The study of long-term fluctuations in the dates of ice formation and ice destruction in the rivers of the Votkinsk reservoir catchment

V G Kalinin and V V Chichagov
Perm State University, 15, Bukireva str., 614990, Perm, Russia
vgkalinin@gmail.com

Abstract. A statistical analysis of the dates of ice formation and ice destruction in the rivers of the Votkinsk reservoir catchment was carried out, including testing the hypothesis of randomness for both the full series of observations and their parts; autocorrelation diagnostics of the time series; verification of the homogeneity hypothesis of the series of observations and the aggregate index. The results indicate that there are statistically significant changes in the ice formation and ice destruction dates occurring in recent years in the rivers of the Votkinsk reservoir catchment.

1. Introduction
Currently, the problem of global climate change is quite actual. An increase in average annual air temperature and a change in precipitation regime inevitably leads to a change in the hydrological regime of water bodies.

The ice regime of rivers, lakes and reservoirs is most sensitive to climate fluctuations. As noted in [1, 2], the effect of changes in global air temperature on long-term fluctuations of the dates of ice formation and ice destruction is characterized by spatial heterogeneity: in some regions there are positive, in others – negative trends, trends increase in autumn, and decrease in spring.

Earlier [3], we performed a study of long-term variability of the dates of ice formation in the rivers of the Votkinsk reservoir catchment, located in the northeast of the European part of Russia. The rivers with the longest observations of ice regime at the gauging stations Kolva – Cherdyn, Kama – Bondyug, Chusovaya – Kyn were chosen.

The main question in the analysis of long-term fluctuations in the dates onset of ice phases in the rivers is as follows: whether it is possible to explain the changes in these dates exclusively by random coincidence, or whether these changes are of some directed natural character and during the observation period on the rivers there were statistically significant changes in the dates onset of ice phases.

As a null hypothesis, which was subjected to verification, we used the statement that there were no statistically significant changes in the dates onset of ice phases. The following two hypothesis were considered as possible alternatives:
- the dates onset of ice phases has changed;
- the dates onset of ice phases shifted to later dates in autumn and earlier dates in spring (due to climate warming in recent decades).
The purpose of this work is to analyse changes in the dates onset of ice phases in autumn and spring periods at 12 gauging stations on different rivers of the Votkinsk reservoir catchment with an assessment of their significance level.

2. Materials and methods

Selection criteria for gauging stations: the longest duration of the observation series and the absence of gaps in them, the relative uniformity of distribution over the territory with coverage the northern, middle and southern parts, mountainous and lowland areas, with different sizes of the catchments, and the absence of factors affecting the dates onset of ice phases (discharges of industrial enterprises, groundwater discharges upstream).

For each gauging station the studies were performed as follows:

1) testing the hypothesis of randomness for the existing series and their individual parts using the reverse arrangement test (Kendall’s test for trend) [4-6];
2) diagnostics of the absence of autocorrelations in the time series using the Ljung – Box test;
3) testing the observation series using the Student, Fisher, Mann – Whitney, Kolmogorov – Smirnov and Wald – Wolfowitz homogeneity tests (hereafter the homogeneity tests) if the hypothesis of randomness is not rejected for the compared parts of the time series.

In the statistical processing of time series, the dates of the onset of ice phases were represented by natural numbers, used as a starting point on September 1 for the dates of ice formation in autumn period and March 1 for the dates of ice destruction in spring. The calculations were performed using the Statistica 10.0 package and the exact distribution of statistics for the reverse arrangement test [7].

Time series analysis was carried out for five periods: 1936–2012 (entire observation period); 1975–2012 (the period of the so-called "climate change") and 1936-1974 (the period preceding the “beginning of climate change”); 1993–2012 (last 20 years) and 1936-1992 (the period preceding the last 20 years).

The choice of periods is based on the following considerations:
1. In papers [8, 9] it was noted that steady global climate warming associated with the restructuring of the atmosphere and confirmed by data on the rise in air and ocean temperatures has been observed since the mid-70s.
2. The fact of the most significant changes in recent decades.

