IDŐJÁRÁS

QUARTERLY JOURNAL
OF THE HUNGARIAN METEOROLOGICAL SERVICE

CONTENTS

VOL. 126 * NO. 4 * OCTOBER – DECEMBER 2022

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service

Editorial Office: Kitaibel P.u. 1, H-1024 Budapest, Hungary
P.O. Box 38, H-1525 Budapest, Hungary
E-mail: journal.idojaras@met.hu

Indexed and abstracted in Science Citation Index Expanded™ and Journal Citation Reports/Science Edition
Covered in the abstract and citation database SCOPUS®
Included in EBSCO's database

Subscription by mail:
IDŐJÁRÁS, P.O. Box 38, H-1525 Budapest, Hungary
E-mail: journal.idojaras@met.hu

DOI:10.28974/idojaras.2022.4.1

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 126, No. 4, October – December, 2022, pp. 425–455

Impact of spatiotemporal land use and land cover changes on surface urban heat islands in a semiarid environment

Alireza Sadeghinia[1,*] and **Mahdi Sedaghat**[2]

[1]*Department of Geography Education*
Farhangian University, Tehran, Iran

[2]*Department of Geography*
Payame Noor University (PNU), Tehran, Iran

**Corresponding author E-mail: alirezasadeghinia@gmail.com*

(Manuscript received in final form August 30, 2021)

Abstract— This study presents the results of research that was conducted in the city of Tehran, located in the subtropics (35° N.) in a semi-desert climate in southwest Asia. The purpose was to analyze the relationship between land use/cover change (LULC) and the spatiotemporal dynamics of surface urban heat islands (SUHIs) and give results regarding the structure of the UHI in the city of Tehran. Using Landsat-5 TM data from 1986 to 2010, we quantified the spatiotemporal variability of the SUHI and LULC in the city of Tehran. The spatial distribution of land surface temperature (LST) showed the most extensive SUHI as spatially located in the western and southwestern areas of Tehran in 1986. In 2010, the spatial extent of SUHI had increased. The occurrence of LULC changes in the southern, southwestern, and especially the western parts of Tehran have played the most important role in expanding and intensifying the SUHI effect. These areas experienced two major alterations: (a) The area lost about 14 km^2 from green cover; and (b) the industrial and commercial land use, and transportation network extended significantly in these areas. Based on LULC and LST distribution patterns, barren lands, industrial and commercial land use, and transportation network have the major roles in the formation and expansion of the SUHI effect in Tehran. The SUHI of Tehran, like that of other arid or semi-arid cities, does not exhibit the classical pattern of SUHI: that is, the hot spots usually are not found in the downtown, as occurs in humid climates. Rather, the SUHI tends to situate over desert areas or barren lands that surround these cities. Therefore, an inversion of the standard SUHI phenomenon during daytime has been observed in Tehran. Research conducted in arid and semi-arid cities suggests that we should refine our point of view on the concept of the UHI in such cities and consider this issue in future studies.

Key-words: surface urban heat island, land use/cover change, land surface temperature, remote sensing, Tehran

1. Introduction

The urban heat island effect (UHI) is a critically important environmental disaster that has attracted researchers' attention for more than 150 years (*Streutker*, 2003). It refers to the phenomenon in which the air and surface temperature in an urban area is higher than that in its surrounding rural areas (*Yuan* and *Bauer*, 2007) and may be accompanied by a series of undesirable environmental effects such as adverse climate, urban diseases, and epidemics, worsened habitability, and greater energy demand for cooling. In other words, the UHI is the clearest expression of the effects of anthropogenic activity on climate at a local level (*Garcia-Cueto et al.,* 2007).

Various methods applied to investigate the UHI effect in later decades generally comprised two approaches: conventional and remote sensing-based methods. The conventional techniques mostly have depended on mathematical and statistical methods (Peterson, 2003; *Arifwidodo* and *Tanaka*, 2015). Two types of indicators, air UHI (AUHI) and surface UHI (SUHI), have been widely used to study UHI (*Li et al.*, 2017). The AUHI was calculated from the weather station network (*Chow* and *Roth*, 2006; *Karl* and *Quayle*, 1988; *Park et al*., 2017). The SUHI has been estimated from thermal infrared remote sensing techniques (*Dickinson et al*., 2010; *Li et al.,* 2017; *Voogt* and *Oke*, 2003; *Zhou et al*., 2014). *Voogt* and *Oke* (2003) described in detail the use of thermal remote sensing in the study of urban climates, and believed that the thermal remote sensing techniques are suitable to study SUHI. First, SUHI were studied using NOAA AVHRR data (*Balling* and *Brazel*, 1988; *Dousset,* 1989; *Gallo et al.,* 1993; Gallo and *Tarpley*, 1996; *Owen*, 1998); more recently, Landsat TM, ETM+, and ASTER data have been utilized for their higher spatial resolution (*Dai et al*., 2010; *Hamdi*, 2010; *Li et al.,* 2009, 2017; *Liu* and *Zhang*, 2011; *Lu* and *Weng*, 2006; *Ranagalage et al*., 2017; *Sun et al*., 2010; *Weng*, 2001, 2003; *Weng et al.,* 2006, 2007; *Weng* and *Lu*, 2008; *Xiao et al*., 2008; *Zhang et al.,* 2007). Considering the close correspondence between the distribution of land surface temperature (LST) and land use and land cover (LULC) characteristics (*Weng* and *Lu*, 2008), numerous research articles have focused on exploring the relationship between LST and urban LULC change (*Amiri et al*., 2009; *Ifatimehin*, 2011; *Lazzarini et al*., 2013; *Lu* and *Weng*, 2006; *Pal* and *Ziaul*, 2017; *Weng*, 2003, *Weng* and *Lu*, 2008; *Zhou et al*., 2014). Understanding the relationships between LULC and LST can assist in urban planning for a better scientific understanding of how the encroachment of the urban land use can form and extend patterns of SUHI (*Weng,* 2003). Urban climate research has investigated the relationship with LST of factors such as vegetation abundance, soil moisture, and the roughness properties of the land surface (*Lo et al*., 1997; *Park et al*., 2017; *Soltani* and *Sharifi,* 2017). Meanwhile, the NDVI-LST relationships more than other indices have been examined (*Kim et al*., 2005; *Liu* and *Zhang*, 2011; *Ranagalage et al*., 2017; *Weng et al*., 2004). Recently, much attention has been paid to the NDBI–LST relationship in urban

areas. For example, *Liu* and *Zhang* (2011) showed that the negative correlation of LST and NDVI means that green land weakens the SUHI effect, whereas the positive correlation between LST and NDBI means that increased development strengthens it (*Liu* and *Zhang*, 2011). Most studies of changes in LST and LULC have emphasized that overdevelopment and degradation of vegetation cover have played an important role in strengthening and intensifying the SUHI effect (*Chen et al.*, 2006).

There is a need to utilize remote sensing data in investigating the LST of cities in dry and semi-dry environment; because these cities experience extremely high temperatures in warm seasons (*Rasul et al*., 2017). The remote sensing techniques have been utilized to study some of the few SUHI studies in arid regions (*Amiri et al*., 2009; *Balling* and *Brazil*, 1988; *Falahatkar*, 2011; *Garcia-Cueto et al*., 2007; *Haashemi et al*., 2016; *Lazzarini et al*., 2013). The study of UHI, both atmospheric and surface, in the city of Mexicali showed a daily cycle in which the AUHI develops during the night but disappears in the daytime, giving way to an urban cold island. Comparing the LST of the urban area of Mexicali with its surrounding environs demonstrated that Mexicali does not show the classical pattern of AUHI; that is, higher surface temperatures are not only found closer to the urban center, as occurs in humid climates, but in this case, are also found in the surrounding desert areas.

The results of some UHI studies in arid regions were similar to those of Garcia-Cueto in the city of Mexicali (*Haashemi et al*., 2016; *Lazzarini et al*., 2013; *Zhou et al*., 2015). *Lazzarani et al*. (2013) used remote sensing data from MODIS, ASTER, and LANDSAT7 to assess land cover–temperature interactions in the Abu Dhabi metropolitan area. Their results showed an inversion of the standard SUHI phenomenon in daytime, where the downtown areas appeared cooler than the suburbs with a daily difference of 5–6 °K in summer and 2–3 °K in winter. *Zhou et al*. (2015) examined the UHI effect in 32 major cities distributed throughout different climatic zones in China using MODIS, TM, ETM+ images during the period 2003–2012. They noted that the LST differences between urban and rural areas were significantly larger than those between urban and suburban areas during the day and night for the cities, except for Lanzhou and Tianjin in the daytime. In particular, Lanzhou demonstrated a cold island effect in comparison with surrounding rural areas.

Many scholars have explored UHI in the Tehran metropolitan area. In general, these studies had the following objectives: to study spatiotemporal variability of the UHI (*Bokaie et al*., 2019; *Haashemi et al*., 2016); to identify the possible causes of UHI in Tehran (*Shahmohamadi et al*., 2015; *Shamsipour et al*., 2012); to assess the relationship between LST and LULC (*Bokaie et al*., 2016; *Rousta et al*., 2018); to investigate the severity and impact of UHI on the environmental conditions of Tehran metropolitan area (*Shahmohamadi et al.,* 2013); and to explore the various models that could be implemented to mitigate the UHI effects in Tehran (Shahmohamadi et al., 2013; *Sodoudi et al*., 2014).

Despite these previous studies, the Tehran SUHI requires more research. Apparently, a more detailed assessment of the spatial patterns of UHI is needed to better understand the structure of the Tehran UHI. Our main objectives were (1) to explore the spatiotemporal variability of the Tehran SUHI between 1986 and 2010; (2) to analyze the spatial distribution of LST and its relationship with LULC; and (3) to analyze the structure of the Tehran SUHI as a semi-desert city.

2. Data and methodology

2.1. Study area

This research was conducted in a semi-desert area located in southwestern Asia, in the city of Tehran, the capital of Iran. Tehran (*Fig. 1*), with 8 million people, has the geographical coordinates of 35°33'10"N to 35°50'12"N and 51°05'17"E to 51°37'36"E with an average elevation of 1600 m and the high Alborz Mountains in its north and northeast. The Alborz Mountains partially reduce the area of dryness. The city has poor vegetation cover. In addition, the distribution of green spaces in the city is heterogeneous (*Haashemi et al*., 2016). This area is on the subtropical high pressure belt during the summer, which makes it warm and dry. Tehran has a cold semiarid climate (Köppen climate classification: BSK) with continental climate characteristics and a Mediterranean climate precipitation pattern. Tehran's climate can be generally described as mild in spring and autumn, hot and dry in summer, and cold and wet in winter. Most of the light annual precipitation occurs from late autumn to mid-spring, but no one month is particularly wet. The hottest month is July, and the coldest is January.

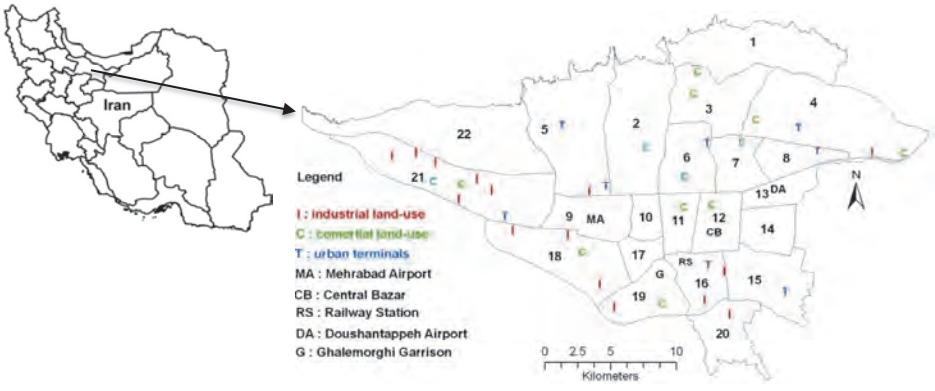


Fig 1. The study area. The most effective land use types for the establishment of UHI are presented on the map by their initials. The numbers represent the 22 districts of Tehran.

Tehran is divided into 22 municipal districts, each with its own administrative center. *Fig. 1* show the spatial distribution of 22 districts, industrial and commercial area, the central bazaar, the Mehrabad and Dooshantapeh airports, urban main terminals of railway stations, and other important phenomena that contribute to the formation of hot spots in the Tehran metropolitan area. Tehran is the most important center of population, culture, industry, commerce, and transportation in Iran, encroaching on the limited surrounding agricultural areas (*Malekpour et al*., 2010). Because of its complicated expansion, structure, and function, temperature variations are very important for its life and existence. Thus, this study has tried to analyze the relationship between the spatiotemporal distribution of SUHI and LULC change in Tehran.

2.2. Data and image preprocessing

Change detection can be defined as the process of identifying differences in the state of an object or phenomenon by observing it at different times. Using multi-sensor images to change detection is a challenge in terms of designing a suitable procedure. Ideally, change detection is conducted with multi-temporal images from the same sensor (*Lu et al*., 2014). In this paper, two landasat-5 TM images acquired on June 2, 1986 and June 4, 2010 were used to extract LST and LULC information. Some assistant data such as the air temperature and the air moisture were collected from five weather stations on June 2, 1986 and June 4, 2010 used as the input parameters to retrieve LST. The data preprocessing and other analyses were performed using ERDAS Imagine 9.2 and ArcGIS 9.3 software.

2.3. Derivation of LST

Two Landsat-5 TM images acquired on June 2, 1986 and June 4, 2010, respectively, were used to extract LST and LULC information. Some auxiliary data such as the air temperature and air moisture were collected from five weather stations on June 2, 1986 and June 4, 2010 used as the input parameters to retrieve LST. Data preprocessing and other analyses were performed using ERDAS Imagine 9.2 and ArcGIS 9.3 software.

A mono-window algorithm was applied to obtain LST from the thermal band (band 6) of Landsat TM images (*Qin et al*., 2001):

$$T_s = \{a(1 - C - D)[b(1 - C - D) + C + D]T_i - DT_a\}/C \qquad (1)$$

With $C = \varepsilon_i \times \tau_i$, $= (1 - \tau_i)[1 + (\varepsilon_i) \times \tau_i]$, a = -67.355351 and b=0.458606, where ε_i is emissivity, τ_i is the total atmospheric transmittance, T_i is the at-sensor brightness temperature (in K), and T_a represents the effective mean atmospheric temperature given by

$$T_a = 16.0110 + 0.92621 \times T_o \quad (2)$$

where T_o is the near surface temperature.

Qin et al. (2001) estimated the atmospheric transmittance from the atmospheric water vapor content (w) according to *Table 1*. Both T_o and w were obtained from local meteorological stations.

Table 1. Estimation of atmospheric transmittance (*Qin et al*., 2001)

Profiles equation (τ_i)	**water vapor (w) (g/cm^2)**	**Transmittance estimation**
High air temperature	0.4 - 1.6	0.974290 – 0.08007 w
	1.6 - 3.0	1.031412 – 0.11536 w
Low air temperature	0.4 - 1.6	0.982007 - 0.09611 w
	1.6 - 3.0	1.053710 - 0.14142 w

The emissivity can be estimated using the Normalized Difference Vegetation Index, *NDVI* (*Van De Griend* and *Owe*, 2003). A complete land surface emissivity estimation method proposed by *Zhang et al*. (2006) was utilized to calculate emissivity for each pixel (*Table 2*).

Table 2. Estimation of emissivity by using NDVI (*Zhang et al*., 2006)

NDVI	**Land surface emissivity (ε_i)**
$NDVI < -0.185$	0.995
$-0.185 \leq NDVI \leq 0.157$	0.97
$0.157 \leq NDVI \leq 0.727$	$1.0094 + 0.047 \ln (NDVI)$
$NDVI > 0.727$	0.99

2.4. Detection of LST spatiotemporal dynamics

To examine the spatial distribution of LST data during the study period and reduce the influence of seasonal difference, the resultant LST images are divided into three levels, high- (L3), normal- (L2), and low-temperature (L1) ranges (three levels), using a robust statistical method proposed by *Zhang et al*. (2007). The proposed method can be applied when the LST images follow a normal distribution. Based on *Fig. 2*, the histogram distribution of LST images for both

1986 and 2010 follow a normal distribution. In this method, the average temperature(u) ±1 standard deviation (std) is set as two threshold values that divide the LST into three ranges. *Table 3* shows the temperature classification method. The area with LST above $u+std$ would be defined as a heat island; by contrast, the area with LST below $u-std$ would be defined as a cold island. The annual difference between the mean surface temperature of the heat and cold islands could be propounded as the SUHI intensity in the study area.

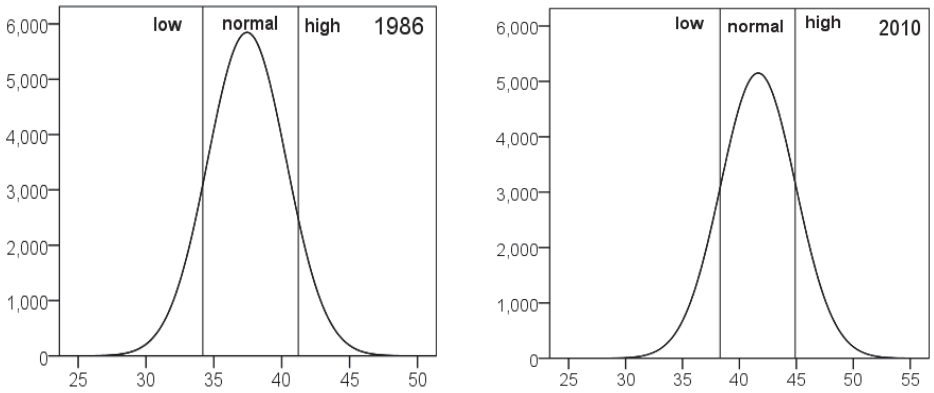


Fig 2. Distribution of LST (C°) for June 2, 1986 and June 4, 2010.

Table 3. Temperature ranges for classification of LST images

temperature classification	**interval of temperature classification**
high temperature area (L_3)	LST > u + std
normal temperature area (L_2)	u – std < LST < u + std
low temperature area (L_1)	LST < u – std

2.5. LULC classification

The proposed method by *Xu* (2007) was used to classify the LULC types in the study area. Based on the previous studies, the urban ecosystem can be broken down into several components, including impervious surface material, green

vegetation, exposed soil, and water (*Chen et al.,* 2006; *Xu,* 2007). Accordingly, the urban area can reasonably be classified into four generalized categories (i.e., built-up land, green cover, bare land, and water). In this study, to characterize these categories, the Normalized Difference Built-up Index (*NDBI*), Soil Adjusted Vegetation Index (*SAVI*), Normalized Difference Bareness Index (*NDBaI*) and Modified Normalized Difference Water Index (*MNDWI*) were applied. *MNDWI* can be employed to detect water features in an urban area. *Xu* (2007) showed that *MNDWI* can enhance the contrast between water and other land use (built-up, bare land and green cover) because other categories reflect *MIR* (*TM5*) radiation much higher than *NIR* (*TM2*) radiation. In the *MNDWI* image, the built-up land, bare land, and green land have negative values, but the water has a positive value. Therefore, *MNDWI* helps distinguish the water class from the others. MNDWI is expressed as follows (*Xu*, 2005):

$$MNDWI = \frac{Green-MIR}{Green+MIR}. \quad (3)$$

Then, a simple logic statement can easily extract the water pixels from other land uses. According to the spatial model maker tools in ERDAS 9.1, the function is as follows:

EITHER 1 IF (MNDWI>0) OR 0 OTHERWISE.

Although *NDVI* is the most important vegetation index, in this study the *SAVI* index was employed to highlight vegetation features because *SAVI* is more suitable in an area with low plant cover such as urban areas (*Ray*, 1994; *Xu*, 2007). SAVI can work in areas with plant cover as low as 15%, whereas *NDVI* can only work effectively in areas with plant cover above 30% (*Ray*, 1994). The *SAVI* can be calculated using the following equation (*Huete*, 1988):

$$SAVI = \frac{(NIR-Red)(1+I)}{NIR+Red+I}, \quad (4)$$

Where *I* is a correction factor that ranges from 0 for very high densities to 1 for very low densities. Given an intermediate vegetation density in the study area, a value of 0.5 was used. As for water, a simple logic statement could be used to extract vegetation pixels from others:

EITHER 1 IF (*SAVI*>0) OR 0 OTHERWISE.

Our studies showed numberless mixed pixels mostly comprising vegetation in combination with bare or built-up land. To prevent noise, these pixels were extracted through determination of a suitable threshold for *SAVI*. The logic calculation can be expressed as follows:

EITHER 1 IF (*SAVI*>-0.1 and *SAVI*<0) OR 0 OTHERWISE.

Another index, *NDBI*, is sensitive to the built-up area (*Zha et al*., 2003). The *NDBI* image is calculated by the following equation:

$$NDBI = \frac{MIR-NIR}{MIR+NIR}. \quad (5)$$

NDBI is a suitable index for extracting the built-up land from urban areas (*Liu* and *Zhang*, 2011), because the built-up lands have higher reflectance in *MIR* than *NIR* (*Xu*, 2007). However, *Gao* (1996) and *Xu* (2007) showed that in some circumstances, drier vegetation and water with high suspended matter concentration can also reflect *MIR* more strongly than *NIR* and, as a result, they will have positive *NDBI* values and present as noise in an *NDBI* image. Additionally, *Chen et al.* (2006) also found that *NDBI* is not sufficient to differentiate the bare land from the built-up area, because both have relatively similar spectral characteristics. Consequently, the contrast of the *NDBI* image is not as good as *SAVI* and *MNDWI* images, because many pixels of vegetation, water and bare land to be mixed with built-up area.

The last index used in this study is *NDBaI*, proposed by *Zhao* and *Chen* (2005) to retrieve bare land from the Landsat imagery. *NDBaI* can be calculated for Landsat imagery using the following equation:

$$NDBaI = \frac{band\ 5-band\ 6}{band\ 5+band\ 6}. \quad (6)$$

In Landsat imagery, the spectral characteristic of band 5 – band 6 > 0 is highly consistent with bare land, so bare land can be distinguished approximately using images with NDBaI>0. However, the proposed threshold (*NDBaI*>0) is not constant; it will change little in different regions or in different conditions of atmosphere and precipitation (*Chen et al*., 2006). Despite *NDBaI*'s relative efficiency, it is not enough to differentiate the bare land from the built-up area. Based on the explanations mentioned, urban built-up and bare land could not be extracted merely based on *NDBI* and *NDBaI* images. Therefore, this study combines the *NDBI*, *NDBaI*, *SAVI*, and *MNDWI* to extract built-up and bare land area. Based on the previous studies (*Xu*, 2007), this method can improve classification accuracy.

To achieve this object, first the *SAVI, MNDWI, NDBI*, and *NDBaI* images were produced, and then a new four band image was created through a layered stack of four images. Two methods were used to extract built-up and bare land from the new images (the new four-band image): principal component analysis (*PCA*) and logic calculation. First, *PCA* was performed on the two new images (June 2, 1986 and June 4, 2010). *Table 4* shows the results of the *PC* transformation on the new images based on the covariance matrix. The values in

the table provide the basis for determining which *PC* has the greatest loadings (values) for *NDBI* or *NDBaI* bands (representing the built-up and bare land classes), and there is also a considerable difference between loadings of bands. It is obvious that in *PC1* and *PC2, MNDWI* and *SAVI* have the greatest loadings, respectively; consequently, built-up and bare land cannot be identified from *PC1* and *PC2*. In *PC4*, although the *NDBI* band has a big loading (0.81, 0.83), the SAVI and MNDWI bands also have positive loadings in the two image and are difficult to separate from one another. Our exploration also confirmed that *PC4* is not suitable for extracting built-up and bare land. Therefore, *PC3* is more suitable than the others for separating built-up and bare land because the *NDBaI* band has a strong positive loading and the *NDBI* band has a small positive loading, in addition to the *MNDWI* band, which has a strong negative loading in *PC3* that entirely helps to differentiate LULC types. Spectral signature analysis represented by the mean of urban land use for 1986/6/2 also confirmed that *PC3* is more efficient than *PC4* for extracting both built-up and bare land, because built-up features have negative values, whereas bare lands have negative values in *PC3* (see *Fig. 3*). Finally, the suitable threshold values were used to extract built-up and bare land from the *PC3* image. According to the spatial model maker tools in ERDAS 9.1 software, the conditional function is as follows:

$$\text{CONDITIONAL } \{(PC3 < 0) <1>, (PC3 >= 0) <2>\},$$

where <1> and <2> are built-up and bare land respectively.

To evaluate classification accuracy, a random sampling method was used, with a total of 150 pixels sampled for each image. Then, the accuracy of the classification maps was verified by field study or by comparing with existing LULC maps. The overall accuracy was 89.1% for 1986 and 90.7% for 2010. The Kappa coefficient was 0.87 and 0.89 for 1986 and 2010, respectively. The precision of the classification results show that they are good enough for further spatiotemporal analysis.

Table 4. Principal component analysis on the two new four-band images

		June 2, 1986 image				**June 4 2010 image**			
		PC1	***PC2***	***PC3***	***PC4***	***PC1***	***PC2***	***PC3***	***PC4***
Eigenvectors	*SAVI*-band	0.47	0.67	0.35	0.45	0.49	0.7	0.27	0.41
	NDBI-band	-0.5	-0.25	0.13	0.81	-0.42	-0.27	0.21	0.83
	NDBaI-band	0.31	-0.55	0.76	-0.10	0.41	-0.49	0.74	-0.14
	MNDWI-band	0.64	-0.41	-0.52	0.35	0.64	-0.41	-0.56	0.32

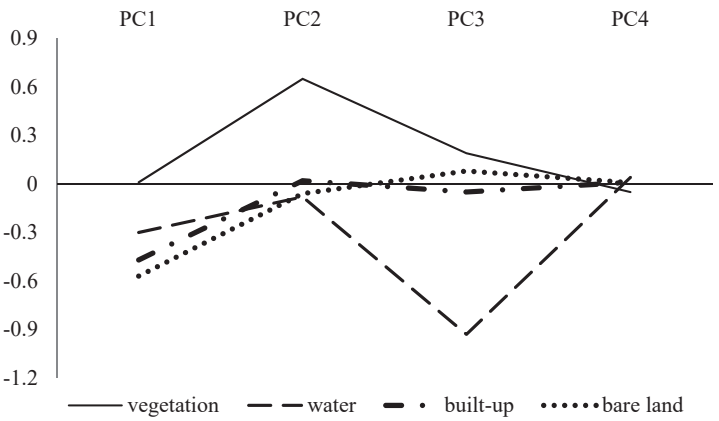


Fig 3. Spectral signatures represented by the mean of urban land use classes of Tehran for June 2, 1986 as an example.

3. Results

3.1. Spatiotemporal distribution of SUHI

Fig. 4 represents the general pattern of LST spatial distribution in Tehran for both 1986 and 2010. In these maps, the LST data were classified into three groups based on the classification scheme of standard deviation that aforesaid in methodology section. The LST patterns show significant differences between the western and southwestern parts of Tehran in comparison with the northern part. The statistics of LST of each image are summarized in *Table 5*.

Table 5. Summary statistics of LST data for Tehran (°C)

Date of image	mean	maximum	minimum	Standard deviation	SUHI intensity
June 2, 1986	37.2	47	11	3.02	10
June 4, 2010	41.6	54	27	3.3	10.15

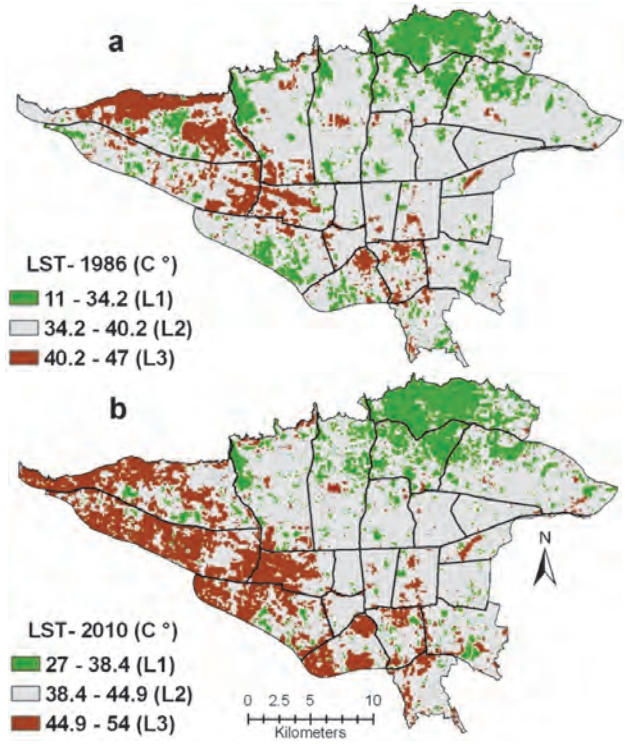


Fig 4. LST distribution images in June 2, 1986 (a) and June 4, 2010 (b).

The statistics of LST on June 2, 1986 indicates that the lowest LST was 11 °C, the highest LST was 47 °C, and the mean was 37.2 °C, with a standard deviation of 3.02. In 1986, some hot spots can be clearly observed (*Fig 5a*). The most extensive hot spots were distributed in the western and southwestern parts of Tehran, over Mehrabad international airport, its surrounding barren land, and a special industrial zone in the west of Tehran. The hottest part of the Tehran SUHI is located over the Mehrabad international airport with approximate area of 900 hectares in district 9. The other hot spots are dispersed all over the city, especially in the old downtown and central bazaar, Dooshantappeh airport, the railway station, bus terminals, grain silo, and factories (see *Fig. 1* for their spatial positions). However, the coldest area of Tehran (cold island) was located in the green land area north of Tehran, especially in the Shemiranat and considerable portions of districts 2, 3 and 4. The other dispersed cold spots mostly correspond to green areas and urban parks, with the parks being the coolest spots. The spatial

pattern of LST on June 4, 2010 differs markedly from that of June 2, 1986, as *Fig. 4* shows. The statistics of LST on June 4, 2010 indicates that the lowest LST was 27 °C, the highest LST was 54 °C, and the mean was 41.6 °C, with a standard deviation of 3.3 (*Table 5*). The most significant difference between 1986 and 2010 is seen in the western, southwestern, and southern parts of Tehran. Although the most extensive SUHI, previously located in the western and southwestern area of the study area, has slightly shifted to the westward in 2010, its extent has increased. Two major hot spots appeared in the south of district 19 (the southern part of Tehran), and a major hot spot was formed in the southwestern part of the city over district 18. Several hot spots also appeared in the western area over districts 21 and 22.

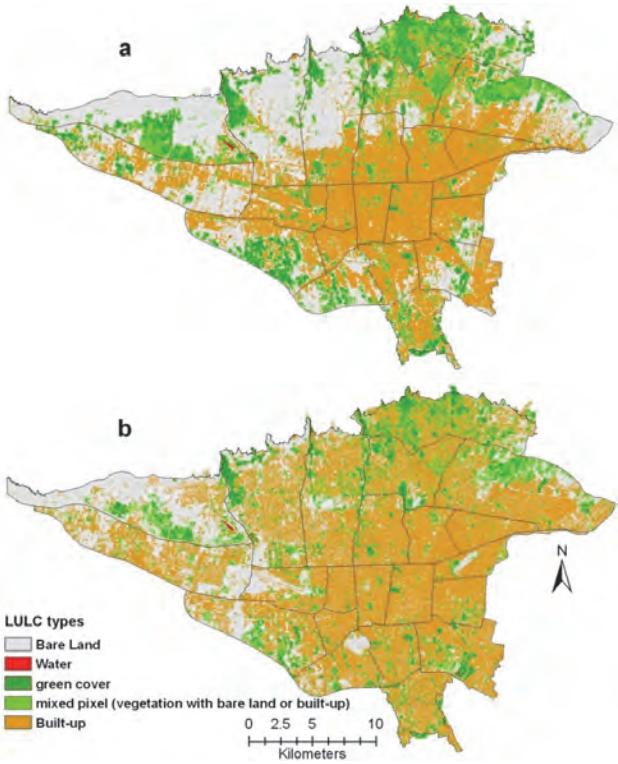


Fig. 5. the LULC distribution in June 2, 1986 (a) and June 4, 2021 (b)

Based on the distribution pattern of hot spots in the period 1986–2010 that characterized the Tehran SUHI, it does not exhibit the classical SUHI pattern like

that of other cities that have been located in arid or semiarid climates; that is, the hot spots usually are not found in central parts of cities, as occurs in humid climates, but instead they tend to situate over desert areas and surrounding bare lands. Therefore, an inversion of the standard SUHI phenomenon has been observed during daytime in Tehran, where the downtown area appears colder than the western, southwestern, and southern areas. This finding confirms the results of Haashemi *et al.* (2016) regarding the structure of the Tehran SUHI.

The result described above has also been observed in other arid and semi-arid cities such as Phoenix (*Balling* and *Brazil*, 1988), Mexicali (*Garcia-Cueto et al*., 2007), and Abu Dhabi (*Lazzarini et al*., 2013).

The results of three temperature areas and the associated changes in each class from 1986 to 2010 are listed in *Table 6*. Class 1 (L_1) comprises values with more than one standard deviation below the mean that shows cold islands; class 2 (L_2) values are normal values situated between the mean and one standard deviation below or above the mean; and class 3 (L3) values are heat islands with more than one standard deviation above the mean. According to the resulting table, the high-temperature area (L3) has increased about 31 km^2 from 1986 to 2010; in contrast, the low-temperature area has decreased about 3 km^2. The values in Table 3 also show that SUHI intensity has increased slightly over time, from 10 °C in 1986 to 10.2 °C in 2010. Furthermore, based on the evolution process of LST data pattern from 1986 to 2010, the area and intensity of the Tehran SUHI has rapidly spread and strengthened, while the cold urban island has weakened.

Table 6. Area in different temperature scales from 1986 to 2010 (km^2)

Range	Area in 1986	Area in 2010	Change between 1986 to 2010
High temperature area (L3)	83	114	+31
Normal temperature area (L2)	449	421	-28
Low temperature area (L1)	82	79	-3

3.2. Relationship between LULC characteristics and SUHI patterns

Table 7 summarizes the average LST value of LULC types. The highest and lowest surface temperature relate to bare lands and water bodies, respectively. After bare lands, which exhibited the highest surface temperature in the study area (38.3 °C in 1986 and 43.3 in 2010), built-up areas had the highest surface temperature (37.8 °C in 1986 and 41.5 °C in 2010), followed by mixed pixels (35.4 °C in 1986 and 39.8 °C in 2010), green cover (32.9 °C in 1986 and 37.6 °C in 2010), and water bodies (30.3 °C in 1986 and 33.2 °C in 2010). The maximum LST difference of LULC types was 8 °C and 10 °C in 1986 and 2010, respectively.

Therefore, the difference between the average value of surface temperature of water bodies (the coldest of land cover) and bare lands (the hottest land cover) has increased about 2 °C. The average values of LULC types for both 1986 and 2010 implies that bare lands and built-up areas have played the major roles in forming and expanding SUHI effects in Tehran. Of course, all the built-up areas did not exhibit the same surface temperature; some land use, such as transport and industrial land use, displayed higher surface temperatures than residential land use. Apparently, residential land use is less effective in promoting hot spots in Tehran. Spatially, the Tehran SUHI has been closely related to bare lands and to transport and industrial land use. These areas mostly have been covered with non-evaporating, non-transpiring materials such as asphalt, metal, concrete, and stone. In contrast, green lands showed considerably lower LST, because vegetation can reduce the amount of stored heat in the land surface through evapotranspiration and shadow. Extensive green spaces in the northern part of Tehran (districts 1, 2, 3, and 4) contributed to lower surface temperatures and the formation of an urban cold island there.

Table 7. The mean value of LST of LULC types

LULC types	Mean temperature (1986)	Mean temperature (2010)
Water	30.3	33.2
Green cover	32.9	37.6
Mixed pixel	35.4	39.8
Built-up	37.8	41.5
Bare land	38.3	43.3

Fig. 5 shows LULC distribution in Tehran for 1986 and 2010. The built-up areas dominated in both 1986 and 2010, occupying about 247.2 km^2 (40%) and 310.7 km^2 (50%), respectively (*Table 8*). Although the built-up area has increased about 64 km^2 from 1986 to 2010, areas of other LULC have decreased. Water bodies can be disregarded, because they have very small surface area. Because built-up areas have expanded in all directions, the bare lands and green lands lost about 37.5 km^2 and 23.9 km^2, respectively. LULC changes in the southern, southwestern, and particularly the western parts have played an important role in expanding and intensifying the SUHI effect in our study area. Because of the expansion of industrial and commercial land use and the development of the transportation network in the western and southwestern parts, several hot spots (or urban heat islands) have appeared. The most expansion of the industrial land use has occurred along the western highways, such as the Karaj highway and the old Karaj road. Similar changes have also occurred in district 18 (southwest of Tehran) and 19 (the south of Tehran). The other important alteration has occurred

in the vegetation cover areas of districts 18, 19, 20, 21, and 22 (the west, south, and southwest of Tehran). Based on our analysis, these districts lost vegetation cover of about 4.6, 1.3, 2.0, 2.8, and 3.3 km^2 respectively. The largest LST pattern alteration occurred in areas in which dark (asphalt, tar, concrete, etc.) and metallic materials (galvanized or aluminum roofs) replaced the concentrated vegetation cover. The industrial and commercial land use areas, transportation network, and warehouses are mostly covered by dark materials and metallic roofs; consequently, they are associated with a high LST. In the bare lands and desert areas, because of the dry nature of non-evaporating materials, the LST is high.

Table 8. Area of LULC types in Tehran from 1986 to 2010

LULC types	Area (1986) km^2	Area (2010) km^2	Change in LULC between 1986 to 2010
Water	0.35	0.34	-0.1
Green cover	76.6	52.7	-23.9
Mixed pixel	49.8	47.6	-2.2
Built-up	247.2	310.7	63.5
Bare land	237	199.5	-37.5

4. Conclusion

In this study, the relationship between LULC changes and the spatiotemporal dynamics of the SUHI of Tehran, Iran were investigated for 1986 and 2010. The results showed that the area of SUHI has increased about 31 km^2, whereas the urban cold island (UCI) area has decreased about 3 km^2. The SUHI intensity (SUHII) also slightly increased. The most extensive SUHI was spatially located in the western and southwestern parts of Tehran in 1986. However, in 2010, SUHI has slightly shifted westward, and its extent has increased; new hot spots have appeared in the western, southwestern and southern parts of Tehran. The highest LST relates to bare lands and built-up areas, followed by mixed pixels, green cover, and water bodies. The precise consideration of LST in the study area indicates that bare lands, industrial and commercial land use areas, and the transportation network have played major roles in the formation and expansion of the SUHI effect in Tehran; in contrast, the residential land use area is less effective in promoting the SUHI effects. In the period 1986–2010, the built-up area increased about 64 km^2, but the bare lands and green lands lost about 37.5 km^2 and 23.9 km^2, respectively. The occurrence of LULC changes in the southern,

southwestern, and especially the western areas of Tehran have played the most important role in expanding and intensifying the SUHI effect. These areas experienced two major alterations: (a) districts 18, 19, 20, 21, and 22 lost about 14 km^2 from green vegetation cover from 1986 to 2010; (b) the industrial and commercial land use areas and the transportation network increased extensively in these areas. This study showed that in arid and semiarid cities, the largest alteration of LST occurs in areas where dark (asphalt, tar, concrete, etc.) and metallic materials (galvanized or aluminum roofs) replace the concentrated vegetation cover (similar to those changes in the western and southwestern areas of Tehran). Therefore, protection and expansion of vegetation cover has great importance in mitigating the SUHI effect. It seems that in semi-desert cities such as Tehran, where dry bare lands, industrial and commercial land use areas, and extensive transportation networks situated together on the city outskirts provide an opportunity for the emergence of an extensive SUHI.

The spatial distribution pattern of the SUHI phenomenon in our study area indicated that the SUHI of arid and semiarid cities (such as Tehran) differ considerably from that of cities situated in humid climates, which usually exhibit the classical SUHI pattern. In other words, the SUHI is usually located downtown, but in arid and semiarid cities, the SUHI tends to be situated over desert areas or extensive bare lands located in the suburbs. Based on the spatial distribution of the LST in Tehran, the highest LST is not observed over the downtown area, but rather increases as we move from the downtown toward the west, southwest, south, and east. The change pattern from the downtown toward the north area differs from that in other directions. Toward the northern area (districts 1, 2, 3, and 4), the LST decreases away from the downtown, because the maximum vegetation rate is concentrated in the northern area of the city. Therefore, an inversion of the standard daytime SUHI phenomenon has been observed in Tehran. Our results regarding the structure of SUHI in Tehran correspond to the results of other research exploring the UHI in dry and semiarid cities (*Balling* and *Brazil*, 1988; *Garcia-Cueto et al*., 2007; *Haashemi et al*., 2016; *Lazzarini et al*., 2013). For instance, *Garcia-Cueto et al*., (2007) examined AUHI and SUHI and their relationship with LULC in the city of Mexicali. Their results showed that during the daytime, in any season of the year, the city becomes a UCI. They found that Mexicali does not show the classical pattern of a UHI; that is, the higher surface temperatures are not only found toward the downtown, as occurs in humid climates (*Garcia-Cueto et al*., 2007). Also, the study of the Abu Dhabi SUHI showed a daytime inversion of the standard SUHI phenomenon in which the downtown areas appeared colder than the suburbs. Abu Dhabi has a hot desert climate (*Lazzarini et al.,* 2013). *Zhou et al*., (2015) investigated the UHII in 32 Chinese cities. Their results showed that the direction of the UHII might be reversed in cities such as Lanzhou located in an arid climate. Lanzhou presented the cold island effect (negative) in comparison with rural areas (*Zhou et al*., 2015). Our results resembled the findings of *Haashemi et al*., (2016), suggesting that, in

semiarid cities such as Tehran, with the urban–rural indicator, a surface urban cool island may be observed in daytime. Therefore, research conducted in arid and semiarid cities suggests the need to refine our point of view on the concept of UHI in arid and semiarid cities and consider this issue in future studies.

References

Amiri, K., Weng, Q., alimohamadi, A., and *Alavipanah, K*., 2009: Spatial–temporal dynamics of land surface temperature in relation to fractional vegetation cover and landuse/cover in the Tabriz urban area, Iran. *Remote Sens. Environ. 113*, 2606–2617. https://doi.org/10.1016/j.rse.2009.07.021

Arifwidodo, S.D. and *Tanaka, T*., 2015: The Characteristics of Urban Heat Island in Bangkok, Thailand. *Procedia – Soc. Behavior. Scie. 195*, 423–428.https://doi.org/10.1016/j.sbspro.2015.06.484

Balling, R.C. and *Brazell, S.W*., 1988: High resolution surface temperature patterns in a complex urban terrain. *Photogram. Engineer.Remote Sens. 54*, 1289–1293.

Bokaie, M., Zarkesh, M.K., Arasteh, P.D., and *Hosseini, A.,* 2016: Assessment of urban heat island based on the relationship between land surface temperature and land use/land cover in Tehran. *Sustain. Cities Soc. 23*, 94–104. https://doi.org/10.1016/j.scs.2016.03.009

Bokaei, M., Shamsipour, A., Khatibi, P., and *Hosseini, A*., 2019: Seasonal monitoring of urban heat island using multi-temporal Landsat and MODIS images in Tehran. *Int. J. Urban Sci. 23*, 269–285. https://doi.org/10.1080/12265934.2018.1548942

Chen, X.L., Zhao, H.M., Li, P.X., and *Yin, Z.Y*., 2006: Remote sensing image-based analysis of the relationship between urban heat island and land use/cover changes. *Remote Sens. Environ. 104*, 133–146. https://doi.org/10.1016/j.rse.2005.11.016

Chow, W.T.L. and *Roth, M.,* 2006: Temporal dynamics of the urban heat island of Singapore. *Int. J. Climatol. 26*, 2243–2260.https://doi.org/10.1002/joc.1364

Dai, X., Guo, Z., Zhang, L., and *Li, D.,* 2010: Spatio-temporal exploratory analysis of urban surface temperature field in Shanghai, China. *Stoch. Environ. Re.s Risk Assess. 24*, 247–257. https://doi.org/10.1007/s00477-009-0314-2

Dickinson, J.L., Zuckerberg, B.J., and *Bonter, D.N*., 2010: Citizen science as an ecological research tool: challenges and benefits. *Ann. Rev. Ecol. Evol. Syst.* 41, 149–172. https://doi.org/10.1146/annurev-ecolsys-102209-144636

Dousset, B., 1989: AVHRR-derived cloudiness and surface temperature patterns over the Los Angeles area and their relationship to land use. Proceedings of IGARSS-89. New York, NY: IEEE, 2132–2137.

Falahatkar, S., Hosseini, S.M., and *Soffianian, A.R.,* 2011: The relationship between land cover changes and spatial-temporal dynamics of land surface temperature. *Indian J. Sci. Technol. 4*(2), 76–81. https://doi.org/10.17485/ijst/2011/v4i2/29937

Gallo, K.P., McNab, A.L., Karl, T.R., Brown, J.F., Hood, J.J., and *Tarpley, J.D.,* 1993: The use of NOAA AVHRR data for assessment of the urban heat island effect. *J. App. Meteorol. 32*, 899–908. https://doi.org/10.1175/1520-0450(1993)032<0899:TUONAD>2.0.CO;2

Gallo, K.P. and *Tarpley, J.D.,* 1996: The comparison of vegetation index and surface temperature composites for urban heat island analysis. *Int. J. Remote Sens. 17*, 3071–3076. https://doi.org/10.1080/01431169608949128

Gao, B.C., 1996: NDWI–A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sens. Environ. 58*, 257–266. https://doi.org/10.1016/S0034-4257(96)00067-3

Garcia-Cueto., Jauregui-Ostos, E., and *Tejeda-Martinez, A*., 2007: Detection of the urban heat island in Mexicali, B.C., Mexico and its relationship with land use. *Atmosfera, 20*(2), 111–131.

Haashemi, S., Weng, Q., Darvishi, A., and *Alavipanah, S.K*., 2016: Seasonal Variations of the Surface Urban Heat Island in a Semi-Arid City. *Remote Sens. 8*, 352. https://doi.org/10.3390/rs8040352

Hamdi, R., 2010: Estimating Urban Heat Island Effects on the Temperature Series of Uccle (Brussels, Belgium) Using Remote Sensing Data and a Land Surface Scheme. *Remote Sens. 2*, 2773–2784. https://doi.org/10.3390/rs2122773

Huete, A.R., 1988: A soil-adjusted vegetation index (SAVI). *Remote Sens. Environ. 25(3),* 295–309. https://doi.org/10.1016/0034-4257(88)90106-X

Ifatimehin, O., Musa, S.D. and *Adeyemi, J.O.,* 2011: Managing Land use transformation and land surface temperature change in Anygiba town, Kogi state, Nigeria. *J. Geograp. Geol. 3*, 77–85. https://doi.org/10.5539/jgg.v3n1p77

Karl, T.R. and *Quayle, R.G*., 1988: Climate change in fact and in theory: Are we collecting the facts? *Climate Change, 13*, 5–17. https://doi.org/10.1007/BF00140159

Kim, H.M., Kim, B.K., and *You, K.S.,* 2005: A statistic correlation analysis algorithm between land surface temperature and vegetation index. *Int. J. Inform. Proc. Syst. 1*, 102–106. https://doi.org/10.23953/cloud.ijaese.204

Lazzarini, M., Marpu, P.R., and *Ghedira, H.,* 2013: Temperature-land cover interactions: The inversion of urban heat island phenomenon in desert city areas. *Remote Sens. Environ. 130*, 136–152. https://doi.org/10.1016/j.rse.2012.11.007

Li, J.J., Wang, X.R., Wang, X.J., Ma, W.C., and *Zhang, H*., 2009: Remote sensing evaluation of urban heat island and its spatial pattern of the Shanghai metropolitan area, China. *Ecol. Complex. 6,* 413–420. https://doi.org/10.1016/j.ecocom.2009.02.002

Li, W., Coa, Q., Lang, K., and *Wu, J.,* 2017: Linking potential heat source and sink to urban heat island: Heterogeneous effects of landscape pattern on land surface temperature. *Sci. Total Environ. 586*, 457–465. https://doi.org/10.1016/j.scitotenv.2017.01.191

Liu, L. and *Zhang, Y.,* 2011: Urban heat island analysis using the Landsat TM data and ASTER data: a case study in Hong Kong. *Remote Sens. 3*, 1535–1552. https://doi.org/10.3390/rs3071535

Lo, C.P., Quattrochi, D.A., and *Luvall, J.C.,* 1997: Application of high-resolution thermal infrared remote sensing and GIS to assess the urban heat island effect. *Int. J. Remote Sens. 18*, 287–304. https://doi.org/10.1080/014311697219079

Lu, D. Li, G., and *Moran, E*., 2014: Current situation and needs of change detection techniques. *Int. J. Image Data Fusion 5*, 13–38. http://dx.doi.org/10.1080/19479832.2013.868372

Lu, D. and *Weng, Q.*, 2006: Spectral mixture analysis of ASTER imagery for examining the relationship between thermal features and biophysical descriptors in Indianapolis, Indiana. *Remote Sens. Environ. 104*, 157–167.https://doi.org/10.1016/j.rse.2005.11.015

Malekpour, P., Taleai, M., and *Rezaei, Y*., 2010: Remote sensing-based spatial-temporal analysis of urban land surface temperature related to urban development: a case study of Tehran, MRSS2010, 6th International Remote Sensing and GIS Conference and Exhibition, April 28–29, 2010, Putra World Trade Centre, Kuala Lumpur, Malaysia.

