OUTLIERS IN BULGARIAN MONTHLY TEMPERATURE SERIES IN THE PERIOD 1900-2000

M. Syrakova[1], *P. Dimitrova*[2]

[1]Sofia University, Faculty of Physics, Dept. of Meteorology and Geophysics, 5, J.Bourchier blvd., Sofia 1164, Bulgaria, e-mail: marsyr@phys.uni-sofia.bg
[2]National Institute of Meteorology and Hydrology, 66, Tzarigradsko chosse, Sofia 1784, Bulgaria, e-mail: prawda.dimitrova@meteo.bg

Abstract. The monthly mean temperature series at 31 stations, representative for the lowlands of Bulgaria, are checked for outliers, i.e. for observations very distant from the mean value. The data cover the period 1900-2000. From November to April, negative outliers dominate, from May to October, positive outliers prevail. The total number of the outliers is greatest in July, but the outliers' magnitudes are greater (by absolute value) in winter. The temporal distribution of the number and magnitude of the outliers over the 20th century shows, that the temperature variability in the beginning of the century can be estimated as moderate, then since 1919 till mid 1950s an increase of the number and magnitude of the outliers (both positive and negative) occurs, the period till late 1980s is relatively calm, and then a new increase of the temperature variability is in progress.

Key words: monthly temperature series, outliers, Bulgaria, 20th century

Introduction

In recent years much attention has been directed at climate variability and climate change. The long-term climate time series, on which such investigations are based, contain some values, which are rather distant from the mean value – the so called "outliers". On the one hand, such values that are not characteristic of the bulk of the data may be erroneous, the errors being caused by corrupted measurements, misprints in the log books or other reasons. That is why the identification of outliers is the basis of the quality control work (e.g. Trenberth and Paolino, 1980; Lanzante, 1996; Peterson and Vose, 1997; Peterson *et al*., 1998). On the other hand, however, the outliers may be due to extreme meteorological events – extremely heavy precipitation or exceptionally cold/warm months in certain year. In these cases they are completely valid values.

Here, the outliers as indicators of extreme meteorological events are considered.

Temperature series at stations, representative of the lowlands of Bulgaria, are inspected. Data and method are described in the next section. Section 3 contains the results – the size, temporal and annual distribution of the outliers. In the end, the conclusions are summed up.

Data and method

In this work, we have used monthly mean temperature series at 31 stations distributed over the lowlands of Bulgaria, as it is shown in Fig.1.

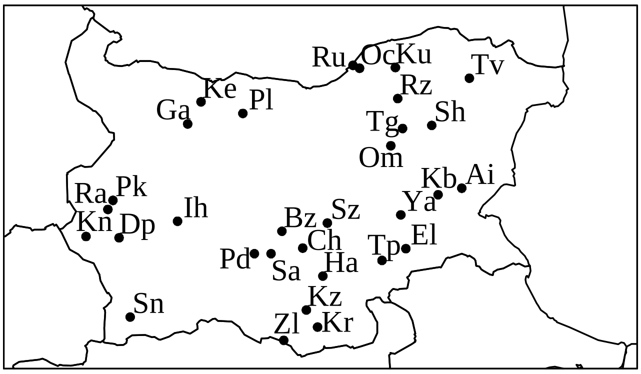

Fig. 1. Spatial distribution of the stations. For the names, see Table 1.

The data are provided by the National Institute of Meteorology and Hydrology (NIMH). Stations with the longest series of temperature observations, containing the least number of missing data and representative for different parts of the territory are chosen. The list of the stations with their geographical position (latitude, longitude, altitude), period of observations and the abbreviations of the names, used in Fig.1 and in some other figures, is given in Table 1 (the arrangement is according to the latitude, from north to south).

Temperature data, covering the entire period of 20th century, are available in very few stations (seven in the list). In fact, observations in some of these stations started in the end of 19th century, but in different years, that is why we put the beginning of these series in 1900. A big part of the records began in 1920-1930, but there are stations where temperature observations were started even later – in 1940s and 1950s. Unfortunately, at some stations, observations were interrupted in 1990s. The time evolution of the number of stations is shown in Fig.2.

As it is said in the Introduction, the meteorological series, especially for long-term periods, contain some values rather distant from the bulk of the data. In our case, these are very low/high temperatures. Examples of the frequency distribution of monthly mean temperatures are given in Fig.3, where one can see such distant values.

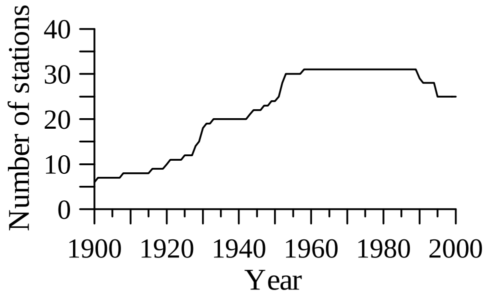

Fig. 2. Time evolution of the number of stations

The identification of the outliers is based on certain threshold values, calculated for each time series – outliers are the values, which trespass the threshold. The threshold values depend on the variability of the series. Examples of using different measures of the variability of a series for calculating the threshold values are given in Peterson *et al.* (1998). The most widely used measure of the variability is the standard deviation and the threshold values are assumes to be $(\bar{x} \pm k\sigma)$, where $\bar{x}$ is the mean value of the series, σ is the standard deviation and different values of k are used in different works. According to Peterson *et al*. (1998), a common practice is using 3σ limit for temperature series. The same limit is used in Trenberth and Paolino (1980) for sea-level pressure series and in Syrakova and Stefanova (2008) for temperature series.

