Stochastic investigation of daily air temperature extremes from a global ground station network

Konstantinos-Georgios Glynis, Theano Iliopoulou, Panayiotis Dimitriadis and Demetris Koutsoyiannis

Department of Water Resources and Environmental Engineering, School of Civil Engineering, National Technical University of Athens, Heroon Polytechneiou 5, 15880 Zographou, Greece

Abstract

Near-surface air temperature is one of the most widely studied hydroclimatic variables, as both its regular and extremal behaviors are of paramount importance to human life. Following the global warming observed in the past decades and the advent of the anthropogenic climate change debate, interest in temperature’s variability and extremes has been rising. It has since become clear that it is imperative not only to identify the exact shape of the temperature’s distribution tails, but also to understand their temporal evolution. Here, we investigate the stochastic behavior of near-surface air temperature using the newly developed estimation tool of Knowable (K-)moments. K-moments, because of their property to substitute higher-order deviations from the mean with the distribution function, enable reliable estimation and an effective alternative to order statistics and, particularly for the outliers-prone distribution tails. We compile a large set of daily timeseries (30 to 200 years) of average, maximum and minimum air temperature, which we standardize with respect to the monthly variability of each record. Our focus is placed on the maximum and minimum temperatures, because they are more reliably measured than the average, yet very rarely analyzed in the literature. We examine segments of each timeseries using consecutive rolling 30-year periods, from which we extract extreme values corresponding to specific return period levels. Results suggest that the average and minimum temperature tend to increase, while overall the maximum temperature is slightly decreasing. Furthermore, we model the temperature timeseries as a filtered Hurst-Kolmogorov process and use Monte Carlo simulation to produce synthetic records with similar stochastic properties through the explicit Symmetric Moving Average scheme. We subsequently evaluate how the patterns observed in the longest records can be reproduced by the synthetic series.

Key words: Stochastics; near-surface air temperature; Extreme temperature; Symmetric Moving Average; Hurst-Kolmogorov dynamics; Monte-Carlo simulation.

1. Introduction

Air temperature is one of the most important hydroclimatic variables and, together with precipitation, it can characterize the climate conditions in a region (e.g., Köppen-Geiger climate classification system, Rubel and Kottek, 2010). During the last decades, global warming, its possible anthropogenic origin and its effects on the environment have been recognized as matters of great political, economic and scientific importance. It has been asserted that, due to the vulnerability of infrastructure, the ecosystem and the entire system of food and energy harvesting, slight disturbances in the very delicate climatic conditions can cause significant problems (Handmer et al., 2012). For this reason, it is imperative to understand, not only the evolution of the average near-surface air temperature, but the changes in its maximum and minimum values as well. These are less studied than the average, yet they are more reliably estimated.

The understanding of the temporal evolution of near-surface air temperature, in terms of the spatial distribution, is also helpful in our effort to recognize the basic drivers of climatic processes, and if and how we can mitigate their negative effects. There are many climatic factors, both internal and external. The internal variability of the Earth's climate includes factors such as the ocean-atmosphere variability (Brown et al., 2015; Hasselmann, 1976), as well as the effects of the biosphere, through the carbon and water cycles. In the external factors we consider drivers such as greenhouse gases (Cronin, 2009), orbital variations, solar activity and volcanic activity. The identification of the exact near-surface air temperature changes, within an interdisciplinary approach, may facilitate the quest of deciphering of the Earth's climate mechanisms.

The aim of the present paper is to identify, based on observations, the temporal evolution of the extremes of near-surface air temperature, i.e., the upper and lower tails of the average temperature, the upper tail of the maximum temperature and the lower tail of the minimum temperature, and to stochastically evaluate the magnitude of the observed changes. To this aim, we study changes of near-surface air temperature both in past and present, and investigate whether these changes fall into the expected range of the formulated stochastic framework of global climatic variations.

Multiple scientific studies have shown that the global average air temperature has increased substantially during the twentieth century (Trenberth et al., 2007; Jones et al., 2012; Sun et al., 2017; Masson-Delmotte et al., 2018). According to the Summary for Policymakers of the Fifth Assessment Report of IPCC (Masson-Delmotte et al., 2018), the 2009-2018 decade was warmer by 0.93 ± 0.07 °C, compared to the pre-industrial baseline (1850-1900). Despite observing a slight deviation between the urban and rural meteorological records (Peterson et al., 1999), the general air temperature trend seems to be increasing, as presented in Figure 1 from data of the Climate Research Unit.

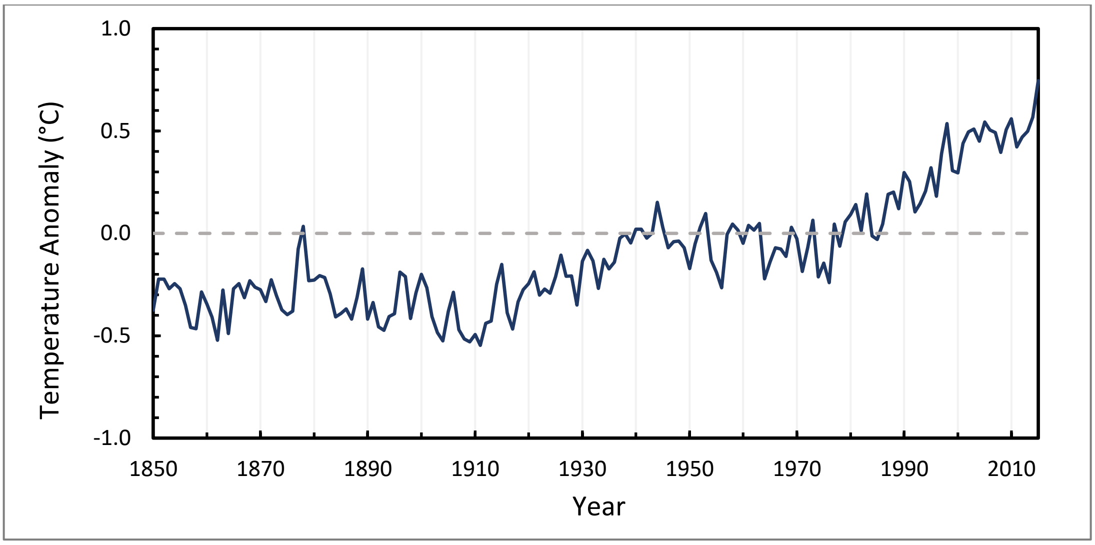

Figure 1: Global annual air temperature anomalies (°C) for the period 1850-2015, relative to the 1961–1990 climatology mean | Source: Jones et al. (2016)

It has been observed that the daily minimum air temperature tends to increase at a faster rate than that of the maximum air temperature (Braganza et al., 2004). As a result of this differential trend behavior, the diurnal temperature range decreases in most areas of the world. According to Easterling et al. (1997) the diurnal temperature range decreases at a rate of about 0.1 °C/decade.

However, several studies of the global near-surface air temperature omit to include into their premises the inherent region-specific seasonal variability of the air temperature. In the present paper, we account for seasonal variability by standardizing the daily air temperature records with respect to each month in order to assess the degree of global variability taking into account the local behavior. Consequently, persistent, yet statistically expected, record entries at certain climatic regions do not skew the general trend significantly.

In addition, we use the Knowable (K)-moments, a variant of probability weighted moments, which are particularly robust, especially in the study of extremes (Koutsoyiannis, 2019a). One of the most important benefits that their use guarantees is that they are knowable even for very high orders, with unbiased estimators. Specifically, their estimation uncertainty is smaller by orders of magnitude (compared to the classical moments) enabling more accurate estimation. Furthermore, the estimators can take into account any existing dependence structure, while, in addition, we can instantly assign return periods to them, as with the use of order-statistics.

In the following section, we introduce the basic tools and theory behind our research. After that, we present the data used and the methodology we follow. Finally, we conclude our findings and discuss how these can be expected from the stochastic viewpoint.

2. Basic tools

2.1 Climacogram

A widely used metric for estimating the second-order properties, including persistence of a stochastic process is the quantification and visualization of the variance of the averaged process vs. scale, else called the climacogram (Koutsoyiannis, 2010). The averaged stochastic process $\underline{z}$ is expressed as:

$$\underline{z}_i^{(k)} = \frac{1}{k} \sum_{j=(i-1)k+1}^{ik} \underline{z}_j \tag{1}$$

where $\underline{z}_i^{(k)}$ is i_{th} element of the averaged stochastic process at scale *k*.

A widely used climacogram estimator, based on the second central moment, can be expressed as:

$$\underline{\hat{\gamma}}(k) = \frac{1}{\lfloor n/k \rfloor - 1} \sum_{i=1}^{\lfloor n/k \rfloor} \left(\underline{z}_i^{(k)} - \overline{\underline{z}} \right)^2 \tag{2}$$

where *n* is the length of the timeseries, $\lfloor n/k \rfloor$ is the integer of *n*/*k*, and $\overline{\underline{z}} = \sum_{l=1}^{n} \underline{z}_l / n$ the unbiased estimator of the mean, μ, of the process.

The quantification of the persistence of a process (or else long-range dependence or long-range change or clustering) can be quantified through the Hurst parameter, *H*, which equals the half of the log-log slope of the climacogram, as scale tends to infinity, plus 1. Depending of the value of the Hurst parameter, a behavioral pattern can be attributed to the studied process. For sufficiently large scales, if $0 \leq H \leq 0.5$ then the process can be characterized as anti-correlated, and if $0.5 \leq H \leq 1$, then the process is positively correlated, which is the most common behavior in geophysical processes, while for $H = 0.5$, the process is purely random (i.e. zero autocorrelation; hence white noise behavior).

A stochastic Gaussian process with persistent behavior is known as Fractional Gaussian noise (fGn; Mandelbrot and Van Ness, 1968) or ARIMA (e.g. Montanari et al., 1997). Specifically, fGn can be defined in discrete time, which is the scope here, in a manner similar to that used in continuous time. It can be defined as a Gaussian process satisfying the condition between the average processes at two scales *k* and *l*:

$$\left(\underline{z}_i^{(k)} - \mu \right) = \left(\frac{k}{l} \right)^{H-1} \left(\underline{z}_j^{(l)} - \mu \right) \tag{3}$$

which is applicable only in (finite-dimensional joint) distribution.

By setting $i = j = l = 1$ in equation (3), it can be obtained (e.g. Koutsoyiannis, 2002):

$$\gamma(k) = k^{2-2H} \gamma(1) \tag{4}$$

This simple equation serves as the basis for estimating the Hurst parameter, since the variance of the average stochastic process at scale k is a power law of k with exponent $2-2H$.

The climacogram presents several advantages, as a stochastic metric, in the identification of both the short-term and the long-term persistent behaviour of a process, as compared to the autocovariance and the power-spectrum, largely because of its simplicity, link to entropy, and statistically more robust estimation properties including bias (Dimitriadis and Koutsoyiannis, 2015; 2019). The latter is of great importance in model identification and fitting from data, which is one of the purposes of this work.

2.2 Hurst-Kolmogorov dynamics

This long-term persistent behavior is also known as the Hurst phenomenon, and is a much-studied subject in engineering and mathematics. Hurst (1951) was the first to identify long-term persistence in natural processes, and specifically in the maximum annual stage of the river Nile. Kolmogorov (1940) was the first who mathematically described it a few years earlier, while working on self-similar processes of turbulent fields (Koutsoyiannis, 2011). To include both contributions this behaviour is also known as Hurst-Kolmogorov (HK) behaviour (Koutsoyiannis, 2010), and it has been expanded to include both the short-term fractal behaviour (Gneiting and Schlather, 2004) and the intermediate-scale behaviour (Koutsoyiannis, 2020), and thus, to express a generalized multi-scale behaviour of the second-order dependence structure. HK dynamics have been observed in various global-scale hydrometeorological and high-resolution turbulent processes (e.g. for a review see O'Connell et al., 2016 and for global-scale applications see Dimitriadis, 2017), including extremes (e.g. Iliopoulou and Koutsoyiannis, 2019), as well as in alternate fields, e.g. rock formations (Dimitriadis et al., 2019), landscapes (Sargentis et al., 2019) and art (Sargentis et al., 2020).

