average of the daily mean, minimum or maximum temperature with the time increment of 1 day ( $d$ ). In contrast to a single year daily temperature record, two consecutive days in an average multi-year record usually are hardly correlated. Figure 1 shows such a synthetic periodic time series for 360 days with

$$T(t) = T_0 + T_s(t) + T_n(t) = 10^\circ\text{C} + 10^\circ\text{C} \cdot \sin\left(\frac{2\pi t}{360d}\right) + T_n(t). \tag{6}$$

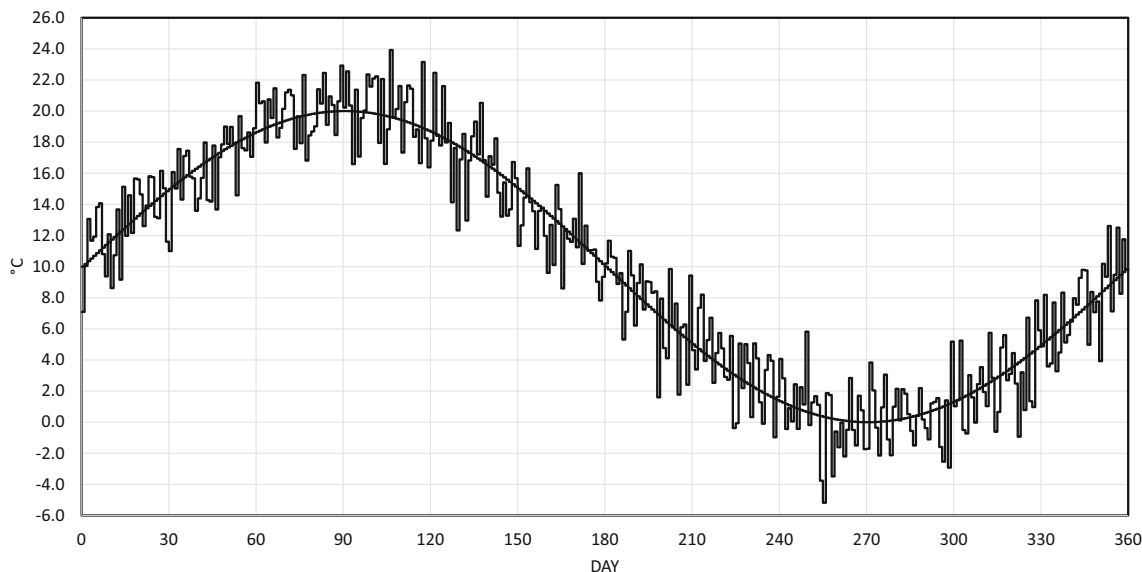
The Gaussian distributed random noise has a standard deviation of 2.

The second synthetic time series consists of an exponential function  $T_e(t)$  superposed by a random Gaussian noise with a standard deviation of 1. This is similar to a very long (360 years) time series of annual mean temperatures with an increasing positive temperature trend at the end and a time increment of 1 year ( $a$ ) (see Fig. 2):

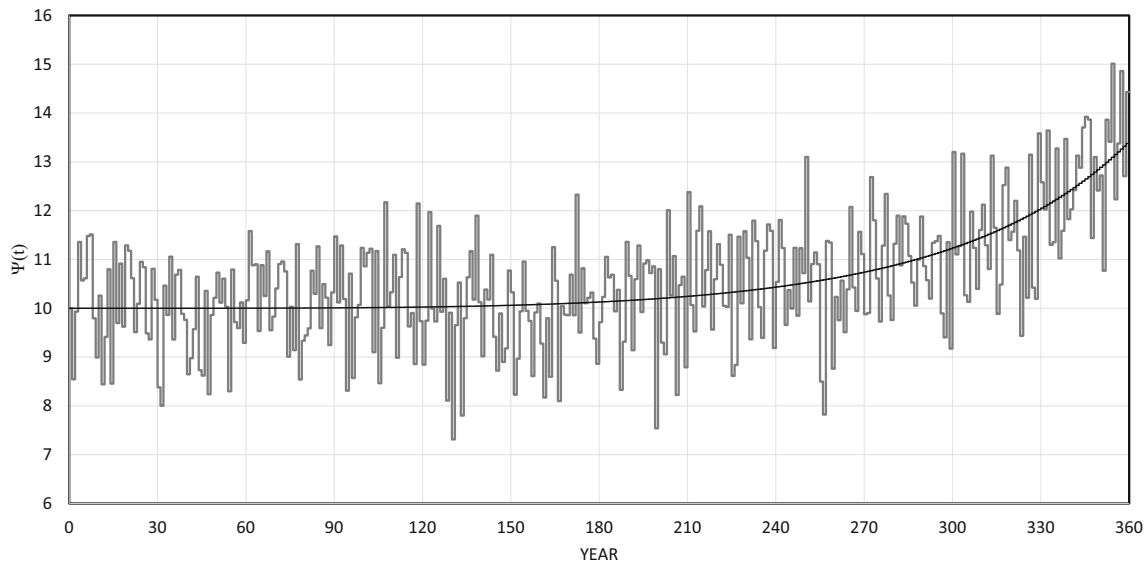
$$T(t) = T_0 + T_e(t) + T_n(t) = 8^\circ\text{C} + 2^\circ\text{C} \cdot \exp[t/360a]^4 + T_n(t) \tag{7}$$

An optimal smoothing filter should keep the signal (sine or exponential) un-modified and remove or damp the noise as much as possible.

To compare the different smoothing methods, we have to select adequate time windows and parameter settings. Figure 3 shows the filter response (weights) of a time series where all data points are set to zero  $T(t \neq 0) = 0$  except at the center, where the value is set to unity  $T(t = 0) = 1$ . The width of the response around the center point should be similar for all applied filters to allow a meaningful comparison. To fulfill this criterion, the smoothing window for MOVAS is set to 31 time increments, the standard deviation for GAUSS is set to 10, and the whole window spans over 61 time increments. LOWESS, LOESS, and LOWESS2 use a half-width of 25 time increments which results according to (1) also in a total smoothing window of 51. For SPECS and SPLIS, the whole domain is considered, for SPECS, the first 6 harmonics are used, and the smoothing parameter for SPLIS  $\gamma = 6 \cdot 10^{-4}$  is



**Fig. 1** Synthetic 360-day time series of temperature  $T(t)$  consisting of a pure sine function  $T_s(t)$  with an amplitude of  $10^\circ\text{C}$  (thin line) and a Gaussian random noise  $T_n(t)$  with a standard deviation of  $2^\circ\text{C}$ , according to Eq. 6