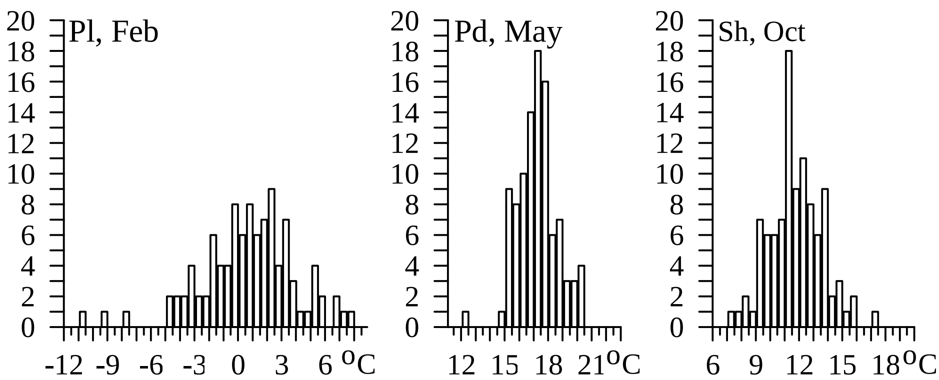

Fig. 3. Examples of frequency distribution of monthly mean temperatures

Table 1. List of the stations. The abbreviations of the names are used in Fig. 1.

	Station		Latitude (N)	Longitude (E)	Altitude (m, a.s.l.)	Period (year)
1.	Russe	(Ru)	43° 51'	25° 57'	37	1928-2000
2.	Obr.chiflik	(Oc)	43° 48'	26° 02'	157	1900-2000
3.	Kubrat	(Ku)	43° 48'	26° 30'	230	1949-1991
4.	Tervel	(Tv)	43° 45'	27° 25'	215	1952-1995
5.	Razgrad	(Rz)	43° 31'	26° 31'	206	1916-2000
6.	Knezha	(Ke)	43° 30'	24° 05'	117	1920-2000
7.	Pleven	(Pl)	43° 24'	24° 36'	156	1900-2000
8.	Gabare	(Ga)	43° 18'	23° 55'	120	1944-1990
9.	Shumen	(Sh)	43° 16'	26° 56'	218	1900-2000
10.	Targovishte	(Tg)	43° 15'	26° 34'	226	1952-2000
11.	Omurtag	(Om)	43° 06'	26° 25'	530	1953-2000
12.	Aitos	(Ai)	42° 42'	27° 16'	90	1925-1995
13.	Karnobat	(Kb)	42° 39'	26° 59'	196	1930-2000
14.	Pernik	(Pk)	42° 36'	23° 02'	694	1947-2000
15.	Radomir	(Ra)	42° 33'	22° 58'	691	1929-1990
16.	Yambol	(Ya)	42° 29'	26° 31'	143	1930-2000
17.	Ihtiman	(Ih)	42° 26'	23° 49'	636	1921-2000
18.	St. Zagora	(Sz)	42° 25'	25° 38'	229	1943-2000
19.	Brezovo	(Bz)	42° 21'	25° 05'	257	1952-1995
20.	Kjustendil	(Kn)	42° 16'	22° 43'	527	1900-2000
21.	Dupnitsa	(Dp)	42° 16'	23° 07'	555	1908-2000
22.	Chirpan	(Ch)	42° 12'	25° 20'	173	1928-2000
23.	Elhovo	(El)	42° 11'	26° 34'	138	1951-2000
24.	Plovdiv	(Pd)	42° 09'	24° 45'	160	1901-2000
25.	Sadovo	(Sa)	42° 09'	24° 57'	155	1900-2000
26.	Topolovgrad	(Tp)	42° 05'	26° 17'	305	1958-2000
27.	Haskovo	(Ha)	41° 57'	25° 34'	230	1900-2000
28.	Kardzhali	(Kz)	41° 39'	25° 22'	331	1930-2000
29.	Sandanski	(Sn)	41° 34'	23° 17'	203	1931-2000
30.	Krumovgrad	(Kr)	41° 30'	25° 30'	231	1953-2000
31.	Zlatograd	(Zl)	41° 22'	25° 06'	430	1933-2000

Another measure of variability, used for example in González-Rouco *et al.* (2001) for identification of outliers in precipitation series, is the interquartile range (IQR) – the difference between the upper $q_{0.75}$ and the lower $q_{0.25}$ quartile. They define the threshold value as $(q_{0.75} + 3IQR)$. This statistic is resistant to outliers, in contrast to the ones based on the standard deviation, which is strongly dependent on them, and this dependence may influence the revealing of the outliers (Lanzante, 1996).

