Check for updates

OPEN Combining stochastic models of air temperature and vapour pressure for the analysis of the bioclimatic comfort through the Humidex

Beniamino Sirangelo[1], Tommaso Caloiero[2], Roberto Coscarelli[3], Ennio Ferrari[4]✉ & Francesco Fusto[5]

Several studies evidenced the importance of the knowledge of the bioclimatic comfort for improving people's quality of life. Temperature and relative humidity are the main variables related to climatic comfort/discomfort, influencing the environmental stress in the human body. In this study, a stochastic approach is proposed for characterizing the bioclimatic conditions through the Humidex values in six sites of Calabria (southern Italy), a region often hit by heat waves in summer months. The stochastic approach is essential, because the available time series of temperature and relative humidity are not long enough and present several missing values. The model allowed the characterization of sequences of extreme values of the Humidex. Results showed different behaviours between inner and coastal stations. For example, a sequence of 20 consecutive days with maximum daily Humidex values greater than 35 has a return period ranging from 10 to 20 years for the inner stations, while it exceeds 100 years for the coastal ones. The maximum yearly Humidex values for the inner stations have a larger range (40–50) than the coastal ones (38–45), reaching higher occurrence probabilities of serious danger conditions. Besides, the different influence of temperature and relative humidity on the Humidex behaviour has been evidenced.

Bioclimatic comfort is a cognitive process that mixes many stimuli influenced by physical, physiological, and psychological factors[1]. Previous researches evidenced the fundamental importance of the knowledge of the bioclimatic comfort for improving people's quality of life. In fact, in the last years temperatures over 36 °C, especially in situations of heat waves, severely affected people's health, mainly if children, elderly and chronically ill[2,3].

An increase of severe heat waves in the current century has been forecasted in areas already affected by heat stress, such as the southern part of the European continent[4–6] and the Mediterranean basin, which is expected to face particularly high impacts from global warming and climate change and to be most vulnerable to their deleterious effects[7]. The effects of heat can be various: increase of mortality[8], working/exercise capacity[9,10] and cognitive performance[11,12].

For the analysis of bioclimatic comfort, in studies conducted at different spatial scales in various parts of the world (e.g. in America[13–16], in Europe[17–19], in Asia[20–22] and in Oceania[23]) several parameters such as temperature, relative humidity of the air, wind and radiation have been analysed. In particular, given their influence on the environmental stress in the human body, temperature and the relative humidity of the air are the key basic parameters related to climatic comfort/discomfort[1]. In fact, as stated by Masterton and Richardson[24], with different humidity conditions the same temperature can provide very different feelings to people. For this reason, in order to detect the effect of heat on human health, an index that combines the effects of temperature and humidity is therefore preferred[25]. In particular, to monitor and assess this effect, several indices have been developed. One of

[1]Department of Environmental Engineering (DIAm), University of Calabria, Ponte P. Bucci, 87036 Rende, CS, Italy. [2]National Research Council of Italy, Institute for Agriculture and Forest Systems in the Mediterranean (CNR-ISAFOM), Via Cavour 4/6, 87036 Rende, CS, Italy. [3]National Research Council of Italy, Research Institute for Geo-Hydrological Protection (CNR-IRPI), Via Cavour 4/6, 87036 Rende, CS, Italy. [4]Department of Computer Engineering, Modeling, Electronics, and Systems Science (DIMES), University of Calabria, Ponte P. Bucci, 87036 Rende, CS, Italy. [5]Multi-Risk Functional Center, Regional Agency for Environmental Protection of Calabria, Viale degli Angioini 143, 88100 Catanzaro, Italy. ✉email: ennio.ferrari@unical.it

the most applied is the humidity index (Humidex) which is a relatively simple thermal comfort index based on air temperature and humidity[24]. It can be considered one of the most robust and effective comfort index, since its results are directly comparable with dry temperature in degrees Celsius and its values are associated with corresponding degrees of thermal comfort, rendering the index widely understandable. Moreover, the Humidex is easier to calculate than more complex indices (e.g. the Universal Thermal Climate Index or the Physiological Equivalent Temperature), as it requires only two input parameters (temperature and humidity) which are commonly measured by meteorological stations, because of their importance for every branch of atmospheric science[26]. In particular, the high accessibility of validation data for Humidex (e.g. in situ air temperature and humidity measurements, in contrast to mean radiant temperature usually needed by other models) has much to recommend it[25]. For these reasons, Humidex has been therefore widely used in several studies. For example, Błażejczyk and Twardosz[27] studied the variability of bioclimatic conditions in Cracow (Poland) during the period of 1826–2006. Changes in extreme heat and extreme cold events represented by various Humidex and wind chill indices were analysed in Canada for the period 1953–2012 at 126 climatological stations by Mekis et al.[28]. The impact of the urban heat island on city residents and visitors was evaluated using the Humidex in Hradec Králové, Czech Republic[29]. Giannopoulou et al.[30] performed an investigation on human thermal comfort in the Greater Athens area, during the period of June–August of 2009. The interactions between urbanization, heat stress, and climate change over the U.S. and southern Canada have been analysed by Oleson et al.[31].

The study of the characteristics of the meteorological processes can be performed using the time series observed at weather stations. However, the analysis of the extreme behaviours of these processes is often unsuitable due to the small amount of data. In the last decades, the use of generation of synthetic data series obtained from the numerical implementation of stochastic models, with statistical properties similar to those of the real processes at the chosen time scale, has largely increased. Concerning the bioclimatic index Humidex, which depends only on temperature and relative humidity, two different numerical approaches can be applied for the stochastic analysis of the joint non-stationary time-series of these variables. Specifically, a first approach considers the diurnal fluctuation of the climatological processes, by assuming them as periodically correlated random processes with a 1-day period. Through a second approach, the meteorological events are thought as non-stationary random processes, to which the method of the inverse distribution functions can be applied for simulating non-Gaussian variables[32]. A similar procedure can also be employed to analyse previously deseasonalized variables in order to remove their cyclical variability[33,34]. A further difference between the two approaches concerns the numerical implementation of the models, which is faster for the models developed with the first approach, as the second one needs the joint time series of the two meteorological variables.

The aim of this paper is to characterize the Humidex distribution in a case study (Calabria region), and to assess the influence of temperature and humidity in determining the Humidex values. With these purposes, as the available time series are not long enough and with several missing values, a stochastic approach to deal with non-stationary random processes has been proposed for the characterization of the occurrence probability of sequences of extreme values of the Humidex, in a region which is usually hit by heat waves especially in summer periods. Through this approach, the parameters of the stochastic models of temperature and relative humidity can be jointly estimated from the observed time series, and then synthetic values of Humidex can be obtained from the generated time series.

Materials and methods

The Humidex.

The Humidex (H) is an index which combines temperature and humidity to better describe the effect of heat on living organisms. The index is expressed by the empirical equation:

$$H = T + \frac{5}{9}(e - 10), \tag{1}$$

where T is the air temperature (°C), and e is the partial vapour pressure (hPa)[24]. As this last variable is not easily available, it can be evaluated by using the relative humidity, U_r (%), and the saturation vapour pressure, e_{sat} (hPa), through the relationship:

$$e = \frac{U_r \cdot e_{sat}}{100}, \tag{2}$$

where e_{sat} depends on the air temperature alone, and can be calculated through the Tetens' formula[35]:

$$e_{sat} = 6.112 \cdot 10^{\frac{7.5 \cdot T}{T + 237.7}}. \tag{3}$$

The Humidex has no specific measurement unit, anyway it can be associated to the same unit of the temperature (°C), though it is not a physical variable. As a result, the temperature perceived by human body can be easily found by using the observed values of temperature and relative humidity in Eqs. (1)–(3), and detecting the discomfort level corresponding to the Humidex value (Table 1).

Stochastic models.

The stochastic approach here proposed analyses the couple temperature—partial pressure of water vapour at the daily scale, whose data have been evaluated starting from an hourly dataset. In particular, given the higher influence of the temperature in the evaluation of the Humidex, for each day the maximum hourly temperature value and the corresponding humidity data have been considered as daily value.

Since the partial pressure of the water vapour is always positive, its natural logarithm has been considered, $Le(t) = \ln(e(t))$, in order to model couples of T-$\ln(e)$ values whose elements can be more easily treated through the linear stochastic modelling. In particular, the sequences of daily temperature, $T(i)$, and partial pressure of

Humidex	State
$H < 29$	Discomfort perceived by a few people
$30 < H < 34$	More or less significant malaise
$35 < H < 39$	Quite intense malaise. Caution. Limit some heavy physical activities
$40 < H < 45$	Sense of general malaise. Danger. Avoid efforts
$46 < H < 53$	Serious danger. Suspend physical activities
$H > 54$	Impending heatstroke (danger of death)

Table 1. Discomfort levels for different classes of the Humidex index[35].

water vapour, $Le(i)$, of the i-day starting from a generic point (i = 0,1,2,…), can be generally recognized as a realization of discrete parameter stochastic processes with cyclostationarity features in a period equal to 1 year.

