Selection of reference stream gauges at the river Ufa catchment

ELENA BELOZEROVA[1,2], NATALIYA KRASNOGORSKAYA[1], ANTONIA LONGOBARDI[2],
ELVIRA NAFIKOVA[1],

[1] Department of Production Safety and Industrial Ecology,
Ufa State Aviation Technical University,
Karl-Marx-Str. 12, 450000, Ufa,
RUSSIAN FEDERATION
[2]Department of Civil Engineering
University of Salerno
University Via Giovanni Paolo II, 132, Fisciano (SA)
ITALY
elena.belozerova.ufa@gmail.com, alongobardi@unisa.it, nk.ufa@mail.ru, vira2006@yandex.ru

Abstract: - The paper is devoted to the problem of selection of reference stream gauges to estimate the average daily streamflow for ungauged catchments. Daily streamflow data are crucial for several fields such as water management, rational resource utilization, hydropower energy production, but the number of ungauged watersheds is still big and it is needed a large scale of regional approaches as a practical response to this problem. In the following, results of an application to the catchments of the rivers: Ufa, Ai, Bolshoj Ik, Tjuj, Sars and Jurjuzan are presented. The map correlation method has been performed [1], where differences is made between the "closest" stream gauge and the "best correlated" stream gauge criterion for reference station selection. It was found that selection of the nearest stream gauges do not usually provide high correlation between their daily streamflow values. Whereas choosing most correlated stream gauge as a reference perform better results in case of goodness of fit parameters (lower errors and higher NSE), compared to the nearest stream gauge.

Key - Words: - Daily streamflow, ungauged catchment prediction, reference stream gauge, Russian Federation.

1 Introduction

Hydrological data, in particular the daily streamflow time series, have a large number of applications, but at the same time the network of hydrological stations is limited. Regional approaches provide practical and challenging solution to this problem, with regional relationship to be used for prediction in ungauged cases. This question is especially actual for Russian Federation, due to a wide area with uneven distribution of gauges throughout the country, in terms of turning to a rational environmental management.

The estimation of daily streamflow at an ungauged catchment commonly based on the drainage area ratio method:

$$Q_{un} = \frac{F_{un}}{F_{ref}} Q_{ref} \qquad (1)$$

where Q_{un} is the streamflow at the ungauged site, m^3/s;

Q_{ref} is the streamflow at the reference stream gauge, m^3/s;

F_{un} is the drainage area of the ungauged catchment, m^2;

F_{ref} is the drainage area to the reference stream gauge, m^2;

Present method assumes that meanings of streamflow at ungauged area change symbatically with the reference one.

If daily streamflow time series of reference stream gauge and ungagged area were well correlated it would mean the matching of timing of water regime phases, regardless of the magnitude of the stream flows. According to Smakhtin et al. [2] correlation is the index by which a reference stream gauge should be selected.

Since the correlation between an ungauged catchment and a reference stream gauge cannot be calculated, it forms several alternative solutions such as:

selecting the nearest stream gauge [3];

several nearby stream gauges [2, 4];

selection of reference stream gauge based on the ratio of drainage area between 0.5 and 1.5 [5].

Current paper proves the statement [1] that nearest stream gauges is not obligatory the best-correlated ones.

Moreover, the analysis results demonstrate that selection of the best – correlated stream gauges as reference stream gauges outperform the estimates of daily stream time series compared with nearest stream gauges and that there is also a relation between correlation and goodness of fit values.

2 The case study

The object under investigation is the catchment area of the river Ufa, within the territory of the Republic of Bashkortostan. The river Ufa is a third order inflow of the river Volga and the largest right-bank inflow of the river Belaya. Apart from the Republic of Bashkortostan, the Ufa catchment area takes place in three more subregions of Russia: Chelyabinsk oblast, Sverdlovsk oblast and Perm Krai. The total drainage area is 51837 km^2, and 24 483 km^2 (thus 47, 2% of the catchment area) fall within the territory of Bashkortostan. along the river network, the Ufa river receives inflows from different tributaries. The largest are the Ai (with inflow the Bolshoj Ik), the Tjuj (with inflow the Sars), the Jurjuzan (Figure 1).