Here, the threshold values are assumed to be $(\bar{t} \pm 2.5\sigma)$, where $\bar{t}$ is the mean temperature. With the narrower 2.5σ limit instead of the 3σ limit, the number of outliers will be increased. Some of them will not be as distant from the bulk of the data as the 3σ outliers, but being at the extreme ends of the distribution they will also indicate exceptionally cold or warm months.

It is obvious that the width of the limits depends only on the standard deviation, which has well manifested seasonal variation and can be different at different stations. Because of the greater intensity of the atmospheric processes in the cold seasons, the values of σ are greater in winter than in summer, being greatest in February. In winter, the spatial distribution of σ shows greater values in North Bulgaria, where in February they are (3.3-3.4)°C at 8 out of 11 stations. At the other stations, the values of σ in February are between 2.4 and 2.9°C. In almost all stations, the summer values are least in June-July – about 1°C, and do not show any spatial patterns.

The limits themselves, however, depend not only on σ, but on the mean temperature $\bar{t}$, as well, which has much higher seasonal and spatial variability. Besides, the mean temperatures depend to a greater extent than σ on the concrete time period over which they are calculated. Having in mind that not all of the temperature series inspected here have equal length, the values of $\bar{t}$ and σ are calculated for the period 1951-1995, when observations are available at most of the stations.

To overcome the dependence of the standard deviation and hence of the threshold values on the outliers themselves, an iterative procedure is applied. First, the threshold values are calculated using the original data. The values outside these limits are substituted by the threshold values, $\bar{t}$ and σ are recalculated and the new limits are defined. This procedure is repeated until no more outliers are found. The number and magnitude of new corrections, however, are small.

The assessment that an outlier is not an erroneous value, but is a valid extreme value is based on the comparison between neighbouring highly correlated stations – the outliers for a given month, if they are not errors, appear in one and the same year at these stations or are in agreement with rather high/low values which lie within the limits. Some values suspected as erroneous are checked in the log books and corrected when necessary.

Results

In Fig.4, the temporal distribution of the number of outliers in different months is shown. Black bars represent the negative outliers, empty bars represent the positive outliers. Because of the different length of the series at different stations, the relative number of outliers instead of the real number is given as percentage of the number of stations in different years. One can argue that this approach is not quite correct, because it is rather formal and the quantities of the outliers in the years with small number of stations may be wrongly estimated, but this is the only way to put the short and long series into accordance.