Deseasonalisation and Gaussianization procedures. To deal with the seasonal features of the variables, each of the $T(i)$ and the $Le(i)$ processes can be separately reduced to a weakly stationary standardized process, respectively termed as $X(i)$ and $Y(i)$, through the transformation:

$$X(i) = \frac{T(i) - \mu_T(i)}{\sigma_T(i)}, \tag{4}$$

$$Y(i) = \frac{Le(i) - \mu_{Le}(i)}{\sigma_{Le}(i)}, \tag{5}$$

where $\mu_T(i)$ and $\sigma_T(i)$ are the mean and the standard deviation functions of the $T(i)$ process, $\mu_{Le}(i)$ and $\sigma_{Le}(i)$ are the analogous functions of the $Le(i)$ process. In particular, the functions $\mu_T(i), \sigma_T^2(i), \mu_{Le}(i)$ and $\sigma_{Le}^2(i)$ can be obtained by means of truncated expansion Fourier series that are composed by one or more harmonics, allowing for different values of the Fourier coefficients, generally estimated through the least squares method. In order to make the better choice about the number of harmonics, also taking into account the parsimony principle, the behaviour of the different Fourier series have to be compared at annual scale to the correspondent observed mean daily values of $\mu_T(i), \sigma_T^2(i), \mu_{Le}(i)$ and $\sigma_{Le}^2(i)$. This can be performed through a simple test, which however requires independent data. To this aim, subsamples from the $X(i)$ and $Y(i)$ series can be extracted for each couple of number of harmonics hypothesized for the mean and the standard deviation functions of $T(i)$ and $Le(i)$ (see Eqs. 4 and 5). Finally, by splitting the subsamples into separate classes with distinct mean and variance values, the hypotheses of equality of the means and the variances of each class can be tested[33].

The sample values of the random variables $X(i)$ and $Y(i)$ have a null mean value and a unit variance, but generally climatic variables, once deseasonalized, show skewness and kurtosis coefficients significantly different from the theoretical values expected for a normal variable[36]. On the other side, the Gaussianization step is needed for developing a coherent linear stochastic model. To cope with such problems, the deseasonalized variables $X(i)$ and $Y(i)$ can be converted into standardized normal variables, $U(i)$ and $V(i)$, respectively, by means of the transformation functions introduced by Johnson[37], whose general equation applied to the random variables $X(i)$ and $Y(i)$ provides the following relationships:

$$U = \eta_u + \theta_u \cdot \ln \left[f_X(x; \alpha_u, \beta_u) \right], \tag{6}$$

$$V = \eta_v + \theta_v \cdot \ln \left[f_Y(y; \alpha_v, \beta_v) \right], \tag{7}$$

where $-\infty < \eta_u, \eta_v < +\infty$, $\theta_u, \theta_v > 0$, $-\infty < \alpha_u, \alpha_v < +\infty$, and $\beta_u, \beta_v > 0$ are the parameters of the transformations. The functions $f_X(x; \alpha_u, \beta_u)$ and $f_Y(y; \alpha_v, \beta_v)$ can assume one of the forms known as unbounded and bounded Johnson transformations, and log-normal law with 3 parameters, depending on the sample values of the skewness ($g_{1,X}$ and $g_{1,Y}$) and the kurtosis coefficients ($g_{2,X}$ and $g_{2,Y}$). For details on the statistical procedure used for the estimation of the parameters, see Sirangelo et al.[33].

Analysis of the correlative structure. Generally, a marked persistence can still characterize the correlative structures of the sample data series $U(i)$ and $V(i)$ obtained from the Johnson transformations. These structures can be explained through *FARIMA* (fractionally differenced autoregressive integrated moving average) models[38–40], that, with reference only to the variable U, can be described by the following expression:

$$\Phi_{u,p_u}(B)(1 - B)^{d_u} u(i) = \Psi_{u,q_u}(B)\varepsilon_u(i), \tag{8}$$

where B is the backward operator, $\Phi_{u,p_u}(B)$ is the p-order polynomial of the autoregressive component of the $u(i)$ process, $\Psi_{u,q_u}(B)$ is the q-order polynomial of its mean average component, $\varepsilon_u(i)$ is a sequence of i.i.d. random variables with mean zero and standard deviation σ_{ε_u}, and d_u is its fractional order of differentiation. The *FARIMA*(p_u, d_u, q_u) model adopted for the $u(i)$ process embodies a fractional filter d_u and an *ARMA*(p_u, q_u) process, which is assumed to describe the intermediate process:

$$u^*(i) = (1 - B)^{d_u} u(i). \tag{9}$$

For each assigned value of the parameter d_u, the sample $u(i)$ can be transformed into a sample $u^*(i)$, which represents a realization of the $ARMA(p_u, q_u)$ process, that can be expressed as:

$$u^*(i) = \sum_{k_p=1}^{p_u} \phi_{u,k_p} u^*(i - k_p) + \sum_{k_q=0}^{q_u} \psi_{u,k_q} \varepsilon_u(i - k_q). \tag{10}$$

The Eqs. (8)–(10) can be implemented also for the $v(i)$ process by adopting analogous parameters ($\Phi_{v,p_v}(B)$, $\Psi_{v,q_v}(B)$, d_v, $\varepsilon_v(i)$ and σ_{ε_v}).

A procedure for the contextual estimation of the parameters of the two processes $u(i)$ and $v(i)$ can now be performed. By setting a value for each fractional filter of the FARIMA models, d_u and d_v, two intermediate realizations $u^*(i)$ and $v^*(i)$ of the processes $ARMA(p_u, q_u)$ and $ARMA(p_v, q_v)$, respectively, can be separately obtained [e.g., Eq. (9) for the $u(i)$ process]. Hence, the parameters $\phi_{u,1},...,\phi_{u,pu}, \psi_{u,1},...,\psi_{u,qu}$ and $\phi_{v,1},...,\phi_{v,pv}, \psi_{v,1},...,$ $\psi_{v,qv}$ of the two ARMA processes can be distinctly evaluated by prefixing the values of the orders (p_u, q_u) and (p_v, q_v) of the autoregressive and mean average components [e.g., Eq. (10) for the $u^*(i)$ process]. Ultimately, varying in a joint way the parameters (p_u, q_u, d_u) for the $u(i)$ process and (p_v, q_v, d_v) for the $v(i)$ process, a trial and error technique allows to obtain a correlation value between the errors ε_u and ε_v of the two ARMA models, $\hat{\rho}_{\varepsilon U*,\varepsilon V*}$, reproducing the sample cross-correlation value observed between the variables U and V, $r_{U,V}$.

Study area and data. Located at the toe of the Italian peninsula, Calabria has a typically Mediterranean climate. It features sharp contrasts due to both its position within the Mediterranean Sea and its orography. Specifically, warm air currents coming from Africa affect the Ionian side, leading to high temperatures, and to short and heavy precipitation. The Tyrrhenian side, instead, is affected by western air currents, which cause milder temperatures and more intense precipitations if compared to the Ionian side. Cold and snowy winters, and fresh summers with some precipitation, are typical of the inner areas of the region[41,42].

In this paper, the original database is composed by a set of hourly time series of air temperatures, T (°C), and relative humidity, Ur (%). In particular, six stations managed by the Multi-Risk Functional Centre of the Regional Agency for Environment Protection (Fig. 1) have been considered. Through a preliminary exploratory analysis focused on the data quality, the hourly values which lead relative humidity to jump from less than 75–100% or viceversa have been detected and consequently discarded from the original data. The main features of the databases are presented in Table 2.

Results

Parameter estimation. First, the functions $\mu_T(i)$, $\sigma_T^2(i)$, $\mu_{Le}(i)$ and $\sigma_{Le}^2(i)$ have been defined by means of the development of truncated Fourier series. In order to verify the adaptation of these functions, characterized by different numbers of harmonics, Fig. 2 shows the comparison at annual scale of the observed mean daily values and the corresponding functions of $\mu_T(i)$ and $\sigma_T^2(i)$ for the Cosenza station, and of $\mu_{Le}(i)$ and $\sigma_{Le}^2(i)$ to for the Torano Scalo station.

Due to parsimony criteria, for each function, the minimum number of harmonics to be used in the truncated Fourier series has been evaluated as the one which allows not to reject the null hypothesis with a significance level of 5%. As an example, with reference to the Cosenza station, Table 3 shows some trials of the test used for the identification of the minimum number of harmonics of the Fourier expansions. As can be seen, for the functions $\sigma_T^2(i)$, $\mu_{Le}(i)$ and $\sigma_{Le}^2(i)$ the hypothesis of one harmonic cannot be rejected for all the classes. As regards the function $\mu_T(i)$, the hypothesis of one harmonic is rejected for the 4th and the 7th classes, thus a further harmonic is needed.

Similarly to the Cosenza station, the number of the harmonics has been estimated for each station. The results, indicated in Table 4, show that in order to remove the periodicity in the mean function $\mu_T(i)$, two harmonics are needed for all the stations. For the functions $\sigma_T^2(i)$, $\mu_{Le}(i)$ and $\sigma_{Le}^2(i)$ the number of harmonics varies between 1 and 2. In particular, 2 harmonics are required for Paola and Torano Scalo as regards $\sigma_T^2(i)$, for Castrovillari, Reggio Calabria and Torano Scalo with respect to $\mu_{Le}(i)$, and for Castrovillari when the function $\sigma_{Le}^2(i)$ is considered.

The gaussianisation procedure, performed through the Johnson transformation, has been applied to the deseasonalised data series $X(i)$ and $Y(i)$. The results show that the Gaussian functions $U(i)$ and $V(i)$ required the use of the unbounded version for all the stations, with the exception of the U function for Torano Scalo, which required the bounded version.

Figure 3 shows the comparisons between theoretical and observed quantiles of the standardized Gaussian laws for the variables X and U, and Y and V evaluated for the Cosenza and Paola stations. Generally, better performances appear for the function U, while for V some discordances between theoretical and sample values have been observed, especially for the extreme ones.

Figure 4 presents the autocorrelograms of U and V evaluated for each station considering a maximum lag value equals to 30. For the variable U, a similar behaviour for all the station has been detected. In particular, the autocorrelation coefficients rapidly decrease, reaching values lower than 0.1 after 10 days. For the variable V the persistency is stronger, especially for the Cosenza and Reggio Calabria stations, for which the autocorrelation coefficients are still about 0.2 after 30 days. The autocorrelation coefficients of the remaining stations decrease faster, achieving a value of about 0.1 just after 10 days.