3. Results

When testing the hypothesis of randomness of the series of observations it became clear whether the observations made were independent and whether it can be assumed that the results of the observations were obtained during the observations of the same random variable. In table 1 shows the p-values of reverse arrangement test obtained as a result of the analysis of the onset of ice phases over the entire observation period 1936–2012 on 12 gauging stations.

| Name of the gauging station | 1. Berezovaya-Buldirya | 2. Velva-Oshib | 3. Vishera-Ryabinino | 4. Inva-Kudimkar | 5. Kosa-Kosa | 6. Lolog-Sergeevasvkiy |
|-----------------------------|------------------------|---------------|----------------------|-----------------|--------------|------------------------|
| Autumn                      | 0.282                  | 0.054         | 0.373                | 0.201           | 0.139        | 0.983                  |
| Spring                      | 0.046                  | 0.018         | 0.109                | 0.001           | 0.141        | 0.139                  |

| Name of the gauging station | 7. Obva-Karagai | 8. Silva-Podkamennoe | 9. Usva-Usva | 10. Kama-Bondug | 11. Kolva-Cherdun | 12. Chusovaya-Kyn |
|-----------------------------|-----------------|-----------------------|-------------|-----------------|-----------------|---------------------|
| Autumn                      | 0.064           | 0.146                 | 0.652       | 0.302           | 0.298           | 0.129               |
| Spring                      | 0.000           | 0.000                 | 0.227       | 0.341           | 0.038           | 0.525               |

The results obtained allow the following interpretation: for all gauging stations with a significance level of 0.054 and above no statistically significant changes in the dates of ice formations on rivers...
over 77 years were revealed; at the same time, for the dates of ice destruction statistically significant changes with a significance level of 0.018 and below were observed on 4 out of 12 gauging stations.

The application of the Ljung – Box criterion [10] showed the absence of statistically significant linear intra-row connections which eliminates the need to construct time series models for the initial data for the entire observation period under study.

Also, we performed a series of observations for normality using the Shapiro – Wilk test, which showed the possibility of using parametric Student and Fisher test along with nonparametric ones when testing the homogeneity hypothesis of time series over the investigated periods of observation.

In table 2 for periods I and II for the dates of ice formation in autumn and for the dates of ice destruction in spring, the average values of these characteristics and p-levels significance (columns 1–5, respectively) of two-sided homogeneity tests are given. For gauging stations where the hypothesis of randomness was rejected, when testing the homogeneity hypothesis preference was given to the Mann – Whitney and Kolmogorov – Smirnov tests. For all other gauging stations the Fisher and Student test were used. In addition, it was taken into account that the Mann – Whitney test compared with the Wald – Wolfowitz test has a higher power when the distributions of the compared samples differ only in the shift [11].

The results of testing the homogeneity of two time series which are given in table 1 are displayed in columns "Period I" and "Period II" table as follows: the time periods to which the acceptance of the hypothesis of randomness with p-levels significance less than 0.025, are shown in italics; the time periods to which the acceptance of the hypothesis of randomness with p-levels significance in the range from 0.025 to 0.075 correspond to underscore; the time periods to which the acceptance of the hypothesis of randomness with p-levels significance greater than 0.075 correspond to a regular font. In columns 1–3 and 5 of the table 2 with p-levels significance less than 0.025 are shown in italics; with p-levels significance in the range from 0.025 to 0.075 correspond to underscore; with p-levels significance greater than 0.075 are reproduced in plain text.

For formal reasons [12], the homogeneity of two time series can be checked using homogeneity test only when both series can be described using a random sampling model, for them the hypothesis of randomness is accepted. In table 2 these periods are highlighted with grey fill. The hypothesis of homogeneity at a given level of significance of 0.05 for the periods of ice formation can only be accepted for gauging stations: No. 1 (for all pairs of periods) and No. 12 (periods 1936–1974 and 1975–2012). In all other cases, the behaviour of the time series in period II differs from the behaviour of the time series in period I. The behaviour of the dates of ice destruction in contrast to the dates of ice formation turned out to be more stable. For six gauging stations No. 2, 3, 5, 6, 10, 11 at the significance level not lower than 0.05 the homogeneity hypothesis is accepted for both pairs of periods and for three gauging stations No. 1, 9, 12 – for one of two pairs periods. It should be noted that for gauging stations No. 2 when analysing the dates of ice destruction for the entire observation period the hypothesis of randomness was rejected.

4. Discussion

The above results indicate changes in the timing of the onset of ice phases at each of the 12 gauging stations separately. We will conduct a similar study for the entire studied region as a whole, taking into account the possible spatial connectivity of the studied water bodies. To this end, we first create an aggregate index that determines the behaviour of each of the time series of observations. Then, based on the values of this index, we draw conclusions about the presence or absence of significant trends in the timing of the onset of ice phases in the rivers of the entire study region.

