variable X, S_{yy} is the sum of the residual squares of the variable Y, S_{xy} is the covariance of the residuals of the random variables X and Y.

The correlation with two independent variables is given by

$$Y = a + b_1(X_1 - \overline{X_1}) + b_2(X_2 - \overline{X_2})$$

where a, b_1, b_2 are regression parameters found by the method of least squares, which is quite standard approach in regression analysis. The regression parameters are given by

$$a = \overline{Y}$$

$$b_1 = \frac{S_{x_2x_2} \cdot S_{x_1y} - S_{x_1x_2} \cdot S_{x_2y}}{S_{x_1x_1} \cdot S_{x_2x_2} - S_{x_1x_2}^2}$$

$$b_2 = \frac{S_{x_1x_1} \cdot S_{x_2y} - S_{x_1x_2} \cdot S_{x_1y}}{S_{x_1x_1} \cdot S_{x_2x_2} - S_{x_1x_2}^2},$$

where $\overline{X_1}$ is the average value of the random variable X_1, $\overline{X_2}$ is the average value of the random variable X_2, $\overline{Y}$ is the average value of the random variable Y, $S_{x_1x_1}$ is the variance of the random variable X_1, $S_{x_2x_2}$ is the variance of the random variable X_2, S_{yy} is the variance of the random variable Y, S_{x_1y} is the covariance of the random variables X_1 and Y, S_{x_2y} is the covariance of the random variables X_2 and Y, $S_{x_1x_2}$ is the covariance of the random variables X_1 and X_2.

The regression equation describes a way how the variable Y is dependent of one or more independent random variables X_i. In defining and obtaining a regression, as was mentioned before, the method of least squares, is one of the most investigated and applied techniques in applications.

The main difference between correlation and regression is that in correlation, you sample both measurement variables randomly from a population, while in regression you choose the values of the independent (X) variable.

The dependence between Y_i and X_i can be written as

$$y_i = a + b \cdot (x_i - \overline{X}) + e_i, \qquad i = 1, 2, ..., n$$

where $\overline{X}$ is the average value of the random variable X, e_i are the errors (errors of the model and errors of the measurement of the variable Y_i), a, b are parameters (coefficients) which can be determined with data obtained by the measurement, n is the number of measurements of X_i and Y_i.

The parameters a and b need to be determined from the variables X_i and Y_i. Usually, when we are dealing with applications, we have limited number of measured data n, and obtained parameters, denoted with a and $\hat{b}$, are their estimates, not exact values of the parameters a and b. Calculated value y_i, obtained by regression, by using exact values x_i and estimated parameters a and $\hat{b}$, is again just an estimate (most probable) value and is denoted by y_i. Therefore, the regression is given by

$$y_i = a + \hat{b} \cdot (x_i - \overline{X}), \quad \text{where } i = 1,2,...,n.$$

The estimation of the parameters a and $\hat{b}$ is made by the method of „least squares", which one guaranties minimal sum of the squares of the differences between real y_i and calculated y_i values of the dependent variable, i.e.

$$\sum_{i=1}^{n} (y_i - y_i)^2 \to \min .$$

In the case of nonlinear regression with two variables (three parameters) for sequences of the measured data (x_i) and (y_i) it is assumed that the function has a form $y = f(x)$ or $y = f(x,a,b,c,...)$. Parameters of the functions are determined in a way that the squares of the differences of the observed points and ordinates of the curve are minimal. In order to get minimal value of the sum of the squares of the differences, it is needed all partial derivatives by all variables to be zero.

3. APPLICATIONS

The discussed techniques above, are going to be applied on hydrological sequences for average annual temperatures on the air for the measuring stations Prilep, Demir Kapija, Bitola, Stip and Skopje. For these measuring stations we have sequences for the period from 1925 to 2000. In these measuring stations there are interruptions in measuring in different time periods. For each sequence it is made statistical analysis by finding basic statistical parameters (see Table 1). The analysis on the homogeneity of the sequences of annual sums of the temperatures for the period $1925-2000$ for measuring stations Prilep, Bitola, Demir Kapija, Stip, using different tests for the hypothesis (normalized z-test, Student's t-test, Fisher's F - test) is given.

Table 1. Statistical parameters for hydrological sequences

Hydr. measure. /Station	**Prilep**	**D. Kapija**	**Bitola**	**Skopje**	**Stip**
Period	1925-2000	1925-2000	1925-2000	1925-2000	1925-2000
Missing data	1925-2000	1925-2000	1925-2000	1925-2000	1925-2000
Number of data- n	69	63	65	67	69
t_a	11,389	13,808	11,281	12,317	12,877

$t_{\min}$	10,125	12,558	10,333	10,825	11,200
$t_{\max}$	13,217	15,300	12,992	14,267	14,300
C_s	0,512	0,455	0,319	0,383	0,307
σ	0,721	0,676	0,565	0,648	0,677
C_v	0,063	0,457	0,319	0,053	0,053

3.1. Measuring station Prilep

The sequence of data for average air temperatures for city Prilep from 1925 till 2000 has 69 entries, and entries missing for 1925,1941,1943,1944,1946,1953 and 1954. According to Fisher's exact test, the sequence of annual sum temperatures for $1925-2000$ at the measuring station Prilep, can be accepted that the sequence is homogeneous with significance level of 5%, meaning that there is 5% risk of concluding that difference exists when there is no actual difference. To fill the gaps for this period, correlative connection with one variable and correlative connection with two variables with measuring stations Stip (X_1) and Demir Kapijia (X_2) are made.

In establishing correlative connections with two variables the order of including the variables is made accordingly the values of the coefficient of correlation between the dependent variable (Y) and the independent variable $(X)-(r_{xy})$. Since $r^2_{x_2y}=0,828>r^2_{x_1y}=0,822$, the first variable which will be included in the correlation is the independent variable (X_2), i.e. it will be used the data for Demir Kapija. The correlation equation with one independent variable (X_2), $y=a'+b\cdot X_2$ for the measuring station Prilep is the following $y=-0,82+0,88X_2$. The correlation equation with two variables (X_1) and (X_2), $y=a+b_1(X_1-\overline{X_1})+b_2(X_2-\overline{X_2})$ for the measuring station Prilep is the following $y=11,39+0,42(X_1-12,88)+0,5(X_2-13,81)$.

The regression equation with two variables (dependent variable Y (Prilep) and independent variable X (Demir Kapija) is given with linear and polynomial degree. The linear regression model between Prilep and Demir Kapija is given by $y=0,9003x-1,085$ and the polynomial regression with degree 2 is given by $y=0,0578x^2-0,7377x+10,307$. Since, these sequences are not continuous, the existing gaps are fulfilled only for the period where measuring data exists (1946,1953 and 1954).

For comparison, the obtained results are given in Table 2 and the graphical data are presented in annual average air temperatures for measuring station Prilep, where the missing data are fulfilled by linear and polynomial regression in Figure 1.

3.2. Measuring station Bitola

The sequence of data for average air temperatures for city Bitola from 1925 till 2000 has 65 entries, and entries missing for $1925-1927, 1934, 1939-1941,$ $1944,\ 1945, 1952$ and 1953. According to Fisher's exact test, the sequence of annual sum temperatures for $1925-2000$ at the measuring station Bitola, it can be accepted that the sequence is homogeneous with significance level of 5%.

To fill the gaps for this period, correlative connection with one variable (X_1) Demir Kapija and correlative connection with two variables with measuring stations Demir Kapija (X_1) and Stip (X_2) are made.

Since $r_{x_1y}^2 = 0,823 > r_{x_2y}^2 = 0,820$, the first variable which will be included in the correlation is the independent variable (X_1), i.e. it will be used the data for Demir Kapija. The correlation equation with one independent variable (X_1), $y = a' + b \cdot X_1$ for the measuring station Bitola is the following $y = 1,79 + 0,69X_1$. The correlation equation with two variables (X_1) and (X_2), $y = a + b_1(X_1 - \overline{X_1}) + b_2(X_2 - \overline{X_2})$ for the measuring station Bitola is the following $y = 11,28 + 0,37(X_1 - 13,81) + 0,34(X_2 - 12,88)$.

The regression equation with two variables (dependent variable Y (Bitola) and independent variable X (Stip) is given with linear and polynomial degree. The linear regression model between Bitola and Stip is given by $y = 0,8116x + 0,84$ and the polynomial regression with degree 2 is given by $y = 0,0199x^2 + 0,2946x + 4,1922$. Since, these sequences are not continuous, the existing gaps are fulfilled only for the period where measuring data exists ($1927, 1934, 1940, 1944, 1945, 1952$ and 1953). For comparison, the obtained results are given in Table 3 and the graphical data are presented in annual average air temperatures for measuring station Bitola, where the missing data are fulfilled by linear and polynomial regression in Figure 2.

3.3. Measuring station Demir Kapija

The sequence of data for average air temperatures for city Demir Kapija from 1925 till 2000 has 63 entries, and entries missing for $1925-1933, 1941,$ $1943-1945$. According to Fisher's exact test, the sequence of annual sum temperatures for $1925-2000$ at the measuring station Demir Kapija, it can be accepted that the sequence is homogeneous with significance level of 5%.

To fill the gaps for this period, correlative connection with one variable (X_1) Stip and correlative connection with two variables with measuring stations Stip (X_1) and Skopje (X_2) are made.

Since $r^2_{x_1y} = 0{,}917 > r^2_{x_2y} = 0{,}841$, the first variable which will be included in the correlation is the independent variable (X_1), i.e. it will be used the data for Stip. The correlation equation with one independent variable (X_1), $y = a' + b \cdot X_1$ for the measuring station Demir Kapija is the following $y = 2{,}01 + 0{,}92X_1$. The correlation equation with two variables (X_1) and (X_2), $y = a + b_1(X_1 - \overline{X_1}) + b_2(X_2 - \overline{X_2})$ for the measuring station Demir Kapija is the following $y = 13{,}81 + 0{,}77(X_1 - 12{,}88) + 0{,}17(X_2 - 12{,}32)$.