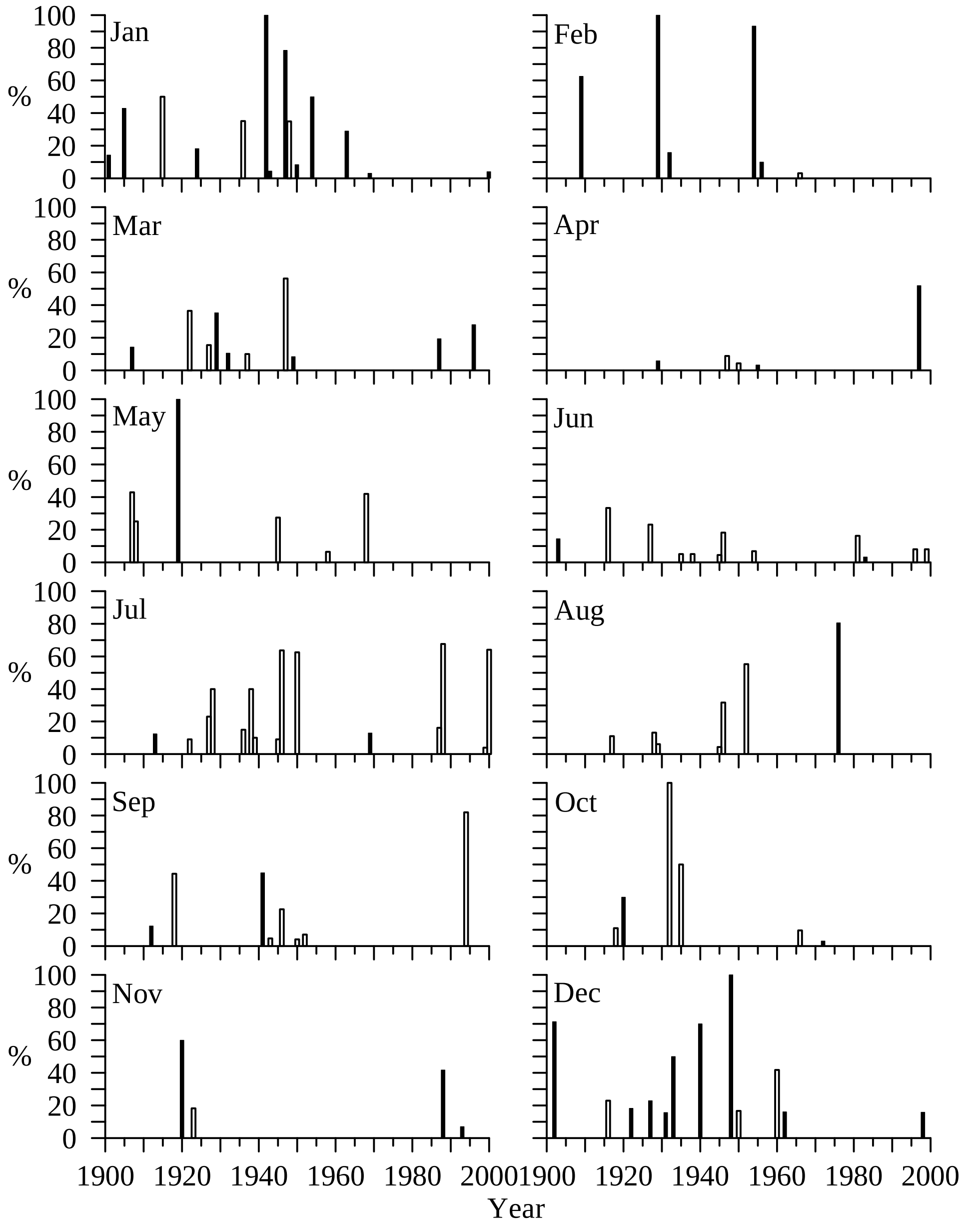

Fig.4. Temporal distribution of the number of outliers in different months (percentage of the number of stations). Positive – empty bars, negative – black bars.

In January, negative outliers dominate, but the positive ones are not negligible. The negative outliers occur in several years, the peak of the number of stations with extremely low temperatures being in 1940s and 1950s. In 1942, the cold spell causes outliers at all 20 available stations, in 1947 they are indicated at 18 out of 23 stations, but the temperatures, which are not outliers, are very close to the limit values. In 1905 and 1954, the cold temperatures are also widespread, but they reach outlying values at about half of the available stations. Exceptionally warm January temperatures occur in 1915 (with outliers at 50% of the stations), 1936 and 1948 (outliers at about 35% of the stations). An interesting fact is that the exceptionally cold January 1947 is followed by exceptionally warm January 1948.

The February temperatures show three well pronounced peaks of cold outliers: in 1909, 1929 (at 100% of the stations), and 1954. In 1954, the cold February follows the cold January, but in February, in contrast to January, the outliers are spread almost all over the country. Positive outliers are practically not indicated. The only one case is at Topolovgrad in 1966, but the temperatures are rather warm at all other stations, yet remaining within the limits.

In the third month which pertains to the winter season, December, negative outliers prevail, as in the other winter months. As it is in January and February, one of the cases with well manifested cold temperatures is in the beginning of the century, in 1902. The other well pronounced cold situations are in 1933, 1940 and especially in 1948 with outliers at all stations. The best pronounced warm situation is in 1960 when the temperatures are beyond the limits at about half of the stations, but at many stations they are not high enough to be close to the limit values.

In March, the years with cold outliers are more than the years with warm outliers, the best pronounced cases being in 1929, 1987, 1996. The outliers, however, emerge outside the limits at maximum 1/3 of the available stations – in 1929, when the cold March follows the exceptionally cold February. The best pronounced warm cases are in 1922 and 1947, with outliers at more than half of the available stations in 1947. Very warm temperatures occur also in 1989 and especially in 1990 at many stations, but they are not warm enough to trespass the limits. The comparison between January and March shows that, in 1947, the exceptionally cold January is followed by exceptionally warm March. After that the temperatures continue to be high in April, but outlying values are reached only at a few stations. Years with widespread warm April temperatures are also 1950 and 1989, but with only one outlier in 1950 and no one in 1989. The cold spell in 1997 is felt at all stations, but the limits are trespassed at about half of them. The low temperatures in 1955 are also widespread, but an outlier emerges only at one station.

The distribution of the outliers in May shows a prevalence of the positive outliers, but the cold spell in 1919 is strong enough to cause negative outliers at all available stations. Two of the cases with warm outliers and widespread high temperatures are in 1907 and 1908. Such high temperatures in the beginning of the century are not typical for the other months. In 1958, the number of outliers is less compared to 1945, but the percentage of the stations where temperatures are close to the upper limits is higher. Exceptionally high temperatures occur also in 1968 and 1969, but while in 1968 they are beyond the limits at about half of the stations, there is no one outlier in 1969.

In June, the outlying values are dispersed in numerous years, but are very few in all cases. Positive outliers prevail. The same is valid for July, together with the abundance of cases (years) with outliers. They are even more than in June and some of them occur in consecutive years. One possible explanation of the abundance of cases with outliers are the narrow limits in June and July compared to the other months. In contrast to June, however, in July, several years with widespread high temperatures, many of which trespassing the limits, can be distinguished: 1928 and 1938 with about 40% outlying values; 1946, 1950, 1988, and 2000 with more than 60 % of outlying values. The warm July temperatures in 1938 follow the few June warm cases. The same and even better pronounced consequence occurs in 1946, when the exceptionally warm July temperatures, following the June ones, continue in August and even in September, but at a diminishing number of stations. As to the negative outliers (1913, 1969), they manifest themselves in very few stations.