The observation of the empirical climacogram constructed from temperature records in fine scales (e.g. hourly or daily), brought to the surface a considerable divergence from the large scales in the area of small scales (Koutsoyiannis et al., 2018). Hence, a more generalized model of the Filtered Hurst-Kolmogorov (FHK) process is used here, which also is shown to maximize entropy production both at small- and large-time scales. The equation of the Filtered Hurst-Kolmogorov (FHK) model (mixed Cauchy-Dagum type) is (Koutsoyiannis, 2017):

$$\gamma(k) = \lambda_1 \left(1 + \left({}^{k}/_{a_1}\right)^2\right)^{H-1} + \lambda_2 \left(1 - \left(1 + \left({}^{k}/_{a_2}\right)^{-2}\right)^{-M}\right) \qquad (5)$$

where $\gamma(k)$ is now the variance of the average process instead of the standard deviation.

The parameter M (in honor of Mandelbrot) is called the smoothness (or fractal) parameter, while H is the Hurst parameter. Both parameters H and M are dimensionless parameters, bounded between zero and one inclusively, while α and λ are scale parameters, with dimensions $[t]$ and $[x^2]$. This form of the modeled climacogram has the advantage of determining the persistence of the process through the first additive term, as well as its smoothness through the second additive term.

2.3 Return period estimation through K-moments

As mentioned in the introduction, K-moments are a fundamental part of metrics we apply in this work. We present basic information about K-moments in this section, while providing extensive background in the relevant appendix.

Let $\underline{x}$ be a stochastic variable and $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_p$ be copies of it, independent and identically distributed, forming a sample, while $F(x)$ is the distribution function of $\underline{x}$.

According to Koutsoyiannis (2020), if we denote the estimator of $\left((F(\underline{x})\right)^{p-1}$ from a random sample of size n as b_{inp} (*i* is the index ranging from 1 to *p*) then an estimator of the noncentral moment K'_{pq} will be:

$$\hat{\underline{K}}'_{pq} = \sum_{i=1}^{n} b_{i,n,p-q+1} \underline{x}_{(i)}^{q} \quad (6)$$

where $\underline{x}_{(i)}$ is the ith element of a sample of $\underline{x}$ of size n, sorted in ascending order; it is stressed that the ordering of the sample is meant in terms of $\underline{x}$ and not $\underline{x}^q$. More precisely, $\underline{x}_{(i)}^{q} := \left(\underline{x}_{(i)}\right)^q$ which can be different from $\left(\underline{x}^q\right)_{(i)}$. The estimator in (18) is unbiased if we choose:

$$b_{inp} = \begin{cases} 0, & i < p \\ \dfrac{p}{n}\dfrac{\Gamma(n-p+1)}{\Gamma(n)}\dfrac{\Gamma(i)}{\Gamma(i-p+1)}, & i \geq p \geq 0 \end{cases} \quad (7)$$

where p can be any positive number (usually, but not necessarily, integer). It is easy to verify that:

$$\sum_{i=1}^{n} b_{inp} = 1 \quad (8)$$

which is a necessary condition for unbiasedness (Koutsoyiannis, 2020).

K-moments constitute an important statistical tool, as does the notion of return period. The return period refers to a time span in which an event (e.g. an extreme one) is expected to happen, and in that way, it is used to associate event occurrences to the likeliness of them happening.

As it can be easily understood, order statistics have a substantial advantage over other statistics in the context of return periods, as we can assign a distinct value of the distribution function to each one of them, hence pair them with the equivalent return period. This turns out to be the case with K-moments as well, since they are closely related to order statistics. Intuitively, we anticipate that the return period corresponding to the non-central K-moment of orders $(p, 1)$, the value $x = K'_{p1}$ will correspond to a return period of about $2p$. This is accurate for a symmetric distribution and for $p = 1$, as K'_{11} is the mean value, which has return period 2, and as explained by Koutsoyiannis (2019a), it cannot be much lower than $2p$ for any p and for any distribution.

Generally, the return period can be expressed by the relationship:

$$\frac{T(K'_{p1})}{D} = \Lambda_p p \tag{9}$$

where D is a time reference for the specification of return period and Λ_p is a coefficient generally depending on the distribution function and the order p.

The precise definition of Λ_p is (Koutsoyiannis, 2019a):

$$\Lambda_p := \frac{1}{p\left(1 - F(K'_{p1})\right)} \tag{10}$$

For given p and distribution function $F(x)$, K'_{p1} is analytically or numerically determined from its definition. Then $T(K'_{p1})$ and Λ_p are determined from their definitions.

In absence of an analytical solution, an exact relationship between p and T has been established by doing numerical calculations for several p. The slight variation of Λ_p with p can be very well approximated if first the specific values Λ_1 and Λ_∞ are accurately determined. The value of Λ_1 is easily determined, as practically is equal to the return period of the mean:

$$\Lambda_1 = \frac{1}{1 - F(\mu)} = \frac{T(\mu)}{D} \tag{11}$$

Developed within extreme value theory, the Generalized Extreme Value distribution is a family of continuous probability distributions, that includes the Extreme Value Type 1 distribution. In a number of customary distributions, specifically those belonging to the domain of the Extreme Value Type 1 distribution, Λ_∞ has a constant value, independent of the distribution. As shown by Koutsoyiannis (2019a), this value is:

$$\Lambda_\infty = e^\gamma = 1.781 \tag{12}$$

where γ is the Euler–Mascheroni constant.

For the approximation of Λ_p, the following simple relationship is used, which is satisfactory for several distributions:

$$\Lambda_p \approx \Lambda_\infty + (\Lambda_1 - \Lambda_\infty)\frac{1}{p} \tag{13}$$

This yields a linear relationship between the return period T and p:

$$\frac{T(K'_{p1})}{D} = p\Lambda_p \approx \Lambda_\infty p + (\Lambda_1 - \Lambda_\infty) \tag{14}$$

For the Normal distribution, which most closely resembles the real distribution of the surface temperature, the approximated values of Λ_1 and Λ_∞ are: $\Lambda_1 = 2$ and $\Lambda_\infty = e^{1/2} = 1.649$.

3. Data

The data used as part of this study were retrieved from the GHCN-D database. GHCN (Global Historical Climatology Network)-Daily is a database of the National Oceanic and Atmospheric Administration of the United States that addresses the critical need for historical daily temperature, precipitation, and snow records over global land areas. GHCN-Daily is a composite of climate records from numerous sources that were merged and then subjected to a suite of quality assurance reviews. It contains temperature records from 106 283 stations in 180 countries and territories (Menne et al., 2012; e.g., see fig. 2). Both the record length and period of record vary by station and cover intervals that extend to more than 200 years.

GHCN-D database has been used in multiple scientific studies of the near-surface air temperature in the past. Studies of both global and regional focus, such as those of Portmann et al. (2009), Cavanaugh and Shen (2014), Dittus et al. (2015), have examined the trends of either the first four moments of the air temperature distribution or just the mean, in the context of statistical significance. In this study, we make use of the same records in the context of the stochastic nature of the air temperature, and how it explains changes in the tails of its distribution.

The stations analyzed, are subjected to multiple quality tests, both from the National Oceanic and Atmospheric Administration, which maintains the database, and the authors. The automated quality tests performed by NOAA resulted in the flagging of faulty data entries. For the purposes of this paper, we utilize only records with no quality flags, thus we dismiss all non-blank quality flagged values from the first stage of data gathering and processing. We isolate, and implemented, timeseries with a first entry prior to 1935, as this limitation enabled the extraction of more than 50 consecutive rolling 30-year periods. Despite the obvious narrowing of the pool of utilizable timeseries by this procedure, it enables us to identify shared large-scale persistence patterns among the stations. This would not have been possible, had we used a constantly changing sample of short-lived timeseries.

From this screening procedure, the number of records that are finally investigated is different for each aspect of temperature. For the study of behavior of the average near-surface air temperature we use 245 stations, while for the study of behavior of the maximum and minimum near-surface air temperature we use 5 006 stations for each one.

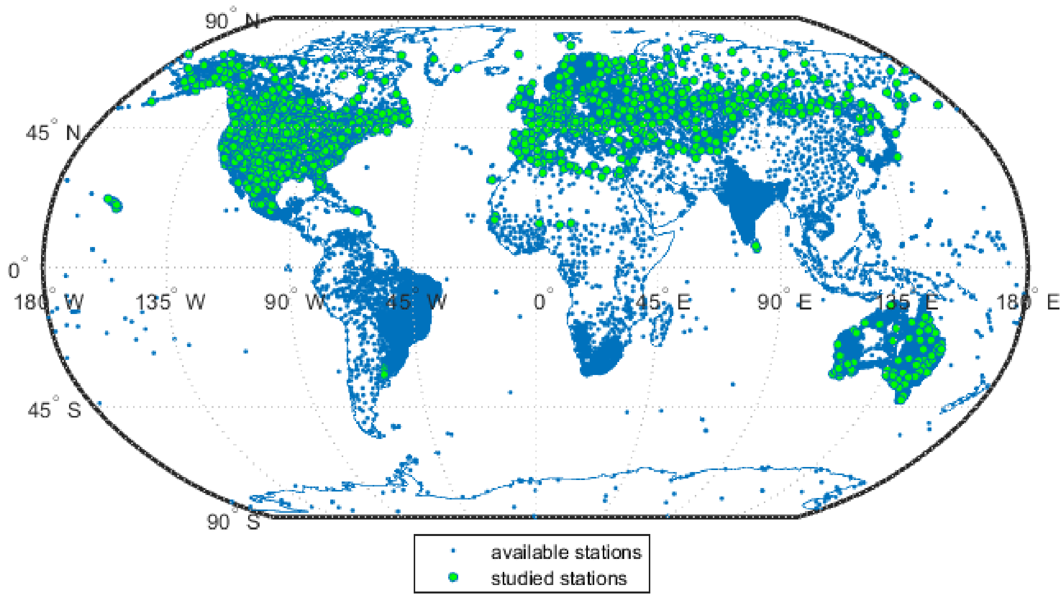

Figure 2: Spatial distribution of GHCN-D stations

It is apparent from the map of Figure 2 that the utilized temperature records originate from weather stations unevenly distributed across the earth surface. There is significant density of the studied stations in the Northern hemisphere, with the notable exception of a large cluster of Australian stations. This spatial limitation is considered inadvertent, since these areas host stations with temperature records of adequate time length.

Table 1: Temporal evolution of air temperature records used

Period	Average Temperature	Maximum Temperature	Minimum Temperature
1880-1899	90	2150	2191
1900-1919	142	4053	4076
1920-1939	245	4942	4943
1940-1959	245	4885	4886
1960-1979	245	4399	4396
1980-1999	245	3705	3705
2000-2018	240	3184	3184

4. Methodology

An overview of the stages followed in the study of the behavior of near-surface air temperature in global scale is presented in Figure 3. It is worth mentioning, that the procedure outlined in Figure 3 is repeated for each of the three variables of air temperature that are studied; i.e., average, maximum and minimum temperature.