The regression equation with two variables (dependent variable Y (Demir Kapija) and independent variable X (Stip) is given with linear and polynomial degree. The linear regression model between Demir Kapija and Stip is given by $y = 0{,}9604x + 1{,}4446$ and the polynomial regression with degree 2 is given by $y = 0{,}0958x^2 - 1{,}5208x + 17{,}473$. Since, these sequences are not continuous, the existing gaps are fulfilled only for the period where measuring data exists ($1927 - 1929, 1931, 1933, 1943 - 1945$).

For comparison, the obtained results are given in Table 4 and the graphical data are presented in annual average air temperatures for measuring station Demir Kapija, where the missing data are fulfilled by linear and polynomial regression in Figure 3.

3.4. Measuring station Stip

The sequence of data for average air temperatures for city Stip from 1925 till 2000 has 70 entries, and entries missing for 1925,1926,1930,1932, 1936,1941 and 1945. According to Fisher's exact test, the sequence of annual sum temperatures for $1925 - 2000$ at the measuring station Stip, it can be accepted that the sequence is homogeneous with significance level of 5%.

To fill the gaps for this period, correlative connection with one variable (X_1) Demir Kapija and correlative connection with two variables with measuring stations Demir Kapija (X_1) and Skopje (X_2) are made.

Since $r^2_{x_1y} = 0{,}917 > r^2_{x_2y} = 0{,}873$, the first variable which will be included in the correlation is the independent variable (X_1), i.e. it will be used the data for Stip. The correlation equation with one independent variable (X_1), $y = a' + b \cdot X_1$ for the measuring station Stip is the following $y = 0{,}19 + 0{,}92X_1$. The correlation equation with two variables (X_1) and (X_2), $y = a + b_1(X_1 - \overline{X_1}) + b_2(X_2 - \overline{X_2})$ for the measuring station Stip is the following $y = 12{,}88 + 0{,}63(X_1 - 13{,}81) + 0{,}36(X_2 - 12{,}32)$.

The regression equation with two variables (dependent variable Y (Stip) and independent variable X (Demir Kapija) is given with linear and polynomial degree. The linear regression model between Stip and Demir Kapija is given by $y = 0,934x - 0,026$ and the polynomial regression with degree 2 is given by $y = -0,0134x^2 + 1,308x - 2,6298$. Since, these sequences are not continuous, the existing gaps are fulfilled only for the period where measuring data exists (1936 and 1947).

For comparison, the obtained results are given in Table 5 and the graphical data are presented in annual average air temperatures for measuring station Stip, where the missing data are fulfilled by linear and polynomial regression in Figure 4.

3.5. Measuring station Skopje

The sequence of data for average air temperatures for city Skopje from 1925 till 2000 has 67 entries, and entries missing for $1934, 1941$ and $1943 - 1949$. According to Fisher's exact test, the sequence of annual sum temperatures for $1925 - 2000$ at the measuring station Skopje, it can be accepted that the sequence is homogeneous with significance level of 5%.

To fill the gaps for this period, correlative connection with one variable (X_2) Demir Kapija and correlative connection with two variables with measuring stations Stip (X_1) and Demir Kapija (X_2) are made.

Since $r^2_{x_2y} = 0,873 > r^2_{x_1y} = 0,841$, the first variable which will be included in the correlation is the independent variable (X_2), i.e. it will be used the data for Demir Kapija. The correlation equation with one independent variable (X_2), $y = a' + b \cdot X_2$ for the measuring station Skopje is the following $y = 1,54 + 0,83X_2$. The correlation equation with two variables (X_1) and (X_2), $y = a + b_1(X_1 - \overline{X_1}) + b_2(X_2 - \overline{X_2})$ for the measuring station Skopje is the following $y = 12,32 + 0,24(X_1 - 13,81) + 0,62(X_2 - 12,88)$.

The regression equation with two variables (dependent variable Y (Skopje) and independent variable X (Stip) is given with linear and logarithmic function. The linear regression model between Skopje and Stip is given by $y = 0,8509x + 1,3671$ and the logarithmic regression is given by $y = 10,976\ln x - 15,708$. Since, these sequences are not continuous, the existing gaps are fulfilled only for the period where measuring data exists ($1934, 1943 - 1946, 1948$ and 1949).

For comparison, the obtained results are given in Table 6 and the graphical data are presented in annual average air temperatures for measuring station Skopje, where the missing data are fulfilled by linear and polynomial regression in Figure 5.

Table 2 – Comparison between obtained results for Prilep with correlation and regression

	Period	1946	1953	1954
Correlation	$y=-0,82+0,88X_2$			
	D. Kapija $=x_2$	15,0	13,5	13,5
	Prilep $=y$	12,4	11,1	11,1
	$y=11,39+0,42(X_1-12,88)+0,50(X_2-13,81)$			
	Stip $=x_1$	14,0	12,6	12,5
	D. Kapija $=x_2$	15,0	13,5	13,5
	Prilep $=y$	12,5	11,1	11,1
	Period	1946	1953	1954
Regression	$y=0,9003x-1,085$			
	D. Kapija $=x$	15,0	13,5	13,5
	Prilep $=y$	12,4	11,1	11,1
	$y=0,0587x^2-0,7377x+10,307$			
	D. Kapija $=x$	15,0	13,5	13,5
	Prilep $=y$	12,5	11,0	11,0

Table 3 – Comparison between obtained results for Bitola with correlation and regression

	Period	1927	1934	1939	1940	1944	1945	1952	1953
Correlation	$y=1,79+0,69X_2$								
	D.Kapija		15,2	14,6	12,7			15,3	13,5
	Bitola $=y$		12,2	11,8	10,5			12,3	11,1
	$y=11,28+0,37(X_1-13,81)+0,34(X_2-12,88)$								
	D.Kapija $=x_1$		15,2	14,6	12,7			15,3	13,5
	Stip $=x_2$		13,9	13,3	11,2			14,2	12,6
	Bitola $=y$		12,2	11,7	10,3			12,3	11,1
Regression	$y=0,8116x+0,84$								
	Stip $=x$	14,3	13,9	13,3	11,2	12,8	13,1	14,2	12,6
	Bitola $=y$	12,4	12,1	11,6	9,9	11,2	11,5	12,4	11,1
	$y=0,0199x^2+0,2946x+4,1922$								
	Stip $=x$	14,3	13,9	13,3	11,2	12,8	13,1	14,2	12,6
	Bitola $=y$	12,5	12,1	11,6	10,0	11,2	11,5	12,4	11,1

Table 4 – Comparison between obtained results for D. Kapija with correlation and regression

	Period	1927	1928	1929	1931	1933	1943	1944	1945
Correlation	$y = 2,01 + 0,92X_1$								
	Stip $= x_1$	14,3	13,6	12,8	12,6	11,8	13,9	12,8	13,1
	D.Kapija $= y$	15,1	14,5	13,7	13,6	12,8	14,7	13,7	14,0
	$y = 13,81 + 0,77(X_1 - 12,88) + 0,17(X_2 - 12,32)$								
	Stip $= x_1$	14,3	13,6	12,8	12,6	11,8			
	Skopje $= x_2$	13,6	12,9	11,6	12,3	11,8			
	D.Kapija $= y$	15,1	14,5	13,6	13,6	12,9			
Regression	$y = 0,9604x + 1,4446$								
	Stip $= x$	14,3	13,6	12,8	12,6	11,8	13,9	12,8	13,1
	D.Kapija $= y$	15,2	14,5	13,7	13,5	12,8	14,8	13,7	14,0
	$y = 0,0958x^2 - 1,5208x + 17,473$								
	Stip $= x$	14,3	13,6	12,8	12,6	11,8	13,9	12,8	13,1
	D.Kapija $= y$	15,3	14,5	13,7	13,5	12,9	14,8	13,7	14,0

Table 5 – Comparison between obtained results for Stip with correlation and regression

	Period	1936	1947
Correlation	$y = 0,19 + 0,92X_1$		
	D. Kapija $= x_1$	14,8	14,4
	Stip $= y$	13,8	13,4
	$y = 12,88 + 0,63(X_1 - 13,81) + 0,36(X_2 - 12,32)$		
	D. Kapija $= x_1$	14,8	
	Skopje $= x_2$	13,0	
	Stip $= y$	13,7	
	Period	1936	1947
Regression	$y = 0,934x - 0,026$		
	D. Kapija $= x$	14,8	14,4
	Stip $= y$	13,8	13,4
	$y = -0,0134x^2 + 1,308x - 2,6298$		
	D. Kapija $= x$	14,8	14,4
	Stip $= y$	13,8	13,4

Table 6 – Comparison between obtained results for Skopje with correlation and regression

	Period	1934	1943	1944	1945	1946	1948	1949
Correlation	$y = 1,54 + 0,83X_2$							
	Stip $= x_2$	13,9	13,9	12,8	13,1	14,0	12,6	12,6
	Skopje	13,2	13,2	12,3	12,5	13,3	12,1	12,1
	$y = 12,32 + 0,24(X_1 - 13,81) + 0,62(X_2 - 12,88)$							
	D.Kapija $= x_2$	15,2				15,0	13,8	13,6
	Stip $= x_1$	13,9				14,0	12,6	12,6
	Skopje $= y$	13,3				13,3	12,1	12,1
Regression	$y = 0,8509x + 1,3671$							
	Stip$= x$	13,9	13,9	12,8	13,1	14,0	12,6	12,6
	Skopje $= y$	13,2	13,2	12,3	12,5	13,3	12,1	12,1
	$y = 10,976 \ln x - 15,708$							
	Stip $= x$	13,9	13,9	12,8	13,1	14,0	12,6	12,6
	Skopje $= y$	13,2	13,2	12,3	12,5	13,3	12,1	12,1