The identification of the $FARIMA(p, d, q)$ process allows to describe the correlative structure of the Gaussian series. The results of the process are summarised in Table 5, in which the $FARIMA$ parameters of both U and V
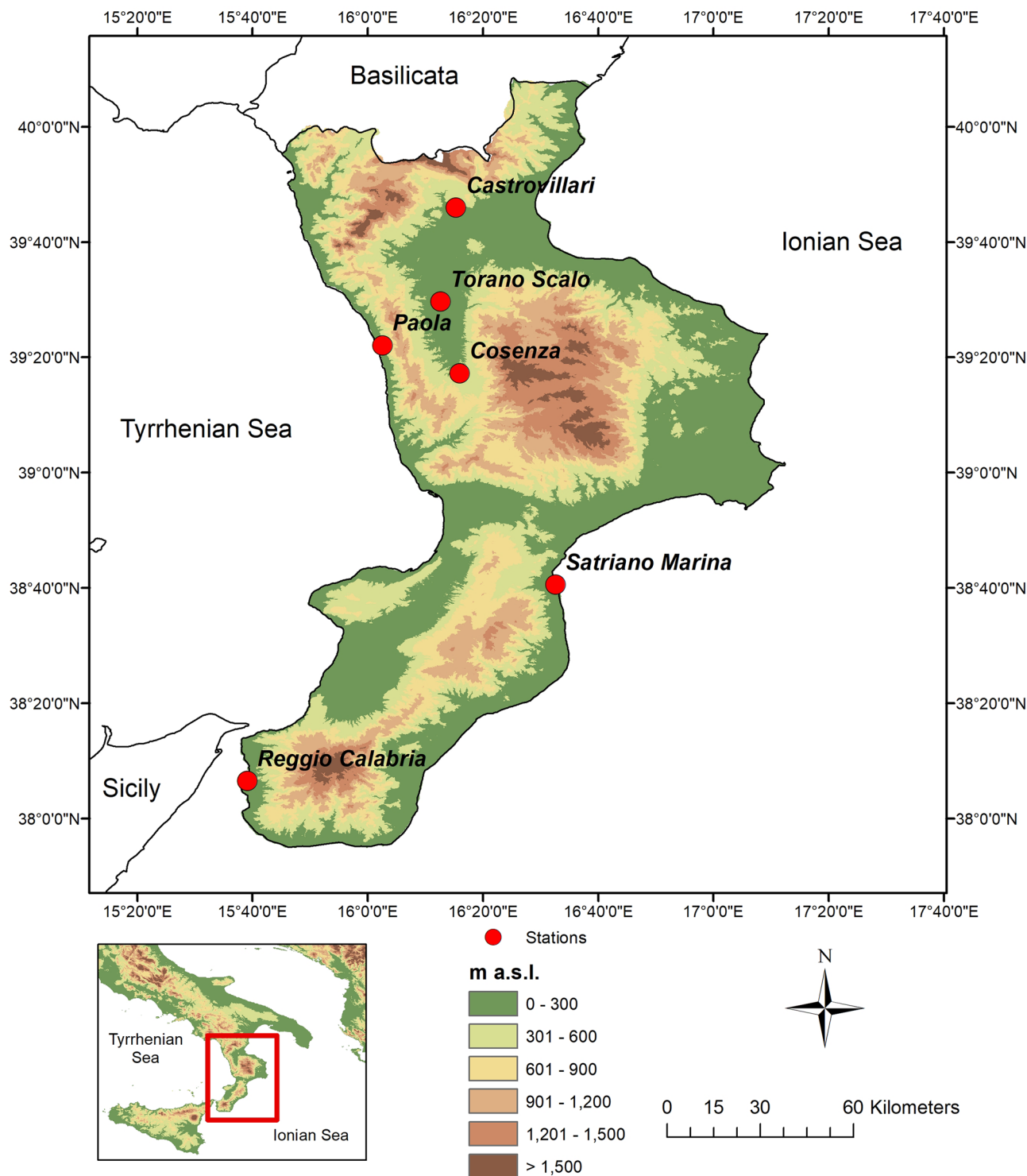
Figure 1. Localization of the stations on a Digital Elevation Model (DEM) of the Calabria region (created with Arcgis 10.4.1, https://desktop.arcgis.com/en/).

are shown for the 6 stations. A *FARIMA*$(1, \hat{d}, 0)$ model was identified in 3 out of 12 cases (Cosenza, Paola, and Torano Scalo for the variable U), while for the other 9 cases a *FARIMA*$(1, \hat{d}, 1)$ model was detected. Figure 5 shows the comparison between the experimental and the theoretical correlograms for the Castrovillari and the Torano Scalo stations for both the functions. Results clearly evidence that the *FARIMA* model well reproduces the long-term memory.

Application of the model. By means of the proposed stochastic model and the estimation of its parameters for each station, 10^5 synthetic annual series of daily values of T and Ur using a Monte Carlo procedure

Station	First day	Last day	N. of data	Missing data (%)
Castrovillari	04/05/2001	13/02/2019	6,494	2.7
Cosenza	01/01/2001	31/12/2018	6,573	1.1
Paola	03/12/2000	13/02/2019	6,646	1.6
Reggio Calabria	03/04/2001	13/02/2019	6,525	3.1
Satriano Marina	03/12/2000	13/02/2019	6,646	1.0
Torano Scalo	03/12/2000	13/02/2019	6,646	0.5

Table 2. Main features of the selected series.
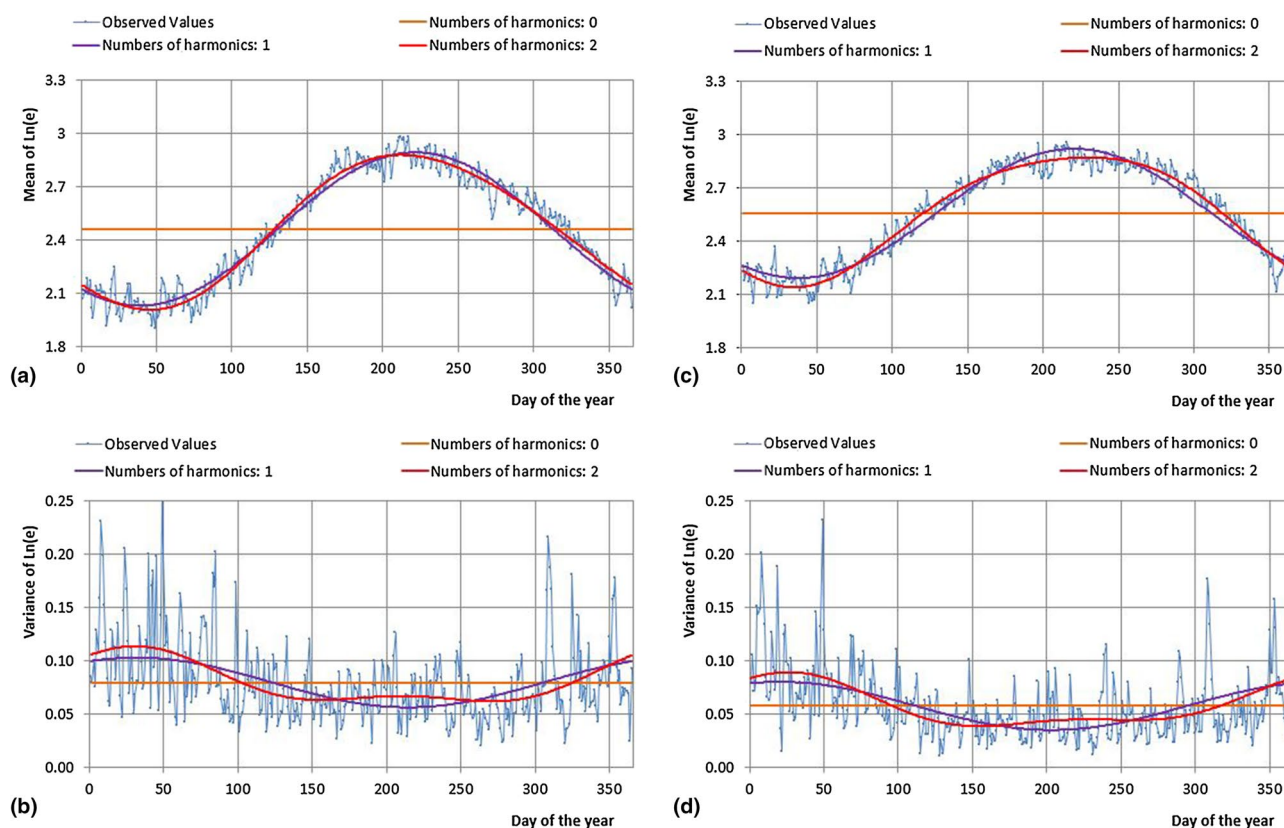
Figure 2. Comparison between observed values and Fourier series expansions (with 0, 1 and 2 harmonics) for the mean (**a**, **c**) and the variance functions (**b**, **d**) of T and Le, respectively, for the Cosenza and Torano Scalo stations.

have been generated, also taking into account leap years. The paired values of these variables have been used to evaluate a corresponding number of annual series of daily Humidex values. In Fig. 6, the average values of the maximum annual Humidex evaluated from the observed data for all the stations are superimposed to the box-plots of the maximum annual values of Humidex obtained from the synthetic series. As a result, all the sample average values fall within the range of the percentiles 25–75%, with the exception of the value observed for the Castrovillari station, which is higher than the 75% percentile.

Figure 7 shows for the Cosenza station the return period of the maximum yearly values of consecutive days with maximum daily Humidex values greater than prefixed thresholds. Obviously, for a fixed return period, the sequences of consecutive days shorten for increasing threshold values. In particular, for a return period of 10 years the sequences vary between 2 and 20 days for a Humidex threshold of 45 and 35, respectively. At the same time, considering a return period of 100 years, the sequences vary from 4 to 35 days for a Humidex threshold of 45 and 35, respectively.