Owen, T.W., Carlson, T.N., and *Gillies, R.R.,* 1998: An assessment of satellite remotely-sensed land cover parameters in quantitatively describing the climatic effect of urbanization. *Int. J. Remote Sens. 19*, 1663–1681. https://doi.org/10.1080/014311698215171

Pal, S. and *Ziaul, S.K*., 2017: Detection of land use and land cover change and land surface temperature in English Bazar urban center. *Egyptian J. Remote Sens. Space Scie 20*, 125–145. https://doi.org/10.1016/j.ejrs.2016.11.003

Park, J., Kim, J.H. Lee, D.K., Park, C.Y., and *Jeong, C.G.,* 2017: The influence of small green space type and structure at the street level on urban heat island mitigation. *Urban Forest. Urban Green. 21,* 203–212. https://doi.org/10.1016/j.ufug.2016.12.005

Peterson, T.C., 2003: Assesment of urban versud rural in situ surface temperature in the contiguous Unites States: No difference found, *J. Climate 18,* 2941–2959. https://doi.org/10.1175/1520-0442(2003)016<2941:AOUVRI>2.0.CO;2

Qin, Z., Karnieli, A., and *Berliner, P*., 2001: A mono-window algorithm for retrieving land surface temperature from Landsat TM data and its application to the Israel-Egypt border region. *Int. J. Remote Sens. 22*, 3719–3746. https://doi.org/10.1080/01431160010006971

Ranagalage, M., Estoque, R.C., Handayani, H.H., Zhang, X., Morimoto, T., Tadono, T., and *Murayama, Y.*, 2017: Relation between Urban Volume and Land Surface Temperature: A Comparative Study of Planned and Traditional Cities in Japan. *Sustainability 10 (7),* 2366. https://doi.org/10.3390/su10072366

Rasul, A., Balzter, H., Smith, C., Remedios, J., Adamu, B., Sobrino, J.A., Srivanit, M and Weng, Q., 2017: A review on remote sensing of urban heat and cool island, Land 6(38), 1–10. https://doi.org/10.3390/land6020038

Ray, T.W., 1994: Vegetation in remote sensing FAQs, Applications, ER Mapper, Ltd., Perth, unpaginated CD-ROM. https://doi.org/10.1155/2017/1353691

Rousta, I., Sarif, M. O., Gupta, R. D., Olafsson, H., Ranagalage, M., Murayama, Y., Zhang, H. and *Mushore, T.D.*, 2018: Spatiotemporal Analysis of Land Use/Land Cover and Its Effects on Surface Urban Heat Island Using Landsat Data: A Case Study of Metropolitan City Tehran (1988–2018: *Sustainability 10*, 4433. https://doi.org/10.3390/su10124433

Shahmohamadi, P., Sodoudi, S., and *Cubasch, U.*, 2013: Mitigation of Urban Heat Island Effects as an Environmental Phenomena in Tehran. *Caspian J. Appl. Sci. Res. 2(8),* 126–135.

Shahmohamadi, P., Chen-Ali, A.I., Abdullah, N.A.G., Maulud, K.N.A., Tahir, M.M., and *Mohd-Nor, M.F.I.,* 2015: The Conceptual Framework on Formation of Urban Heat Island in Tehran Metropolitan, Iran: A Focus on Urbanization Factor. Select. *Topics Power Syst. Remote Sens.* 251–259.Corpus ID: 73707678

Shamsipour, A., Mahdian Mahforouzi, M., and *Akhavan, H.*, 2012: An Integrated Study on the Urban Heat Island Formation in Tehran. *Geophys. Res. Abstr., EGU2012-PREVIEW* 14. https://doi.org/10.1016/j.buildenv.2010.04.001

Sodoudi, S., Shahmohamadi, P., Vallack, K., Cubasch, U., and *Chi-Ani, A.I.*, 2014: Mitigating the Urban Heat Island Effect in Megacity Tehran. *Adv. Meteorol.,* Article ID 547974, 1–19. https://doi.org/10.1155/2014/547974

Soltani, A. and *Sharifi, E.,* 2017: Daily variation of urban heat island effect and its correlations to urban greenery: A case study of Adelaide. *Front. Architect. Res. 6*, 529–538. https://doi.org/10.1016/j.foar.2017.08.001

Streutker, D.R., 2003: Satellite-measured growth of the urban heat island of Houston, Texas. *Remote Sens. Environ 85,* 282–289. https://doi.org/10.1016/S0034-4257(03)00007-5

Sun, Q., Tan, J. and *Xu, Y.*, 2010: An ERDAS image processing method for retrieving LST and describing urban heat evolution: A case study in the Pearl River Delta Region in South China. *Environ. Earth Sci 59*, 1047–1055. https://doi.org/10.1007/s12665-009-0096-3

Van de Griend, A.A., and *Owe, M.*, 2003: On the relationship between thermal emissivity and the normalized difference vegetation index for natural surfaces. *Int. J. Remote Sens. 14*, 1119–1131. https://doi.org/10.1080/01431169308904400

Voogt, J.A. and *Oke, T.R.*, 2003: Thermal remote sensing of urban climates. *Remote Sens. Environ. 86*, 370–38. https://doi.org/10.1016/S0034-4257(03)00079-8

Weng, Q., 2001: A remote sensing-GIS evaluation of urban expansion and its impact on surface temperature in the Zhujiang Delta, China. *Int. J. Remote Sens. 22*, 1999–2014. https://doi.org/10.1080/713860788

Weng, Q., 2003: Fractal analysis of satellite-detected urban heat island effect. *Photogram. Engin. Remote Sens 69,* 555–566. https://doi.org/10.14358/PERS.69.5.555

Weng, Q., Lu, D., and *Schubring, J.*, 2004: Estimation of land surface temperature–vegetation abundance relationship for urban heat island studies. *Remote Sens. Environ 89*, 467–483. https://doi.org/10.1016/j.rse.2003.11.005

Weng, Q., Lu, D., and *Liang, B.,* 2006: Urban Surface Biophysical Descriptors and Land Surface Temperature Variations. *Photogram. Engin. Remote Sens 72*, 1275–1286. https://doi.org/10.14358/PERS.72.11.1275

Weng, Q., Liu, H., and *Lu, D.*, 2007: Assessing the effects of land use and land cover patterns on thermal conditions using landscape metrics in city of Indianapolis, United States. *Urban Ecosyst. 10*, 203–219. https://doi.org/10.1007/s11252-007-0020-0

Weng, Q. and *Lu, D.,* 2008: A sub-pixel analysis of urbanization effect on land surface temperature and its interplay with impervious surface and vegetation coverage in Indianapolis, United States. *Int. J. Appl. Earth Observ. Geoinform.* 10, 68–83. https://doi.org/10.1016/j.jag.2007.05.002

Xiao, R., Weng, Q., Ouyang, Z., Li, W., Schienke, E.W., and *Zhang, W.*, 2008: Land surface temperature variation and major factors in Beijing, China. *Photogram. Engin. Remote Sens 74*, 451–461. https://doi.org/10.14358/PERS.74.4.451

Xu, H., 2005: A study on information extraction of water body with the Modified Normalized Difference Water Index (MNDWI). *J. Remote Sens. 9*, 511–517. Corpus ID: 133201908

Xu, H., 2007: Extraction of urban built-up land features from Landsat imagery using thematic-oriented index combination technique. *Photogram. Engin. Remote Sens 73*, 1381–1391. https://doi.org/10.14358/PERS.73.12.1381

Yuan, F. and *Bauer, M.E.,* 2007: Comparison of impervious surface area and normalized difference vegetation index as indicators of surface urban heat island effects in Landsat imagery. *Remote Sens. Environ. 106*, 375–386. https://doi.org/10.1016/j.rse.2006.09.003

Zha, Y., Gao, J., and *Ni, S.*, 2003: Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *Int. J. Remote Sens. 24*, 583–594. https://doi.org/10.1080/01431160304987

Zhao, H.M. and *Chen, X.L.*, 2005: Use of normalized difference bareness index in quickly mapping bare areas from TM/ETM+. *Geosci. Remote Sens Symp 3*(25–29), 1666–1668. https://doi.org/10.1109/IGARSS.2005.1526319

Zhang, J., Wang, Y., and *Li, Y*., 2006: A C++ program for retrieving land surface temperature from the data of Landsat TM/ETM+ band6. *Comput. Geosci. 32*, 1796–1805. https://doi.org/10.1016/j.cageo.2006.05.001

Zhang, J., Wang, Y., and *Wang, Z*., 2007: Change analysis of land surface temperature based on robust statistics in the estuarine area or Pearl River (China) from 1990 to 2000 by Landsat TM/ETM+ data. *Int. J. Remote Sens. 28*, 2383–2390. https://doi.org/10.1080/01431160701236811

Zhou, D., Zhao, S., Liu, S., Zhang, L., and *Zhu, C.,* 2014: Surface urban heat island in China's 32 major cities: Spatial patterns and drivers, *Remote Sens. Environ. 152*, 51–61. https://doi.org/10.1016/j.rse.2014.05.017

Zhou, D., Zhao, S., Zhang, L., Sun, G., and *Liu Y.*, 2015: The footprint of urban heat island effect in China. *Sci. Rep.5*, 11160. https://doi.org/10.1038/ srep11160

DOI:10.28974/idojaras.2022.4.2

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 126, No. 4, October – December, 2022, pp. 457–480

Investigation of a supercell merger leading to the EF4 tornado in the Czech Republic on June 24, 2021 using radar data and numerical model outputs

Kornél Komjáti[1,2,5,*], **Ákos János Varga**[2], **Ladislav Méri**[3,4], **Hajnalka Breuer**[2], and **Sándor Kun**[5]

[1]*Hungarian Meteorological Service*
Kitaibel Pál u. 1, Budapest H-1024, Hungary

[2]*ELTE Eötvös Loránd University*
Institute of Geography and Earth Sciences
Department of Meteorology
Pázmány Péter sétány 1/A, Budapest H-1117, Hungary

[3]*Slovak Hydrometeorological Institute*
Jeséniova 17, 833 15 Bratislava, Slovakia

[4]*Department of Astronomy, Physics of the Earth and Meteorology*
Comenius University in Bratislava, 842 48 Bratislava, Slovakia

[5]*Hungarian Association of Stormchasers and Storm Damage Surveyors*
Fiastyúk u. 57. 3/3., Budapest H-1139, Hungary

**Corresponding author E-mail: komjati.k@met.hu*

(Manuscript received in final form January 24, 2022)

Abstract— An unprecedented deadly and destructive EF4 tornado struck the Czech Republic across Břeclav and Hodonín districts on June 24, 2021. On this day, several supercells developed in Central Europe, however, in Austria and the Czech Republic region only one cell produced a tornado. For this reason, in addition to the macrosynoptic setup, it is also worth exploring the small-scale cell interactions that can lead to the formation of a devastating EF4 tornado. We use ECMWF analysis and forecast fields, sounding profiles, and radar measurements to examine the synoptic weather situation and convective processes. Moreover, to investigate the evolution and structure of convection, two Weather Research and Forecasting (WRF) model simulations were carried out at 1.5 km grid spacing with one-moment and two-moment microphysical parameterizations. WRF captures the overall spatial distribution and supercellular nature of thunderstorms, although discrepancies exist in the magnitude and spatial location of individual cells. The low-reflectivity region accompanying the thunderstorms is better represented by the one-moment microphysics scheme.

Key-words: supercell, tornado, cell merger, convection, thunderstorm, WRF model, numerical simulation

1. Introduction

On June 24, 2021, an unusually strong tornado formed in southeast Czechia, resulting in at least 6 deaths, and injuring more than 200 people. Based on the available information and the caused damage, the European Severe Storm Laboratory (ESSL) rated the tornado as category 4 on the Enhanced Fujita Scale (EF4). The specialty of the case was that before the tornadogenesis, the tornado-producing mother supercell merged with another supercell, similarly to the historic May 22, 2011 Joplin (USA, Missouri) tornado (*Van Leer*, 2013*, Knupp et al.*, 2014).

The importance of the storm merger in tornadogenesis has been discussed in several papers (*Bluestein* and *Weisman*, 2000; *Lee et al.*, 2006, *Wurman et al.*, 2007, *Van Leer*, 2013). The definition of cell merging is generally based on radar observations and describes the union of two, initially independent radar echoes (*Westcott* and *Kennedy* 1989; *Lee et al.*, 2006), or the merging of the updraft region (*Wescott*, 1994; *Bluestein* and *Weisman*, 2000; *Hastings* and *Richardson*, 2016). The success of a merger strongly depends on the angle at which the cells interact with each other, namely if the merger occurs in the inflow region of the mother cell, the downdraft might cut off the main updraft of the mother cell, destroying the storm (*Jaret et al.*, 2008). In addition, the strength of the outflow and the distance between the cells are also important (*Hastings* and *Richardson*, 2016). A typical sign of an effective merger is the reflectivity cloud bridge between the cells (*Simpson et al.*, 1980), created by the downdraft outflow boundaries.

In recent years, much attention has been paid to studying multiple gust front convergence zones and their role in tornadogenesis (*Marquis et al.*, 2008; *Beck* and *Weiss*, 2013; *Orf et al.*, 2017; *Betten et al.*, 2018; *Schueth et al.*, 2021). There are many questions about the dynamic processes that might be relevant in the production of the secondary rear flank gust front and its adverse or advantageous effect on near-surface stretching. Although, in several cases, it is seen that the secondary boundary on the surface inside of the rear flank downdraft (RFD) region might contribute to the low-level mesocyclone intensification through the multiple convergent zones and the horizontal wind-shift generated vorticity. However, a secondary rear flank gust front may also appear not only inside a supercell, but with a connection of different downdraft regions as well during a cell-merger process (*Van Leer*, 2013), or cell interaction with a remnant outflow boundary (*Markowski et al.*, 1998).

The Weather Research and Forecasting (WRF) model is an increasingly popular tool for the numerical simulation of weather-related phenomena in both operational and academic applications (*Powers et al.*, 2017). It has been used extensively to model tornado-producing supercell thunderstorms (e.g., *Miglietta et al.*, 2017; *Scheffknecht et al.*, 2017; *Pigluj et al.*, 2019; *Spiridonov et al.*, 2021). Numerical studies of supercells require convection-allowing (< 4 km) grid sizes,

where the role of the microphysical parameterization becomes crucial (*Johnson et al.*, 2016). It has been argued that two-moment schemes that also predict the number concentration of hydrometeor species can improve on the results of one-moment parameterizations when modeling convection-related processes (e.g., *Dawson et al.*, 2010; *Jung et al.*, 2012).

In the current study, we aim to examine the effect of environmental conditions, particularly the potential impact of the cell-merger on the tornadogenesis with ECMWF (European Centre for Medium-Range Weather Forecasts) IFS (Integrated Forecasting System) model products, atmospheric soundings requested from the Hungarian Meteorological Service, and real-time radial base velocity measurements and CAPPI (Constant Altitude Plan Position Indicator) planes from the Slovak Hydrometeorological Institute. Additionally, we utilize two WRF simulations with one- and two-moment microphysics schemes to study the evolution and structure of convection on the day of tornadic supercell occurrence at the Slovakian-Czech border. The aim is to investigate the capability of WRF to capture the spatiotemporal pattern and supercellular nature of thunderstorms.

2. Synoptic and mesoscale overview and storm formation

2.1. Forecasted synoptic and mesoscale conditions

In the afternoon of June 24, 2021, a strong, extended frontal boundary was located in Central Europe, which separated the Atlantic air mass from the unstable, moist air of southern and eastern Europe (*Fig. 1, top*). As the frontal zone crossed the Alps, a warm frontal wave formed on it, which caused a surface low on the lee side of the Alps. Above the warm frontal stage of the boundary in the upper levels, a short-wave trough spread northeast (*Fig. 1, bottom*) with a mid-level jet, which extended from the Mediterranean Sea to Poland (*Fig. 2, top*). In the lower levels, alongside the boundary, a strengthening low-level jet was forecasted for 18:00 UTC, which started to spread up from the Mediterranean Sea through the Czech-Slovakian border to the Baltic states (*Fig. 2, bottom*).

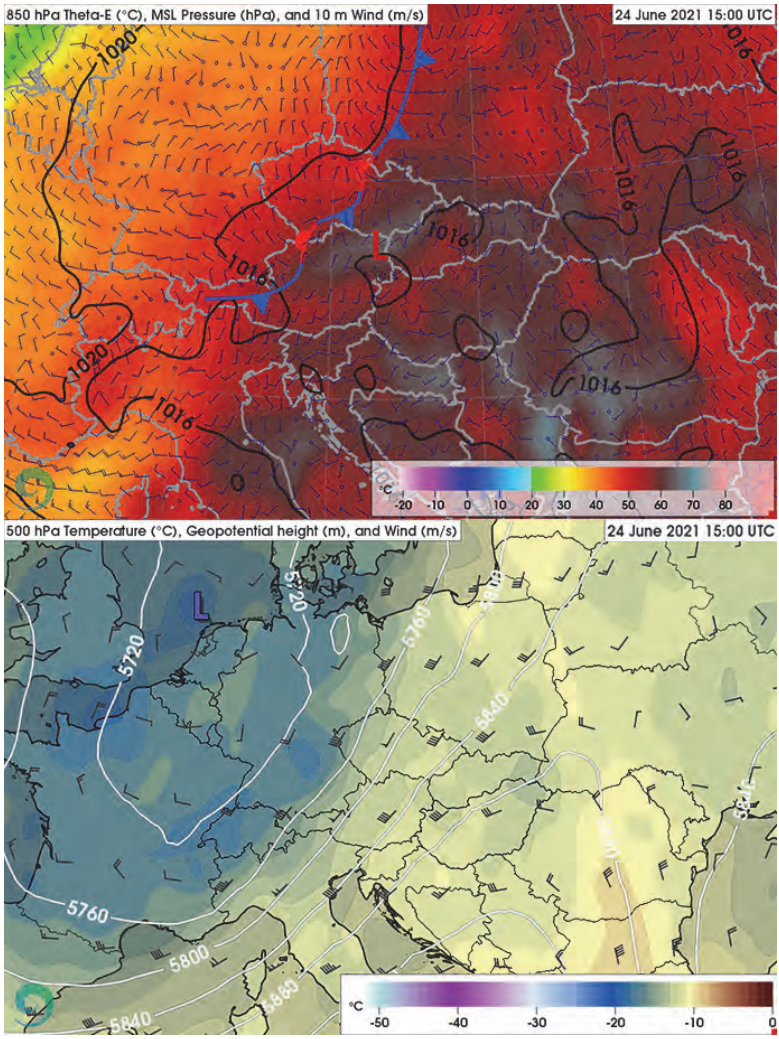


Fig. 1. ECMWF June 24, 2021 15:00 UTC forecast of 850 hPa equivalent potential temperature (shaded), surface pressure (solid black lines), fronts, and 10 m wind (blue barbs) (top). ECMWF June 24, 2021 15:00 UTC forecast of 500 hPa temperature (shaded), geopotential height (solid white lines), and wind (black barbs) (bottom).

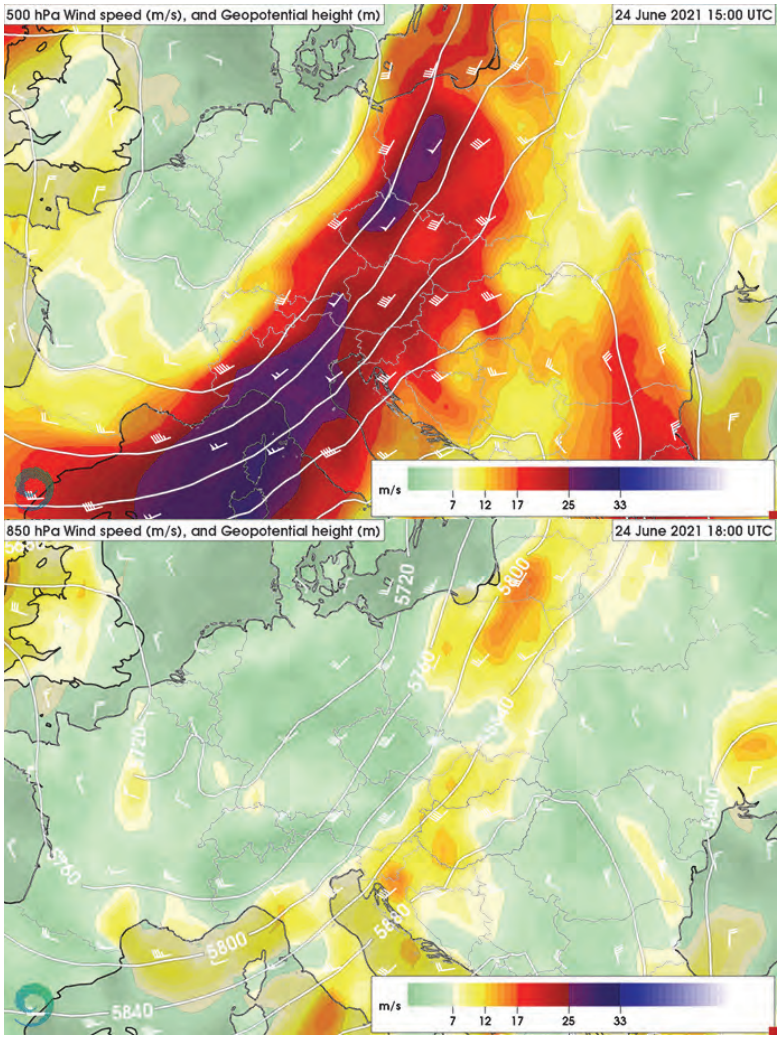


Fig. 2. ECMWF June 24, 2021 15:00 UTC forecast of 500 hPa wind speed (shaded), geopotential height (solid white lines), and wind (white barbs) (top). ECMWF June 24, 2021 18:00 UTC forecast of 850 hPa wind speed (shaded), geopotential height (solid white lines), and wind (white barbs) (bottom).

The convective initiation started in a very unstable and moist environment as predicted by the ECMWF IFS model (with 60–62 °C equivalent potential temperature and 2500–3000 J kg^{-1} CAPE (convective available potential energy) maxima) in Central Austria. In the warm sector, a near-surface confluent flow

(caused by the above-mentioned developing surface low and the orography) triggered the deep convective activity. Besides the convergent zones, the cyclonic flow resulted in northeastern wind components at the backside of the pressure minima at 15:00 UTC, which induced a strong storm-relative inflow and a notable curvature in the wind profile in the lowest 1000 m for the developing thunderstorms (*Fig. 3*). Thus, the developing surface low and the strengthening mid-level flow, together with the increasingly curved hodograph and the high environmental bulk shear (25–30 ms^{-1} for the 0–6 km layer), supplemented by the unstable, humid air mass resulted in especially favorable conditions for supercells. The storm-relative helicity for the right-moving cells (SREH-R) in the 0–3 vertical layer, and the supercell composite parameter (SCP) also showed that the conditions were ideal for intense supercells (*Fig. 4*). These favorable parameters particularly aligned with each other at around 15:00 UTC in the forecast over the central and western parts of the Lower Austria region. However, increased values of tornadic parameters (significant tornado parameter (STP), 0–1 km SREH-R, and the 0–1 km bulk shear) were predicted only at around 18:00 UTC (*Fig. 5*), when the low-level jet started to strengthen.

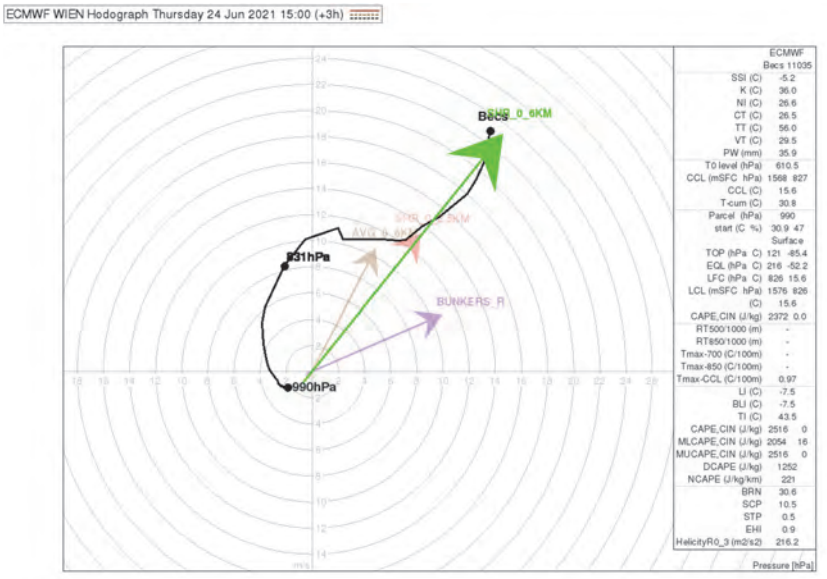


Fig. 3. ECMWF June 24, 2021 15:00 UTC forecast of hodograph over Wien-Hohe Warte. Wind shear profile between 0 and 500 hPa (solid black line), 0–6 km bulk shear vector (green arrow), 0–2.5 km bulk shear vector (red arrow), 0–6 km bulk mean wind vector (brown arrow), and the Bunkers storm motion vector for right-moving supercells (purple

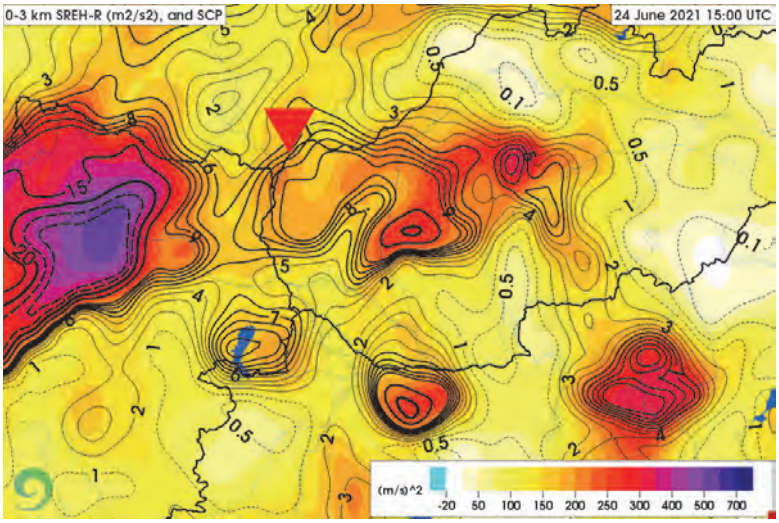


Fig. 4. ECMWF June 24, 2021 15:00 UTC forecast of 0–3 km storm relative helicity for right-moving supercells (SREH-R; shaded), and the supercell composite parameter (SCP; solid and dashed black lines). The red triangle depicts the observed position of the tornado.

2.2. Storm formation and evolution

The 12:00 UTC sounding over Wien – Hohe Warte revealed that the forecasted unstable environment mentioned above was indeed accomplished: 2228 J kg^{-1} CAPE, –3.2 °C SSI, 54.7 TT (*Fig. 6, top*). The soundings showed a classical Great Plains Type setup (*Gordon* and *Albert*, 2000) with a mid-level dry air bulge, some capping at 850 hPa, and relatively high, ≈19°C dew point temperature with a steep profile in the lowest 100 hPa. The Prostejov (Czech Republic) soundings showed a more unstable environment but slightly drier mid-level conditions (*Fig. 6, bottom*).

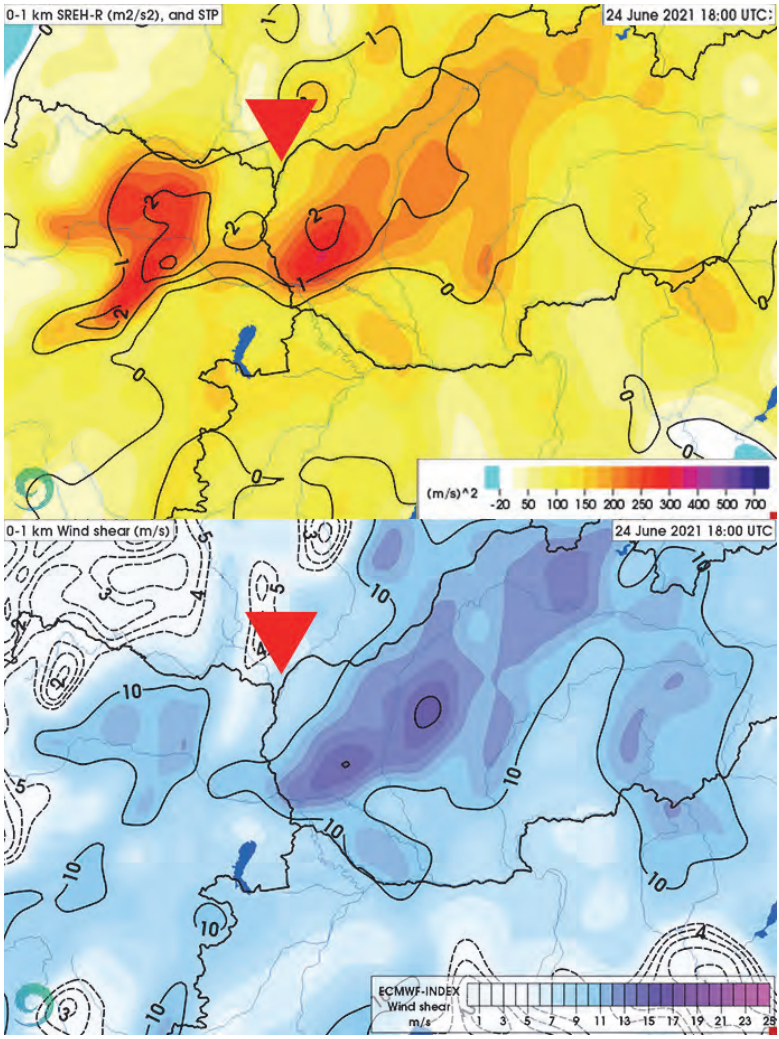


Fig. 5. ECMWF June 24, 2021 18:00 UTC forecast of 0–1 km storm-relative helicity for right-moving supercells (SREH-R; shaded), and the significant tornado parameter (STP; solid and dashed black line) (top). ECMWF June 24, 2021 18:00 UTC forecast of 0–1 km bulk shear (shaded, and solid black lines), and the observed position of the tornado (red triangle) (bottom).

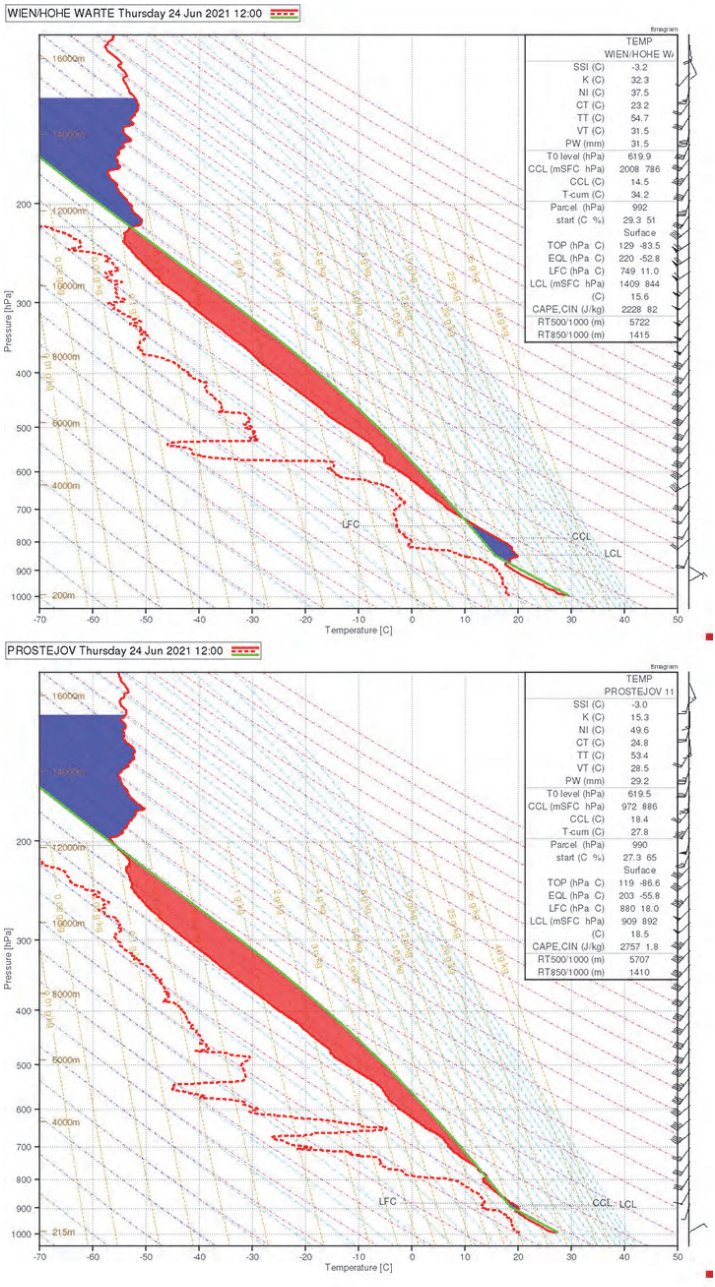


Fig. 6. Upper air data at the initiation time (June 24, 2021 12:00 UTC) in Wien – Austria (top), and Prostejov – Czech Republic (bottom). The stable (unstable) area of the sounding is shaded by blue (red).

The first thunderstorm of the day initiated at 12:00 UTC over Austria, triggered by the orographic lifting effect, and started to move to the northeast. At 13:05 UTC, a supercell (C1) started to form at the boundary of the left member of the splitting supercell at the border of Styria and Lower Austria regions (not shown). After the first thunderstorms, at around 15:00 UTC gradually more and more cells initiated over the central and the western part of Lower Austria.

Over these areas, as shown in *Fig. 4.,* the forecasted SCP and SREH-R values guaranteed exceptionally suitable conditions for intensive supercells, and as a result, a cell (C2) appeared at 14:30 UTC over Krems an der Donau. The C2 thunderstorm became a strong supercell which was indicated by the well-defined hook echo as well at 15:30 UTC (*Fig. 7*).

During the development of the C2 supercell, the C1 supercell started to split under favorable conditions and the left-mover member (C1/L) showed up on radar at 15:20 UTC. The deviantly moving C1/L cell gradually approached the C2 cell toward its RFD (rear flank downdraft) region. The merger of C1/L and C2 occurred at a nearly perfect angle, thus the downdraft region of C1/L penetrated the RFD of C2. This process might have created an external secondary gust front that provided a new source of surface convergence for the main updraft (*Fig. 8*). This transport may have contributed to the intensification of the low-level mesocyclogenesis, resulting in an even more definite right turn in C2's movement.

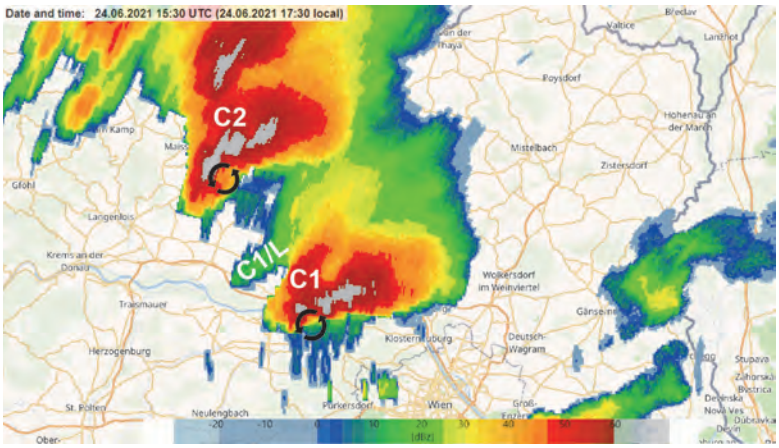


Fig. 7. 2 km CAPPI radar reflectivity (dBz) plane valid for June 24, 2021 15:30 UTC. The black rotating arrows represent the low-level mesocyclones of C1 and C2 supercells. C1/L is the left-mover member of the splitting C1 supercell.

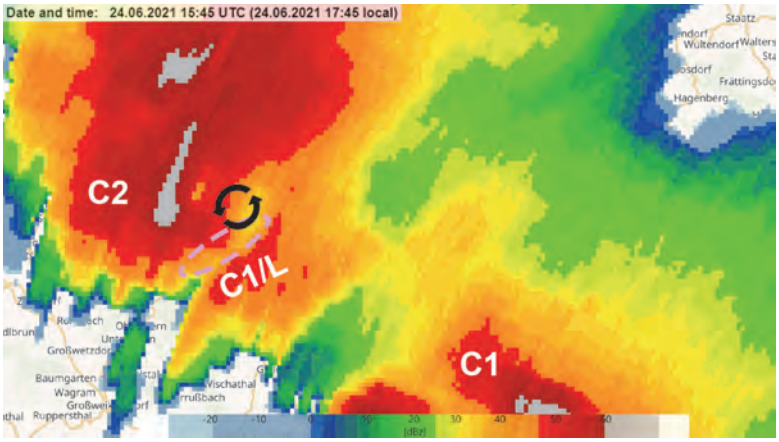


Fig. 8. 2 km CAPPI radar reflectivity (dBz) plane valid for June 24, 2021 15:45 UTC. The black rotating arrows represent the low-level mesocyclone of the C2 supercell. C1/L is the left-mover member of the splitting C1 supercell. The purple dashed ellipse depicts the effective merging area.

The rapid evolution of C2's mesocyclone indicated an interaction with the surrounding C1 supercell. The faster moving C2 started to approach C1 progressively, and at 16:00 UTC a reflectivity bridge cell appeared between the merging cells generated by the downdraft regions (*Fig. 9, top*). At 16:10 UTC, C2's more intensive and faster RFD gust front spread out and started to connect with the C1's RFD (*Fig. 9, bottom*). In a similar way to the interaction between C1/L and C2, the RFD regions combined and presumably resulted in an external secondary gust front in C1's RFD near-surface flow field. The merger process was completed at around 16:20 UTC. Based on the radar images, the cell interaction was especially beneficial for the supercell, and the regenerating low-level mesocyclone became very intense in a short time. Approximately 20 minutes after the merging, at 16:50 UTC[1], the C1 supercell reached the border of the Czech Republic with a noticeable hook echo (*Fig. 10, top*) and possibly a TVS (Tornado Vortex Signature) inside the mesocyclone. The neighboring pixels showed -30 ms^{-1} inbound and +30 ms^{-1} outbound values on the base velocity field (*Fig. 10, bottom*).

[1] At this time in addition to the effect of the cell-merger, the strengthening low-level jet probably also aided the intensification of the supercell.

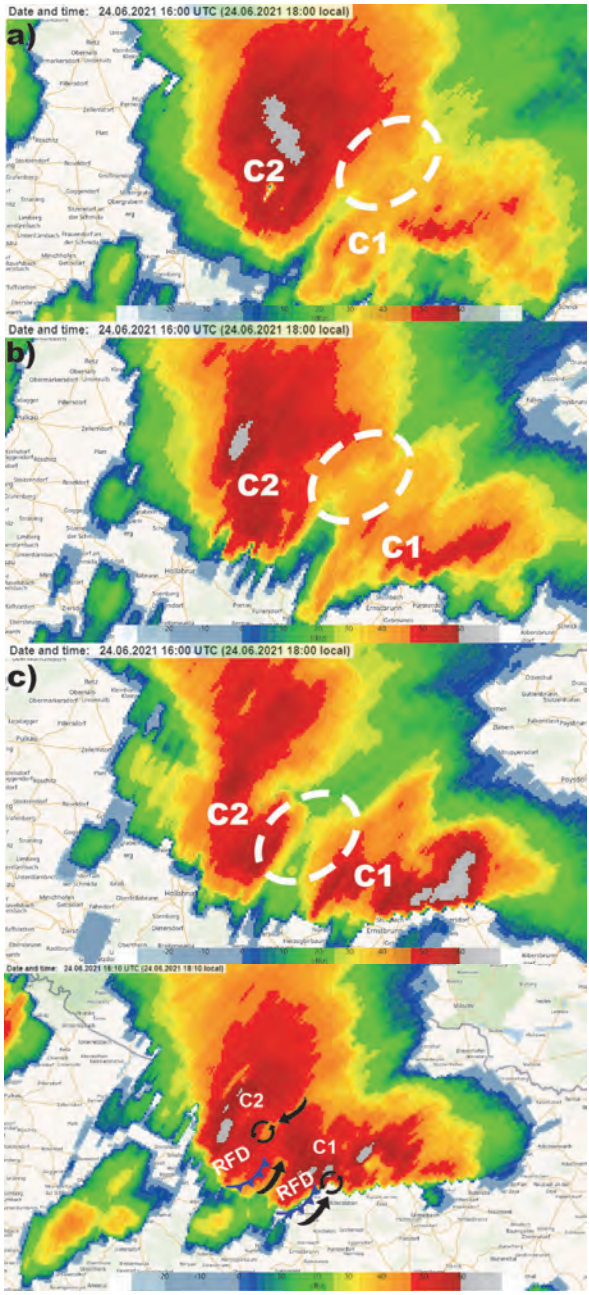


Fig. 9. 2 km (a), 4 km (b), and 6 km (c) CAPPI radar reflectivity (dBz) plane valid for June 24, 2021 16:00 UTC. The dashed white ellipses depict the reflectivity bridge between the merging storms (top). 2 km CAPPI radar reflectivity (dBz) plane is valid for 24 June 2021 16:00 UTC (bottom). The black rotating arrows represent the low-level mesocyclone of supercells C1 and C2. The black curved arrows show the inflow notches, and the blue fronts represent the rear inflow downdraft (RFD) gust fronts.

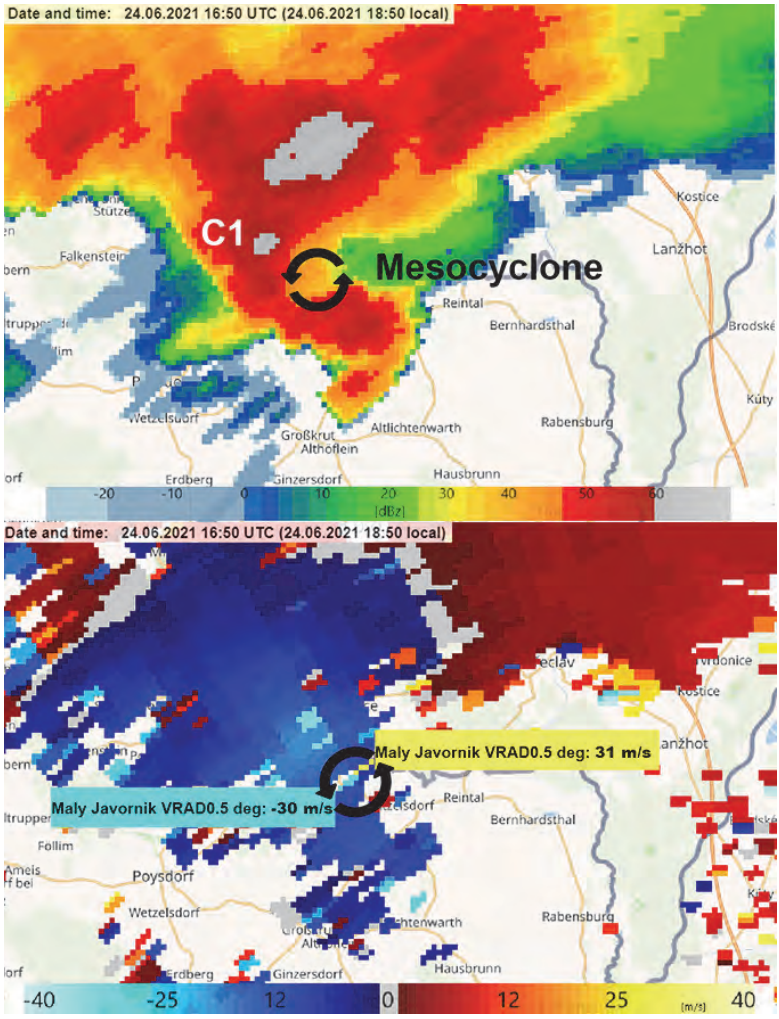


Fig. 10. 2 km CAPPI radar reflectivity (dBz) plane valid for 24 June 2021 16:50 UTC (top). Possible tornado vortex signature can be seen on the 0.5 degrees radial base velocity measurement (ms^{-1}) valid for June 24, 2021 16:50 UTC (bottom). The black rotating arrows represent the low-level mesocyclone, the bluish shades represent the inbound motions, and the reddish and yellow shades depict the outbound movement.

At 17:05 UTC in the Czech Republic over Břeclav, a tornado-like vortex appeared on the 2 km CAPPI with a donut-shaped signature (*Fig. 11*), which refers to a low-reflectivity eye with an intensive updraft region (*Wood et al.*, 2009). This donut hole signature was continuously present when the first

touchdown was observed in Hrušky at 17:20 UTC (*Fig. 12*). After the first observation, the tornado continued its path along the Slovakian and Czech border causing serious damages in Moravská Nová Ves, Lužice, and Hodonín towns, causing at least 6 deaths, and injuring more than 200 people. According to the reports, the tornado left Hodonín and dissipated at around 17:45 UTC.

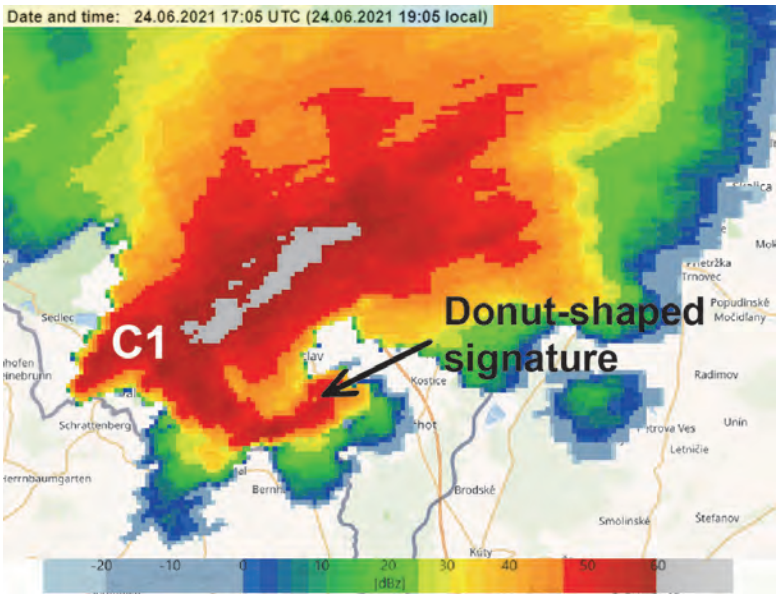


Fig. 11. Potential tornado-like vortex on the 2 km CAPPI radar reflectivity (dBz) plane valid for June 24, 2021 17:05 UTC.

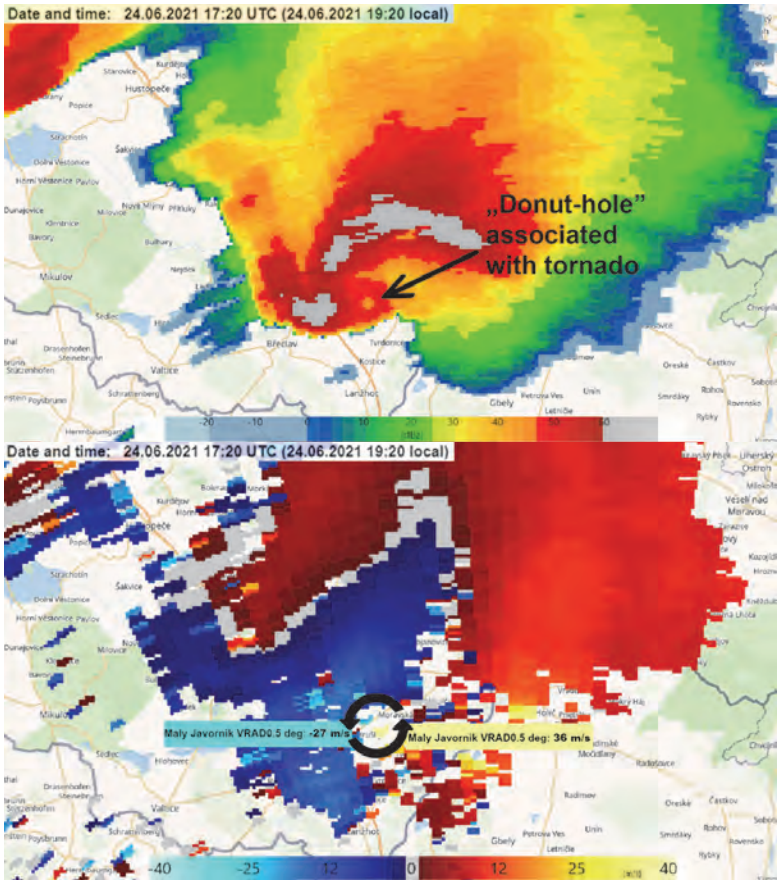


Fig. 12. Donut-shaped radar signature associated with a tornado on the 2 km CAPPI radar reflectivity (dBz) plane valid for June 24, 2021 17:20 UTC (top). Possible tornado vortex signature can be seen on the 0.5 degrees radial base velocity measurement (ms^{-1}) valid for June 24, 2021 17:20 UTC (bottom). The black rotating arrows represent the low-level mesocyclone, the bluish shades represent the inbound motions, and the reddish and yellow shades depict the outbound movement.