The aggregate index and the corresponding time series will be performed using the principal component analysis [13, 14], implemented in the Statistica 10.0, according to the following algorithm:

1) calculate the correlation matrix of the timing of the onset of ice phases for all gauging stations for the given time intervals of observations;
2) determine the aggregate index – the 1st principal component of the correlation matrix and cumulative percentage of total variation of the onset of ice phases across all gauging stations, explained by this component;

3) find the values of the time series of the aggregate index as the value of the 1st principal component.

The proposed approach to constructing an aggregate index seems reasonable since in our case the information content of the 1st principal component turned out to be significantly higher than the information content of the other principal components of the correlation matrix.

Table 2. The results of the comparison of the dates of the onset of ice phases by periods I and II.

| Name of the gauging station | Season | Period       | Mean | p-levels |
|-----------------------------|--------|--------------|------|----------|
|                             |        | I 1936-1992  |      |          |
|                             |        | II 1993-2012 |      |          |
|                             |        | I 1936-1992  |      |          |
|                             |        | II 1993-2012 |      |          |
|                             | 3      | 4            | 5    | 6        | 7    | 8    | 9    | 10   | 11   |
| 1. Beregovaya-Buldarya      | Autumn | 1936-1992    | 54.5 | 59.4    | 0.087 | 0.780 | 0.099 | p > .10 | 0.387 |
|                             | Spring | 1936-1992    | 55.2 | 56.3    | 0.084 | 0.894 | 0.639 | p > .10 | 0.999 |
| 2. Velva-Oshib              | Autumn | 1936-1992    | 59.5 | 55.8    | 0.044 | 0.430 | 0.047 | p < .05 | 0.974 |
|                             | Spring | 1936-1992    | 59.5 | 57.6    | 0.242 | 0.178 | 0.284 | p > .10 | 0.251 |
| 3. Vishera-Ryabininino      | Autumn | 1936-1992    | 56.2 | 64.4    | 0.001 | 0.860 | 0.001 | p < .025 | 0.527 |
|                             | Spring | 1936-1992    | 57.5 | 59.3    | 0.432 | 0.528 | 0.353 | p > .10 | 0.490 |
| 4. Inv-Kudimkar             | Autumn | 1936-1992    | 52.0 | 50.1    | 0.295 | 0.833 | 0.189 | p > .10 | 0.527 |
|                             | Spring | 1936-1992    | 52.9 | 50.1    | 0.075 | 0.473 | 0.142 | p > .10 | 0.999 |
| 5. Kosa-Kosa                | Autumn | 1936-1992    | 55.8 | 61.6    | 0.021 | 0.912 | 0.021 | p < .10 | 0.790 |
|                             | Spring | 1936-1992    | 57.0 | 57.6    | 0.775 | 0.638 | 0.632 | p > .10 | 0.399 |
| 6. Lolog-Sergeevskiy        | Autumn | 1936-1992    | 58.7 | 56.4    | 0.204 | 0.474 | 0.156 | p > .10 | 0.974 |
|                             | Spring | 1936-1992    | 58.9 | 57.2    | 0.282 | 0.255 | 0.364 | p > .10 | 0.252 |
| 7. Obva-Karagai             | Autumn | 1936-1992    | 58.5 | 59.4    | 0.697 | 0.776 | 0.691 | p > .10 | 0.168 |
|                             | Spring | 1936-1992    | 54.5 | 52.2    | 0.227 | 0.498 | 0.197 | p > .10 | 0.351 |
| 8. Silva-Podkamennooe       | Autumn | 1936-1992    | 54.9 | 53.1    | 0.469 | 0.284 | 0.475 | p > .10 | 0.359 |
|                             | Spring | 1936-1992    | 54.3 | 53.1    | 0.469 | 0.284 | 0.475 | p > .10 | 0.359 |
| 9. Usva-Usva                | Autumn | 1936-1992    | 55.8 | 55.2    | 0.406 | 0.675 | 0.460 | p > .10 | 0.999 |
|                             | Spring | 1936-1992    | 55.5 | 55.2    | 0.748 | 0.030 | 0.275 | p > .10 | 0.143 |
| 10. Kama-Bondug             | Autumn | 1936-1992    | 56.0 | 55.2    | 0.660 | 0.111 | 0.414 | p > .10 | 0.490 |
|                             | Spring | 1936-1992    | 58.4 | 67.4    | 0.002 | 0.007 | 0.014 | p < .025 | 0.527 |

p < .005, p > .10
The calculation of the correlation matrix of the timing of the onset of ice phases was carried out over 5 periods: 1936–2012, 1936–1974, 1975–2012, 1936–1992 and 1993–2012 both for all 12 gauging stations and for two groups of gauging stations: 4 and 8 respectively (Tab. 3). The division into two groups was carried out by the factor analysis. Both for the dates of ice formation and ice destruction the first group included Obva – Karagay, Sylva – Podkamennoye, Usva – Usva, Chusovaya – Kyn posts, and the second group – the remaining 8 gauging stations. The factor analysis confirmed the influence of latitudinal zonality: the posts of the 1st group are located on the rivers of the southern part of the Votkinsk reservoir catchment, and the posts of the 2nd group are located on the rivers of the northern part of the catchment.