Figure 8 is similar to Fig. 7 but the return periods have been estimated for all the stations and only for a Humidex threshold value equal to 35. Results show that, for their behaviour, the six stations can be divided into two groups: the first one includes the inner stations (Cosenza, Torano Scalo and Castrovillari), while the second one contains the stations near to the sea (Satriano Marina, Reggio Calabria and Paola), thus evidencing the possible influence of sea proximity. As an example, the occurrence of sequence of consecutive 20 days has a return

	Classes									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Percentiles of $\mu_T(i)$										
$m_{T,0.025}$	11.66	12.57	16.73	21.45	27.37	30.21	29.54	23.80	18.37	12.75
$m_{T,0.975}$	13.17	14.13	18.31	22.97	28.98	31.50	30.92	25.27	19.98	14.21
0 Harmonics										
$\mu_T(i)$	21.28	21.28	21.28	21.28	21.28	21.28	21.28	21.28	21.28	21.28
H_0 rejected	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
1 Harmonic										
$\mu_T(i)$	11.80	13.08	17.48	23.32	28.37	30.71	29.43	25.02	19.19	14.13
H_0 rejected	N	N	N	Y	N	N	Y	N	N	N
2 Harmonics										
$\mu_T(i)$	12.25	13.81	17.49	22.60	27.93	31.16	30.15	25.03	18.47	13.69
H_0 rejected	N	N	N	N	N	N	N	N	N	N
Percentiles of $\mu_{Le}(i)$										
$m_{ln(e),0.025}$	1.97	1.96	2.11	2.40	2.66	2.82	2.79	2.64	2.48	2.17
$m_{ln(e),0.975}$	2.11	2.11	2.23	2.50	2.77	2.91	2.90	2.74	2.60	2.30
0 Harmonics										
$\mu_{ln(e)}(i)$	2.47	2.47	2.47	2.47	2.47	2.47	2.47	2.47	2.47	2.47
H_0 rejected	Y	Y	Y	N	Y	Y	Y	Y	Y	Y
1 Harmonic										
$\mu_{ln(e)}(i)$	2.06	2.05	2.20	2.45	2.70	2.87	2.88	2.73	2.49	2.23
H_0 rejected	N	N	N	N	N	N	N	N	N	N
Percentiles of $\sigma_T^2(i)$										
$s^2_{T,0.025}$	10.16	10.89	11.76	10.64	11.82	7.63	8.32	9.60	9.81	10.21
$s^2_{T,0.975}$	19.09	20.45	22.09	20.00	22.21	14.34	15.63	18.04	18.43	19.19
0 Harmonics										
$\sigma_T^2(i)$	11.61	11.61	11.61	11.61	11.61	11.61	11.61	11.61	11.61	11.61
H_0 rejected	N	N	Y	N	Y	N	N	N	N	N
1 Harmonic										
$\sigma_T^2(i)$	12.13	13.10	13.50	13.18	12.25	11.08	10.11	9.72	10.04	10.96
H_0 rejected	N	N	N	N	N	N	N	N	N	N
Percentiles of $\sigma_{Le}^2(i)$										
$s^2_{ln(e),0.025}$	0.093	0.096	0.059	0.052	0.058	0.041	0.055	0.051	0.068	0.079
$s^2_{ln(e),0.975}$	0.174	0.179	0.112	0.098	0.110	0.076	0.104	0.096	0.127	0.149
0 Harmonics										
$\sigma_{ln(e)}^2(i)$	0.080	0.080	0.080	0.080	0.080	0.080	0.080	0.080	0.080	0.080
H_0 rejected	Y	Y	N	N	N	Y	N	N	N	N
1 Harmonic										
$\sigma_{ln(e)}^2(i)$	0.103	0.101	0.092	0.078	0.065	0.057	0.058	0.068	0.082	0.095
H_0 rejected	N	N	N	N	N	N	N	N	N	N

Table 3. Results of the rejection tests for different number of harmonics for the functions $\mu_T(i)$, $\sigma_T^2(i)$, $\mu_{Le}(i)$ and $\sigma_{Le}^2(i)$ of the Cosenza gauge.

Station	$\mu_T(i)$	$\sigma_T^2(i)$	$\mu_{Le}(i)$	$\sigma_{Le}^2(i)$
Castrovillari	2	1	2	2
Cosenza	2	1	1	1
Paola	2	2	1	1
Reggio Calabria	2	1	2	1
Satriano Marina	2	1	1	1
Torano Scalo	2	2	2	1

Table 4. Number of harmonics chosen for the functions $\mu_T(i)$, $\sigma_T^2(i)$, $\mu_{Le}(i)$ and $\sigma_{Le}^2(i)$ of all the stations.
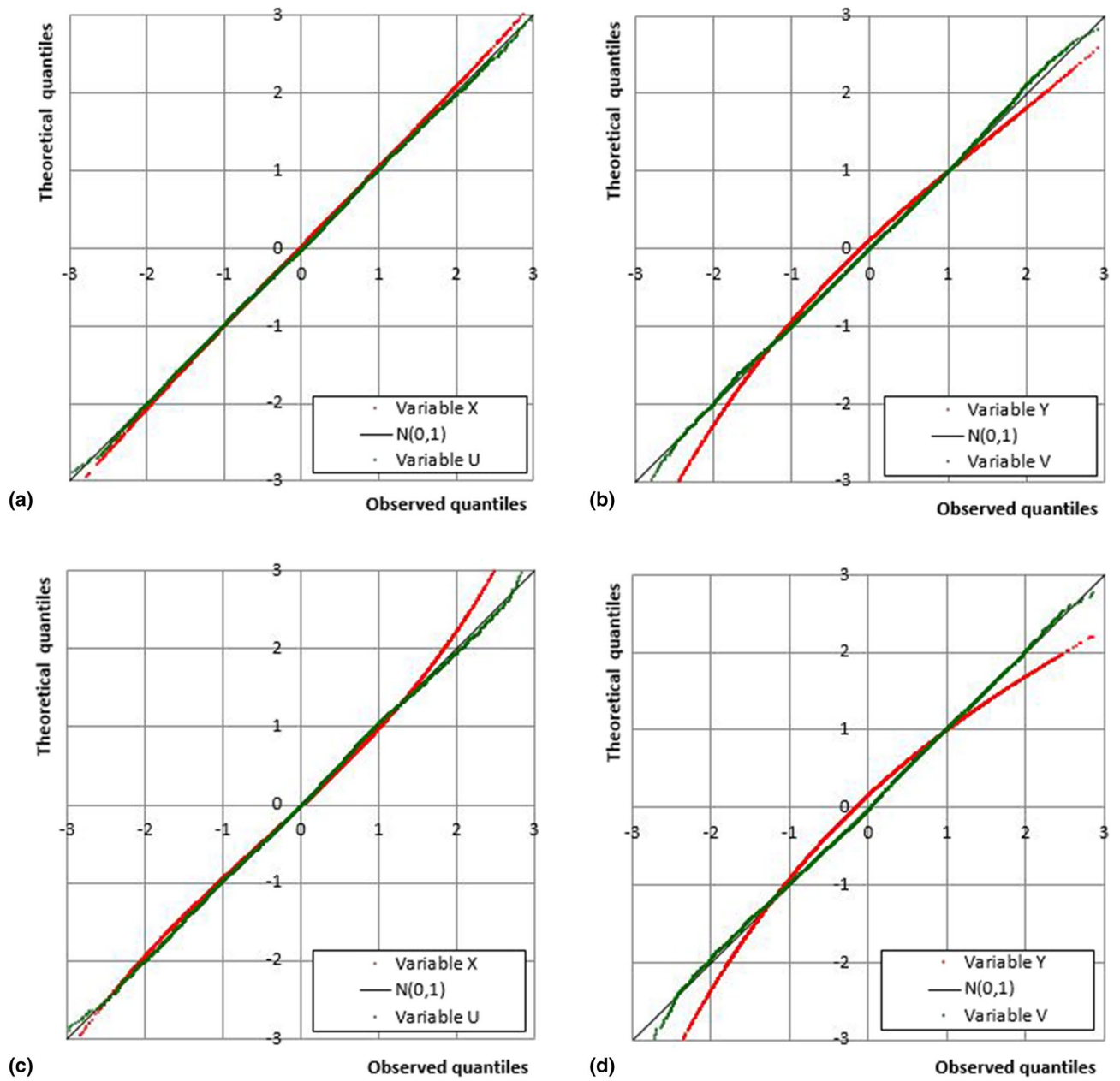
Figure 3. Q–Q plot of *X* vs *U*, and of *Y* vs *V* for the Cosenza (**a**, **b**) and Paola (**c**, **d**) gauges.
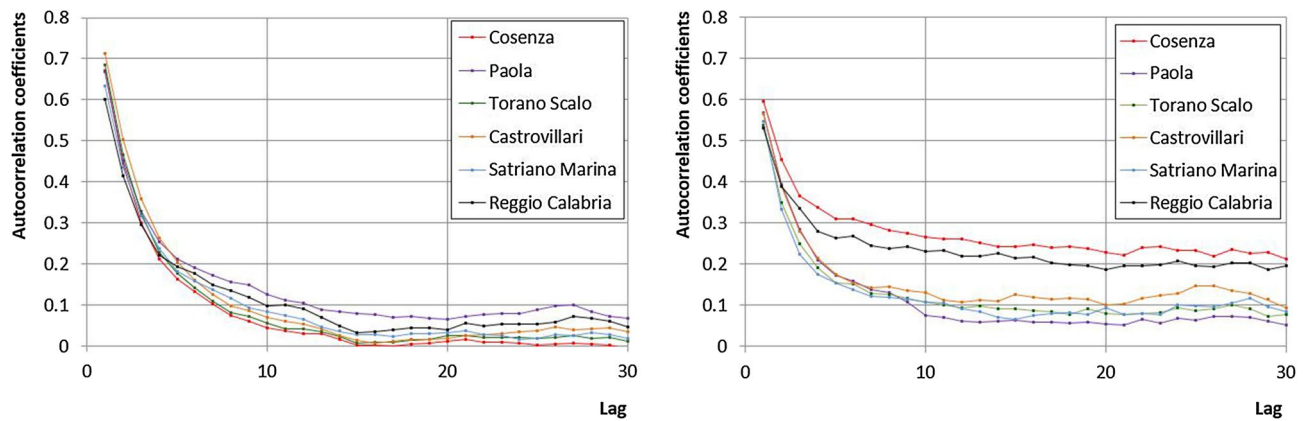

Figure 4. Autocorrelograms of the *U* (left) and *V* (right) variables for all the considered stations.