Figure 1 – Average yearly air temperatures in Prilep with results obtained linear and nonlinear correlation and regression

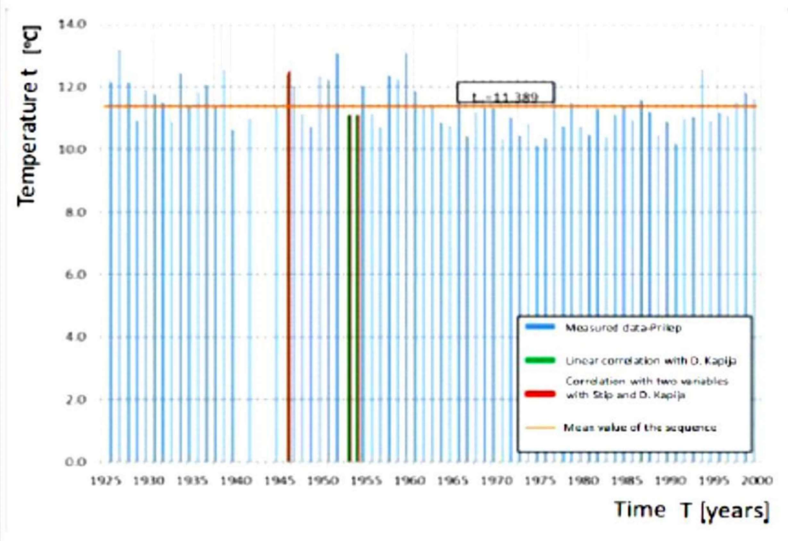


Figure 2 - Average yearly air temperatures in Bitola with results obtained linear and nonlinear correlation and regression

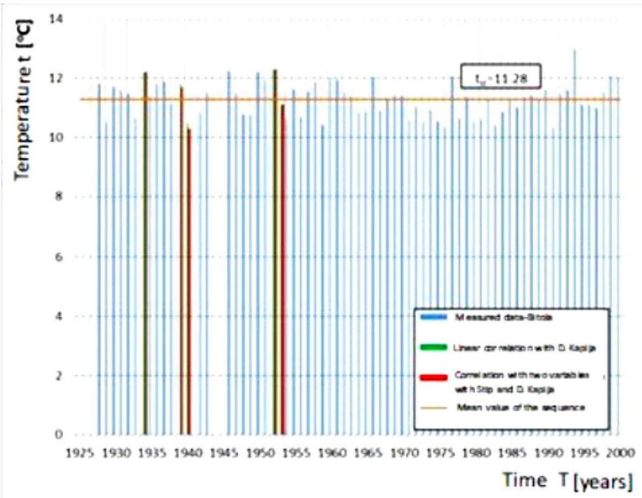


Figure 3 - Average yearly air temperatures in D. Kapija with results obtained linear and nonlinear correlation and regression

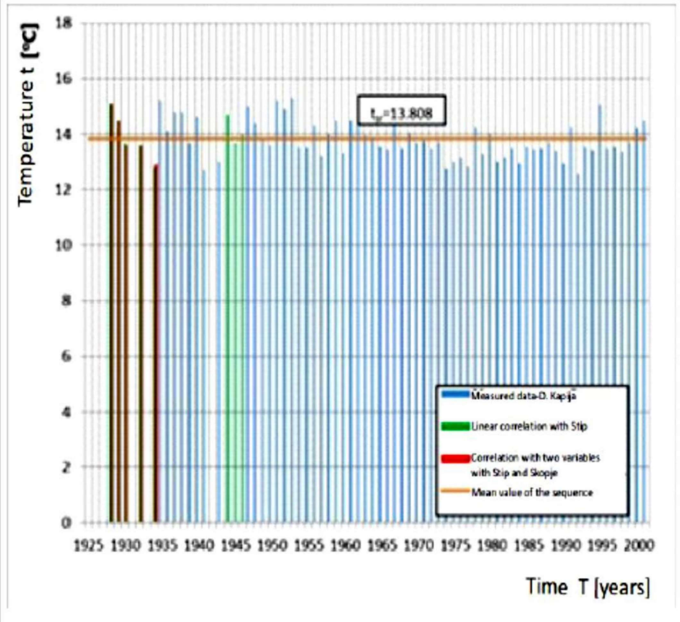


Figure 4 - Average yearly air temperatures in Stip with results obtained linear and nonlinear correlation and regression

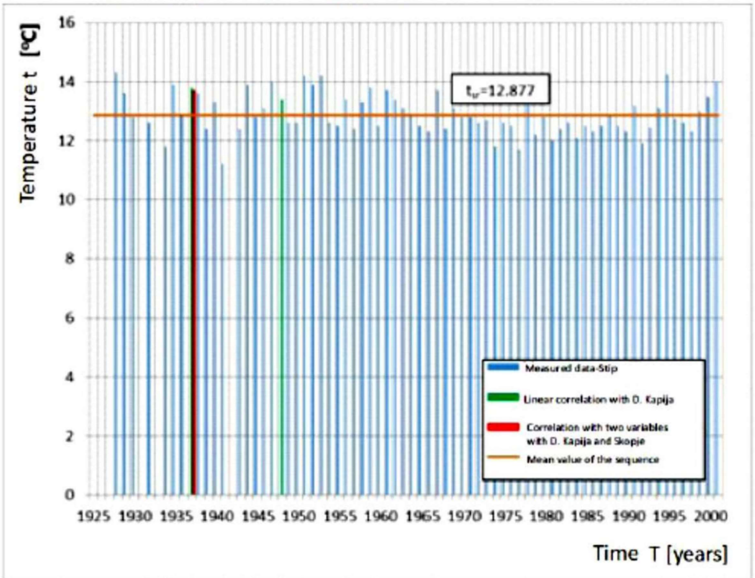
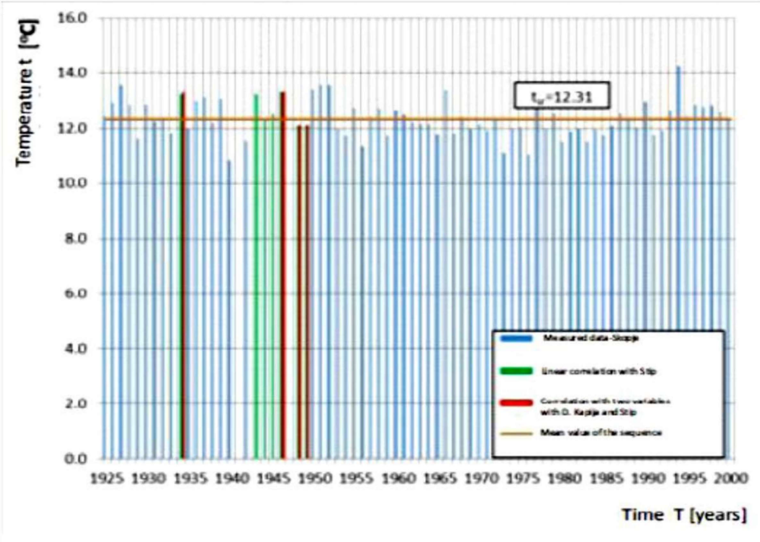


Figure 5 - Average yearly air temperatures in Skopje with results obtained linear and nonlinear correlation and regression



4. RESULT ANALYSIS

- ✓ Homogeneity analysis of the sequences of the annual temperature sums for the period $1925-2000$ for the measuring stations Prilep, Bitola, Demir Kapija and Stip, using different statistical tests (normalized z-test, Student's test and Fisher's test) is made. These parametric testes show that certain sequences, using certain test can be accepted, while using another test can not be accepted, with significance level of 5%. In hydrology, usually it is used the Fisher's test and his properties meets the needs and level of accuracy in this scientific field, so here only this test is presented. According to this test the sequences of the annual temperature sums for the period $1925-2000$ for the measuring stations Prilep, Bitola, Demir Kapija and Stip, can be accepted that these sequences are homogeneous with significance level of 5%.
- ✓ The choice of the measuring stations using to explore the correlation and regression connections is made according to smallest distance of the measuring stations.
- ✓ In establishing the correlative connection between two variables it is taken care of the values of the coefficient of correlation for dependent variable (Y) and independent variable $(X)-(r_{xy})$.
- ✓ Fulfilling the sequence of measured data for the measuring station Prilep is made through a correlation connection with one variable (using data from measuring station Demir Kapija) and correlative connection with two variables (using data from measuring stations Stip and Demir Kapija). The linear and polynomial regression between data sequences from Prilep and Demir Kapija is given. Obtained values by correlation with one variable and with two variables, as well as with regression, are very close (minimal difference in values). Also, we have very close values of the coefficients of regression in polynomial regression between sequences from Demir Kapija and Prilep ($r^2 = 0,848$) and linear regression between the same sequences ($r^2 = 0,847$).
- ✓ Fulfilling the sequence of measured data for the measuring station Bitola is made through a correlation connection with one variable (using data from measuring station Demir Kapija) and correlative connection with two variables (using data from measuring stations Demir Kapija and Stip). The obtained values with correlation with one or two variables, as well as the values with regression, are very close (there minimal differences in obtained values, except for 1940, where the temperature varies from $9,9°C$ (linear regression) to $10,5°C$ (correlation with one city). Also, we have very close values of the coefficients of regression in polynomial regression between sequences from Stip and Bitola ($r^2 = 0,878$) and linear regression between the same sequences ($r^2 = 0,877$).