3. WRF simulations

3.1. Model settings

The non-hydrostatic mesoscale Advanced Research WRF (ARW) version 4.2 (*Skamarock et al.*, 2019) was applied to investigate the spatiotemporal evolution of convective processes and cell structure. The WRF model was set up on a Lambert conformal projection comprising 720 and 666 grid points in the west-east and south-north directions, respectively, with a horizontal grid spacing of

1.5 km and 61 hybrid σ–p levels in the vertical. The domain focuses on the Central European region. The initial and boundary conditions were derived from 6-hourly analysis fields of the operative IFS model (Cycle 47r2) by ECMWF. The integration period begins at 00:00 UTC on June 24, 2021 and covers 24 hours.

Two numerical experiments were carried out that only differ in the complexity of the microphysical scheme used. One WRF run utilizes the parameterization of *Thompson et al.* (2008), which is two-moment for rain and ice particles, but single-moment for cloud water, snow, and graupel. The other simulation makes use of the *Morrison et al.* (2009) scheme, which is additionally two-moment for cloud water, snow, and graupel, thus representing a more advanced class of microphysics parameterizations. Other physical processes are represented identically in the two simulations: the radiative transfer by the RRTMG scheme (Rapid Radiative Transfer Model for General Circulation Models; *Iacono et al.*, 2008), the land-surface interactions by the Noah-MP land-surface model (*Niu et al.*, 2011), the planetary boundary layer and surface layer exchange processes by the Yonsei University nonlocal closure (*Hong et al.*, 2006) together with the MM5 model's Monin-Obukhov scheme (*Jiménez et al.*, 2012). The deep convection parameterization is turned off in both experiments.

3.2. Simulation results

The WRF-simulated convective cell at the Slovakian-Czech border at the time of the tornado occurrence (at around 17:20 UTC) is considerably weaker than its observed counterpart, regardless of the microphysical parameterization used (*Fig. 13*). In addition, the cell is misplaced to the east, especially in the Morrison run. Although the WRF configuration utilized in this study did not capture the magnitude of the analyzed tornadic supercell in terms of the simulated reflectivity and missed the preceding storm merger, the overall mesoscale spatial pattern is in good agreement with radar observations (*Fig. 14*).

Comparing the two microphysical parameterizations, the Thompson scheme (*Fig. 13, top*) produces smaller and more isolated high-reflectivity regions and larger stratiform precipitation areas than the Morrison scheme (*Fig. 13, bottom*). An extensive region of relatively low (20–30 dBZ) reflectivity can be observed on radar imagery as well (*Fig. 14*), suggesting the suitability of the Thompson scheme to better capture the widespread, moderate precipitation accompanying the convective cells.

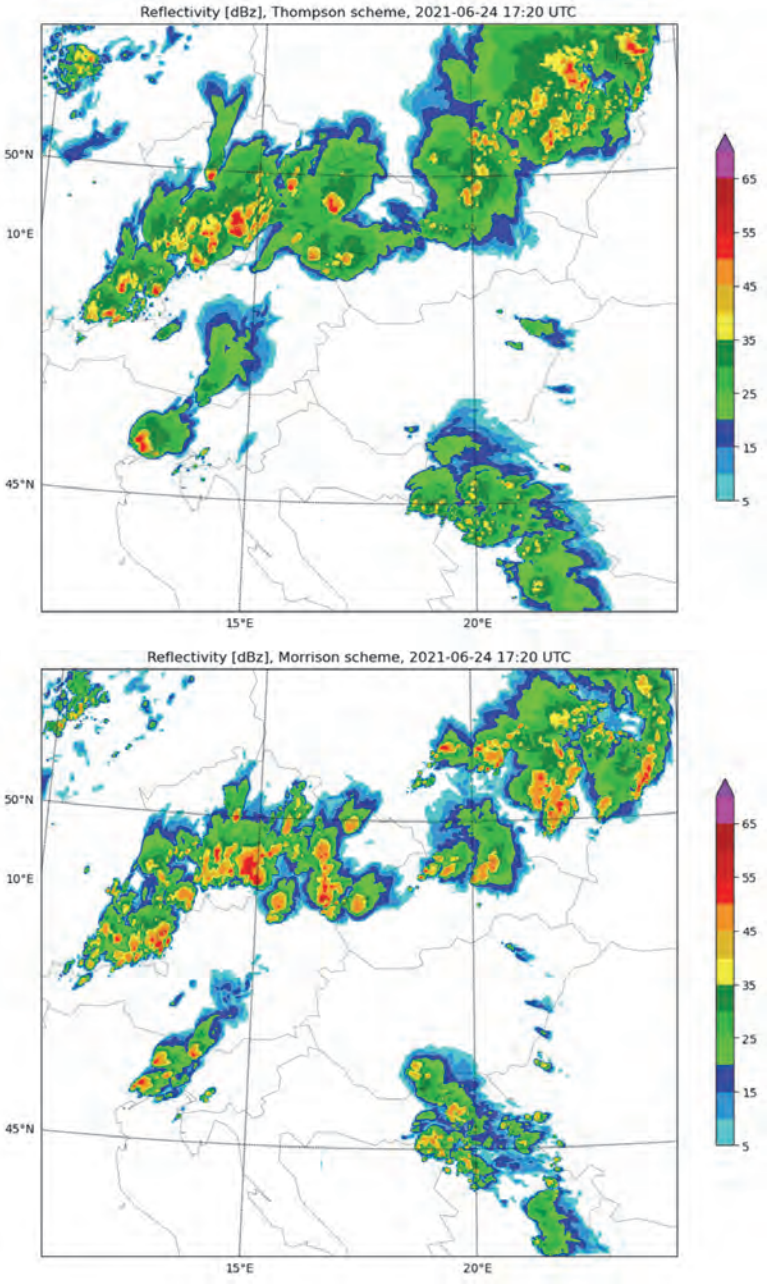


Fig. 13. WRF-simulated composite radar reflectivity valid for June 24, 2021 17:20 UTC, using the Thompson (top) and Morrison (bottom) microphysics parameterization shemes.

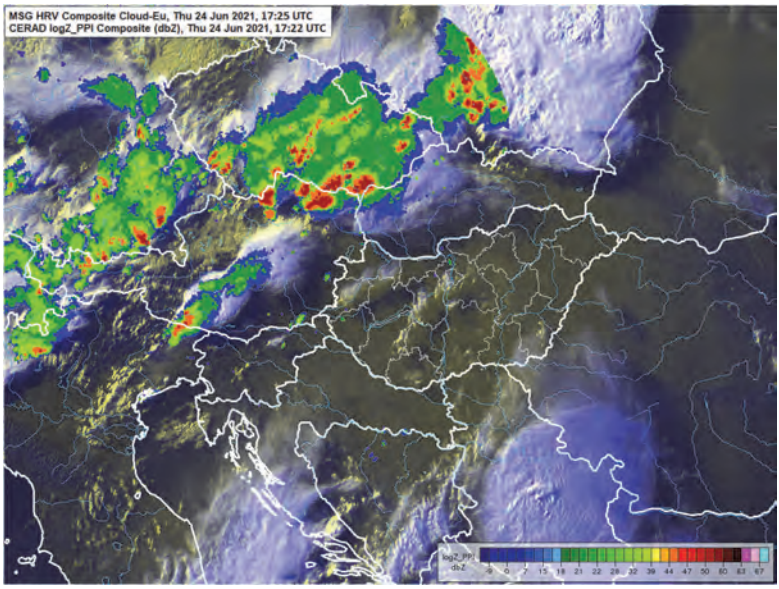


Fig. 14. Supercells over the Central European region: column maximum reflectivity (dBz) of the Central European Radar Network (CERAD), valid for June 24, 2021 17:22 UTC, and the Meteosat Second Generation (MSG) satellite High Resolution Visible (HRV) channel image, valid for 24 June, 2021 17:25 UTC.

In summary, despite requiring more than twice as much computational time, the full two-moment Morrison scheme does not remarkably improve the spatial pattern of simulated radar reflectivity compared to the Thompson parameterization in this specific case. Moreover, the Thompson scheme is used for operational weather prediction purposes even in leading American institutions such as the National Oceanic and Atmospheric Administration (*Benjamin et al.*, 2016). Therefore, results from the Thompson scheme will be presented in the upcoming discussion about storm structure.

Evidence of supercellular convection will be inferred from an arbitrarily selected storm present on the model-derived composite reflectivity field at 16:30 UTC, June 24, 2021 (*Fig. 15*).

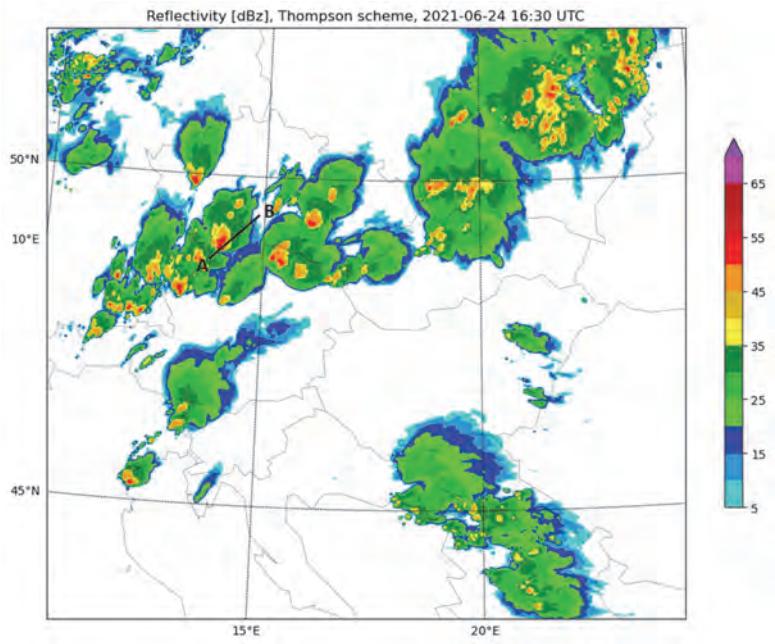


Fig. 15. WRF-simulated composite radar reflectivity valid for June 24, 2021 16:30 UTC, using the Thompson microphysics parameterization. The black line indicates the location of the vertical cross-sections presented in *Fig. 16* and *Fig. 17*.

The vertical cross-sections of reflectivity and vertical velocity (*Fig. 16*) clearly show a typical supercell structure with a bounded weak echo region (BWER) corresponding to the updraft axis. The maximum value of reflectivity and vertical velocity exceeds 55 dBZ and 35 ms^{-1}, respectively. These values, however, refer to this particular cross-section plane and might be higher for the entirety of the convective cell.

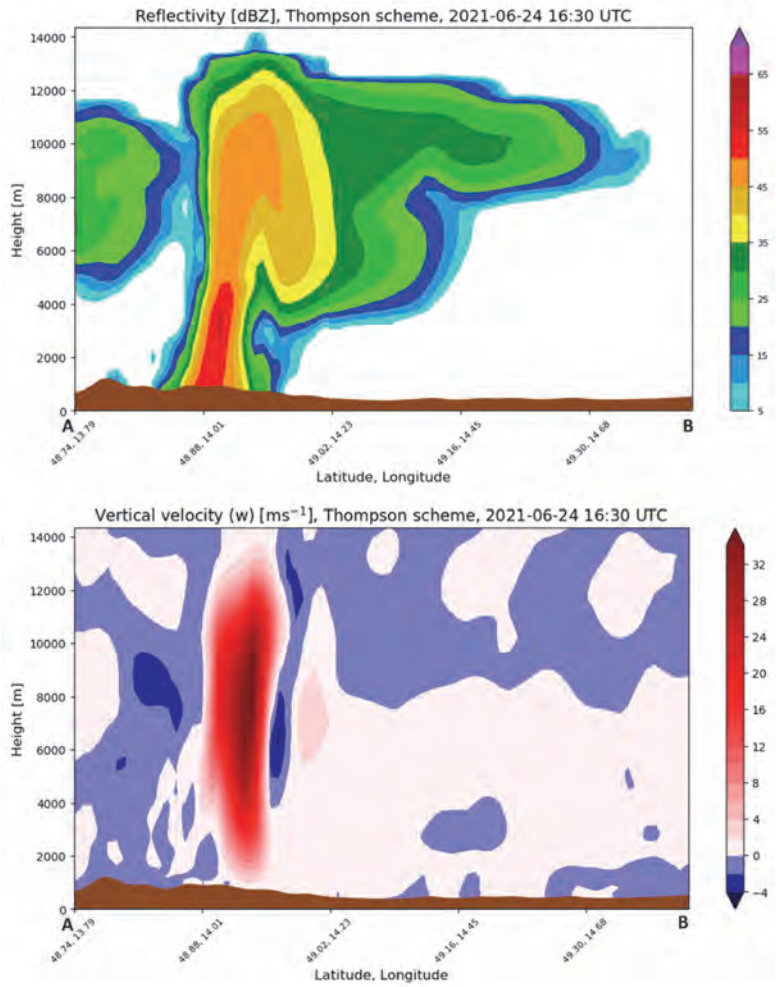


Fig. 16. Vertical cross-section of WRF-simulated radar reflectivity (top) and vertical velocity (bottom) valid for June 24, 2021 16:30 UTC, using the Thompson microphysics parameterization. The location of the vertical cross-sections is indicated by the black line in *Fig. 15*.

The absolute vorticity cross-section implies a rotating updraft with a cyclonic (counter-clockwise) vorticity maximum of ≈0.015 s^{-1} (*Fig. 17*). This is indicative of a mesocyclone, which is a characteristic feature of supercell thunderstorms. The highest values of absolute vorticity can be found at a height of ≈6 km, just below the updraft velocity maxima.

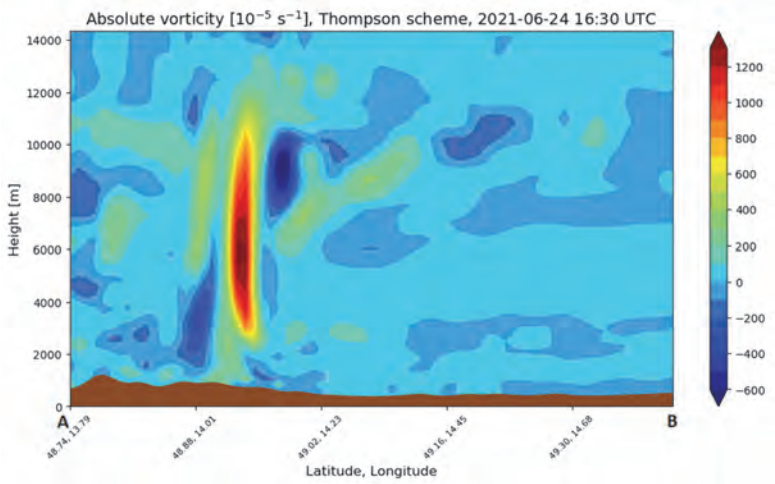


Fig. 17. Vertical cross-section of WRF-simulated absolute vorticity valid for June 24, 2021 16:30 UTC, using the Thompson microphysics parameterization. The location of the vertical cross-section is indicated by the black line in *Fig. 15*.

4. Concluding remarks

Based on the available data in this study, it can be stated, that the suitable environment forecasted by the ECMWF IFS model was approximately realized and aided the development of strong, long-lived supercells. With the strengthening of the low-level jet and deepening of the surface low, the low-level wind shear profile became more favorable for the near-surface vortices. However, only one supercell (marked C1) produced a tornado, namely a destructive EF4 one. Thus, additional effects may have contributed to this local, devastating phenomenon. The most likely contributing factor may have been the cell merger. Based on the radial wind measurements and CAPPI planes from the Radar Malý Javorník (SHMU), two, initially separated right-mover supercells (C1 and C2) merged between 16:00 and 16:20 UTC in Lower Austria resulting in a much stronger supercell structure with an impressive hook echo in a short time. The more intensive and larger C2 cell caught up with the smaller C1 supercell. The faster moving C2 RFD gust front penetrated to the C1 RFD and created an external secondary gust front that may have contributed to the vorticity transport towards the C1 mesocyclone through the emerging secondary surface convergent zone in the RFD. However, the description of the external secondary gust front on the tornadogenesis in this paper is only theoretical, there were no adequate measurements available to justify the process.

Numerical experiments were carried out with the WRF model to study the evolution and structure of convective phenomena on the day of the supercell outbreak at the Slovakian-Czech border. The overall pattern of simulated radar reflectivity is in accordance with radar observations, although the magnitude of the tornadic supercell in focus is considerably weaker in the model. The storm merger was also missed by the simulations. Nevertheless, based on vertical cross-sections of radar reflectivity, vertical velocity, and absolute vorticity from an arbitrarily selected thunderstorm, the WRF model successfully captures the supercellular convection and the corresponding mesocyclone structure. Accordingly, short-term weather forecasts and severe weather warnings might greatly benefit from such high-resolution WRF simulations. The extensive low-reflectivity (20–30 dBZ) area accompanying the convective cells is better captured by the Thompson microphysical parameterization than the Morrison scheme. Therefore, it is suggested that the complexity and thus higher computational demand of a full two-moment microphysical parameterization do not necessarily improve model performance, which is important from an operative numerical weather prediction perspective.

In the future, WRF simulations with finer grid spacing (at the order of 100 m) could be carried out to successfully capture the storm merger process and the fine-scale details of the tornado-producing supercell. An extensive analysis of the physical-dynamical settings of the model is also recommended.

Acknowledgement: The authors are grateful to the Hungarian Meteorological Service (OMSZ) and the Slovak Hydrometeorological Institute for making the data available for research. The research leading to this paper was supported by the Hungarian Scientific Research Fund under the grant FK132014. Hajnalka Breuer's work was additionally financed by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences.

References

Beck, J. and *C. Weiss*, 2013: An assessment of low-level baroclinity and vorticity within a simulated supercell. *Month. Weather Rev. 141*, 649– 669. https://doi.org/10.1175/MWR-D-11-00115.1

Benjamin, S.G., Weygandt, S.S., Brown, J.M., Hu, M., Alexander, C.R., Smirnova, T.G., Olson, J.B., James, E.P., Dowell, D.C., Grell, G.A., and *Lin, H.*, 2016: A North American hourly assimilation and model forecast cycle: The Rapid Refresh. *Month. Weather Rev. 144*, 1669–1694. https://doi.org/10.1175/MWR-D-15-0242.1

Betten, D.P., Biggerstaff, M.I., and *Ziegler, C.L.*, 2018: Three-Dimensional Storm Structure and Low-Level Boundaries at Different Stages of Cyclic Mesocyclone Evolution in a High-Precipitation Tornadic Supercell. *Adv. Meteorol. 2018,* 1–24. https://doi.org/10.1155/2018/9432670

Bluestein, H.B. and *M. L. Weisman*, 2000: The interaction of numerically simulated supercells initiated along lines. *Month. Weather Rev. 128*, 3128–3149. https://doi.org/10.1175/1520-0493(2000)128%3C3128:TIONSS%3E2.0.CO;2

Bunkers, J.M., Klimowski, B.A., Zeitler, J.W., Thompson, R.L., and *Weisman, M.L.*, 2000: Predicting Supercell Motion Using a New Hodograph Technique. *Weather and Forecast. 15*, 61–79. https://doi.org/10.1175/1520-0434(2000)015%3C0061:PSMUAN%3E2.0.CO;2

Dawson, D.T., Xue, M., Milbrandt, J.A., and *Yau, M.K.*, 2010: Comparison of evaporation and cold pool development between single-moment and multimoment bulk microphysics schemes in idealized simulations of tornadic thunderstorms. *Month. Weather Rev. 138*, 1152–1171. https://doi.org/10.1175/2009MWR2956.1

Gordon, J.D. and *Albert D.*, 2000: A Comprehensive Severe Weather Forecast Checklist and Reference Guide. US Department of Commerce, National Oceanic and Atmospheric Administration, National Weather Service, Scientific Services Division, Central Region: Missouri.

Hastings, R. and *Richardson, Y.*, 2016: Long-Term Morphological Changes in Simulated Supercells Following Mergers with Nascent Supercells in Directionally Varying Shear. *Month. Weather Rev. 144,* 471–499. https://doi.org/10.1175/MWR-D-15-0193.1

Hong, S.–Y., Noh, Y., and *Dudhia, J.*, 2006: A new vertical diffusion package with an explicit treatment of entrainment processes. *Month. Weather Rev. 134*, 2318–2341. https://doi.org/10.1175/MWR3199.1

Iacono, M.J., Delamere, J.S., Mlawer, E.J., Shephard, M.W., Clough, S.A., and *Collins, W.D.*, 2008: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *J. Geophys. Res-Atmos. 113(D13)*. https://doi.org/10.1029/2008JD009944

Jaret, W., Rogers, A., and *Weiss, C.C.*, 2008*:* The association of cell mergers with tornado occurrence. Poster Presentation 24th Conference on Severe Local Storms. Savannah, Georgia.

Jiménez, P.A., Dudhia, J., González-Rouco, J.F., Navarro, J., Montávez, J.P., and *García-Bustamante, E.*, 2012: A revised scheme for the WRF surface layer formulation. *Month. Weather Rev. 140*, 898–918. https://doi.org/10.1175/MWR-D-11-00056.1

Johnson, M., Jung, Y., Dawson, D.T., and *Xue, M.*, 2016: Comparison of simulated polarimetric signatures in idealized supercell storms using two-moment bulk microphysics schemes in WRF. *Month. Weather Rev. 144*, 971–996. https://doi.org/10.1175/MWR-D-15-0233.1

Jung, Y., Xue, M., and *Tong, M.*, 2012: Ensemble Kalman filter analyses of the 29–30 May 2004 Oklahoma tornadic thunderstorm using one-and two-moment bulk microphysics schemes, with verification against polarimetric radar data. *Month. Weather Rev. 140*, 1457–1475. https://doi.org/10.1175/MWR-D-11-00032.1

Knupp, K.R., Murphy, T.A., Coleman, T.A., Wade, R.A., Mullins, S.A., Schultz, C.J., Schultz, E.V., Carey, L., Sherrer, A., McCaul Jr., E.W., Carcione, B., Latimer, S., Kula, A., Laws, K., Marsh, P.T., and *Klockow, T.*, 2014: Meteorological Overview of the Devastating 27 April 2011 Tornado Outbreak. *Bull. Amer. Meteorol. Soc. 95*, 1041–1062. https://doi.org/10.1175/BAMS-D-11-00229.1

Lee, B.D., Jewett, F., and *Wilhelmson, R. B.*, 2006: The 19 April 1996 Illinois tornado outbreak. Part II: Cell Mergers and associated tornado incidence, *Weather Forecast. 21*, 449−446. https://doi.org/10.1175/WAF943.1

Markowski, P., Rasmussen, E.N., and *Straka, J.M.*, 1998: The Occurrence of Tornadoes in Supercells Interacting with Boundaries during VORTEX-95. *Weather Forecast 13*, 852–859. https://doi.org/10.1175/1520-0434(1998)013%3C0852:TOOTIS%3E2.0.CO;2

Marquis, J., Richardson, Y., Wurman, J., and *Markowski, P.*, 2008: "Single- and dual-Doppler analysis of a tornadic vortex and surrounding storm-scale flow in the Crowell, Texas, supercell of 30 April 2000," *Month. Weather Rev. 136*, 5017–5043. https://doi.org/10.1175/2008MWR2442.1

Miglietta, M.M., Mazon, J., and *Rotunno, R.*, 2017: Numerical simulations of a tornadic supercell over the Mediterranean. *Weather Forecast. 32*, 1209–1226. https://doi.org/10.1175/WAF-D-16-0223.1

Morrison, H., Thompson, G., and *Tatarskii, V.*, 2009: Impact of Cloud Microphysics on the Development of Trailing Stratiform Precipitation in a Simulated Squall Line: Comparison of One– and Two–Moment Schemes. *Month. Weather Rev. 137*, 991–1007. https://doi.org/10.1175/2008MWR2556.1

Niu, G.Y., Yang, Z.L., Mitchell, K.E., Chen, F., Ek, M.B., Barlage, M., Kumar, A., Manning, K., Niyogi, D., Rosero, E., and *Tewari, M.*, 2011: The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. *J. Geophys. Res-Atmos. 116(D12)*. https://doi.org/10.1029/2010JD015139

Orf., L., Wilhelmson, R., Lee, B., Finley, C., Houston, A., 2017: Evolution of a Long-Track Violent Tornado within a Simulated Supercell. *Bull. Amer. Meteorol. Soc. 98*, 45–68. https://doi.org/10.1175/BAMS-D-15-00073.1

Pilguj, N., Taszarek, M., Pajurek, Ł., and *Kryza, M.*, 2019: High-resolution simulation of an isolated tornadic supercell in Poland on 20 June 2016. *Atmosph. Res. 218*, 145–159. https://doi.org/10.1016/j.atmosres.2018.11.017

Powers, J.G., Klemp, J.B., Skamarock, W.C., Davis, C.A., Dudhia, J., Gill, D.O., Coen, J.L., Gochis, D.J., Ahmadov, R., Peckham, S.E., and *Grell, G.A.*, 2017: The weather research and forecasting model: Overview, system efforts, and future directions. *Bull. Amer. Meteorol. Soc. 98*, 1717–1737. https://doi.org/10.1175/BAMS-D-15-00308.1

Scheffknecht, P., Serafin, S., and *Grubišić, V.*, 2017: A long-lived supercell over mountainous terrain. *Quart. J. Roy. Meteorol. Soc. 143(709)*, 2973–2986. https://doi.org/10.1002/qj.3127

Schueth, A., Weiss, C., and *Dahl, J.M.L.*, 2021: Comparing Observations and Simulations of the Streamwise Vorticity Current and the Forward-Flank Convergence Boundary in a Supercell Storm. *Month. Weather Rev. 149*, 1651–1671. https://doi.org/10.1175/MWR-D-20-0251.1

Simpson, J.,Westcott, N.E., Clerman, R.J., and *Peilke, R.A.*, 1980: On cumulus mergers. *Archiv für Meteorologie, Geophysik Bioklimatol. Ser. A, 29*, 1–40. https://doi.org/10.1007/BF02247731

Skamarock, W.C., Klemp, J.B., Dudhia, J., Gill, D.O., Liu, Z., Berner, J., Wang, W., Powers, J.G., Duda, M.G., Barker, D.M., and *Huang, X-Y.*, 2019: A Description of the Advanced Research WRF Model Version 4. NCAR Tech Note NCAR/TN-556+STR, Mesoscale and Microscale Meteorology Division, Boulder CO, USA. https://doi.org/10.5065/1dfh-6p97

Spiridonov, V., Ćurić, M., Velinov, G., and *Jakimovski, B.*, 2021: Numerical simulation of a violent supercell tornado over Vienna airport initialized and initiated with a cloud model. *Atmosph. Res. 261*, 105758. https://doi.org/10.1016/j.atmosres.2021.105758

Thompson, G., Field, P.R., Rasmussen, R.M., and *Hall, W.D.*, 2008: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization. *Month. Weather Rev. 136)*, 5095–5115. https://doi.org/10.1175/2008MWR2387.1

Van Leer, K.W., 2013: Storm mergers and their role in tornado genesis during the 2011 Joplin storm. *Graduate Thesis, 1–77.*, Department of Atmospheric Sciences, University of Illinois Urbana-Champaign, Illinois. http://hdl.handle.net/2142/44134

Westcott, N. and *Kennedy, P.C.*, 1989: Cell development and merger in an Illinois thunderstorm observed by Doppl radar. *J. Atmosph. Sci. 46*, 117–131. https://doi.org/10.1175/1520-0469(1989)046%3C0117:CDAMIA%3E2.0.CO;2

Westcott, N., 1994: Merging of convective clouds: Cloud initiation, bridging, and subsequent growth. *Month. Weather Rev. 122*, 780–790. https://doi.org/10.1175/1520-0493(1994)122%3C0780:MOCCCI%3E2.0.CO;2

Wood, V.T., Brown, R.A., and *Dowell, D.C.*, 2009: Simulated WSR-88D Velocity and Reflectivity Signatures of Numerically Modeled Tornadoes. *J. Atmosph. Ocean. Technol. 26*, 876–893. https://doi.org/10.1175/2008JTECHA1181.1

Wurman, J., Y. Richardson, C. Alexander, S. Weygandt, and *P.F. Zhang*, 2007: Dual-Doppler and Single-Doppler Analysis of a Tornadic Storm Undergoing Mergers and Repeated Tornadogenesis. *Month. Weather Rev. 135*, 736–758. https://doi.org/10.1175/MWR3276.1

DOI:10.28974/idojaras.2022.4.3

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 126, No. 4, October – December, 2022, pp. 481–510

Statistical method for estimating average daily wind speed during the day

Károly Tar[1,*], **István Lázár**[1], and **István Hadnagy**[1,2]

[1]*Department of Meteorology, University of Debrecen,*
Egyetem Square 1, H-4010 Debrecen, Hungary

[2]*Department of Biology* and *Chemistry*
Ferenc Rákóczi II. Transcarpathian Hungarian Institute
Kossuth Square 6, UA-90202 Berehove, Ukraine

*Corresponding author E-mail: tarko47@gmail.com

(Manuscript received in final form January 12, 2022)

Abstract— Meteorologists keep searching and running models to provide the most accurate forecast of wind speed in addition to gaining a more detailed understanding of the wind conditions in Hungary. Wind speed and wind energy estimates, forecasts, and their verification are based on wind statistics from a longer or shorter previous period. Consequently, in addition to dynamic methods, purely statistical models also play an important role, i.e., findings that can be obtained from the statistical analysis of the existing database of measured data. The successive phases of the statistical method for producing scientific or operational information that can be extracted from measured, corrected, and stored meteorological data are generally: statistical analysis/processing, creating, verification, and application of the model, recording of the required information. The targeted information in this paper is the daily average of hourly wind speeds. The exact average of this time series can only be determined after the last measurement. To estimate this average during the day, however, the so-called sliding average model has been developed, which can be applied to any climatic element if its measured values are recorded at regular times over a certain period of time. The results presented in this paper are recommended for the preparation of the so-called "timetable", which is one of the most difficult problems for wind farm operators. This is basically the estimation of the amount of electricity produced the following day over short periods. It would be a significant help in the above if we can determine the probability of a decrease or increase in the average wind speed on the next day (and with it, the average daily wind power), or which of these two probabilities is greater. This requires an estimate of average wind speed of the next day. In addition, the results of one of our previous studies on the statistical structure of day-to-day changes in average daily wind speeds were also used. According to the results of the monthly testing of the model over a given period, the frequency of good estimates is between 80.6 % and 54.8%.

Key-words: sliding average model, wind statistics, wind farms, daily wind power, event frequency, Hungary

1. Introduction, antecedents

Wind speed and wind energy estimates, forecasts, and their verification are based on wind statistics from a longer or shorter previous period. Consequently, in addition to dynamic methods, purely statistical models also play an essential role. An overview of these can be found in the work of *Aggarwal* and *Gupta* (2013).

A statistical method for producing scientific or operational information that can be extracted from measured, corrected, and stored meteorological data is presented here. The successive phases of this are: statistical analysis/processing, modeling, model verification, model application, recording of the desired information.

The targeted information here is the average of the values of a climatic element measured at regular times. The exact average of this time series can only be determined after the last measurement. However, in some cases, it may be necessary to estimate this value with an acceptable error before the last measurement.

When wind energy is harnessed, this particular climatic element is wind speed. With the integration of wind energy into electricity grids, it is becoming increasingly important to obtain accurate wind speed/power forecasts. Accurate wind speed forecasts are necessary to schedule dispatchable generation and tariffs in the day-ahead electricity market *(Bremnes et al.,* 2002; *Kavasseri* and *Seetharaman*, 2009; *Shukur* and *Lee*, 2015). A very important element of this process is the preparation of a so-called "timetable", a difficult problem for wind power plant operators. This is basically the estimation of the amount of electricity produced the following day over short periods. It would be a significant help in the above if we can determine the probability of a decrease or increase in the average wind speed on the next day (and with it, the average daily wind power), or which of these two probabilities is greater. This requires an estimate of average wind speed of the next day.

In two previous studies (*Tar* and *Lázár*, 2018, *Tar*, 2021), we described the process of building a mathematical statistical model that is ultimately suitable for estimating the sign of the next day's average wind speed change and the magnitude of the average wind speed of the next day from today's average wind speed. The most important steps in the construction of this model are briefly summarized in the following.

The model is based on a time series of observed average daily wind speeds transformed to the height of 10 metres. Analyses were performed on the entire time series and its subsets of days for cyclone and anticyclone macrosynoptic situation groups (*Péczely*, 1961) and their transitions.

The changes of daily average wind speed from day to day were characterized by the relative value of

$$\Delta v_r = \frac{v_n - v_p}{v_p}, \quad (1)$$

where v_p is the average wind speed of the present day and v_n is that of the following day. The value of Δv_r is more or less independent of the height of the measurement, i.e., the height of the anemometer, error is made only for the day preceding the change of the measurement height. When given in percentage, it shows the change of the daily average wind speed of the next day in relation to that of the previous day.

Since Δv_r is the observed value of a random variable with a special structure, its most important statistical functions have been analyzed in more detail, primarily in relation to the situational groups and their transitions, using the average daily wind speeds of nine Hungarian meteorological stations over 10 years (1991–2000).

We looked at the relationship of the sign of the relative change and the average wind speed of the present day. According to Eq. (1), Δv_r has a very complex function relationship with the average wind speed (v_p) of the present day, given that the average wind speed (v_n) of the next day also depends on it. Therefore, it is advisable to consider the (v_p,Δv_r) relationship as stochastic. Logarithmic regression was the closest correlation. The correlation index i(v_p,Δv_r) showing the closeness of logarithmic regression varies between 0.404 and 0.592, with its highest valuemeasured at the meteorological station Győr at the transition from the day in the cyclone situation group to the same situation group (CG/CG). *Fig. 1* shows the regression calculated for Debrecen station for the entire 10-year period.

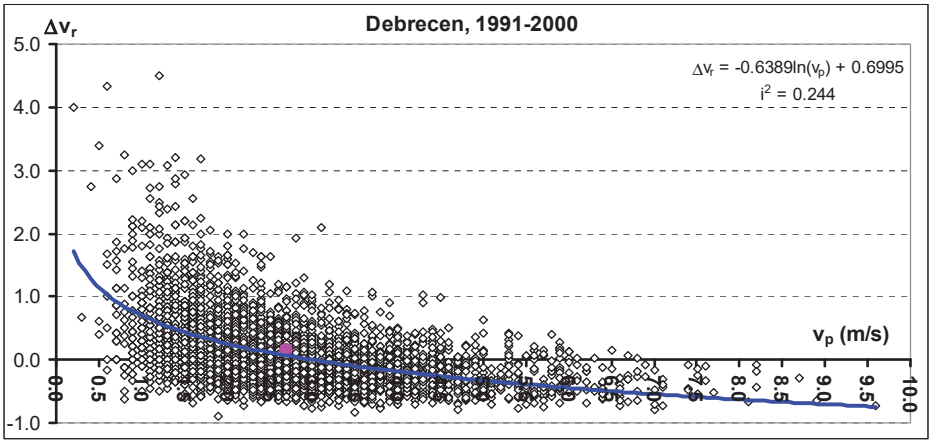


Fig. 1. Logarithmic regression between today's average wind speed (v_p) and next day's relative change (Δv_r) in Debrecen.

In all cases, the regression curve intersects the horizontal axis, i.e., the v_p axis. Let this be the zero point, v_{p0}. For the (x,y) coordinate points of the regression curve, $y>0$ before the zero point and $y<0$ after the zero point. Therefore, it can be assumed that the sign of the observed values of Δv_r may also be associated with zero points. Zero points can therefore be considered as threshold values for the examination of the sign of Δv_r. Detailed analysis confirmed this, but calculating v_{p0}, as we have seen, is not simple, consequently it is advisable to use a statistic that is easier to calculate or may be known already as a threshold instead. Citing previous studies, we presumed that average wind speeds in the categories ([v]) could be used as thresholds instead of zero points.

Based on the results of the detailed analysis related to this, it can be stated that if the average wind speed of the present day is less than the average speed of the category, the increase in the average wind speed of the following day is 1.4–2.3 times more likely, on average 1.9 times, than its decrease. If, on the other hand, the daily average wind speed is higher than the average speed in the category, the probability of a decrease in the next daily average wind speed is 1.6 to 5.2 times, on average 2.4 times, greater than that of an increase. Therefore, only [v] depends on the weather situation.

Therefore, in order to make our model usable operationally for estimating the sign of the change in the average wind speed by the next day, the following conditions must be met:

– The average of the long-term wind speed of the site at an altitude of 10 m for the whole period and its selected subsets has to be known.
– The exact average wind speed of the present day, transformed to 10 m height has to be known.

However, the exact average daily speed can only be determined from hourly data at the end of the day. In order to use the estimate, this data shall be known sooner, therefore, an approximate value that can be calculated earlier has to be applied. The method intended to determine this value is presented in the following.

2. The sliding average model

The problem can be generally stated as follows: The measured values of a climatic element are recorded at regular times (e.g., hourly, daily) during periods i=1, 2, ..., (n–1), n. The exact average of this time series can only be determined after measurement n. However, in some cases, it may be necessary to estimate this average before date n with an acceptable error.

2.1. Structure of the model

The statistical model to be presented in the following was designed to solve the above problem. The bases of the model were published by: *Tar*, 1990, 1993, 1995ab, 2004, 2019, *Tar* and *Kircsi*, 2001, *Tar* et al., 2001, 2007, *Tar* and *Szegedi*, 2011.

The database of the model is composed of a statistically sufficient measurement data matrix for a given climate element, the elements of which shall be $x_{i,j}$. The general form of the matrix is: j: row index, j = 1, 2, ..., (N-1), N, i: column index, i = 1, 2, ..., (n–1), n. Thus, N can represent the number of days involved in processing, and n can be the number of measurements at equal intervals (e.g., hourly) per day.

At each measurement time i, the $[x_{i,j}]$ elements of the so-called *sliding averages* matrix are counted per row j:

$$[x_{i,j}] = \frac{1}{i}\sum_{k=1}^{i} x_{k,j} \ . \tag{2}$$

Thus, $[x_{i,j}]$ represents the average calculated up to the measurement time i of row j, i.e., $[x_{n,j}]$ gives the total average of the row j. Knowing this, the so-called *relative sliding averages* are obtained as

$$R_{i,j} = \frac{[x_{i,j}]}{[x_{n,j}]}, \tag{3}$$

which shows that the average until time i is the proportion of the average of the complete row. Their average – the so-called *average relative sliding average* – has to be calculated at each measurement time:

$$[R_i] = \frac{1}{N}\sum_{j=1}^{N} R_{i,j} \ . \tag{4}$$

An example of deriving the above parameters is given in *Table 1*, where $x_{i,j}$ is the hourly (i) wind speed on day j of the data matrix measured at Debrecen meteorological station on July 20, 1991. The process on one day is presented in *Fig. 2*.

Table 1. Relative sliding average parameters of the wind speed data matrix measured at Debrecen station on July 20, 1991

i	$x_{i,j}$	$[x_{i,j}]$	$R_{i,j}$
1	2.9	2.90	0.78
2	2.9	2.90	0.78
3	1.8	2.53	0.68
4	2.9	2.63	0.71
5	3.4	2.78	0.75
6	3.3	2.87	0.77
7	4.2	3.06	0.83
8	4.7	3.26	1.88
9	4.9	3.44	0.93
10	4.6	3.56	0.96
11	5.4	3.73	1.01
12	5.6	3.88	1.05
13	6.1	4.05	1.10
14	6.1	4.20	1.14
15	5.8	4.31	1.16
16	4.7	4.33	1.17
17	4.2	4.32	1.17
18	2.8	4.24	1.15
19	1.4	4.09	1.11
20	0.8	3.93	1.06
21	0.8	3.78	1.02
22	2.8	3.73	1.01
23	2.9	3.70	1.00
24	3.5	3.69	1.00

$R_{i,j}$ and thus $[R_i]$ do not depend on the height of the device, because they are relative quantities. Their value is also not disturbed by changes in the height of the device, if it can be taken as constant within a series (e.g., one day). However, $[R_i]$ depends on the selected climate element and is presumably dependent on the location of the observation, the weather situation, as well as the season. Therefore, it is advisable to produce the average relative sliding average at a given location in addition to the entire database for certain subsets of this, e.g., by macrosynoptic position group or situation, for the growing season, seasonally, etc.

The $[R_i]$parameter is used for testing the model and of course for its operative running, i.e., for the estimation of the exact series average outside database of the given climatic element used in the formation of the model. The row averages $[x_{n,j}]$ (e.g., the daily averages) are estimated from the sliding average $[x_{i,j}]$ at measurement time i. The estimation is made by using the average of $R_{i,j}$ instead of $R_{i,j}$ in Eq.(3), i.e.,

$$[x_{n,j}]_{estim,i} = \frac{[x_{i,j}]}{[R_i]} \ . \tag{5}$$

Therefore, the estimated value of the total average also depends on the time point from which the estimate is made.

Of course, a different parameter of R_i distribution selected for the particular goal (i.e., mode) can also be used instead of $[R_i]$ in the course of the estimation.

In the course of the *verification*, Eq.(5) is performed at all times of all series, then – as $[x_{n,j}]$ is known – the $[E_i]$ average of the relative error of the estimations (in %) at the times is calculated from the relative error per estimate, $E_{i,j}$

$$E_{i,j} = 100 \frac{|[x_{n,j}]_{estim,i} - [x_{n,j}]|}{[x_{n,j}]} \ , \tag{6}$$

hence

$$[E_i] = \frac{1}{N} \sum_{j=1}^{N} E_{i,j}. \tag{7}$$

Eq.(6) measures the magnitude of the daily relative error, which is always positive or zero. If we also want to examine the sign of the relative error, we use the form of Eq.(6) without absolute value.

The deviation of the approximated or modeled values from the actual values is most often measured by the RMSD (root mean square deviation) parameter (*Armstrong* and *Collopy*, 1992; *Olaofe* and *Folly*, 2012). This number is actually the so-called residual standard deviation, that is, the square root of the mean of the square errors. RMSD is sensitive to outliers, which means that larger errors disproportionately affect its value. On the other hand, RMSD is a measure of accuracy, to compare forecasting errors of different models for a particular dataset and not between datasets, as it is scale-dependent (*Armstrong* and *Collopy*, 1992). Because of these, the relative errors defined by Eq.(6) were chosen to verify the model. Since the estimation given by Eq.(5) is performed at each measurement time point (i), the magnitude of the first few errors would disproportionately increase the RMSD value. For this reason, the trend of changes in errors over time (necessary reduction) would not be clear either. On the other hand, due to the scale dependence of RMSD, it is not possible to compare the usability of the model for different climatic elements. However, the use of relative values reduces the dependence of the error rate on the size and number of sample elements. This makes the $[E_i]$ parameter more comparable and increases the information content of the conclusions that can be drawn.

For actual series-by-series estimates, errors (Eq.(7)) are determined after calculating the last (n) sliding average, $x_{n,j}$ giving the actual average so that $[x_{n,j}]_{estim,i}$ estimates are stored.

If estimating the sum of the data matrix per series is the aim (e.g., monthly precipitation or global radiation), we use sliding sums instead of sliding averages.

2.2. *Database of model development*

The database of model development is now composed of the hourly wind speeds of five Hungarian meteorological stations: Szombathely, Budapest-Pestszentlőrinc, Debrecen, Szeged (non-mountain stations), and Kékestető (height a.s.l. is 1011.3 m) in the period 1991–2000. In the case of the diagrams showing daily runs, we use the 24-hour schedule accepted in Hungary to mark the times.
The statistical parameters of the model have been determined and verified for the entire period above and for its following subsets: the anticyclone and cyclone situational group of Péczely's macrosynoptic situations (*Péczely*, 1961; *Károssy* 1993, 1998, 2001) and seasonally.

The total number of days in the period 1991–2000 is 3653. Measurements of 78 days are missing in Szombathely and those of 5 days are missing in Budapest. This, however, causes no significant difference between the proportions of neither the situation groups nor the number of days of the seasons compared to the other three stations. The proportion of days in the anticyclone situational group varies between 67.1 and 67.4%, and the proportion of days in each season varies between 24.4% and 25.5%.

2.3. *Specifics of the daily changes of the average relative sliding average*

The hourly sliding averages shall therefore be calculated first from the hourly wind speeds for each day of the whole period or of the above subsets, based on Eq.(2). The 24th hour sliding average is the average wind speed per day *j*; dividing the sliding averages by this, the hourly value of the Eq.(3) relative sliding averages is obtained. The process on one day is shown in *Fig. 2*. This is followed by averaging the relative sliding averages by the hour, i.e., the determination of [R_i] values.

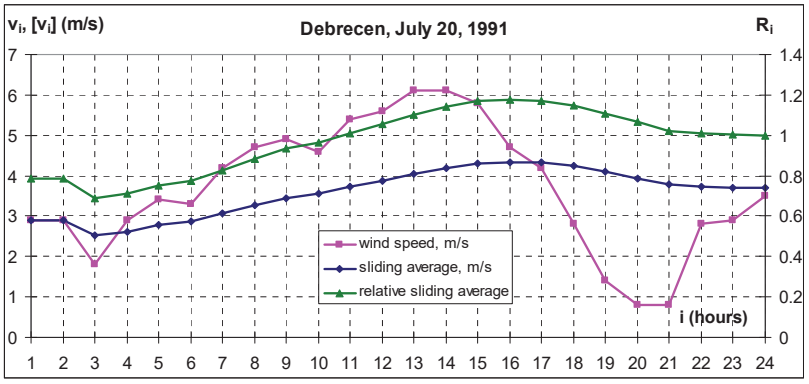


Fig. 2. The values of hourly wind speed (v_i), sliding average ([v_i]) and relative sliding average (R_i) in Debrecen on July 20, 1991.

Fig. 3 shows the daily changes of [R_i] at the five stations for the entire period, as well as for the anticyclone and cyclone situation groups and seasons. Hourly differences exceed 0.05 only in the anticyclone situation group in the early hours after midnight. Between about 1pm and 8pm, the decreasing order of Szeged, Debrecen, Szombathely, Budapest is formed in all three cases. This is also more or less observed in the seasons, most notably in autumn.

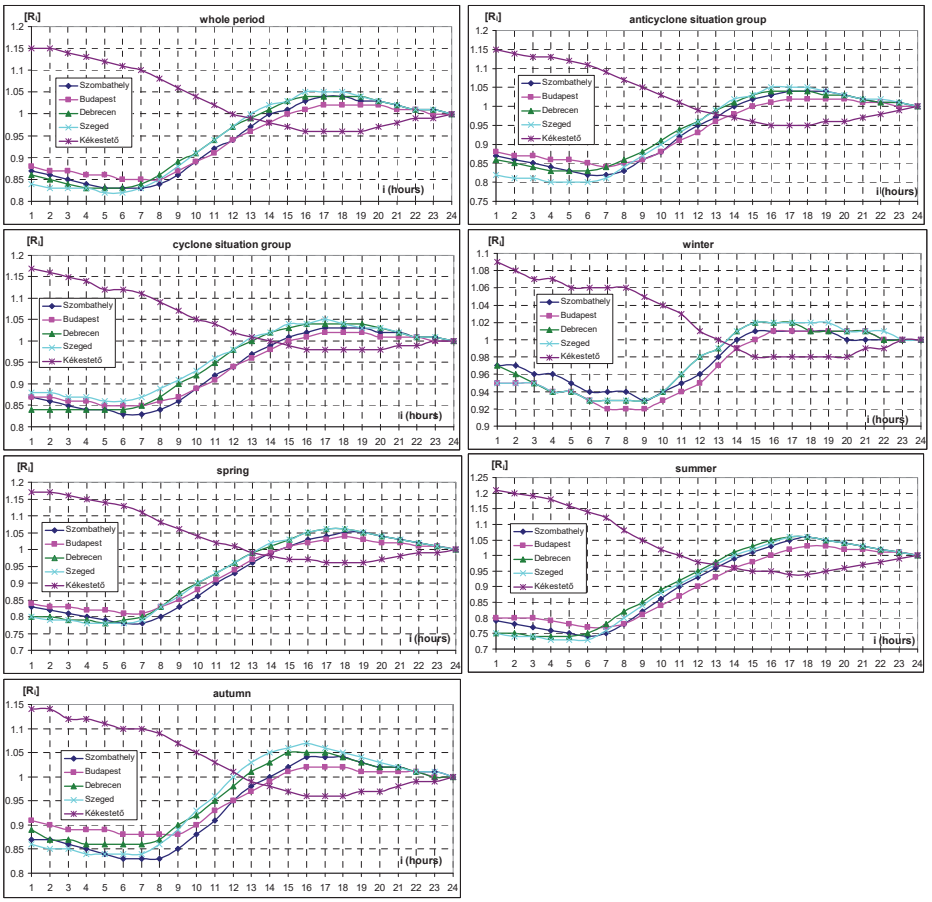


Fig. 3. Daily course of mean relative sliding averages ([R_i]).

At the four non-mountain stations, the waves of [R_i] take their lowest values between 4am and 7am ranging from 0.82 to 0.86. The maximums occur at 5pm and 6pm, with values between 1.02 and 1.05. In all three cases, the amplitude (maximum-minimum) of the [R_i] waves is the smallest in Budapest (0.17, 0.18). The highest amplitude was observed in Szeged (0.23, 0.25) during the entire period and in the anticyclone situation group, however, in the cyclone situation group, in addition to the Budapest minimum, amplitudes can be considered equal.

In winter, extreme values persist for a long time. The maximum value (1.01, 1.02) can be observed for 4–6 hours between 3pm and 9pm, and the minimum value (0.92, 0.93) can be observed for 3–4 hours between 6am and 9am. Szombathely is an exception to the latter, with the minimum value occurring at 9am. Amplitudes are practically equal (0.08, 0.09). In spring, the maximum value (1.04-1.06) occurs between 5pm and 7pm and can be observed at all four stations at 6pm. The minimum value (0.78-0.80) occurs between 4am and 7am, and can be observed at 5am or 6am at all four stations. The amplitude is the smallest in Budapest with 0.23, and can be considered equal at the other three stations (0.27, 0.28). In summer, maximums (1.03–1.06) occur everywhere at 6pm, while the minimums (0.73-0.74) occur between 5am and 6am at the four stations.