Station	Variable	FARIMA	$\hat{d}$	$\hat{\phi}_1$	$\hat{\psi}_1$	$\hat{\sigma}_\varepsilon$	$\hat{\rho}_{\varepsilon_{U_*}, \varepsilon_{V_*}}$	$r_{U,V}$
Castrovillari	U	$(1, d, 1)$	0.093	0.631	− 0.019	0.706	0.180	0.173
	V	$(1, d, 1)$	0.212	0.483	− 0.220	0.827		
Cosenza	U	$(1, d, 0)$	0.084	0.588	–	0.741	0.202	0.183
	V	$(1, d, 1)$	0.381	0.456	− 0.466	0.792		
Paola	U	$(1, d, 0)$	0.205	0.442	–	0.741	0.374	0.371
	V	$(1, d, 1)$	0.181	0.447	− 0.090	0.807		
Reggio Calabria	U	$(1, d, 1)$	0.186	0.522	− 0.175	0.791	0.140	0.136
	V	$(1, d, 1)$	0.340	0.060	− 0.047	0.834		
Satriano Marina	U	$(1, d, 1)$	0.126	0.594	− 0.140	0.771	0.202	0.199
	V	$(1, d, 1)$	0.205	0.367	− 0.096	0.839		
Torano Scalo	U	$(1, d, 0)$	0.087	0.601	–	0.728	0.137	0.132
	V	$(1, d, 1)$	0.221	0.286	− 0.014	0.833		

Table 5. Estimated values of the parameters of the *FARIMA* processes applied to all the stations.
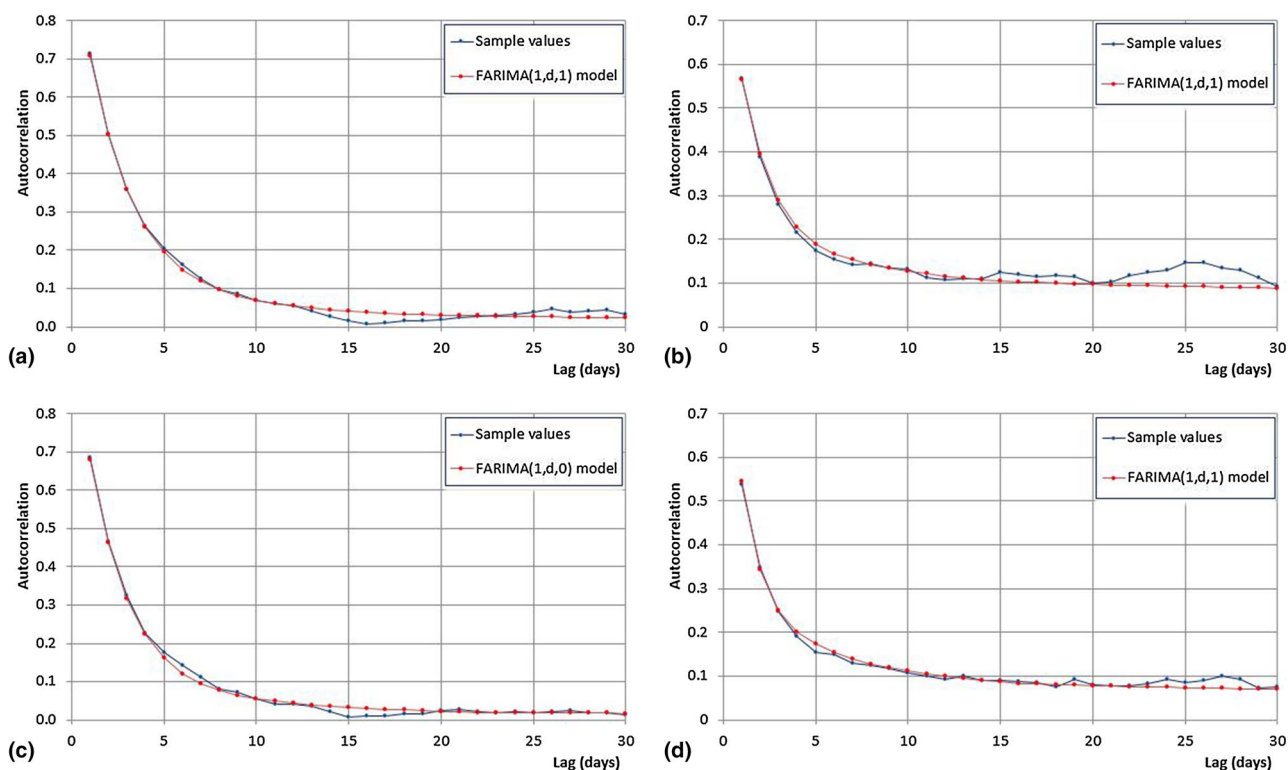
Figure 5. Comparison between theoretical and observed autocorrelograms of the variables *U* (**a**, **c**) and *V* (**b**, **d**) respectively, for the Castrovillari and Torano Scalo stations.

period ranging from 10 years (Cosenza) to about 20 years (Castrovillari). Conversely, for the stations nearer to the sea this occurrence has return periods much higher than 100 years.

It is interesting to estimate in which day of the year the highest occurrence probability to start a sequence of consecutive days with the daily maximum Humidex value higher than a prefixed threshold occurs. As an example, for the Cosenza and Satriano Marina stations (Fig. 9), results show that the lower the Humidex threshold is, later in the year the highest value of probability is registered. Moreover, the diagrams show also a tendency to a decrease of the highest probability: the higher the Humidex threshold is, the lower the occurrence probability. In fact, for the Cosenza station the highest probabilities during a year occur from the 205th (Humidex = 45; P_{max} = 0.017) to the 235th day (Humidex = 30; P_{max} = 0.019) while for the Satriano Marina station little higher probabilities than the Cosenza station have been observed, e.g. for Humidex = 30 the highest probability is 0.021.

In order to detect the influence of the temperature value in the evaluation of the Humidex, it is interesting to analyse the occurrence probability of temperatures for fixed values of the maximum daily Humidex. As an example, for the Cosenza and the Paola stations, the results of this analysis show a tendency to a decrease of the maximum probabilities when the Humidex and temperature values increase (Fig. 10). In particular, for Cosenza,
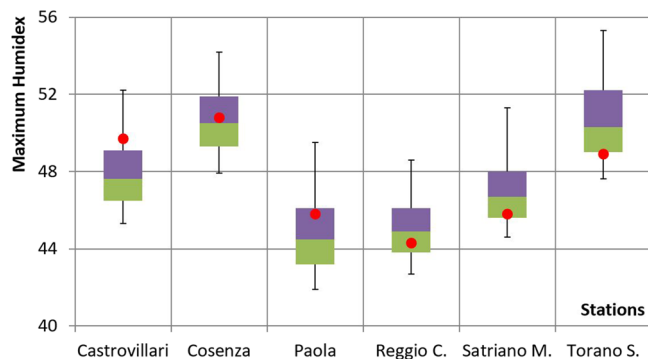
Figure 6. Comparison between the average values of the sample maximum annual Humidex (red points) and the maximum annual values of Humidex obtained from the synthetic series (box-plots). (Bottom and top of the box: 25th and 75th percentiles. Band inside the box: 50th percentiles. Ends of the whiskers: 5th and 95th percentiles. Green colour: values below the median. Violet colour: values above the median).
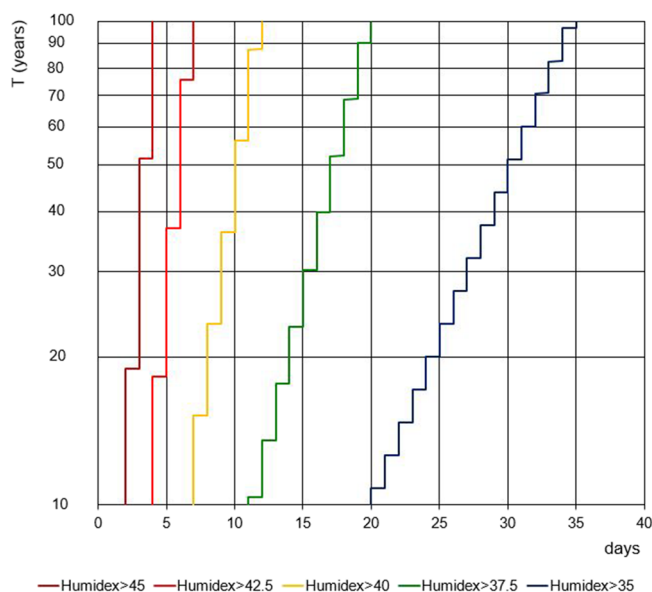

Figure 7. Annual maximum values of the number of consecutive days with maximum daily value of Humidex greater than specified thresholds for the Cosenza station.

with a Humidex value of 35 the highest probability (0.220) is reached with a temperature of 31°, while with a Humidex equal to 40, the highest probability is lower (0.208) and corresponds to a temperature of 34°. Similarly, for the Paola station, maximum probability values of 0.291 (T = 30°) and of 0.245 (T = 34°) have been observed for Humidex equal to 35 and 40, respectively.