- ✓ Fulfilling the sequence of measured data for the measuring station Demir Kapija is made through a correlation connection with one variable (using data from measuring station Stip) and correlative connection with two variables (using data from measuring stations Stip and Skopje). The linear and polynomial regression between the sequences data from the measuring station Demir Kapija and Stip is given. It can be noticed that obtained values with correlation with one or two variables, as well as with regression, are very close. We have very close values of the coefficients of regression in polynomial regression between sequences from Stip and Demir Kapija ($r^2 = 0,95$) and linear regression between the same sequences ($r^2 = 0,947$).
- ✓ Fulfilling the sequence of measured data for the measuring station Stip is made through a correlation connection with one variable (using data from measuring station Demir Kapija) and correlative connection with two variables (using data from measuring stations Demir Kapija and Skopje). The linear and polynomial regression between the sequences data from the measuring station Stip and Demir Kapija is given. It can be noticed that obtained values with correlation with one or two variables, as well as with regression, are very close. We have same values of the coefficients of regression in polynomial regression between sequences from Stip and Demir Kapija ($r^2 = 0,9471$) and linear regression between the same sequences ($r^2 = 0,9471$).
- ✓ Fulfilling the sequence of measured data for the measuring station Skopje is made through a correlation connection with one variable (using data from measuring station Stip) and correlative connection with two variables (using data from measuring stations Stip and Demir Kapija). The linear and logarithmic regression between the sequences data from the measuring station Skopje and Stip is given. It can be noticed that obtained values with correlation with one or two variables, as well as with regression, are very close. We have similar values of the coefficients of regression in linear regression between sequences from Stip and Skopje ($r^2 = 0,865$) and logarithmic regression between the same sequences ($r^2 = 0,864$).
- ✓ In establishing the correlative connection with one variable the biggest coefficient of correlation has the connection between the sequences of data from the measuring station Demir Kapija and measuring station Stip and vice versa ($r^2_{x_1 y} = 0,917$).
- ✓ In establishing the correlation connection with one variable the smallest coefficient of correlation has the connection between the sequences of data from the measuring station Bitola and measuring station Prilep and vice versa ($r^2_{x_1 y} = 0,691$).

- ✓ In establishing correlative connection with two variables, the biggest coefficient of correlation have the connections of Prilep with Stip and Demir Kapija ($r^2_{x_1x_2}=0,917$), Bitola with Stip and Demir Kapija ($r^2_{x_1x_2}=0,917$) and Skopje with Demir Kapija and Stip ($r^2_{x_1x_2}=0,917$).

Correlative connections with two variables, the least coefficient of correlation have the connections of Stip with Prilep and Bitola ($r^2_{x_1x_2}=0,691$), Skopje with Prilep and Bitola ($r^2_{x_1x_2}=0,691$) and Demir Kapija with Prilep and Bitola ($r^2_{x_1x_2}=0,691$).

5. CONCLUSIONS

In a hydrological point of view, the measured data as incoming parameters are very important for qualitative hydrological analysis. We must be careful with the type and quality of the sequence used into forming the relation. Also, it is very important to be careful about the type and strength of the established connection. Having coefficient of correlation less than $0,75$, we can state that connection is very weak or we state that there is no connection. When we have coefficient of correlation above $0,75$ of the correlation, we can state that the connection is good. Hence they can be used for fulfilling the sequences.

The application of the correlation and regression in hydrology is important for analysis of measured data examination of their applicability and continuation of certain sequences of data. The most used statistical techniques in hydrology are correlation and regression analysis. These techniques are complement each other, but there are significantly different. When using correlation, it does not matter whether certain phenomena is dependent or independent, the result is the same. Using correlation between three or more variables, one of the variables must be specified as dependent in advance and other must be specified as independent variables. The main goal of the correlation is to check and quantify whether exists consent between the variables (observed variables). In the regression analysis, we need to have information in advance about the (in)dependence of the variables. The main goal of the regression is to determine the type of the connection, i.e. dependence between observed phenomena.

The correlation and regression are methods for describing the mutual relation between two or more variables. Using these methods in any scientific field impose the following conditions: (1) Correlation and regression lines give linear and nonlinear connection, (2) Correlation and regression lines obtained by the least square method are susceptible on other influences. During the calculations always should be taken into consideration all possible parameters influencing the final result, (3) Correlation between two variables can be better understand if we take into consideration other variables. Lurking variables can make wrong correlation or regression.

References

[1] М. Георгиева, *Случајни грешки,* Математичка школа, Скопје, 1982.
[2] S. E. Edgell, S.M. Noon, *Effect of violation of normality on the* t*–test of the correlation coefficient,* Psychological Bulletin 95 (1984), 576-583.
[3] D. C. Gayou, *Effects of temperature on the mating call of Hyla versicolor,* Copeia (1984), 733-738
[4] Z. A. Ivkovic, *Matematicka statistika,* Naucna knjiga, Beograd, 1982.
[5] С. Јадровски, *Примена на статистички методи во хидрологијата,* Магистерска работа, Градежен факултет, Скопје, 2018.
[6] Р. Малчески, В. *Малческа, Теорија на веројатност,* Армаганка, Скопје, 2019.
[7] J. H. McDonald, K. W. Dunn, *Statistical tests for measures of colocalization in biological microscopy,* Journal of Microscopy 252 (2013), 295-302.
[8] B. H. McCardle, *Lines, models, and errors: Regression in the field,* Limnology and Oceanography 48 (2003), 1363-1366.
[9] D. S. Moore, *The Basic Practice of statistics,* W.H. Freedman and Company, New York, 2010.
[10] I. Pavlic, *Statisticka teorija i primjena,* Tehnicka knjiga, Zagreb, 1977
[11] A. Papoulis, Probability, *Random variables and Stochastic Processes,* International Student Edition, New York, 1965.
[12] Ц. Поповска, В. Ѓешовска, *Хидрологија-теорија со решени задачи,* Градежен факултет-Скопје, 2012.
[13] S. Prohaska, V. Ristić, *Hidrologija kroz teoriju i praksu,* Univerzitet u Beogradu, Rudarsko-geološki fakultet, Beograd, 1996.
[14] S. Prohaska, *Hidrologija I,* Univerzitet u Beogradu, Rudarsko-geološki fakultet, Beograd, 2006.
[15] S. Prohaska, *Hidrologija II,* Univerzitet u Beogradu, Rudarsko-geološki fakultet, Beograd, 2006.
[16] D. Srebrenović, *Primijenjena hidrologija,* Tehnička knjiga, Zagreb, 1986.
[17] *Стратегија за развој и унапредување на хидрометеоролошката служба во Македонија (2009-2011),* Министерство за земјоделство, шумарство и водостопанство на РМ.
[18] N. Kannan, J. P. Keating, R. L. Mason, *A comparison of classical and inverse estimators in the calibration problem,* Communications in Statistics: Theory and Methods 36 (2007), 83-95.
[19] R. G. Krutchkoff, *Classical and inverse regression methods of calibration,* Technometrics 9 (1967), 425-439.

[20] R. G. Krutchkoff, *Classical and inverse regression methods of calibration in extrapolation,* Technometrics 11 (1969), 605-608.
[21] Р. Ѓ. Кукутанов, *Метеорологија,* Универзитет „Гоце Делчев ", Штип, 2014.
[22] I. Ličina, D. Lisjak, *Primjena statističkih metoda s naglaskom na trend, korelaciju i regresiju*, Ekonomski faklutet u Osjeku, 2014.
[23] T. Lwin, J. S. Maritz, *An analysis of the linear-calibration controversy from the perspective of compound estimation,* Technometrics 24 (1982), 235-242.

Faculty of Civil Engineering, Ss. Cyril and Methodius University, Skopje, Macedonia
E-mail address: violetag@gf.ukim.edu.mk
Faculty of Civil Engineering, Ss. Cyril and Methodius University, Skopje, Macedonia
E-mail address: velinov.daniel@gmail.com
BALKAN ENERGY GROUP, Skopje, Macedonia
E-mail address: sjadrovski@hotmail.com

FLOOD FORECASTING USING ARTIFICIAL NEURAL NETWORKS

UDC: 004.89.032.26:519.248]:556.166.06(497.7)

Drenushe Fidani, Violeta Gjeshovska and Silvana Petrusheva

Abstract. Floods, as natural disasters, cause huge material damages and often result in loss of human lives. Their early prediction is necessary in order to take appropriate actions to reduce economic losses and risks for people. Albeit there is no universal method for flood modeling, significant advances in the technology of flood modeling techniques provide opportunities for flood prediction. Their modeling usually requires a large amount of data. In cases where only a specific part of the river basin is explored for more accurate modeling, the time and the effort to implement such complicated models is not justified. Therefore, the use of intelligent techniques such as Artificial Neural Networks (ANN) can be a practical alternative. The purpose of the investigations presented in this paper has been to make flood forecast for part of the Polog region using ANN. The forecast has been based on a model developed for this purpose. Modeling has been performed by use of four artificial neural networks in time series: Support Vector Machine (SVM), Radial Basis Function Neuron Network (RBFNN), General Regression Neural Network (GRNN) and Multilayer Perception (MP). Data on maximum annual flows of Vardar river, recorded at the Radusha measuring station throughout a period of 58 years, have been used as an input for the models and the output of the ANN is the maximum annual flow forecast for a 5-year (1951-2008) period. The results presented show that the ANN method, in this case the GRNN, can be useful and can provide sufficient accuracy in solving problems related to hydrological extremes.

1. INTRODUCTION

Water is vital for life, health and safety on Earth. There is no life without water. Since water resources are unevenly distributed in time and space, they need to be managed in order to avoid occurrence of floods and droughts. Floods are hydrological extremes known as natural hazards. Often, they can cause huge material damages including loss of human lives. Taking the dramatic climate changes into account, almost every country must be concerned with floods. As it is the case with many countries, floods are one of the most common natural phenomena in the Republic of North Macedonia, as well. The country is exposed to two types of floods: river - regional and flash- local sudden floods [1].

2010 *Mathematics Subject Classification.* Primary: 92B20, Secondary: 68T05, 62P30.

Key words and phrases. Artificial neural networks, maximum annual flow, time series, general regression neural network, floods.