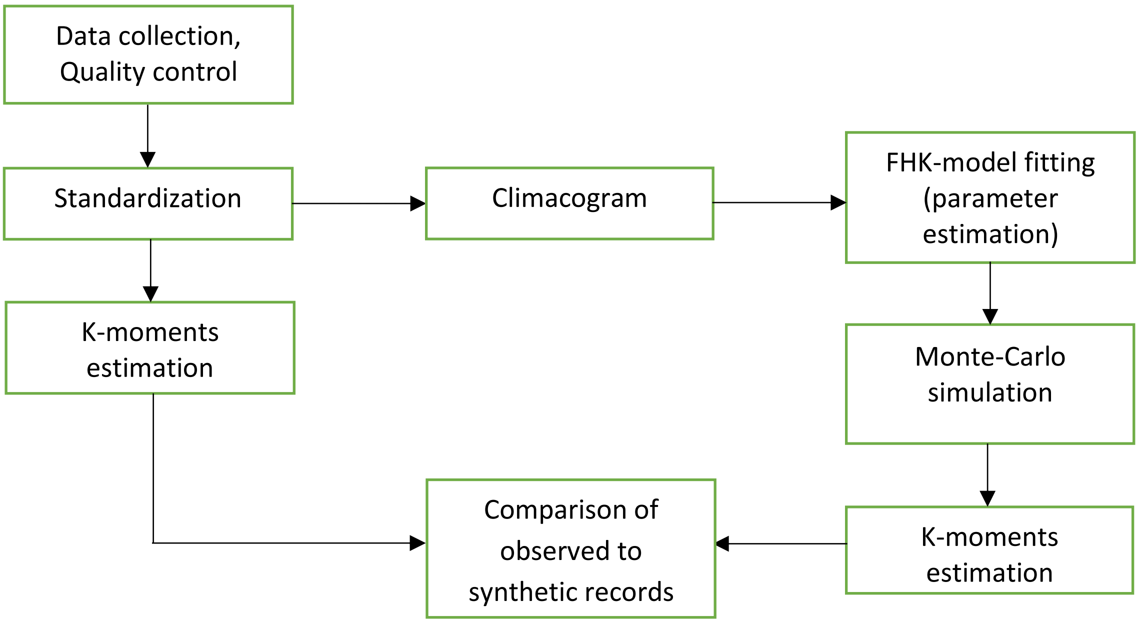

Figure 3: Methodology Layout

4.1 Initial data analysis

At the first stage, we download the daily average, maximum and minimum temperature daily records from the GHCN-D database, remove the flagged values and short length timeseries, and standardize the remaining utilizable timeseries. The standardization of the timeseries is performed in order for the input data, as well the results, to be comparable. Since multiple studies conclude that the distribution of the near-surface air temperature closely resembles the Gaussian, it was determined to standardize the timeseries according to the Gaussian distribution.

Moreover, since the study is focused on the behavior of the temperature on global scale, it is deemed reasonable to proceed with the standardization in a multi-year time frame. This is because of the fact that many weather stations around the world are located in climate zones with great variance of temperature among the different seasons.

Spatial variability plays a very important role when examining temperature dynamics, as the climatic conditions can affect to a great degree the persistence of extremes. For purposes of justifying the unified treatment of all the timeseries independently of the climatic conditions, we conduct a preliminary examination of the geographical homogeneity or heterogeneity of the climacograms of the timeseries for the different Köppen-Geiger climatic zones. Specifically, we estimate the mean climacogram for each of the three temperature metrics (average, maximum and minimum) and for each of the five (i.e., A, B, C, D, E) major climatic zones as defined by the Köppen-Geiger classification system (Geiger, 1954).

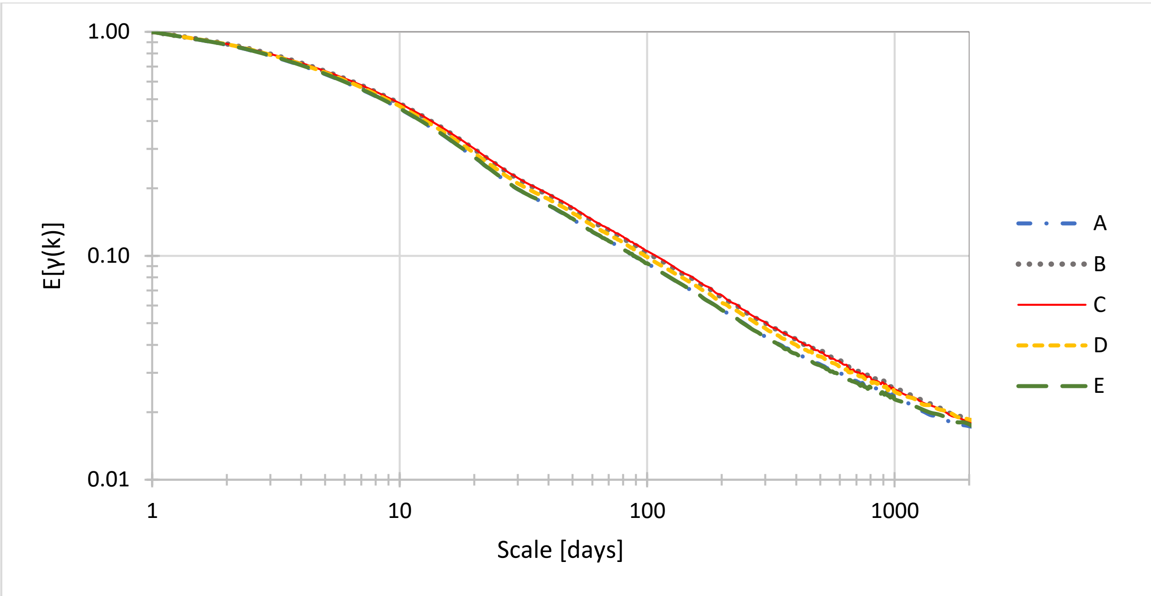

Figure 4: Climacograms of observed timeseries of the average near-surface air temperature for different climatic zones.

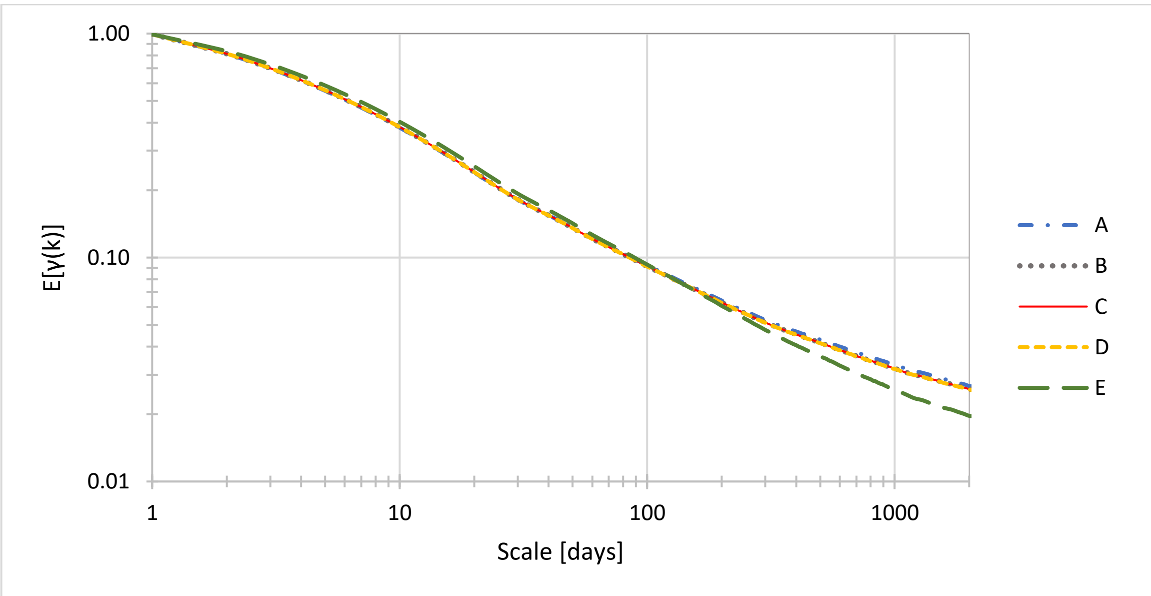

Figure 5: Climacograms of observed timeseries of the maximum near-surface air temperature for different climatic zones.

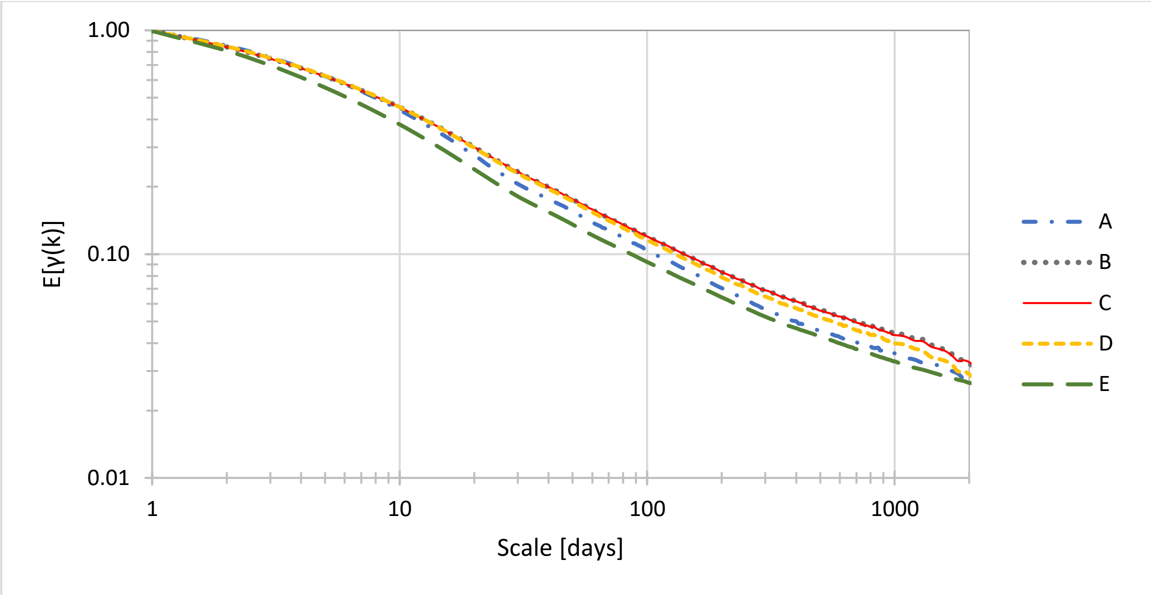

Figure 6: Climacograms of observed timeseries of the minimum near-surface air temperature for different climatic zones.

As can be seen in the Figures 4- 6, the differences of the standardized climacograms of the observed timeseries are minimal among the different climatic zones. This means that, if the seasonal variations of each climatic zone are removed (in terms of the first two moments), the resulting standardized timeseries behave similarly in terms of the dependence structure, irrespective of their location. Therefore, it is a reasonable choice to treat the ensemble of the standardized timeseries as a whole, even though in a further study more options of sub-setting based on geographical location could be explored.

4.2 Rolling 30-year periods

At the next stage, we use the rolling 30-year periods (see figure 7), as separate timeseries, from which specific extreme values corresponding to pre-selected return periods are extracted. We determine that the time length of each sub-series should be 30-years long, since three decades is an adequate time period to characterize the climate regime of an area. Moreover, 30 years is a time span that is equivalent to a human generation interval; hence, it is important to identify changes in such time scales. Longer time frames (e.g., 50 years) would significantly minimize the number of available rolling periods that could be extracted from each primary timeseries. Shorter time frames (e.g. 10 years) would inhibit the extraction of valuable extreme values occurring at larger time intervals, which are the ones most of interest.

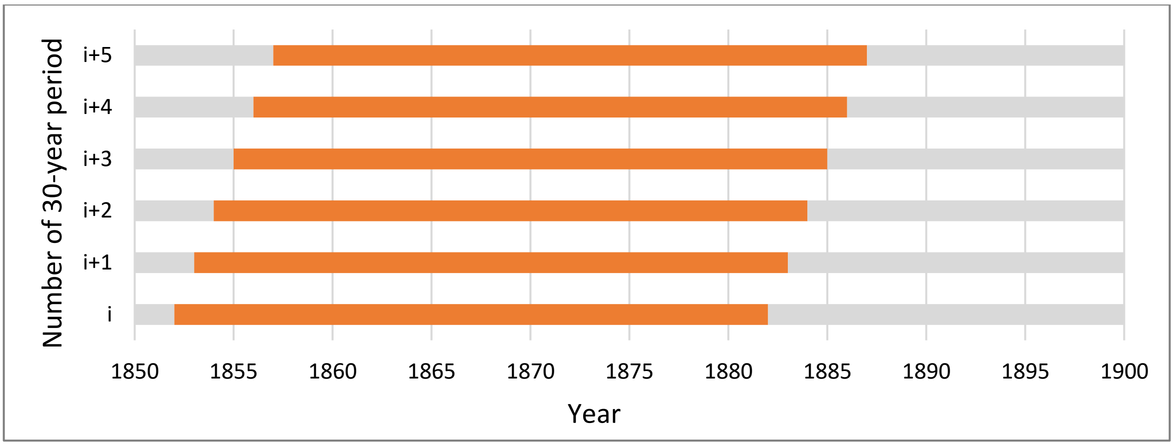

Figure 7: Example of consecutive 30-year periods

4.3 K-moments

After each 30-year long timeseries section is standardized, in relation to corresponding months, we apply the K-moments framework as follows. For the study of the upper tail, we sort the timeseries in ascending order, while for the study of the lower tail, we sort the timeseries in descending order. At this point, we calculate the return periods, in terms of days (not years) following the resolution of the temperature available data. For computational efficiency, the K-moments are extracted from the maxima of the 3, 10, 20 and 30-year periods, as shown in Table 2. However, since the exact definition of the return period T yields a linear relationship with the coefficient Λ_P, which depends on the distribution function (assumed Gaussian), the theoretical return period for the number of selected days is different, as described in Equation (14). Hence, the studied return periods are shown in columns three and four in Table 2.