The amplitude is again the smallest in Budapest, with 0.26, and can be considered equal again at the other three stations (0.32, 0.33). In autumn, the maximum value (1.02–1.07) occurs at 4pm everywhere except Szeged, where it occurs at 5pm and 6pm as well. The minimums (0.83–0.88) occur at 6pm or 7pm at all four stations. Amplitudes are now more diverse: 0.14 in Budapest, 0.19 in Debrecen, 0.21 in Szombathely, and 0.23 in Szeged.

At the altitude of Kékestető, the daily wind speed shows its minimum at early afternoon, consequently, the daily changes of $[R_i]$ are the opposite of that of the other four stations, which are located at much lower altitudes. The maximums occur at 1am or 2am (1.21–1.09) for each of the seven cases. The time of the minimums is spread between 3pm and 8pm, but occurs at 5pm and 6pm in all seven cases. The amplitude is the largest (0.27) in summer and the smallest (0.11) in winter.

2.4. Verification

Using the known time series of sliding averages and average relative sliding averages produced from hourly wind speeds of the present day, the average daily wind speed can be estimated at any hour of the day based on Eq.(5). According to the above, the estimation can be made based on the time series of $[R_i]$ for the whole period and for the macrosynoptic situational group, and also on its seasonal time series.

For the verification of the model, the wind speed database involved in the modeling was used. Since the daily average wind speed is now known, the error of the estimate can be calculated per hour and then the average of these can be calculated as well. The process is illustrated in *Fig. 4*. Average hourly estimate errors calculated on the basis of Eq.(7) are shown in *Fig. 5*.

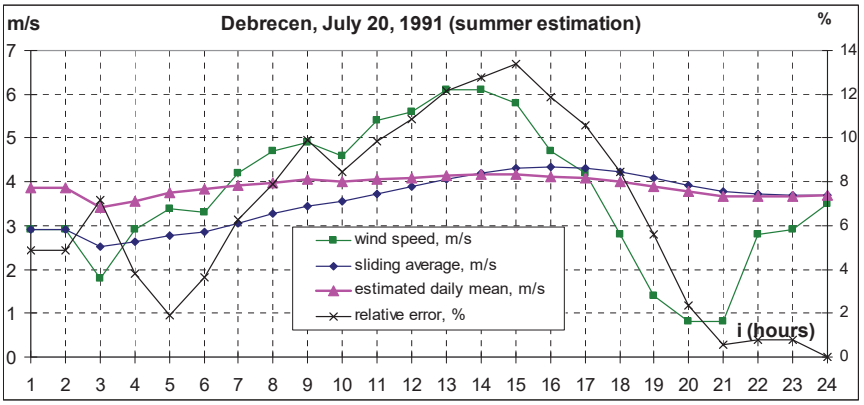


Fig. 4. The estimation process to verify the model.

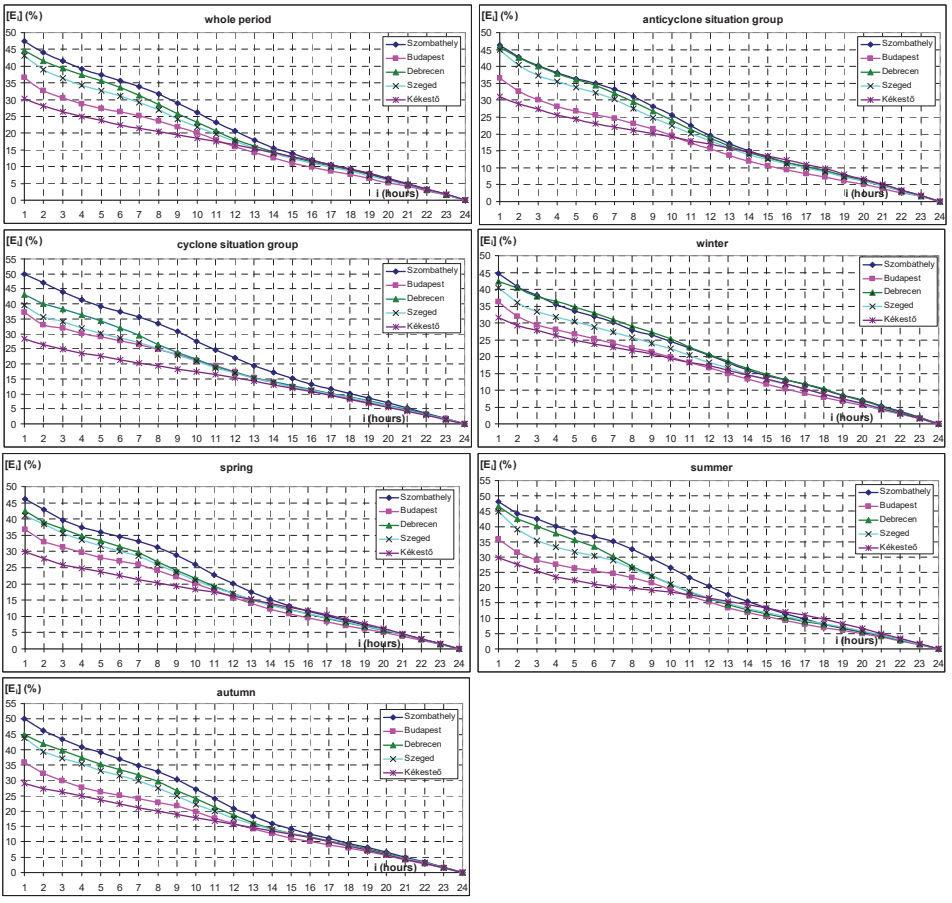


Fig. 5. The daily course of the hourly average relative error of the estimate.

According to *Fig. 5*, the average relative error ([E_i]) decreases rapidly as the time of the estimate approaches the end of the day. The values of the [E_i] time series can be approached well with the linear trend. The steepness of the linear trend specifies the value of the average daily decrease, and the absolute value defines the measure. The average hourly decrease has the highest absolute value at Szombathely and the lowest at Kékestető, and the order between them is Debrecen, Szeged, Budapest in all categories. A certain orographic effect can therefore be assumed, since the last three stations are located in a lowland environment, and Kékestető has an even more open horizon. The maximum values at Szombathely are between 2.1 and 2.2%/hour except in spring, and the maximum in spring is 1.85%/hour. The minimum values at Kékestető are between 1.2 and 1.3%/hour. In the other three stations, the average daily decreases are between 2.0 and 1.5 %/h in any category. As a result of the rapid decrease, the values of [E_i] fall below 20% in all cases after 1pm.

Fig. 5 also indicates that if the estimation of the daily averages during the early afternoon is enough, it will not be necessary to subdivide the studied period.

3. Testing the model

For the operational application of the model, the user must have a time series of average relative sliding averages ([R_i]) produced from long-standing hourly wind speeds at that location or at a nearby weather station for at least the entire (annual) period. (Under the entire period, we mean at least one year, i.e., in this case we get the annual averages of [R_i].) On a given day, the estimation of the average wind speed of the next day has to be performed based on one or more sliding averages of the period after 12 o'clock. The average of these can also be considered as a good estimate.

It is assumed that the estimated value of the daily average wind speed will approach the true value with the smallest error if the estimate is performed from time i where [R_i] ≈1. These times before 4pm – at least 8 hours before the end of the day when it is still worth to perform the estimate – allowing an absolute deviation of 0.01 ($0.99 \leq [R_i] \leq 1.01$) are shown in *Table 2*. The most frequent times listed in the table are 1pm, 2pm, and 3pm.

Table 2. The times (hours) in which $[R_i] \approx 1$

Period	Szombathely	Budapest	Debrecen	Szeged	Kékestető
whole	14, 15	15, 16	13, 14	13	12, 13
anti-cyclonal	14	15, 16	13, 14	13	11, 12
cyclonal	14, 15	15, 16	13	13	13, 14, 15
winter	14, 15, 16	14, 15, 16	13, 14	13, 14	12, 13, 14
spring	14, 15	14, 15	13, 14	13	12, 13
summer	14, 15	16	14	14	11
autumn	14	15, 16	13	12	12, 13

In the following, model simulations using annual SODAR data measured at 30 m height, in Debrecen, in 2013, are discussed. The input – similarly to wind speed measured on wind power plants – is a sequential file, the records of which include the date (year, month, day) and time of measurement (0, 10, 20, 30, 40, and 50 minutes of every hour). The estimate is made at 1pm, 2pm, and 3pm using the average relative sliding averages ($[R_i]$) determined for the whole period (year). The daily pattern of $[R_i]$ values in Debrecen is presented in *Fig. 6*. It can be seen that the values are somewhat larger (by 0.02–0.03) at the three times in autumn than in other cases, where they are nearly equal. This, however, does not mean that estimating the daily change of the average daily wind speed would be independent of the season, because the other parameter of the estimate is the average wind speed for the period.

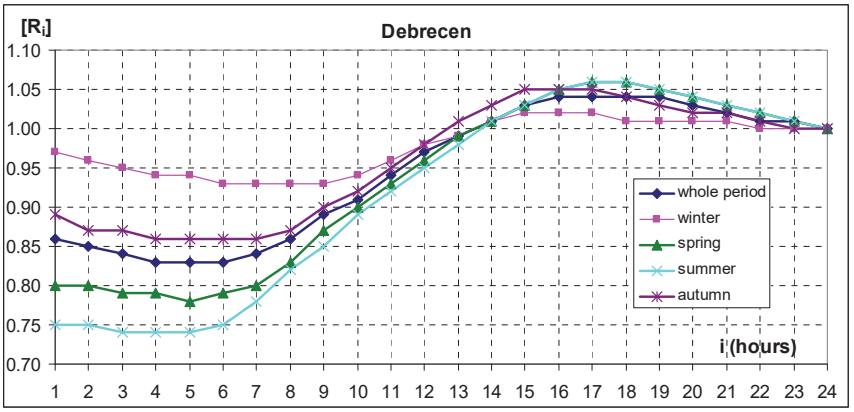


Fig. 6. Annual and seasonal daily course of hourly average relative sliding averages in Debrecen based on SODAR data measured at 30 m height in 2013.

Monthly analyses are also carried out throughout the whole period (year). From these we can infer the reality of the annual pattern of each characteristics (the significance of the differences between them).

3.1. Frequency of accurate, under-, and overestimates

First, we analyze the frequency of the sign of the estimation error, i.e., the frequency of accurate, under-, and overestimation is analyzed. The estimate is considered accurate if its difference from the real daily average is 0.0 to one decimal point. Underestimation and overestimation mean that the difference is negative or positive, respectively. The annual pattern of the proportion of accurate estimates and the differences of underestimates and overestimates are shown in *Fig. 7*.

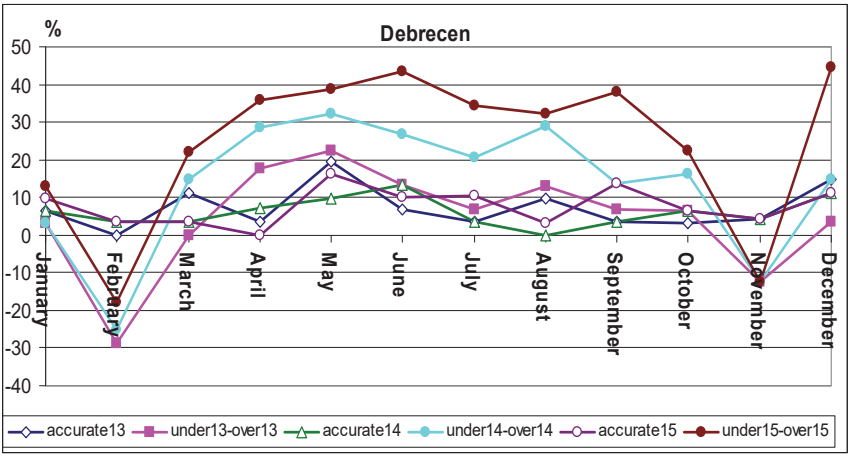


Fig. 7. The annual course of the proportion of accurate estimates and the difference between under- and overestimates. (Note, that 13 refers to 1pm, 14 refers to 2pm, and 15 refers to 3pm.)

Surprisingly, the maximum of the monthly relative frequency of accurate estimates is divided between 1pm and 3pm in almost 50%–50% of the cases. The difference between the monthly ratio of underestimates and overestimates is negative in February and November (i.e., the number of overestimates is higher in these months), but with the exception of November, this difference is the largest for the 3pm estimate. Therefore, the number of underestimates increased as the estimate time increases in the present case. The evaluation of the full-year results of the estimates also shows this (see *Table 3*).

Table 3. Accurate, underestimation and overestimation rates throughout the year

% at	accurate	under-estimation	over-
1 pm	7.2	48.8	43.9
2 pm	6.1	54.0	39.9
3 pm	7.8	58.7	33.5

3.2. Statistics of simple estimate error

The simple measure of the signed estimation error is the difference between the estimated and the actual daily average wind speed, i.e., $[x_{n,j}]_{estim,i} - [x_{n,j}]$, and now i=13, 14, 15, j=1, 2, …., N.

The main statistical characteristics of this error are briefly analyzed for the three estimates for the whole year, paying attention to the fact that the real difference is the absolute value of the error. The values of the most important characteristics are given in *Table 4*.

Table 4. The most important statistical characteristics of daily simple estimation errors (m/s, Debrecen, 2013)

m/s, at	1pm	2pm	3pm
mean	0.00	-0.07	-0.14
st. deviation	0.71	0.64	0.56
maximum	2.80	2.30	1.80
minimum	-2.20	-2.00	-1.80
range	5.00	4.30	3.60
skewness	0.32	0.25	0.16
kurtosis	0.99	0.86	0.76
mode	-0.20	-0.30	-0.20
median	0.00	-0.10	-0.20

The characteristics of the simple estimation error decrease over time, except for the minimum value and mode, and the minimum increases. Modes can also be considered equal. This can be decided based on the frequency distribution of errors, which is prepared by classifying the errors into 0.2 m/s long intervals taking the extreme values into account as well.

The columns of *Fig. 8* show that errors are classified into the interval of (-0.4:0.0) m/s with highest frequencies, 27.7, 31.2, and 32.9% similarly to at all three times, averages and medians. In addition, the values of the skewness and

kurtosis coefficients suggest the possibility of an approach with normal distribution. To determine the goodness of the fit, χ^2 test was used. Accordingly, the hypothesis that the frequency distribution of the magnitude of the estimation errors is normal, is not rejected at a significance level of 0.05 for any of the hourly estimates. This means that differences between -0.4 and 0.0 m/s are most likely to occur at all three estimation times.

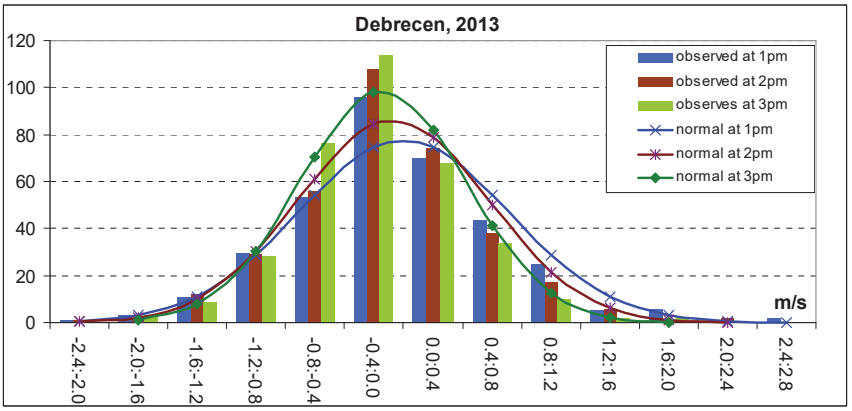


Fig. 8. Frequency distribution (days) of the magnitude of the daily simple estimation errors and their approximation to the normal distribution.

3.3. Relative error of the estimations

We examine the more important statistical properties of the exact relative deviation of the results of the estimates from the known daily average wind speed. For this, we use the absolutely value-free form of Eq.(6) and Eq.(7).

Fig. 9 shows the average monthly relative errors. It can be seen that their value decreases every month as the time of the estimate increases. The average monthly relative error is positive in a total of 7 cases: for all three cases in February and only for the estimates at 1pm and 2pm in November. The primary and secondary maximums, i.e., when the number of overestimates is greater than that of underestimates are found in these months.

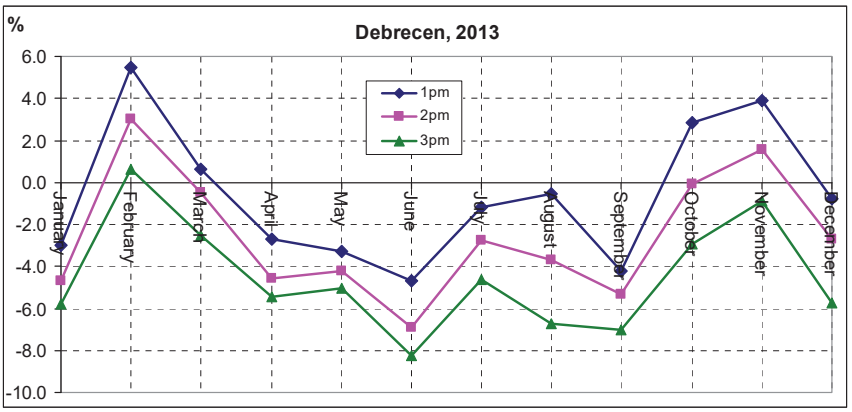


Fig. 9. Annual course of the monthly average relative error of the estimates.

The most important statistical characteristics of the annual process of day-to-day relative error are shown in *Table 5.*

Table 5. The most important statistical characteristics (%) of the annual course of the relative error per day

at	1pm	2pm	3pm
mean	-0.74	-2.68	-4.62
st. deviation	18.56	16.75	14.87
maximum	58.30	47.90	37.50
minimum	-53.10	-53.10	-53.10
range	111.50	101.00	90.60
skewness	-0.03	-0.08	-0.12
kurtosis	0.15	0.20	0.18
mode	0.00	0.00	0.00
median	0.00	-3.00	-5.30

The characteristics of the relative error per day decrease with time except for the minimum value and the mode, which do not change. The comparison of the mean, mode, and median values once again raises the possibility that the studied data originate from a normal distribution. To determine this, the frequency distribution of the magnitude of the studied errors are exemined taking into account the extreme values with classifying them into intervals of 10%.

Fig. 10 shows the observed frequencies and those approached with normal distribution at the three estimation times. According to the χ^2 test, the hypothesis that the frequency distribution of the daily relative estimation errors is also normal is not rejected at 0.05 significance level for the estimates at either hour. This means that relative errors between -10% and 0% are most likely to occur at all three estimation times (1pm, 2pm, and 3pm).

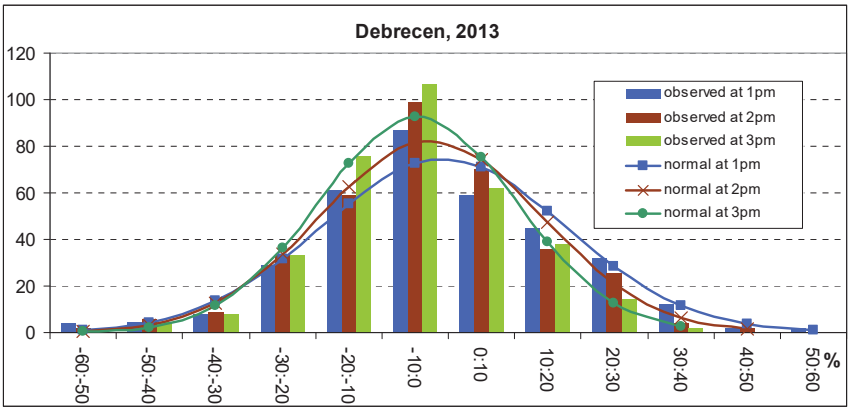


Fig. 10. Frequency distribution (days) of daily relative estimation errors and their approximation to the normal distribution.

Examining the statistical characteristics of the daily simple and relative estimation errors does not give an accurate picture of the difference between estimated and true values. Errors with different signs may, for example, balance each other in the course of averaging, therefore, an average error close to zero may be obtained. The real differences, i.e., the absolute values of these errors, are a more pronounced indicators of the reliability of our model. Therefore, we now examine the absolute value (magnitude) of the daily relative errors, as you see in Eq.(6). Eq.(7) gives the average of these for different periods. From *Fig. 5* it can be concluded that in Debrecen, this error is on average between 15% and 12% at the three selected estimation times in the whole period in the case of the 10-year-long time series used in model construction.

Based on *Fig. 11*, monthly averages are the highest in November and the smallest in May at all three estimation times. In May, August, and December, the average of the three estimates is almost the same, and in the other months – except March – average errors decrease as the estimation times increase.

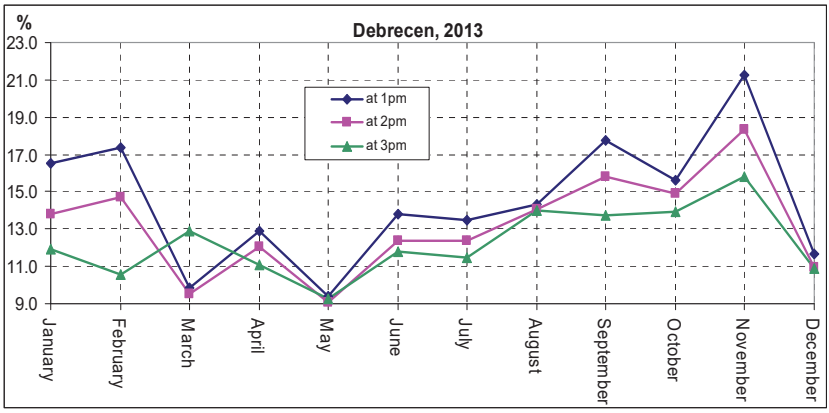


Fig. 11. Monthly averages of the magnitude of the daily relative estimation errors.

Looking at the annual pattern of the magnitude of the daily relative estimation error, its most important statistical characteristics and frequency distribution were determined (*Table 6*). Averages are a very good match to the values shown in *Fig. 5*. To determine the mode, intervals of 5% were used to create frequency distributions shown in the columns in *Fig. 12*. Most data can be categorised into the 0-5% interval at 1pm and 3pm while into the 5–10% interval at 2pm. The centre of these can be considered the value of the mode.

Table 6. The most important statistical characteristics (%) of the annual course of the magnitude of the relative estimation error per day

at	1pm	2pm	3pm
mean	14.68	13.35	12.31
st. deviation	11.36	10.44	9.52
maximum	58.33	53.13	53.13
minimum	0.00	0.00	0.00
skewness	1.06	1.11	1.12
kurtosis	1.17	1.28	1.60
mode	2.50	7.50	2.50
median	12.50	10.53	10.34

It follows from both the minimum and the mode values that the frequencies can be approximated by a monotonous descending theoretical distribution. Exponential and gamma distributions were tested. Despite the fact that the parameter determining the shape of the gamma distribution is in no case smaller than 1, in which case the mode would fall in the 0-5% interval (*Dévényi* and *Gulyás*, 1988), the latter proved to be successful. Thus, the maximum of the theoretical distributions in all three cases is within the 5-10 % interval, not only

at the 14-hour estimate (*see Fig. 12*). However, based on the χ^2 test, the hypothesis that the frequency distribution of the magnitude of the daily relative estimation errors has gamma distribution at a significance level of 0.05 is not rejected for any of the hourly estimates. This means that real differences between 0 % and 10% are most likely to occur at all three estimation times.

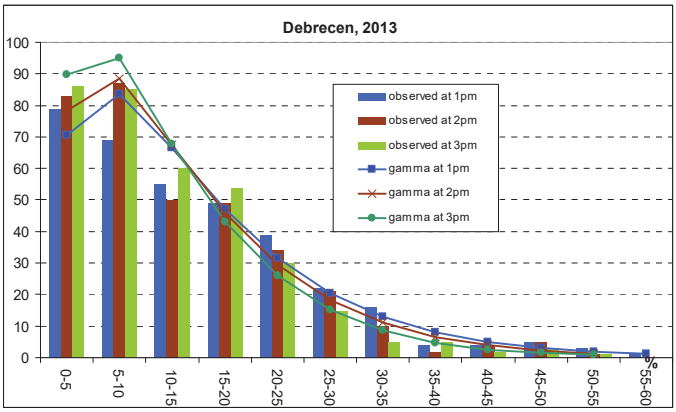


Fig. 12. Frequency distribution (days) of the magnitude of daily relative estimation errors and their approximation to the gamma distribution.

3.4. The estimated present-day and the accurate previous-day daily average wind speed

Based on of above results, if we have decided which estimate to accept as the average wind speed of the present day, this can be compared with the known average wind speed of the previous day, it can be decided whether the average wind speed of the present day and with it the average daily wind power decreased or increased compared to the previous day.

Considering the difference between the average wind speeds estimated for the present day and known for the previous day, if the difference is negative, the average wind speed of the present day decreases compared to that of the previous day. If the sign of this difference is compared with the sign of the difference in real average wind speeds, very valuable information can be obtained regarding the reliability of the model.

Let us have, as an example, the average wind speed of the present day, estimated at 1pm, $\bar{v}_{te}$ =4.3 m/s, and that of the previous day, $\bar{v}_{yr}$ =5.3 m/s. Their difference is negative, therefore, the average wind speed of the present day is expected to be smaller than 5.3 m/s. At the end of the day, the average wind speed of the present day is found to be $\bar{v}_{tr}$ =3.7 m/s, which means that the estimate was

correct. Therefore, reliability of the model can also be tested so that the sign of the $\bar{v}_{te} - \bar{v}_{yr}$ and $\bar{v}_{tr} - \bar{v}_{yr}$ differences are compared. Their similarity indicates that the estimation is correct.

Similarity of the signs is tested monthly, taking into account the missing days as $\bar{v}_{yr}$ is missing on the following days. The number of days when this was the case – i.e., sign($\bar{v}_{te}$–$\bar{v}_{yr}$)=sign($\bar{v}_{tr}$– $\bar{v}_{yr}$) – compared to the number of days can be considered in the model (in %) are presented in *Fig. 13*.

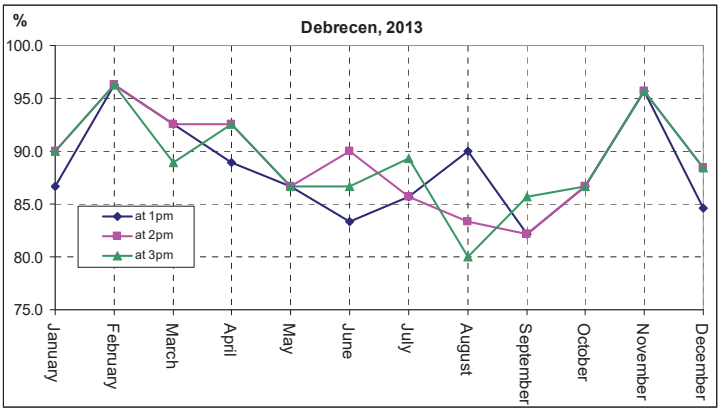


Fig. 13. The identical signs of the difference between the estimated current and yesterday's real, and today's real and yesterday's real average wind speeds per month in relation to the number of days to be taken into account (in %).

According to the chart, the maximum of the proportions is set in February at 96.3% for all three estimates. This is also the case in November with a secondary maximum of 95.7%. The minimum value is 82.1% in September at 1pm and 2pm and 80.0% at 3pm in August. In other words, the sign of the differences examined is between 80.0% and 96.3%. Our assumption, that these rates do not decrease as the estimate date increases, is true with the exception of March, June, and August.

For the whole year, these rates are 88.6, 89.5, and 89.2, i.e., with a slight maximum at 2pm. On average, therefore, there is a nearly 90% probability that if the average wind speed of the present day is smaller or greater than the real average wind speed of the previous day, then the real wind speed of the present day will change accordingly by the end of the day.

3.5. Change in the average wind speed of the next day compared to that of the present day based on the estimated present-day average wind speed

In two previous articles (*Tar* and *Lázár*, 2018; *Tar*, 2021), the average daily wind speeds observed at different stations were transformed to 10 m. In addition to comparability, this was also justified by the fac,t that the height of the anemometer at certain stations changed during the studied period, and therefore, a reference level was also required. The main conclusion of the detailed analysis carried out is that if the average wind speed of the present day is less than the average speed of the category (year, season, macrosynoptic situation group) over many years, then the increase of the average wind speed of the following day is 1.4–2.3 times more likely, on average, 1.9 times more likely than its decrease. However, if the average daily wind speed is greater than the average speed of the category, the decrease of the average wind speed of the following day is 1.6–5.2 times, on average, 2.4 times more likely than its increase.

The change in the average wind speed of the next day compared to that of the present day is determined from the average wind speeds of the present day estimated as detailed above using the annual 30m SODAR data of 2013. According to the above, these observed and estimated daily average wind speeds should be transformed to 10 m. This was performed using the so-called WMO formula (*Mezősi* and *Simon*, 1981), according to which the values at the 10 m height is about 80 % of those at 30 m. The average wind speed at 10 m during the modeled (studied) period (1991–2000) was 2.8 m/s.

The procedure of the estimation is detailed in *Table 7*. Here $[v]_e$ represents the average wind speed at 10 m estimated at 1pm of the given day. The next column shows the difference from the average wind speed of the studied category, i.e., the total period (2.8 m/s). If this is negative or 0, the average wind speed of the next day, based on the above, is most likely to increase or not to decrease (I), otherwise it will decrease (D). Based on the data in the I/D_e column, we estimate that in 32.3% of the days of the month, the average wind speed of the next day will increase compared to that of the previous day. This can be controlled, as daily averages at 10m are available for the test period. These are indicated by [v] in the table, and Δv indicates their differences on the following days. The increase is now the $\Delta v \geq 0$. The data in column I/D show that the true frequency of increase (I) is 58.1%, which was therefore underestimated by about 25%. If the same event (I or D) occurs with an estimation and observation on a given day, it can be considered a good estimate. The total number of good estimates (II+DD) in the table is 17, i.e., in January the estimate at 1pm and the observation are in accordance with each other in the case of 54.8% of the days.

Table 7. Illustration of estimation process (bold-italic letters indicate good estimate)

	estimation at 1pm			observed		
January	**$[v]_e$**	**$[v]_e$-2.8**	**I/D_e**	**[v]**	**Δv**	**I/D**
1.	0.9	-1.9	***I***	0.9	0.2	***I***
2.	0.6	-2.2	***I***	1.1	3.1	***I***
3.	3.1	0.3	***D***	4.3	-1.3	***D***
4.	3.5	0.7	D	3.0	1.7	I
5.	3.8	1.0	***D***	4.7	-1.7	***D***
6.	3.5	0.7	D	3.0	2.7	I
7.	6.6	3.8	***D***	5.6	-2.9	***D***
8.	3.2	0.4	***D***	2.7	-0.8	***D***
9.	2.5	-0.3	***I***	1.9	0.3	***I***
10.	2.1	-0.7	***I***	2.3	1.1	***I***
11.	3.1	0.3	***D***	3.4	-0.6	***D***
12.	2.4	-0.4	***I***	2.8	0.3	***I***
13.	2.3	-0.5	***I***	3.1	0.4	***I***
14.	3.8	1.0	***D***	3.5	-0.8	***D***
15.	2.8	0.0	D	2.7	0.2	I
16.	3.6	0.8	D	2.9	1.9	I
17.	4.7	1.9	D	4.8	1.4	I
18.	6.7	3.9	***D***	6.2	-3.4	***D***
19.	2.7	-0.1	I	2.8	-0.1	D
20.	2.9	0.1	D	2.7	0.2	I
21.	2.9	0.1	D	2.9	0.6	I
22.	3.1	0.3	***D***	3.5	-0.2	***D***
23.	3.6	0.8	D	3.2	1.7	I
24.	3.4	0.6	D	4.9	2.3	I
25.	7.7	4.9	***D***	7.2	-3.6	***D***
26.	4.3	1.5	***D***	3.6	-1.4	***D***
27.	1.7	-1.1	N	2.2	-0.7	D
28.	1.8	-1.0	***I***	1.4	2.7	***I***
29.	2.3	-0.5	I	4.1	-0.3	C
30.	3.1	0.3	D	3.8	1.4	I
31.	4.3	1.5	D	5.2	1.6	I

The analysis detailed above was carried out monthly for all three estimate times. Therefore, the reality of the assumed periods (half-year, seasonal) and the best estimate time can be identified. Finally, results for the whole year are also provided.

Fig. 14 shows the monthly frequency of the occurrence of event N in the percentage of the applicable days in the given month.

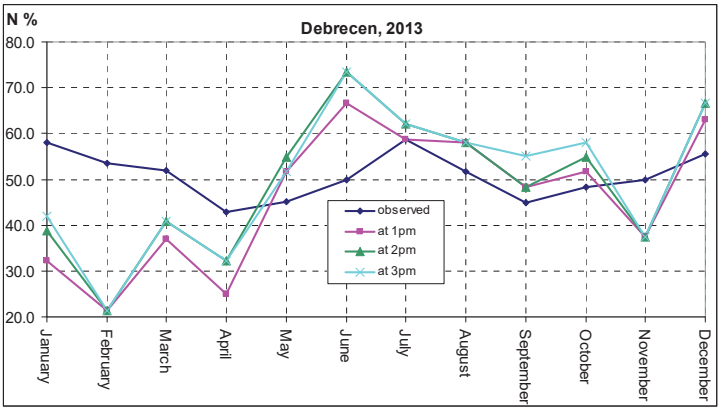


Fig. 14. Observed and estimated monthly frequencies of occurrence of event N in % of usable days at the three time points.

The figure shows that the difference between the estimates at the three different times is unlikely to be significant. On average, the estimates at 3pm and 2pm differ from each other the least, by 0.8%, and those at 3pm and 1pm differ the most, by 4.0%. However, all three estimates approximate the observed frequency with statistically reasonable accuracy only during the July-October period.

The regularity in the annual pattern of the curve showing the observed frequency suggests the presence of a real period. According to the period analysis (*Dobosi* and *Felméry*, 1971; *Tar* and *Kircsi*, 2001; *Tar et al.*, 2002; *Matyasovszky*, 2002; *Tar*, 2007, 2008ab), the wave with a half-year period has a real, non-random annual pattern. This means that the probability of the occurrence of event N during the winter and summer months is significantly higher than in the other two (transitional) seasons (see *Fig. 15*).

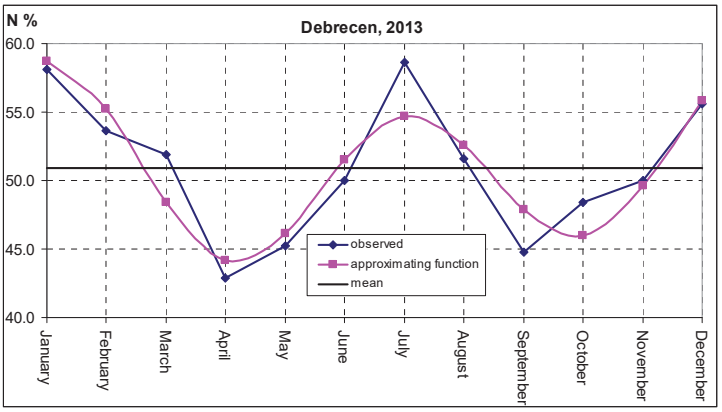


Fig. 15. The annual course of the observed frequency of event N and the trigonometric function with a realistic period of half a year approaching it.

Regarding the whole year, on the basis of the observed (real) values, the average daily wind speed decreased in 49.1% of all days compared to the previous day and increased (not decreased) in 50.9%. The probability of the two events can therefore be considered roughly equal in the studied year. The values calculated from the estimates at 3pm are the closest to the values of 49.7% and 50.3%. Regarding the other two estimate times, the frequency of event C is 7.5% and 1.2% higher than that of event N.

Monthly frequencies of good estimates are shown in *Fig. 16* as a % of the days that can be used in a given month.

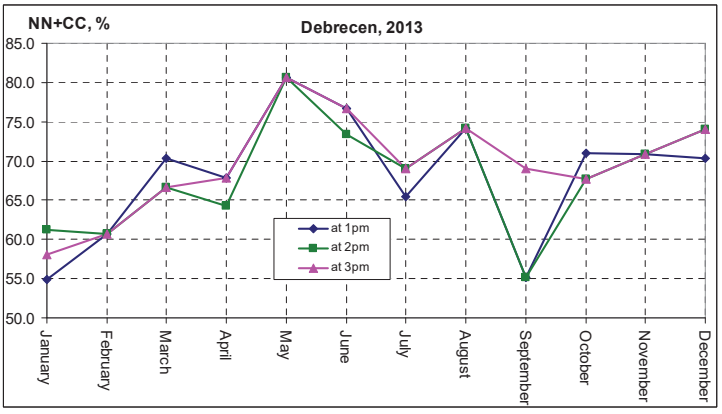


Fig. 16. Monthly frequencies of coincidence of estimated and observed N and C events.

The frequency of good estimates ranges from 80.6% (May, all three times) to 54.8% (January, 1pm). The curves run well together except for September, when there is a nearly 14% difference between the frequencies at 3pm and also at 1pm and 2pm. On an annual average, however, the frequency of good estimates for 1pm and 2pm is only 1.5% behind the estimates at 3pm, which is 69.7%.

The September estimate was also carried out with the autumn estimation parameters: with 1.01, 1.03 and 1.05 [R_i] values and a seasonal average of 2.5 m/s. However, the number of good estimates decreased at 2pm and 3pm compared to the values shown in *Fig. 16*.

4. Summary

According to our previous study, the day-to-day change of the average daily wind speed has an appreciable stochastic relationship with the average wind speed of the previous day. The conclusion based on the regression function is that if the average wind speed of the present day is less than the average speed of the studied period (e.g., year, season), average increase in the average wind speed of the next day is nearly twice as likely as the decrease. However, if the average daily wind speed is greater than the average speed of the category, decrease in the average wind speed of the next day is, on average, nearly two and a half times more likely than its increase. Thus, knowing the average speed of the present day, makes it possible to estimate the sign of a change for the next day, which can help to prepare a timetable. However, in order to perform this estimate, the average wind speed of the present day has to be estimated as well at the time of day that can be used for the preparation of the timetable. The sliding average model was established to face this problem.

After the detailed analysis of the daily pattern of the annual and seasonal average hourly sliding averages ([R_i], i=1,2,...,24) produced from the ten-year hourly wind speeds of five Hungarian meteorological stations, the average daily wind speed was estimated from each value of the above using a simple ratio, and the relative error of the estimate was also determined.

The average hourly relative error ([E_i]) decreases rapidly as the estimate time approaches the end of the day. The average hourly decrease is the highest in Szombathely and the lowest at Kékestető regarding absolute values, and the order between them is Debrecen, Szeged, Budapest in all categories (years, seasons). As a result of rapid decreases, the values of [E_i] fall below 20% after 1pm in all cases.

For the testing of the model, the annual SODAR data measured at 30 m height in Debrecen in 2013 were used. Daily average wind speed is estimated at 1pm, 2pm, and 3pm using the average relative sliding averages ([R_i]) determined for the whole period (year) monthly. The estimate is considered accurate if its difference from the real daily average is 0.0 to one decimal point. Underestimation and overestimation mean that the difference is negative or positive, respectively.

The maximum of the monthly relative frequency of accurate estimates is divided between 1pm and 3pm in almost 50%–50% of the cases. The difference between the monthly ratio of underestimates and overestimates is negative in February and November (i.e., the number of overestimates is higher in these months), but with the exception of November, this difference is the largest for the 3pm estimate. Therefore, the number of underestimates increased as the estimate time increases in the present case. The evaluation of the full-year results of the estimates also shows this.

The sign estimation error is the difference between the estimated and the real daily average wind speed. The characteristics of the estimation error decrease over time, except for the minimum value and mode, and the minimum increases. Modes can be considered equal. This can be decided by the frequency distribution of the magnitude of errors. According to this, errors occur with the highest frequency within the interval (-0.4:0.0) m/s at all three times, the same way as averages and medians. Other parameters also suggest the possibility of approaching with normal distribution. According to the χ^2 test, the hypothesis that the frequency distribution of the magnitude of estimation errors has normal distribution at a significance level of 0.05 is not rejected for either estimates. This means that deviations between -0.4 and 0.0 m/s are most likely to occur at all three estimation times.

Then the main statistical properties of the relative differences of the results of the estimates in % relative to the exact average daily wind speed are known at the time of testing. The comparison of the mean, mode, and median values raises again the possibility that the studied samples are from a normal distribution. To decide this, the frequency distribution of the magnitude of the studied errors, with an interval of 10% taking into account the extreme values is exemined. According to the χ^2 test, the hypothesis that the frequency distribution of the magnitude of estimation errors has normal distribution at a significance level of 0.05, is not rejected for either estimates. This means that relative errors between -10% and 0m/s are most likely to occur at all three estimation times.

Examining the statistical characteristics of daily simple and relative estimation errors gives no accurate picture of the difference between estimated and real values. Real differences, i.e., the absolute values of these errors, are more pronounced indicators of the reliability of our model. Therefore, the monthly characteristics of the absolute value of daily relative errors were also examined. In Debrecen, at the three selected estimates, this error is on average between 15% and 12% in Debrecen at the three selected estimates for the 10-year-long time series used in model construction in the whole period. With this, the monthly averages of the values calculated here are in a very good match. Intervals of 5% were used to create frequency distributions for determining the mode. Most elements of the sample classify into the interval 0–5% at 1pm and 3pm, however, at 2pm most data belong to the 5–10% interval. Frequencies can therefore be approximated by a monotonous descending theoretical distribution. Based on the

χ^2 test, the hypothesis that the frequency distribution of the magnitude of the daily relative estimation errors has gamma distribution is not rejected at a significance level of 0.05 for any estimation time. This means that real differences between 0% and 10% are most likely to occur at all three estimation times.

Comparing the present-day average wind speed estimate with the known, real, accurate average wind speed of the previous day, it can be assumed whether the average wind speed, and with it, the average daily wind power of the present day will decrease or increase compared to those of the previous day. Let us take the difference between the estimated present-day and the known average wind speed of the previous day. If this difference is negative, then that of the present day will decrease compared to that of the previous day. If the sign of this difference is compared to the sign of the difference between the real, accurate average speeds, their matching indicates correct estimation. The maximum of correct estimates occurs in February, which is 96.3% for all three estimates. This is also the case in November with a secondary maximum of 95.7%. The minimum value is 82.1% in September at 1pm and 2pm, while at 3pm, it is 80.0% in August. In other words, the sign of the differences examined is between 80.0% and 96.3%. For the whole year, these rates are 88.6, 89.5, and 89.2, i.e., with a slight maximum at 2pm. On average, therefore, there is a nearly 90% probability that the average wind speed of the present day is less than or greater than the real average wind speed of the previous day, then the real wind speed of the present day will change accordingly by the end of the day.

Finally, based on the results of the testing the model was applied to solve the original problem. To do this, the estimated average wind speed of the present day had to be compared with the long-term average wind speed of the studied period (category). As described at the beginning of this section, if the estimated average wind speed of the present day is less than or greater than the average speed of the period, which is currently 2.8 m/s, then the increase (N) or decrease (C) of the average wind speed of the next day is more likely. The number and proportion of good estimates can be determined by comparing estimates with real average daily wind speeds.

The analyses were carried out monthly and throughout the year for all three estimation times. The monthly occurrences of event N indicate, that the difference between the three estimates is unlikely to be significant. On average, the estimates at 3pm and 2pm differ the least, by 0.8%, and the estimates at 3pm and 1pm differ the most, by 4.0%. However, all three estimates approximate the observed frequency with statistically reasonable accuracy only during the July-October period. For the whole year, on the basis of the observed (real) values, the average daily wind speed decreased in 49.1% of all days compared to the previous day and increased (not decreased) in 50.9%. The probability of the two events can therefore be considered to be equal with a good approximation in the studied year. The values calculated from the estimates at 3pm approximate best the above

values: 49.7% and 50.3%. Regarding the other two estimates, the frequency of event C is 7.5% and 1.2% higher than that of event N.

Events N and C are assigned to the present day and show that the average wind speed for the next day will increase or decrease. The time series of these events can be produced by comparing the estimated average daily wind speeds with the long-term average. A good estimate is obtained when the event assigned to the given day based on the estimation is the same as the event that can be determined from the real (known at testing) daily average wind speeds. These are events NN and CC, thus the number of good estimates is NN+CC. Curves in the annual pattern look quite the same except for September, when there is a nearly 14% difference between the frequencies related to 3pm, 1pm, and 2pm. On an annual average, however, the frequency of good estimates at 1pm and 2pm is only 1.5% behind those at 3pm, the latter is 69.7%.

From the detailed analyses above, it can be seen that by merging the model describing the daily average wind speed with the sliding average model, more information about the wind climate in Hungary can be revealed, in addition to the hopes that it will help to prepare the timetable. For a given wind power plant, the method is considerably easier to apply, since the long-term wind speeds measured there can produce the characteristics needed to operate both models.

Acknowledgements: The authors would like to express their thanks to the Hungarian Meteorological Service for providing data for the analysis.

References

Aggarwal, S.K. and *Meenu Gupta*, 2013: Wind Power Forecasting: A Review of Statistical Models. *Int. J. Energy Sci.3*(1).

Armstrong, J.S. and *Collopy, F.*, 1992: Error measures for generalizing about forecasting methods: Empirical comparisons. *Int. J. Forecast 8*, 69–80.

Bremnes, J.B., Villanger, F., and *AS, K.V.,* 2002: Probabilistic forecasts for daily power production. Proceedings of the Global Wind Power Conference, Paris.

Dévényi, D. and *Gulyás, O.,* 1988: Matematika statisztikai módszerek a meteorológiában. Tankönyvkiadó, Budapest. (in Hungarian)

Dobosi, Z. and *Felméry*, *L.,* 1971: Klimatológia. Egyetemi jegyzet, Tankönyvkiadó, Budapest. (in Hungarian)

Kavasseri, R.G. and *Seetharaman, K.*, 2009: Day-ahead wind speed forecasting using f-ARIMA models. *Renew. Energy 34,* 1388–1393. *https://doi.org/10.1016/j.renene.2008.09.006*

Károssy, Cs., 1993: A Péczely-féle makroszinoptikus tipizálás és a helyzetek katalógusa (1951–1992). In (Ed.: *Nowinszky L.*) A fénycsapdás rovargyűjtést módosító abiotikus tényezők. I. kötet, OSKAR Kiadó, Szombathely,113–126. (in Hungarian)

Károssy, Cs., 1998: Péczely's classification of macrosynoptic types and catalogue of weather situations (1992–1997). In (Ed.: *Nowinszky L.*) Light trapping of insects influenced by abiotic factors. Part II, Savaria University Press, 117–130.

Károssy, Cs., 2001: Characterisation and catalogue of the Péczely's macrosynoptic weather types (1996–2000). In (Ed.: *Nowinszky L.*) Light trapping of insects influenced by abiotic factors. Part III. Savaria University Press, 75–86.

Matyasovszky I., 2002: Statisztikus klimatológia. Idősorok elemzése. ELTE Eötvös Kiadó, Budapest. (in Hungarian)

Mezősi, M. and *Simon, A.,* 1981: A meteorológiai szélmérés elmélete és gyakorlata. Meteorológiai Tanulmányok 36. (in Hungarian)

Olaofe, Z.O., Folly, K.A., 2012: Statistical Analysis of the Wind Resources at Darling for Energy Production. *Int. J. Renew. Energ. Res. 2*, 250–261.

Péczely, Gy., 1961: Magyarország makroszinoptikus helyzeteinek éghajlati jellemzése. Az Országos Meteorológiai Intézet Kisebb Kiadványai 32. (in Hungarian)

Shukur, O.B. and *Lee, M.H.*, 2015*:* Daily wind speed forecasting through hybrid KF-ANN model based on ARIMA. *Renew. Energy 76,* 637–647. *https://doi.org/10.1016/j.renene.2014.11.084*

Tar, K., 1990: Statistical Investigation on the 130-year Time Series of precipitation in Debrecen. Climatic Change in the Historical and Instrumental Periods. Masaryk University-Brno, 275–279.

Tar, K., 1993: Investigation of the Time Series of the Monthly Relative Sums of Precipitation. Early Meteorological Instrumental Records in Europe. Uniwersytet Jagiellonski, Kraków, 183–191.

Tar, K., 1995a: A havi relatív csapadékösszegek idősorának tulajdonságai Magyarországon. Berényi Dénes professzor születésének 95. évfordulója tiszteletére tudományos emlékülés előadásai. KLTE Debrecen, 158–165. (in Hungarian)

Tar, K., 1995b: Investigation of the time series of the monthly relative sums of precipitation in Hungary. Proceeding of Conference in Atmospheric Physics and Dynamics in the Analysis and Prognosis of Precipitation Fields, Rome, 405–410.

Tar, K. and *Kircsi, A*., 2001: Kísérlet a szélenergia statisztikai becslésére. Szélenergia konferencia, Gödöllő, 28–34. (in Hungarian)

Tar, K., Kircsi, A., and *Szegedi, S.*, 2001: A possible statistical estimation of wind energy. Proceedings of the European Wind Energy Conference, Copenhagen, Denmark, 886–889.

Tar, K., Kircsi, A., and *Vágvölgyi, S*., 2002: Temporal changes of wind energy in connection with the climatic change. Proceedings of the Global Windpower Conference and Exhibition, Paris, France, 2-5 April, CD-ROM.