**Fig. 2** Synthetic 360-year time series of temperature  $T(t)$  consisting of an exponential function  $T_e(t)$  and a superimposed Gaussian noise  $T_n(t)$  with a standard deviation of 1 °C according to Eq. 7

chosen. The averaging window for MEVSS has been selected as 30 time increments. In Table 1, some statistical results are shown to compare the quality of the different smoothing methods. Four different smoothing intervals were chosen, 0.5, 1, 2, and 3 times the window settings of Fig. 1. For the sine function, periodic boundaries are used and for the exponential function, reduced windows are used; i.e., for MOVAS, GAUSS, LOWESS, LOESS, and LOWESS2, one-sided smoothed values are computed at both ends. For MEVSS, it is not necessary to compute the smoothed values for all possible time increments, because the results for time windows shifted by one time increment are quite similar. Only three different smoothing settings, shifted by 1/3 of the averaging window, were used here for the full domain and the resulting values have then been averaged. For the 30 time increment window of the periodic sine (left) and the non-periodic exponential function (right), only the mean values for the following windows were computed for the constraint (Eq. 3) of MEVSS:

Window setting	periodic sine	non – periodic exponential
1 <sup>st</sup>	$\sum_{t=1}^{30} T_t, \sum_{t=31}^{60} T_t, \dots, \sum_{t=331}^{360} T_t$	$\sum_{t=1}^{30} T_t, \sum_{t=31}^{60} T_t, \dots, \sum_{t=331}^{360} T_t$
2 <sup>nd</sup>	$\sum_{t=1}^{20} T_t + \sum_{t=351}^{360} T_t, \sum_{t=21}^{50} T_t, \dots, \sum_{t=321}^{350} T_t$	$\sum_{t=1}^{20} T_t, \sum_{t=21}^{50} T_t, \dots, \sum_{t=321}^{360} T_t$
3 <sup>rd</sup>	$\sum_{t=11}^{40} T_t, \sum_{t=41}^{70} T_t, \dots, \sum_{t=341}^{360} T_t + \sum_{t=1}^{10} T_t$	$\sum_{t=1}^{40} T_t, \sum_{t=41}^{70} T_t, \dots, \sum_{t=341}^{360} T_t$

(8)

This means that for an open domain at both ends, time windows with different lengths are being chosen.

It comes out clearly from Table 1 that for MOVAS, GAUSS, and LOWESS, the smoothing leads to a damping of the (sine) amplitude, especially for wide smoothing intervals. Hence, for the smoothing of diurnal or annual (periodic) time series, these methods are not recommendable. LOESS,

LOWESS2, SPECS, and MEVSS do much better in this perspective. The same is true for the sine fit, i.e., that the whole sine function is not much modified by the latter four smoothing methods. According to the damping of the noise, however, the latter methods behave differently. Although the noise generated for this test is only one possible realization, a comparison between the different smoothing algorithms seems to be nevertheless possible. Especially for short smoothing periods, LOESS and LOWESS2 do not filter the noise very well. It might be surprising that SPECS, which is conserving the sine function perfectly, does not filter the noise too well. This is due to random wave components, which are contained in the noise. For short smoothing windows, GAUSS and for long smoothing windows MEVSS are performing best. This can be understood, because GAUSS and MEVSS are using mean values, which are determined by setting the positive and negative deviations equal. LOESS, LOWESS, LOWESS2, SPECS and SPLIS use squared deviations (regression) and are hence more sensitive to extremes. The same is true for filtering the sine function plus noise. The interincremental RMSE, i.e., the day to day variation of the smoothed sine plus noise is much higher as compared to the pure sine especially for MOVAS, LOESS, and LOWESS2 with short averaging windows. For long windows, GAUSS and LOWESS even show a smaller RMSE than the pure sine. This is due to the significant damping of the sine function. For an open domain it comes out clearly, that the smoothing of the exponential function does not give reasonable values for the end of the time series by MOVAS and GAUSS. The offset becomes increasingly large for long smoothing intervals. If we consider the performance of the smoothing methods against noisy data at the ends of an open domain, it is surprising, how strong the noise—with an RMS of only 30% of the total exponential increase—changes the offset. The impact of noise

**Table 1** Comparison of the quality of different smoothing methods. All numbers are given in percent. Sine damping is defined as  $100 * \left[ 1 - \frac{(\tilde{T}_{s,max} - \tilde{T}_{s,min})}{(T_{s,max} - T_{s,min})} \right]$ , a value of < 1 means an amplitude error < 0.1 °C. Sine fit is defined as  $100 * \left[ 1 - \frac{RMSE(\tilde{T}_s(t) - T_s(t))}{0.5 * (T_{s,max} - T_{s,min})} \right]$ , a value of > 99 means an RMSE < 0.1 °C. Noise damping is defined as  $100 * \left[ 1 - \frac{RMSE(\tilde{T}(t) - T_s(t))}{RMSE(T(t) - T_s(t))} \right]$ , a value > 90 means a reduction of the RMSE of the noise to < 0.2 °C. Sine noise fit is defined as  $100 * \left[ 1 - \frac{RMSE(\tilde{T}(t) - T_s(t))}{0.5 * (T_{s,max} - T_{s,min})} \right]$ , a value of > 98 means an RMSE of  $(T_t - T_{s,t}) < 0.2$  °C. The interdiurnal RMSE is defined as  $100 * \left[ 1 - RMSE\left(\tilde{T}(t+1) \frac{\tilde{T}(t)}{RMSE(T_s(t+1))} - T_s(t)\right) \right]$ , a value of ± 8.0 means an RMSE of ± 0.01 °C with respect to the interdiurnal RMSE of the pure sine

function. Exp fit is defined as  $100 * \left[ 1 - \frac{RMSE(\tilde{T}_e(t) - T_e(t))}{(T_{e,max} - T_{e,min})} \right]$ , a value > 97 means an RMSE of < 0.1 °C. Exp + noise fit is defined as  $100 * \left[ 1 - \frac{RMSE(\tilde{T}(t) - T_e(t))}{(T_{e,max} - T_{e,min})} \right]$ , a value > 94 corresponds to an RMSE of  $(T(t) - T_e(t)) < 0.2$  °C. End offset is defined as  $100 * \tilde{T}(t=360) \frac{-T_e(t=360)}{(T_{e,max} - T_{e,min})}$ , a value within ± 7 means an error of the last data point less than ± 0.2 °C against the pure exponential function. End trend is defined as  $100 * \left[ 1 - \tilde{T}(t=360) \frac{-\tilde{T}(t=359)}{T_e(t=360)} - T_e(t=359) \right]$ , a value of ± 10 means an error of ± 0.2 °C/30 years. The interannual RMSE is defined as  $100 * \left[ 1 - RMSE\left(\tilde{T}(t+1) \frac{\tilde{T}(t)}{RMSE(T_e(t+1))} - T_e(t)\right) \right]$ , a value of ± 25.0 means an RMSE of ± 0.01 °C with respect to the interannual RMSE of the pure exponential function