In August, as in the other summer months, positive outliers dominate, but are not spread in so many years as in June and July. The peak of warm temperatures is in 1952 with outliers at about half of the stations. Only one year (1976) is exceptionally cold, to an extent that 25 of 31 temperatures are outliers and the remaining are close to the limit values.

In the autumn months September and October, positive outliers continue to prevail. In September, the high temperatures in 1946 and 1952, characteristic for the previous month(s) continue to be widespread, yet the number of outliers is small. The number of positive outliers is greater in 1918 (at about 45% of the stations), and especially in 1994 (at about 80% of the stations). In 1941 September is cold all over the country with outlying values at about half of the available stations. In October, the negative outliers are very few. Well pronounced positive outliers occur in 1930, followed by 1935. In the last autumn months, November, the proportion of the positive to negative outliers is reversed, with a prevalence of negative outliers, as it is in the winter months. The two well pronounced cases of low temperatures are in 1920 and 1988, the cold spell in 1988 following the exceptionally high temperatures in July.

The cases, in which the outliers are manifested at all available stations, are rather rare: January 1942, February 1929, May 1919, October 1932, and December 1948. In only one of them, October 1932, the outliers are positive. In four cases the situation is close to these “ideal” cases: January 1947, February 1954, August 1976, and September 1994. In many cases, when the outliers are not spread over all stations, one can not find any patterns in their spatial distribution. In some cases, however, it is possible to determine regions, affected to a greater extent by the cold/warm spells. Here some examples are given. Cases when the cold temperatures affect the northern part of the country to a less extent than the other regions are January 1905, February 1909, March 1929 and 1996, November 1988. On the contrary, in January 1963, the cold spell is felt to a greater extent in North Bulgaria; in February 1956 cold temperatures emerge as outliers only at the group of three stations located in more western part of North Bulgaria (see Fig.1). In June 1981 the positive outliers are concentrated in the central regions of the country.

The seasonal variation of the proportion of negative to positive outliers, mentioned in the description of the outliers month by month, is more clearly illustrated in Fig.5, which shows the annual distribution of the outliers with opposite signs. Here, in contrast to Fig.4, the real (not relative) number of outliers is given. The number of outliers would have been greater and the estimates would have been more correct if the observations at all stations have covered the entire period since 1900 to 2000. For example, in January 1942 the temperatures trespass the respective lower limits at all 20 available stations, but it will not be correct to assert that this have occurred at the other 11 stations from the list, where temperature measurements were not taken at that time.

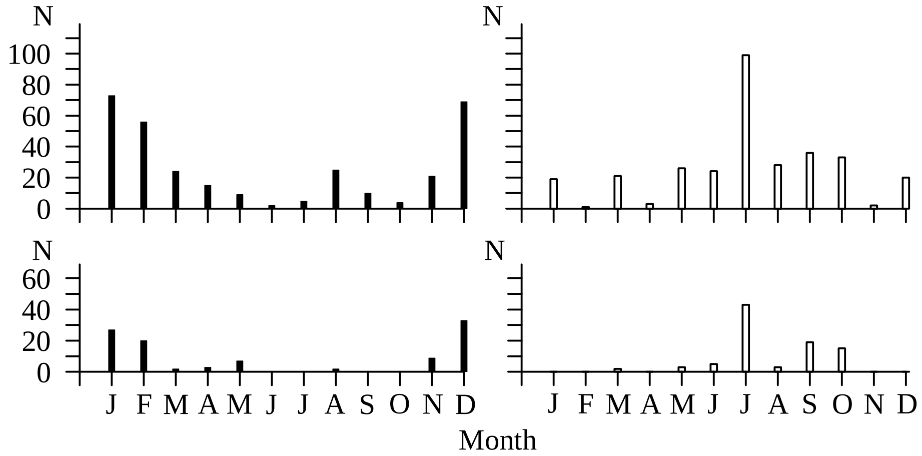

Fig. 5. Annual distribution of he number of outliers (N). Left – negative, right – positive outliers. Upper plots – with respect to the 2.5σ limits, lower plots – with respect to the 3σ limits.

Six months, from the end of spring (May) to the middle of autumn (October), the number of positive outliers is much greater than the number of negative ones with a high peak in July. Another interesting feature in these months is the rather great number of negative outliers in August. From the end of autumn (November) to the middle of spring (April), negative outliers dominate with a well pronounced increase from November to December-January and a decrease to April. As a whole (over all months) the number of positive outliers is almost equal to the number of negative ones.

The month with the greatest total number of outliers, 95% of which are positive, is July, followed by January and December with about equal number of outliers, almost 80% of which are negative. The month with the least number of outliers is April, preceded by November.