Table 3. The main characteristics of the aggregate index and its time series.

| Observation period, years | Number of g/s | Correlation coefficients | % Total variance | p-levels significance |
|--------------------------|---------------|--------------------------|-----------------|----------------------|
|                          |               | Autumn                  | Spring          | Autumn               | Spring               | Autumn | Spring | Autumn | Spring | Autumn | Spring | Autumn | Spring |
| 1936–2012                | 12            | 0.77 (0.54)             | 0.47 (0.66)     | 0.96 (0.95)         | 0.97 (0.96)          | 79     | 73     | 0.071  | 0.068  |
|                           | 8             | 0.87 (0.72)             | 0.61 (0.75)     | 0.96 (0.95)         | 0.97 (0.97)          | 87     | 84     | 0.076  | 0.136  |
|                           | 4             | 0.85 (0.58)             | 0.56 (0.79)     | 0.92 (0.83)         | 0.81 (0.91)          | 78     | 74     | 0.063  | 0.023  |
| 1936–1974                | 12            | 0.72 (0.42)             | 0.32 (0.52)     | 0.95 (0.98)         | 0.97 (0.96)          | 78     | 72     | 0.516  | 0.289  |
|                           | 8             | 0.86 (0.68)             | 0.68 (0.81)     | 0.97 (0.98)         | 0.97 (0.98)          | 87     | 87     | 0.923  | 0.718  |
|                           | 4             | 0.83 (0.50)             | 0.53 (0.73)     | 0.96 (0.89)         | 0.89 (0.93)          | 80     | 76     | 0.268  | 0.008  |
| 1975–2012                | 12            | 0.80 (0.57)             | 0.51 (0.76)     | 0.97 (0.97)         | 0.97 (0.97)          | 81     | 76     | 0.000  | 0.305  |
|                           | 8             | 0.86 (0.73)             | 0.51 (0.73)     | 0.98 (0.97)         | 0.97 (0.98)          | 87     | 83     | 0.000  | 0.329  |
|                           | 4             | 0.84 (0.62)             | 0.63 (0.85)     | 0.91 (0.85)         | 0.85 (0.92)          | 78     | 78     | 0.000  | 0.409  |
| 1936–1992                | 12            | 0.74 (0.49)             | 0.43 (0.64)     | 0.96 (0.98)         | 0.96 (0.96)          | 79     | 73     | 0.246  | 0.320  |
|                           | 8             | 0.89 (0.74)             | 0.61 (0.76)     | 0.97 (0.98)         | 0.96 (0.97)          | 89     | 85     | 0.287  | 0.588  |
|                           | 4             | 0.85 (0.57)             | 0.59 (0.78)     | 0.95 (0.83)         | 0.80 (0.91)          | 79     | 77     | 0.269  | 0.049  |
| 1993–2012                | 12            | 0.69 (0.31)             | 0.42 (0.74)     | 0.96 (0.98)         | 0.98 (0.97)          | 73     | 76     | 0.002  | 0.233  |
|                           | 8             | 0.75 (0.55)             | 0.51 (0.74)     | 0.96 (0.98)         | 0.98 (0.99)          | 78     | 83     | 0.001  | 0.209  |
|                           | 4             | 0.84 (0.54)             | 0.57 (0.83)     | 0.86 (0.78)         | 0.84 (0.92)          | 73     | 76     | 0.007  | 0.288  |

Note. In the column “Correlation coefficients” in the column “Min.” first the minimum element of the correlation matrix of the ice phase occurrence periods for the corresponding group of gauging stations is given. In brackets the minimum of the correlation coefficients between the dates of the ice phase occurrence at the gauging stations and the value of the aggregate index is given. In the “Max.” Column the maximum values of the correlation coefficients are reproduced in the same way.