Method, Parameter setting	360-day sine function (periodic)					360-year exponential function (open)				
	Sine damping	Sine fit	Noise Damping	Sine + noise fit	Interdiurnal RMSE	Exp fit	Exp + noise fit	End offset	End trend	Interannual RMSE
Acceptable range	<1.0	>99.0	>90.0	>98.0	-8.0 to 8.0	>97.0	>94.0	<10.0	-10.0 to 10.0	-25.0 to 25.0
MOVAS, window = 15 d/a	0.3	99.8	76.2	95.3	86.7	99.6	93.0	1.2	-82.4	605.3
GAUSS, σ = 5 d/a	0.4	99.7	78.3	95.7	10.7	99.4	93.5	6.9	-38.7	203.3
LOWESS, Δτ = 12 d/a	0.3	99.8	77.9	95.6	78.5	99.8	93.3	25.2	-5.2	112.0
LOESS, window = 25 d/a	0.0	100.0	72.9	94.6	84.3	99.1	92.3	19.2	-114.6	213.6
LOWESS2, Δτ = 12 d/a	0.0	100.0	68.7	93.7	64.4	99.1	90.1	9.3	0.8	123.6
SPECS, wave numbers 1-12	0.0	100.0	75.2	95.0	11.2	79.5	77.2	-42.0	-78.3	100.9
SPLIS, γ = 3.2*10 <sup>-3</sup>	0.0	100.0	77.1	95.4	11.7	100.0	93.1	22.8	37.2	119.6
MEVSS, window = 15 d/a	0.0	100.0	77.0	95.4	11.8	99.9	94.9	9.6	-12.7	43.6
MOVAS, window = 31 d/a	1.2	99.1	83.1	96.6	22.7	98.6	94.9	-23.1	-181.4	333.8
GAUSS, σ = 10 d/a	1.5	99.0	84.2	96.8	0.5	98.6	95.5	-11.4	-59.3	22.8
LOWESS, Δτ = 25 d/a	1.4	99.0	84.1	96.8	30.8	98.6	95.4	-11.4	-63.6	-31.4
LOESS, window = 51 d/a	0.0	100.0	80.4	96.1	24.7	99.4	94.0	21.0	109.9	41.5
LOWESS2, Δτ = 25 d/a	0.0	100.0	78.0	95.6	8.6	99.4	93.3	32.7	184.9	-0.8
SPECS, wave numbers 1-6	0.0	100.0	80.4	96.1	2.6	78.9	77.3	-36.0	-17.7	14.4
SPLIS, γ = 2.0*10 <sup>-4</sup>	0.0	100.0	82.9	96.6	2.1	99.9	94.9	18.0	11.9	35.5
MEVSS, window = 30 d/a	0.0	100.0	82.2	96.4	1.2	99.6	97.2	1.5	-8.8	7.4
MOVAS, window = 61 d/a	4.7	96.7	88.8	96.2	3.4	96.9	95.6	-40.2	-88.5	85.9
GAUSS, σ = 20 d/a	5.8	95.9	89.9	95.6	-4.2	96.7	96.1	-37.5	-67.5	-21.8
LOWESS, Δτ = 50 d/a	5.4	96.2	89.9	96.0	-4.6	98.9	96.6	17.1	6.9	16.3
LOESS, window = 101 d/a	0.2	99.9	85.5	97.1	6.8	99.4	95.4	29.1	46.3	-11.1
LOWESS2, Δτ = 50 d/a	0.1	100.0	83.5	96.7	1.7	99.4	95.1	18.3	-1.5	-36.1
SPECS, wave numbers 1-3	0.0	100.0	84.1	96.8	1.4	78.1	77.0	-57.9	7.1	-29.3
SPLIS, γ = 1.2*10 <sup>-5</sup>	0.8	99.5	89.2	97.8	0.0	99.3	96.7	3.3	-13.8	8.1
	0.0	100.0	90.9	98.2	0.8	98.1	97.2	-21.9	-47.4	-16.2



**Table 1** (continued)

Method, Parameter setting	360-day sine function (periodic)					360-year exponential function (open)				
	Sine damping	Sine fit	Noise Damping	Sine + noise fit	Interdiurnal RMSE	Exp fit	Exp + noise fit	End offset	End trend	Interannual RMSE
MEVSS, window = 60 d/a										
MOVAS, window = 91 d/a	10.2	92.8	<b>93.1</b>	92.9	<b>-6.2</b>	94.9	<b>96.1</b>	-64.5	-10.1	27.5
GAUSS, $\sigma = 30$ d/a	12.5	91.1	<b>93.9</b>	91.3	-12.1	94.8	93.5	-60.3	-73.1	-45.2
LOWESS, $\Delta\tau = 75$ d/a	11.8	91.7	<b>94.2</b>	91.8	-11.3	<b>98.1</b>	<b>97.3</b>	<b>6.0</b>	<b>-2.6</b>	<b>2.4</b>
LOESS, window = 151 d/a	1.0	<b>99.3</b>	<b>90.1</b>	98.0	<b>2.4</b>	<b>99.9</b>	<b>96.4</b>	19.8	-20.9	-27.5
LOWESS2, $\Delta\tau = 75$ d/a	<b>0.3</b>	<b>99.8</b>	86.8	97.4	<b>0.6</b>	<b>99.4</b>	<b>95.6</b>	24.6	43.2	-36.7
SPECS wave numbers 1-2	<b>0.0</b>	<b>100.0</b>	89.5	97.9	<b>0.9</b>	77.2	76.2	-64.5	18.3	-40.0
SPLIS, $\gamma = 2.5 \cdot 10^{-6}$	3.7	97.4	<b>94.0</b>	97.4	<b>-3.3</b>	<b>98.2</b>	<b>98.2</b>	-13.8	-33.6	<b>-10.8</b>
MEVSS, window = 90 d/a	<b>0.1</b>	<b>99.9</b>	<b>95.9</b>	<b>99.2</b>	<b>0.0</b>	96.0	<b>96.0</b>	-49.5	-67.1	-36.1