If the outliers are defined using the 3σ limits, their number considerably diminish. The lower plots in Fig.5 give an idea of this reduction. In no one case the outliers are spread over all stations. The negative 3σ outliers are 33%, the positive are 29% of the 2.5σ outliers of the respective sign. As a result of this proportion, the negative 3σ outliers have a slight prevalence to the positive ones (53 to 47%). The reduction of the number of outliers, however, is not uniform all over the year and does not depend on the annual distribution of the standard deviation. It is most considerable in March and August when the 3σ outliers are less than 10% of the 2.5σ outliers. The percentage is about 30-35% in winter (December, January, February) and about 40% in autumn (September, October, November) and July.

With the 3σ limits, one can reveal more distinctly the extremely cold/warm situations. In most months the outliers are separated, being of only one sign in certain month. The well pronounced cold cases are in January, February, November, and December; the well pronounced warm cases are in July, September, and October. Only in three months (March, May, and August) both positive and negative outliers occur.

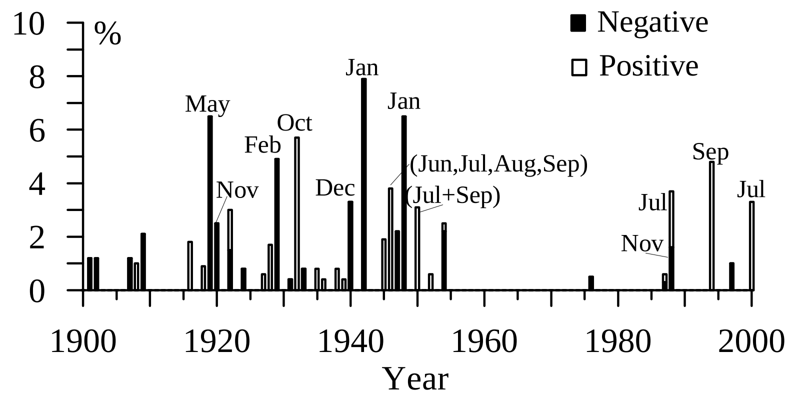

Fig. 6. Temporal distribution of the annual number of outliers, calculated with respect to the limit values $(\bar{t} \pm 3\sigma)$ (relative number)

The distribution of the annual number of the 3σ outliers over the period 1900-2000 is visualized in Fig.6. As in Fig.4, the relative number of outliers is given, as percentage of the number of stations in different years multiplied by 12 (the number of months in a year). In most cases, all outliers in a given year occur in one month, e.g. the cold May 1919, the cold February 1929, the warm October 1932, the cold December 1940, 1948, the cold January 1942, 1947, the warm September 1994, the warm July 2000. In a few years, the number of the positive outliers is a result of more than one warm month, the best pronounced case being in 1946, when exceptionally high temperatures occur consecutively in June, July, August, and September. There are also a few cases with both negative and positive outliers registered in one and the same year, e.g. in 1922 with exceptionally warm March and exceptionally cold December or in 1988 with exceptionally warm July and exceptionally cold November. In Fig.6, we have marked the months in which high number of outliers is registered in some years.

An interesting feature of the temporal distribution of the temperature outliers, which is clearly seen in Fig.6, is that they are not evenly distributed over the 20th century. In the beginning of the century, the number of the outliers is not very high. Until 1910, negative outliers prevail; after that warm temperatures produce outliers in a few years. Since 1919 till mid 1950s, there is an increase of the number and frequency of the outliers, both positive and negative. The period since mid 1950s till late 1980s is relatively calm. After that, till the end of the century, there is a new increase of the outliers' number with a dominance of the high temperatures.

In Fig.7, the temporal distribution of the magnitude of the outliers is given. These are not the temperatures themselves, but the deviations Δt from the respective lower (the negative values) and upper (the positive values) limits $(\bar{t} \pm 2.5\sigma)$. As there are coinciding values of Δt in some months/years, many of the points represent more than one case with the respective value of Δt .

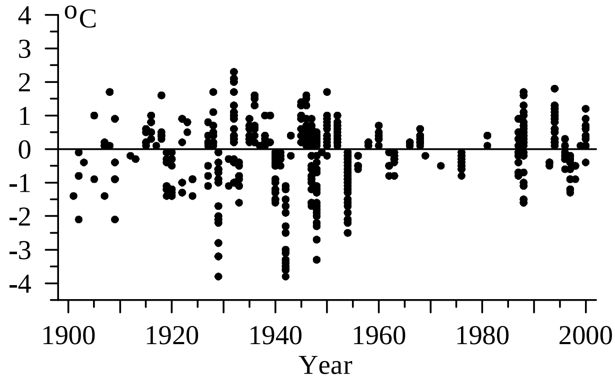

Fig. 7. Temporal distribution of the magnitude of outliers, calculated with respect to the limit values $(\bar{t} \pm 2.5\sigma)$

First, it should be mentioned that the negative outliers reach greater absolute values than the positive outliers. As one can see in the next Fig.8, the greatest deviations from the respective limits are reached in winter: up to -3.8°C in January and February and up to -3.3°C in December. The greatest negative outliers prove to be in the years with the greatest number of such outliers in the respective month (compare Fig.7 and Fig.4). These are: January – 1942 followed by 1947; February – 1929 followed by 1954; December – 1948 followed by 1940. Of course, at many stations the magnitudes of the outliers remain much smaller in these years. For example, in January 1942 when the greatest deviation of -3.8°C from the limit is registered (Yambol), the smallest nonzero deviation is -1.1°C (Kardzhali), while at Zlatograd $\Delta t = 0$, i.e. the temperature coincides with the limit value. In February 1929 the difference is from -3.8°C (Knezha) to -0.1°C (Radomir, Haskovo).