The total length of the river Ufa is 918 km. There are 85 tributaries of less than 10 kilometers, flowing directly into it; the total length of the river network is 863 km. River valley in the upper part is mainly V-shaped. In the lower reaches the valley becomes wide, with the channel width ranging between 300 - 400 m). Sinuosity ratio is 1.7 on average [6, 7].
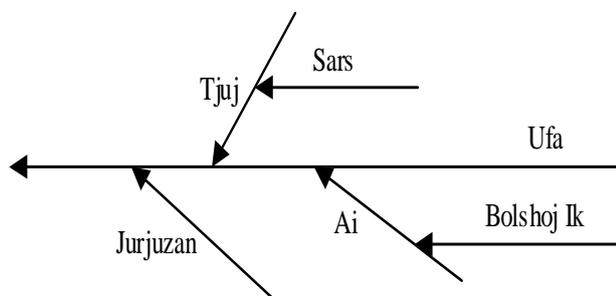
Fig 1. –Scheme of the river Ufa and main tributaries

In this study, 9 stream gauges stations located along the river Ufa are considered. Streamflow values at these stream gauges have a common 4-year period of record extending from 1 January 2009 till 31 December 2012, the results of more than 13 000 daily runoff records have been analyzed (Figure 2, Table 1).
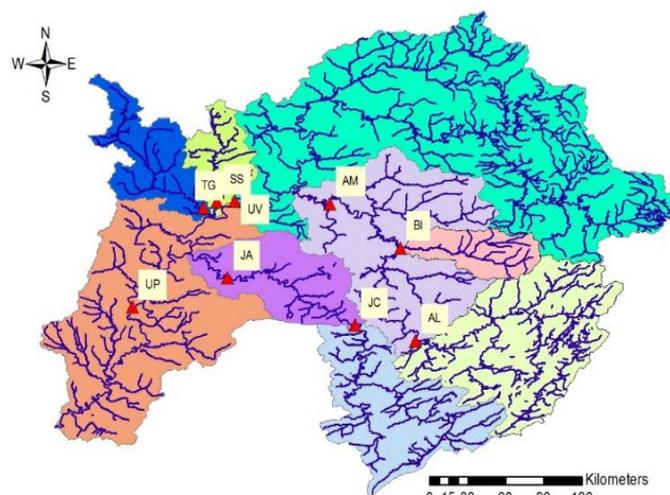
Fig 2. – Scheme of location of stream gauges under study.

Table 1. – Information about stream gauges under study [8].

Stream gauge name	Code	Drainage area (km^2)
1.Ai, Lakli	AL	6 440
2.Ai, Meteli	AM	14 200
12.Bolshoj Ik, Taishevo	BI	1 450
24.Sars, Sultanbekovo	SS	1 300
29.Tjuj, Gumbino	TG	2 180
32.Ufa, Verhnij Sujan	UV	32 400
33.Ufa, Pavlovskaja Hydropower Plant	UP	47 100
36.Jurjuzan, Atnjash	JA	6 930
37.Jurjuzan, Chulpan	JC	4 850

2.1 Hydrological regime

Annual average water flow in the estuary is 388 m^3/s. The hydrological regime is dominated by snow melting. The proportion of spring runoff is about 55-65% of annual runoff, while summer-autumn is up to 26-44%. During the low flow period, rivers of mountainous regions (the Ai with tributaries) receive a significant amount of summer-autumn rainfall. In some cases, this would causes flooding.

During the summer low water, the discharge of the Ufa basin is usually 1.6-2 times higher than in

winter period. The winter runoff depends on the total area moisture and the regulating ability of watercourses. Rather high values of winter streamflow is specific for the Yuryuzan, flowing through the hydrated territory of the Ural region (7-12% annual). Winter low flow starts in the second half of November, the minimum runoff observed in February and March, sometimes in January. The maximum streamflow of the Ufa basin occurs in April and May [6]. Water discharge in the period of spring floods may exceed the winter low flow water discharge up to 20 times.

There are several factors, affecting the hydrological regime, such as: geological and climate properties, land cover and hydrography and the anthropogenic features. They are described in the following

2.2 Geological properties

The geological structure of the Ufa river catchment is a platform area represented by the Ufa plateau of the Yuryuzan - Ai plain [9].