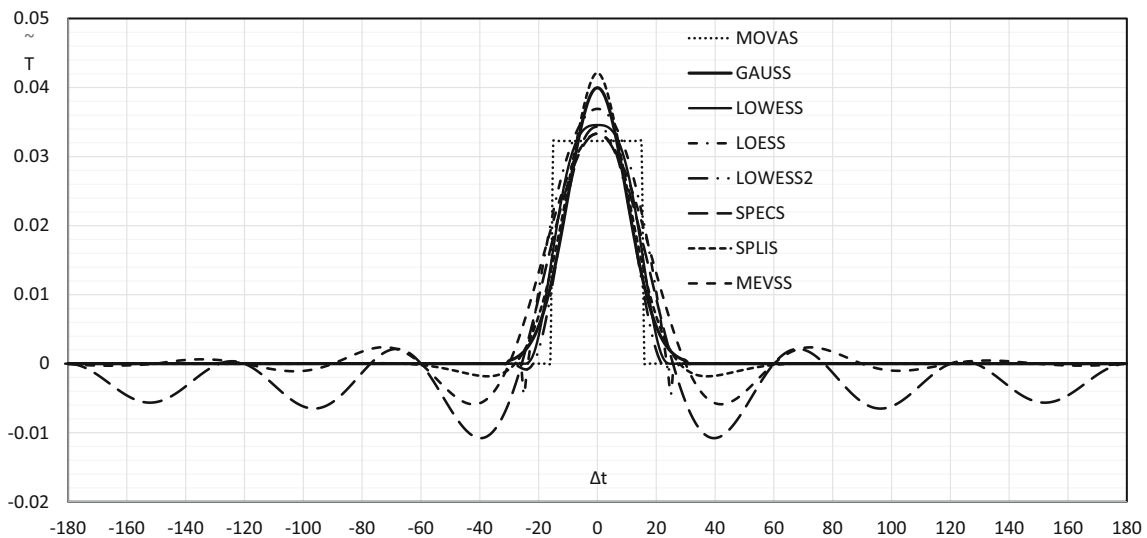
Bold means that the value lies within the acceptable range

seems to be least for the MEVSS method especially for medium smoothing windows. The trend at the end of an open domain, i.e., the difference between the last and the second last date, is reasonably captured only by LOESS, LOWESS, LOWESS2, and by MEVSS for short to medium averaging intervals. Noisy data do not allow to get reasonable estimates for the end trend of a time series, except for short to medium smoothing windows by MEVSS or for long windows by LOWESS2. For wide smoothing intervals and periodic domains, MEVSS outperforms all other methods because it smooths the noise best and

keeps the signal rather unchanged. Hence, we have chosen MEVSS for smoothing the long-term annual temperature time series and for the mean daily temperature series.

### 3 Application of MEVSS to long-term time series

The 244-year time series of annual mean temperatures of Vienna (see Fig. 4) has been smoothed with the 30-year



**Fig. 3** Weights (spectral response) for different smoothing methods of a time series with unity value at time  $t = 0$  and zero at all other times. The smoothing time window for MOVAS is 31 years, for GAUSS 60 years (with  $\sigma = 10$  years) and for LOWESS, LOESS and LOWESS2 51 years. For SPECS only the first 6 harmonics are considered and the smoothing

parameter for SPLIS has been set to  $\gamma = 6 \cdot 10^{-4}$ . The smoothing time window for MEVSS is 30 years. The sum of all weights is 1 and the setting was selected such that the standard deviation of weights is as close as possible at 10 years for all methods

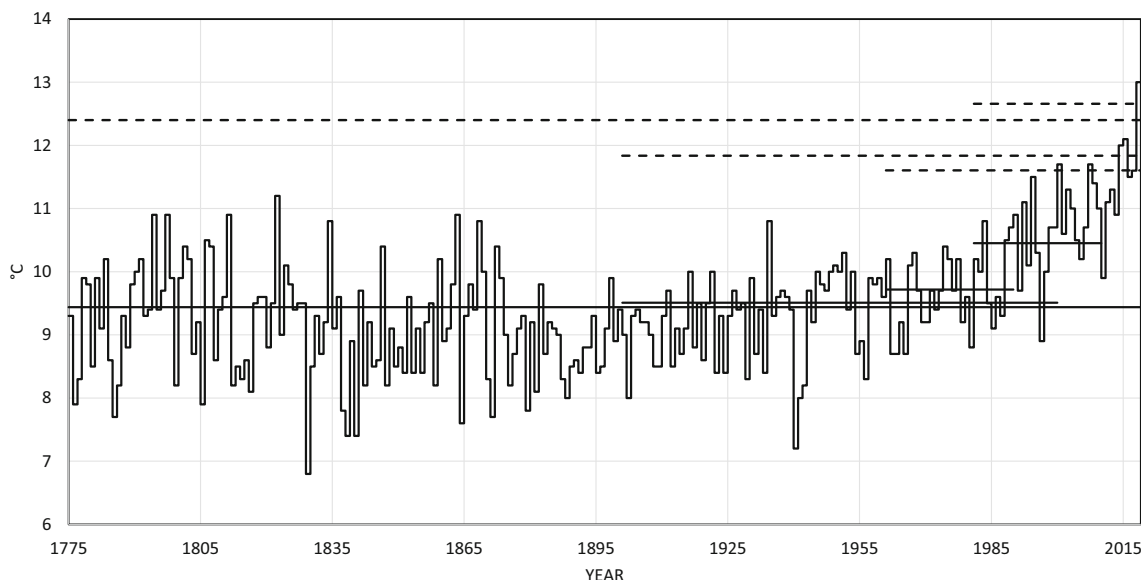
MEVSS method (see Fig. 5). In addition, the 25-year (half-width) LOWESS smoothing has been carried out, which agrees to less than 0.1 °C to the MEVSS curve except at the very beginning and end of the series. Due to the very high value in 2018, the LOWESS curve ends with a 0.26 °C higher value at the end and due to the two low values in 1776 and 1777 it begins with a 0.17 °C lower value. A more conservative value at the beginning and the end of an open time series, like produced by MEVSS, seems to be the preferable way to avoid an exaggerated trend.

If we compute the mean temperature of all climate normal periods 1781–1810, 1811–1840, ... and ending with the only 28-year period of 1991–2018 (Table 2), we see that until 1931–1960, all values were close to 9 °C. 1961–1990 and especially 1991–2018 show an increasing mean temperature due to global warming. Looking at the SDEV of the periods, we see somewhat larger values up to 1870 and in the recent period 1990–2018. Does that mean the variability of annual mean temperatures was higher (more extreme) in the early and recent years of the series? Not necessarily, because to determine the true climatic SDEV, we need to de-trend time series. If we do so and take the deviations of the annual mean temperature from the smooth MEVSS curve (lower part of Fig. 5) and compute the root mean square differences (RMSD (MEVSS), Table 2), we see that up to 1960, they lie very close to the SDEV values. This means, we do not need to de-trend time series when trends are small. The increasing difference between SDEV and RMSD (MEVSS) from 1960 onwards tells us however, that during periods of significant trends we need to de-trend the series to get a realistic climatic measure of

dispersion. Instead of taking the SDEV directly, RMSD (MEVSS) can be called de-trended SDEV and is hence the measure of choice. After de-trending, only during the first three CLNPs the variability of annual mean temperatures stays somewhat higher (around 0.9 °C) than after 1870 with roughly 0.6 °C. It is interesting to note that the station location of the Zentralanstalt für Meteorologie und Geodynamik in Vienna has changed in 1872 to its present location (Hammerl 2001). Could there be a discontinuity in the interannual temperature variation before and after the change of the station location, although the temperature series itself was carefully homogenized (Auer et al. 2001)? Climatic temperature time series and trends of the nineteenth century and earlier should be treated carefully in any case, because instrumental biases were in the order of 0.5 to 1 °C (Middleton 1966; Bernhard 2004; Winkler 2009).