The main characteristics of the aggregate index and the results of the analysis of the corresponding time series using the reverse arrangement test for each of the groups of gauging stations for the studied time periods are given in Tab. 3. The cumulative percentage of total variation explained by the 1st principal component was from 73 to 89% for the dates of ice formation, and from 72 to 87% for the dates of ice destruction. This means that these terms at each of the gauging stations are largely determined by the value of the aggregate index. Note that the dividing of 12 posts into 2 groups
allowed us to select 8 posts with a higher cumulative percentage of total variation. In the last column of the tab. 3 shows the values of the level of significance of the reverse arrangement test applied to the analysis of the time series of the aggregate index for different groups of the gauging stations and time periods.

When analysing the dates of ice formation, the hypothesis of randomness is rejected at a significance level of 0.05 only for the periods 1975–2012 and 1993–2012. At the same time, for the dates of ice destruction for the same periods the hypothesis of randomness is accepted. These conclusions are valid both for all 12 hydrological stations and for the two selected groups of the rivers of the northern and southern parts of the Votkinsk reservoir catchment. At the same time, an interesting fact emerged: for the periods 1936–1974, 1936–2012 and for gauging stations of the southern group the hypothesis of randomness for the dates of ice destruction is also rejected at the 0.05 significance level. It may indicate the presence of earlier statistically significant changes in this part of the Votkinsk reservoir catchment compared to its northern part.

5. Conclusion
A statistical analysis of the dates of the onset of ice phases was carried out for each of the 12 gauging stations on the rivers of the Votkinsk reservoir catchment, as well as for the entire studied region as a whole using statistical test of randomness and homogeneity and an aggregate index. The results indicate that there are statistically significant changes in the dates of the onset of ice phases occurring in recent years on the rivers of the entire study area.

References
[1] Ginzburg B M and Soldatova I I 1996 Mnogoletnie kolebaniya srokov zamerzaniya i vskrytiya rek v razlichnyh geograficheskikh zonah [Long-term fluctuations in terms of freezing and opening of rivers in various geographical areas] Meteorologiya i gidrologiya 6 pp 101–108
[2] Ginzburg B M and Gurova I N 2007 Ledovyj rezhim rek Rossii v usloviyah sovremennogo klimata i ego vozmozhnye izmeneniya [Ice regime of Russian rivers in the current climate and its possible changes] Vestnik Moskovskogo universiteta. Ser. geograficheskaya 6 pp 31–35
[3] Kalinin V G and Chichagov V V 2014 Long-term variability of ice formation dates on the rivers of the Votkinsk reservoir catchment Russian Meteorology and Hydrology 39 7 pp 491–497
[4] Bendat J S, Piersol A G 2010 Random Data. Analysis and Measurement Procedures Wiley & Sons p 604
[5] Meylan P, Favre A–C and Musy A 2011 Predictive Hydrology. A Frequency Analysis Approach Taylor and Francis Group p 212
[6] Becka T W, Housha T J, Weibr J P, Cramer J T, Vardaxis V, Johnson G O, Coburn J W, Malek M H and Mielke M 2006 An examination of the Runs Test, Reverse Arrangements Test, and modified Reverse Arrangements Test for assessing surface EMG signal stationarity Journal of Neuroscience Methods 156 pp 242–248
[7] Ivchenko G I and Medvedev Y I 1992 Matematicheskaya statistika [Mathematical statistics] M: Vysshaya shkola shkola p 304
[8] Shimaraev M N 2007 Cirkulyacionnye faktory izmenenij ledovo-termicheskogo rezhima Bajkala [Circulation factors of changes in the ice-thermal regime of the Lake Baikal] Geografiya i prirodnye resursy 4 pp 54–60
[9] Frolov A V 2013 Vklad gidrometeorologicheskogo nauki v razvitie vodohozyaystvennogo kompleksa strany VII Gidrologicheskij s”ezd: tez. plenarnyh dokl. Sant-Peterburg [Contribution of hydrometeorological science to the development of the country's water management complex] pp 1–5
[10] Box G E P, Jenkins G M and Reinsel G C 1994 Time Series Analysis: Forecasting and Control Prentice-Hall, Englewood Cliffs, NJ
[11] Bagdonavicius V, Kruopis J and Nikulin M 2011 Nonparametric tests for complete data Great Britain and the United States by ISTE Ltd and John Wiley & Sons, Inc p 321
[12] Rozhdestvensky A V and Chebotaryov A I 1974 Statisticheskie metody v gidrologii [Statistical Methods in Hydrology] p 424
[13] Meshcherskaya A V, Rukhovets L V, Yudin M I and Yakovleva N I 1970 Estestvennyye sostavlyayushchie meteorologicheskih polej [Natural components of meteorological fields] p 199 (L.: Gidrometeoizdat)
[14] Jolliffe I T 2002 Principal Component Analysis (NY: Springer) p 518