The surface of the Ufa plateau (with an average height of 380 - 460 m), is densely dissected by rivers Ufa, Yuryuzan and Ai. A karst phenomena is developed, that promotes the increase of natural flow regulation. In this regard, on the plateau surface runoff and the local river network is very poorly developed. Surface water is quickly transferred to the underground karst system.

The Yuryuzan - Ai plain is located between the Ufa plateau and the western slope of the Urals. Flat relief characterizes the western part of the plain. Here, karst phenomena influences the hydrological regime of rivers. The average height of the surface of Yuryuzan - Ai plain within the catchment area is 250 - 400 m, increasing from 180 m in the Ai river catchment area to 530 m on the north-east of plain Belokataysky upland).

2.3 Land cover and hydrography

Distribution of soils in the catchment area of the river Ufa agreed with latitudinal-altitudinal zonation [10]. Dark coniferous forests on sod-podzolic and light-gray forest soils grow in the western part of Yuryuzan - Ai plain. On the Ufa plateau forest covers over 90% (mainly grow the spruce-fir forests with few linden). Such a high value of forest cover leads to the lowest values of surface runoff - 13 - 15 mm [11]. Northeast of the Yuryuzan - Ai plain is fully covered by forests.

Soils within the forest-steppe zone represented by sod-podzolic, sod-carbonate, light gray and humus.

The level of development of the river network is determined by the drainage density, which is the ratio of the sum of the lengths of all the rivers over the area of the related drainage basin. The lowest value of the coefficient (0.21 - 0.40 km / km^2) is typical for the northern part of the Ufa plateau, due to the development of karst phenomena, the highest values were observed in the upper reaches of rivers the Ai and the Yuryuzan (0.51 - 0.60 km / km^2) [7, 12].

2.4 Climate properties

The climate is continental, characterized by long cold winters, warm, sometimes hot summers, with high amplitude of air temperature fluctuations. The Ural Mountains prevent inflow of western air masses carrying moisture from the Atlantic Ocean, from Siberia, which provides favorable conditions for precipitation in the Urals area. The tropical air masses incoming from the sea bring thaw in winter and coolness in the summer. Invasion of arctic air in summer and continental air from Siberia in the winter cause sharp temperature drop. The average annual temperature in the river Ufa basin is 0.7 - 2.5 C. The coldest month is January, the hottest - July. The average duration of the frost-free period is of 80 - 130 days [6, 7]. Annual precipitation changes on the territory within 550-800 mm. The bulk of precipitation (60-70% of the annual amount) falls in the warm season - from April to October. The minimum amount of rainfall recorded in February, the maximum - in July, heavy rains fall mainly in July and August [6,13].

The distribution of solid precipitation on the territory of the river Ufa basin is considerably variable. First sign of the snow cover appears in the mountains in the middle of October, and in the plain by the end of this month. The snow cover sets usually in early November. Within the Ufa plateau the highest values of snow depth (200-250 mm), is observed in Yuryuzan - Ai plain is inherent by the lowest rates (100-150 mm). The depth of soil freezing is 110-120 mm. Evaporation loss from the river Ufa catchment surface is coherent with the laws of geographic zoning. The average annual evaporation from the catchment surface is 460-480 mm [7].

2.5 Anthropogenic condition

The main anthropogenic changes of the total water flow in the river Ufa are: the increase of water loss by evaporation from the water surface, changes of the conditions of surface runoff formation, water intake for

industrial, domestic and agricultural needs, urbanization.

Thus, the catchment under study is of practical interest by several reasons. The river Ufa provides water to several big cities, industries, agricultural complex, at the same time it is a source of hydroenergy, an area for recreation and tourism, waterways and fisheries.

3 The correlation structure for the Ufa river basin

The most commonly - used measure of correlation is Pearson's r. It is also called linear correlation coefficient because r measures the linear association between two variables. If the data lie exactly along a straight line with positive slope, then r = 1.