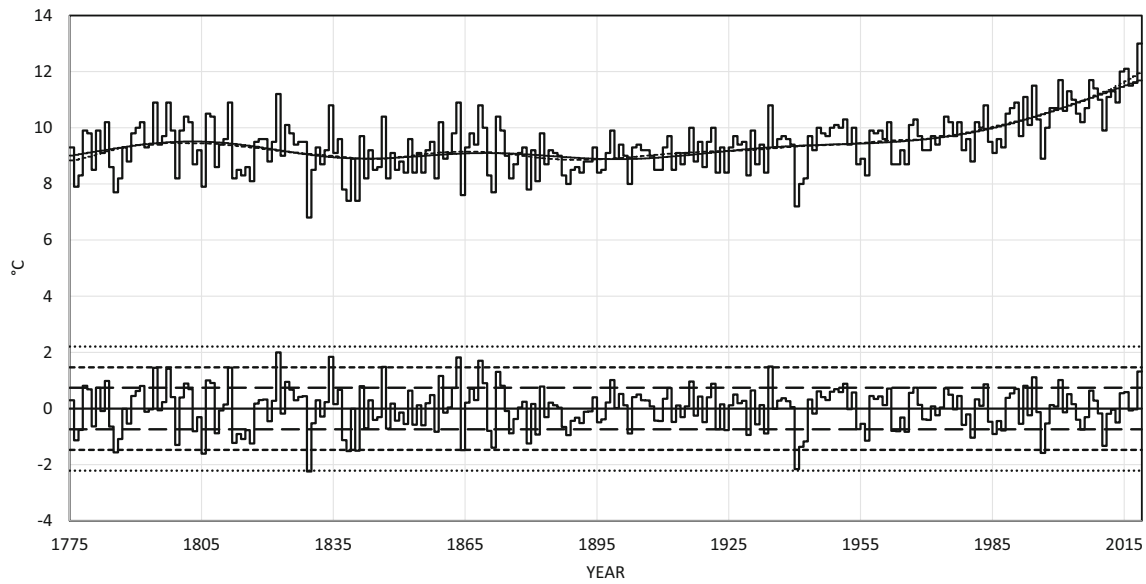
When we use statistics of past CLNPs to relate it to actual observations—what is normally done—measures of dispersion may become meaningless. Table 2 shows an enormous increase of RMSD (prev CLNP) above RMSD (MEVSS) in the period 1990–2018 by a factor of more than two (from 0.66 to 1.48). Now it becomes clear, why the 2018 mean temperature of 13 °C, related to the statistics of the previous CLNP 1961–1990 is an outlier with more than four times the SDEV above the mean whereas it stays below two times the RMSD (MEVSS), which is still rather normal.

We can conclude, that not the 2018 mean temperature was so extreme according to the present climate but the increase of the trend of the climatological mean, which has reached almost 0.06 °C per year (or 1.8 °C per 30 years!), a value never



**Fig. 4** Time series of annual mean temperatures for Vienna, Hohe Warte, from 1775 to 2018. Mean values are plotted for the periods 1775–2018, 1901–2000, 1961–1990, and for 1981–2010 (continuous lines). Dashed

lines are the corresponding positive 3σ levels for the different periods, each extended until 2018



**Fig. 5** Annual mean temperature at Vienna, Hohe Warte, smoothed with 30-year MEVSS (top, continuous) and 51-year LOWESS (top, dashed). Deviation from the MEVSS-smoothed temperature (bottom) with  $\pm 1$  SDEV (dashed),  $\pm 2$  SDEV (short dashed), and  $\pm 3$  SDEV (dotted) lines

observed before in the 244-year temperature series of Vienna. The  $1.8^{\circ}\text{C}$  increase of the trend of 1988–2018 was more than 3 times the SDEV of all 215 30-year trends between 1776 and 2018.

A problem with the determination of the actual climatic mean temperature is, that we just know the observations of the last 15 years but we do not know the observations of the next 15 to come of the 30-year period centered to the actual year. How can we proof, that the determination of actual climate statistics by the endpoint of an actual smoothed temperature series is better than taking the mean from past periods? We can do at least a hindcast. For all years from 1875 to 2003, the end values of the MEVSS, LOESS, and LOWESS curves ( $T_{MEVSS}$ ,  $T_{LOESS}$ ,  $T_{LOWESS}$ ) based on a 100-year period have been computed; i.e., for 1875, the years 1776 to 1875 have been used to predict the climatic mean temperature for 1875 by the above smoothing methods. For 1876, the years 1777–1876 have been used and so forth. The values have then been

compared with the 30-year mean temperature, e.g.,

$$\bar{T}_{1875} = \frac{1}{30} \left( \frac{T_{1860} + T_{1890}}{2} + \sum_{t=1861}^{1889} T_t \right), \quad (9)$$

which would not have been known in 1875. Furthermore, a comparison with the last available CLNP in the respective year (for 1875 this would have been 1841–1870) has been carried out. Finally, a comparison with the corresponding MEVSS temperature  $\tilde{T}$  derived of the whole series from 1875 to 2018 was carried out. The results are shown in Table 3. On average, the endpoint-temperature of the smoothed MEVSS, LOESS, and LOWESS temperatures fit very well to the centered 30-year mean temperatures. The maximum differences, however, all dating from the last few years with the strong positive temperature trend, is with  $0.48$  for MEVSS and LOESS and with  $0.65^{\circ}\text{C}$  for LOWESS not really satisfying. If we take the standard methodology, to use the last available CLNP statistics, the result becomes much

**Table 2** Mean temperatures and measures of dispersion for climate normal periods (CLNP) in Vienna. For the last period, only the 28 years between 1991 and 2018 have been used. Besides the SDEV of the annual mean temperatures within the CLNP, the root mean square differences

between the observed annual temperatures and the respective 30-year MEVSS temperatures (RMSD (MEVSS)) and the root mean square differences between the observed annual temperatures and the mean temperature of the previous CLNP (RMSD (prev CLNP)) are shown

CLNP	1781–1810 °C	1811–1840	1841–1870	1871–1900	1901–1930	1931–1960	1961–1990	1991–2018
Mean	9.5	9.1	9.1	8.9	9.1	9.4	9.7	10.9
SDEV	0.84	1.00	0.81	0.63	0.52	0.77	0.63	0.83
RMSD (MEVSS)	0.83	0.99	0.80	0.64	0.51	0.76	0.57	0.66
RMSD (prev CLNP)		1.09	0.81	0.68	0.57	0.82	0.71	1.48



**Table 3** Basic statistics (mean difference, root mean square difference, and maximum absolute difference) of the differences between the endpoints of the MEVSS, LOESS, and LOWESS temperatures ( $T_{MEVSS}$ ,  $T_{LOESS}$ ,  $T_{LOWESS}$ ) and the observed 30-year mean temperatures  $\bar{T}$  (according to Eq. 9) centered at the endpoint for all 100-year periods