Technically and financially, it is impossible to prevent all causes of flooding. However, early forecasting can provide good preparedness from the aspect of minimizing damages and reducing the consequences to nil.

Albeit there is no universal method for flood modeling, significant advances in the technology of flood modeling techniques provide opportunities for flood prediction. Their modeling usually requires a large amount of data. In cases where only a specific part of the river basin is explored for more accurate modeling, the time and the effort to implement such complicated models is not justified. Therefore, the use of intelligent techniques such as ANN can be a practical alternative.

Due to the above, the purpose of the investigations presented in this paper has been to predict the occurrence of hydrological extremes - floods based on the Maximum Annual Flow (MAF) of part of the Vardar river basin, the biggest river in North Macedonia. Actually, this has been done as part of doctoral thesis analyses performed for the Vardar river. Using four different ANNs, namely Support Vector Machine (SVM), Radial Basis Function Neuron Network (RBFNN), General Regression Neural Network (GRNN) and Multilayer Perception (MP) forecasting of the Vardar river flow has been done for four hydrological stations (Radusha, Gevgelija, Skopje, Jegunovce) for a 5 years' period of time. The forecasting has been done using the DTREG software for time series in combination with ANN. Presented in this paper is only the model of one station established by applying the General Regression Neural Network. For this model, data on flows of Vardar river measured at hydrological stations in the village of Radusha, Polog region, over a period of 58 years (1951–2008), have been used. MAF has been used as an alternative to detect changes in order to identify possible floods early enough for the purpose of responding appropriately.

The study has been based on recent scientific studies of the implementation of ANN in hydrology. The use of ANN models is becoming increasingly common in hydrological analyses and solving problems with water resources [2]. This is mainly because of the ability of ANN to model both linear and nonlinear systems without the need to make assumptions like those in implicit traditional statistical approaches [3].

The ANN technology is an alternative software approach inspired by studies of the brain and nervous system. Like the human brain, ANN behavior demonstrates the ability to learn and generalize from training data [2] and is a flexible structure capable of making nonlinear mapping between input and output layers [4]. ANN has already been successfully applied in some hydrological problems such as: rainfall forecasting, flood disaster prediction, modeling of engineering variables for water resources, river sediments and flow, prediction of river water quality, prediction of river flow, etc. Numerous studies show that ANN can offer a promising alternative to hydrological river flow forecasting and flood forecasting: Elsafi has used ANN to predict flood levels along the Nile river. According to him, this method is advantageous because only one variable can be used as a predictor, while other models require several variables to produce accurate predictions [5]. To predict floods in Indonesia, an ANN has been used by Sanubari et al. In this case, the Radial Basis Function neural network has been used. It is a network whose architecture consists of three layers, namely, an input layer, a hidden layer and an output. The analysis has been based on the water level from the 2015 rainfall data recorded at Dayeuhkolot, with results for the mean absolute percentage error-MAPE in the process of training and testing amounting to 4,97% and 29,1% for the rainfall and 0,047% and 1,05% for the water level [6]. In India, by selection of two different networks, namely, the feed forward network and the recurrent neural network, ANNs have been used to predict monthly river flows. The recurrent neural network has given better results and has therefore been recommended as a tool for predicting river flows [7].

Although the application of ANN is well developed, predicting time series events still remains the most challenging task for many engineers and scientists, "forcing engineers to constantly try to optimize existing solutions in order to obtain more accurate results" [8]. Hrnjica and Bonaci, in their paper on Vrana lake, located on the island of Cres in Croatia, present results from models for predicting from extending time series. In the paper, based on monthly measurements of the lake level during the last 38 years, ANN has been used to predict the levels for 6 and 12 months. Two types of ANN have been used: the Long-Short Term Memory (LSTM) recurrent neural network (RNN) and the Feed Forward Neural Network (FFNN). The investigations presented in this paper have confirmed the possibility and efficiency of ANN in forecasting hydrological phenomena [8], etc.

Annual maximum flow modeling is a key tool for early warning of flood hazards [5]. This has been proved by the following studies: In a study performed by Singo et al., MAF data from 8 stations involving hydrological data recorded in the course of 50 years were used to analyze the flood frequencies in the river basin. To rule out the likelihood of flooding, frequency distributions have been tested and have best described the past characteristics and magnitude of such floods [9]. Similar research has been done by Seyam and Othman. They have conducted a long-term analysis of variations in the annual river flow regime over a period of 50 years. The purpose of their analysis has been to identify long-term variations in the annual flow regime of the Selangor river, which is one of the major rivers in Malaysia, over a 50 -year period [10]. From the given examples of scientific research, it can be concluded that ANNs that use annual maximum flow can be a very useful tool that can be used with satisfactory accuracy for certain forecasts in solving water resources problems.

2. METHODOLOGY

Artificial Neural Network. An artificial neural network (ANN) is a flexible mathematical tool, inspired by the biological neural networks of human brain.

As in the human brain, in form of signals neurons receive external information, in the same way artificial neural networks receive external data or input. Consistently the neurons are arranged in a layer, with the output of one layer serving as the input to the next layer and possibly other layers. Different layers may perform different transformations on their inputs. From here, neural networks consist of input and output layers, and in most case a hidden layer [11].

A single layer neural network is called a Perceptron. It gives a single output as shown in Figure 1.

In the Figure 1, $x_0, x_1, x_2, x_3, ..., x_n$ represents various inputs, independent variables, to the network. Each of these inputs is multiplied by a connection weight or synapse. The weights are represented as $w_0, w_1, w_2, w_3, ..., w_n$. Weight shows the strength of a particular node. In the simplest case, these products are summed, fed to a transfer activation function to generate a result, and this result is sent as output [12].

Mathematically it can be written as

$$x_1 w_1 + x_2 w_2 + x_3 w_3 + x_4 w_4 + ... + x_n w_n = \sum_{i=1}^{n} x_i w_i .$$

Activation function which is applied $f\left(\sum_{i=1}^{n} x_i w_i\right)$.

Activation function decides whether a neuron should be activated or not by calculating the weighted sum. The motive is to introduce non-linearity into the output of a neuron. Neural Network is considered Universal Function Approximators which means they can learn and compute any function at all [12]. Due to this feature, they are used to identify complex nonlinear relationships between input and output data sets.

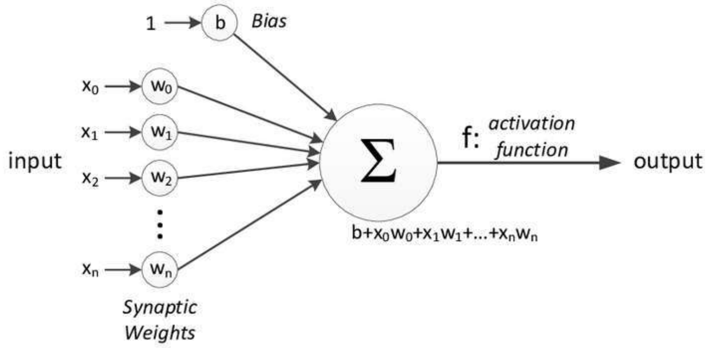


Figure 1. A single layer neural network: Perceptron (Source [12])

There are many types of neural network, each with their own specific architecture and levels of complexity. In Figure 2 is presented a typical multilayer artificial neural network showing the input layer for ten different inputs, the hidden layer, and the output layer having three outputs. Generally, the neurons in the input layer receive an input from the external environment and without any transformations upon the inputs they send their weighted values to the neurons in the hidden layer. The neurons of the hidden layer receive the transferred weighted inputs from the input, perform the needed transformations on it, and pass the output to the next hidden layer or the output layer. The output layer consists of neurons that receive the hidden layer output and send it to the user [11].

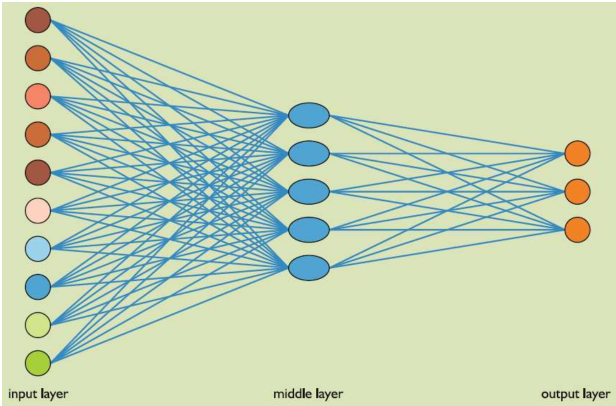


Figure 2. A typical multilayer artificial neural network with input layer, hidden layer and output layer (Source [11])

GRNN. GRNN networks have four layers input layer, pattern layer, summation layer and output layer. They do not require an iterative training procedure. Categorized by a layer that feeds back upon itself using adaptable weights, even with a constant input, they do not necessarily settle to a constant output. They can exhibit limit cycles and even chaotic behavior [13]. As schematically given in Figure 3, the input layer is connected to the pattern layer where ach neuron in the pattern layer represents a training pattern. The pattern layer performs a nonlinear transformation on the input data [14]. There are only two neurons in the summation layer. One neuron is the denominator summation unit, the other is the numerator summation unit. The denominator summation unit (in pink) figures out the weight values coming from each of the hidden neurons while the numerator summation unit (in green) figures out the weight values multiplied by the actual target value for each hidden neuron [11]. The neuron in the output layer divides the value accumulated in the numerator summation unit by the value in the denominator summation unit to yield the predicted result.

GRNN-Performance Measures: Some important performance measures for the NN model are: Mean Square Error (MSE), Normalized Mean Square Error (NMS) and correlation coefficient (r)

Mean Square Error can be determined by the following equation

$$MSE = \frac{\sum_{j=0}^{P}\sum_{i=0}^{N}(d_{ij}-y_{ij}{}^{2})}{N \cdot P},$$

where (P) is number of outputs, (N) is number of exemplars in the data set, (y_{ij}) is network output for exemplar (i) at processing elements (PE_j), (d_{ij}) is desired output for exemplar (i) at (PE_j).