Tar, K., 2004: Becslési módszerek a magyarországi szélenergia potenciál meghatározására. *Magyar Energetika 12*(4), 37–48. (in Hungarian)

Tar, K., 2007: Diurnal course of potential wind power with respect to the synoptic situation. *Időjárás 111*, 261–279.

Tar, K., *Maghiar, T., Bondor, K.* and *Szegedi, S*., 2007: Statistical estimation of diurnal average potential windpower. Proceedings of the 9th International Conference on Engineering of Modern Electric System, Oradea, Romania, 86–90.

Tar, K., 2008a: Az időjárási helyzetek szélenergiájáról. Tanulmánykötet Dr. Göőz Lajos professzor 80. születésnapjára. Nyíregyháza, 267–276. (in Hungarian)

Tar, K. 2008b: Energetic characterization of near surface windfield in Hungary. *Renew. Sustain. Energ. Rev. 12*, 250–264.

Tar, K. and *Szegedi, S*., 2011: A statistical model for estimating electricity produced by wind energy. *Renew. Energ. 36,* 823–828. DOI: 10.1016/j.renene.2010.06.032.

Tar, K. and *Lázár, I.*, 2018: Statistical structure of day by day alteration of daily average wind speeds. *Időjárás 122*, 285–304.

Tar, K., 2019: Statisztikai módszer a napi átlagos szélsebesség napközbeni becslésére. Tiszteletkötet Puskás János 65. születésnapjára. Szombathely, 115–127. (in Hungarian)

Tar, K., 2021: Az átlagos szélsebesség napi változásának statisztikai becslése. *Légkör 66*(1), 27–32. (in Hungarian)

DOI:10.28974/idojaras.2022.4.4

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 126, No. 4, October – December, 2022, pp. 511–543

Features of climatic temperature over Saudi Arabia: A Review

Hosny M. Hasanean* and **Abdulhaleem H. Labban**

Department of Meteorology, King Abdulaziz University,
P.O. Box 80234, Jeddah, Saudi Arabia

**Corresponding Author e-mail: hhasanean@kau.edu.sa*

(Manuscript received in final form January 24, 2022)

Abstract—The climate around the world including Saudi Arabia has been fluctuating from cold to warm during different periods. The climate of the earlier period of the 650 ka BP was warmer than the present time in Saudi Arabia due to greenhouse gases in the atmosphere. The current climate of Saudi Arabia is arid to semi-arid with different climate classes. The seasonal surface air temperatures (SATs) are high in the central and northern regions compared to the southern region. The summer of Saudi Arabia is the warmest around the globe with the exception of the coastal region. Due to different air masses that invade the regions of Saudi Arabia, there are different SATs in different seasons. Depending upon seasonal and annual basis, the frequency of the extreme cold SAT is less than the extreme warm SAT.

The circulation pattern of high and low pressures plays an important role in the climatic SAT of Saudi Arabia. The coldest year is associated with the Siberian high-pressure during winter and early spring, especially in the central and northern areas, while the warmest year is related to the Indian monsoon low-pressure during summer and early autumn especially on the northeastern parts, majority of the east coast, and central regions of Saudi Arabia. On the other hand, the Icelandic low pressure extended to the southern region causes cooling air over the area, especially, the northern part of Saudi Arabia, while the Sudan low-pressure causes warming and moisture from the southern and southwestern regions in the winter season. The synoptic situation in the spring season is almost similar to the autumn season. During the spring and autumn seasons, the synoptic circulation over Saudi Arabia is Siberian high-pressure from the east, subtropical high-pressure from the west, Mediterranean depression from the north, and Sudan low and/or Asian monsoon low from south.

Key-words: surface air temperature, Saudi Arabia, climate past and present, synoptic circulation, Siberian high pressure, subtropical high pressure, Icelandic low, Sudan monsoon low

1. Introduction

Saudi Arabia is characterized by complex topographical features and occupies nearly eighty percent of the Arabian Peninsula, which is approximately 2,250,000 km^2. Saudi Arabia extends from 15.5°N to 32.5°N in latitude and from 32°E to 55°E in longitude (*Fig. 1*). The country is distinguished by different climatic regions because of high spatial and temporal SATs variability. Moreover, the distribution of the observed SATs dataset is also distinguished by high inter-annual variability (*Almazroui et al.,* 2009).

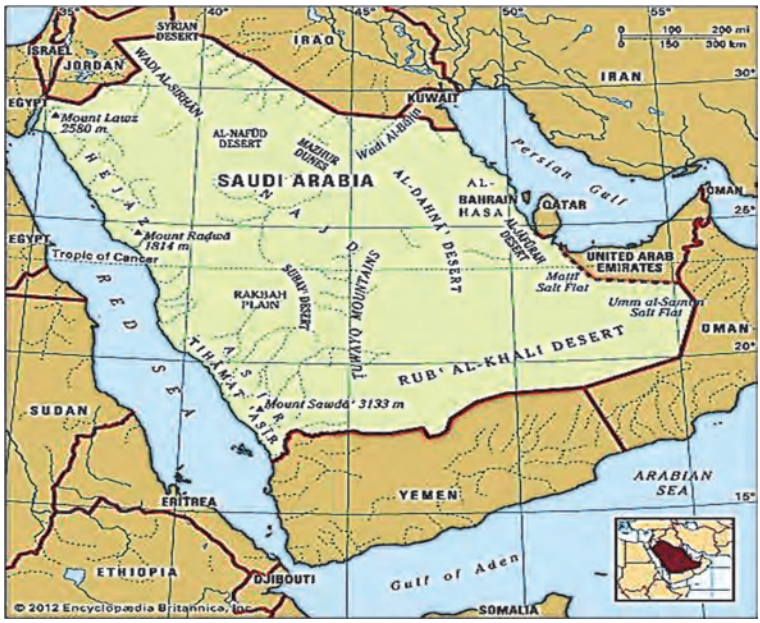


Fig. 1. Location of Saudi Arabia

Surface air temperature over Saudi Arabia in the past

Between 870 Ma and at the end of the Precambrian time (roughly around ~542 Ma), the structure of the Arabian-Nubian Shield and the East African Orogen developed (*Stern et al.,* 2006; *Kotwicki et al.,* 2007*; Kotwicki* and *Al Sulaimani,* 2009), and the main structural and tectonic events shaped the area in the late Cretaceous. The Arabian Peninsula, as the Arabian Plate, stayed in continuous contact with Africa up to the formation of the Red Sea rift in the late Tertiary era. More than 2 Ma ago, the great Al-Rub Al-Khali desert began to form, with its climate oscillating from hyper-arid to temperate.

From the time series of atmospheric trace gases and aerosols for the period from about 650 ka BP to the present (*Jansen et al.,* 2007), one can find that earlier periods were warmer than the present. The world climate was cold and dry by 150,000 BCE and started to increase by 140,000 BCE (*Edmiston,* 2007). By 130,000 BCE, the climate was almost analogous to the current era of the world, while it remained cold and dry during 120,000 BCE. At the end of the initial cooling (110,000 BCE), temperature changed dramatically. However, the climate remained stable for thousands of years. The climate became slightly warmer around sixty thousand years ago followed by a long phase of climate oscillation. Around thirty thousand years ago, the climate returned to cooling, which is termed as late-glacial cold stage, and extended to the successive twenty-one thousand and seventeen thousand years. At the beginning of the Holocene phase (9500 BCE), the cold period ended abruptly, while warming continued to prevail gradually. Similarly, there were cold and dry periods for the next thousand years, although significant areas of the world were under high moisture and warmer conditions. Therefore, the Arabian and Saharan deserts disappeared with most of their area covering vegetation. During the Holocene Climate Optimum (HCO), which is the period between nine thousand and five thousand BCE, the Sahara regions were greatly influenced by increased rainfall. However, Eastern Sahara began to return desert during 5300 BCE due to decrease in rainfall.

McClure (2007) established the presence of the human population in the Empty Quarter with the existence of hand axes that have been found from more than 100 ka on the fringes of the region. However, the origin of the human population in the region has not been firmly established yet. *Majewski et al.,* (2004) categorized six periods of significant rapid climate change, that is, 9,000–8,000, 6,000–5,000, 4,200–3,800, 3,500–500, 1,200–1,000, and 600–150 BP. Changes in atmospheric circulation, tropical aridity, and polar cooling are the most frequent occurrences of climate change events in these global records, even though polar cooling together with rise in the moisture in some areas of the tropics happened in the most recent period from 600 to 150 BP. Numerous eras correspond with the main disturbances of civilization, showing the significance of Holocene climate changes on humans. Across the Pleistocene/Holocene boundary, the Arabian Peninsula climate changed from temperate to arid. *Edmiston* (2007) estimated that a 200-year cooling period during 8,000 BP had reduced the amount of rainfall by 30 percent in the Arabian Peninsula.

Kotwicki and *Al Sulaimani* (2009) found four successive ice ages, alternating with hot periods in the Pleistocene due to climate change. Throughout the greatest recent glaciations, most of the earth's water insulated in the ice sheets of North America and Asia, and the sea level declined by approximately 125 m, uncovering the bottom of the Arabian Gulf (mean depth about 35 m). The Tigris-Euphrates glacial era stream extended from Shatt-al-Arab to the Strait of Hormuz into the Arabian Sea.

An overview: present climates of Saudi Arabia

In Saudi Arabia, the climate varies from place to place due to the changes in dynamics of serious weather and climatic elements. Climatic elements are liable for provincial dispersal because of the influences of radiation, atmospheric circulation cells, different water and landmasses, latitude, winds, altitude, and air masses. Climate controls in combinations and interactions modify the short-term averages of temperature and precipitation giving special climate zones (*Masatoshi and Urushibara,* 1981). The leading elements controlling SAT over Saudi Arabia are latitude and height. SATs are highly seasonal dependent, with being relatively cool in the period from December to February and warm in the period from June to September. However, the annual variability in temperature is very low in contrast to variability in annual rainfall amount.

Most of Saudi Arabia lies in the tropics (15.5°N - 32.5°N; 32°E - 55°E). It is a huge area that lies between the vast continental landmasses of Africa and Asia. The vast area of the country (80% of the Arabian Peninsula) and its complex topography encompass significant climate variations in space and time. Saudi Arabia is one of the warmest countries in the world except for its coastal regions during the summer season. Saudi Arabia is among the few countries in the globe, where SAT during the summer season exceeds 50 °C. According to *Krishna* (2014), the SAT is in between 27 °C to 43 °C during the summer season over the country except for the coastal regions, where it lies between 27 °C to 38 °C. SAT over Saudi Arabia increases gradually from north to south (*Almazroui*, 2019). Moreover, highlands play an important role in the local climate in the country.

The Saudi Arabian climate is known for its deserts except for the mountainous areas in the southwest. It is characterized by warming in the daytime and sudden cooling in the nighttime. Warm SATs take place in the north of the Tropic of Cancer as well as in the ultimate northern and interior regions of Saudi Arabia. However, the mountains in the southwestern region experience temperate temperatures. Away from the topographical effect, the climate of Saudi Arabia is also influenced by the latitude, tropical winds, vicinity to the sea, elevations, etc.

The two of the most famous world climate classification have been drawn by *Köppen* and *Geiger* (1928), *Koppen* (1936), and *Troll* and *Paffen* (1980). The world map of the Köppen-Geiger climate classification has been updated by *Peel et al.* (2007). However, the Arabian climates have been classified by both authors on a broad scale without consideration of regional differences. One reason associated with the broad-scale climate classification of the Arabian Peninsula is the limited number of weather stations operated at that time. According to the *Köppen* and *Geiger* (1928) classification, the climate of the whole Arabian Peninsula is 'desert climate' except for the mountainous regions in the southwest of Saudi Arabia. Also, they classified the climate of Saudi Arabia as BWh climate; a dry desert climate with an average SAT above 18°C, whereas *Köppen* (1936) used different climate classes for the Arabian Peninsula. *Walter* and *Lieth* (1967),

Muller (1982), and *Moore* (1986) provided brief and sound climatological information on the Arabian Peninsula. They showed a wide spectrum of climatic change in the Arabian Peninsula, from the ice in the Asir mountain to the great humidity of the Arabian Gulf, and from the rising heat of Al-Rub Al Khali to the monsoon rains in the Qari Province in Dhofar.

The climate of Saudi Arabia is described as semi-arid and arid with spatiotemporal variations of temperature and rainfall noticed across the region (*Almazroui et al.,* 2012a, *Almazroui,* 2013). *Almazroui et al.* (2012a) and *Almazroui et al.* (2012b) demonstrated that air temperature is low in the northwestern and southwestern highlands regions of the country and high in the central area towards the Arabian Gulf and western coastal areas along the Red Sea. Additionally, the seasonal mean temperature variability is high in the northern and central regions compared to the south of the country (*Almazroui et al.,* 2012a). Except for the province of Asir on the western coast, Saudi Arabia has a desert climate characterized by extreme heat during the day, an abrupt drop in temperature at night, and very low annual rainfall. Because of the influence of a subtropical high-pressure system, there is significant variation in temperature and humidity in the country. The two major distinctions in the climate of Saudi Arabia can be felt between the coastal part and the interior regions.

Almazroui et al. (2014) used the principal component and correlation methods to regionalize the climate of Saudi Arabia into several homogenous groups. They selected twenty-seven stations across the country for regionalization of climate during a period of twenty-six years (1985 to 2010). They identified five groups: group "A" for the northern region, group "B" for the Red Sea coast, group "C" for the interior region, group "D" for the highlands, and group "E" for the southern region. Moreover, the interannual variability and the degree of seasonality are the same for each climatic group.

2. Air masses that affect the surface air temperature over Saudi Arabia

The movement and transformation of air masses, and interactions between air masses along the fronts give the regional climate of Saudi Arabia. So, the climate of the country can be clarified in terms of regional air masses. According to *Abdullah* and *Al-Mazroui* (1998), four different types of air mass (*Fig. 2*) influence the weather elements in Saudi Arabia. *Fig. 2* reveals the source and direction of these air masses.

1. The polar continental air mass affects the region during the winter season from December to February, sometimes to mid-March. The domination of the Siberian high in central Asia is responsible for this effect. Polar continental air masses are cold (-35 °C to -20 °C), very stable with strong temperature inversions, and dry. They extend through the Siberian high

pressure and dominate the weather of the entire continent. Generally, polar continental air masses generate cold air and clear sky weather.

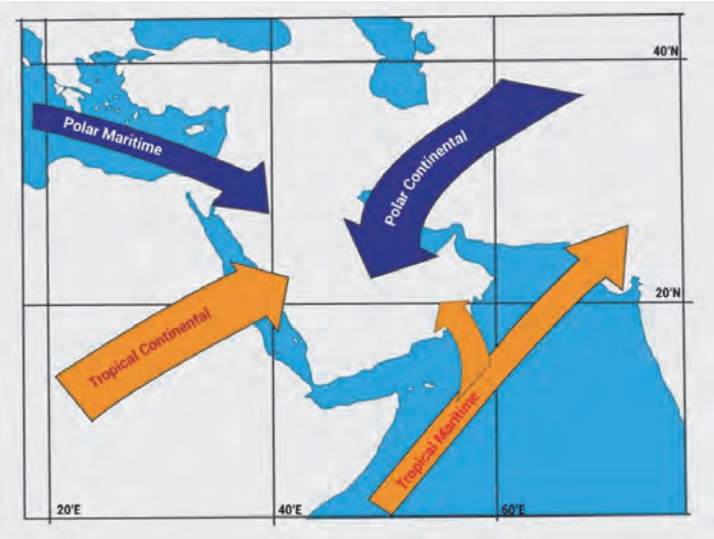


Fig. 2. Air masses influencing the weather elements in the Saudi Arabian.

2. Occasionally, Saudi Arabia is influenced by polar maritime air masses coming from the North Atlantic. It reaches the area because of mid-latitude depressions that move from the west during the winter season. Polar maritime air masses are cool (0 °C to 10 °C), basically conditionally stable, and moist.
3. The tropical continental air mass affects Saudi Arabia in the spring and early summer seasons. This air mass allows the region to become a stable high-pressure zone and a source of tropical continental air. This air mass is hot and dry (>45 °C) and can create strong heating on the surface and dusty weather system (*Fisher* and *Membery,* 1998). In general, in southwestern Asia, summer is very hot, the temperature ranges from 30 °C to 42 °C. A tropical continental air mass is conditionally stable, dusty, and holds moisture. Tropical continental air mass flows north and northwest during the summer season in western Asia. When the Indian monsoon system moves toward the region of the Arabian Peninsula, certain activities occur in summer, particularly in southwestern Saudi Arabia. On the other hand, its influence is limited by the strong tropical continental air mass, which prevails over Saudi Arabia at the time. Overall, the southwestern region has

a lower temperature because of elevation, whereas the other areas are influenced by the tropical air mass during the spring, summer, and autumn seasons (*Tarawneh* and *Chowdhury*, 2018).

4. The tropical maritime air mass invades from the north of the Indian Ocean and the Arabian Sea in the summer season. It impacts principally the southwestern region of Saudi Arabia and produces hot and humid air (*Fisher* and *Membery*, 1998; *Abdullah* and *Al-Mazroui,* 1998). Initially, tropical maritime air masses are hot, conditionally stable, and very moist.

3. Climatology of the surface air temperature in different seasons

The SAT over Saudi Arabia is influenced by many factors including mountain barriers, altitude above sea level, ocean currents, solar radiation, and pressure cells. A cooling trend from south to north in the region of Saudi Arabia is found due to the impact of the cold continental air masses during the winter season. The SAT is low (<15 °C) over the northern, moderate (15–20 °C) over the central and western, and high (up to 20–25 °C) over the southeastern parts of Saudi Arabia during the winter season. Therefore, the climate of the region during the winter season has three different climate zones. Due to the impact of the cyclones in the Mediterranean region, northern Saudi Arabia is cooler than the other regions during the winter season, while southern Saudi Arabia is much warmer due to less precipitation. Moreover, the SAT in the southwestern regions, which are associated with the local topography, is lesser than those in the southeastern regions of Saudi Arabia (*Langodan et al.,* 2014; *Viswanadhapalli et al.,* 2017; *Attada et al.,* 2019). In the winter season, the SAT rarely goes below 0 °C, however, the virtual absence of moisture and the high wind-chill factor create a somewhat cold atmosphere.

During spring, in general, the SAT increases over Saudi Arabia; however, it increases faster in the northern region (between 20 and 25 °C), while the eastern desert region is warmer with SAT between 25 and 27 °C (*Khan* and *Alghafari* 2018; *Attada et al.,* 2019). The spring season is sometimes called the cool moderate season of the post-winter and the summer season (*Khan* and *Alghafari,* 2018).

In the summer season, the SAT becomes intense quickly after sunrise up to sunset, then suddenly becomes cool at night. In early summer (May and June), the distribution of SAT is steadier across Saudi Arabia. It is observed above 35 °C in the eastern region and around 25 °C in the southern part of Saudi Arabia, see *Attada et al.* (2019). The average SAT in the summer season is around 45°C, except for an unusual increase of around 54 °C especially over desert areas due to strong descent motion (*Babu* 2016; *Hasanean* and *Almazroui,* 2017). According to the study of *Rodwell* and *Hoskins* (1996, 2001), this descent motion is related to the mechanism of the desert-monsoon over the region and stronger summer

insolation. Additionally, *El Kenawy et al.* (2014, 2016) exhibited heating during the summer season with thermal lows.

In the autumn season, SATs over Saudi Arabia is between 20 to 25 °C. Moreover, SAT remains below 25 °C in the western and northwestern regions of Saudi Arabia, while the highest values of SATs are found in the eastern regions of (*Attada et al.,* 2019). In the spring and autumn, the SAT is moderate with mean temperatures approximately 29 °C.

The Climate Research Unit (CRU) is a Global Climate Dataset, available through the IPCC DDC, consists of a multi-variate 0.5º latitude by 0.5º longitude resolution mean monthly climatology for global land areas, excluding Antarctica. The CRU TS3.1 dataset derived from the gridded climate datasets, interpolated using the original information from the stations since 1901. In the interpolation records, the anomaly with respect to the average for the period 1961–1990 is calculated. The anomaly is interpolated using thin-plate splines as a function of latitude/longitude. The CRU dataset gridded values are obtained by applying a smooth fitting (in 3D space) to the available surface station observations (*New et al*., 2000).

Almazroui et al., (2012b) studied the spatial distribution of the mean average temperature for the period 1979–2009 for the wet season (from November to April) and dry season (from May to October) by using ground-based meteorological data measured at the meteorological stations located in Saudi Arabia and Climate Research Unit (CRU) data.

They found that the mean temperature is low (below 15 °C) in the northern part of Saudi Arabia for the wet season. In addition, low temperatures in the southwestern coastal region of Saudi Arabia are also found. However, the highest temperature (> 24 °C) is found over the west coast region. On the other hand, during the dry season, mean temperature over Saudi Arabia is high (*Almazroui et al.,* 2012b). The highest temperature ranges from 33–36 °C from inland of the Arabian Gulf to the center of the country. Low-temperature regions appear over the northwestern and southwestern zones. However, the lowest temperature (>24 °C) is observed over the southwestern regions of Saudi Arabia.

The distribution pattern of the wet season averaged maximum temperature for the period 1979–2009 is similar to the mean temperature pattern (*Almazroui et al.,* 2012b). The highest maximum temperature over Saudi Arabia is observed along the Red Sea coast (> 30 °C). The mean maximum temperature over Al-Rub Al-Khali is between 27 °C and 30 °C. However, during the dry season, the highest maximum temperature (>42 °C) is found in the northeast region of the country. In the dry season, the maximum heat (temperature < 42 °C) is observed over northern Saudi Arabia. The minimum heat (temperature > 27 °C) is found over southwestern Yemen. At most stations of Saudi Arabia, high temperatures are also observed.

Almazroui et al., (2012b) studied average minimum temperature of the wet and dry seasons using the CRU and stations dataset for the study period (1979–2009). The minimum temperature (> 9 °C) is investigated over each of the northwestern and southwestern areas of Saudi Arabia, Iran, Iraq, Jordan, and Egypt. The minimum temperature over the west coast and Al-Rub Al-Khali of Saudi Arabia is around 18 °C and 24 °C, respectively. The results arising from the spatial distribution of the minimum temperature are supported by the results arising from the CRU dataset (*Almazroui et al.,* 2012b). The highest value of the average minimum temperature for the dry season (> 33 °C) is found in a narrow part of the eastern part of Saudi Arabia, while the lowest value (<18 °C) is evident over the southwestern part of Saudi Arabia extending to the western region of Yemen. For the minimum temperature during the dry season, the lower temperature is investigated over the northern parts of Saudi Arabia. The observed datasets confirm the pattern distribution of the Climate Research Unit datasets. In comparison to the wet season, the region of high temperature is transmitted to the north.

4. Annual surface air temperature climatology

Temperature is an important element for any vulnerability evaluation in a climate change, and its analysis depends on the variations in its mean, maximum, and minimum values. Saudi Arabia is a warm country with a mean annual temperature of exceedingly more than 30 °C in some areas. *Almazroui et al.* (2012a) and *Attada et al.* (2019), studied the climatological behavior of the mean, minimum, and maximum annual temperatures over Saudi Arabia. They used the gridded dataset from CRU and the observed datasets retrieved from various ground stations in different parts of Saudi Arabia. From the gridded CRU dataset for the period from 1979 to 2009, the highest mean annual temperature of the SAT is noticed over the Al-Rub Al-Khali (around 27–30 °C), and southwest of Saudi Arabia mainly along the Red Sea coast (*Almazroui et al.,* 2012a). Moreover, the mean annual SAT over most regions of Saudi Arabia is between 24–30 °C. Whereas, in the northern part of Saudi Arabia, the mean annual SAT is lower than 21 °C. The study of *Almazroui et al.,* (2012a) showed that the mean annual temperature over most areas in Saudi Arabia is around 20 °C.

Almazroui et al., (2012a) used CRU gridded dataset as well as the ground-based observational dataset to study the annual mean minimum temperature pattern from the period 1979 to 2009 over Saudi Arabia. The low value of temperature from 6 to 15 °C is observed over the northwestern and southwestern regions of Saudi Arabia. However, a relatively higher temperature of 21–24 °C over the southeastern regions, mainly over Al-Rub Al-Khali is observed. The south-eastern regions are the highest temperature zones in the country. It is also noted that the results of the ground-based observational data are like those of the gridded CRU data.

The annual mean maximum temperature pattern for Saudi Arabia averaged over the period from 1979 to 2009 is studied by *Almazroui et al.* (2012a) using both the CRU and ground-based observational datasets. The maximum value of temperature peaks observed in the middle of the country is due to the influence of the land-sea contrast. Moreover, the highest temperature (33–36 °C) over the southwestern region of Saudi Arabia is also observed. On the other hand, a slightly lower temperature is found along the Red Sea coast. A moderate temperature between 27 °C and 33 °C is found in most of Saudi Arabia. Once again, the results of the ground-based observational data are similar to those of the gridded CRU data. *Almazroui et al.* (2012a) concluded that the gridded CRU datasets of the annual mean, minimum, and maximum temperature over Saudi Arabia fairly represent the climatological features of Saudi Arabia, which is similar to the observed dataset.

5. Climate extremes of temperature over Saudi Arabia

Climatic extreme events may have main influences on the economy, ecosystems, society, and human health. Climatic extreme events drive natural systems compared to average climate (*Parmesan et al.,* 2000). At the global scale, the anthropogenic influences may have led to warming of extreme daily minimum and maximum temperatures and contributed to the intensification of extreme precipitation (*Handmer et al.,* 2012). Moreover, *Peterson et al.* (2012) illustrated how anthropogenic climatic change is changing the probability of occurrence of extreme events. Extreme cold and hot temperature can have huge impact on society. Seasonal changes in temperature both extreme cold and/or hot can have serious negative impacts on farming, tourism, etc. Prominently, what is 'normal for one area may be extreme for another area that is less well adapted to such temperatures.

Over the northwestern area of Saudi Arabia, the extreme temperature is in overall agreement with the prognostic warming over the Mediterranean region (*UNFCCC,* 2011). During the summer of 2010, extreme warming influenced the Saudi Arabia region, with outstanding readings of 52.0 °C at Jeddah. In the second report of the UNFCCC, one can conclude that the anomalies of the annual mean temperature of the coolest year (1992) are found at all stations. All anomalies are negative ranging from -1.8 ° to -0.6 °C, in which extreme cooling occurred over the central and northern regions of Saudi Arabia in association with a deep invasion of the Siberian high pressure in winter and early spring. On the other hand, anomalies of the annual mean temperature of the warmest year (1999) are found at all stations. All anomalies are positive ranging from 1.5 ºC to 0.2 ºC, in which the extreme warming occurred over the central part and most of the eastern coasts and northeastern regions of Saudi Arabia in association with an extension of the Indian monsoon in the summer and early autumn.

Extreme heat and cold waves arise from any changes/variations in climate (*Kotwicki* and *Al Sulaimani,* 2009). Any variations and/or changes in climate cause variations and/or changes in extreme weather events, such as heat and cold waves (*Kotwicki* and *Al Sulaimani,* 2009; *Almazroui et al.,* 2012b). For example, the climatic extreme event in 2010 was considered as the warmest year of the observational period over Saudi Arabia (*Almazroui et al.,* 2012a). On June 2010, the temperature over Saudi Arabia was recorded at 52 °C. Based on the observed daily temperature datasets from 19 stations for the thirty-year period from 1979 to 2008, *Athar* (2014) studied the extreme climate over Saudi Arabia using climatic indices. According to his findings, the frequency of the extreme cold temperatures is less than the frequency of the extremely warm temperatures on seasonal and annual basis. In addition, the daily extreme temperature is warmer during the summer season. Moreover, the southwest coastal region has displayed more warming than the inland region. Also, during the thirty-year-long period, the statistically significant increase of the daily regional temperature is 0.21 °C per decade. Lastly, warming in general is noticed by him over Saudi Arabia, especially in maximum temperatures. The results of *Athar* (2014) agreed with the results of *AlSarmi* and *Washington* (2014), who studied the changes in temperature extremes over the Arabian Peninsula based on daily data over 23 stations covering six countries (Saudi Arabia, Oman, Bahrain, Qatar, Kuwait, and UAE) at different time periods. *AlSarmi* and *Washington* (2014) found a significant increase in very warm nights during the period from 1986 to 2008. Moreover, they noticed in general, that extreme temperatures over the northern Arabian Peninsula in the daytime are high, while for the nighttime, extreme temperatures are observed over the southern region, especially during the last period of the study. From the spring to autumn, strong warming above 5 days per decade is found (*Islam et al.,* 2015). *Almazroui* (2020) studied the extremes in the temperature over Saudi Arabia for the period 1978–2019 and found an increase in the average temperature extremes over the second period (2000–2019) as compared to the first period (1980–1999). The results show a considerable increase in the number of warm days/nights in the second period compared to the first period. The results also exhibit a considerable decrease in cold days/nights. Moreover, in the summer season, an increase in warm days/nights is observed, while in the winter season, a decrease in the number of cold days/nights is found.

6. Circulation weather types and their influence on the surface air temperature

6.1. Overview

The atmospheric circulation such as low and high pressure plays an important role in determining weather and climate. Centers of high or low pressure over Asia generally immigrate from east to west, and these centers change radically from

wintertime to summertime. The various systems that control pressure over Saudi Arabia during different seasons are the Siberian high over central Asia, the Indian low and monsoon Asiatic (*Hastenrath,* 1985), and the Mediterranean secondary low pressure (*Kendrew,* 2012). In wintertime, a thermal high pressure called Siberian high pressure is developed over Mongolia, which controls the eastern and southern climates of Asia. This thermal high pressure is shallow, cold, and dry which extends to southwest Asia through southern and eastern regions of Asia. The Siberian high pressure exists from November to March, and it reaches the maximum value in February over Asia. February marks the peak of the Siberian high's overriding of the winter circulation over Asia, while a similar pattern, depicted on the January map, exists from November to March. February indicates the maximum of the Siberian high's control of the circulation in the wintertime over Asia, though the same distribution pattern has been observed in January, and from November to March. This is related to the advection of dry, cool, and subsiding polar continental air mass. Moreover, in April, the Siberian high pressure weakens and shifts to central Asia, then it disperses in May, when intense thermal low dominates over the tip of the Arabian Peninsula (*Davydova et al.,* 1966).

From March to May (spring season) and from September to November, the Saudi Arabian (autumn season) tropical disturbances (may intensify into tropical storms, but it is rare to grow into tropical cyclones) are invaded in the Arabian Peninsula. Tropical cyclones occur in the period from October to December, but their impact is limited to the southern part of the Arabian Peninsula (*Taha et al.,* 1981). Also, very dry and hot continental tropical air from the Sahara arrives in the area during early autumn and late spring.

From June to August (summer season), the Asian monsoon (Indian monsoon) low pressure dominates over the Arabian Peninsula (*Hastenrath*, 1985). Moreover, the Azore's high pressure (subtropical high pressure) affects the western rim of the Arabian Peninsula. Very hot and dry air masses arrive irregularly in the region from northern India after the tropical maritime air loses its moisture. Moreover, the neighboring water surface of the Gulf Sea, the Gulf of Oman, and the Arabian Sea contributed as the sources of moisture (*Taha et al.,* 1981). Additionally, in summer, the intertropical front forms from the meeting of converging air masses over the Arabian Peninsula region. Moreover, the water surface near the Arabian Sea and the Gulf of Oman works as a source of moisture and heat (*Taha et al.,* 1981). The main characteristics to be considered in the climate of Saudi Arabia are the topographic and the interior regions, especially in the western region along the Red Sea (*Al-Jerash* 1985; *Ahmed* 1997; *Subyani et al.,* 2010; *Almazroui et al.,* 2012a).

6.2. Synoptic circulation in the winter season

In the winter season, the synoptic circulation systems that affect SAT of Saudi Arabia are the Siberian high pressure extended from the east, subtropical high pressure extended from the west, Sudan low pressure extended from the south, and Mediterranean depressions and/or Icelandic low extended from the north (*Fig. 3*). These synoptic circulation systems generate weather activity over the Saudi Arabian region, which is summarized below.

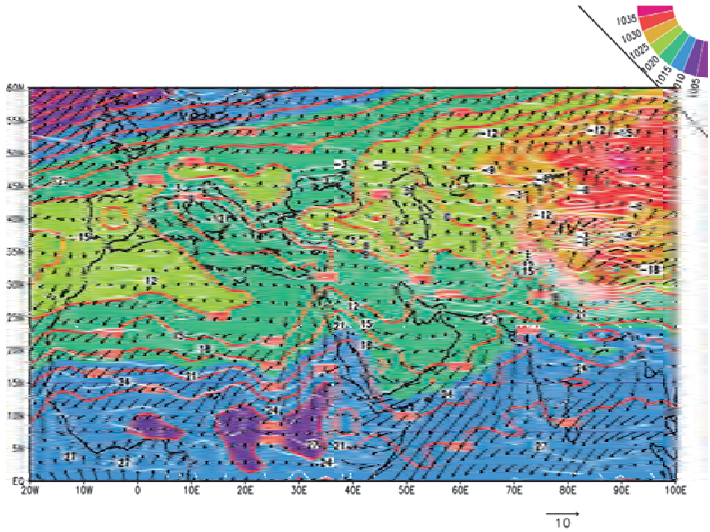


Fig. 3. The distribution pattern of sea level pressure (solid red curve), surface air temperature (small dash curve), and vector wind in the winter season.

A) The Mediterranean depressions on the surface are in relationship with upper troughs coming from the west to east, the active subtropical jet, and the polar jet making precipitation during their transit over Saudi Arabia region. Their potential of producing weather activity decreases generally from north to south over the area except for the mountainous areas, where the uplift acts as an exterior factor. Consequently, the distribution of winter weather activity shows maximum movement in the northern part of the main plateau with a gradual decrease in the lowlands in the eastern and western sides. The trough emanates from the Icelandic low extends southward influencing the weather in Saudi Arabia (*Hasanean et al.,* 2015). A positive relationship exists between the Icelandic low and SAT over the northern parts of Saudi Arabia (*Fig. 4*, *Hasanean et al.,* 2015; *Attada et al.,* 2019). Therefore, the SAT decreases (cooling) during the deepening of the Icelandic low (*Fig. 5a*), while it increases (warming) during the weakening of the Icelandic low (*Fig. 5b*).

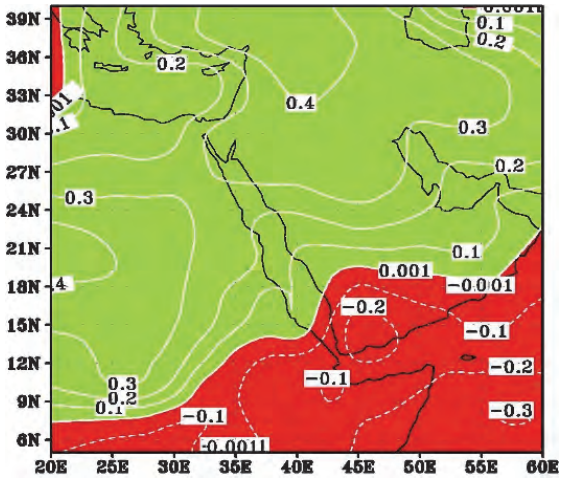


Fig. 4. Horizontal distribution of the correlation coefficient between the SAT and the Icelandic low pressure (*Hasanean et al.*, 2015).

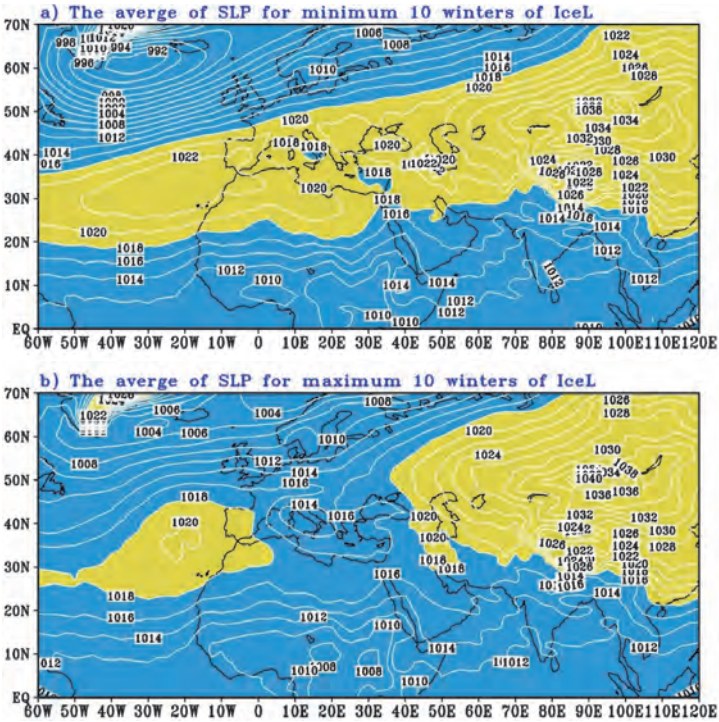


Fig. 5. Composite sea level pressure pattern for the ten winters, when the Icelandic low was (a) deepening and (b) weakening (*Hasanean et al.* 2015).

B) The major pressure system influencing the climate of Saudi Arabia in winter is the Siberian high pressure over Asia, that tends to merge with centers of high pressure over the Arabian Peninsula (Saudi Arabian high) and Sahara Desert (Saharan high). The Siberian high pressure in winter centered over northern Mongolia extends southward to cover Saudi Arabia and the Arabian Gulf carrying cold air to these regions. The Siberian high-pressure influences the climate of mid-high latitudes (*Guo,* 1996; *Zhu et al.,* 1997; *Gong* and *Wang* 1999; *Miyazaki et al.,* 1999; *Yin,* 1999; *Hasanean et al.,* 2013, 2015). A positive relationship between the Siberian high pressure and SAT is found over the northern region of Saudi Arabia extending to the central regions (*Fig. 6*, *Hasanean et al.,* 2013, 2015; *Attada et al.,* 2019). Therefore, the SAT warms in the northern and central regions of Saudi Arabia during the intensification of the Siberian high pressure in the winter season.

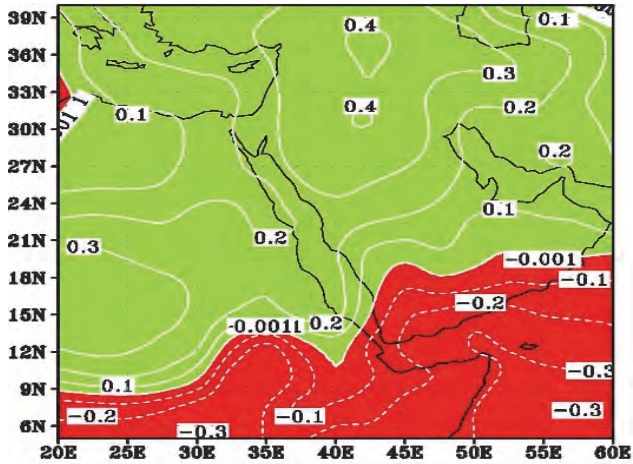


Fig. 6. Horizontal distribution of the correlation coefficient between the SAT and the Siberian high pressure (*Hasanean et al.,* 2015).

Fig. 7a reveals the composite 10 years of sea level pressure difference between the maximum and minimum of the Siberian high pressure. Cooling of SAT over Saudi Arabia, Egypt, the northeast of the Mediterranean, and Sudan (*Fig. 7a*) is found during the minimum of the Siberian high pressure. This cooling of SAT is between 0.5 ºC in the southern region of Saudi Arabia and 2.0 ºC over the northern region of Saudi Arabia and the Turkish area. This is because the Siberian high dominates over the Saudi Arabia region providing very cold and dry air during the deepening of the Siberian high pressure. Two sources of relatively

warm moist air, one from southern Europe and the other from the southern and southeast parts of Saudi Arabia are invaded during the years of weakening Siberian high pressure (*Fig. 7b*). *Fig. 7b* shows the composites of 10 winters of the sea level pressure when the Siberian high was strong. In this case, the Siberian high is separated from the subtropical high pressure, and the pressure over Saudi Arabia and the southern Mediterranean regions becomes weak. Therefore, the distribution pattern of the sea level pressure, in this case, leads to interlinkage among two different air masses, first with the cold moist air mass coming from the Mediterranean depression, which is traveling from west to east, and second with the warm moist air coming from the northward oscillation of the Sudan low. The interlinkage between these pressure systems leads to an increase of the SAT over Saudi Arabia, the east of Africa, southeast Europe, and the east Mediterranean.

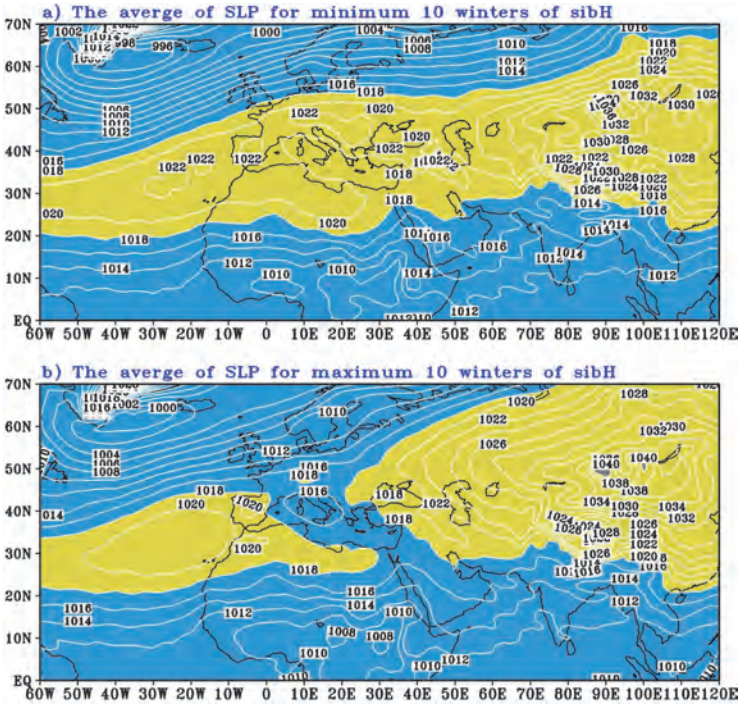


Fig. 7. Composite sea level pressure pattern for the ten winters, when the Siberian high was (a) weakening and (b) strongest (*Hasanean et al.* 2015).

C) The dominant subtropical high pressure (Azores High) is situated around the latitudes of 30° N in the Northern Hemisphere and forms a broad

continuous ridge in the mid-latitudes and the subtropical region. The movement and oscillation of the surface subtropical high are associated with the movement and oscillation of the Siberian high (*Hasanean et al.,* 2015). *Barry and Chorley* (1992) explained that the movement of the subtropical high toward the equator through the winter is due to the increasing temperature differences between the tropics and the poles. Saudi Arabia is influenced by the extension of the subtropical high pressure. *Hasanean et al.* (2015) investigated the relationship between the subtropical high and SAT and found, that in the greatest parts of Saudi Arabia with the exception of the southwest area, an inverse relationship between the subtropical high-pressure and SAT exists (*Fig. 8*). Therefore, intensifying (weakness) in the subtropical high-pressure leads to a decrease (increase) in the SAT. The SAT over Saudi Arabia is warming during the weakness of the subtropical high pressure, where the warm air mass comes from the southern parts of Europe (relatively warm) and the southern/southeastern parts of Saudi Arabia. During the subtropical high pressure intensity, SAT is cooling due to very cold and dry air masses coming from Siberian high pressure (*Hasanean et al.,* 2015).

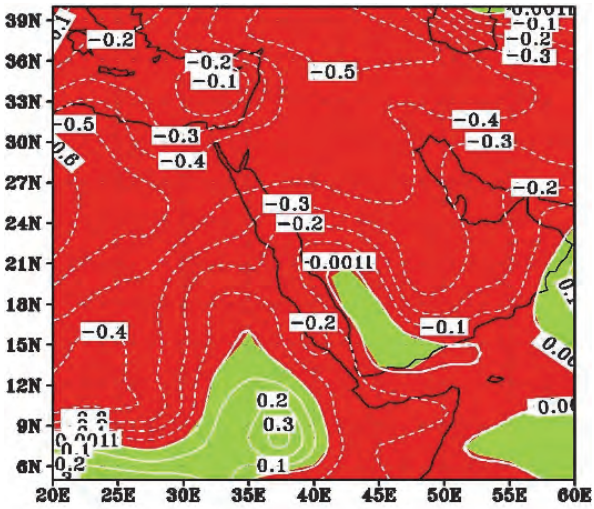


Fig. 8. Horizontal distribution of the correlation coefficient between the SAT and the subtropical high pressure (*Hasanean et al.,* 2015).

D) *Vorhees et al.* (2006) and *Hasanean et al.* (2013) noticed that in winter, the Saudi Arabian high pressure plays an important role in the climate of southwest Asia. *Fig. 9* reveals the typical circulation distribution patterns over the Arabian Peninsula and northeast Africa in winter. The Saudi Arabian high is on the northern, Mediterranean low pressure systems are on the

northwest parts of the Arabian Peninsula. Moreover, when the Saudi Arabian high pressure is prevailing, the air is more stable and relatively warmer within the area. At the same time, the Siberian high pressure prevents the Mediterranean depression, and consequently, the northwest wind (cold air) comes into the area.

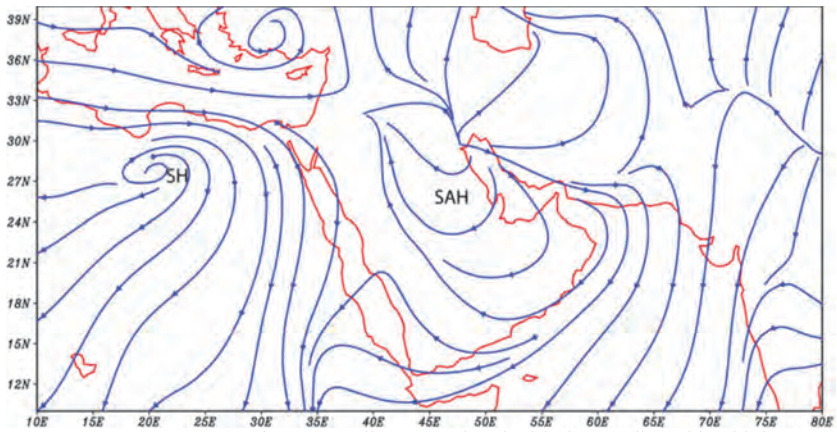


Fig. 9. Long-term mean surface streamlines associated with the Saudi Arabian high (SAH) and the North African high (Sahara high, SH; (*Hasanean et al.,* 2015)).

E) The Sudan low is a dynamic low pressure that carries humid and warm air to Saudi Arabia and the south of the Mediterranean Sea in the cold and rainy seasons (*Rasuly et al.,* 2012). The Red Sea trough (RST) is counted as an expansion of the Sudan low (*El-Fandy,* 1948). So, one can say that the southward and/or northward oscillations of the Red Sea trough arises from the southward and/or northward oscillations of the Sudan low. Usually, the movements of the Sudan low can be classified into two different kinds of oscillations. The first is the movement of its center from near to the Abyssinian Plateau and its return, twice over the year. The second comprises a series of quite small oscillations superimposed on the annual track. These small oscillations are most evident in the two transitional seasons (*El Fandy,* 1940). A strong negative relationship between the Sudan low and SAT over the northern and central areas of Saudi Arabia is found (*Fig. 10*, *Hasanean et al.,* 2015). Therefore, the SAT warms over all regions of Saudi Arabia except for the southern region during deepening (decrease in mean sea level) of the Sudan low, and the SAT colds during the weakening of the Sudan low (increase in mean sea level).

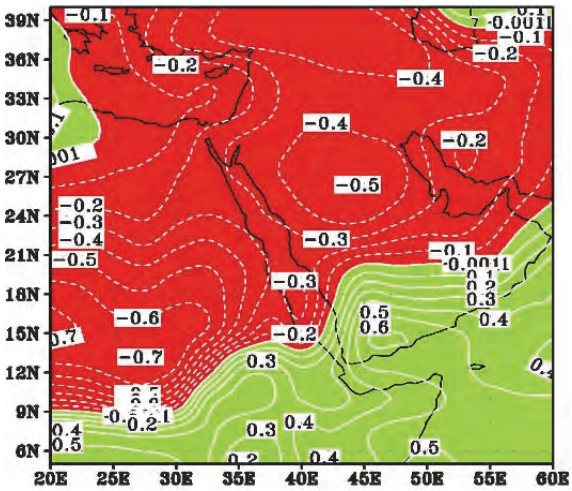


Fig. 10. Horizontal distribution of the correlation coefficient between the SAT and the Sudan low pressure (*Hasanean et al.,* 2015).

According to *Hasanean et al.* (2015), two main distribution patterns of pressure affect the SAT of Saudi Arabia. The first pattern happens when the subtropical high pressure is strengthening, Siberian high is weakening, Icelandic low is deepening, and Sudan low is weakening. In this situation, the subtropical high pressure merges with the Siberian high pressure to prevent the interaction between the extratropical and midlatitude systems that leads to cooling SAT. The second situation occurs when the subtropical high pressure is weakening, Siberian high pressure is strengthening, Icelandic low is weakening, and Sudan low is deepening. This pattern causes interaction between two different air masses, the first is a cold moist air mass that is related to the traveling Mediterranean depression from west to east and the second is a warm moist air mass related to the northward oscillation of the Sudan low pressure and its inverted V-shape trough.