°C	$T_{MEVSS}-\bar{T}$	$T_{LOESS}-\bar{T}$	$T_{LOWESS}-\bar{T}$	$T_{CLNP}-\bar{T}$	$T_{MEVSS}-\tilde{T}$
Mean	0.0	0.0	0.0	-0.2	0.0
RMSD	0.22	0.22	0.23	0.29	0.18
Max abs diff	0.48	0.48	0.65	1.11	0.40

between 1776–1875 and 1904–2003.  $T_{CLNP}$  is the last completed climate normal period (1781–1810, 1811–1840, ..., 1961–1990) before the corresponding year,  $\tilde{T}$  is the MEVSS temperature for the corresponding year based on the whole temperature series from 1775 to 2018

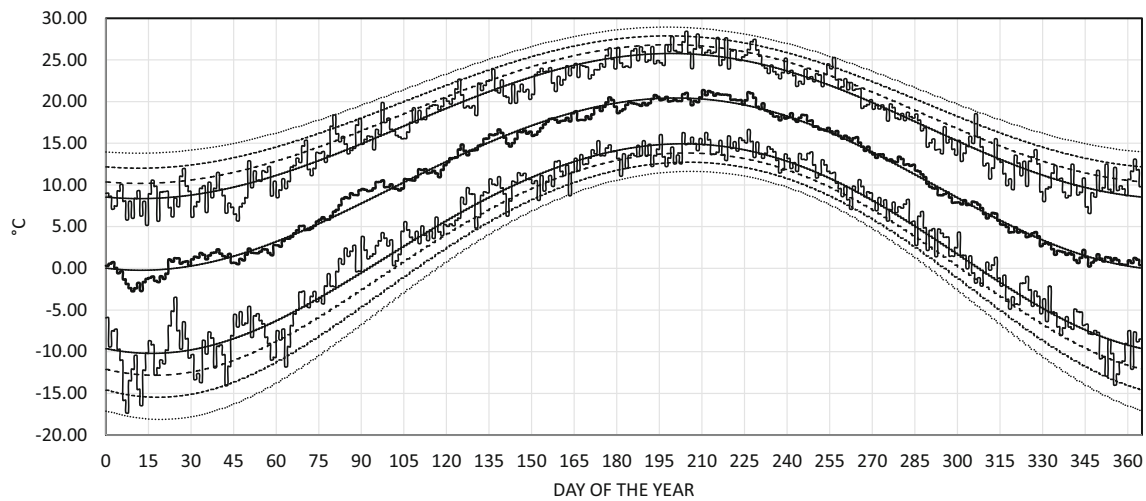
worse. Now even the mean difference is  $-0.2^\circ\text{C}$  and the maximum, again dating from one of the last few years, gives an intolerable  $1.11^\circ\text{C}$  difference. It is important to see that the difference between the MEVSS endpoint value and the temperature of the complete MEVSS series gives the best result. No mean difference and a  $0.18^\circ\text{C}$  RMSD with a maximum difference of  $0.4^\circ\text{C}$  is not perfect, but it is probably the best what we can do, to estimate the actual climate statistics.

The steep ascent of the climatological mean temperature has also consequences with respect to the judgment of monthly or daily temperature series. Often, the deviations of these temperatures are discussed in the public, using mean values of the last climatological normal period. In Fig. 6, besides the mean, maximum, and minimum daily mean temperatures and their MEVSS-smoothed curves also 1, 2, and 3 times SDEV of the maximum daily temperature with respect to the smoothed values are plotted above the maximum and below the minimum curves. It is important to note, that the smoothed curves are not pure sine functions, albeit they are based on only 4 (three-monthly) mean values. The mean time distance between the annual minimum (on January 12) and the maximum (on July 19) is 188 days, the cooling period is only lasting 177 days. Hence, also the maximum smoothed interdiurnal cooling rate in autumn ( $-0.187^\circ\text{C}/\text{day}$ ) is somewhat larger than the smoothed maximum warming rate ( $+0.169^\circ\text{C}/\text{day}$ ) in spring. If the distribution function is Gaussian, the mean equals the median and the smooth maxima/minima curves can be used to indicate quantiles and to derive return periods. An extended Wilk-Shapiro test (Shapiro and Wilk 1965; Royston 1982) has been applied to the daily extreme temperature distribution and a proof for a Gaussian distribution was found. Hence, the smooth curve of the maximum temperature indicates the 60-year return period, because on average each second day of the 30-year series lies above/below this curve. This means that for each individual day the probability of being higher/lower than the mean smoothed maximum/minimum curve is  $1/60$ . Statistically there should be one day within a 2-month period and roughly 6 days within a single year above/below the curves. Due to the

nearly Gaussian distribution, the maximum/minimum curves plus/minus 1, 2, and 3 SDEV can also be used to indicate other return periods. The 1/2/3 SDEV curves indicate the 189/1316/ and 23,095-year return periods. For the 1/2/3 SDEV curves there should be statistically 2/0.3/0.02 days within a year above/below the respective curves. For the 30 years of a climate normal period, the statistical numbers are 60/8/0.5 days for the 1/2/3 SDEV lines. The observed values for the period 1961–1990 are 51/15/0 days for the maximum and 53/15/1 for the minimum, which are quite close to the expectation.

As has been shown by the annual mean temperature time series with a  $2^\circ\text{C}$  increase of the mean over the last 40 years, the statement “above” or “below average” for a concrete current year should not be made with historic averages, unless we want to compare apples and oranges. They must be made with averages according to the actual climate. Let us take the daily mean temperature series of Vienna for 2018. If we take the MEVSS-smoothed annual curve, based on the last climate normal period (1961–1990, see Fig. 7), we come to the conclusion, that the 284 out of 365 days, having temperatures above the climatological mean 1961–1990, represent an exceptional high number of “warmer than average” days. The number of “colder than average” days (81 out of 365) is exceptional low. In Fig. 7, only the 1 in 100 years lines for the minimum and maximum are indicated, which equal the 1st and 99th percentile. The 1 in 100 years line is equal to the smoothed maximum/minimum line  $\pm 0.476$  SDEV according to a Gaussian distribution. In a whole year the statistically expected number of days, exceeding the 1 in 100 years maximum and minimum curves should be roughly 4 ( $0.01 \cdot 365$ ). The observed number of days for the maximum in Fig. 7, however, was 39, nearly 10 times as much as expected, an exceptional value. The observed number of days exceeding the 1 in 100 minimum line was 4. But this comparison is not adequate, we are comparing temperatures of the present climate with past climate data.