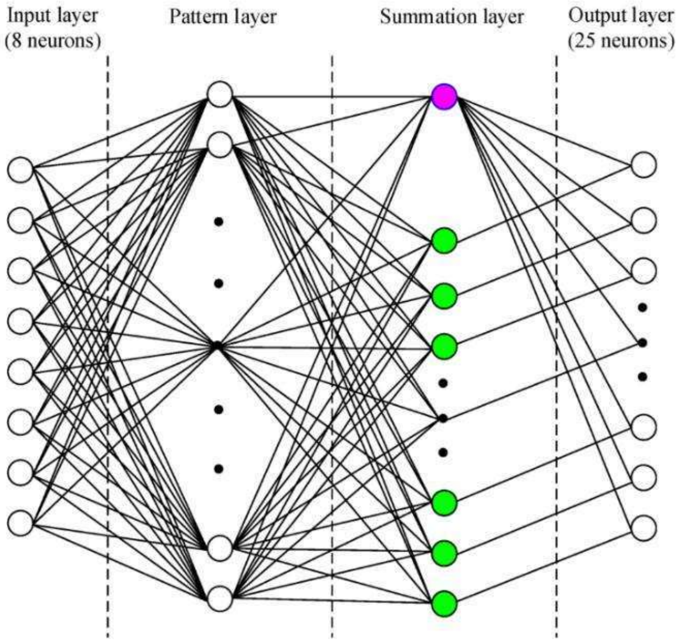


Figure 3. General regression neural network-GRNN architecture (Source [14])

Normalized Mean Square Error can be determined by the following equation

$$NMSE = \frac{P \cdot N \cdot MSE}{\sum_{j=0}^{P}\frac{N\sum_{i=0}^{N}d_{ij}^{2}-\left(\sum_{i=0}^{N}d_{ij}\right)^{2}}{N}},$$

where (P) is number of output processing elements, (N) is number of exemplars in the data set, (MSE) is mean square error, (d_{ij}) = desired output for exemplar (i) at processing element (j).

The equation for determining the correlation coefficient is

$$r = \frac{\frac{\sum_{i}(x_i-\bar{x})(d_i-\bar{d})}{N}}{\sqrt{\frac{\sum_{i}(d_i-\bar{d})^{2}}{N}} \cdot \sqrt{\frac{\sum_{i}(x_i-\bar{x})^{2}}{N}}},$$

where $\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$ and $\bar{d} = \frac{1}{N}\sum_{i=1}^{N} d_i$.

Time Series. Forecasting presents the transformation of information across time. The time series is a chronological sequence of excitations for a particular variable, usually, observed in regular intervals days, periods, months, years. In the time series with regular pattern, a value of the series is a function of the previous values. If (Y) is the value we are trying to model and predict, and (Y_t) is the value of (Y) at time t , then the goal is to create a model of the form

$$Y_t = f(Y_{t-1}, Y_{t-2}, Y_{t-3}, ..., Y_{t-n}) + et$$

where (Y_{t-1}) is the value of (Y) for the previous observation, (Y_{t-2}) is the value two observations ago, etc., and *et* represents noise that does not follow a predictable pattern [11]. Time series *forecasting* is the use of a model to predict future values based on previously observed values. The purpose for constructing a time series model is to create a model such that the error between the predicted value of the target variable and the real value is as small as possible.

The ANN approach does provide a viable and effective time series approach for developing input- output simulation and forecasting models. A proper design of the architecture of Artificial Neural Network (ANN) models can provide a robust tool in water resources modeling and forecasting.

3. STUDY AREA AND DATA SET

The Basin of River Vardar. The Vardar River is the largest river in the Republic of North Macedonia which provides 75% of the total water resources in the country [15]. Its spring is found in the Polog valley in the village of Vrutok, municipality of Gostivar. It passes through the cities of Gostivar and Tetovo, and then passes through Skopje. It flows through the central part of the country, enters Greece and finally reaches the Aegean Sea. The total annual average flow is estimated as $4289 \times 106\ m^3$ / year [16].

In this paper are presented the analyzes of the flow of the river Vardar measured at the measuring hydro station Radusa. For that purpose, the watershed upstream of this station is first analyzed, Figure 4. The section on the river Vardar is located in the watershed of Goren Vardar or in the lower part of the Polog valley. Dominant participation in the outflow and formation of flows in the river Vardar have surface waters coming from the northwestern massif, i.e. from the left tributaries of the river Vardar. Larger tributaries of the river Vardar in this part of the watershed are the rivers Pena, Vratnicka and Bistrica. The total area of the watershed upstream of the station in Raduse is about 1489 km^2. There are several categories of land use in the watershed, the so-called forests and semi-natural areas, artificial areas (urban areas, industrial, commercial and transport facilities, then mines, landfills and construction sites), agricultural areas, wetlands and swamps, water bodies, etc. In general, the studied region belongs to a modified type of Mediterranean climate which is a result of the influence of the continental climate from the central and eastern regions of Europe [17].

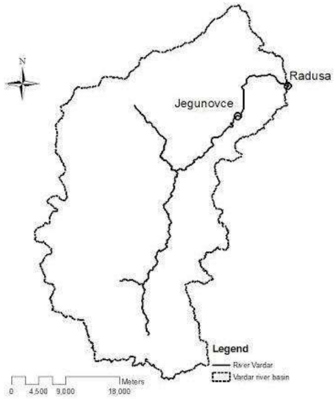


Figure 4. The watershed of the river Vardar upstream from Radusa

Data. The data used in the paper, are measured at the hydrological station Radusa, measuring station on the river Vardar for which there is a sufficiently long and quality series of data on the flow. The data were measured by the authorized institution for monitoring of hydrological stations – National Hydro meteorological service of the Republic of North Macedonia (УХМР). Data for maximum, minimum and average flow for the river Vardar in the village of Radusa for the period from 1951 to 2008 [18]. In this paper the analyzes are made with the maximum annual flows.

Using ANN, with the help of forecasting modeling software DTREG [19] a forecast of maximum annual flow for 5 years was obtained. The data were analyzed with several types of neural networks, namely Support Vector Machine (SVM), Radial Basis Function Neuron Network (RBFNN), General Regression Neural Network (GRNN) and Multilayer Perception (MP). To improve the accuracy of the model, the data can be normalized (the values of all data may be mapped at some intervals, usually [0,1]). The most accurate results were obtained with the General Regression Neural Network (GRNN). In this case, the data have been logarithmized, Figure 5.

AA1 fx Ln(max VQ)

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	year	1	Ln(1)	2	Ln(2)	3	Ln(3)	4	Ln(4)	5	Ln(5)	6	Ln(6)
2	1951	40.5	3.701302	31	3.433987	49	3.89182	47.4	3.858622	94	4.543295	68.3	4.22391
3	1952	45	3.806662	32.3	3.475067	78.4	4.361824	78.4	4.361824	28	3.332205	19.1	2.949688
4	1953	38.2	3.642836	62	4.127134	23	3.135494	74.7	4.31348	58.4	4.067316	72.9	4.289089
5	1954	9.6	2.261763	15.3	2.727853	102	4.624973	92	4.521789	100	4.60517	51.4	3.939638
6	1955	51.4	3.939638	70.1	4.249923	66.5	4.197202	51.4	3.939638	49.8	3.908015	33	3.496508
7	1956	31	3.433987	30.4	3.414443	59.3	4.082609	134	4.89784	72.9	4.289089	56.6	4.036009

Figure 5. Logarithmized data set

4. RESULTS

The data have been analyzed by use of several types of neural networks: Support Vector Machine (SVM), Radial Basis Function Neuron Network (RBFNN), General Regression Neural Network (GRNN) and Multilayer Perception (MP), using DTREG software [19]. The most accurate results have been obtained with the General Regression Neural Network (GRNN). An optimal solution has been obtained. The accuracy of the model, represented by the standard accuracy assessors, is: the Mean Absolute Percentage Error (MAPE) is $3,25\%$, while the Coefficient of Determination (R^2), which expresses the general convenience of the model, is $85,153\%$. The correlation coefficient is $0,962$, Table 1. The forecasting has been done for a 5-year period, as given in Figure 6.

Table 1. Validation Data

Validation Data	
CV -Coefficient of variation	0.032869
NMSE -Normalized mean square error	0.148471
Correlation between actual and predicted	0.962000
Maximum error	0.1682587
RMSE-Root Mean Squared Error	0.1396308
MSE- Mean Squared Error	0.0194968
R2- Coefficient of determination	0.85153 (85.153%)
MAE - Mean Absolute Error	0.1370405
MAPE- Mean Absolute Percentage Error	3.2503804

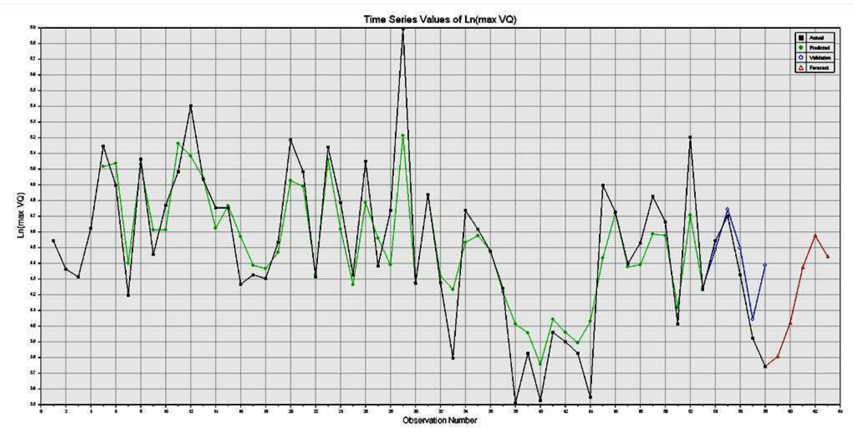


Figure 6. Time series of the maximum annual flow with a forecast for a 5-year period.