There are exchanges between air masses arising from high and low pressure systems that influence the SAT over Saudi Arabia. For an instant, there is an exchange of air mass between the subtropical high pressure and Icelandic low to build up the North Atlantic oscillation, which is linked to SAT variability (*Rogers* and *van Loon* 1979; *Rogers,* 1984). When the North Atlantic oscillation index becomes negative (subtropical high pressure and Icelandic low are weak), the relatively warm moist air moves to lower latitudes. On the other hand, during positive North Atlantic oscillation index, the stronger clockwise current around the subtropical high center is remarkable, and the temperature over the Middle East is cooling (*Hurrell et al.,* 2003). In winter, the subtropical high-pressure retreats southwestward, and the Red Sea trough oscillates northward over the

eastern Mediterranean area (*Alpert et al.,* 2004), and the Siberian high pressure moves eastward to bring the mild temperature to the Arabian Peninsula. On the other hand, during the weakening of the Siberian high pressure, few depressions moves to the Mediterranean (*Makrogiannis et al.,* 1991; *Sahsamanoglou* and *Makrogiannis* 1992; *Wilby* 1993), which affects Saudi Arabia, especially its northern region.

6.3. Synoptic circulation in the spring season

The weather of Saudi Arabia during the spring season is affected by the Sudan monsoon low, which is entering from the southeast (*Fig. 11*). *Attada et al.,* (2019) investigated the surface air temperature variability over the Arabian Peninsula and its links to circulation patterns. They found a negative (positive) relationship between the SAT and the subtropical high pressure (Siberian high pressure) over the Arabian Peninsula during the spring season (March, April, and May). Therefore, the SAT warms during the strengthening of the subtropical high pressure and the weakening of the Siberian high pressure. They suggested that the variability of temperature in the spring season is related to the strongness/ weakness of the Siberian high pressure. Therefore, when the Siberian high pressure is strong and extends to the Saudi Arabia region, the SAT becomes cold, and vice versa, when the Siberian high pressure is weak. The Arabian Peninsula is influenced by northeasterly winds blowing from the Indian region, proposing that the spring temperature over the Arabian Peninsula is related to the Mediterranean depression and Siberian high pressure (*Attada et al.,* 2019). Moreover, at this time, the SAT is influenced by the weak westerly windsof the upper trough in the mid-troposphere because of the change taking place over Asia. These changes take place in a relatively short time due to the thermal equator that moves northward. In addition, in the spring season, strong contrast in temperature between sea and land and between valleys and mountains occurs. There exists a contrast in temperature, where cool moist air (polar maritime and polar continental air) comes from the eastern Mediterranean, and warm air temperature exists over the desert (tropical continental air). The inverse relationship between sea level pressure and SAT over Saudi Arabia is found (*Attada et al.,* 2019).

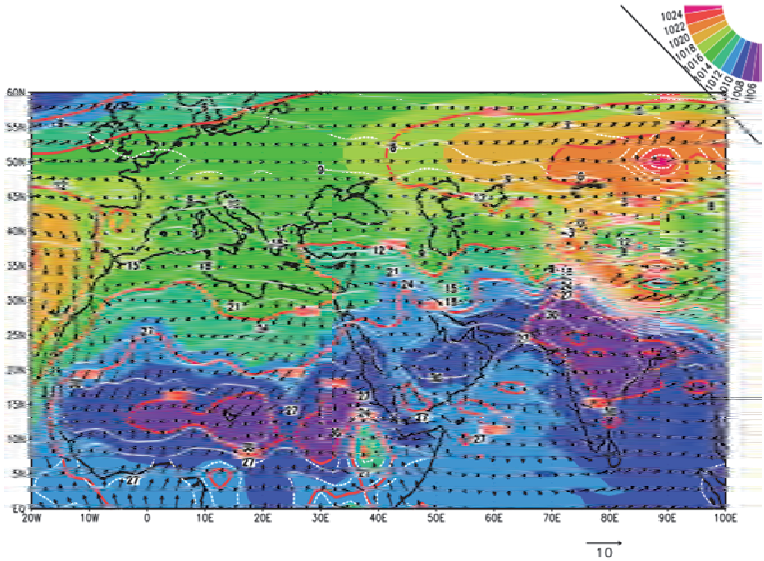


Fig. 11. The distribution pattern of the sea level pressure (solid red curve), surface air temperature (small dash curve), and wind vector in the spring season.

6.4. Synoptic circulation in the summer season

The distribution patterns of SAT become uniform over the Arabian Peninsula with intense temperatures (above 36 ºC) which are found in the eastern region, and minimum temperatures (30 ºC) over the northern part and the mountains of the southern region of the Arabian Peninsula in the summer season (*Fig. 12*). These results agree with the study of *Attada et al.* (2019). *Izumo et al.* (2008) and *Yao* and *Hoteit* (2015) illustrating that the SATs over the mountain of the southern region of the Arabian Peninsula decrease due to strong cross-equatorial flow along the Somalian coast, which cools the sea surface temperature of the western Arabian sea. The highest SATs in the summer season over the desert regions arise from the descent motion, which causes adiabatic warming (*Khan* and *Alghafari*, 2018; *Babu et al.*, 2018; *Attada et al.*, 2019). *Rodwell* and *Hoskins* (1996) and *Rodwell and Hoskins* (2001) found, that the descent motion is due to the monsoon-desert mechanism and stronger summer solar isolation. *El Kenawy et al.* (2014) and *El Kenawy* and *McCabe* (2016) illustrated, that the high temperature creates a thermal low that increases the cyclonic activities over the Arabian Peninsula. A strong and positive relationship is found between the SAT over the Arabian Peninsula and the mean sea level pressure over the regions of the Indian subcontinent and East Asia, but a negative relationship is found with the Saharan

heat low and the equatorial Atlantic Ocean (*Attada et al.,* 2019; *Babu et al.,* 2018). Also, *Attada et al.* (2019) found a cyclonic circulation over the Sahara in North Africa and the western Mediterranean regions that intensifies the Saharan thermal low and influences the SAT of the Arabian Peninsula. The convergence of air arising from the easterly wind from the equatorial Indian ocean and the westerly wind from Africa moves toward the Arabian Peninsula. The Indian summer monsoon and the Sudan low are the major synoptic circulations that influence Saudi Arabia from the south (*Almazroui et al.,* 2012a).

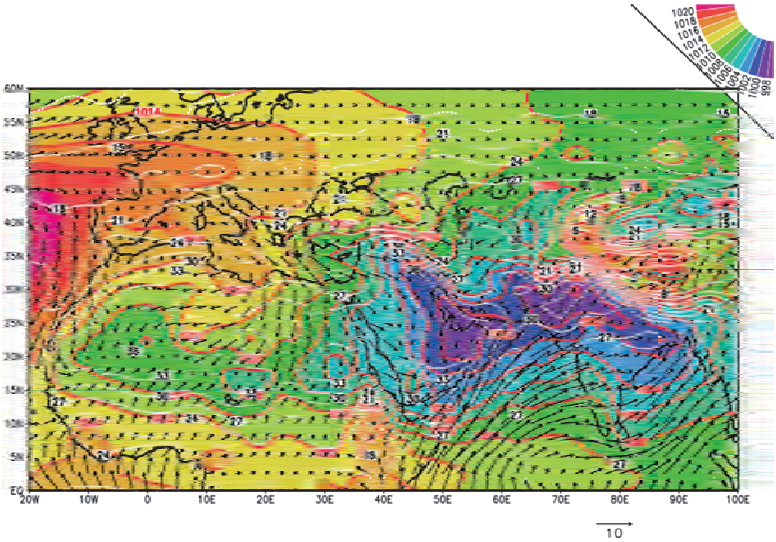


Fig. 12. The distribution pattern of the sea level pressure (solid red curve), surface air temperature (small dash curve), and wind vector in the summer season.

The orography and the distances from main bodies of water in the western parts of Saudi Arabia along the Red Sea coast have effects on the climatic distribution pattern of temperatures (*Al-Jerash,* 1985; *Ahmed,* 1997; *Subyani et al.,* 2010; *Attada et al.,* 2019). The main structures to be considered are the difference between the shore and inner areas; and the difference between the two shore zones and the windward and leeward sides of the mountains ranging from the south to north. The thermal low of the Indian monsoon is centered over northwestern India, Pakistan, and Baluchistan, which expands to the Arabian Peninsula and creates a strong heat on the surface. Another thermal low develops along the plain of southern Iraq that leads to the "shamal" winds (northwest and hot dusty winds), which blow southeastward across Kuwait and down into the Gulf (*Aurelius,* 2008). Above sea level from one to two kilometers of the Arabian Sea, the wind direction is shifted from southwest to northwest with an increase in

temperature with elevation (inversion), while northwesterly wind adverts tropical continental air mass from the "empty quarter". Upwelling is arising from the southwesterly wind when blowing over cold coastal water. Therefore, the outcome of the cold sea is to cool the maritime air mass of the lowest layers.

Usually, the shamal winds ease the SAT in summer. In some cases, the shamal winds arising from a dry cold front in the Arabian Peninsula decrease the maximum temperature by 5 ºC (*Aurelius,* 2008). In mid-July, the pressure gradient becomes weak in the area leading to a period of light northwesterly winds, and the shamal winds become weak. Usually, over Iraq, a thermal low develops that leads to a weak pressure gradient over the eastern region of Saudi Arabia and near Kuwait (*Preusser et al.,* 2002). *Fig. 12* illustrates the distribution pattern of sea level pressure and the wind vector in the summer season. Southeast Asia is dominated by the Indian monsoon. The Indian monsoon during a strong period (mid-August) extends to the Mediterranean Sea that breaks the frontal system from invading the area and pushing the subtropical high pressure east of Saudi Arabia. Southern Arabia is affected by the western margin of monsoon circulation, which controls the wind field and the amount of precipitation acting in terrestrial environments (*Preusser et al.,* 2002). At present, atmospheric circulation in summer is dominated by a series of low pressure cells that are located along the intertropical convergence zone (ITCZ) and across southern Arabia.

6.5. Synoptic circulation in the autumn season

In Saudi Arabia, the SAT, in general, is between 20 °C to 25 °C. The highest values of SAT are observed over the eastern and southeastern areas of Saudi Arabia, while the SAT is below 25 °C in the western and northwestern areas (*Fig. 13*). These results agree with the study of *Attada et al.* (2019). The mean SAT is about 29 °C in the autumn season in Saudi Arabia. During the autumn season, the weather of Saudi Arabia is influenced by the Sudan monsoon low, such as in the spring season, which is entering from the southeast (*Fig. 13*). Also, Saudi Arabia is influenced by the early Mediterranean depressions, which cause activity in weather when the southerly hot air mass meets the cold air mass over the northern region of Saudi Arabia. At the end of November, the troughs moving to eastward in the upper troposphere and the yearly cycle renews (*Ghazanfar* and *Fisher,* 1998). Strong negative correlations between the SAT and the mean sea level pressure in autumn over the Sahara and North Africa are observed (*Attada et al.,* 2019). Due to the intensification of the Sahara heat low and its movement towards the south of the Arabian Peninsula, surface air temperature increases over Saudi Arabia.

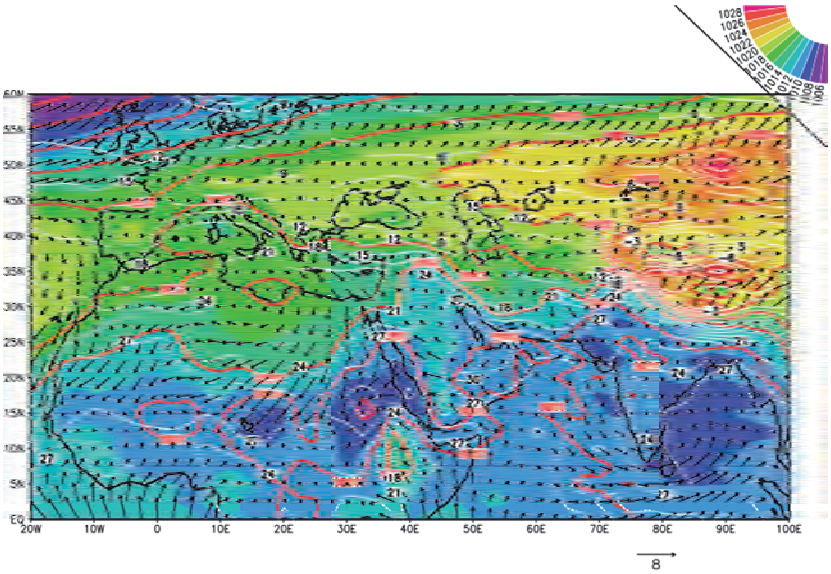


Fig. 13. The distribution pattern of the sea level pressure (solid red curve), surface air temperature (small dash curve), and wind vector in the autumn season.

7. Conclusions

The main findings of this review article can be summarized as follows:
For the past climate, from about 650 ka BP to the present, the SAT over Saudi Arabia during the earlier period is higher than in the present period due to atmospheric greenhouse gases (*Jansen et al.,* 2007). The climate over the world fluctuates/changes from cold to while it was warm from time to time. By the date 150,000 BCE, the climate was cold, warm in 140,000 BCE, similar to the present climate in 130,000 BCE, cold in 120,000 BCE, the gradual warming continued in 9500 BCE, and the cold periods prevailed for the next thousand years (*Edmiston,* 2007). The climate in the six periods (9,000–8,000, 6,000–5,000, 4,200–3,800, 3,500–500, 1,200–1,000, and 600–150 BP) are featured by polar cooling, tropical aridity, and main atmospheric circulation changes (*Mayewski et al.,* 1981). During the 8,000s BP, a 200-year-long cooling period occurred (*Edmiston,* 2007). Due to climate change, during four successive ice ages, hot periods in the Pleistocene were observed over Saudi Arabia (*Kotwickiet al.,* 2007).

For the present situation, the climates over Saudi Arabia vary from place to place due to the impact of climatic controls such as the difference between landmass and water-body, air masses, mountains barriers, elevations above mean sea level, atmospheric circulations, and ocean currents. Moreover, Saudi Arabia has occupied a vast area (about 80% of the Arabian Peninsula); therefore, the

temperature is different in space and time. The climate of Saudi Arabia is a desert climate (warming in the daytime and sudden cooling in nighttime) with the exception of the mountainous region in the southwest. Saudi Arabia is the warmest region in the summer season (above 50 °C) around the globe except for the coastal region. Over the north of the Tropics of Cancer and the inland of Saudi Arabia, high temperatures are observed. On the other hand, moderate SATs are found in the southwest highlands and northwest regions. The features of the climate of the Saudi Arabia are arid to semi-arid (*Köppen* and *Geiger,* 1928; *Walter* and *Lieth,* 1967; *Troll* and *Paffen,* 1980; *Muller,* 1982; *Moore,* 1986; *Almazroui et al.,* 2012a, 2012b). In addition, the seasonal temperatures are high in the central and northern regions compared to the southern region. Moreover, the diurnal temperature is characterized as high during the day and has an abrupt drop at night.

Four different types of air masses affect the climate elements of Saudi Arabia, as follows.

During wintertime, the polar continental air mass affects Saudi Arabia that comes from the Siberian high, which is dominant in central Asia and extends to the Arabian Peninsula. In addition, during winter, polar maritime air masses coming from the mid-latitude depression from the west influence the climate of Saudi Arabia. In spring and early summer, the tropical continental air mass influences the climate of Saudi Arabia. Also, during summer and autumn, the tropical continental air mass prevails over Saudi Arabia through the Indian monsoon. Moreover, during summer, the tropical maritime air mass comes from the north of the Indian Ocean and the Arabian Sea, which affects mainly the southwestern region of Saudi Arabia bringing hot and humid air.

SATs over Saudi Arabia are different from season to season due to different air masses that invade the region. During the winter season, SAT over the southeastern region is high, and it is higher than in the southwestern region due to the local topographical effect, it is moderate over western and central regions, and low over the northern region due to Mediterranean depression. During spring in the northern region, the SAT increases faster than over other regions of Saudi Arabia except for the eastern region. During the summer season, the eastern region of Saudi Arabia is warmer than the southern region. Unusually high SAT over desert areas exists due to the strong descent motion. The SAT in the summer season is the warmest due to the dominant Indian monsoon low over Saudi Arabia. The SAT during the autumn season is moderate (around 29 °C) over Saudi Arabia, while in the western and northwestern regions are lower than the eastern region of Saudi Arabia.

Mean annual SATs over Saudi Arabia are between 24 °C–30 °C. The highest mean annual values of the SAT are observed over the Al-Rub Al-Khali and southwest mainly along the Red Sea coast. Whereas, the lowest SATs (around 21 °C) are observed in the northern area. Low annual mean minimum of SAT

(6 °C–15 °C) is found over the northwestern and southwestern areas of Saudi Arabia. The SATs in the southeastern region (mainly over Al-Rub Al-Khali) are relatively high. The annual mean maximum SATs (between 27 °C and 33 °C) are found over the most regions of Saudi Arabia. The highest mean maximum SAT (33–36 °C) is observed over the central parts of the country and in the southwest region, while a slightly lower mean maximum SAT is found along the Red Sea coast.

Over Saudi Arabia, the warmest year is 1999 and the coldest year is 1992, according to the second report of the UNFCCC. The summer 2010 is the warmest season with an outstanding temperature (52.0 °C) observed in the city of Jeddah. During the coldest year, in winter and early spring, central and northern regions of Saudi Arabia have extreme cooling that associates with the invasion of the Siberian high, while during the warmest year, in summer and early autumn, the central and most of the eastern coasts and northeastern regions of Saudi Arabia have extreme warming, that associates with an extension of the Indian monsoon. Based on daily observations, the extreme SAT is warmer during the summer season (*Athar,* 2014; *AlSarmi* and *Washington,* 2014). *AlSarmi* and *Washington* (2014) noticed, that the northern region of the Arabian Peninsula has high SAT in the daytime, while the southern region has high SAT in the nighttime. In addition, the frequency of the extreme warm SAT is greater than the extreme cold SAT based on seasonal and annual observations. Moreover, the southwest coastal area has shown more warming than the inland area. Over most stations of Saudi Arabia, strong warming above 5 days/decade is found during the spring to autumn seasons (*Islam et al.,* 2015). In addition, it is observed during these seasons, that the number of warm daytimes increases rapidly compared to the number of warm nighttimes during spring, summer, and autumn seasons over most stations of Saudi Arabia.

The atmospheric circulations (low and high pressure) play a vital role with regard to weather and climate. In general, it can be concluded that the Siberian high pressure is controlling the SAT of Saudi Arabia in the winter season, and the Indian monsoon low is controlling the SAT in the summer season (*Davydova et al.,* 1966; *Vorhees* 2006; *Hasanean et al.,* 2015; *Babu et al.,* 2018). In addition, the synoptic circulation patterns in spring and autumn are complicated, when centers of high and low dominate the regions of Saudi Arabia. During the spring and autumn seasons, the extended Mediterranean low pressure influences the northern region and the Sudan monsoon low/ Red Sea trough influence the southern region of Saudi Arabia, while subtropical high pressure impacts the western region, and the Siberian high pressure impacts the eastern region of Saudi Arabia.

During the spring season (March, April, and May),the subtropical high pressure from the west and the Siberian high pressure from the east effect the SAT of Saudi Arabia. The subtropical high pressure and the Siberian high pressure

affect each other. When Siberian high pressure is strong and extends to the Saudi Arabian region, the subtropical high pressure is weak and the SAT is cold and vice versa; when the Siberian high pressure is weak, the subtropical high pressure is strong, and the SAT is warm. The SAT of Saudi Arabia is influenced by the Mediterranean depression from the north and may extend to the central region, which brings cold air, and the Sudan monsoon low from the south may extend to the north that brings warm and moist air (*Attada et al.,* 2019). The circulation pattern in the autumn season is similar to the circulation pattern in the spring season.

The main features to be considered in the synoptic circulation patterns of Saudi Arabia in the different seasons are as follows.

During the winter season, the synoptic circulation features that influenced the SAT of Saudi Arabia are: the trough emanating from the Icelandic low (cold moist air mass), which extends southward influencing the weather in Saudi Arabia (*Hasanean et al.,* 2015; *Attada et al.,* 2019). The SAT decreases (cooling) during the deepening of the Icelandic low, while it increases (warming) during the weakening of the Icelandic low. In addition, the Siberian high pressure is dominant over Asia (cold air mass), and it tends to merge with high pressure centers over the Arabian Peninsula (Saudi Arabian or Saharan high pressure) (*Guo,* 1996; *Gong* and *Wang,* 1999; *Miyazaki et al.,* 1999; *Yin,* 1999; *Hasanean et al.,* 2013, 2015). The SAT warms in the northern and central regions of Saudi Arabia during the intensification of the Siberian high pressure in the winter season (*Hasanean et al.,* 2013, 2015; *Attada et al.,* 2019).

In the winter season, the SAT is cooling when the Siberian high pressure is deepening and dominating over Saudi Arabia that prevents the Mediterranean depression, while the SAT is relatively warm when the Siberian high pressure is weakening, where the relatively warm moist air is blowing from southern Europe and from the southern and southeastern parts of Saudi Arabia. The SAT of Saudi Arabia is influenced by the extension of the subtropical high pressure, which is related to the movement of the Siberian high pressure to the east and/or to the west (*Barry,* 1992; *Hasanean et al.,* 2015). The SAT is relatively warming when the subtropical highpressure is weakening, and the SAT is cooling when the subtropical high-pressure is deepening. The Saudi Arabian high pressure, as a relatively small feature, plays an important role in the SAT in the winter season (*Vorhees*, 2006; *Hasanean et al.,* 2013). When the Saudi Arabian high pressure dominates over the region, the air is more stable, and the SAT is relatively warm across the region. Moreover, the Sudan low and the Red Sea trough are influencing the SAT of Saudi Arabia in the winter season (*Rasuly et al.,* 2012), carrying humid air to the region. When the Sudan low is weakening, the SAT is cooling, and when Sudan low is deepening, the SAT is relatively warming (*Hasanean et al.,* 2015; *El-Fandy,* 1940, 1948; *Hasanean et al.,* 2013). Two circulation patterns dominate the SAT of Saudi Arabia in the winter season, the

first pattern is when the subtropical high pressure merges with the Siberian high pressure to prevent the interaction between extratropical and mid-latitude systems that leads to cooling of the SAT (*Hasanean et al.,* 2015). The second pattern is when the subtropical high pressure is weakening, the Siberian high pressure is strengthening, Icelandic low is weakening, and Sudan low is deepening (*Hasanean et al.,* 2015) that allow two different air masses to invade Saudi Arabia (cold moist air mass from the Mediterranean depression from west, and warm moist air mass from the Sudan low pressure from south), (*Hasanean et al.,* 2015).

During the spring season, when the Siberian high pressure is deepening, the SAT is coolingm while when the subtropical is deepening, the Siberian high pressure is weakening and the SAT is warming (*Attada et al.,* 2019). In addition, the Mediterranean depression can influence the SAT when extending to the northern region of Saudi Arabia and cooling the region. From the southern region, the Sudan low is extending to the northern region and/or Asian monsoon low is extending from the east thereby warming Saudi Arabia.

During the summer season, the Indian summer monsoon and the Sudan low are the major synoptic circulations influencing the SAT of Saudi Arabia (*Almazroui et al.,* 2012a; *Babu et al.,* 2018; *Attada et al.,* 2019). Another thermal low develops along the plain of southern Iraq leading to the "shamal" winds over the eastern parts of Saudi Arabia (*Aurelius,* 2008). The Indian monsoon may extend to the Mediterranean Sea, which cuts the frontal system and pushes the subtropical high pressure to the east of Saudi Arabia. When the subtropical high pressure is strengthening and extending eastward, it influences the SAT in the western region of Saudi Arabia relatively cooling the region. In addition, when the Saharan thermal low is deepening over North Africa, it influences the SAT of Saudi Arabia.

During the autumn season, the Sudan monsoon low coming from the southeast influences the SAT of Saudi Arabia, thereby warming the region. In addition, the early autumn Mediterranean cyclone is cooling SAT in the northern region of Saudi Arabia. The SAT of the northern region of Saudi Arabia may be influenced by the subtropical high pressure thereby cooling the region.

Acknowledgments: The authors are thankful to the Faculty of Meteorology, Environmental and Arid Land Agriculture, King Abdulaziz University for making available the computer and other facilities in this work. Also, the authors are thankful to the colleagues of the Department of Meteorology for their suggestions and advices. We thank the reviewers for their comments on earlier versions of this article.

References

Abdullah, M.A., and *Al-Mazroui, M.A.,* 1998: Climatological study of the southwestern region of Saudi Arabia. I. Rainfall Analysis. *Clim. Res. 9,* 213–223. https://doi.org/10.3354/cr009213

Ahmed, B.Y.M., 1997: Climatic classification of Saudi Arabia: An application of factor - cluster analysis. *GeoJournal 41,* 69–84. https://doi.org/10.1023/A:1006827322880

Al-Jerash, M.A., 1985: Climatic subdivisions in Saudi Arabia: an application of principal component analysis. *J. Climatol. 5,* 307–323. https://doi.org/10.1002/joc.3370050307

Almazroui, M., 2019: Temperature changes over the CORDEXMENA domain in the 21st century using CMIP5 data downscaled with RegCM4: a focus on the Arabian peninsula. *Adv. Meteorol.* 2019, Article ID 5395676. https://doi.org/10.1155/2019/5395676

Almazroui, M., Dambul R., Islam, N., and *Jones, P.D.,* 2014: Principal components-based regionalization of the Saudi Arabian climate. *Int. J. Climatol. 35,* 2555–2573. https://doi.org/10.1002/joc.4139

Almazroui, M.A., 2013: Simulation of present and future climate of Saudi Arabia using a regional climate model (PRECIS). *Int. J. Climatol. 33,* 2247–2259. https://doi.org/10.1002/joc.3721

Almazroui, M.A., 2020: Changes in Temperature Trends and Extremes over Saudi Arabia for the Period 1978–2019. *Adv. Meteorol. 2020,* Article ID 8828421. https://doi.org/10.1155/2020/8828421

Almazroui, M.A., Al Khalaf, A.K., Abdel Basset, H.M., and *Hasanean, H.M.,* 2009: Detecting Climate Change Signals in Saudi Arabia Using Surface Temperature. Project Number (305/428) is supported from King Abdelaziz University, Kingdom of Saudi Arabia.

Almazroui, M.A., Islam N.; Athar H., Jones P.D., and *Rahmana M.A.,* 2012b. Recent climate change in the Arabian Peninsula: Seasonal rainfall and temperature climatology of Saudi Arabia for 1979–2009. *Atmos. Res. 111,* 29–45. https://doi.org/10.1016/j.atmosres.2012.02.013

Almazroui., Islam N., Athar H., Jones P.D., and *Rahmana M.A.,* 2012a: Recent climate change in the Arabian Peninsula: annual rainfall and temperature analysis of Saudi Arabia for 1978–2009. *Int. J. Climatol. 32,* 953–966. https://doi.org/10.1002/joc.3446

Alpert, P., Osetinsky, I., Ziv, B., and *Shafir, H.,* 2004: Semi-objective classification for daily synoptic systems: application to the eastern Mediterranean climate change. *Int. J. Climatol. 24,* 1001–1011. https://doi.org/10.1002/joc.1036

AlSarmi, S.H., and *Washington, R.,* 2014: Changes in climate extremes in the Arabian Peninsula: analysis of daily data. *Int. J. Climatol. 34,* 1329–1345. https://doi.org/10.1002/joc.3772

Athar, H., 2014: Trends in observed extreme climate indices in Saudi Arabia during 1979–2008. *Int J Climatol. 34,* 1561–1574. https://doi.org/10.1002/joc.3783

Attada, R., Dasari, H.P., Chowdary, J.S., Yadav, R.K., and *Knio, O. et al.,,* 2019: Surface air temperature variability over the Arabian Peninsula and its links to circulation patterns. *Int. J. Climatol. 39,* 445–464. https://doi.org/10.1002/joc.5821

Aurelius, L., 2008: The impact of Shamal winds on tall building design in the Gulf. Dubai Building, Government of Dubai.

Babu, C.A., Jayakrishnan, P.R., and *Varikoden, H.,* 2016: Characteristics of precipitation pattern in the Arabian Peninsula and its variability associated with ENSO. *Arabian J. Geosci. 9,* 186. https://doi.org/10.1007/s12517-015-2265-x

Barry, R.G. and *Chorley, R.J.,* 2006: Atmosphere, weather &climate, 6th edn. London: Routledge, 392.

Davydova, M., Kamenskii, A., Nekiukova, N., and *Tushinskii, G.,* 1966: Fizicheskaia Geografiia SSSR. Moskva: Prosveshchenie. Encyclopedia of Earth Sciences Series, Encyclopedia of World Climatology, 57–82, 439–704.

Edmiston, R., 2007. Hot time on the old globe tonight. available at: www.dcr.net/,-stickmak/JOHT/joht32hottime.htm (accessed November 30, 2018).

El Kenawy, A.M. and *McCabe, M.F.*, 2016: A multi-decadal assessment of the performance of gauge- and model-based rainfall products over Saudi Arabia: climatology, anomalies and trends. *Int. J. Climatol. 36*, 656–674. https://doi.org/10.1002/joc.4374

El Kenawy, A.M., McCabe, M.F., Stenchikov, G., and *Raj, J.*, 2014: Multi-decadal classification of synoptic weather types, observed trends and links to rainfall characteristics over Saudi Arabia. *Frontiers Environ. Sci. Engineer. 2*, 37. https://doi.org/10.3389/fenvs.2014.00037

El Kenawy, A.M., McCabe, M.F., Vicente-Serrano, S.M., Robaa, S.M., and *Lopez-Moreno, J.I.*, 2016: Recent changes in continentality and aridity conditions over the Middle East and North Africa region, and their association with circulation patterns. *Climat. Res. 69*, 25–43. https://doi.org/10.3354/cr01389

El-Fandy, M.G., 1940: The formation of depressions of khamasin type. *Quart. J. Royal Meteorol. Soc. 66*, 323–335. https://doi.org/10.1002/qj.49706628607

El-Fandy, MG., 1948: The effect of the Sudan monsoon low on the development of thundery conditions in Egypt, Palestine and Syria. *Quart. J. Royal Meteorol. Soc. 74*, 31–38.

Fisher, M. and *Membery, D.A.*, 1998: Climate. In Vegetation of the Arabian Peninsula. (eds. *Ghazanfar SA, Fisher M*). Kluwer Academic Press: Netherlands, 5–38. https://doi.org/10.1002/qj.49707431904

Ghazanfar, S.A. and *Fisher, M.*, 1998: Vegetation of the Arabian Peninsula. Kluwer Academic Publishers. https://doi.org/10.1007/978-94-017-3637-4

Gong, D.Y. and *Wang, S.W.*, 1999: Long term variability of the Siberian High and the possible influence of global warming. *Acta Geographica Sinica 54*, 125–133 (in Chinese) https://doi.org/10.5194/gh-54-125-1999

Guo, Q.Y., 1996: Climate change in China and East Asian monsoon. In: (ed. Yafeng Shi) Historical climate change in China. Shandong Science and Technology Press, Ji'nan, 468–483.

Handmer, J. et al., (34 authors), 2012: Changes in impacts of climate extremes: human systems and ecosystems. In book: Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation Publisher: Editors: *Field, C B and Barros, V and Stocker, T F and Qin, D and Dokken, D J and Ebi, K L and Mastrandrea, M D and Mach, K J and Plattner, G K and Allen, S K and Tignor, M and Midgley, P M.* A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK, and New York, NY, USA.

Hasanean H.M., Abdel Basset H.M., and *Husseein M.A.A.*, 2015. On the Relationship between Climatic Variables and Pressure Systems over Saudi Arabia in the Winter Season. *Adv. Atmos. Sci. 32*, 690–703. https://doi.org/10.1007/s00376-014-4149-5

Hasanean, H.M. and *Almazroui, M.A.*, 2017: Teleconnections of the tropical sea surface temperatures to the surface air temperature over Saudi Arabia in summer season. *Int. J. Climatol. 37*, 1040–1049. https://doi.org/10.1002/joc.4758

Hasanean, H.M., Almazroui M.A., and *Jones, P.D.; Alamoudi, A.A.*, 2013. Siberian high variability and its teleconnections with tropical circulations and surface air temperature over Saudi Arabia. *Climat. Dynam. 41*, 2003–2018. https://doi.org/10.1007/s00382-012-1657-9

Hastenrath, S., 1985: Climate and the Circulation of the Tropics. Boston (D.Reidel). https://doi.org/10.1007/978-94-009-5388-8

Hurrell, J.W., Kushnir, Y., Ottersen, G., and *Visbeck, M.,* 2003: The North Atlantic Oscillation: Climatic Significance and Environmental Impact. Geophysical Monograph Series, 134. Washington, DC: AGU. https://doi.org/10.1029/GM134

Islam, N., Almazroui, M.A., Dambul, R., Jones, P.D., and *Alamoudi, A.O.,* 2015: Long-term changes in seasonal temperature extremes over Saudi Arabia during 1981–2010. *Int. J. Climatol., 35,* 1579–1592. https://doi.org/10.1002/joc.4078

Izumo, T., Montégut Cd.B., Luo, J-J., Behera, S.K., Masson, S., and *Yamagata, T.,* 2008: The role of the western Arabian Sea upwelling in Indian monsoon rainfall variability. *J. Clim. 21,* 5603–5623. https://doi.org/10.1175/2008JCLI2158.1

Jansen, E., Overpeck, J., Briffa, K.R., Duplessy, J.-C., Joos, F., Masson-Delmotte, V., Olago, D., Otto-Bliesner, B., Peltier, W.R., Rahmstorf, S., Ramesh, R.; Raynaud, D., Rind, D., Solomina, D.; Villalba, R., Zhang, D. Palaeoclimate, in Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K.B., Tignor, M., and *Miller, H.L.* 2007: In: Climates of the Arabian Peninsula, Climate Change The Physical Science Basis, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge.

Kendrew, W.G., 2012: The climate of the continent. Oxford University Press, New York, 5th edition. KISR. Assorted publications on impact of dust storms in Kuwait, available at http://www.kisr.edu.kw

Khan, S. and *Alghafari, Y.,* 2018: Temperature, Precipitation and Relative Humidity Fluctuation of Makkah Al Mukarramah, Kingdom of Saudi Arabia (1985-2016). *Transact. Machine Learn. Artificial Intellig. 6,* 42–58.

Köppen, W., 1936: Das geographisca System der Klimate, in: Handbuch ¨ der Klimatologie, edited by: Koppen, W. and Geiger, G., 1. C. ¨ Gebr, Borntraeger, 1–44.

Köppen, W., and *Geiger, R.,* 1928: Klimate der Erde, Justus Perthes, Gotha.

Kotwicki, V., Al Sulaimani, Z., and *Al Khatri, A.,* 2007: Effect of long-term weather patterns and climate change on water resources. Proceedings of the IV Wadi Conference, Muscat.

Kotwicki, V. and *Al Sulaimani, Z.,* 2009: Climates of the Arabian Peninsula - past, present, future. *Int. J.f Climate Change Strat. Manage. 1,* 297–310.https://doi.org/10.1108/17568690910977500

Krishna, L.V., 2014: Long Term Temperature Trends in Four Different Climatic Zones of Saudi Arabia. *Int. J. Appl. Sci. Technol. 4,* 233–242.

Langodan, S., Cavaleri, L., Yesubabu, V., and *Hoteit, I.,* 2014: The Red Sea: a natural laboratory for wind and wave modeling. *J. Phys. Oceanogr. 44,* 3139–3159. https://doi.org/10.1175/JPO-D-13-0242.1

Makrogiannis, T.T., Sahsamanoglou, H. S., Flocas, A. A., and *Bloutosos, A.A.,* 1991: Analysis of monthly zonal index values and long-term changes of circulation over the North-Atlantic and Europe. *Int. J. Climatol. 11,* 493–503. https://doi.org/10.1002/joc.3370110503

Masatoshi, M., and *Urushibara, R.,* 1981: Regionality of climatic change in East Asia. *GeoJourna, 5,* 121–132. https://doi.org/10.1007/BF02582045

Mayewski, P.A., Rohling, E., Stager, E., Karlen, J.C., Maasch, W., Meeker, A., Meyerson, L.D.; Gasse, E.A., van Kreveld, F., Holmgren, S., Lee-Thorp, K., Rosqvist, J., Rack, G., Staubwasser, F., Schneider, M., and Steig, R.R., 2004: Holecene climate variability. Quaternary Res. 62, 243–255. https://doi.org/10.1016/j.yqres.2004.07.001

McClure, H.A., 2007: A new Arabian stone tool assemblage and notes on the Aterian industry of North Africa. *Arabian Archaeol. Epigrap. 5,* 1–16. https://doi.org/10.1111/j.1600-0471.1994.tb00052.x

Miyazaki, S., Yasunari, T., and *Adyasuren, T.,* 1999: Abrupt seasonal changes of surface climate observed in northern Mangolia by an automatic weather station. *J. Meteorol. Soc. Japan 77,* 583–593https://doi.org/10.2151/jmsj1965.77.2_583.

Moore, E., 1986: Gardening in the Middle East. London: Stacey International.

Muller, M.J., 1982: Selected climatic data for a global set of standard stations for vegetation science, In: (Ed. *Lieth, H.*), Tasks for Vegetation Sciences, 5. Dordrecht: Kluwer Academic Publishers Group. https://doi.org/10.1007/978-94-009-8040-2

New, M., Hulme, M., and *Jones, P.,* 2000: Representing twentieth–century space–time climate variability. Part II: development of 1901–1996 monthly grids of terrestrial surface climate. *J. Climate 13,* 2217–2238. https://doi.org/10.1175/1520-0442(2000)013<2217:RTCSTC>2.0.CO;2

Parmesan, C., Root T.L., and *Willig, M.R.,* 2000: Impacts of Extreme Weather and Climate on errestrial Biota. *Bull. Amer. Meteorol. Soc., 81,* 443–450. https://doi.org/10.1175/1520-0477(2000)081<0443:IOEWAC>2.3.CO;2

Peel, M.C., Finlayson, B.L., and *McMahon, T.A.,* 2007. Updated world map of the Köppen-Geiger climate classification. *Hydrol. Earth Syst. Sci., 11,* 1633–1644. https://doi.org/10.5194/hess-11-1633-2007

Peterson, T.C., Stott, P.A., and *Herring, S.,* 2012: Explaining Extreme Events of 2011 from a Climate Perspective. *Bull. Amer. Meteorol. Soc., 93,* 1041–1067. https://doi.org/10.1175/BAMS-D-12-00021.1

Preusser, F., Radies, D., and *Matter, A.,* 2002: A 160,000-Year Record of Dune Development and Atmospheric Circulation in Southern Arabia. *Science, 296,* 2018–2020. https://doi.org/10.1126/science.1069875

Rasuly, A.A., Babaeian, I., Ghaemi, H., and *ZawarReza, P.,* 2012: Time series analysis of the pressure of the synoptic pattern centers affecting on seasonal precipitation of Iran. Geography and Development 10nd Year, No. 27, Summer 2012, 18–21. [Available online at http://www.sid.ir/en/VEWSSID/J pdf/ 98920122706.pdf]

Rodwell, M.J., and *Hoskins, B.J.,* 1996: Monsoons and the dynamics of deserts. *Quart. J. Royal Meteorol. Soc., 122,* 1385–1404. https://doi.org/10.1002/qj.49712253408

Rodwell, M.R., and *Hoskins, B.J.,* 2001: Subtropical anticyclones and monsoons. *J. Climate 14,* 3192–3211. https://doi.org/10.1175/1520-0442(2001)014<3192:SAASM>2.0.CO;2

Rogers, J. C., 1984: The Association between the North Atlantic Oscillation and the Southern Oscillation in the Northern Hemisphere. *Mon. Weather. Rev. 112,* 1999–2015. https://doi.org/10.1175/1520-0493(1984)112<1999:TABTNA>2.0.CO;2

Rogers, J.C., and *van Loon H.,* 1979: The seasaw in winter temperature between Greenland and northern Europe. Part II: Some oceanic and atmospheric effects in middle and high latitudes. *Mon. Weather Rev. 106,* 324–330.

Sahsamanoglou, H.S., and *Makrogiannis, T. J.,* 1992. Temperature trends over the Mediterranean regions, 1950–88. *Theor. Appl. Climatol. 45,* 183–192. https://doi.org/10.1007/BF00866191

Stern, R.J., Avigadb, D., Miller, N.R., and *Beyth, M.,* 2006: Evidence for the Snowball Earth hypothesis in the Arabian next term-Nubian Shield and the East African Orogen. *J. African Earth Sci. 44,* 1–20. https://doi.org/10.1016/j.jafrearsci.2005.10.003

Subyani, A.M., Al-Modayan, A.A., and *Al-Ahmadi, F.S.,* 2010: Topographic, seasonal and aridity influences on rainfall variability in western Saudi Arabia. *J. Environ. Hydrol. 18,* 1–11.

Taha, M.F., Harb, S.A., Nagib, M.K., and *Tantawy, A.H.,* 1981: The Climate of the Near East, In: Takahashi, K. and H. Arakawa (eds.), Climate of Southern and Western Asia, Elsevier, 183–241.

Tarawneh, QY. and *Chowdhury, S.,* 2018: Trends of Climate Change in Saudi Arabia: Implications on Water Resources. *Climate, 6,* 8. https://doi.org/10.3390/cli6010008

Troll, C., and *Paffen, K.H.,* 1980: Jahreszeitenkarte der Erde, Berlin.

United Nations Framework Convention on Climate Change (UNFCCC, 2011. Second report, National Communication Kingdom of Saudi Arabia Report, 2011

Viswanadhapalli, Y., Dasari, H.P., Langodan, S. Challa, V.S., and *Hoteit, I.,* 2017: Climatic features of the Red Sea from a regional assimilative model. *Int. J. Climatol. 37,* 2563–2581, https://doi.org/10.1002/joc.4865

Vorhees, C.D., Murphree, T., and *Pfeiffer, L.K.* 2006: Tropical Climate Variations and Their the Northwest Indian Ocean-Northeast Africa- Southwest Asia Region, Tropical Met Conference, Apr 06, 2006.

Walter, H., and *Lieth, H.,* 1967. Klimadiagramm-Weltatlas, Jena: VEB Gustav Fischer Verlag.

Wilby, R., 1993: Evidence of ENSO in the synoptic climate of the British Isles since 1880. *Weather 48,* 234–239. https://doi.org/10.1002/j.1477-8696.1993.tb05897.x

Yao, F. and *Hoteit, I.,* 2015: Thermocline regulated seasonal evolution of surface chlorophyll in the Gulf of Aden. PLoS ONE, 10. https://doi.org/10.1371/journal.pone.0119951

Yin, Z.Y., 1999: Winter temperature anomalies of the North China Plain and macroscale extratropical circulation. *Int. J. Climatol. 19,* 291–308. https://doi.org/10.1002/(SICI)1097-0088(19990315)19:3<291::AID-JOC334>3.0.CO;2-B

Zhu, Q.G.; Shi, N.; Wu, Z.H., 1997. Low frequency variation of winter ACAs in north hemisphere and climate change in China during the past century. *Acta Meteorlogica Sinica 55*, 750–758.

DOI:10.28974/idojaras.2022.4.5

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 126, No. 4, October – December, 2022, pp. 545–566

Synoptic circulation patterns associated with foehn days in Sofia in the period 1979–2014

Krasimir Stoev [1,2,*], **Piia Post** [3], and **Guergana Guerova** [1]

[1] *Sofia University "St. Kliment Ohridski"*
Physics Faculty, Department Meteorology and Geophysics
5 James Bourchier blvd., BG-1164 Sofia, Bulgaria

[2] *National Institute of Meteorology and Hydrology*
66 Tsarigradsko shose Blvd., 1784 Sofia, Bulgaria

[3] *Institute of Physics University of Tartu, Estonia*

**Corresponding author E-mail: krasimir.stoev@gmail.com*

(Manuscript received in final form September 12, 2021)

Abstract— Foehn is a well-known example of local atmospheric circulation and is an extreme weather event for wind gusts. It can cause rapid snowmelt in spring or spread of forest fire in summer, as well as significant economic losses. The foehn in Bulgaria is observed on the northern slopes of the mountains. For the period 1979–2014, 261 foehn days are registered north of the Vitosha mountain, where Sofia valley is located. The average annual number of foehn days is 8.1, 8.3, and 4.5 for the periods 1985–1994, 1995–2004, and 2005–2014. After 2004, the average annual number of days with foehn decreases, and the lowest maximum wind gusts are registered. To check whether atmospheric circulation changes, could be the reason for this change two objective circulation classifications and a manual one are used to study the foehn occurrence in the Sofia valley. Based on the GrossWettertypen (GWT) catalogue of circulation patterns produced by the COST Action 733 and ten circulation types of the Jenkinson-Collison (JCT) catalogue of weather types the largest number of foehn days occur at the SW, W, and NW circulation patterns. GWT and JCT classifications with 26 types confirm the foehn occurrence during the W and SW flows, but add two more cyclonic types, the CW and CSW. For the foehn days in March 1979–2014, best agreement was found with the manual circulation classification with 26 types. A comparison between the decades 1995–2004 and 2005–2014 shows a substantial decrease in western and northwestern circulation types during the foehn days. An analysis of circulation types for all days confirms an overall reduction of W and NW circulation types after 2004.

Key-words: foehn climatology, manual classification, objective circulation classification

1. Introduction

Foehn is a well-known example of local atmospheric circulation with warm and dry wind blowing on the leeward side of the mountains. This severe weather event caused by wind gusts can trigger rapid snowmelt in spring, forest fire spread in summer, or disruption in air traffic all year. On the Balkan Peninsula, the foehn events are closely linked to the development of cyclonic vortices in the presence of a strong flow. The Mediterranean Sea is a well-known area, where cyclones form. Mediterranean cyclones and their formation, frequency, trajectories, and influence on the weather and the climate of the Balkan Peninsula have been studied by many authors (*Stanchev*, 1954; *Pisarski*, 1955 (a,b); *Blagoev*, 1961; *Martinov*, 1967; *Peiter,* 1975; *Radinovic,* 1987). *Bocheva et al*. (2007) report that the monthly and annual distribution of the cyclones generated over the Mediterranean for the period 1980–2001 show the highest cyclone frequency during winter (from December to March). The cyclonic activity decreases in April, when cyclones appear three times less frequently than in March. From May to September, cyclogenesis is a rare phenomenon. The number of cyclones strongly increases in November, being four times higher than in October. *Marinova et al*. (2005) find that: 1) after 1990, the cyclogenesis season over the Mediterranean Sea is two months shorter, and it is located mainly over the central part of the sea and 2) the largest number of Mediterranean cyclones pass over the southern part of the Balkan peninsula. The first study of the Mediterranean cyclone trajectories was by *Van Bebber* (1891). Since then, the main paths of the Mediterranean cyclones and their seasonal variability have been studied by many authors (*Pisarski,* 1955 (a,b); *Popova et al*., 1975; *Martinov*, 1983; *Jansa et al.,* 2001; *Bartholy et al.,* 2009, *Catrina et al.,* 2019). Three main trajectories are reported, namely: 1) through Croatia and Hungary (NW), 2) through the Adriatic Sea and the Balkan Peninsula toward the Black Sea (E), and 3) southern parts of the Balkan Peninsula toward Asia Minor (SE then E).

The foehn in Bulgaria is observed on the northern slopes of the mountains, as a result of advection of the warm air from the south and southwest. Its occurrence is the highest north of the Vitosha and the Balkan Mountains. Among the highest foehn occurrence is the Sofia valley, located at the foot of the north slopes of the Vitosha mountain. *Stoev* and *Guerova* (2020) present a manual classification of the meteorological conditions leading to foehn in the central meteorological station in Sofia. For the period 1975–2014, there are 298 foehn days classified in four manual circulation types. The circulation type associated with a Mediterranean cyclone has the highest frequency (52%). Foehn climatology gives an average annual number of 7.5 foehn days, and the lowest annual number of days (4.5) is registered in the last decade 2005–2014. Wind gusts over 14 m/s require issuing of a warning code, following the Meteoalarm warning system. For the period 1993–2014, in 79.6% of the foehn cases, wind gusts in Sofia range between 14 and 20 m/s (which corresponds to the yellow

Meteoalarm code), 15.8% are with gusts between 20–30 m/s (orange code), and 4.6% are with gusts 30 m/s (red code). A mediterranean cyclone trajectory over Hungary results in 52, 65, and 100% of the days with yellow, orange, and red codes, respectively.

Manual circulation classifications have two major limitations, namely: 1) they are labor intensive, i.e., require many person hours, and 2) the results are not consistent among investigators, i.e., they are not reproducible (*Yarnal,* 1993). The COST Action 733 "Harmonization and Applications of Weather Types Classifications for European Regions" (2005–2010, http://cost733.met.no/index.htm) aimed to develop a classification technique scalable to any European region for a wide range of applications. A general numerical approach was required to assess and compare atmospheric circulation classification and typical weather regimes in Europe to reach this goal. The COST Action 733 produced an extensive, consistent catalogue of atmospheric circulation type classifications based on different methodological concepts. A set of 33 methods or algorithms, ranging from manual classifications to data mining and machine learning methods, have been used to classify daily circulation patterns (*Philipp et al.*, 2016; *Tveito et al.*, 2016). A comprehensive dataset of classification catalogues, time series of type numbers or names representing atmospheric states for 12 European domains has been compiled using a specially developed open-source software package (*Tveito* et al., 2016). They are categorized into 1) weather type classifications - grouping several weather variables to create classes of atmospheric states and 2) circulation classifications - based only on circulation variables like atmospheric pressure, geopotential height, or large-scale wind components (*Philipp et al.,* 2016). The COST 733 weather type classifications have been applied to various phenomena, including heavy precipitations, freezing rains, droughts, floods, snow avalanches, and many more (*Tveito et al*., 2016). Two weather type classifications have been conducted in Bulgaria for heavy precipitations (*Neykov et al.,* 2016) and freezing rains (*Nikolov et al.,* 2016).