What can we do, to seriously compare the daily temperature series of 2018 with the actual climate? To get the smooth mean climatological temperature for 2018 we can first compute the three-monthly (Jan–Mar, Feb–Apr, ..., Dec–Feb)

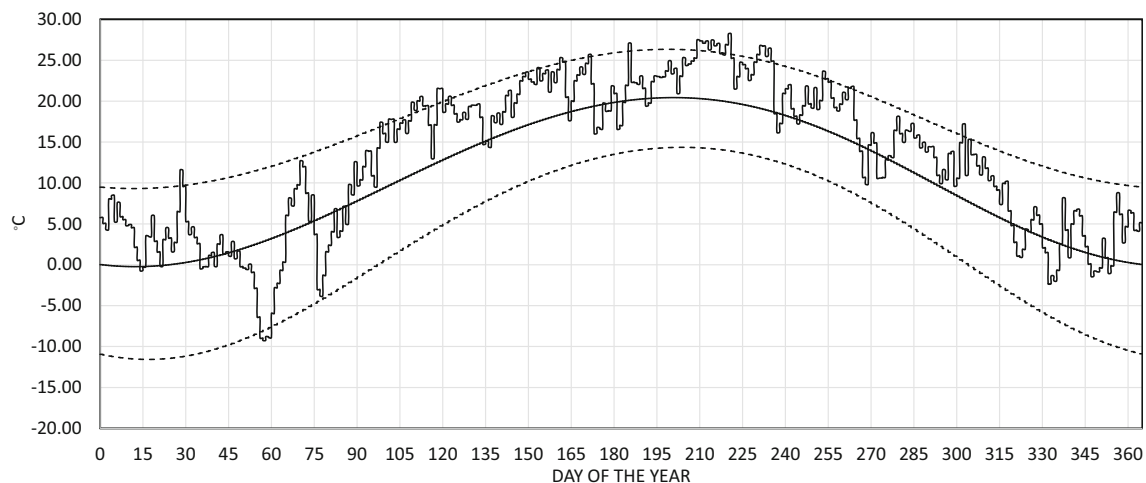


**Fig. 6** Mean (bold), maximum, and minimum daily mean temperatures of the climate normal period 1961–1990 with 3-month MEVSS (smooth continuous) in Vienna, Hohe Warte. The dashed lines show the smoothed

maximum/minimum  $\pm 1$  STDV, the short dashed lines the smoothed maximum/minimum  $\pm 2$  STDV, and the dotted lines the smoothed maximum/minimum  $\pm 3$  STDV

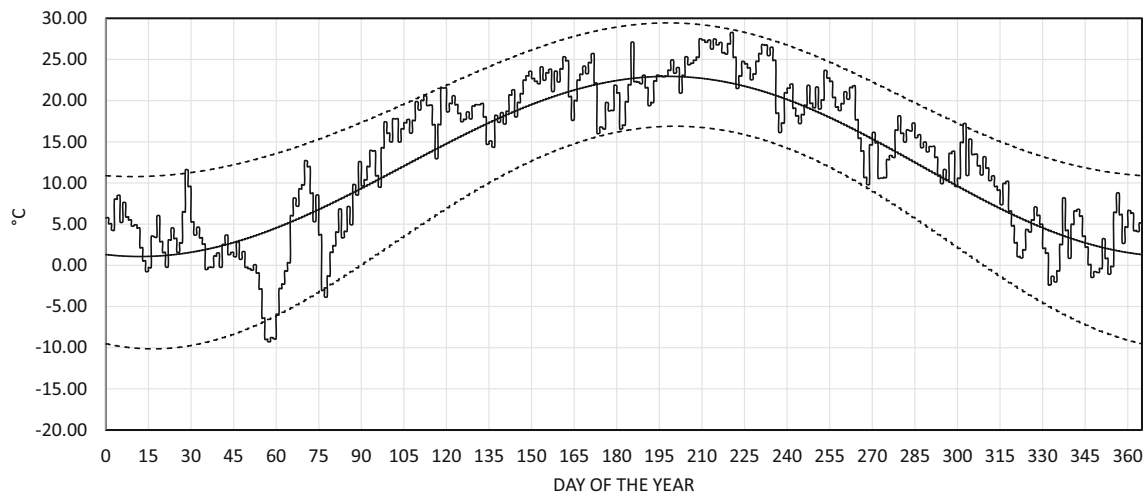
smooth temperature time series from 1863–2018, averaging the time series e.g. for the periods 1863–1880, 1881–1910, ..., 2001–2018, the periods 1863–1890, 1891–1920, ..., 1981–2018, and the periods 1863–1900, 1901–1930, ..., 1991–2018 with MEVSS. Then, we can further average the MEVSS-smoothed daily temperature series of 2018 for the periods Jan–Mar, Apr–June, ..., Oct–Dec, for the periods Feb–Apr, May–Jul, ..., Nov–Jan, and for the periods Mar–May, June–Aug, ..., Dec–Feb. The so generated smooth daily average temperature for the year 2018 (“climate of 2018”) is plotted in Fig. 8. Now, the number of days warmer than the 2018-climate daily average is reduced to 241 and colder than average is increased to 124. The still much larger number of positive anomalies is due to the fact, that the year 2018 was warmer by 1.5 °C than the climatological mean of 2018 (compare Fig. 4). The smooth 1 in 100 years curves for the

2018 climate cannot be computed in the same manner as for the smooth mean daily temperatures, because the values of the maximum and minimum time series depend on the length of the time series. Therefore, in Fig. 7, the 1 in 100 years curves of the smooth daily maxima and daily minima have been taken with the same distance from the smooth mean curve like for the 30-year time series 1989–2018. This can be justified by the fact that the smoothed mean amplitude between the minimum and maximum curves and the STDEV of the minima and maxima did not change very much between the last climate normal period 1961–1990 and 1989–2018. Now, the number of days above the 1 in 100 years maximum curve (99th percentile) in 2018 reduces to 3 and the number of days below the 1 in 100 years minimum curve (1st percentile) in 2018 is 5, both very close to the statistical expectation of 4 days. One interesting finding of the 2018 time series with the



**Fig. 7** Daily mean temperatures of the year 2018 in Vienna, Hohe Warte, and the 3-month smoothed MEVSS curve for the climate normal period 1961–1990. The dashed lines represent the 1 in 100 years curves for the

maximum (99th percentile) and minimum (1st percentile) daily mean temperatures 1961–1990. For further explanations, see text



**Fig. 8** Same as Fig. 7 but with mean and 1 in 100 years curves for the “climate of 2018”

smooth curves adjusted to the “climate of 2018” is that the most extreme event of the daily mean temperatures was not in the summer half year with the many warmer than average periods, but in winter, when the temperature at the end of February fell way below the 1 in 100 years curve. Again, the single days of 2018 were not exceptional compared to the actual climate but the increase of the mean values between the climate of 1961–1990 and 2018 was extreme, especially in summer by more than 2.5 °C. Whereas in the period 1961–1990, the climatological number of days with a mean temperature of 20°C or greater was 35; in the present climate (2018), this number has increased to 87 days. With a further temperature increase, the number of days with high temperatures will rise further sharply, because the probability to exceed a threshold will increase and the seasonal extent, when this can occur will also increase. With respect to future heat waves, this means an exponential rather than a linear increase.