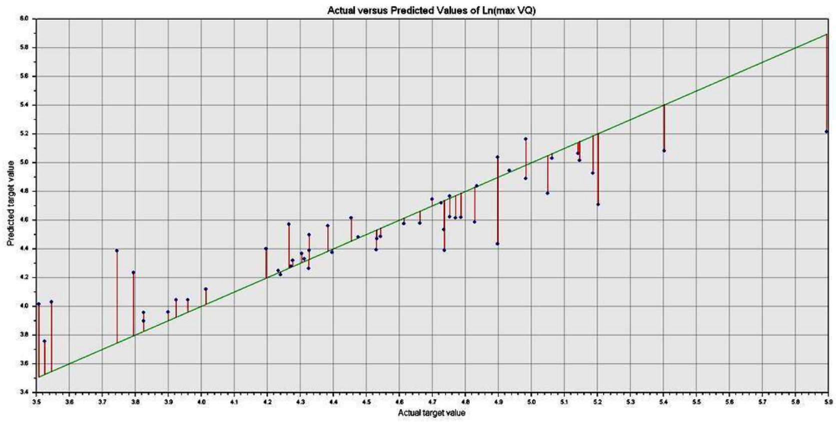


Figure 7. Actual target values vs Predicted target values

5. CONCLUSION

Artificial neural networks have shown a good ability to model a hydrological process. With artificial neural networks, using DTREG software for time series, MAF has been forecasted for 5 years. The results that have been obtained in this study, namely, the mean absolute percentage error is $3,25\%$, the coefficient of determination R^2, which expresses the general suitability of the model is $85,153\%$ and the correlation coefficient is $0,962$ prove that the General Regression Neural Network can be used to obtain a model with a satisfactory accuracy in forecasting maximum river flows that can be used as a basis for prediction of floods. For further research, it is recommended to develop a model of a greater accuracy using data on maximum monthly or daily flows.

References

[1] N. Dragovic, R. Ristic, H. Pulsel, B. Wolfschner, *Natural resources management in SEE: forests, soils and waters,* Deutsche Gesellschaft für Internationale, 2018.

[2] S. Riad, J. Mania, L. Bouchaou, Y. Najjar, *Rainfall-runoff model usingan artificial neural network approach,* Mathematical and Computer Modelling, 40(7-8), (2004), 839-846.

[3] I. Aichouri, A. Hani, N. Bougherira, L. Djabri, H. Chaffai, S. Lallahem, *River Flow Model Using Artificial Neural Networks,* Energy Procedia, 74, (2015), 1007-1014.

[4] M. Zakermoshfegh, M. Ghodsian, S.A.A. Salehi Neishabouri, M. Shakiba, *River Flow Forecasting Using Neural Networks and Auto-Calibrated NAM Model with Shuffled Complex Evolution,* Journal of Applied Sciences, 8: (2008), 1487-1494.

[5] S. Elsafi, *Artificial Neural Networks (ANNs) for flood forecasting at Dongola Station in the River Nile, Sudan,* Alexandria Engineering Journal, 53(3), (2014), 655-662.

[6] A. R. Sanubari, P. D. Kusuma, C. Setianingsih, *Flood Modelling and Prediction Using Artificial Neural Network,* IEEE International Conference on Internet of Things and Intelligence System (IOTAIS), (2018), 227-233.

[7] K. D. Nagesh, R. K. Srinivasa, T. Sathish, *River Flow Forecasting using Recurrent Neural Networks,* Water Resources Management, 18(2), (2004), 143-161.

[8] B. Hrnjica, O. Bonacci, *Lake Level Prediction using Feed Forward and Recurrent Neural Networks,* Water Resources Management, 33(7), (2019), 2471-2484

[9] L. Singo, P. Kundu, J. Odiyo, F. Mathivha, T. R. Nkuna, *Flood Frequency Analysis of Annual Maximum Stream Flows for Luvuvhu River Catchment,* Limpopo Province, South Africa, 2012.

[10] M. Seyam, F. Othman, *Long-term variation analysis of a tropical river's annual streamflow regime over a 50-year period,* Theoretical and Applied Climatology, 121(1-2), (2014), 71-85.

[11] D. Loucks, E. van Beek, *Water Resource Systems Planning And Management*, Gewerbestrasse , Springer International Publishing AG, Switzerland, 2017.

[12] N. Singh Chauhan, *Introduction to Artificial Neural Networks (ANN),* Medium, 2019.

[13] R. Deshmukh, A. Ghatol, *Short term flood forecasting using General Recurrent neural network modeling a comparative study,* International Journal of Computer Applications, 8(12), (2010), 5-9.

[14] L. Li, T. Xu, Y. Chen, *Improved Urban Flooding Mapping from Remote Sensing Images Using Generalized Regression Neural Network-Based Super-Resolution Algorithm,* Remote Sensing, 8(8), (2016), 625-636.

[15] Wfd-twinning.info. (2019). *Draft Vardar River Basin Management Plan (VRBMP),*2019. [online]

[16] S. Popov, T. Stafilov, R. Šajn, C. Tănăselia, K. Bačeva, Applying of Factor Analyses for Determination of Trace Elements Distribution in Water from

River Vardar and Its Tributaries, Macedonia/Greece, The Scientific World Journal, (2014), 1-11

[17] *Technical documentation on Vardar River Regulation in Polog region, Book 1-Hydrology,* Faculty of Civil Engineering, Skopje, 2016.

[18] *The Study on Integrated Water Resources Development and Management Master Plan in the Republic of Macedonia- Final Report,* NIPPON KOEI CO., LTD. KRI INTERNATIONAL CORPORATION, Japan International Cooperation Agency, 1999.

[19] P. Sherrod, *DTREG Predictive Modeling Software and Tutorial,* 2014, (www.dtreg.com)

Department of Architecture, International Vision University, Gostivar, Macedonia
E-mail address: drenushefidani@gmail.com
Faculty of Civil Engineering, Ss. Cyril and Methodius University, Skopje, Macedonia
E-mail address: violetag@gf.ukim.edu.mk
Faculty of Civil Engineering, Ss. Cyril and Methodius University, Skopje, Macedonia
E-mail address: silvana@gf.ukim.edu.mk

SEVERAL RELATIONS BETWEEN THE ROOTS OF POLYNOMIALS, THE ROOTS OF THEIR DERIVATIVES AND THE FOCI OF IN-ELLIPSES

UDC: 512.622:514.122

Sava Grozdev[1], Veselin Nenkov[2]

Abstract. It is a well known fact, that if the roots of a polynomial of third degree are in the vertices of a triangle, then the roots of its derivative are in the foci of the in-ellipse which is tangent to the sides of the triangle at their midpoints. This assertion motivates looking for answers to the following two basic questions: 1) Is there another type of polynomials that have roots in the vertices of a triangle, while their derivatives have roots in the foci of other in-ellipses of the triangle? 2) Is it possible to discover similar relations between the roots of polynomials in the vertices of polygons, the roots of their derivatives and the foci of in-ellipses of these polygons? We propose a short survey of some results which give answers to the above questions. We consider relations between polynomials with roots in the vertices of quadrilaterals or pentagons and the corresponding derivatives with roots in the foci of in-ellipses. It is noted a special kind of polynomials (medi-tangential) for which a relation exists between the roots of such a polynomial, the corresponding derivative and the foci of all in-ellipses. Thus, a generalization of the theorem for the triangles is obtained.

A remarkable relation between the roots of some polynomials of third degree, the roots of the corresponding derivatives and the foci of special ellipses is presented by the following theorem.

Theorem 1. *If the roots of a polynomial* $P(z)$ *of third degree are in the vertices of a triangle* Δ*, then the roots of its derivative* $P'(z)$ *are in the foci of the in-ellipse which is tangent to the sides of* Δ *at their mid-points.*

A proof of theorem 1 is presented in [1]. If $P(z)=P(x)$ is a polynomial of third degree, of a real variable and with real coefficients, the result of theorem 1 has a geometric interpretation according to which the roots of $P(x)$ are in the vertices of an isosceles triangle (Fig. 1).

Besides the ellipse that is mentioned in theorem 1, infinitely many ellipses could be inscribed in each triangle. For this reason a question arises about the possibility that some of the ellipses also realize geometric relations between polynomials and their corresponding derivatives. It turns out indeed, that some of these ellipses realize geometric relations between some kinds of polynomials with multiple roots in the vertices of a given triangle and the roots of their derivatives.

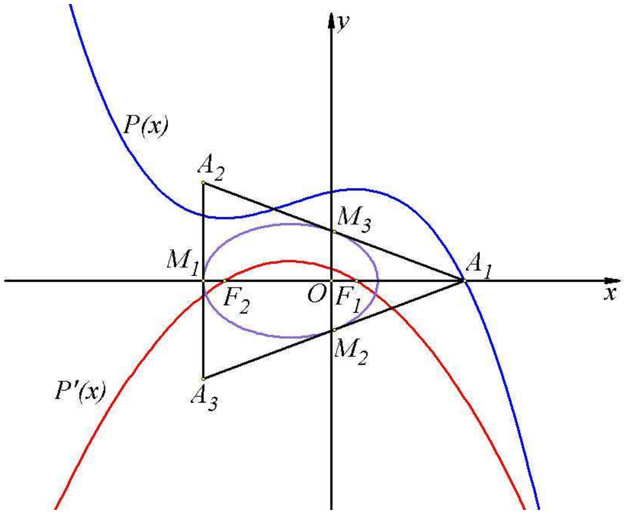


Fig. 1.