For the day intervals (shown in the third column of Table 2) corresponding to these return periods, we use an iterative procedure to calculate both the fixed and added terms of each K-moment. The fixed terms depend only on p and the length of the sample, while the added terms depend on the index (see equations 6 and 7).

After we calculate the K-moments for the entirety of the suitable time-series and for all the 30-year periods of their time span, we perform a basic statistical analysis to summarize the information. For illustration, it is decided to isolate the distribution of each return period and each 30-year time-frame, and extract the values corresponding to the 25th, 50th (median) and 75th percentiles.

Table 2: The studied return periods of maximum or minimum observations in a time window defined in terms of day and yearly intervals assuming a Gaussian distribution.

Notation	Time window (years)	Time window (days)	Return period (years for Gaussian distribution)
T1	3	1096	5.30
T2	10	3653	16.84
T3	20	7305	33.33
T4	30	10958	49.82

4.4 Climacogram structure

For all the standardized timeseries we estimate the climacogram, for scales 1 to $n/10$, as suggested as a rule of thumb for the robust estimation of the long-term persistent parameter *H* (Dimitriadis and Koutsoyiannis, 2015), where *n* is the length of each timeseries. Then, we sum the values of the respective scales of the climacograms, so as to produce the arithmetic mean (average) of the climacogram for each scale. As a common maximum scale of the ensemble of the timeseries, we select the arithmetic mean of the lengths n_i of all timeseries. The estimated climacogram are juxtaposed with the theoretical expected values of the climacogram for a timeseries of a theoretical length, equal to the average of the lengths n_i of all timeseries.

The theoretical values of the climacogram for a Filtered Hurst-Kolmogorov process can be derived from equation (5). We fit equation (5) to the climacogram of the observed timeseries by estimating the parameters *H*, *M*, α_1, α_2, λ_1 and λ_2, through the minimization of the root mean square error ($\mathrm{E_{RMS}}$), which equals to:

$$\mathrm{E_{RMS}} = \sqrt{\sum_{i=1}^{n}(X_\mathrm{o} - X_\mathrm{m})^2 / n} \qquad (15)$$

where X_o is the observed value of the climacogram, X_m is its theoretical (modeled) value, and n is the total number of scales.

The optimization problem of minimizing the value of the $\mathrm{E_{RMS}}$ is handled by a combination of the Generalized Reduced Gradient (GRG2) algorithm and the Evolutionary algorithm. We use the GRG method to quickly identify the global minimum of the domain, while the Evolutionary algorithm is used, so as to improve even further the margin of error (limiting the number of iterations to 100,000).

From the above optimization, the *H* and *M* parameters are estimated for the average, maximum and minimum climacograms of the three air-temperature processes (Table 3). We observe that all processes exhibit a long-term persistent behaviour (H > 0.5), with an average value of the Hurst parameter equal to H_{ave} = 0.783.

Table 3: Hurst and Mandelbrot coefficients of optimized air temperature Climacograms

Air Temperature	H	M
Average	0.745	0.180
Maximum	0.766	0.077
Minimum	0.839	0.024

4.5 Stochastic synthesis

A rigorous and parsimonious method to produce synthetic timeseries for a physical process, like temperature, is by preserving its marginal and second-order dependence structures through the symmetric-moving average (SMA) scheme introduced by Koutsoyiannis (2000), further improved by Koutsoyiannis (2016) and implemented within the Castalia computer package (Efstratiadis et al., 2014). The SMA algorithm has the advantage of fully preserving in an exact way any second-order structure of a process and, simultaneously, the complete multivariate distribution function. As extended by Dimitriadis and Koutsoyiannis (2018), the SMA generation scheme can simulate a stochastic process by preserving explicitly its second-order dependence structure and its marginal structure through the first four central moments, which is found to be sufficient for various distributions applied in geophysical processes.

As explained in Dimitriadis and Koutsoyiannis (2018), high-order moments are extremely hard to calculate reliably from data, while the non-Gaussian distributions can be easily substantiated empirically, as well as derived in theory (Koutsoyiannis, 2014). One way to simulate the effect of the second-order dependence structure on the marginal structure is by explicitly preserving the high-order moments, as estimated from the distribution model and not from data. In most situations, the preservation of just four moments is a sufficient approximation of the distribution function. The fourth moment, in particular, has been deemed very important for some applications, e.g., in turbulence intermittency (Batchelor and Townsend, 1949).

In the SMA scheme, the simulated process is represented as the sum of products of coefficients a_j and white noise terms $\underline{v}_i$, (Koutsoyiannis, 2000):

$$\underline{x}_i = \sum_{j=-l}^{l} a_{|j|} \underline{v}_{i+j} \tag{16}$$

where, for simplicity and without losing generality, it is assumed that $\mathrm{E}[\underline{x}] = \mathrm{E}[\underline{v}] = 0$ and $\mathrm{E}[\underline{v}^2] = \mathrm{Var}[\underline{v}] = 1$, where index j ranges from 0 to infinity.

The SMA generation scheme can be employed for the stochastic generation of any type of second order structure, as represented through the climacogram, and this is pivotal in its selection in the present study of the near-surface air temperature. This scheme presents several advantages over other models, such as the backwards moving average (BMA). Particularly, for $l \to \infty$ or l finite, the coefficients can be analytically calculated through the Fourrier transform of the discrete power spectrum of the coefficients, which is directly related to the analytically expressed discrete power spectrum of the process (Koutsoyiannis, 2000):

$$s_{a_\mathrm{d}}(\omega) = \sqrt{2 s_\mathrm{d}(\omega)} \tag{17}$$

where s_{a_d} and s_d are the SMA coefficients and process power spectra in discrete time, respectively. For instance, for an HK process with $H > 0.5$, the SMA coefficients can be easily estimated from the expression (Koutsoyiannis, 2016):

$$a_j = \frac{1}{2}\sqrt{2\Gamma(2H+1)\sin(\pi H)\,\gamma_\Delta \Gamma^2(2H+1)(1+\sin(\pi H))} \times \left(|j+1|^{H+\frac{1}{2}} + |j-1|^{H+\frac{1}{2}} - 2|j|^{H+\frac{1}{2}}\right) \tag{18}$$

The algorithm to produce timeseries with the SMA scheme, developed in Dimitriadis and Koutsoyiannis (2018), requires the first four central moments, the climacogram model for each process (average, maximum and minimum temperature), and the length of the synthetic timeseries. Therefore, for each observed timeseries that passed the multiple quality checks, we calculate the first four central moments, and the climacogram model parameters.

Since the generation of a synthetic timeseries may be time-consuming for very large lengths, we produce for each of the three processes (average, maximum and minimum temperature) only a fixed number of synthetic timeseries. This number of produced timeseries is equal to the least observed, yet utilizable, timeseries for each of the three temperature processes. The number of observed timeseries for the average process of the air-temperature is 245, while for the maximum and minimum processes is 5 006. Thus, 245 synthetic timeseries are created for each process. For the synthetic timeseries, we follow the same methodology as in the analysis of the observed timeseries. The synthetic data are first standardized, then separated into rolling 30-year periods, from which the K-moments of the selected return period levels were extracted.

4.6 Longest individual records

As part of the study of the behavior of air temperature, we compare the aggregate behavior of all the available timeseries, with the behavior of the longest timeseries, to identify possible similarities of the variability among them. The longest air temperature records for each aspect of near-surface air temperature are shown in Table 4. For the study of these records, we apply a similar methodology as the one with the sum of observed records.

Table 4: Longest recording individual stations

Air Temperature	Station ID	Location	Record Length
Average	RSM00026063	St. Petersburg, Russia	136 years
Maximum	ITE00100554	Milan, Italy	246 years
Minimum	ITE00100554	Milan, Italy	246 years