This study aims to find the synoptic-scale atmospheric circulation patterns for foehn days in Sofia, with an ultimate goal to understand better the role of atmospheric circulation changes in the decreased number of foehn days. Former manual classification of foehn days' weather charts provides the grounded choice of two objective classifications for this purpose. In Section 2, the circulation classifications are presented. Foehn climatology and objective circulation classification using two threshold-based methods are applied for Sofia valley. Summary and outlook are given in Section 4.

2. Method and data sets

2.1. Objective circulation classifications

The Cost733class software version 1.2 (*Philipp et al.,* 2016) is used to classify atmospheric circulation. The Cost733class is a software package to create,

compare, visualize, and evaluate weather and circulation type classifications. The software includes more than 20 automated methods and is licensed under GPL with documentation and software download available on the cost733wiki. For this study, two circulation classifications are selected: the GrossWettertypen (GWT) and the Jenkinson-Collison Type (JCT) classification. Both GWT and JCT are threshold-based classification methods with circulation types defined by using a numerical threshold for circulation indices. The main advantage of threshold-based classifications is that they discriminate between synoptically significant types (*Philipp et al*., 2016). The assignment of cases to the classes is done objectively and is automated by using threshold values for pre-defined indices discriminating between the types. Use of three indices result in three states (low, intermediate, and high) and 27 circulation types. For 9 or 18 circulation types, a smaller number of indices/states is required (*Philipp et al.,* 2016). The indices represent: 1) large-scale flow directions, i.e., zonal and meridional and 2) vorticity, i.e., high/low central pressure. The indices are derived from one circulation map, thus, the method is not suitable for multi-parameter datasets requiring two or more circulation maps to be classified together (*Philipp et al.,* 2016).

The GWT circulation types are derived by calculating correlations with raw or normalized correlation coefficients (*Beck*, 2000; *Beck et al*., 2007) and prototype fields of a zonal flow, meridional flow, and cyclone in the center of the domain (*Beck et al.,* 2007). The JCT circulation types (*Jenkinson* and *Collison*, 1977) are derived using an automated scheme for the Lamb weather types (*Lamb,* 1972). Daily 850 hPa geopotential height fields from ERA5 reanalysis with horizontal resolution of 31 km (*Hersbach et al.,* 2020) are classified into circulation types for the period 1979–2014. The calculations produced catalogues of classifications, where each day is specified by one circulation type from every classification. At first prototypical patterns of an idealized zonal and meridional flow and a low pressure in the center of the field are defined. The correlations of the real geopotential fields with these patterns are calculated. These correlations give the indices of zonality, meridionality, and vorticity of the atmospheric flow at that day. The spatial resolution of the fields is 0.5°. The circulation classifications are computed for a regional domain covering Southeast Europe (D10, 7° E – 30° E and 34° N – 49° N).

Both classifications are calculated for 10 and 26 types. In *Figs. 1* and *2,* maps of GWT and JCT with 10 types (GWT10, JCT10) and 26 types (GWT26, JCT26) are shown, respectively. Circulation classification with ten types (*Fig. 1a* and *1b*) includes 1) eight main directional types (W, SW, NW, N, NE, E, SE, and S), 2) one cyclonic type (C), and 3) one anticyclonic type (AC). For the 26 circulation types (*Figs. 2a* and 2*b*), there are two additional groups: 1) eight anticyclonic directional types (AW, ASW, ANW, AN, ANE, AE, ASE, and AS), and 2) eight cyclonic directional types (CW, CSW, CNW, CN, CNE, CE, CSE, and CS).

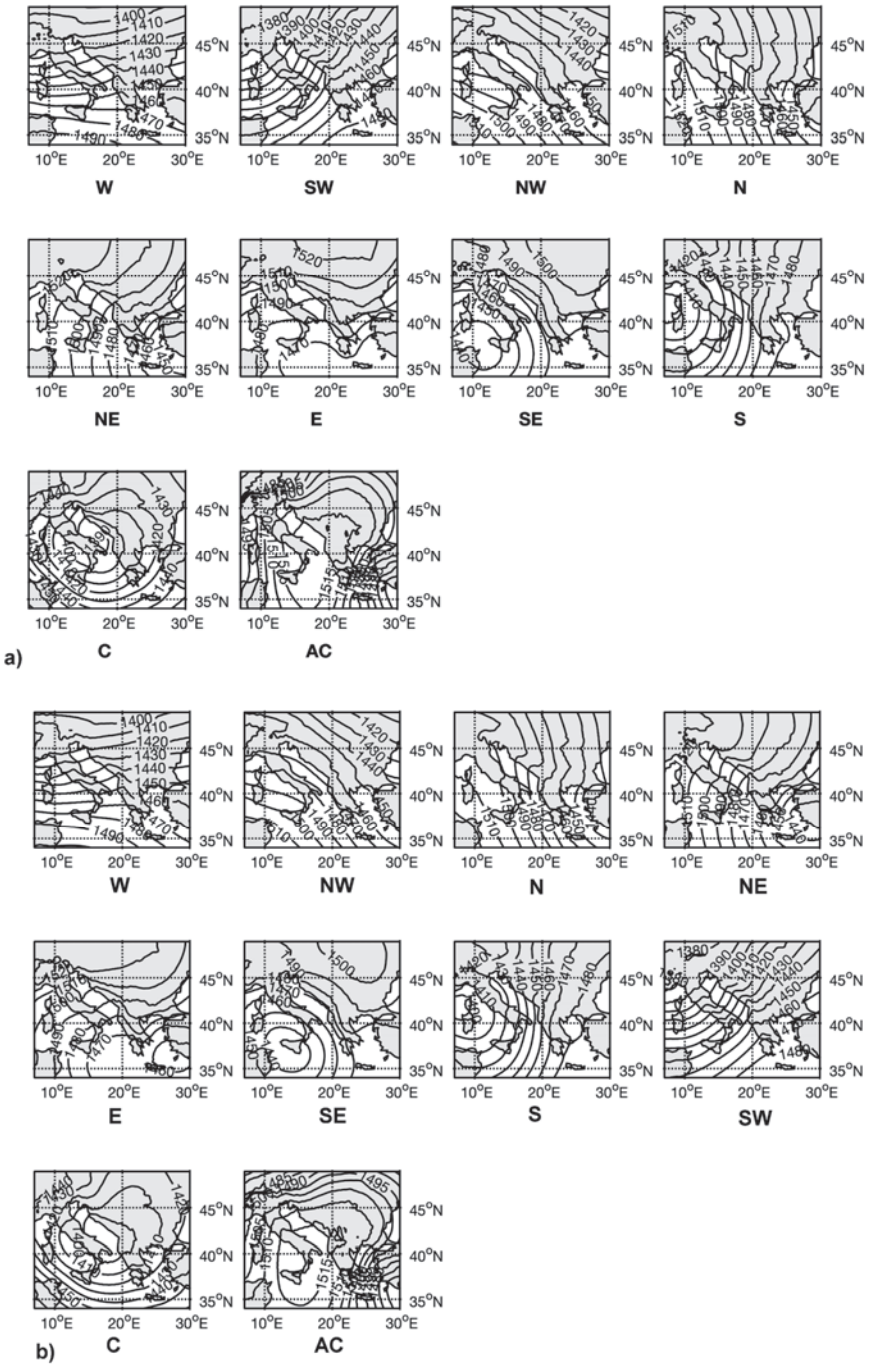


Fig. 1. Maps of the 850 hPa geopotential height of the GWT10 (a) and JCT10 (b) circulation types.

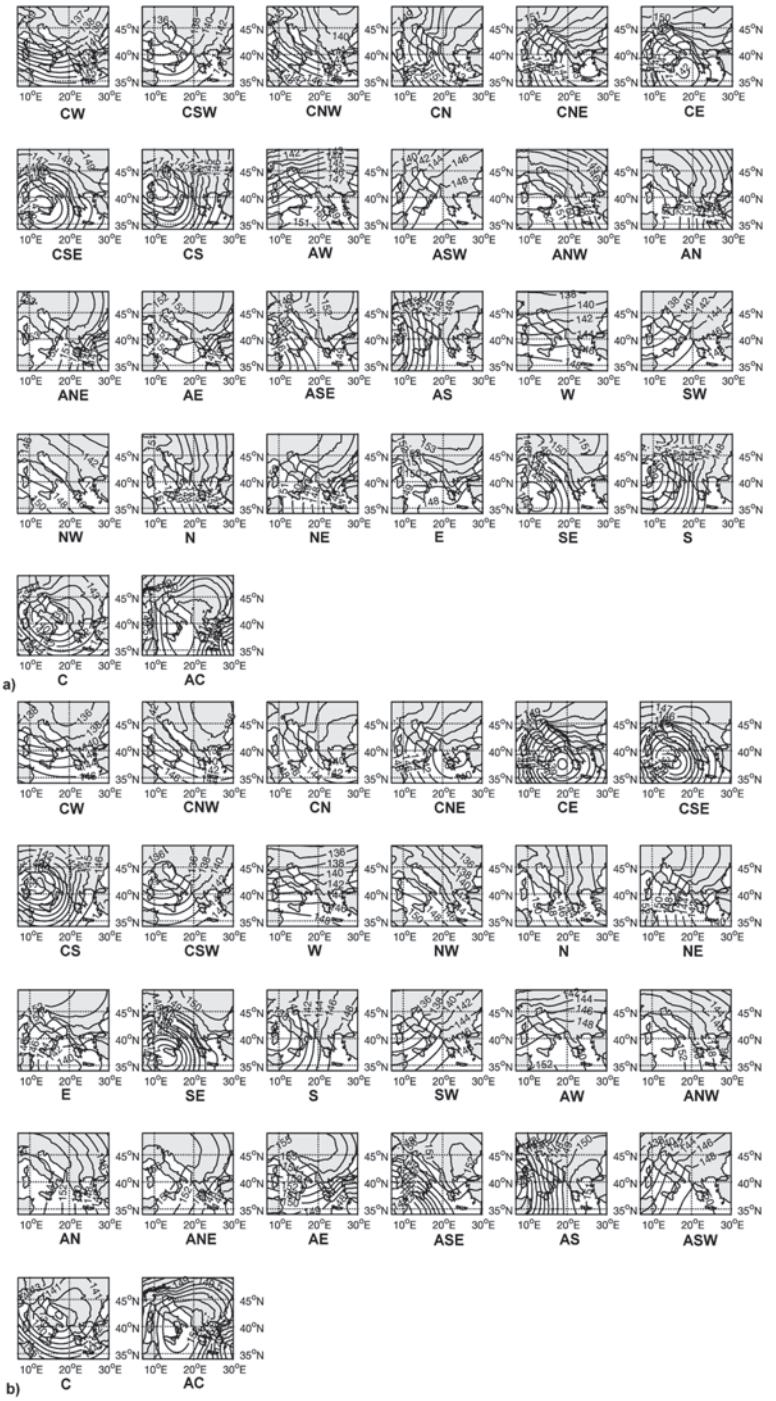


Fig. 2. Maps of the 850 hPa geopotential height of the GWT26 (a) and JCT26 (b) circulation types.

2.2. Manual circulation classification for foehn in Sofia

A manual circulation classification titled Bulgaria Foehn Type (BFT, *Stoev* and *Guerova*, 2020) is developed for foehn situations in Sofia for the period 1975–2014. The BFT classification is produced by an experienced operational forecaster using the National Centres for Environmental Prediction reanalysis data archive (NCEP, www.wetterzentrale.de/topkarten/fsres.2eur.html). Analyzed fields are: 1) surface pressure maps, 2) 500 hPa geopotential height maps (AT500), 3) 850 hPa geopotential height and temperature maps (AT850), and 4) 700 hPa relative humidity maps (AT700). The BFT has four major types: 1) type I: a Mediterranean cyclone that moves west of Bulgaria, 2) type II: cyclogenesis over Hungary and a cold front approaching the western part of the Balkan Peninsula, 3) type III: a cyclone centered over the North Sea, its periphery moves over Central Europe and 4) type IV: a cyclone centered over Scandinavia, moves through the Baltic area towards the southern part of European Russia also known as a *diving* cyclone.

2.3. Surface observations

To determine the days with foehn, observations of 35 years from the central meteorological station of the National Institute of Meteorology and Hydrology in Sofia are analyzed. Synoptic observations are made every 3 hours, at 00, 03, 06, 09, 12, 15, 18, and 21 UTC. Observations are made according to the standards and recommendations of the World Meteorological Organization (WMO). The following meteorological elements are used in this work: 1) wind speed and direction, 2) air temperature and maximum temperature at 2 m, and 3) relative humidity at 2 m. Quantitative criteria proposed by *Hristov* and *Tanev* (1970) are applied for foehn occurrence, such as 1) wind direction between SE and SW and wind speed increase by more than 5 m/s in 3 hours, 2) sharp rise of the temperature by 5 °C or more in 3 hours, and 3) relative humidity decrease with more than 20% in 3 hours.

2.4. Conditional foehn probability

The conditional foehn probability (F) for a selected circulation type is the probability that the foehn will occur given the knowledge of all events for this circulation type (A). If events A and F are dependent, then the probability of the intersection of A and F (the probability that both events occur) is defined by

$$P(A \text{ and } F) = P(A)P(F|A).$$

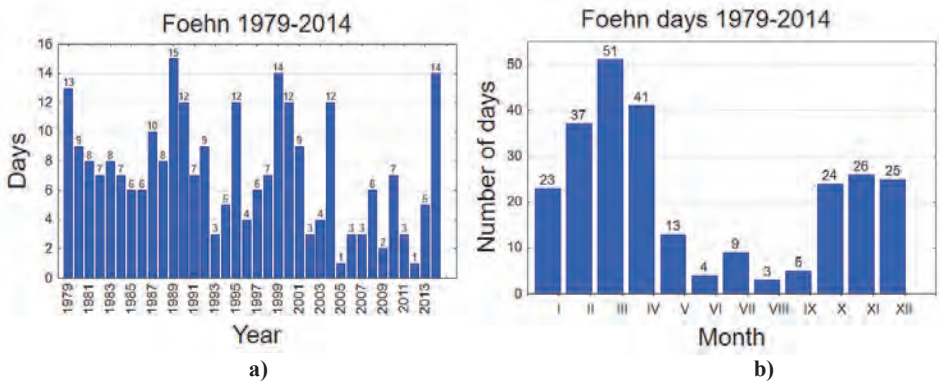
Thus the conditional probability for foehn circulation type $P(F|A)$ is $\in(0,1)$ and is given with

$$P(F|A) = \frac{P(A \text{ and } F)}{P(A)}.$$

3. Results

3.1. Foehn climatology of the period 1979–2014

For the period 1979–2014, there are 261 foehn days in Sofia. The annual foehn climatology presented in *Fig. 3a* shows that the number of foehn days per year ranges from 1 to 15. As reported in *Stoev* and *Guerova* (2020), the average annual number of foehn days is 8.1, 8.3, and 4.5 for the period 1985–1994, 1995–2004, and 2005–2014, respectively. After 2004, a decrease of the average annual number of days with foehn is found and the lowest maximum wind gusts are registered. In Section 3.3, the reason for this decrease is studied in detail. Monthly foehn climatology for the period 1979–2014 is seen in *Fig. 3b*. The largest number of foehn days are in March (51) followed by April (41) and February (37). The period September-May has 92% of the foehn days, while 17% of the foehn days are alone in March. This motivates the detailed discussion of those selected periods in Section 3.2.



a) b)

Fig. 3. a) Annual number of foehn days and b) monthly number of foehn days for the period 1979–2014 for Sofia, Bulgaria.

3.2. Atmospheric circulation patterns associated with foehn in the period 1979–2014

Two objective circulation classifications are used, namely GWT and JCT, with 10 and 26 classification types for the region of Southeast Europe. Due to the topography of the Balkan Peninsula, the circulation types are based on the field of the 850 hPa geopotential height. Mean sea level pressure classification has been computed, but it was difficult to distinguish the predominant synoptic patterns,

therefore it was not used in this study. To facilitate the comparison, the results are presented based on the number of classification types.

3.2.1. Foehn in the period 1979-2014: 10 circulation types

In this section, a comparison between circulation classifications with ten types is offered. It is to be noted, that the comparison is made for all days and foehn days only for the period September-May. This period is selected based on the foehn climatology (*Stoev* and *Guerova*, 2020), which shows a minimum of foehn days in the period June-August of 1975–2014.

For the period 1979–2014, GWT10 gives four main types (not shown), namely: 1) type NW (2747 days), 2) type N (2565 days), 3) type W (2365 days), and 4) type SW (1590 days). A characteristic feature of the type NW is that the cyclone center is located northeast of the Balkan Peninsula. A cold atmospheric front passes over Bulgaria, and the flow is from the northwest. For type N, the air mass advection is from the north, with the cyclonic vortex located east of the Balkans. In types W and SW, the transport of air masses over the Balkan Peninsula is from the west-southwest. For type W, the cyclone center is over Central Europe, and for type SW – over the Alps. As seen from the blue bars in *Fig. 4a*, for the period September-May of 1979–2014, there are four main circulation types, namely: 1) W with 20% of occurrence (1918 days), 2) NW with 19% (1836 days), 3) N with 16% (1555 days), and 4) SW with 14% (1431 days). For March (*Fig. 4b*) the main GWT10 atmospheric circulation types are: 1) W (21%, 244 days), 2) NW (20%, 228 days), 3) N (17%, 128 days), and 4) SW (11%, 129 days).

For the foehn days, GWT10 gives only three main circulation types, namely: 1) W with 5 % of occurrence (132 days), 2) SW with 32% (83 days), and 3) NW with 14% (36 days). A common feature of the atmospheric circulation, which specifies the weather over the Balkan Peninsula to type W and SW, is the cyclonic circulation with warm air advection from W or SW. As seen from the red bars in *Fig. 4a*, for the period September-May of 1979–2014, three main circulation types are leading to foehn in Sofia, namely: 1) W with 50% of occurrence (122 days), 2) SW with 34% (83 days), and 3) NW with 12% (32 days). The main circulation type with foehn in March (red bars in *Fig. 4b*) is W with 60% (31 days). Visual comparison of *Figs. 4a* and *4b* leads to the conclusion that the days with foehn in Sofia are associated with three main GWT10 circulations types W, SW, and NW. It is to be noted, that the N type of GWT10 circulation classification does not lead to foehn.

For classification scheme JCT10, there are four main general atmospheric circulation types for the period 1979–2014 (not shown): 1) type W with 2053 days, 2) type N with 2020 days, 3) type AC with 1906 days, and 4) type NW with 1889 days. The characteristic feature of JCT10 type W is a cyclone situated over Central Europe. For JCT10 type N, the air mass advection is from the north with a cyclone east-southeast of the Balkan Peninsula, over Asia Minor. The JCT10 type NW circulation is a cyclonic field with a cold atmospheric front passing over

the Balkan peninsula and connecting to a cyclone over Belarus and Ukraine. The JCT10 type AC is an anticyclonic circulation over the Balkan Peninsula. As seen from *Fig. 4c* (blue bars), the main JCT10 circulation types for the period September-May of 1979–2014 are: 1) C with 11% (1264 days), 2) AC with 11% (1215 days), 3) E with 11% (1197 days), 4) NE with 11% (1191 days), and 5) type SE with 10% (1066 days). The characteristic feature of JCT10 type C is a Mediterranean cyclone crossing the Balkan Peninsula. For type E, the cyclone is located above the Eastern Mediterranean. For March in 1979–2014 (blue bars on figure 4d), the two main JCT10 types with about 15% of occurrences are C (168 days) and NE (159 days). The SE type is with 11% (119 days), while N, NW, and E types are with 10% and 109, 112, and 116 days, respectively.

For the period 1979-2014, the JCT10 circulation leading to foehn in Sofia is W with 42% of occurrence (109 days) and SW with 31% (80 days). The general characteristic feature of the atmospheric circulation, which determines the weather over the Balkan Peninsula in these types, is cyclonic circulation, the predominant flow of warm air masses from west-southwest with a large pressure gradient. For September-May of 1979–2014 (*Fig. 4c*), the main circulation types leading to foehn in Sofia are type W with 40% of occurrence (101 days) and type SW with 30% (78 days). The main JCT10 circulations leading to foehn for March of 1979–2014 (red bars on *Fig. 4d*) are type W with 60%, type SW with 20%, and type C with 10%.

From *Figs. 4a* and *c*, an interesting comparison can be made between GWT10 and JCT10 for the foehn cases. Both classifications have the largest number of foehn days for type W but with a 10% difference. For SW and NW types of GWT10 and JCT10, this agreement is within 5%. For JCT10, the fourth significant type is the C with about 10% difference. This pattern is best seen for the visual comparison of the red bars in *Figs. 4b* and *d*. In *Table 1,* conditional probability for the foehn event for circulation classification GWT10 and JCT10 is presented for the period 1979–2014. Classification types W and SW have the largest probability for foehn. However, there is a difference for the third major circulation type, which is NW for GWT10 and C for JCT10. NW type for JCT10 comes as the fourth most probable type.

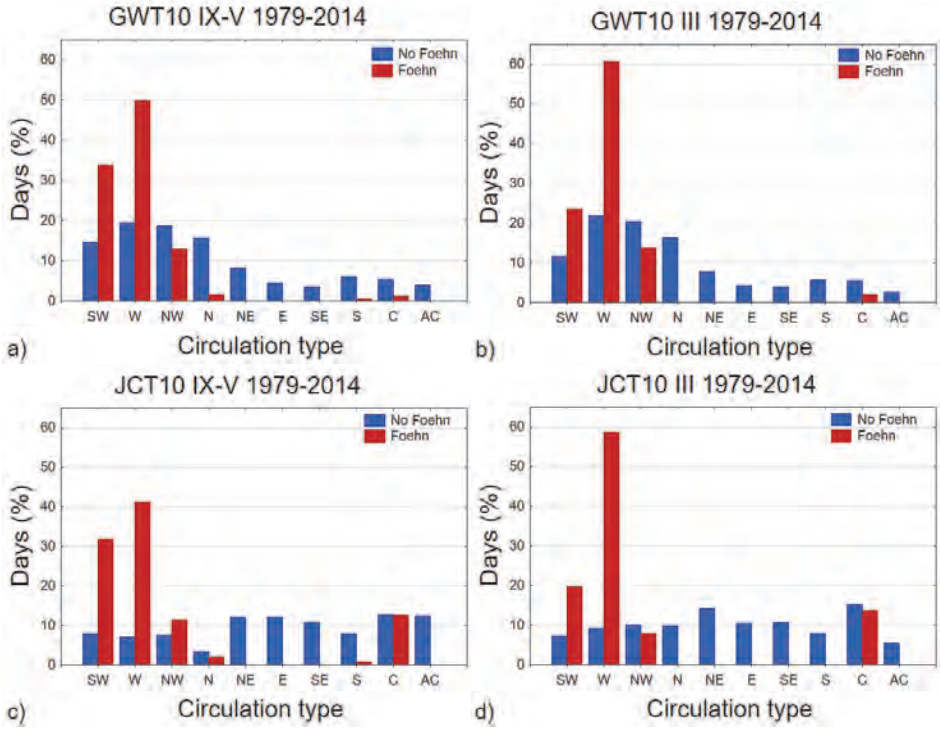


Fig. 4. Relative frequencies of circulation types for all days (blue bars) and foehn days (red bars). GWT10 a) September-May of 1979–2014 and b) March of 1979–2014. JCT10 c) September-May of 1979–2014, and d) March of 1979–2014.

Table 1. Conditional probability of the foehn for circulation classification GWT10 (column 2) and JCT10 (column 3) for the period 1979–2014

Type	**P(GWT10)**	**P(JCT10)**
SW	0.052	0.056
W	0.056	0.053
NW	0.013	0.016
N	0	0.004
NE	0	0
E	0	0
SE	0	0
S	0.002	0.003
C	0.005	0.022
AC	0	0

3.2.2. Foehn in the period 1979–2014: 26 circulation types

For the period 1979–2014, GWT circulation classification with 26 types for the 850 hPa geopotential height gives three significant types: 1) ANW with 1348 days, 2) AN with 1220 days, and 3) AW with 970 days. A characteristic feature of AN and ANW types is that the main cyclonic center is above Northern Europe, and an anticyclone is above the Western and Central Mediterranean, respectively. For AW type, the cyclone is over Western Europe and the anticyclone is over the Central Mediterranean. As seen from the blue bars on *Fig. 5a*, for the period September-May of 1979–2014, three main circulation types stand out: ANW with 9% (882 days), AW with 8% (756 days), and AN with 7% (710 days). GWT26 for March gives three main types of circulation: ANW with 10% (107 days), W with 8% (94 days), and AW with 8% (87 days). For GWT26 circulation type W, the cyclone is located over Central Europe and the Baltics (not seen in *Fig. 2a*), with westerly flow over the Balkan Peninsula (seen in *Fig. 2a*).

On the other side, the synoptic conditions leading to foehn in Sofia give four main GWT26 types: 1) CW with 68 days (26%), 2) W with 47 days (18%), 3) CSW with 44 days (17%), and 4) SW with 35 days (14%). A common feature of the atmospheric circulation, which determines the weather over the Balkan Peninsula, is cyclonic circulation with warm air mass flow from the southwest-west. During the September-May period, the largest number of foehn days is registered, and four main GWT26 types are leading to foehn: CW, W, CSW, and SW with 25%, 19%, 18%, and 14%, respectively. As seen from red bars in *Fig. 5b*, for March only, the maximum foehn frequency of 31% is for GWT26 circulation type W, 16% is for CW, and 14% is for CSW and AW.

JCT26 circulation types registered in the period 1979–2014 are as follow: 1) 1906 days (15 %) for type AC, 2) 1402 days (11%) for type C, 3) 1380 days (10%) for type W, and 4) 1330 days (10%) for type N. For the September-May period (blue bars in *Fig. 5c*), type AC has 1316 days (13%), type C – 1172 days (12%), and type W – 1051 days (11%). As seen from *Fig. 5d*, for March only, type W has 156 days (14%), type C – 144 days (13%), and type AC – 106 days (9%).

JCT26 circulation classifications for foehn days is: 30% (77 days) for type W, followed by 20% (53 days) for type SW, 12% (31 days) for type C, and 11% (29 days) for type CW. As in *Fig. 5c*, for September-May type, W has 29% of occurrence (72 days), type SW – 21% (52 days), type C – 13% (31 days), type CW – 11% (27 days), and type CSW – 11% (26 days). For March (*Fig. 5d*), type W is with 51% (26 days), type C with 14% (7 days), and type SW with 12% (6 days).

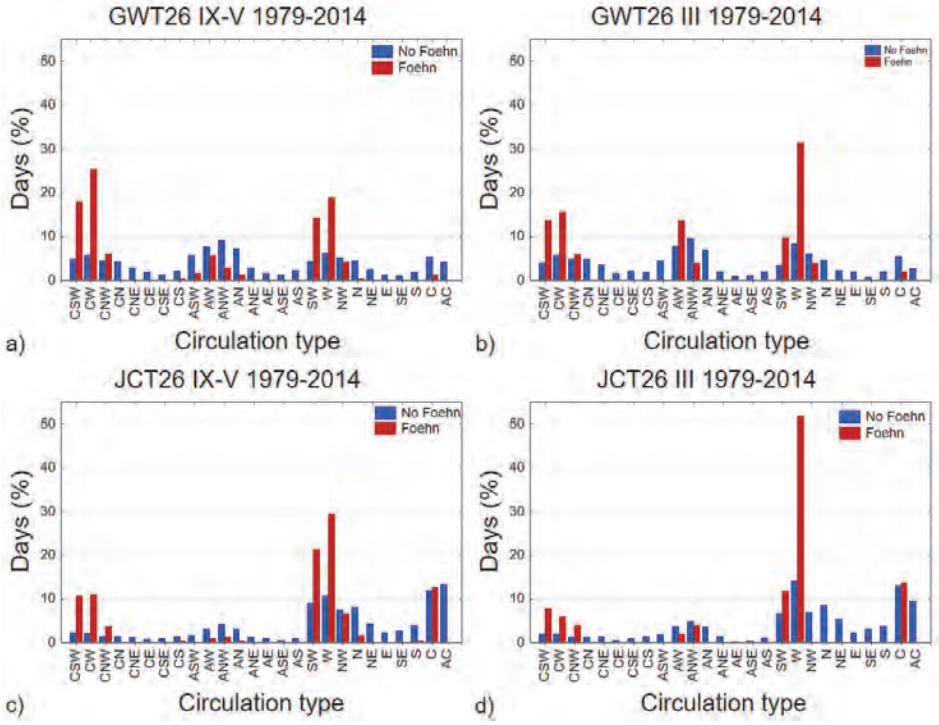


Fig. 5. Relative frequencies of circulation types for all days (blue bars) and foehn days (red bars). GWT26 a) September-May of 1979–2014 and b) March of 1979–2014. JCT26 c) September-May of 1979–2014 and d) March of 1979–2014.

To conclude, the GWT26 classification shows the foehn occurrence for circulation types W, SW, CW, and CSW (red bars in *Figs. 5a* and *b*). A comparison between *Figs. 5a* and *b* show that GWT26 has a similar distribution between eight cyclonic, eight anticyclonic, and eight directional circulation types (blue bars), with dominant SW, W, NW, and N types. This, however, is not the case for JCT26 (*Figs. 5c* and *d*). The prevalent circulation types for September-May and March (blue bars on *Figs. 5c* and *d*) are W, SW, NW, N, C, and AC types. The remaining eight cyclonic and eight anticyclonic types are with contributions less than 5%. JCT26 circulation types that result in foehn are W, SW, C, CSW, and CW with 29, 21, 13, 11, and 11%, respectively, for the period September-May of 1979–2014. The same holds for March, but with the JCT26 W type dominating the foehn days in 51% of the cases. For circulation classification, GWT26 conditional foehn probability (column 2 in *Table 2*) is above 0.05 for types CW, CSW, W, and SW. As seen from *Table 2* for JCT26 with conditional probability above 0.05, types (CW, CSW, W and SW) are the same. There are,

however, differences between GWT26 and JCT26 type CNW with conditional probability 0.26 and 0.05, respectively. Notably the largest difference is for type C, where GWT26 probability is 0.002 versus the 0.022 value for JCT26.

Table 2. Conditional probability of the foehn type (column 1) for circulation classification GWT26 (column 2) and JCT26 (column 3)

Type	**P(GWT26)**	**P(JCT26)**
CSW	0.088	0.107
CW	0.103	0.106
CNW	0.026	0.050
CN	0	0
CNE	0	0
CE	0	0
CSE	0	0
CS	0.005	0.007
ASW	0.006	0
AW	0.017	0.007
ANW	0.006	0.005
AN	0.004	0.006
ANE	0	0
AE	0	0
ASE	0	0
AS	0	0
SW	0.077	0.053
W	0.064	0.056
NW	0.015	0.015
N	0.001	0.003
NE	0	0
E	0	0
SE	0	0
S	0	0.002
C	0.002	0.022
AC	0	0

3.2.3. Comparison of objective and manual circulation classifications

In this section comparison between objective and manual foehn classifications is presented. The motivation for this is to select the best objective classification that corresponds to the manual classification, and thus, benefits from the experienced forecaster's analysis. Clearly, manual classification for other extreme events will be very time consuming, thus objective classifications are the best possible option for operational forecasting. *Table 3* presents the comparison between the manual circulation classification BFT and objective classifications GWT and JCT with 10 and 26 types. Here again, for JCT10 and JCT26, the C type stents out for BFT-I, BFT-II, and BFT-III. In addition, the 26 circulation types differ from the 10 types by including CW and CSW types.

Table 3. BFT and corresponding GWT10 (column 2), JCT10 (column 3), GWT26 (column 4), and JCT26 (column 5) circulation types

BFT	GWT10 (%)	JCT10 (%)	GWT26 (%)	JCT26 (%)
I	W (22%) SW (22%) NW (5%)	W (16%) SW (22%) NW (5%) C (7%)	CW (14%) CSW (12%) SW (9%) W (8%) C (1%)	CW (6%) CSW (9%) SW (13%) W (10%) C (7%) NW (3%)
II	W (22%) SW (4%) NE (6%)	W (9%) SW (4%) NW (5%) C (3%)	CW (8%) CSW (2%) SW (1%) W (2%) CNW (3%)	CW (3%) CSW (1%) SW (3%) W (6%) C (3%) NW (3%)
III	W (15%) SW (6%) NW (2%)	W (16%) SW (5%) NW (1%) C (1%)	W (7%) CW (4%) CSW (2%) SW (3%) AW (4%)	CW (2%) SW (5%) W (14%) C (1%) AW (1%)
IV	NW (1%) N (1%)	NW (1%) N (1%)	ANW (1%) AN (1%)	AN (1%)

The foehn days for March of 1979–2014 are used to compare the objective circulation classifications GWT and JCT with the in-house developed manual classification BFT. It is to be noted, that the BFT has four main circulation types, and their corresponding GWT and JCT types are shown in *Table 3*. The largest number of foehn days is with circulation types BFT-III and BFT-I, which are

associated with the cold front approaching Bulgaria and the Mediterranean cyclone, respectively. As seen from *Table 4*, the GWT10 and JCT10 objective classifications tend to have large differences with the BFT types. For example, BFT-I has 16 days while GWT10 has 1 day, but JCT10 has 17 days. BFT-II, however, has 12 foehn days, and GWT10 agrees, while JCT10 has none. A much better agreement is obtained between BFT and GWT26, and BFT and JCT26. JCT26 has the best agreement with BFT for BFT-II and BFT-IV types. GWT26 has 11 days for circulation type corresponding to BFT-IV.

Table 4. Comparison of number of the foehn days from manual classification Bulgaria Forecast Type (BFT column 2) and corresponding GWT10 (column 3), JCT10 (column 4), GWT26 (column 5) and JCT26 (column 6) for March 1979-2014.

Type	BFT	GWT10	JCT10	GWT26	JCT26
BFT-I	16 (31%)	1	17	13 (25%)	11 (22%)
BFT-II	12 (23%)	12	-	11 (22%)	11 (22%)
BFT-III	21 (42%)	31	30	16 (31%)	26 (51%)
BFT-IV	2 (4%)	7	4	11 (22%)	3 (6%)
Total	51	51	51	51	51

3.3. Foehn circulation types in the periods 1995–2004 and 2005–2014

As a next step, we investigated the decrease in the number of foehn days from 83 in the period 1995–2004 to 45 in 2005–2014. The number of foehn days for circulation types GWT10 and JCT10 are presented in *Table 5*. For each classification type, the percentage decrease/increase is calculated. For the period 2005–2014, both GWT10 and JCT10 classifications show that the foehn days decrease due to the decreased circulation types W by -65% and -80%, respectively and NW with -25 and -27%, respectively. For GWT10, the circulation type SW also decreased by -25% for the 2005–2014 period. The percentage is calculated considering the number of days from the period 1995–2004 as 100%, subtracting the number of days from the period 2005–2014, and then dividing by the number of days from the period 1995–2004.

Table 5. Circulation types for the foehn days for the periods 1995–2004 and 2005–2014. GWT10 1995–2004 (column 2) vs 2005–2014 (column 3) and percent of decrease/increase (column 4). JCT10 1995–2004 (column 5) vs 2005–2014 (column 6) and percent of decrease/increase (column 7)

Type	**GWT10 1995–**	**GWT10 2005–**	**GWT10 %**	**JCT10 1995–**	**JCT10 2005–**	**JCT10 %**
SW	24	18	-25%	19	19	
P(SW)	0.056	0.039		0.048	0.047	
W	46	16	-65%	45	9	-80%
P(W)	0.046	0.026		0.076	0.017	
NW	12	9	-25%	11	8	-27%
P(NW)	0.015	0.012		0.019	0.017	
N	-	1	+100%	-	1	+100%
P(N)		0.001			0.002	
C	1	1		7	8	+14%
P(C)	0.006	0.005		0.019	0.018	
AC				1	-	
P(AC)				0.002		

Comparison of circulation types GWT26 and JCT26 for the periods 1995–2004 and 2005–2014 is shown in *Table 6*. Both GWT26 and JCT26 confirm the foehn days decrease for the 2005–2014 period due to a decrease of circulation types W by -76% and -82%, respectively and NW with -100% and -71%, respectively. Circulation type CW decreased for the 2005–2014 period by -57% for GWT26 and by -73% for JCT26. CSW and AW also show a decrease for both GWT26 and JCT26.

In *Tables 5* and *6* the conditional foehn probability for circulation classification GWT and JCT with 10 and 26 types are presented, and they also confirm the decrease of the above mentioned types for the period 2005–2014.

Table 6. Circulation types for the foehn days for the periods 1995-2004 and 2005–2014. GWT26 1995–2004 (column 2) vs 2005–2014 (column 3) and percent of decrease/increase (column 4). JCT26 1995–2004 (column 5) vs 2005–2014 (column 6) and percent of decrease/increase (column 7)

Type	GWT26 1995–	GWT26 2005–	GWT26 %	JCT26 1995–	JCT26 2005–	JCT26 %
CSW	15	10	-33%	8	4	-50%
P(CSW)	0.110	0.069		0.121	0.057	
CW	23	10	-57%	11	3	-73%
P(CW)	0.112	0.057		0.132	0.041	
CNW	5	4	-11%	4	3	-25%
P(CNW)	0.04	0.021		0.059	0.052	
ASW	-	2	+100%			
P(ASW)		0.011				
AW	6	2	-67%	1	-	-33%
P(AW)	0.022	0.008		0.009		
ANW	2	5	+60%	-	3	+100%
P(ANW)	0.005	0.015			0.022	
AN	-	1		-	1	
P(AN)		0.003			0.009	
SW	9	6	-33%	11	15	+36%
P(SW)	0.08	0.046		0.040	0.054	
W	17	4	-76%	33	6	-82%
P(W)	0.08	0.022		0.084	0.017	
NW	5	-	-100%	7	2	-71%
P(NW)	0.024			0.212	0.007	
C	1	-		7	8	+14%
P(C)	0.006			0.019	0.018	
AC	-	1		1	-	
P(AC)		0.006		0.002		

3.4. Changes in synoptic-scale circulation in Southeast Europe in three decades

As a next step, we investigated the general atmospheric circulation types over Southeast Europe for three decades: 1985–1995, 1995–2004, and 2005–2014. *Figs. 6a* and *b* show GWT10 and JCT10 for each decade separately. It is seen that for GWT10, there are four main types with more than 10% of the days: NW, N, W, and SW. Interestingly, for the 2005–2014 period there is a decrease of NW and W types in the range of 1–2.5%. It is to be noted, that the two types also show a reduction for the foehn days as reported in Section 3.3. For the 2005–2014

period, the increase of types C, E, and SW have a compensating effect for GWT10. A decrease of W and NW types between 2 and 4% is also seen for JCT10 classification for the 2004–2015 period. This is compensated by a 2% increase of C and 1% increase of AC type.

For the period 2005–2014, GWT26 classification also shows a decrease of the W type and the second type with reduction is ANW. In *Figs. 7a* and *7b* an increase is seen in the percentage of days with the C and AC types for both the GWT26 and JCT26 classifications. For JCT26, there is about a 2% decrease in the number of days with W, NW, and ANW.

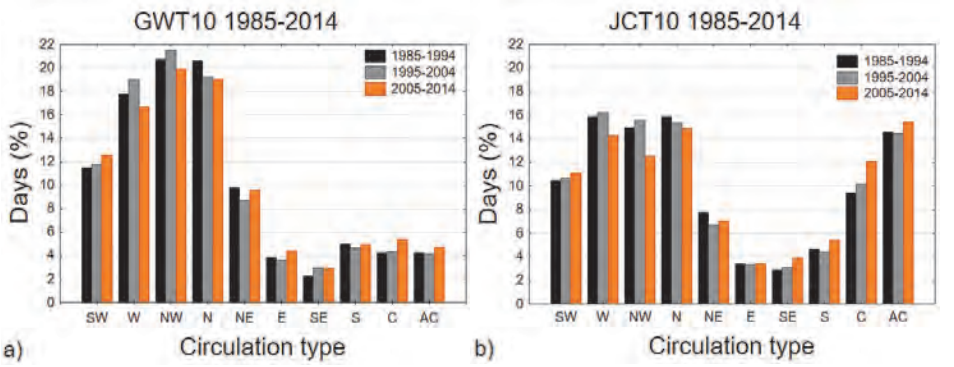


Fig. 6. Relative frequencies of circulation types for the periods 1985–1994 (black bars), 1995–2004 (grey bars), and 2005–2014 (orange bars) for a) GWT10 and b) JCT10.

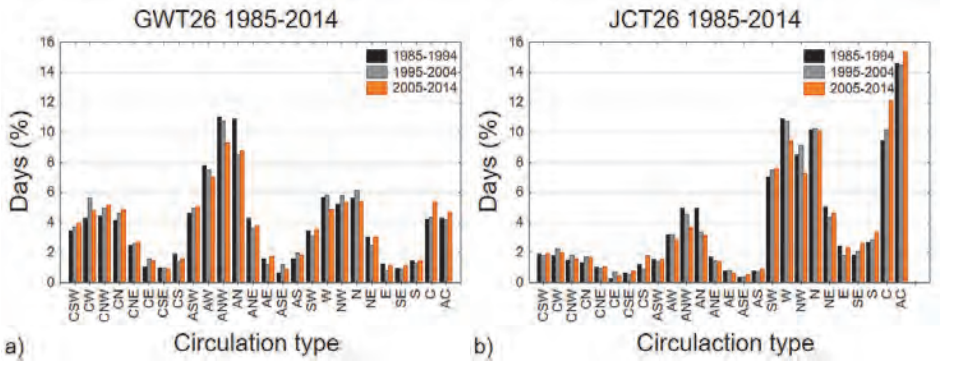


Fig. 7. Relative frequencies of circulation types for the periods 1985–1994 (black bars), 1995–2004 (grey bars), and 2005–2014 (orange bars) for a) GWT26 and b) JCT26.

4. Conclusion

In this study, two threshold-based objective circulation classifications (GWT and JCT) and one manual classification (BFT) were applied for 261 foehn days in the Sofia valley in the period 1979–2014. Threshold-based objective classifications are selected as they are based on the concept of subjectively pre-defined types similar to the manual classification method. Two variants of the objective classifications are selected with 10 (GWT10, JCT10) and 26 types (GWT26, JCT26). A comparison between GWT10 and JCT10 types for the foehn days in the period September-May gives the largest number for types SW, W, and NW. However, they differ by 10% for SW type and 5% for W and NW types. GWT26 confirms the foehn occurrence for circulation types W (25%), SW (19%), CW (18%), and CSW (14%). JCT26 circulation types for foehn days are W (29%), SW (21%), C (13%), CSW (11%), and CW (11%). It is to be noted, that GWT26 has evenly distributed cyclonic, anticyclonic, and directional circulation types with dominant SW, W, NW, and N types in each group. On the other hand, for JCT26, the dominant circulation types are W, SW, NW, N, C, and AC. Eight cyclonic and eight anticyclonic JCT26 types occur in less than 5% of the days. For the foehn days in March of 1979–2014, a comparison between objective circulation classifications and manual classification show a tendency for GWT10 and JCT10 to underestimate or overestimate the BFT types. A much better agreement is obtained for GWT26, and the best agreement is with JCT26.

A previous study showed a decrease in the average annual number of days with foehn in Sofia after 2004. A comparison between the periods 1995–2004 and 2005–2014 gives a decrease of circulation types W by -65% (GWT10) and -80% (JCT10) and NW with -25% (GWT10) and -27% (JCT10). GWT26 and JCT26 confirm the reduction of the foehn days due to the decrease of circulation types W by -76% (GWT26) and -82% (JCT26) and NW with -100% (GWT26) and -71% (JCT26). A general circulation analysis for the period 1985–2014 confirms the decrease of W and NW circulation types for all days in the period 2005–2014.

The findings of this study show that the objective circulation classifications can be used to quantify the foehn occurrence in the Sofia valley. This work is the first step towards developing an objective foehn warning system for operational weather forecasting at the National Institute of Meteorology and Hydrology in Bulgaria.

Acknowledgments: We are very grateful to the National Institute of Meteorology and Hydrology for providing the data from synoptic station Sofia. The COST 733 core group provided the synoptic catalogues and the software used in this study. This research has been supported by the Estonian Research Council (grant no. PSG202).

References

Bartholy, J., Pongrácz, R., and *Pattantyús-Ábrahám, M.,* 2009: Analyzing the genesis, intensity, and tracks of western Mediterranean cyclones. *Theor. Appl. Climatol. 96,* 133–144.

Van Bebber, W., 1891: Die Zugstraßen der barometrischen Minima. *Meteor. Zeit. 8,* 361–366. (In German)

Beck, C., 2000: Zirkulationsdynamische Variabilität im Bereich Nordatlantik-Europa seit 1780. *Würzburger Geographische Arbeiten 95*. (In German)

Beck, C., Jacobeit, J., and *Jones, P.D.*, 2007: Frequency and within-type variations of large-scale circulation types and their effects on low-frequency climate variability in Central Europe since 1780. *Int. J. Climatol. 27,* 473–491.

Blagoev, H., 1961: Patishta na Sredizemnomorskite Tsikloni, *Hydrol. Meteorol*. **1:** 43–51. (in Bulgarian)

Bocheva, L., Georgiev, C.G., and *Simeonov, P.,* 2007: A climatic study of severe storms over Bulgaria produced by Mediterranean cyclones in 1990–2001 period. *Atmos. Res. 83,* 284–293.

Catrina, O., Ştefa, S., and *Crăciun, C.,* 2019: Objective identification of Mediterranean cyclones and their trajectories towards Romania. *Meteorol. Appl. 26,* 429–441.

Hersbach, H., Bell, B., Berrisford, P., et al., 2020: The ERA5 global reanalysis. *Quart. J. Roy. Meteorol. Soc. 146*, 1999–2049. https://doi.org/10.1002/qj.3803

Hristov, P. and *Tanev, A.,* 1970: Nauka i izkustvo, Sofia. (In Bulgaarian)

Jansa, A., Genoves, A., Picornell, M., Campins, J., Riosalido, R., and *Carretero, O.,* 2001: Western Mediterranean cyclones and heavy rain. part 2: Statistical approach. *Meteorol. App. 8*, 43–56.

Jenkinson, A.F. and *Collison, F.P.,* 1977: An initial climatology of gales over the North Sea. Synoptic Climatology Branch Memorandum 62, Meteorological Office, Bracknell.

Lamb, H.H., 1972. British Isles weather types and a register of the daily sequence of circulation patterns 1861–1971. *Geophys. Memoir. 116,* 85L.

Marinova, T., Bocheva, L., and *Sharov, V.,* 2005: On some climatic changes in the circulation over the mediterranean area. *Időjárás 109*, 55–67.

Martinov, M., 1967: Niakoi osobednosti na tsiklogeneza v raiona na Sredizemno more i Balkanskia poluostrov *Hydrol. Meteorol. 6,* 7–21. (in Bulgarian)

Martinov, M., 1983: Synoptical and statistical processing of historical data for Mediterranean depression. WMO, PSMP Report Series 3: 121–144.

Neykov, N., Trifonova, L., Gospodinov, I., and *Neytchev, P.,* 2016: Circulation types and associated precipitations over Bulgaria. In (Eds. *Tveito, O.E., Huth, R., Philipp, A., Post, P., Pasqui, M., Esteban, P., Beck, C., Demuzere, M., Prudhomme, C.*) COST Action 733: Harmonization and application of weather type classifications for European Regions; final scientific report. Augsburg, University of Augsburg. 270–276.

Nikolov, D., Beck, C., Philipp, A., and *Neychev, P.*, 2016: Circulation types associated with freezing precipitation over Bulgaria. In (Eds. *Tveito, O.E., Huth, R., Philipp, A., Post, P., Pasqui, M., Esteban, P., Beck, C., Demuzere, M., Prudhomme, C.*) COST Action 733: Harmonization and application of weather type classifications for European Regions; final scientific report. Augsburg, University of Augsburg. 277–280.

Peiter, E., 1975: Handbook for forecasters in the Mediterranean, Tech. Rep. 344, Environmental Prediction Research Facility, Naval Postgraduate School, Monterey, California.

Philipp, A., Beck, C., Huth, R., and *Jacobeit, J.,* 2016: Development and comparison of circulation type classifications using COST733 dataset and software. *Int. J. Climatol. 36,* 2673–2691. doi: 10.1002/joc.3920

Pisarski, A., 1955a: Sredizemnomorskite Tskiloni i tyahnoto vilyanie na vremeto v Bulgaria, part I, *Hidrologia i Meteorologia .5*, 33-50. (in Bulgarian)

Pisarski, A., 1955b: The Mediterranean cyclones and their influence on the weather in Bulgaria. *Hydrol. Meteorol.* **6**: 3–15. (in Bulgarian).

Popova, T., Runcanu, T., Tanczer, T., and *Sharov, V.*, 1975: Sredizemnomorskie t̂siklony v pole oblachnosti. Gidrometeoizdat, Leningrad. (in Russian)

Radinovic, D., 1987: Mediterranean cyclones and their influence on the weather and climate. WMO, PSMP Report Series 24, 131.

Stanchev, K., 1954: The South cyclones, their passage over the Balkan Peninsula and the weather in Bulgaria. *Hydrol. Meteorol. 5,* 19–39. (in Bulgarian)

Stoev, K. and *Guerova, G.,* 2020: Foehn classification and climatology in Sofia for 1975–2014. *Időjárás 124*, 483–497.

Tveito, O.E., Huth, R., Philipp, A., Post, P., Pasqui, M., Esteban, P., Beck, C., Demuzere, M., and *Prudhomme, C.,* 2016: COST Action 733: Harmonization and application of weather type classifications for European Regions; final scientific report.

Yarnal, B., 1993: Synoptic Climatology in Environmental Analysis. Belhaven Press: London.