Figure 11 shows diagrams analogous to those of Fig. 10, but referred to Ur, for the Cosenza and Castrovillari stations. Differently from the results obtained for the temperature, the curves are very close each other with the highest probability values falling within a short range of Ur values. Moreover, an opposite behaviour with respect to the temperature has been observed, with the maximum values of the probabilities that tend to fall when Humidex decreases. Specifically, for the Cosenza station, the highest probabilities range between 0.037 (Humidex = 45) for $Ur = 35\%$ to 0.029 (Humidex = 30) for $Ur = 45\%$. For the Castrovillari station, higher maximum probabilities values and lower corresponding Ur values than the Cosenza stations have been observed: these data vary from 0.047 (Humidex = 40) for $Ur = 25\%$ to 0.037 (Humidex = 32.5) for $Ur = 35\%$.

A significant analysis can be performed to evaluate the excursion of the maximum daily Humidex values from one day to another. Figure 12 shows the maximum yearly increase of this excursion in 1–7 consecutive days for different probability values (90%, 95%, and 99%), for the Cosenza and Reggio Calabria stations. For Cosenza, this increase of Humidex values ranges between 12.9 (P = 90%) and 15.3 (P = 99%) for a lag of 1 day and can reach values between 20.2 (90%) and 24.2 (99%) for a lag of 7 days. For Reggio Calabria, the increases are lower than the Cosenza station. In fact, for a lag equal to 1 day, the Humidex values can span from 10.1 (90%) to 12.7 (99%), and, for a lag of 7 days, can range between 14.6 (90%) and 18 (99%).
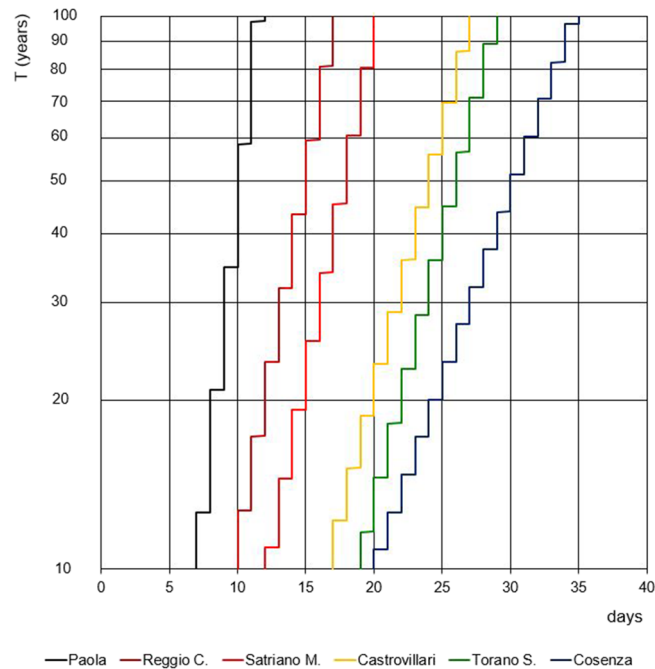
Figure 8. Comparison among annual maximum values of the number of consecutive days with maximum daily value of Humidex greater than 35 °C for all the stations.
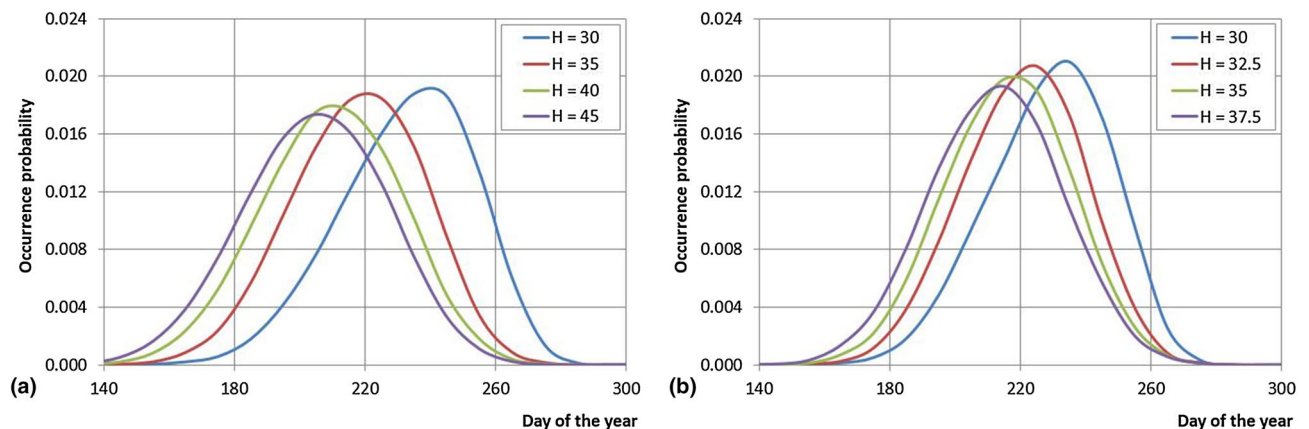

Figure 9. Occurrence distribution of the starting day of the sequences of consecutive days with maximum length of daily maximum Humidex greater than specified thresholds for the Cosenza and Satriano Marina stations.

Similarly, in Fig. 13 a comparison of the maximum yearly Humidex increases for lags 1–7 consecutive days is shown, with reference to the 95% probability value and for all the stations. The curves of the Cosenza and Torano Scalo stations are very close and present the highest values of rise. Conversely, the Reggio Calabria station shows the lowest increases. In particular, Cosenza and Torano Scalo evidenced a ΔH of about 14 in just one day, reaching an increase value of about 22 in 7 consecutive days. On the contrary, for Reggio Calabria, ΔH values of about 11 and 14 have been obtained for a lag of 1 and 7 days, respectively.

In order to evidence the combined influence of the variables T and Ur on the Humidex behaviour, the contour lines corresponding to different values of probability of the maximum yearly values of Humidex have been evaluated through the synthetic series. Figure 14 shows the results for three different probability values, i.e. 50%, 75%, and 95%, for the Cosenza and Reggio Calabria stations. For the Cosenza station the greatest part of the maximum yearly Humidex values ranges from 40 to 50, while for Reggio Calabria the curves are very close to theoretical 40-curve and ranges from about 38 and 45, thus evidencing higher probabilities of reaching serious danger conditions for Cosenza.

With reference only to the 95% probability value, the different behaviour of the maximum yearly Humidex values among all the stations is shown in Fig. 15. As a result, a marked difference has been detected between the
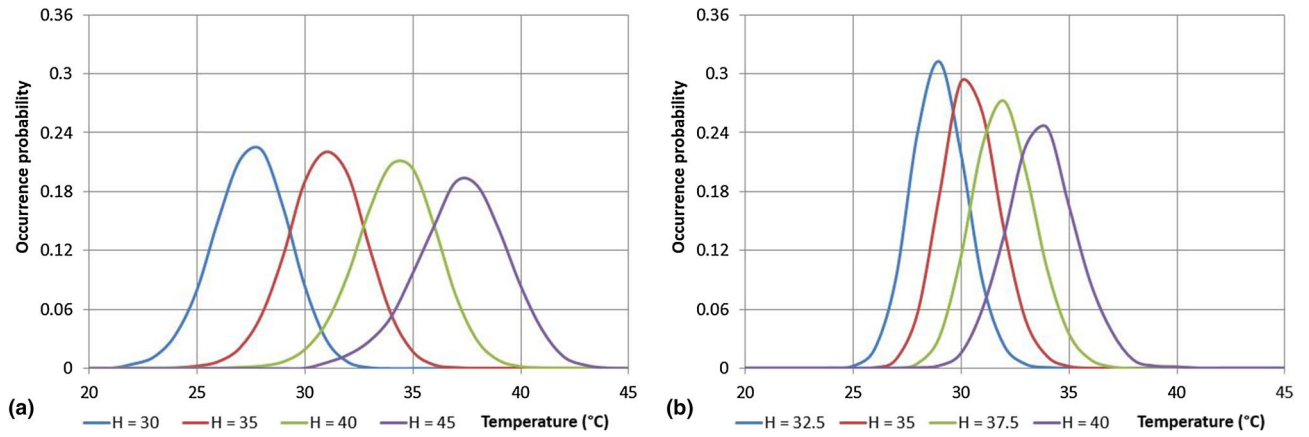
Figure 10. Distribution of the temperature for assigned values of the daily maximum Humidex for the Cosenza and Paola stations.
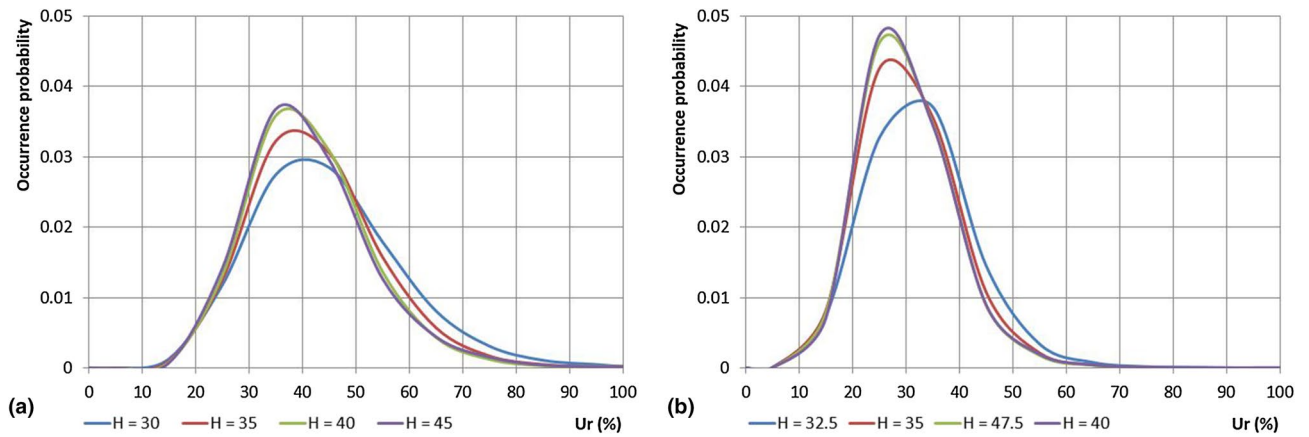