5. Results and Discussion

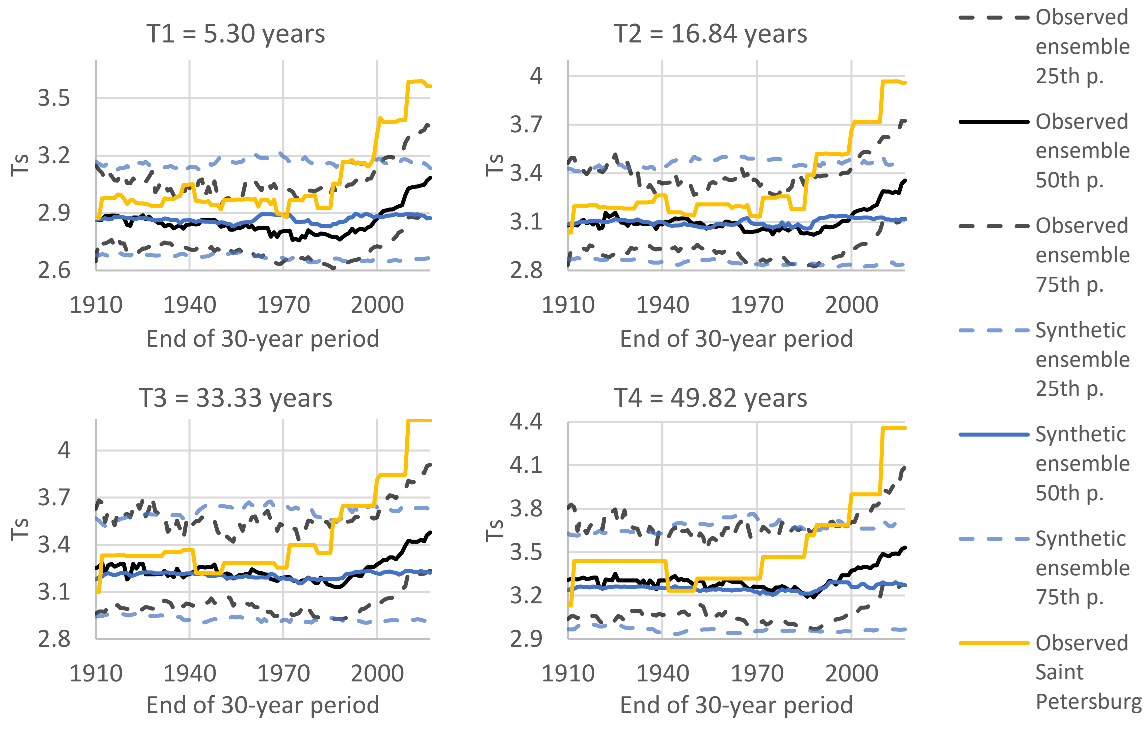

Figure 8: Upper tail of the standardized average air temperature (Ts) over time

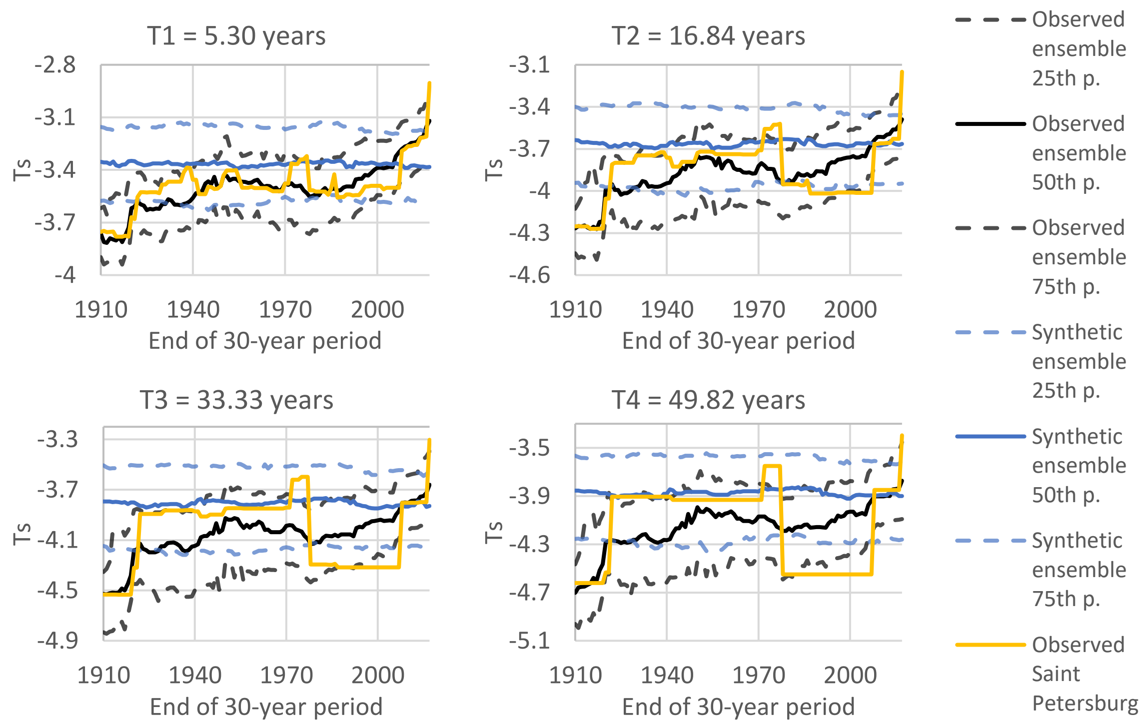

Figure 9: Lower tail of the standardized average air temperature (Ts) over time

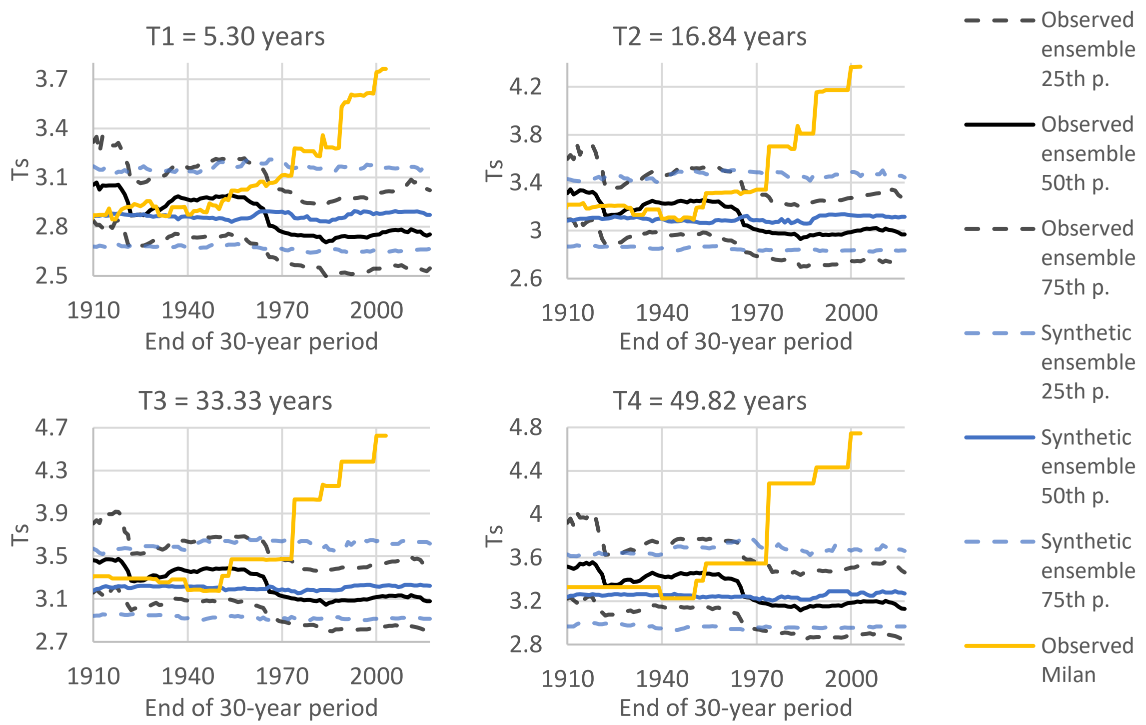

Figure 10: Upper tail of the standardized maximum air temperature (Ts) over time

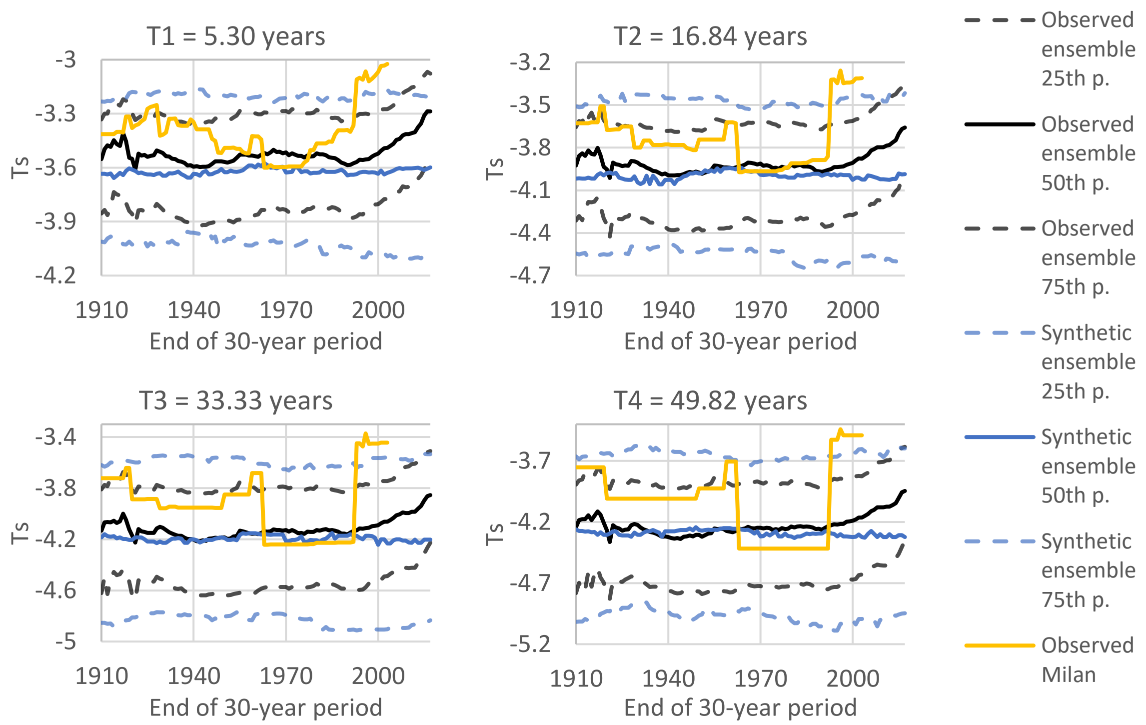

Figure 11: Lower tail of the standardized minimum air temperature (Ts) over time

The study of the near-surface air temperature records brought interesting facts to light. All three air temperature metrics (namely average, maximum and minimum) show an unstable behavior, with prominent fluctuations at the climatic scale throughout the years. Most noticeable changes include the fattening and thinning of the tails. Fattening can be witnessed as an increase of the standardized air temperature for the upper tails and decrease of the standardized air temperature for the lower tails. Thinning corresponds to exactly the opposite changes; i.e., decrease of the standardized air temperature for the upper tails and increase of the standardized air temperature for the lower tails.

The average near-surface air temperature exhibits the most coherent behavior with a progressive warming evident in both tails. The upper tail tends towards becoming thicker (witnessed by increasing Ts), whereas the lower tail appears to become thinner (increasing Ts) as time progresses (Figures 8 and 9). This confirms the expectation that the global average air temperature increases, a result that may have important social and environmental effects. Yet in terms of the lower tail of the average temperature, this increase takes place from the start of the 20th century, whereas in terms of the upper tail it concerns only the past 30 years.

Maximum and minimum air temperature present an even more complex behavior. An interesting and somehow unexpected finding is the thinning (decreasing Ts) of the upper tail of the maximum near-surface air temperature. Contrary to our expectations, we find that the behavior of the maximum temperature diverges from that of the average one, suggesting that temperature is a more complex climatic variable than previously thought. Average temperature is by definition calculated as the mean of multiple observations within a certain time-span, which obviously includes the maximum and minimum as well. Even though, the maximum recorded temperature is an integral part of the set of data from which the average temperature is derived, it appears that its effect on the average is not so intelligible.

As for the lower tail of the minimum near-surface air temperature, it is shown that despite the increasing trend of the average air temperature (even of its lower values), it remains surprisingly steady, at all return period levels, and only in the last 10-20 years presents an increasing trend. This shows that the temperature changes are not consistent throughout the range of its variability, and a form of asynchronicity is present among the different temperature metrics. Thus, climate dynamics is characterized by a sort of “stamina” and is probably able to mitigate, to some extent, changes in the atmosphere. The fact that the upward trend is almost equally evident in all the return period levels of the lower tail, suggests that it is more probable that the resulting change stems from a change in the average of the distribution rather than a change in its standard deviation, assuming that the minimum temperature presents a nearly-Gaussian distribution.

The average near-surface air temperature synthetic records produced present an ambiguous behavior with respect to the two tails. On the one hand, the upper tail presents a similar pattern to the observed data up to 2000, but when the beginning of the 21st century is included in the analysis, a divergence of the observed and synthetic records is evident. The variance of the trend, as expressed through the interquartile range in each return period level, is almost the same, which suggests that the synthetic series reproduce well the variability range. On the other hand, the lower tail of the synthetic series, despite having the same variance at all the return period levels, is much thinner than in reality, while it does not reproduce the thinning trend of the past century. This means that the extreme cold waves, affecting the lower tail of the average temperature, are much more common than anticipated by the reproduction of the observed stochastic behavior, although this trend tends to reverse as time progresses (Figure 9).

Despite having a uniquely fitted Filtered Hurst-Kolmogorov process to estimate the persistence of the average air temperature, it seems inadequate, at first glance, to decipher, and consequently reproduce, the complex temporal behavior of the average air temperature in terms of both its tails. However, that is not really the case if one considers the very strict percentile margins we have depicted. Namely, the depicted 25th and 75th percentile range contains only half (i.e., 50%) of the range standardized air temperature fluctuates, meaning that the other half of the ensemble is outside of these margins.

Concerning the synthetic records that reproduce the maximum near-surface air temperature, the variance of the trend, expressed through the interquartile range at each return period level, is greater than that of the observed data; hence, proving that any upward trend of the upper tail is within the stochastically expected boundaries. Furthermore, the slightly increasing trend of the upper 75th percentile of the interquartile range, which is present at higher return periods, is completely the opposite from the limit performance of the observed data (Figure 10).

The minimum near-surface air temperature synthetic records produced (see Figure 11), present many comparative similarities with the maximum air temperature. Specifically, the size of the interquartile range is greater than the one derived from the observed data, and in fact, overspreads it. This means that any changes present in the observed timeseries can be explained, and thus anticipated, through the study of their statistical behavior. Moreover, the slightly increasing trend of the extreme values of the lower tail may suggest a return to stability and not a spiraling towards global overheating.

Individual records though may present a markedly different behavior, from the average of the ensemble of observed records. Specifically, both the average temperature records of Saint Petersburg, and the maximum and minimum temperature records of Milan show substantial warming. At some return period levels this warming is a multiple of the warming present on all the other records. One possible reason for this divergence of results is the location of these weather stations in relation to the urban agglomerations (see also the similar work of Sigourou et al., 2018) and the increasing scarcity of green, open spaces that mitigate the urban heat island effect (see the works of Bernatzky, 1982 and Aram et al., 2019). According to the coordinates obtained from the GHCN-D station directory (Menne et al., 2012), both Milan and Saint Petersburg stations are deep within the center of the urban areas, meaning that the urban heat island effect has profound implications on the temperature measurement.