Person's r is computed assuming that the data follow a bivariate normal distribution [14]. With this distribution, not only do the individual variables x and y follow a normal distribution, but their joint variation also follow a specified pattern. Thus, r is often not useful for describing the correlation between untransformed hydraulic variables such as runoff. Pearson's r is invariant to scale changes, as in converting streamflows in cubic feet per second into cubic meters per second, etc. This dimensionless property is obtained by standardizing, dividing the distance from the mean by the sample standard deviation, as shown in the formula 2.

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2]\cdot[n\sum y^2 - (\sum y)^2]}} ; \quad (2)$$

Person's correlation coefficient was calculated from daily streamflow for each couple of stream gauges.

In a case of direct dependency between distance and correlation (the lower distance the higher correlation), it is likely to conclude that distance is a good criteria to choose the reference stream gauge. The relations between distance and r for the study stream gauges are presented in Figure 3.
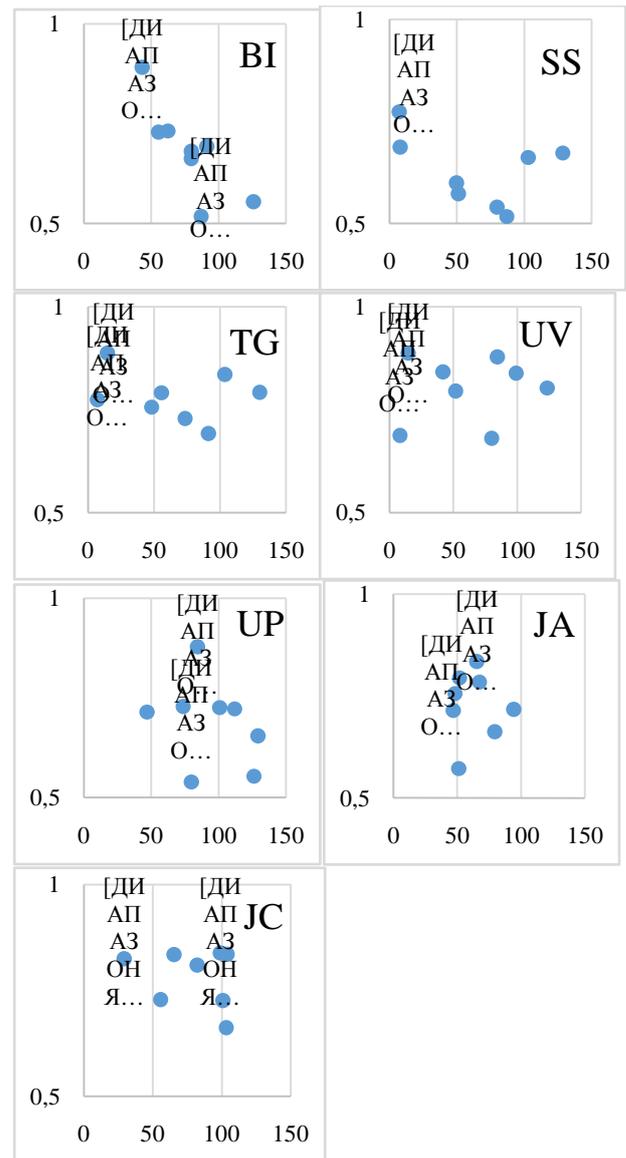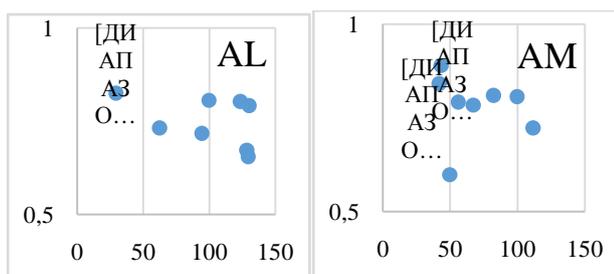
Fig 3. - Relation between distance and Pearson's r correlation coefficient