Figure 11. Distribution of the relative humidity for assigned values of the daily maximum Humidex for the Cosenza and Castrovillari stations.
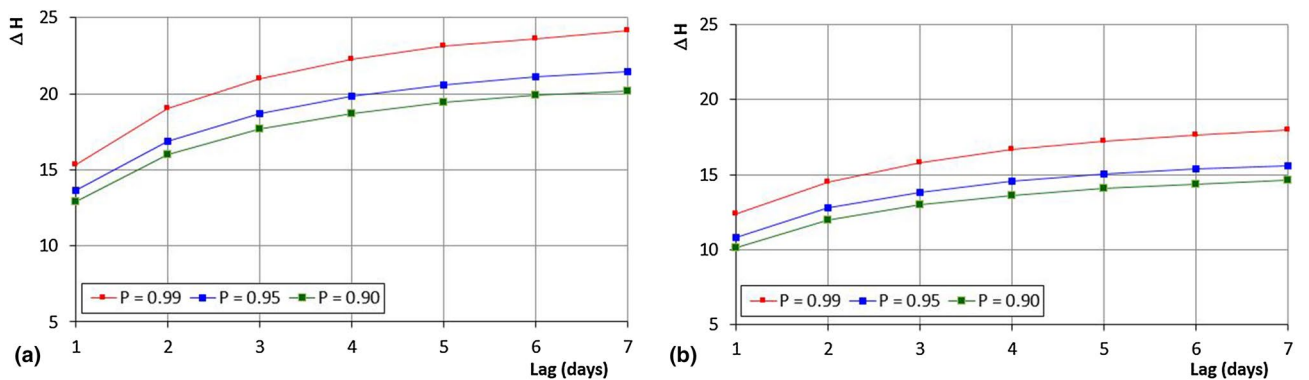

Figure 12. Quantiles of the annual maximum rises of the daily maximum Humidex for the Cosenza and Reggio Calabria stations.

stations nearer to the sea (Reggio Calabria, Paola, Satriano Marina) and the inner stations (Cosenza, Castrovillari, Torano Scalo), confirming the results shown in Fig. 13.

Discussion and conclusions

Human-perceived equivalent temperature under extremely warm periods depends on the humidity conditions. Specifically, the relative humidity can compromise the body's evaporative cooling mechanism, thus inducing a lower degree of wellness[43]. The occurrence of high value of an index such as the Humidex, which is based on
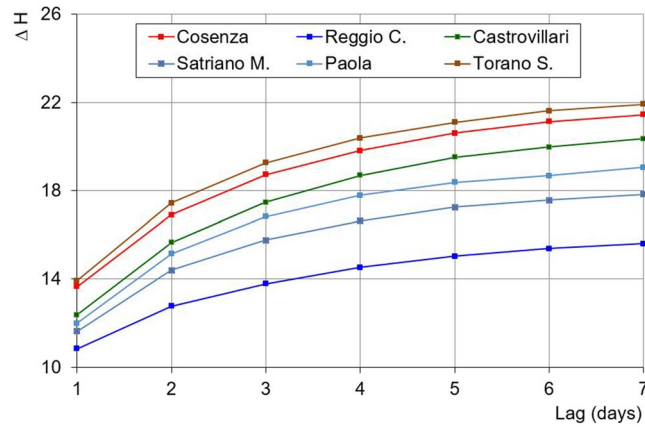
Figure 13. Comparison of the 95%-percentile of the annual maximum rises of the daily maximum Humidex for all the stations.
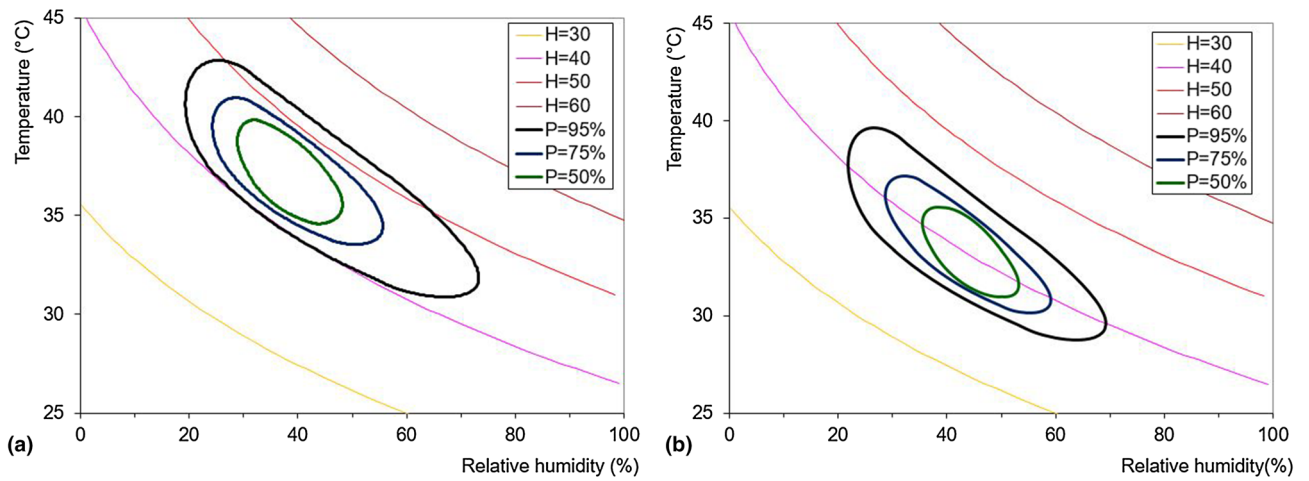

Figure 14. Behaviour of the maximum yearly Humidex values with varying temperature and relative humidity for the Cosenza and Reggio Calabria stations.
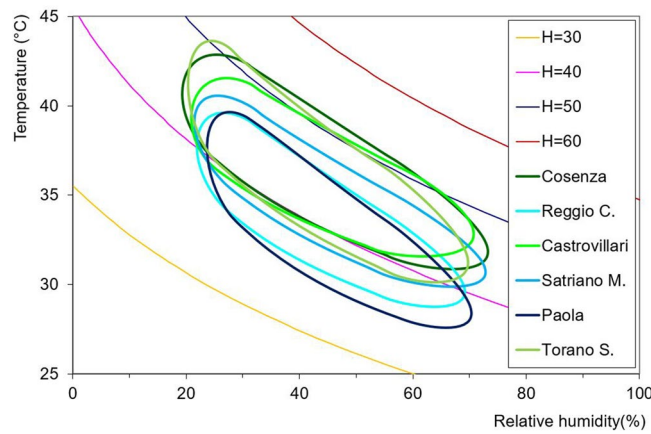

Figure 15. Comparison of the behaviour of the maximum yearly Humidex values for varying temperature and relative humidity for all the stations.

temperature and humidity, is thus paramount for its impact on human health[44–47]. In fact, extreme heat conditions are characterized by values of Humidex beyond specified thresholds, which also take into account peculiar humidity conditions increasing the impact of temperature on people's health[48].

Usually, the quality of the time series of temperature and relative humidity is not good enough for statistical purposes, presenting missing values or too low years of observation. To overcome this problem, a stochastic model has been proposed and applied to six stations of Calabria, an Italian region often coping with summer periods characterized by very high temperature[41]. The model required the use of *FARIMA* processes to describe the correlative structures of temperature and relative humidity series, once duly deseasonalized and normalized. The goodness of the stochastic model has been assessed from the comparison of extreme observed Humidex values to the simulated Humidex series obtained by synthetic generation through a Monte Carlo procedure.

Focusing on the results provided by the application of the stochastic model, the maximum yearly values of consecutive days with maximum daily Humidex value greater than prefixed thresholds, for a fixed return period, shorten for increasing threshold values. In this way, a different behaviour has been recognized between inner stations and stations located near the coast. In fact, return periods corresponding to the same sequences of consecutive days with prefixed Humidex values are much higher for the coastal stations than for the interior ones. This evidences the possible influence of sea proximity, which appears to relieve uncomfortable conditions. This result confirms the outcomes of Cannistraro et al.[49], which revealed better comfort conditions in coastal cities subject to constant ventilation than in inner areas, using different approach and index and comparing the data of various stations.

Moreover, the statistical investigation of the synthetic series evidenced the different influence of temperature and relative humidity on the Humidex behaviour. In fact, confirming what is widely stated in the literature[50], the analyses show that the discomfort conditions are significantly related to air temperature, while the impacts of humidity are of less importance. More specifically, our results also confirm that the relative humidity contributes to the most dangerous discomfort conditions only within a narrow range of percentile values, due to a lower influence of relative humidity with higher temperature values[51].

The consequences of the analyses confirm the importance of taking account of the Humidex, in order to detect promptly the occurrence of potential, unusual heat-related discomfort conditions. In other terms, the results obtained for specific communities could provide help to local health agencies in inferring about the possible occurrence of discomfort conditions.