## 4 Conclusion and outlook

In times of significant climate change, what we just experience, the comparative statistical climatology has to take into account trends of meteorological parameters. Considering trends means, that events, which would have been characterized as extreme events under a past climate, can be rather “normal” under present climatic conditions. Heatwaves, which have been experienced more often in the past summers, should not be called extreme events anymore, because they become more “normal” under present conditions. This would also improve the communication with the public, because the term “extreme events” is intuitively related to rare events. We must communicate that some extreme events of the past are becoming “normal events” nowadays. It is more than questionable, if the traditional way of using 30-year average values as “climate normals” should be continued. If the trend within

30 years is in the order of the STDEV or even higher, which is the case for the 1991–2020 CLNP, the climate of a year at the beginning and the end of this CLNP cannot be set equal. There is nothing like the “climate of 1991 to 2020”. The 30-year smoothing allows to define “climate normals” virtually for each year. In this paper, it was shown that in the frame of rapid global change this makes sense.

The methodology of MEVSS smoothing is seemingly the best way how to produce smooth climatological time series, because it smooths the short term (random) variations to a high degree, whereas the long-term variations are kept nearly undisturbed and better than with other smoothing methods. This is especially true for periodic time series like annual or daily variations. This methodology seems to allow also a wide application to other climate elements. An application to non-Gaussian distributed variables should be checked in the future.

Another concrete application of the MEVSS method, which by no other existing method may be achieved alike, is the refinement of the annual variation of an element, when only monthly mean- or extreme values are given. For many stations worldwide such information is easily available online (e.g., Weather-Online 2019). Instead of plotting an 11-month pseudo-annual temperature-“curve” by connecting the 12 monthly mean values by straight lines, the MEVSS smoothing gives a much better estimate for daily time series of the whole annual curve, including “true” minima and maxima.

**Acknowledgements** Thanks are due to the Zentralanstalt für Meteorologie und Geodynamik in Vienna for providing the daily and yearly temperature data for this paper.

**Availability of data and material** All data published in this article are available upon request except the raw daily and yearly data from Vienna, which have to be requested from the Zentralanstalt für Meteorologie und Geophysik, Vienna.

**Code availability** Most evaluations have been carried out by own Excel programs, a few by Matlab-programs, which are available upon request from the author.

**Funding** Open access funding provided by University of Vienna.

## Declarations

**Conflict of interest** The authors declare no competing interests.

**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

- Angel JR, Easterling WR, Kirtsch SW (1993) Towards defining appropriate averaging periods for climate normals. *Clim Bulletin* 27:29–44
- Arguez A, Yu P, O'Brien JJ (2008) A new method for time series filtering near endpoints. *J Atmos Ocean Technol* 25:534–546
- Auer I, Böhm R, Schöner W (2001) Austrian long-term climate 1767–2000. In: Multiple instrumental climate time series from Central Europe. Österreichische Beiträge zu Meteorologie und Geophysik 25. Zentralanstalt für Meteorologie und Geodynamik, Wien. <https://doi.org/10.1002/joc.754>
- Bernhard F (2004) Technische Temperaturmessung. Springer Vieweg. <https://doi.org/10.1007/978-3-642-24506-0>
- Box GEP, Jenkins GM (1970) Time series analysis, forecasting and control. Cambridge University Press, Oakland
- Cleveland WS (1981) LOWESS: A program for smoothing scatterplots by robust locally weighted regression. *Am Stat* 35(1):54. <https://doi.org/10.2307/2683591>. JSTOR2683591
- Hammerl C (2001) Die Zentralanstalt für Meteorologie und Geodynamik 1851–2001: 150 Jahre Meteorologie und Geophysik in Österreich. Leykam, Graz
- Hamming RW (1989) Digital Filters, 3rd edn. Hall, Prentice 284 pp
- Hiebl J, B Chimani, M Ganekind, A Orlik (2019) Österreichisches Klimabulletin 2018. Zentralanstalt für Meteorologie und Geodynamik, Wien. <https://www.zamg.ac.at/zamgWeb/klima/bulletin/2018/bulletin-2018.pdf>, accessed August 1, 2019
- Livezey RE, Vinnikov KY, Timofeyeva MM, Tinker R, van den Dool HM (2007) Estimation and extrapolation of climate normals and climatic trends. *J Appl Meteorol Climatol* 46:1759–1776
- Mann ME (2004) On smoothing potentially nonstationary climate time series. *Geophys Res Lett* 31:L07214. <https://doi.org/10.1029/2004GL019569>
- Mann ME (2008) Smoothing of climate time series revisited. *Geophys Res Lett* 35:L16708. <https://doi.org/10.1029/2008GL034716>
- Middleton WEK (1966) A history of the thermometer and its use in meteorology. Johns Hopkins Press, Baltimore
- Royston JP (1982) An extension of Shapiro and Wilk's W test for normality to large samples. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 31(2):115–124
- Sasaki Y (1955): A fundamental study of the numerical prediction based on the variational principle, *Journal of the Meteorological Society of Japan, Ser. 2, Vol. 33, No. 6,*
- Savitzky A, Golay MJE (1964) Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry* 36(8):1627–1639. <https://doi.org/10.1021/ac60214a047>
- Schönwiese CD (2013) Praktische Statistik für Meteorologen und Geowissenschaftler. 5. vollständig überarbeitete und erweiterte Auflage. Bornträger
- Shapiro SS, Wilk MB (1965) An analysis of variance test for normality (complete samples). *Biometrika* 52(3–4):591–611. <https://doi.org/10.1093/biomet/52.3-4.591>
- Shumway R, Stoffer D (2006) Time series analysis and its applications: with R examples. Springer, New York
- Weather-Online (2019) Worldwide Climate Diagrams: <https://www.weatheronline.co.uk/reports/climate/Diagrams.htm>, accessed August 1, 2019
- Winkler P (2009) Revision and necessary correction of the long-term temperature series of Hohenpeissenberg, 1781–2006. *Theor Appl Climatol* 98:259–268. <https://doi.org/10.1007/s00704-009-0108-y>
- WMO (1989) Calculation of monthly and annual 30-year standard normals. WCDP-No. 10, WMO-TD/No. 341, World Meteorological Organization, Geneva
- WMO (2007) The role of climatological normals in a changing climate. WCDMP-No. 61, WMO-TD/No. 1377, World Meteorological Organization, Geneva
- ZAMG (2019) Histalp Jahresbericht 2018, Wien, [https://www.zamg.ac.at/cms/de/dokumente/klima/dok\\_news/dok\\_histalp/jahresbericht-2018/bericht](https://www.zamg.ac.at/cms/de/dokumente/klima/dok_news/dok_histalp/jahresbericht-2018/bericht), accessed August 1, 2019

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