As can be seen from the graph Fig 3, streamgauges: Ai- Lakli, Ai- Meteli, Tjuj - Gumbino, Sars- Sultanbekovo, Jurjuzan - Chulpan, Bolshoj Ik -Taishevo show apparent relation between distance and r coefficient, however one third of stream gauges under study demonstrated weak or nearly no relation between distance and r: Ufa - Verhnij Sujan, Ufa - Pavlovskaja HPP, Jurjuzan - Atnjash. Significantly, that for the 5 stream gauges from 9 the best correlated stream gauge is not the nearest one. Thereby for this particular area the principle of choosing the closest stream gauge as a reference is not always valid. The stream gauge with the highest r value typically was the next to the nearest stream gauge. The relations between the highest values of correlation and distance are presented on Fig. 4.
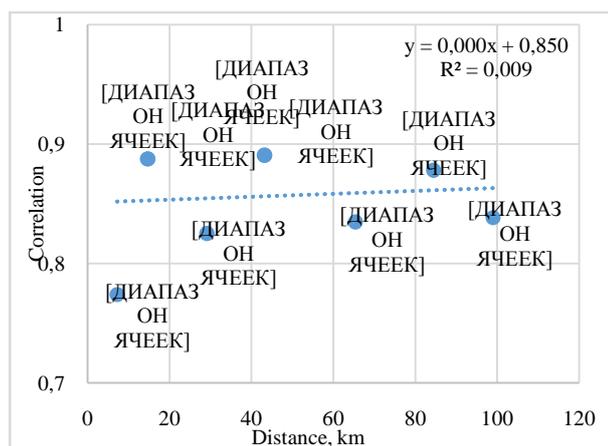
Fig. 4- Relation between the highest values of correlation and distance

4. Validation of the map correlation method for the Ufa river basin

To compare the two selection criteria for reference stream gauge (the nearest stream gauge versus the most correlated stream gauge), the drainage area ratio method (1) was applied to each of the study stream gauges using both the nearest stream gauge as the reference stream gauge and the most correlated stream gauge as the reference stream gauge (the stream gauge having the highest value of r). To evaluate goodness of fit, the Nash-Sutcliffe efficiency value E [15] and mean percentage error were calculated from both the observed and the estimated streamflows.

The efficiency E proposed by Nash and Sutcliffe is defined as one minus the sum of the absolute squared differences between the predicted and observed values normalized by the variance of the observed values during the period under investigation. It is calculated as:

$$E = 1 - \frac{\sum_{t=1}^{T}\left(Q_o^t - Q_m^t\right)^2}{\sum_{t=1}^{T}\left(Q_o^t - \overline{Q_o}\right)^2}; \qquad (3)$$

where Q_o is the mean of observed discharges, and Q_m is modeled discharge. $Q_o^{\,t}$ is observed discharge at time t.

The normalization of the variance of the observation series results in relatively higher values of E in catchments with higher dynamics and lower values of E in catchments with lower dynamics. To obtain comparable values of E in a catchment with lower dynamics the prediction has to be better than

in a basin with high dynamics. The range of E lies between 1.0 (perfect fit) and $-\infty$. An efficiency of lower than zero indicates that the mean value of the observed time series would have been a better predictor than the model [16].

Results are presented in Tables 2 and 3.

Table 2. Stream gauge with the highest correlation

Study Stream gauge	Most correlated stream gauge	Distance, km	Corre lation	Error in Q, %	NSE
AL	**JC**	**29,14**	**0,83**	**4,30**	**0,68**
AM	BI	43,23	0,89	6,76	0,47
BI	**AM**	**43,23**	**0,89**	**6,33**	**0,75**
SS	**TG**	**7,28**	**0,77**	**20,24**	**0,18**
TG	UV	14,73	0,89	34,83	0,66
UV	TG	14,73	0,89	53,44	0,23
UP	UV	84,40	0,88	1,48	0,198
JA	JC	65,32	0,83	61,93	0,21
JC	UV	98,93	0,84	24,14	0,62

Table 3. Nearest stream gauges

Study Stream gauge	Nearest Stream gauge	Distance, km	Corre lation	Error in Q, %	NSE
AL	**JC**	**29,14**	**0,83**	**4,30**	**0,68**
AM	UV	41,79	0,84	22,98	0,57
BI	**AM**	**43,23**	**0,89**	**6,33**	**0,75**
SS	**TG**	**7,28**	**0,77**	**20,24**	**0,18**
TG	SS	7,28	0,77	25,38	0,59
UV	SS	7,95	0,69	92,38	-0,22
UP	JA	46,87	0,72	108,5	-4,60
JA	UP	46,87	0,72	52,04	0,10
JC	AL	29,14	0,83	4,49	0,54