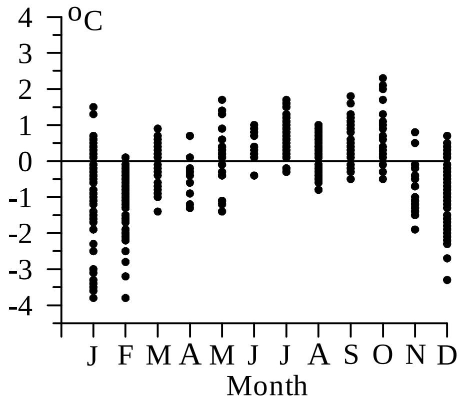

Fig. 8. Annual distribution of the magnitude of outliers calculated with respect to the limit values $(\bar{t} \pm 2.5\sigma)$

From winter to summer, the magnitudes of the negative outliers diminish together with their number and since May till October the positive values of Δt are greater by absolute value than the negative ones. In August, however, the negative outliers in 1976 reach absolute values almost as great as the positive outliers.

The correlation between the magnitude and the number of outliers is lost in July, when in spite of the greatest number the positive outliers do not reach the highest values. The highest positive values (up to 2.5°C at Dupnitsa) are reached in October 1932, the year with the greatest number of outliers in this month. The highest value of Δt in July is 1.7°C and it is almost the same in January, May, and September.

As a whole, in winter the extreme situations are better pronounced than in summer – not only the limits defined by the standard deviation are wider (see Part 2), but the values trespassing these limits (Δt) are greater, and this is also a manifestation of the greater intensity of the atmospheric processes in the cold seasons.

The variation of the outliers' magnitudes throughout the 20th century shows the same peculiarities as the variation of the outliers' number (compare Fig.6 and Fig.7, but having in mind that Fig.6 refers to the 3σ outliers, while Fig.7 represents the 2.5σ outliers). In the period since 1920s till mid 1950s the greatest number correlates with the greatest magnitudes of the outliers. During the calm period till late 1980s, the magnitudes of the outliers decrease considerably, and increase again till the end of the century together with the frequency and the number of the extreme situations. The temporal distribution of the number and magnitude of the outliers, however, does not show any well pronounced periods of prevailing negative or positive outliers. Something more, there are cases when both positive and negative outliers are registered in one and the same year, as it is mentioned above. Even in the end of the century, together with the numerous cases of extremely high temperatures, many cases of exceptionally cold temperatures are registered (e.g. March 1987, 1996; December 1988; May 1997 etc.).

It should be noticed, that the magnitudes of Δt are not informative about the real temperatures that prove to be outliers, because $\bar{t}$ and σ, which define the limits, are different in different months and at different stations. That is why a temperature that produces an outlier at a given station may be within the limits at other stations. In the first place, this is a result of the temperature difference north-south, which is best pronounced in the cold seasons. There are also differences caused by the altitude or other peculiarities of the terrain. Here some examples are given. In February 1954, the temperature of -5.5°C that produces a 2.5σ outlier with magnitude -1.7°C at Sadovo, is well within the limits at Pleven (where the lower limit is -7.3°C), Shumen (lower limit -6.4°C), and other stations. Or, the temperature of 21.4°C at Tervel, that produces a 3σ outlier of 0.5°C in September 1994, does not reach the upper limit at Shumen (21.6°C) and is well under the limits at Yambol (22.5°C), Chirpan (23.5°C) and other stations.

In the end, it is interesting to check for outliers the parts of the series before 1900. Data that can be used for that purpose are available at Obr. chiflik, Sadovo (measurements since 1891), Pleven and Kjustendil (measurements since 1896). A well pronounced case of cold temperatures is January 1893 – the magnitude of the outliers both at Obr. chiflik and Sadovo is about -2.0°C. Another case of good coincidence is the warm temperature in October 1896. It gives outliers at Obr. chiflik and Kjustendil and is on the limit at Pleven. The temperature remains within the limit only at Sadovo, but is high enough to be close to the limit. The situation is similar in July 1894, when the warm temperature produces an outlier of 2.0°C at Obr. chiflik, but does not reach the upper limit at Sadovo by 0.4°C. In other cases, e.g. September 1992, November 1897 or December 1899 the coincidence is not so good and the situation is not very clear.