As shown in Figure 12, a similar study undertaken by Koutsoyiannis (2019b) reveals that the weather station in the city center of Milan presents a different behavior than that of the suburban weather stations of Monte Cimone and Paganella, which are both in the vicinity of Milan. In more detail, while the weather records of Monte Cimone and Paganella show a relatively steady level of the maximum air temperature, the Milan station presents a clearly warming trend, even though it refers to the same return period as the other two weather stations. This strongly supports the assumption of the great impact that the heat island effect may have on the Milan temperature.

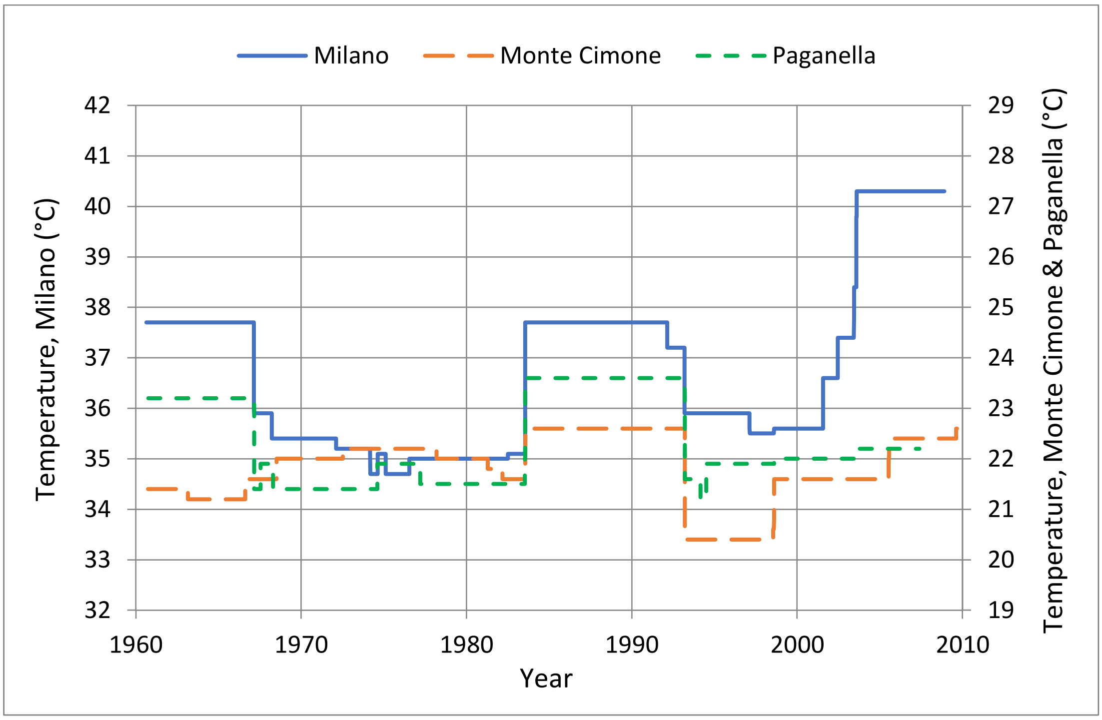

Figure 12: Comparison of Lombardy temperature extremes (Koutsoyiannis, 2019b)

Another potent reason is the location bias of the longest individual records, in comparison with the great scattering of the ensemble of air temperature records. Since both Milan and Saint Petersburg are in the European continent (and even close to the sea front), their warming behavior could be a characteristic of coastal or near-coastal regions that is not shared with all the other temperature records. However, it is intuitively important to compare them with the rest of the shorter-length records as a point of reference.

Last, but not least, it is clearly visible in Figures 6 and 8, that there are significant drops (i.e., fattening of the lower tail of the minimum temperature) in the individual records of Saint Petersburg and Milan respectively during the latter half of the 20th century. These drops correspond to cold weather extremes, which in terms of our methodology, are the absolute minima for a period of 30-years (see Section 4.2). The winter of 1962/63 was one of the harshest for Western Europe, where Milan is, with recorded temperatures up to -5 degrees Celsius below the expected average for the season (Hirschi and Sinha, 2007). As far as the Saint Petersburg drop is concerned, it corresponds to the extremely cold winter of 1978/79 that affected European Russia (Khasanov, 2013).

6. Conclusions

This work performs a global stochastic investigation of the extremes of near-surface air temperature, employing a set of advanced stochastic tools, i.e., the climacogram and the K-moments for the estimation thereof. From this worldwide survey on the near-surface air temperature extremes, it is revealed that the air temperature presents a counter-intuitive and much more complex behavior than usually modeled by classical statistics. It particularly exhibits temporal changes in magnitude, in variability and in shape of the tail distribution. A divergence between observed and synthetic series was noted in some cases, which can possibly be explained by the preservation of solely the first four moments. Nevertheless, this simple stochastic model is still able to adequately reproduce the observed variability range. Also, the assumption of a common worldwide behavior is justified based on the similarities of the different climate zones.

Yet what was less expected is the fact that the observed temporal changes in the average, maximum and minimum temperatures are neither synchronous nor consistent to each other. In particular, the lower tail of the distribution of average shows a prominent increasing trend in the first half of the 20th century, whereas, on the contrary, the upper tail of the distribution of average and the upper extremes exhibit notable stability over the same period. On the other hand, the increase in the average and minimum temperature over the past 20-30 years is not followed by an increase in the maximum temperature, but rather by a decrease. These observed peculiarities among the different indices of air temperature (namely average, maximum and minimum) can be, in part, attributed to the deviation of the air temperature distribution from Gaussianity, but should be mostly regarded as evidence of the pronounced inherent variability.

Overall, the observed changes of the air-temperature behavior correspond to a probability distribution whose upper tail (i.e., high temperature extremes) tends to become slightly thinner whereas its lower tail (i.e., the low temperature extremes) tends to become even more thinner. Hence, the high and low temperature extremes tend to become, more or less, scarcer than in the past, especially the ones of the lower tail. The average temperature, however, which corresponds to the main body of the distribution, increases substantially. All these in combination, according to the authors' perception, create a shift towards a temperature distribution with seemingly smaller variance but with a higher average (see Figure 13).

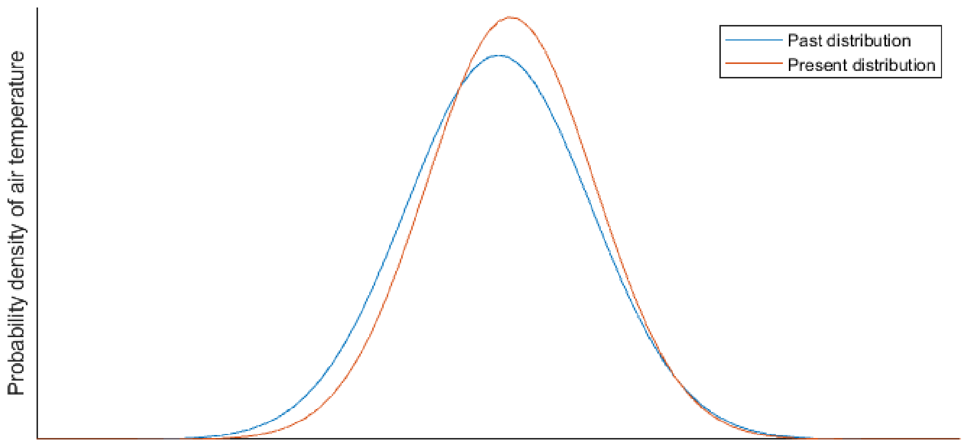

Figure 13: One possible historical evolution of the air temperature probability density function.

This conclusion may seem counter-intuitive and inconsistent with previous research (e.g., Coumou and Rahmstorf, 2012 and IPCC, 2014). However, this shifted and altered distribution should not be misconstrued as directly leading to less weather extremes, since the relationship between temperature and weather extreme phenomena is much more complex. In droughts for instance, a major factor of their occurrence is a prolonged period of higher than usual temperatures coupled with less, or none at all, precipitation. This may very well be linked to higher average temperatures, as found to be the result of the present study.

Furthermore, when comparing the aforementioned results' divergence, one has to consider that a major differentiation point of the current study in comparison to previous studies has been the use of K-moments in estimating the past and present extremes of the temperature's tails. The very powerful statistical properties of K-moments (see Appendix 1) and their supreme performance in reducing the estimation bias (see Koutsoyiannis, 2020) may be the reason for the difference in the results. Overall, it is the authors' aspiration that the results of this study will shed some light into the complicated near-surface temperature extremes changes over time, in order to facilitate future research on temperature dynamics.

Declarations

Acknowledgements

We are grateful to the Editor in Chief George Christakos, the anonymous Associate Editor and the two anonymous reviewers for their efforts, useful comments and suggestions that helped us improve the paper.

Funding information

No funds have been available for this work.

Code availability

The scripts and functions used, all of which were implemented in Matlab can be downloaded from www.itia.ntua.gr/2079/. Also, a readme file, in txt format, on the same repository contains explanatory information on the operations each code script performs.

Conflicts of interest

All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

References

Aram, F., García, E. H., Solgi, E., & Mansournia, S. (2019). Urban green space cooling effect in cities. *Heliyon*, *5*(4), e01339.

Batchelor, G. K., & Townsend, A. A. (1949). The nature of turbulent motion at large wave-numbers. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, *199*(1057), 238-255.

Bernatzky, A. (1982). The contribution of tress and green spaces to a town climate. *Energy and buildings*, *5*(1), 1-10.

Braganza, K., Karoly, D. J., & Arblaster, J. M. (2004). Diurnal temperature range as an index of global climate change during the twentieth century. *Geophysical research letters*, *31*(13).

Brown, P. T., Li, W., Cordero, E. C., & Mauget, S. A. (2015). Comparing the model-simulated global warming signal to observations using empirical estimates of unforced noise. *Scientific reports*, *5*, 9957.

Cavanaugh, N. R., & Shen, S. S. (2014). Northern Hemisphere climatology and trends of statistical moments documented from GHCN-daily surface air temperature station data from 1950 to 2010. *Journal of climate*, *27*(14), 5396-5410.

Coumou, D., & Rahmstorf, S. (2012). A decade of weather extremes. *Nature climate change*, *2*(7), 491-496. doi-org.tudelft.idm.oclc.org/10.1038/nclimate1452

Cronin, T. M. (2009). *Paleoclimates: understanding climate change past and present*. Columbia University Press.

Dimitriadis, P. (2017). Hurst-Kolmogorov Dynamics in Hydrometeorological Processes and in the Microscale of Turbulence. Ph.D. Thesis, Department of Water Resources and Environmental Engineering—National Technical University of Athens, Athens, Greece.

Dimitriadis, P., & Koutsoyiannis, D. (2015). Climacogram versus autocovariance and power spectrum in stochastic modelling for Markovian and Hurst–Kolmogorov processes. *Stochastic environmental research and risk assessment*, *29*(6), 1649-1669.

Dimitriadis, P., & Koutsoyiannis, D. (2018). Stochastic synthesis approximating any process dependence and distribution. *Stochastic environmental research and risk assessment*, *32*(6), 1493-1515.43.

Dimitriadis, P., and D. Koutsoyiannis (2019). The mode of the climacogram estimator for a Gaussian Hurst-Kolmogorov process, Journal of Hydroinformatics, doi:10.2166/hydro.2019.038.

Dimitriadis, P., K. Tzouka, D. Koutsoyiannis, H. Tyralis, A. Kalamioti, E. Lerias, and P. Voudouris (2019). Stochastic investigation of long-term persistence in two-dimensional images of rocks, Spatial Statistics, 29, 177–191, doi:10.1016/j.spasta.2018.11.002.

Dittus, A. J., Karoly, D. J., Lewis, S. C., & Alexander, L. V. (2015). A multiregion assessment of observed changes in the areal extent of temperature and precipitation extremes. *Journal of Climate*, *28*(23), 9206-9220.