As seen from the tables, for couples AL-JC, BI-AM, SS-TG the closest and most correlated criteria do correspond. In the case of stream gauges AM and JC difference in correlations between most correlated and closest are not inessential (0,89 versus 0,84, 0,84 versus 0,83) respectively. For table 2 the correlation coefficient varies from 0,89 (BI – AM) to 0,77 (SS - TG) on average - 0,85, which is high, according to the correlation scale. The lowest mistake is noted for couple UP-UV. And the highest for UV-TG, JA-JC. Unsurprisingly that highest NSE meaning are typical for pairs with both high correlation and low mistake. However, it is uncommon that SS-TG, UP-UV show undersized NSE values. Interestingly that streamflow gauges AM and BI, TG and UV generate double couples herewith UV being the most frequently considered

reference stream gauge, belonging to different pairs. These strong relations may be explained by the central position of the recording station fig 2.

In regard to table 3 there are 3 double couples. AL and JC in both combinations show stably low errors and high NSE with a minor distinction between each coefficient. Errors in SS and TG pairs are lowest among all and almost similar, only NSE coefficients are deviated Whereas UP and JA couples presented dramatic gaps in all indicators, the most significant errors up to 100% and even negative NSE, either in UV-SS case. Compare to the previous table, correlation coefficients, errors and NSE of nearest stream gauges are always in agreement among themselves and change unidirectionally. The correlation coefficient for the closest stream gauges is fluctuating within 0,89 (BI-AM) and 0,69 (UV-SS) with average 0,78, which is less than for most correlated couples, but still high. It can be noted, that only AL-JC, SS-TG, UP-UV, JA-JC and JC-AL fulfill the Hortness's [5] condition about drainage catchments ratio within the 0.5-1.5 range.

For a graphic interpretation of NSE value distribution, the box-plot (Fig. 3) was built. So providing that the reference stream gauge is a stream gauge with the highest r value (1) the median is slightly lower, than for the nearest stream gauges (2), but in first case the range of data is lower and they are evenly distributed above and below the median. That makes box-plot more compact and symmetric. Oppositely, for the second the disproportion of 2^d quartile is obvious and an outlier UP-JA takes place.
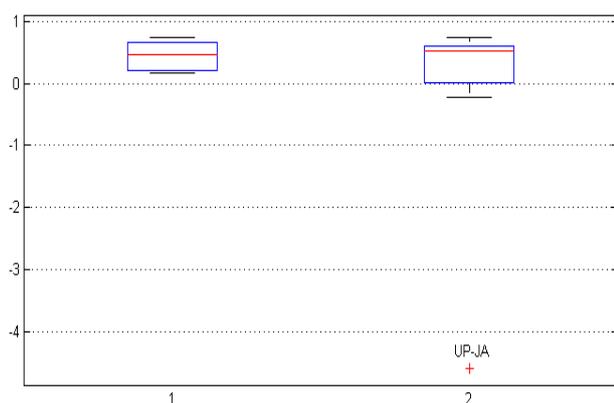
Fig 3. Box – plot NSE

5 Discussions

In the previous part is was already mentioned that, selection of most correlated stream gauges as reference results in better values of goodness of fit measures, but not the best possible. Table 4 presents the stream gauge couples with lowest errors and highest NSE criteria.

Table 4 – Stream gauges with lowest errors and highest NSE

Study Stream gauge	Lowest error	Error, Q, %	Highest NSE	NSE
AL	BI	3,13	JC	0,680119
AM	**JC**	**0,93**	**JC**	**0,578924**
BI	AL	3,04	AM	0,750991
SS	JA	10,00	TG	0,179730
TG	SS	25,38	UV	0,662158
UV	**UP**	**1,50**	**UP**	**0,699908**
UP	**UV**	**1,48**	UV	0,198366
JA	SS	9,09	TG	0,431641
JC	**AM**	**0,94**	**AM**	**0,644238**

Bold type is used to highlight the stream gauges combination with both lowest values of errors and highest NSE, and the gray cells indicate that these couples are also most correlated. So, NSE coefficient mostly agrees with the highest r value. Surprisingly, that stream gauges AM and LC form a double couple and show error less than 1%, but they are nor most correlated, neither closest. Oppositely, stream gauges TG and SS presented one of the weakest relations.