References

1. García, M. C. Thermal differences, comfort/discomfort and Humidex summer climate in Mar del Plata, Argentina. In *Urban Climates in Latin America* (eds Henríquez, C. & Romero, H.) (Springer, Cham, 2019).
2. Díaz, J., Carmona, R., Mirón, I. J., Ortiz, C. & Linares, C. Comparison of the effects of extreme temperatures on daily mortality in Madrid (Spain), by age group: The need for a cold wave prevention plan. *Environ. Res.* **143**, 186–191 (2015).
3. Tejedor, E et al. Islas de calor y confort térmico en Zaragoza durante la ola de calor de julio de 2015. In *X Congreso Internacional AEC: Clima, Sociedad, Riesgos y Ordenación del Territorio* 141–151 (2016).
4. Sterl, A. *et al.* When can we expect extremely high surface temperatures?. *Geophys. Res. Lett.* **35**, L14703 (2008).
5. Jacob, D. *et al.* EURO-CORDEX: New high-resolution climate change projections for European impact research. *Reg. Environ. Change* **14**, 563–578 (2014).
6. Zampieri, M. *et al.* Global assessment of heat wave magnitudes from 1901 to 2010 and implications for the river discharge of the Alps. *Sci. Total Environ.* **571**, 1330–1339 (2016).
7. Segnalini, M., Bernabucci, U., Vitali, A., Nardone, A. & Lacetera, N. Temperature humidity index scenarios in the Mediterranean basin. *Int. J. Biometeorol.* **57**(3), 451–458 (2013).
8. Basu, R. & Samet, J. M. Relation between elevated ambient temperature and mortality: A review of the epidemiologic evidence. *Epidemiol. Rev.* **24**, 190–202 (2002).
9. Galloway, S. D. R. & Maughan, R. J. Effects of ambient temperature on the capacity to perform prolonged cycle exercise in man. *Med. Sci. Sports Exerc.* **29**, 1240–1249 (1997).
10. Kjellstrom, T. *et al.* Heat, human performance, and occupational health: A key issue for the assessment of global climate change impacts. *Annu. Rev. Public Health* **37**, 97–112 (2016).
11. Hocking, C., Silberstein, R. B., Lau, W. M., Stough, C. & Roberts, W. Evaluation of cognitive performance in the heat by functional brain imaging and psychometric testing. *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* **128**(4), 719–734 (2001).
12. Hancock, P. A. & Vasmatzidis, I. Effects of heat stress on cognitive performance: The current state of knowledge. *Int. J. Hyperth.* **19**, 355–372 (2003).
13. Scott, D. & McBoyle, G. Using a 'tourism climate index' to examine the implications of climate change for climate as a natural resource for tourism. In *Proceedings of the First International Workshop on Climate, Tourism and Recreation* (eds. Matzarakis, A. & de Freitas, C.) 69–98 (Halkidiki, Greece, 2001).
14. Vecchia, F. *Clima y confort humano. Criterios para la caracterización del régimen climático* (Universidad de São Paulo, São Paulo, 2000).
15. García, M. C. *El clima urbano costero de la zona atlántica comprendida entre 37° 40'0 y 38° 50' S y 57° 00' y 59° 00' W*. Dissertation, Universidad Nacional del Sur (2009).
16. García, M. C. *Clima urbano costero de Mar del Plata y Necochea-Quequén* (BM Press, Buenos Aires, 2013).
17. Olcina, J. & Miró, J. Influencia de las circulaciones estivales de brisas en el desarrollo de tormentas convectivas. *Papeles de Geografía* **28**, 109–132 (1998).
18. Andrade, H. Microclimatic variations of thermal comfort in a Lisbon city district. In *Fifth International Conference on Urban Climate, Poland* (2005).
19. Gulyas, A. & Matzarakis, A. Selected examples of bioclimatic analysis applying the physiologically equivalent temperature in Hungary. *Acta Climatol. Chorol.* **40–41**, 37–46 (2007).
20. Givoni, B. *et al.* Outdoor comfort research issues. *Energ Buildings* **35**, 77–86 (2003).
21. Lin, T. P. & Matzarakis, A. Tourism climate and thermal comfort in Sun Moon Lake, Taiwan. *Int. J. Biometeorol.* **52**, 281–290 (2008).

22. Lin, T. P. Thermal perception, adaptation and attendance in a public square in hot and humid regions. *Build. Environ.* **44**, 2017–2026 (2009).
23. Skinner, C. & Dear, R. Climate and tourism—An Australian perspective. In *Proceedings of the First International Workshop on Climate, Tourism and Recreation. International Society of Biometeorology, Commission on Climate Tourism and Recreation, Report of a Workshop Held at Porto Carras, Neos Marmaras* (eds Matzarakis, A. & de Freitas, C.) (Halkidiki, Greece, 2002).
24. Masterton, J. M. & Richardson, F. A. Humidex, A Method of Quantifying Human Discomfort Due to Excessive Heat and Humidity, CLI 1–79, Environment Canada, Atmospheric Environment Service, Downsview, Ontario (1979).
25. Geletič, J., Lehnert, M., Savić, S. & Milošević, D. Modelled spatiotemporal variability of outdoor thermal comfort in local climate zones of the city of Brno, Czech Republic. *Sci. Total Environ.* **624**, 385–395 (2018).
26. Charalampopoulos, I. A comparative sensitivity analysis of human thermal comfort indices with generalized additive models. *Theor. Appl. Climatol.* **137**, 1605–1622 (2019).
27. Błażejczyk, K. & Twardosz, R. Long-term changes of bioclimatic conditions in Cracow (Poland). In *The Polish Climate in the European Context: An Historical Overview* (ed. Przybylak, R.) 235–246 (Springer, Netherlands, 2010).
28. Mekis, É, Vincent, L. A., Shephard, M. W. & Zhang, X. Observed trends in severe weather conditions based on HUMIDEX, wind chill, and heavy rainfall events in Canada for 1953–2012. *Atmos. Ocean* **53**, 383–397 (2015).
29. Středová, H., Středa, T. & Litschmann, T. Smart tools of urban climate evaluation for smart spatial planning. *Moravian Geogr. Rep.* **23**, 47–57 (2015).
30. Giannopoulou, K. *et al.* The influence of air temperature and humidity on human thermal comfort over the greater Athens area. *Sustain. Cities Soc.* **10**, 184–194 (2014).
31. Oleson, K. W. *et al.* Interactions between urbanization, heat stress, and climate change. *Clim. Change* **129**, 525–541 (2015).
32. Kargapolova, N., Khlebnikova, E. & Ogorodnikov, V. Numerical study of properties of air heat content indicators based on stochastic models of the joint meteorological series. *Russ. J. Numer. Anal. Math. Model.* **34**(2), 95–104 (2019).
33. Sirangelo, B., Caloiero, T., Coscarelli, R. & Ferrari, E. A stochastic model for the analysis of maximum daily temperature. *Theor. Appl. Climatol.* **130**, 275–289 (2017).
34. Sirangelo, B., Caloiero, T., Coscarelli, R. & Ferrari, E. A combined stochastic analysis of mean daily temperature and diurnal temperature range. *Theor. Appl. Climatol.* **135**, 1349–1359 (2019).
35. D'Ambrosio Alfano, F. R., Palella, B. I. & Riccio, G. Thermal Environment assessment reliability using temperature-humidity indices. *Ind. Health* **49**(1), 95–106 (2011).
36. Caloiero, T., Sirangelo, B., Coscarelli, R. & Ferrari, E. An analysis of the occurrence probabilities of wet and dry periods through a stochastic monthly rainfall model. *Water* **8**(2), 39 (2016).
37. Johnson, N. L. Systems of frequency curves generated by methods of translation. *Biometrika* **36**, 149–176 (1949).
38. Hosking, J. R. M. Fractional differencing. *Biometrika* **68**, 165–176 (1981).
39. Montanari, A., Rosso, R. & Taqqu, M. S. Fractionally differenced ARIMA models applied to hydrologic time series: Identification, estimation, and simulation. *Water Resour. Res.* **33**, 1035–1044 (1997).
40. Montanari, A., Rosso, R. & Taqqu, M. S. A seasonal fractional ARIMA model applied to the Nile River monthly flows at Aswan. *Water Resour. Res.* **36**, 1249–1259 (2000).
41. Caloiero, T., Coscarelli, R., Ferrari, E. & Sirangelo, B. Trend analysis of monthly mean values and extreme indices of daily temperature in a region of southern Italy. *Int. J. Climatol.* **37**, 284–297 (2017).
42. Caroletti, G. N., Coscarelli, R. & Caloiero, T. Validation of gridded observational and modelled monthly rainfall data in Calabria (southern Italy). *Remote Sens.* **11**(13), 1625 (2019).
43. Ashcroft, F. *Life at the Extremes: The Science of Survival* (University of California Press, Berkeley, 2000).
44. Steadman, R. G. The assessment of sultriness: Part I: A temperature-humidity index based on human physiology and clothing science. *J. Appl. Meteorol.* **18**, 861–873 (1979).
45. Steadman, R. G. The assessment of sultriness: Part II: Effect of wind, extra radiation and barometric pressure on apparent temperature. *J. Appl. Meteorol.* **18**, 874–884 (1979).
46. Basu, R. High ambient temperature and mortality: A review of epidemiologic studies from 2001 to 2008. *Environ. Health* **8**, 40 (2009).
47. Massetti, L., Petralli, M., Brandani, G. & Orlandini, S. An approach to evaluate the intra-urban thermal variability in summer using an urban indicator. *Environ. Pollut.* **192**, 259–265 (2014).
48. Gaughan, J. B., Mader, T. L., Holt, S. M. & Lisle, A. A new heat load index for feedlot cattle. *J. Anim. Sci.* **86**, 226–234 (2008).
49. Cannistraro, G., Cannistraro, M. & Restivo, R. A. Smart thermo hygrometric global index for the evaluation of particularly critical urban areas quality: The City of Messina Chosen as a case study. *Smart Sci.* **2**(1), 29–35 (2014).
50. Barnett, A. G., Tong, S. & Clements, A. C. A. What measure of temperature is the best predictor of mortality?. *Environ. Res.* **110**, 604–611 (2011).
51. Scoccimarro, E., Fogli, P. G. & Gualdi, S. The role of humidity in determining scenarios of perceived temperature extremes in Europe. *Environ. Res. Lett.* **12**, 114029 (2017).

Author contributions

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to E.F.

Reprints and permissions information is available at www.nature.com/reprints.

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