Summary and conclusion

The monthly mean temperature series at 31 stations, representative for the lowlands of Bulgaria, are checked for outliers. These observations that are very distant from the mean value are due to exceptionally cold/warm temperatures in some years. The threshold values for identification of the outliers are assumed to be $(\bar{t} \pm 2.5\sigma)$, the mean temperature $\bar{t}$ and the standard deviation σ being calculated over the period 1951-1995, which is common for almost all stations.

The cases in which the outliers in a given month are manifested at all available stations in a given year are rather rare: January 1942, February 1929, May 1919, October 1932 and December 1948. In the other cases, the temperatures remain within the limits at some of the stations but are either very close to the limits or just correspond with the high/low temperatures that produce outliers. From the end of autumn (November) to the middle of spring (April), negative outliers dominate; from May to October, the number of positive outliers is much greater with a high peak in July.

The cases, in which the outliers are very well pronounced, i.e. are widespread and reach high absolute values, are mainly the cases of negative outliers. They are typical for the cold seasons and are as follows: January 1942, followed by 1947; February 1929, followed by 1954; November 1988; December 1948, followed by 1940. Well pronounced cold spells, however, occur also in May 1919 and even in summer – August 1976. The warm spells are best pronounced in September 1994 and October 1932, followed by 1935. In July, the month with the greatest number of outliers, predominantly positive, the values are not as high as in October (not the temperatures' but the outliers' values!) and are not arranged in a way that would allow to determine a year with the highest temperatures all over the country.

The greatest total number of outliers in July, many of which do not reach high magnitudes, can be explained by the narrow limits (a result from the least standard deviations in this month). In winter, when the extreme situations are better pronounced not only the limits are wider but the values trespassing these limits are greater and this is a manifestation of the greater intensity of the atmospheric processes in the cold seasons.

The temporal distribution of the number and magnitude of the outliers shows that they are not evenly distributed over the 20th century. The temperature variability in the beginning of the century can be estimated as moderate. Since 1919 till mid 1950s there is an increase of the number, frequency and magnitude of the outliers, both positive and negative. The period since mid 1950s till late 1980s is relatively calm. After that, in the end of the century, a new increase of the extreme temperature events is in progress. One can not determine, however, well pronounced periods of prevailing negative or positive outliers. Even in the end of the century, together with the numerous cases of warm temperatures (including widespread cases when the temperatures, yet being rather high, remain within the limits) many cases of exceptionally cold temperatures are registered, mainly in late autumn (November), December and spring.

In the end, we must notice, that a possible reason for some isolated cases of outliers may be the inhomogeneity of some temperature series. In many cases, however, this can influence to a greater extent the magnitude of the outliers, and does not change considerably the overall picture of the distribution of the outliers throughout the 20th century.

Acknowledgments. The authors wish to thank the National Institute of Meteorology and Hydrology (NIMH), Bulgaria, for providing the temperature data and Prof. D. Syrakov for the computer programs used in this work.

References

González-Rouco, J.F., J.L. Jimenes, V. Quesada, and F. Valero, 2001. Quality control and homogeneity of precipitation data in the Southwest of Europe, *J. Climate*, **14**, 964-978.

Lanzante, J.R., 1996. Resistant, robust and non-parametric techniques for the analysis of climate data: Theory and examples, including applications to historical radiosonde stations data, *Int. J. Climatol.*, **16**, 1197-1226.

Peterson, T.C., and R.S. Vose, 1997. An overview of the Global Historical Climatology Network temperature database, *Bull. Am. Meteorol. Soc*. **78**, 2837-2849

Peterson, T.C., R. Vose, R. Schmoyer, and V. Razuvaev, 1998. Global Historical Climatology Network (GHCN) quality control of monthly temperature data, *Int. J. Climatol.*, **18**, 1169-1179.

Syrakova, M., and M. Stefanova, 2008. Homogenization of Bulgarian Temperature Series, *Int. J. Climatol.*, (accepted)

Trenberth K.E., and D.A. Paolino, Jr, 1980. The Northern Hemosphere sea-level pressure data: Trends, errors and Discontinuities, *Mon. Wea. Rev.* **108,** 855-872.

Екстремно големи стойности в температурни редици на територията на България за периода 1900-2000 г

М. Сиракова, П. Димитрова

Резюме. Редиците от средните месечни температури в 31 станции, представителни за извънпланинската част на България са проверени за екстремно големи стойности, т.е. за стойности, твърде отдалечени от средната. Данните обхващат периода 1900-2000 г. От ноември до април доминират отрицателните, от май до октомври – положителните отклонения от граничните стойности. Общият брой на екстремно големите стойности е най-голям през юли, но абсолютните стойности на разликите между температурата и граничната стойност са по-големи през зимата. Разпределението на броя и големината на екстремно големите стойности показва, че в началото на XX век температурната променливост може да се оцени като умерена, след това, от 1919 г. до средата на 1950-те години се наблюдава нарастване на броя и големината на екстремно големите стойности (както на положителните, така и на отрицателните отклонения), периодът до втората половина на 1980-те години е относително спокоен, след което се наблюдава ново усилване на температурната променливост.