This brief analysis brings out a limitation of only pro-correlation view point and encourages to further investigation. In particular, to test other goodness of fit measures and to look in more detail to conditions of runoff formation.

Thus, the results of current research firstly indicate that selection of nearest stream gauge as reference as well as selection based on drainage area ratio are not provide high correlation between daily streamflow values. Secondly, choosing most correlated stream gauge as a reference usually demonstrate better results relatively to goodness of fit parameters (lower errors and higher NSE), compared to the nearest stream gauge.

References:

[1]. Archfield, S. A., Vogel, R. M., Map-correlation method: Selection of a reference streamgage to estimate daily streamflow at ungaged catchments, *Water Resources Research,* Vol. 46, (10), 2010, W10513.

[2]. Smakhtin, V. U., Hughes D. A., Creuse-Naudin E., Regionalization of daily flow characteristics in part of the Eastern Cape, South Africa, *Hydrological Sciences Journal,* Vol. 42(6), 1997, pp. 919–936.

[3]. Mohamoud, Y. M., Prediction of daily flow duration curves and streamflow for ungauged catchments using regional flow duration curves, *Hydrological Sciences Journal,* Vol. 53(4), 2008, pp. 706–724.

[4]. Smakhtin, V. U., Generation of natural daily flow time series in regulated rivers using a nonlinear spatial interpolation technique, *Regulated Rivers: Research & Management*, Vol. 15, 1999, pp. 311–323.

[5]. Hortness, J. E., *Estimating low-flow frequency statistics for unregulated streams in Idaho*, U.S. Geological Survey Scientific Investigations Report 2006-5035, 2006.

[6]. Gareev, A. M., *Rivers and lakes of Bashkortostan*, Kitap, 2001. [Russian]

[7]. Rogozina, T. A. Issledovanie kachestva rechnyh vod v uslovijah antropogennogo vozdejstvija :na primere Ufimskogo bassejna. PhD dissertation, [The study of river water quality under the anthropogenic influence: applied to the river Ufa basin, PhD dissertation], Saint Petersburg, 2008, 202 p. [Russian]

[8]. *Surface water resources of the USSR*. Vol.11, no. 1, Gidrometeoizdat, 1973, [Russian]

[9]. Balkov, V.A. *Water resources of Bashkiria*, Ufa, 1978. [Russian]

[10]. Brehovskih, V.F., Volkova, Z.V., Modern ecological status of some water bodies of Central Russia, *Water problems in the turn of the century, Russian Academy of Sciences. Institute of water issues,* 1999, pp. 208-216. [Russian]

[11]. Molchanov, A.A., *The hydrological role of forest,* Publishing house of Academy of Sciences of the USSR, 1970. [Russian]

[12]. *Atlas of the Republic of Bashkortostan*, Omsk cartographic factory, 1992. [Russian]

[13]. Afanasev, I, Volkova, T., Elizaryev, A, Longobardi, A., Analysis of interpolation methods to map the long-term annual precipitation spatial variability for the Republic of Bashkortostan, Russian Federation, *WSEAS Transactions on Environment and Development*, Vol. 10 (1), 2014, pp. 405-416.

[14]. Helsel, D.R., Hirsch, R. M., *Statistical Methods in Water Resources Techniques of Water Resources Investigations*, Book 4, chapter A3. U.S. Geological Survey, 2002.

[15]. Nash, J.E., Sutcliffe, J.V., River flow forecasting through conceptual models part I—A discussion of principles, *Journal of Hydrology,* Vol. 10(3), 1970, pp. 282–290.

[16]. Krause P., Boyle D. P., Base F., Comparison of different efficiency criteria for hydrological model assessment, *Advances in Geosciences*, Vol. 5, 2005, pp. 89-97.