Easterling, D. R., Horton, B., Jones, P. D., Peterson, T. C., Karl, T. R., Parker, D. E., ... & Folland, C. K. (1997). Maximum and minimum temperature trends for the globe. *Science*, *277*(5324), 364-367.

Efstratiadis, A., Dialynas, Y. G., Kozanis, S., & Koutsoyiannis, D. (2014). A multivariate stochastic model for the generation of synthetic time series at multiple time scales reproducing long-term persistence. *Environmental Modelling & Software*, *62*, 139-152.

Geiger, R. (1954). Klassifikation der klimate nach W. Köppen. *Landolt-Börnstein–Zahlenwerte und Funktionen aus Physik, Chemie, Astronomie, Geophysik und Technik*, *3*, 603-607.

Gneiting T.; Schlather M. (2004). Stochastic Models That Separate Fractal Dimension and the Hurst Effect. SIAM Review, 46, 269-282, 10.1137/s0036144501394387.

Handmer, J., Honda, Y., Kundzewicz, Z. W., Arnell, N., Benito, G., Hatfield, J., ... & Takahashi, K. (2012). Changes in impacts of climate extremes: human systems and ecosystems. In *Managing the risks of extreme events and disasters to advance climate change adaptation special report of the intergovernmental panel on climate change* (pp. 231-290). Intergovernmental Panel on Climate Change.

Hasselmann, K. (1976). Stochastic climate models part I. Theory. *tellus*, *28*(6), 473-485.

Hirschi, J. J. M., & Sinha, B. (2007). Negative NAO and cold Eurasian winters: how exceptional was the winter of 1962/1963?. *Weather*, *62*(2), 43-48.

Hurst, H. E. (1951). Long-term storage capacity of reservoirs. *Trans. Amer. Soc. Civil Eng.*, *116*, 770-799.

Iliopoulou, and D. Koutsoyiannis (2019). Revealing hidden persistence in maximum rainfall records, Hydrological Sciences Journal, 64 (14), 1673–1689, doi:10.1080/02626667.2019.1657578.

IPCC. (2014). *Climate Change 2013 – The Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge: Cambridge University Press. doi:10.1017/CBO9781107415324

IPCC. (2018). *Summary for Policymakers. In: Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty*. World Meteorological Organization, Geneva

Jones, P. D., Lister, D. H., Osborn, T. J., Harpham, C., Salmon, M., & Morice, C. P. (2012). Hemispheric and large-scale land-surface air temperature variations: An extensive revision and an update to 2010. *Journal of Geophysical Research: Atmospheres*, *117*(D5).

Jones, P. D., Parker, D. E., Osborn, T. J., & Briffa, K. R. (2016). *Global and Hemispheric Temperature Anomalies: Land and Marine Instrumental Records (1850-2015)*. Environmental System Science Data Infrastructure for a Virtual Ecosystem; Carbon Dioxide Information Analysis Center (CDIAC), Oak Ridge National Laboratory, Oak Ridge, TN (USA). doi: 10.3334/CDIAC/cli.002

Khasanov, B. F. (2013). Severe winter rings of oak trees (Quercus robur L.) from Central European Russia. *International journal of biometeorology*, *57*(6), 835-843.

Kolmogorov, A. N. (1940). Wienersche spiralen und einige andere interessante kurven in hilbertscen raum, cr (doklady). *Acad. Sci. URSS (NS)*, *26*, 115-118.

Koutsoyiannis, D. (2000). A generalized mathematical framework for stochastic simulation and forecast of hydrologic time series. *Water Resources Research*, *36*(6), 1519-1533.

Koutsoyiannis, D. (2002). The Hurst phenomenon and fractional Gaussian noise made easy. *Hydrological Sciences Journal*, *47*(4), 573-595.

Koutsoyiannis, D. (2010). HESS Opinions" A random walk on water". *Hydrology and Earth System Sciences*, *14*(3), 585-601.

Koutsoyiannis, D. (2011). Hurst-Kolmogorov Dynamics and Uncertainty 1. *JAWRA Journal of the American Water Resources Association*, *47*(3), 481-495.

Koutsoyiannis, D. (2014). Entropy: from thermodynamics to hydrology. *Entropy*, *16*(3), 1287-1314.

Koutsoyiannis, D. (2016). Generic and parsimonious stochastic modelling for hydrology and beyond. *Hydrological Sciences Journal*, *61*(2), 225-244.

Koutsoyiannis, D. (2017). Entropy production in stochastics. *Entropy*, *19*(11), 581.

Koutsoyiannis, D. (2019a). Knowable moments for high-order stochastic characterization and modelling of hydrological processes. *Hydrological sciences journal*, *64*(1), 19-33.

Koutsoyiannis, D. (2019b) *Advances in stochastics of hydroclimatic extremes*. Presentation. Conference: Giornata di studio in memoria di Baldassare Bacchi , University of Brescia, Italy. doi: 10.13140/RG.2.2.30655.05282/1

Koutsoyiannis, D. (2020). *Stochastics of Hydroclimatic Extremes*. National Technical University of Athens. Access Date: 20 December 2020. Available at: http://itia.ntua.gr/2000/

Koutsoyiannis D.; Dimitriadis P.; Lombardo F.; Stevens S. (2018), From fractals to stochastics: Seeking theoretical consistency in analysis of geophysical data, Advances in Nonlinear Geosciences, edited by A.A. Tsonis, 237–278, doi:10.1007/978-3-319-58895-7_14, Springer.

Mandelbrot, B. B., & Van Ness, J. W. (1968). Fractional Brownian motions, fractional noises and applications. *SIAM review*, *10*(4), 422-437.

Masson-Delmotte, T. W. V., Zhai, P., Pörtner, H. O., Roberts, D., Skea, J., Shukla, P. R., ... & Connors, S. (2018). IPCC, 2018: Summary for Policymakers. In: Global warming of 1.5 C. An IPCC Special Report on the impacts of global warming of 1.5 C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global. *World Meteorological Organization, Geneva, Tech. Rep*.

Menne, M. J., Durre, I., Korzeniewski, B., McNeal, S., Thomas, K., Yin, X., ... & Houston, T. G. (2012). Global historical climatology network-daily (GHCN-Daily), Version 3. *NOAA National Climatic Data Center*, *10*, V5D21VHZ. Access Date: 15 April 2019

Montanari, A., Rosso, R., & Taqqu, M. S. (1997). Fractionally differenced ARIMA models applied to hydrologic time series: Identification, estimation, and simulation. *Water resources research*, *33*(5), 1035-1044.

O'Connell, P.E.; Koutsoyiannis, D.; Lins, H.F.; Markonis, Y.; Montanari, A.; Cohn, T.; The scientific legacy of Harold Edwin Hurst. Hydrol. Sci. J., 61, 1571–1590, 10.1080/02626667.2015.1125998, 2016.

Papoulis, A. (1990). *Probability & statistics* (Vol. 2). Englewood Cliffs: Prentice-Hall.

Peterson, T. C., Gallo, K. P., Lawrimore, J., Owen, T. W., Huang, A., & McKittrick, D. A. (1999). Global rural temperature trends. *Geophysical Research Letters*, *26*(3), 329-332.

Portmann, R. W., Solomon, S., & Hegerl, G. C. (2009). Spatial and seasonal patterns in climate change, temperatures, and precipitation across the United States. *Proceedings of the National Academy of Sciences*, *106*(18), 7324-7329.

Rubel, F., & Kottek, M. (2010). Observed and projected climate shifts 1901–2100 depicted by world maps of the Köppen-Geiger climate classification. *Meteorologische Zeitschrift*, *19*(2), 135-141.

Sargentis G.-F.; Dimitriadis P.; Ioannidis R.; Iliopoulou T.; Koutsoyiannis D. (2019). Stochastic Evaluation of Landscapes Transformed by Renewable Energy Installations and Civil Works, Energies, 12, 2817, 10.3390/en12142817.

Sargentis G.-F.; Dimitriadis P.; Koutsoyiannis D. (2020); Aesthetical Issues of Leonardo Da Vinci's and Pablo Picasso's Paintings with Stochastic Evaluation, Heritage, 3, 283-305, 10.3390/heritage3020017.

Sigourou, S., Dimitriadis, P., Iliopoulou, T., Ioannidis, R., Skopeliti, A., Sakellari, K., & Koutsoyiannis, D. (2018, April). Statistical and stochastic comparison of climate change vs. urbanization. In *EGU General Assembly Conference Abstracts* (Vol. 20, p. 18608).

Sun, X., Ren, G., Xu, W., Li, Q., & Ren, Y. (2017). Global land-surface air temperature change based on the new CMA GLSAT data set. *Sci. Bull*, *62*(4).

Trenberth, K. E., Jones, P. D., Ambenje, P., Bojariu, R., Easterling, D., Klein Tank, A., ... & Soden, B. (2007). Observations: surface and atmospheric climate change. Chapter 3. *Climate change*, 235-336.

Appendix 1

To facilitate the understanding of the theory behind K-moments, we explain some basic notions of statistics in this appendix.

Let $\underline{x}$ be a stochastic variable and $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_p$ be copies of it, independent and identically distributed, forming a sample. The maximum of all, which is identical to the pth order stochastic, is by definition:

$$\underline{x}_{(p)} := \max\,(\underline{x}_1, \underline{x}_2, \dots, \underline{x}_p) \tag{19}$$

It is readily obtained that if $F(x)$ is the distribution function of $\underline{x}$ and $f(x)$ its probability density function, then those of $\underline{x}_{(p)}$ are distributed by:

$$F^{(p)}(x) = (F(x))^p, \qquad f^{(p)}(x) = pf(x)\big(F(\underline{x})\big)^{p-1} \tag{20}$$

where the former is the product of p instances of $F(x)$ (justified by the independent and identically distributed assumption), while the latter is the derivative of $F^{(p)}(x)$ with respect to x. The *expected maximum order of* p of $\underline{x}$, i.e. the expected value of $\underline{x}_{(p)}$, is therefore:

$$\mathrm{E}\big[\underline{x}_{(p)}\big] = \mathrm{E}\big[\max\big(\underline{x}_1, \underline{x}_2, \dots, \underline{x}_p\big)\big] = p\mathrm{E}\left[\big(F(\underline{x})\big)^{p-1}\,\underline{x}\right] \tag{21}$$

It is worth to stress that the variables $\underline{x}_1$, $\underline{x}_2$, … , $\underline{x}_p$ considered here, are not meant in temporal succession and, in this respect, do not form a stochastic process, but are rather regarded to be an ensemble of copies of $\underline{x}$. In other words, the possible dependence in time of a stochastic process is not considered to be prerequisite for the application.

In geophysical processes, it is justifiable to assume that the variance $\mu_2 \equiv \sigma^2$ is finite, because an infinite variance would translate to an infinite amount of energy to materialize, which is absurd. However, high-order classical moments μ_p diverge to infinity beyond a certain p (i.e., in heavy-tailed distributions). That is not the case for the K-moments, where a significant part of the moment is calculated using the always finite distribution function (Koutsoyiannis, 2019a), which is the reason from which their knowability stems.