DOI:10.28974/idojaras.2022.4.6

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 126, No. 4, October – December, 2022, pp. 567–582

Estimation of seasonal and annual river flow volume based on temperature and rainfall by multiple linear and Bayesian quantile regressions

Sajjad Modabber-Azizi, **Meysam Salarijazi***, and **Khalil Ghorbani**

Department of Water Engineering
Faculty of Water and Soil Engineering
Gorgan University of Agricultural Sciences and Natural Resources
Golestan Province, Gorgan, Shahid Beheshti, RCVQ+6V2, Iran

**Corresponding author E-mail: meysam.salarijazi@gau.ac.ir*

(Manuscript received in final form June 1, 2021)

Abstract— Investigation of river flow volume in different conditions as a function of temperature and rainfall variables can be quite effective in understanding the hydrological and hydro-climatic conditions of the watershed. Multiple linear regression models were applied in estimating river flow in several studies due to their straightforwardness and appropriate interpretation of results. In this study, to overcome the limitations of the multiple linear regression model, the Bayesian quantile regression model was used to estimate the river flow volume as a function of rainfall and temperature, and the results were compared. The data and information used for the Qareh-Sou basin in northern Iran are of substantial environmental and socio-economic importance. Five data series, including spring, summer, autumn, winter, and annual series, were created and used for this study. It was found that the Bayesian quantile regression model has considerable flexibility to model the volume of flow for different quantiles, predominantly upper and lower quantiles, and can be used to model high and low flows. With increasing the values of quantiles, a limited decreasing pattern in the effect of rainfall on the volume of flow was identified, which can be due to increasing the effect of other factors in the formation of extreme flows of the river. For summer data in high quantiles, the effect of rainfall on river flow volume shows an increasing pattern. This pattern is different from the other studied series, which may be due to the low base flow in summer. The results confirm that the application of Bayesian quantile regression compared to multiple linear regression leads to much more valuable information on the impact of rainfall and temperature on river flow volume.

Key-words: Qareh-Sou basin, modeling, quantile, extreme events

1. Introduction

One of the critical components of the hydrological cycle is river flow (*Ansarifar et al*., 2020a). This component can interact with other components such as groundwater (*Ansarifar et al*., 2020b). Surface water, which is the result of rainfall-runoff responses in a basin, is a potential source that, if properly managed, can meet agricultural (*Steinfeld et al*., 2020), industrial (*More et al*., 2020), and environmental (*Karimi et al*., 2021) demands. The increase in water demand in different regions, especially in arid and semi-arid regions, shows the need for optimal water resources management. Therefore, the estimation of river flow resulting from climatic factors is the basis for studying many different plans to develop and exploit water resources (*Bahrami et al*., 2019). Estimating river flow in a basin is a complex one, in which human knowledge, understanding, and knowledge of the physical laws governing it are incomplete. Several factors affect the river flow pattern in the basin area (*Salarijazi* and *Ghorbani*, 2019). These factors include topographic features, river morphology, rainfall dynamics, temperature, and human activities. Estimating river flow under the influence of hydroclimatic variables is possible using different approaches (*Mudbhatkal et al.,* 2017). In general, there are two major approaches to modeling river flow. The first approach is knowledge-based, known as modeling, based on the basin area's characteristics and physical laws (*Kavian et al*., 2020). This approach requires a wide range of different information and data that, in most cases, may not be available (*Bahremand et a*l., 2021). In most parts of the world, especially in developing countries, this approach is limited. The second approach is data-driven, which involves analyzing the data set recorded over a historical period (*Chadalawada et al*., 2017). There is a need for more limited data and information in this approach than the first approach (*Nourani et al.,* 2019). The use of data-driven models has developed in recent years (*Sezen et al*., 2019). Although Modeling with a data-driven approach may not be sufficient to interpret the physical processes within the basin, it can accurately estimate the amount of river flow (*Mishra et al*., 2018: The multiple linear regression model is one of the basic and well-known models in the data-driven approach (*Niedzielski et al*., 2019). This model has several advantages. The multiple linear regression model is fast and straightforward and leads to specific mathematical equations. Also, by interpreting these equations, we can understand the effect of each of the model inputs on the output (*Cho* and *Lee*, 2018: A multiple linear regression model has been used in meteorology, climatology, hydrology, and water resources due to the stated advantages (*Niu et al*., 2019: Using data from 33 catchments in Iowa, *Schilling* and *Walter* (2005) used multiple linear regression modes to predict total flow, base-flow, and flood flow. The results of this study indicate a significant effect of rainfall over other input variables of the model. A multiple linear regression model was developed using principal component analysis and discrepancy ratio modified by *Noori et al*. (2010). This study showed that the

developed model has better performance than the standard model for predicting river flow. The multiple linear regression model was developed using bootstrap resampling and wavelet analysis, and was evaluated to predict the daily flow of the river. This study indicates that it is substantial that the developed model has better accuracy in estimating the peak flow of river flow in flood conditions than the standard model of multiple linear regression (*Sehgal et al*., 2014). *Latt* and *Wittenberg* (2014) studied artificial neural networks and stepwise multiple linear regression models to simulate the flow of the Chindwin River in Myanmar. They showed that the multiple linear regression model has good accuracy in predicting river flow, but it is weak in estimating extreme values. The results of a study in India showed that the multiple linear regression model could be used as a suitable option to assemble different hydrological models to predict river runoff (*Kumar et al.,* 2015). Using 14 years of Wainganga River runoff data in India, the efficiency of the multiple linear regression method to simulate river flow using rainfall and temperature data was studied. The study results indicate the appropriate efficiency of the multiple linear regression model in rainfall-runoff modeling and the effect of different inputs on increasing the accuracy of the results (*Patel et al.,* 2016). *Tsakiri et al*. (2018), in their research on river flow modeling in the Mohawk River in New York, concluded that the use of a multiple linear regression model has the advantage that it can lead to a physical interpretation of the river flow time series. He also pointed out that the development of a standard model can significantly improve the model's accuracy. *Popat et al.* (2020) used a multiple linear regression model to predict river flow in the Wernersbach catchment, Germany. In this study, rainfall, runoff, and soil moisture information were used for modeling. The results show that the multiple linear regression model is not accurate enough to predict extreme flows.

The quantile regression model has been considered in meteorology, climatology, and hydrology in recent years (*Nguyen et al*., 2021). This model has far fewer limitations than the multiple linear regression model (*Hossain et al.,* 2021). *Shiau* and *Chen* (2015) used the quantile regression model to estimate the uncertainty of river sediment load as an appropriate model. *Sa'adi et al.* (2017) used the quantile regression model to estimate changes in the variable probability distribution function of rainfall in Sarawak, Malaysia. They described this method as a suitable tool in this field. In another study, the quantile regression method was used to investigate changes in extreme rainfall in South Korea. Based on the results, the study areas were classified according to the type of changes, and the use of this method was recommended to classify rainfall changes (*Uranchimeg et al*., 2020). In another study, the quantile regression model was used to predict dissolved oxygen concentrations considering land use and soil cover (*Ahmed* and *Lin,* 2021).

A review of the research using the multiple linear regression model to predict river flow shows that this model has relatively good accuracy for predicting the mean values of river flow. At the same time, it should be developed for extreme

flow modeling. The Bayesian quantile regression model has also been developed to be suitable for the modeling of extreme flows. This research investigates the Bayesian quantile regression model in predicting river flow volume in different time scales and compares it with the multiple linear regression model. Moreover, the impacts of inputs and modeling results in different standard and extreme flow conditions are compared and analyzed for better interpretation. The Qareh-Sou River in northern Iran is of significant environmental importance, and in this study, the effect of rainfall and temperature on the volume of this river flow is studied.

2. Materials and methods

2.1. The Qareh-Sou basin

The Qareh-Sou basin, with an area of 1670 square kilometers, forms a significant part of Golestan province in northern Iran. This basin area is limited to the Gorgan-Roud basin from the north and east, the Naka-Roud basin from the south, and the Gorgan Bay basin and the Great Caspian Sea from the west. The Qareh-Sou River discharges into the bay near Qareh-Sou village. The main Qareh-Sou basin area is covered by forest in the south, while in the north, an alluvial plain with agricultural and residential uses forms the basin. The differences in elevation between the southern heights and northern alluvial plain, besides heavy rainfall, have caused very young south-north rivers to flow with severe erosion. After reaching the plain, these rivers leave their primary sediment by forming large-grained alluvial fans. Due to a sudden change of direction, the rivers upstream of this basin discharge most of their sediments in the river after joining the main river of the Qareh-Sou basin. The Qareh-Sou River is vital in supplying agricultural water resources in the region, and therefore, it has socio-economic importance.

Another point is that this river is the leading supplier of freshwater resources for Gorgan Bay. Gorgan Bay is of enormous environmental and ecological importance. Due to the quite effective role of the Qareh-Sou River, any changes in the flow volume of this river can be the source of severe effects on this water body. The data of Siah-Ab and Gorgan hydrometric meteorological stations were used to investigate the effect of rainfall and temperature variables on the flow volume of the river. The location of the studied basin, and the hydrometric and meteorological stations are shown in *Fig. 1*.

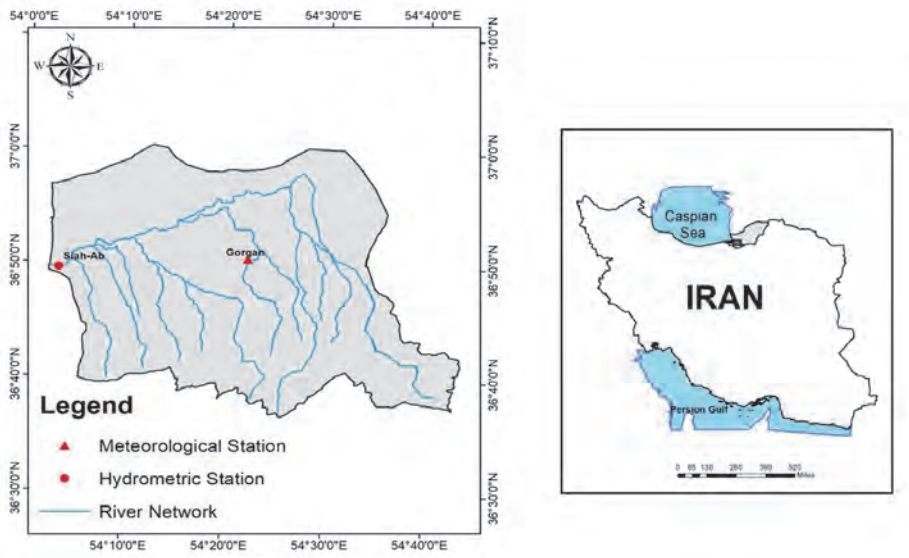


Fig. 1. Location of Gorgan meteorological and Siah-Ab hydrometric stations in the Qareh-Sou basin.

2.2. The multiple linear regression model

One of the standard methods in multivariate analysis is the multiple linear regression model (*Kadam et al*., 2019). A linear relationship is established between the independent variable and one or more dependent variables (*Jolánkai* and *Koncsos,* 2018). In the multiple linear regression, the parameters of a linear model are estimated using an objective function and the values of the variables (*Zhang et al*., 2020). In the linear regression, the considered model is a linear relationship between the model parameters (*Ali et al*., 2020). Thus, if we have n observations of x independent variable with p dimension and want to establish a linear relationship with the dependent variable y, we can use the following linear regression model (*Li et al*., 2019):

$$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip} + \varepsilon_i \ , \quad i = 1, \ldots, n \ , \tag{1}$$

where β is the model parameter. Index i shows the observation number and ε is considered a regression model error. If two independent variables are linearly

related to a dependent variable in multiple linear regression, the relationship will form a plane (*Fig. 2*).

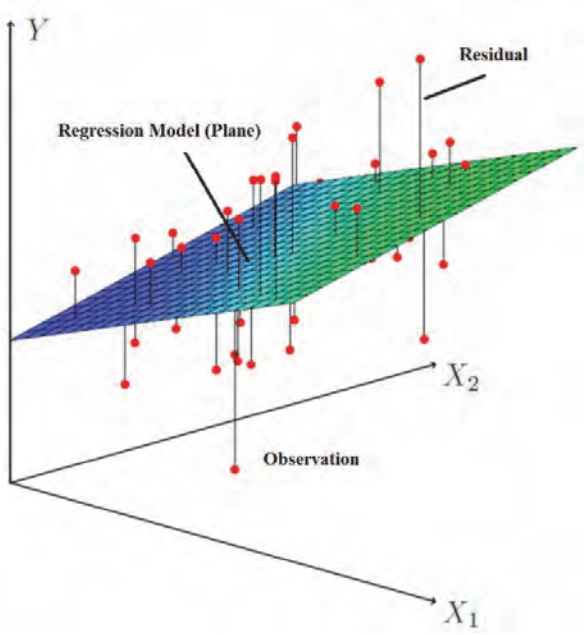


Fig. 2. Multiple linear regression model for two independent variables.

2.3. The Bayesian quantile regression model

Research on changes in hydrological and hydroclimatic variables has been mainly based on models that examine the median or average changes (*Hu et al.*, 2020; *Ali et al*., 2019). An important point to note is that in hydrological and hydroclimatic events, the upper and lower quantile, which can represent extreme events, are extremely important (*MacLeod et al*., 2021). Simultaneously, it should be considered that conventional models in this field do not have good performance (*Shiau* and *Huang*, 2015). The study of changes in hydrological and hydroclimatic variables in the upper tail of the probability distribution function is of great importance for studies related to risk and uncertainty in design related to hydrology, climatology, meteorology, and the environment (*Shiau* and *Chen*, 2015). The Bayesian quantile regression model can be a suitable and practical tool to study the upper and lower quantiles (*Uranchimeg et al*., 2020). Estimating the changes in the upper and lower quantiles can be used to study wet and dry seasons

and extreme floods, which shows the importance of this type of analysis (*Kalisa et al*., 2021). In the quantile regression, the values of conditional quantiles of dependent variables estimate for changes in independent variables (*Wan* and *Liew*, 2020). Therefore, the quantile regression model is entirely different from the known model of linear regression and multiple linear regression that examines the conditional mean changes of the dependent variable (*Bogner et al.,* 2017). The Bayesian regression has been developed to overcome the limitations of quantile regression. More information on quantile regression and Bayesian quantile regression are available from sources such as *Acharya et al*. (2020), *He et al.* (2021), and *Shin et al*. (2021). The following function is minimized in the quantile regression model to estimate regression lines for different quantiles (*Wang et al.,* 2018):

$$\hat{\beta}_\tau = argmin \sum_{i=1}^{n} \rho_\tau \quad (y_i - x_i^T \beta)^2 \quad , \tag{2}$$

where $\hat{\beta}_\tau$ is τth quantile regression line. The $\rho_\tau(x) = x(\tau - I(x < 0))$ is also considered a loss function, and i is defined as an indicator function. The maximization of a regression likelihood function generated by asymmetric Laplace densities, presented by *Yu* and *Zhang* (2005), is the same as the minimization of the previous equation:

$$f(x|\mu, \delta, \tau) = \frac{\tau(1-\tau)}{\sigma} \exp[-\sigma = \rho_\tau \left(\frac{x-\mu}{\sigma}\right)] \quad , \tag{3}$$

The Bayesian inference can estimate the studied parameter's entire posterior probability distribution function, including parameter uncertainty, based on this inference (*Yang*, 2019). In this study, a Bayesian quantile regression model was used to investigate the relationship between the river flow volume as a dependent variable and the rainfall and temperature as independent variables. The calculations were performed using the "bayesQR" package (*Benoit* and *Van den Poel*, 2017) developed in the R environment.

3. Results and discussion

The data were divided into five series: annual, spring, summer, autumn, and winter. The reason of division is that the relationship between rainfall and temperature with the volume of river flow experience changes in different seasons. According to the generated series, the relationship between hydroclimatic variables and river flow volume was investigated using multiple linear regression and Bayesian quantile regression models, as reported below.

Examination of the slope values of Bayesian quantile regression lines in the annual data shows that the relationship between the annual rainfall and annual flow volume with the slope range (7–156) is direct, which increases with increasing the values of quantiles (*Fig. 3*). In the upper quantiles, this incremental pattern disappears, which may be because a set of other factors can also have significant effects on the annual extreme flows. The annual temperature effect on the annual flow volume with a slope range of ((-3) –15) is also direct in some quantiles and indirect in others. The maximum effect of temperature on the annual flow volume is in the upper quantiles. Comparing the effect of rainfall and temperature on the annual flow volume confirms that rainfall has greater effect than temperature, so that with increasing the values of quantiles, the difference between the effect of rainfall and temperature increases.

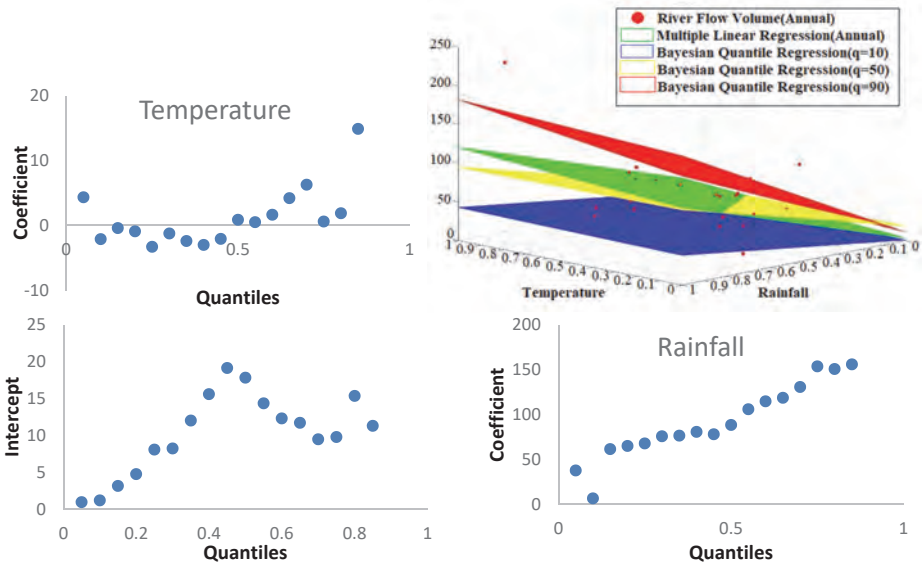


Fig. 3. Results of Bayesian quantile and multiple linear regression models for the annual series.

The slope obtained from the multiple linear regression method for rainfall and temperature is 97 and 20. In the Bayesian quantile regression model, there is a negative slope for temperature in some quantiles. In contrast, in multiple linear regression, there is a positive slope sign. The slope value obtained in the multiple linear regression model for rainfall is in the range of slopes obtained in the Bayesian quantile regression model. For temperature, the slope value obtained in the multiple linear regression method is outside the slope range obtained in the Bayesian quantile regression model.

The slope of Bayesian quantile regression lines in the spring data for the rainfall variable is in the range of ((-5) –4), indicating that the relationship between the spring rainfall and spring flow volume is direct in some quantiles and indirect in others (*Fig. 4*). A remarkable effect of spring rainfall on spring flow volume is detectable in the upper quantiles. The slope value associated with different quantiles for the temperature variable is in the range ((-28) – (- 3)), which means that the relationships between the spring temperature and spring flow volume in all quantiles are indirect. The remarkable effect of spring temperature on the volume of spring flow is in the middle quantile, while in the upper and lower quantiles, this effect is significantly reduced. In spring, the effect of temperature on runoff volume is more than the effect rainfall. The most remarkable difference between the magnitude of the effects of these two variables can be seen in the middle quantiles. The slope obtained from the multiple linear regression model for temperature and rainfall is estimated to be -18 and -17, respectively. Therefore, it can be seen that in the Bayesian quantile regression model, in some quantiles, rainfall has a positive slope, but in multiple linear regression, the slope sign is negative. The value of the slope obtained in the multiple linear regression method for rainfall is outside the range of the slopes obtained in the Bayesian quantile regression model, while for temperature, there is the opposite behavior.

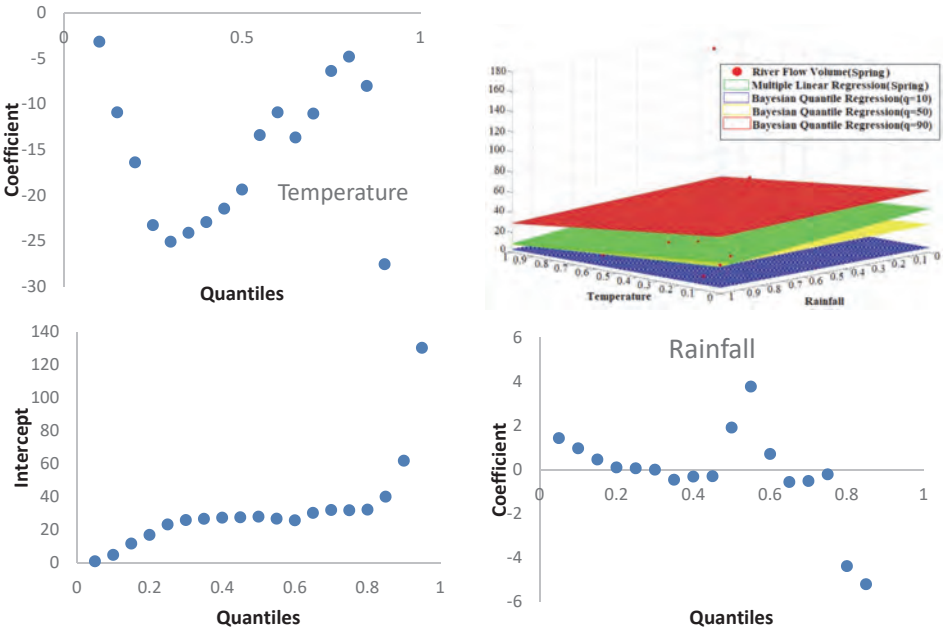


Fig. 4. Results of Bayesian quantile and multiple linear regression models for the spring series

Summer data show that the slope value for the temperature variable is in the range (0–9). Therefore, it can be said that the relationships between the temperature and flow volume in all quantiles are direct (*Fig. 5*). In general, with increasing the values of quantiles, the magnitude of the effect of temperature also increases, and experience a decrease only in the last quantile. This result is because in upper quantiles, the influence of other factors on flow volume increases. For the rainfall variable, the slope value was in the range (0–46), and with increasing the quantile value, the slope magnitude increases significantly. This result is due to the predominant effect of rainfall on the volume of flow in summer, because in this season, according to the river conditions, the river flow in most conditions is the base flow. In the lower and middle quantiles, the effect of temperature is greater than that of rainfall, although this difference is not remarkable. In upper quantiles, the magnitude of the effect of rainfall is dramatic compared to temperature increases, which is different from the other studied series. The value of the slope calculated for temperature and rainfall using multiple linear regression model is 7 and 9, respectively, which in terms of sign and the values are consistent with the results of the Bayesian quantile regression model.

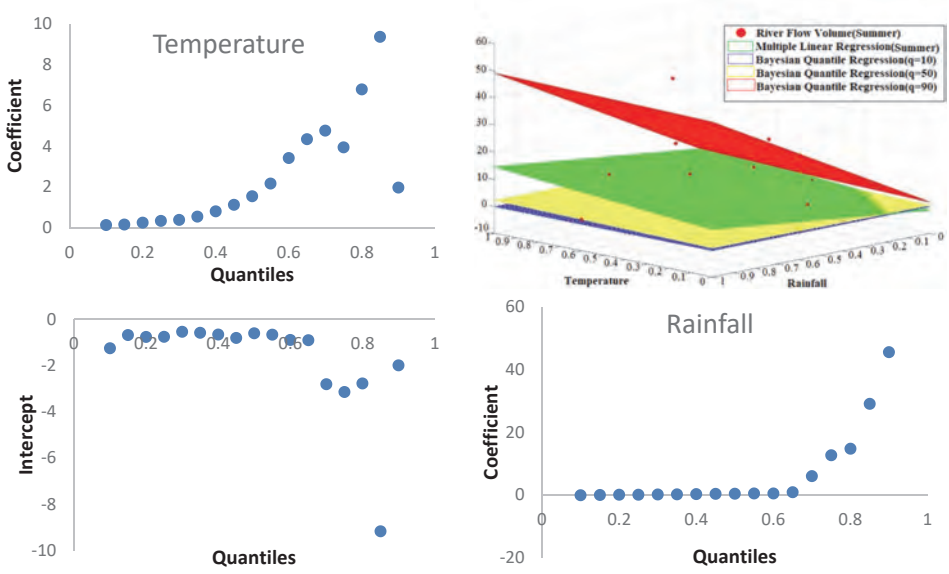


Fig. 5. Results of Bayesian quantile and multiple linear regression models for the summer series.

In autumn, the slope for rainfall is in the range (6–25), which means that the relationship between the rainfall and flow volume is direct in all quantiles, and a remarkable amount of impact is observed in the upper quantile (*Fig. 6*). It is important to note that the increasing trend of the rainfall-related slope disappears in the upper quantiles, which may be due to the significant impact of other variables on autumn flow volume. The slope range for temperature in this season is (1–4), which means that in autumn, the relationship between the temperature and flow volume is generally similar to the relationship between the rainfall and flow volume, with the difference that the intensity of the impact of rainfall is far greater than that of the temperature. The differences between the magnitudes of rainfall and temperature in the middle and upper quantiles are far more significant than in the lower quantiles. In a multiple linear regression model, the slopes for temperature and rainfall are 7 and 20, respectively. The linear regression model's slope sign in autumn is similar to the Bayesian quantile regression model results. It should be noted that the slope values for temperature and rainfall in the multiple linear regression model are outside and inside the range obtained from the Bayesian quantile regression method, respectively.

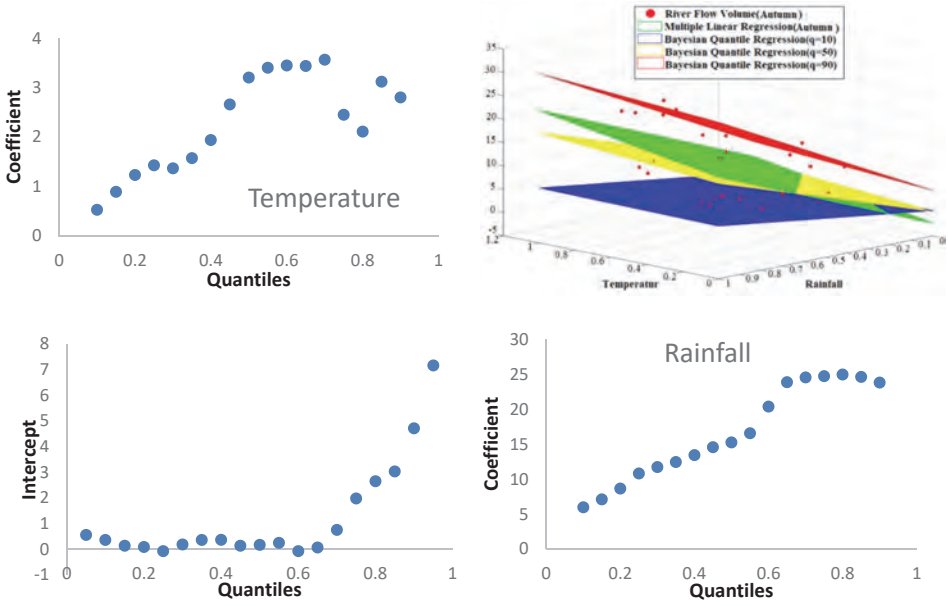


Fig. 6. Results of Bayesian quantile and multiple linear regression models for the autumn series.

Examination of winter data reveals that the values of slopes for rainfall are in the range (9–41), and in other words, the effect of rainfall on flow volume is direct (*Fig. 7*). However, the magnitude of this effect in the middle quantiles is significantly higher than that of the upper and lower quantiles in this respect, and it behaves almost like spring. The values obtained for the temperature slopes are also in the range ((-12) –0). The effects of temperature on flow volume are direct in the lower quantiles and indirect in the upper quantiles. The intensity of this effect increases with increasing the values of quantiles. Comparison between the magnitude of the effect of rainfall and temperature on the volume of winter flow shows that rainfall has more effect than the temperature, and the critical point is that the most significant difference between the magnitude of the effect of these two variables occurred in the middle quantiles, which behave similarly to spring data. The slope obtained from the multiple linear regression model for temperature and rainfall in winter is -8 and 28, respectively. Comparison of these values with the range of values recorded in the Bayesian quantile regression model indicates a quantitative agreement between the results of these two models.

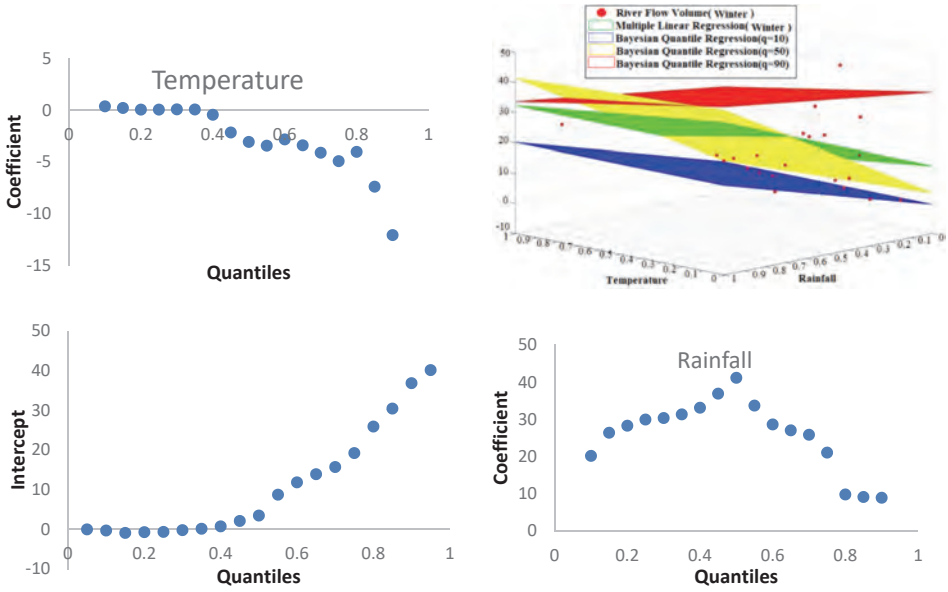


Fig. 7. Results of Bayesian quantile and multiple linear regression models for the winter series.

A comparison between the multiple linear regression and Bayesian quantile regression results was presented in the section above. Investigating these results indicates that the behavior between the flow volume and temperature and rainfall variables in different quantiles may be quite different. This difference can be seen in the magnitude of the slope value and in the slope sign of the regression lines. This issue is fundamental in hydrological estimates, because it shows that the value of rainfall and temperature variables varies in different quantiles on the volume of flow, and this difference is significant in some cases.

4. Conclusion

The volume of river flow is significantly affected by hydroclimatic factors such as rainfall and temperature. The multiple linear regression model is a well-known model in hydrological and climatological studies used to investigate the effect of independent variables on dependent variables, but this model has its limitations. In this study, multiple linear regression and Bayesian quantile regression models were used to investigate the effect of rainfall and temperature on river flow volume. Data belonging to the Qareh-Sou basin area in northern Iran were used in five (annual, spring, summer, autumn, and winter) series for this study. According to the results of the calculations, the following can be considered a general conclusion of this research.

Comparison between the magnitude of the effect of rainfall and temperature in different series indicates that in spring, the effect of temperature on flow volume is greater than the effect of rainfall, while in the annual, autumn, winter, and summer series, the effect of rainfall on flow volume is much greater than that of the temperature .The effect of rainfall and temperature variables on flow volume in different quantiles in terms of value and sign can significantly change.
The results obtained from the multiple linear regression model differ from the results obtained from the application of Bayesian quantile regression for the quantile 0.5 in value and in some cases in the sign, which means that the only application of multiple linear regression models alone can lead to erroneous analysis. The differences between the plane fitted by multiple linear regression with the planes fitted by Bayesian quantile regression in the upper and lower quantiles are enormous. Therefore, the multiple linear regression model has many limitations in studies related to extreme river flows. In the annual, autumn, winter, and spring series, with increasing the values of quantiles, the effect of rainfall on flow volume decreases, which may be because of extreme flows. Other variables such as previous soil moisture, soil cover, and land use are influential. In summer, a different pattern is seen so that with increasing the values of quantiles, the effect of rainfall on flow volume increases. This result may be since river flow in summer is generally of the base-flow type, and therefore, the amount of rainfall has a significant effect on flow volume in upper quantiles.

References

Acharya, S.C., Babel, M.S., Madsen, H., Sisomphon, P., and *Shrestha, S*., 2020: Comparison of different quantile regression methods to estimate predictive hydrological uncertainty in the Upper Chao Phraya River Basin, Thailand. *J. Flood Risk Manage. 13*, e12585. https://doi.org/10.1111/jfr3.12585

Ahmed, M.H. and *Lin, L.S*., 2021: Dissolved oxygen concentration predictions for running waters with different land use land cover using a quantile regression forest machine learning technique. *J.Hydrology*, 126213. https://doi.org/10.1016/j.jhydrol.2021.126213

Ali, M., Prasad, R., Xiang, Y., and *Deo, R.C.*, 2020: Near real-time significant wave height forecasting with hybridized multiple linear regression algorithms. *Renew. Sustain. Energy Rev. 132*, 110003. https://doi.org/10.1016/j.rser.2020.110003

Ali, R., Kuriqi, A., Abubaker, S., and *Kisi, O.*, 2019: Long-term trends and seasonality detection of the observed flow in Yangtze River using Mann-Kendall and Sen's innovative trend method. *Water 11*(9), 1855. https://doi.org/10.3390/w11091855

Ansarifar, M.M., Salarijazi, M., Ghorbani, K., and *Kaboli, A.R*., 2020a: Simulation of groundwater level in a coastal aquifer. *Marine Geores. Geotechnol. 38*, 257–265. https://doi.org/10.1080/1064119X.2019.1639226

Ansarifar, M. M., Salarijazi, M., Ghorbani, K., and *Kaboli, A.R.*, 2020b: Spatial estimation of aquifer's hydraulic parameters by a combination of borehole data and inverse solution. *Bull. Engineer. Geolog. Environ. 79*, 729–738. https://doi.org/10.1007/s10064-019-01616-w

Bahrami, E., Mohammadrezapour, O., Salarijazi, M., and *Jou, P.H*., 2019: Effect of base flow and rainfall excess separation on runoff hydrograph estimation using gamma model (case study: Jong catchment: *KSCE J. Civil Engin. 23*, 1420–1426. https://doi.org/10.1007/s12205-019-0591-3

Bahremand, A., Ahmadyousefi, S., Sheikh, V., and *Komaki, C.B*., 2021: A parameter allocation approach for flow simulation using the WetSpa-Python model. *Hydrol. Process. 35*(1), e13992. https://doi.org/10.1002/hyp.13992

Benoit, D.F. and Van den *Poel, D.*, 2017: bayesQR: A Bayesian approach to quantile regression. *J. Stat. Software 76*, 1–32. https://doi.org/10.18637/jss.v076.i07

Bogner, K., Liechti, K., and *Zappa, M.*, 2017: Combining quantile forecasts and predictive distributions of streamflows. *Hydrol. Earth Syst. Sci. 21*, 5493–5502. https://doi.org/10.5194/hess-21-5493-2017

Chadalawada, J., Havlicek, V., and *Babovic, V*., 2017: A genetic programming approach to system identification of rainfall-runoff models. *Water Resour. Manage. 31* 3975–3992. https://doi.org/10.1007/s11269-017-1719-1

Cho, J.H. and *Lee, J.H*., 2018: Multiple linear regression models for predicting nonpoint-source pollutant discharge from a highland agricultural region. *Water 10*(9), 1156. https://doi.org/10.3390/w10091156

He, Y., Fan, H., Lei, X., and *Wan, J.*, 2021: A runoff probability density prediction method based on B-spline quantile regression and kernel density estimation. *Appl. Math. Model. 93*, 852–867. https://doi.org/10.1016/j.apm.2020.12.043

Hossain, S., Biswas, R.K., and *Hossain, M.A.*, 2021: Body mass index of women in Bangladesh: comparing Multiple Linear Regression and Quantile Regression. *J. Biosoc. Sci. 53*, 247–265. https://doi.org/10.1017/S0021932020000176

Hu, Z., Liu, S., Zhong, G., Lin, H., and *Zhou, Z.*, 2020: Modified Mann-Kendall trend test for hydrological time series under the scaling hypothesis and its application. *Hydrol. Sci. J. 65*, 2419–2438. https://doi.org/10.1080/02626667.2020.1810253

Jolánkai, Z. and *Koncsos, L.*, 2018: Base flow index estimation on gauged and ungauged catchments in Hungary using digital filter, multiple linear regression and artificial neural networks. *Periodica Polytechnica Civil Engin. 62*, 363–372. https://doi.org/10.3311/PPci.10518

Kadam, A.K., Wagh, V.M., Muley, A.A., Umrikar, B.N., and *Sankhua, R.N.*, 2019: Prediction of water quality index using artificial neural network and multiple linear regression modelling approach in Shivganga River basin, India. *Model. Earth Syst. Environ. 5*, 951–962. https://doi.org/10.1007/s40808-019-00581-3

Kalisa, W., Igbawua, T., Ujoh, F., Aondoakaa, I.S., Namugize, J.N., and *Zhang, J*., 2021: Spatio-temporal variability of dry and wet conditions over East Africa from 1982 to 2015 using quantile regression model. *Nat. Hazards 106*, 2047–2076. https://doi.org/10.1007/s11069-021-04530-1

Karimi, S., Salarijazi, M., Ghorbani, K., and *Heydari, M.*, 2021: Comparative assessment of environmental flow using hydrological methods of low flow indexes, Smakhtin, Tennant and flow duration curve. *Acta Geophysica 69*, 285–293. https://doi.org/10.1007/s11600-021-00539-z

Kavian, A., Javidan, N., Bahrehmand, A., Gyasi-Agyei, Y., Hazbavi, Z., and *Rodrigo-Comino, J*., 2020: Assessing the hydrological effects of land-use changes on a catchment using the Markov chain and WetSpa models. *Hydrol. Sci. J. 65*, 2604–2615. https://doi.org/10.1080/02626667.2020.1797046

Kumar, A., Singh, R., Jena, P.P., Chatterjee, C., and *Mishra, A.,* 2015: Identification of the best multi-model combination for simulating river discharge. *J. Hydrology 525*, 313–325. https://doi.org/10.1016/j.jhydrol.2015.03.060

Latt, Z. and *Wittenberg, H*., 2014: Improving flood forecasting in a developing country: a comparative study of stepwise multiple linear regression and artificial neural network. *Water Res. Manage. 28*, 2109–2128. https://doi.org/10.1007/s11269-014-0600-8

Li, W., Zhou, J., Chen, L., Feng, K., Zhang, H., Meng, C., and *Sun, N.*, 2019: Upper and lower bound interval forecasting methodology based on ideal boundary and multiple linear regression models. *Water Res. Manage. 33*, 1203–1215. https://doi.org/10.1007/s11269-018-2177-0

MacLeod, D.A., Dankers, R., Graham, R., Guigma, K., Jenkins, L., Todd, M.C., ... and *Mwangi, E*., 2021: Drivers and subseasonal predictability of heavy rainfall in equatorial East Africa and relationship with flood risk. *J. Hydrometeorol. 22*, 887–903. https://doi.org/10.1175/JHM-D-20-0211.1

Mishra, S., Saravanan, C., Dwivedi, V.K., and *Shukla, J.P.,* 2018: Rainfall-Runoff Modeling using Clustering and Regression Analysis for the River Brahmaputra Basin. *J. Geol Soc India 92*, 305–312. https://doi.org/10.1007/s12594-018-1012-9

More, K.S., Wolkersdorfer, C., Kang, N., and *Elmaghraby, A.S.,* 2020: Automated measurement systems in mine water management and mine workings–A review of potential methods. *Water Res. Industry*, 100136. https://doi.org/10.1016/j.wri.2020.100136

Mudbhatkal, A., Raikar, R. V., Venkatesh, B., and *Mahesha, A*., 2017: Impacts of climate change on varied river-flow regimes of southern India. *J. Hydrol. Engin. 22*, 05017017. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001556

Nguyen, H.H., Cho, S., Jeong, J., and *Choi, M.,* 2021: A D-vine copula quantile regression approach for soil moisture retrieval from dual polarimetric SAR Sentinel-1 over vegetated terrains. *Remote Sens. Environ. 255*, 112283. https://doi.org/10.1016/j.rse.2021.112283

Niedzielski, T., Szymanowski, M., Miziński, B., Spallek, W., Witek-Kasprzak, M., Ślopek, J., ... and *Leszczyński, L.*, 2019: Estimating snow water equivalent using unmanned aerial vehicles for determining snow-melt runoff. *J. Hydrology 578*, 124046. https://doi.org/10.1016/j.jhydrol.2019.124046

Niu, W.J., Feng, Z.K., Feng, B.F., Min, Y.W., Cheng, C.T., and *Zhou, J.Z.,* 2019: Comparison of multiple linear regression, artificial neural network, extreme learning machine, and support vector machine in deriving operation rule of hydropower reservoir. *Water 11*(1), 88. https://doi.org/10.3390/w11010088

Noori, R., Khakpour, A., Omidvar, B., and *Farokhnia, A.*, 2010: Comparison of ANN and principal component analysis-multivariate linear regression models for predicting the river flow based on developed discrepancy ratio statistic. *Exp. Syst. Appl. 37*, 5856–5862. https://doi.org/10.1016/j.eswa.2010.02.020

Nourani, V., Tajbakhsh, A. D., and *Molajou, A.,* 2019: Data mining based on wavelet and decision tree for rainfall-runoff simulation. *Hydrol. Res. 50*, 75–84. https://doi.org/10.2166/nh.2018.049

Patel, S., Hardaha, M.K., Seetpal, M.K., and *Madankar, K.K.*, 2016: Multiple linear regression model for stream flow estimation of Wainganga River. *Amer. J. Water Sci. Engin. 2*(1), 1–5.

Popat, E., Kuleshov, A., Kronenberg, R., and *Bernhofer, C.*, 2020: Data-driven discharge analysis: a case study for the Wernersbach catchment, Germany. *Meteorol. Hydrol. Water Manage. Res. Oper. Appl. 8*. https://doi.org/10.26491/mhwm/112284

Sa'adi, Z., Shahid, S., Ismail, T., Chung, E.S., and *Wang, X.J.*, 2017: Distributional changes in rainfall and river flow in Sarawak, Malaysia. *Asia-Pacific J. Atmosph Sci. 53*, 489–500. https://doi.org/10.1007/s13143-017-0051-2

Salarijazi, M. and *Ghorbani, K.*, 2019: Improvement of the simple regression model for river'EC estimation. *Arabian J. Geoscie. 12*(7), 1–14. https://doi.org/10.1007/s12517-019-4392-2

Schilling, K.E. and *Walter, C.F.*, 2005: Estimation of streamflow, base flow and nitrate-nitrogen loads in Iowa using multiple linear regression 1. *J. Amer. Water Res. Assoc. 41*, 1333–1346. https://doi.org/10.1111/j.1752-1688.2005.tb03803.x

Sehgal, V., Tiwari, M.K., and *Chatterjee, C.*, 2014: Wavelet bootstrap multiple linear regression based hybrid modeling for daily river discharge forecasting. *Water Res. Manage. 28*, 2793–2811. https://doi.org/10.1007/s11269-014-0638-7

Sezen, C., Bezak, N., Bai, Y., and *Šraj, M.*, 2019: Hydrological modelling of karst catchment using lumped conceptual and data mining models. *J Hydrology 576*, 98–110. https://doi.org/10.1016/j.jhydrol.2019.06.036

Shiau, J.T., and *Huang, W.H.*, 2015: Detecting distributional changes of annual rainfall indices in Taiwan using quantile regression. *J. Hydro-environ. Res. 9*, 368–380. https://doi.org/10.1016/j.jher.2014.07.006

Shin, J., You, H., Kaown, D., Koh, E. H., Lee, S., Lim, C.Y., and *Lee, K.K.*, 2021: Investigating distribution of nitrate concentration using ensemble nonparametric quantile regression. *Sci. Total Environ. 777*, 146098. https://doi.org/10.1016/j.scitotenv.2021.146098

Steinfeld, C.M., Sharma, A., Mehrotra, R., and *Kingsford, R.T.*, 2020: The human dimension of water availability: Influence of management rules on water supply for irrigated agriculture and the environment. *J. Hydrology 588*, 125009. https://doi.org/10.1016/j.jhydrol.2020.125009

Tsakiri, K., Marsellos, A., and *Kapetanakis, S.*, 2018: Artificial neural network and multiple linear regression for flood prediction in Mohawk River, New York. *Water 10*(9), 1158. https://doi.org/10.3390/w10091158

Uranchimeg, S., Kwon, H.H., Kim, B., and *Kim, T.W.* 2020: Changes in extreme rainfall and its implications for design rainfall using a Bayesian quantile regression approach. *Hydrology Res. 51*, 699–719. https://doi.org/10.2166/nh.2020.003

Wan, J.S. and *Liew, E.C.*, 2020: Genus-level change in aggressiveness with continuous invasions: a phylogenetically-informed Bayesian quantile regression. *Biol. Invasions 22*, 1931–1946. https://doi.org/10.1007/s10530-020-02229-1

Wang, H.J., McKeague, I.W., and *Qian, M.*, 2018: Testing for marginal linear effects in quantile regression. *J. Roy. Stat. Soc.. Ser. B, Stat. Method. 80*(2), 433. https://doi.org/10.1111/rssb.12258

Yang, Y., 2019: Spatial and Temporal Variabilities of Climate Extremes over Canada in a Changing Climate. A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Water Resources Engineering Department of Civil and Environmental Engineering University of Albert. https://doi.org/10.7939/r3-kz2b-5r47

Yu, K. and *Zhang, J.*, 2005: A three-parameter asymmetric Laplace distribution and its xtension. *Commun. Stat.—Theory Methods 34*, 1867–1879. https://doi.org/10.1080/03610920500199018

Zhang, G., Liu, X., Lu, S., Zhang, J., and *Wang, W.*, 2020: Occurrence of typical antibiotics in Nansi Lake's inflowing rivers and antibiotic source contribution to Nansi Lake based on principal component analysis-multiple linear regression model. *Chemosphere 242*, 125269. https://doi.org/10.1016/j.chemosphere.2019.125269

IDŐJÁRÁS

VOLUME 126 * 2022

BUDAPEST, HUNGARY

AUTHOR INDEX

TABLE OF CONTENTS

I. papers

SUBJECT INDEX

V

W

INSTRUCTIONS TO AUTHORS OF *IDŐJÁRÁS*

The purpose of the journal is to publish papers in any field of meteorology and atmosphere related scientific areas. These may be

- research papers on new results of scientific investigations,
- critical review articles summarizing the current state of art of a certain topic,
- short contributions dealing with a particular question.

Some issues contain "News" and "Book review", therefore, such contributions are also welcome. The papers must be in American English and should be checked by a native speaker if necessary.

Authors are requested to send their manuscripts to

Editor-in Chief of IDŐJÁRÁS
P.O. Box 38, H-1525 Budapest, Hungary
E-mail: journal.idojaras@met.hu

including all illustrations. MS Word format is preferred in electronic submission. Papers will then be reviewed normally by two independent referees, who remain unidentified for the author(s). The Editor-in-Chief will inform the author(s) whether or not the paper is acceptable for publication, and what modifications, if any, are necessary.

Please, follow the order given below when typing manuscripts.

Title page should consist of the title, the name(s) of the author(s), their affiliation(s) including full postal and e-mail address(es). In case of more than one author, the corresponding author must be identified.

Abstract: should contain the purpose, the applied data and methods as well as the basic conclusion(s) of the paper.

Key-words: must be included (from 5 to 10) to help to classify the topic.

Text: has to be typed in single spacing on an A4 size paper using 14 pt Times New Roman font if possible. Use of S.I. units are expected, and the use of negative exponent is preferred to fractional sign. Mathematical formulae are expected to be as simple as possible and numbered in parentheses at the right margin.

All publications cited in the text should be presented in the *list of references*, arranged in alphabetical order. For an article: name(s) of author(s) in Italics, year, title of article, name of journal, volume, number (the latter two in Italics) and pages. E.g., *Nathan, K.K.,* 1986: A note on the relationship between photosynthetically active radiation and cloud amount. *Időjárás 90,* 10–13. For a book: name(s) of author(s), year, title of the book (all in Italics except the year), publisher and place of publication. E.g., *Junge, C.E.,* 1963: *Air Chemistry and Radioactivity.* Academic Press, New York and London. Reference in the text should contain the name(s) of the author(s) in Italics and year of publication. E.g., in the case of one author: *Miller* (1989); in the case of two authors: *Gamov* and *Cleveland* (1973); and if there are more than two authors: *Smith et al.* (1990). If the name of the author cannot be fitted into the text: *(Miller,* 1989); etc. When referring papers published in the same year by the same author, letters a, b, c, etc. should follow the year of publication. DOI numbers of references should be provided if applicable.

Tables should be marked by Arabic numbers and printed in separate sheets with their numbers and legends given below them. Avoid too lengthy or complicated tables, or tables duplicating results given in other form in the manuscript (e.g., graphs). *Figures* should also be marked with Arabic numbers and printed in black and white or color (under special arrangement) in separate sheets with their numbers and captions given below them. JPG, TIF, GIF, BMP or PNG formats should be used for electronic artwork submission.

More information for authors is available: journal.idojaras@met.hu

Published by the Hungarian Meteorological Service

Budapest, Hungary

ISSN 0324-6329 (Print)

ISSN 2677-187X (Online)