To derive *knowable* moments for high orders p, in the expectation defining the pth moment, we raise $(\underline{x} - \mu)$ to a low power $q < p$ and for the remaining $(p - q)$ multiplicative terms, we replace $(\underline{x} - \mu)$ with $\big(2F(\underline{x}) - 1\big)$, where $F(x)$ is the distribution function. This leads to the following definition of central *K-moment* of order (p, q) (Koutsoyiannis, 2019a):

$$K_{pq} := (p - q + 1)\mathrm{E}\left[\big(2F(\underline{x}) - 1\big)^{p-q}(\underline{x} - \mu)^q\right], \qquad p \geq q \tag{22}$$

Likewise, the non-central K-moment of order (p, q) is defined (Koutsoyiannis, 2019a):

$$K'_{pq} \coloneqq (p-q+1)\mathrm{E}\left[\left(F(\underline{x})\right)^{p-q} \underline{x}^q\right], \qquad p \geq q \tag{23}$$

The quantities $\left(F(\underline{x})\right)^{p-q}$ and $\left(2F(\underline{x})-1\right)^{p-q}$ are estimated from a sample, without the use of powers of $\underline{x}$, thus making the estimation more reliable. Specifically, for the ith element of a sample $x_{(i)}$ of size n, sorted in ascending order, $F(x_{(i)})$ and $\left(2F(x_{(i)})-1\right)$ are estimated as:

$$\hat{F}(x_{(i)}) = \frac{i-1}{n-1},\ 2\hat{F}(x_{(i)}) - 1 = \frac{2i-n-1}{n-1} \tag{24}$$

taking values in [0,1] and [-1,1], respectively, irrespective of the values $x_{(i)}$. Hence, the estimators of K-moments are:

$$\hat{K}'_{pq} = \frac{p-q+1}{n} \sum_{i=1}^{n} \left(\frac{i-1}{n-1}\right)^{p-q} \underline{x}_{(i)}^{q} \tag{25}$$

$$\hat{K}_{pq} = \frac{p-q+1}{n} \sum_{i=1}^{n} \left(\frac{2i-n-1}{n-1}\right)^{p-q} (\underline{x}_{(i)} - \hat{\mu})^{q} \tag{26}$$

The rationale of the definition is relatively easy to grasp. Assuming that the distribution mean is close to the median, so that $F(\mu) \approx {}^{1}\!/_{2}$ (this is precisely true for a symmetric distribution), the quantity whose expectation is taken from the definition of the central K-moment of order (p,q) is: $A(\underline{x}) \coloneqq (2F(\underline{x})-1)^{p-q}(\underline{x}-\mu)^q$ and its Taylor expansion is:

$$\begin{aligned} A(\underline{x}) = {} & (2f(\mu))^{p-q}(\underline{x}-\mu)^p + (p-q)\left(2f(\mu)\right)^{p-q-1} f'(\mu)\left(\underline{x}-\mu\right)^{p+1} \\ & + O(\left(\underline{x}-\mu\right)^{p+2}) \end{aligned} \tag{27}$$

where $f(x)$ is the probability density function of $\underline{x}$. Clearly then, K_{pq} depends on μ_p as well as on classical moments of $\underline{x}$ of order higher than p. The independence of K_{pq} from classical moments of order smaller than p is the reason why it is a competent surrogate of the unknowable μ_p.

In addition, as p becomes large, by virtue of the multiplicative term $(p-q+1)$ in the definition of K-moments, K_{pq} shares similar asymptotic properties with $\hat{\mu}_p{}^{q/p}$ (the estimate, not the true $\mu_p{}^{q/p}$). To illustrate this for $q=1$ and for independent variables $\underline{x}_i$, we consider the variable $\underline{z}_p \coloneqq \max_{1 \leq i \leq p} \underline{x}_i$ and denote $f(\,)$ and $h(\,)$ the probability densities of $\underline{x}_i$ and $\underline{z}_i$ respectively. Then (Papoulis, 1990):

$$h(z) = pf(z)\left((F(z)\right)^{p-1} \tag{28}$$

and thus, by virtue of the definition of non-central *K-moment* of order (p,q):

$$\mathrm{E}[\underline{z}_p] = p\mathrm{E}\left[\left((F(\underline{x})\right)^{p-1} \underline{x}\right] = K'_{p1} \tag{29}$$

On the other hand, for positive $\underline{x}$ and large $p \to n$,

$$\left(\mathrm{E}\left[\underline{\widehat{\mu}'}_p\right]\right)^{1/p} = \left(\mathrm{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n}\underline{x}_i{}^p\right)\right]\right)^{1/p} \approx \left(\mathrm{E}\left[\left(\frac{1}{n}\max_{1\le i\le n}(x_i^p)\right)\right]\right)^{1/p} \tag{30}$$

$$\approx n^{-1/p}\mathrm{E}\left[\max_{1\le i\le n}(\underline{x}_i)\right] \approx \mathrm{E}[\underline{z}_n]$$

It is also worth noting that the multiplicative term $(p - q + 1)$ in the definitions of central and non-central K_{pq} and K'_{pq} makes K-moments generally increasing functions of p.

Appendix 2

The Climacograms of the three parameters of the near-surface air temperature (average, maximum and minimum) are presented in the following figures. Note that the climacogram derived from the empirical data is depicted in blue color, while the climacogram of the synthetic data is in green color respectively. Solid lines represent the mean of each dataset (empirical and synthetic), while dashed lines represent the 5th and 95th percentile (90% confidence levels) of the respective distributions. The climacogram derived from the optimally fitted theoretical model is depicted in red colored solid line.

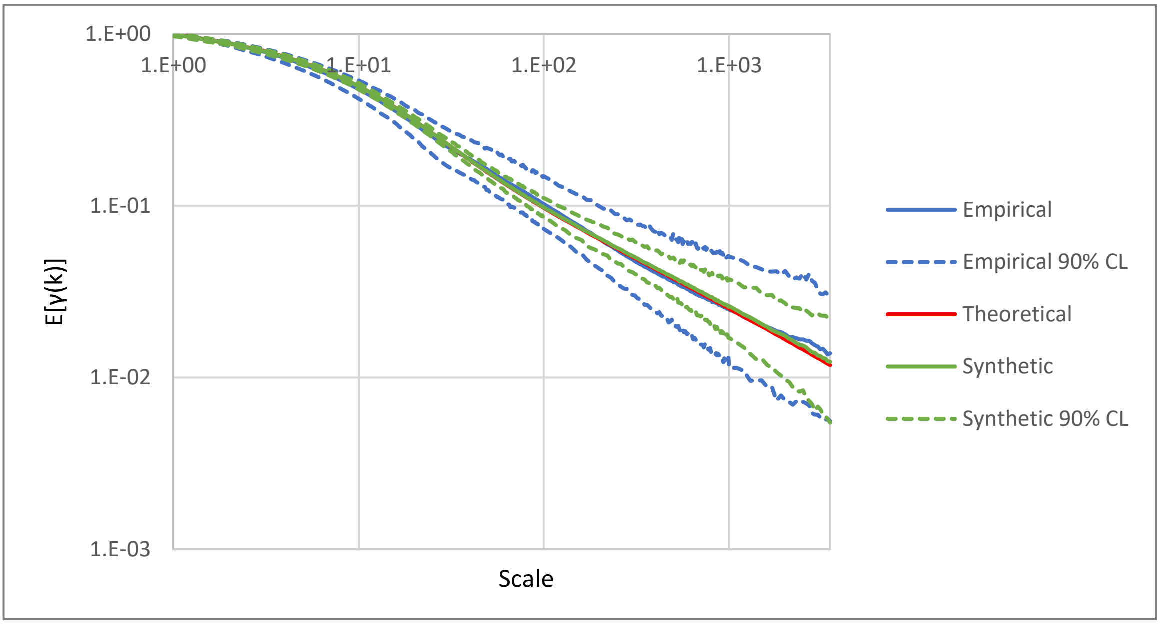

Figure 14: Climacogram of the average air temperature

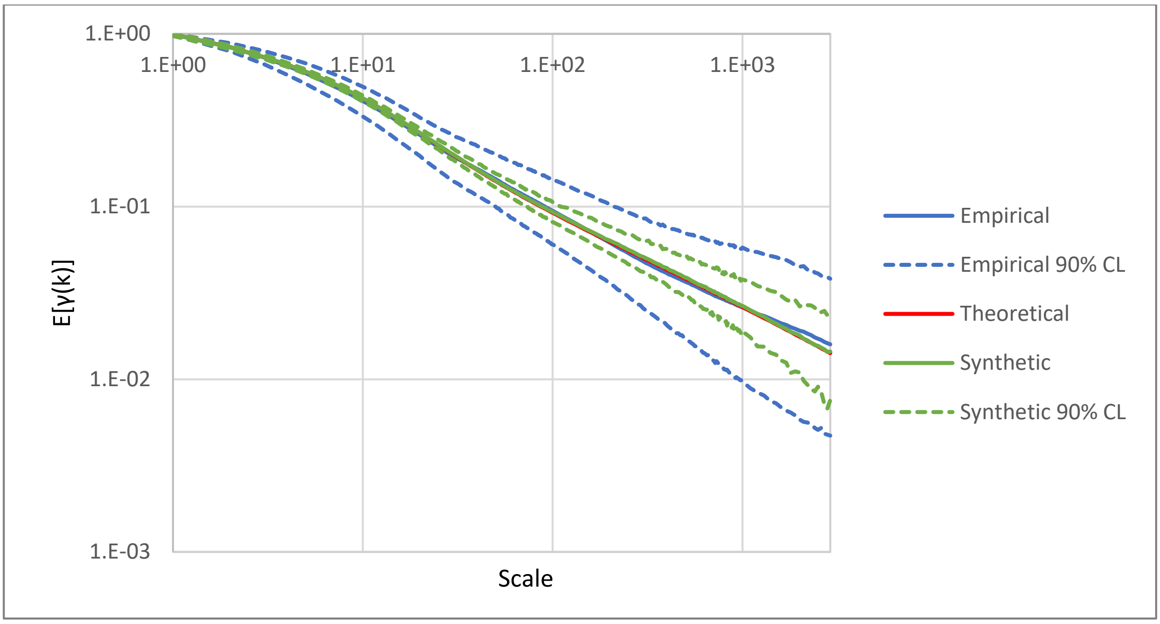

Figure 15: Climacogram of the maximum air temperature

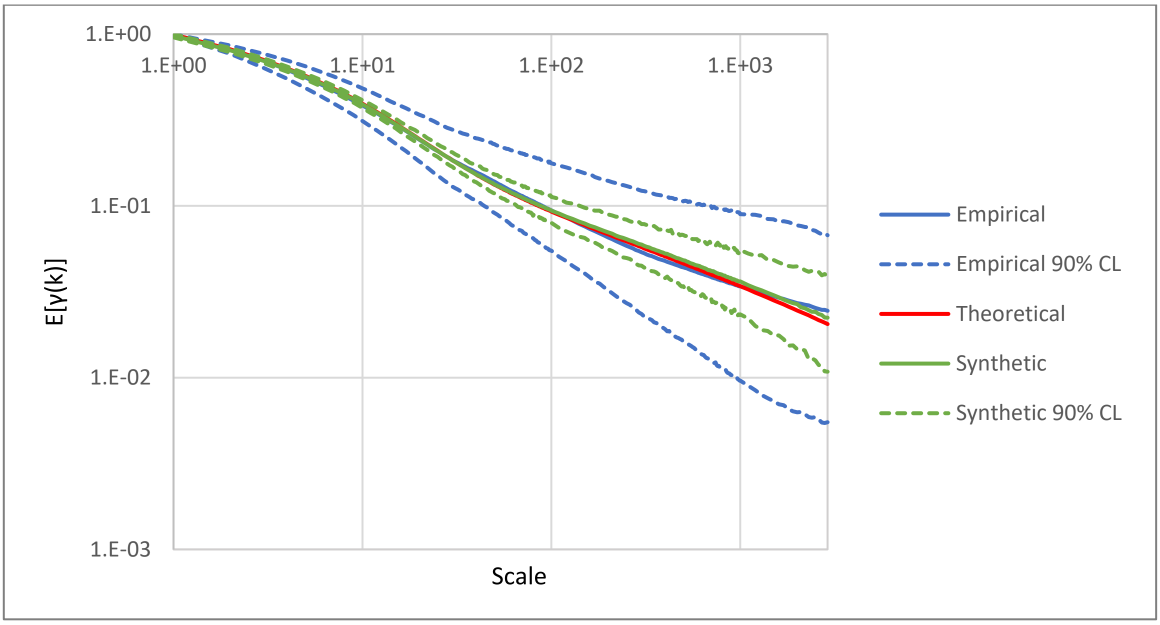

Figure 16: Climacogram of the minimum air temperature

It is worth noting that the range between the 5th and 95th percentiles of the synthetic data in each of the three climacograms is narrower than the expected one from the respective empirical data. This is probably caused by the use of the same model (imposed by the same Hurst and Mandelbrot parameters) in the production of the synthetic timeseries for each of the three parameters of near-surface air temperature (Figures 14, 15, 16).