Theorem 2. *If a polynomial* $P(z)$ *of degree* $n = k_1 + k_2 + k_3$ *has a* k_j*-multiple root in the vertex* A_j $(j = 1,2,3)$ *of* $\Delta A_1 A_2 A_3$*, then the derivative* $P'(z)$ *of* $P(z)$ *has roots in the foci of the in-ellipse* k*, which is tangent to the lines* A_2A_3*,* A_3A_1 *and* A_1A_2 *at the points* P_1*,* P_2 *and* P_3*, respectively, verifying the equations*

$$\overline{A_2P_1} : \overline{A_3P_1} = -k_2 : k_3 ,\ \overline{A_3P_2} : \overline{A_1P_2} = -k_3 : k_1 ,\ \overline{A_1P_3} : \overline{A_2P_3} = -k_1 : k_2 .$$

A proof of theorem 2 is presented in [2]. The polynomial $P'(z)$ has a $(k_j - 1)$ - multiple root in the vertex A_j $(j = 1,2,3)$ [3], while the other two roots are described by theorem 2. In such a way a full geometric picture of the roots of $P'(z)$ is obtained.

The next consequence follows from theorem 2 directly:

Corollary 1. *If a polynomial* $P(z)$ *of degree* $n = 3.m$ *has a* m *- multiple root in the vertex* A_j $(j = 1,2,3)$ *of* $\Delta A_1 A_2 A_3$*, then the derivative* $P'(z)$ *of* $P(z)$ *has roots in the foci of the in-ellipse* k*, which is tangent to the segments* A_2A_3*,* A_3A_1 *and* A_1A_2 *at their mid-points.*

It is obvious, that theorem 1 follows from corollary 1 when $m = 1$.

If $P(z) = P(x)$ is a polynomial of a real variable and with real coefficients, then some geometric interpretations of theorem 2 and its consequence could be presented. In fig. 2 the polynomial $P(x)$ has a double root in A_1 and two simple complex conjugated roots in A_2 and A_3. The derivative $P'(x)$ has three simple real roots in the vertex A_1 and the foci F_1 and F_2 of the ellipse k. In fig. 3 the polynomial $P(x)$ has a simple real root in A_1 and two double complex conjugated roots in A_2 and A_3. The derivative $P'(x)$ has two simple real roots in the foci F_1 and F_2 of the ellipse k. The polynomials, which correspond to the corollary are presented in fig. 4 and 5. These polynomials have double and 3-multiple roots in the vertices of $\Delta A_1A_2A_3$, correspondingly. They are obtained when $m=2$ and $m=3$.

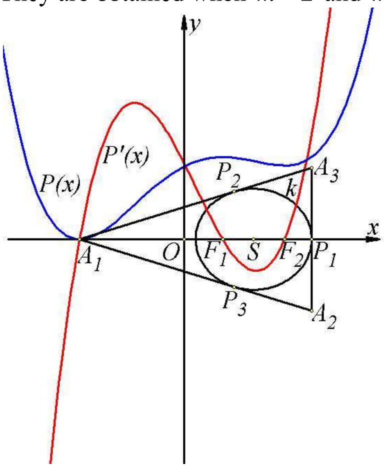
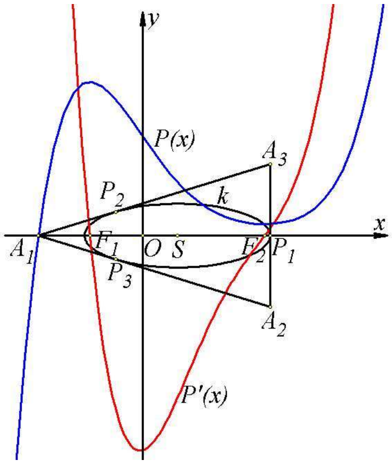


Fig. 2



6Fig. 3

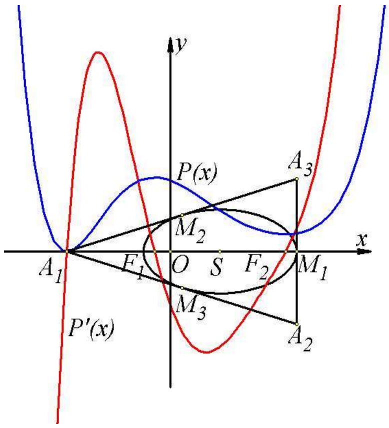


Fig. 4

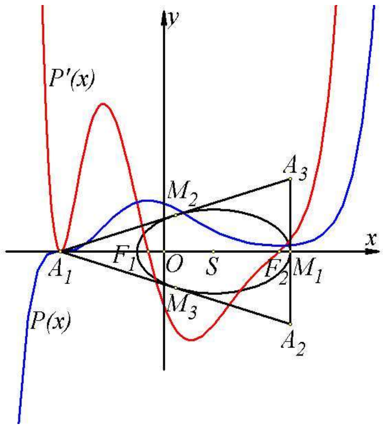


Fig. 5

It seems, that among the other geometric figures the parallelogram is mostly close to the triangle. There is an ellipse in the parallelogram too, which is tangent at the mid-points of its sides. Thus, the following proposal appears in a natural way;

Theorem 3. *If the roots of a polynomial $P(z)$ of forth degree are in the vertices of a parallelogram P, then the roots of the derivative $P'(z)$ of $P(z)$ are in the foci and the center of the ellipse, which is tangent to the sides of P at their mid-points.*

A proof of theorem 3 is presented in [4]. If $P(z)=P(x)$ is a polynomial of forth degree of a real variable and with real coefficients, then a geometric interpretation of theorem 3 exists. It could be applied if the polynomial under consideration is one of the following two kinds:

1) $P(x)$ has two real and two complex conjugated roots. This means that the parallelogram is a rhombus;

2) $P(x)$ has two pairs complex conjugated roots. This means that the parallelogram is a rectangle.

In case 1) the graphical representation of $P(x)$, as an image of a function of a real variable, passes through two opposite vertices of the rhombus only, which corresponds to the two real roots of $P(x)$. The graphical representation of $P'(x)$ is a cubic parabola whose center of symmetry is in the intersection point S of the diagonals of $A_1A_2A_3A_4$, which is always on the abscise axis.

The ellipse κ, which is tangent to the sides of $A_1A_2A_3A_4$ at their mid-points, has foci on the abscise axis or on the diagonal of $A_1A_2A_3A_4$, parallel to the ordinate axis. In the first case the cubic parabola $P'(x)$ intersects the abscise axis in the center S and the foci F_1 and F_2 of κ. In this case the roots of $P'(x)$ are real and coincide with the real points S, F_1 and F_2 (Fig. 6). In the second case the cubic parabola $P'(x)$ intersects the abscise axis in the point S only and does not pass through the points F_1 and F_2. In this case the points F_1 and F_2 correspond to non-real roots of $P'(x)$. However, always the point S corresponds to a real root and for this reason the graphical representation of $P'(x)$ passes through it always (Fig. 7).

In case 2) the graphical representation of $P(x)$, as an image of a function of real variable, does not pass through any vertex of the rectangle $A_1A_2A_3A_4$ in the general case. The graphical representation of $P'(x)$ is a cubic parabola with center of symmetry in the intercept point S of the diagonals of $A_1A_2A_3A_4$, which is on the abscise axis always.

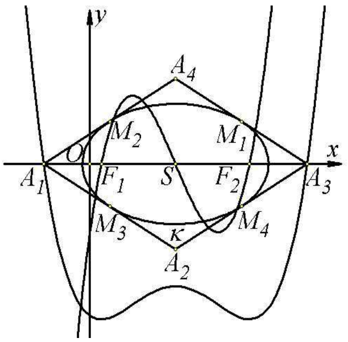


Fig.

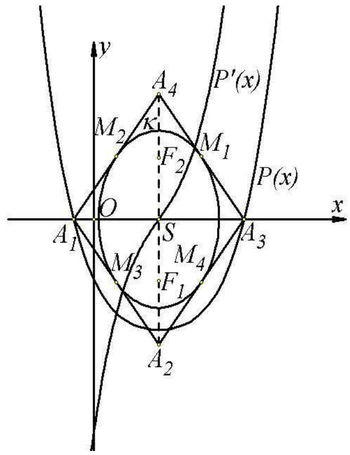


Fig. 7

The ellipse κ, which is tangent to the sides of $A_1A_2A_3A_4$ at their mid-points, has foci on the abscise axis or on a line, parallel to the ordinate axis.

In the first case the cubic parabola $P'(x)$ intersects the abscise axis in the center S and the foci F_1 and F_2 of κ. In this case the roots of $P'(x)$ are real and coincide with the real points S, F_1 and F_2 of κ (Fig. 8).

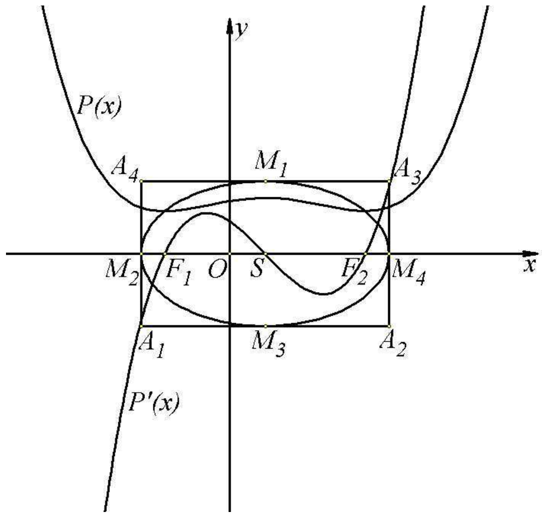


Fig. 8

In the second case the cubic parabola $P'(x)$ intersects the abscise axis in the point S only and does not pass through the points F_1 and F_2. In this case the points F_1 and F_2 correspond to non-real roots of $P'(x)$. However, always the point S corresponds to a real root and for this reason the graphical representation of $P'(x)$ passes through it always (Fig. 9).

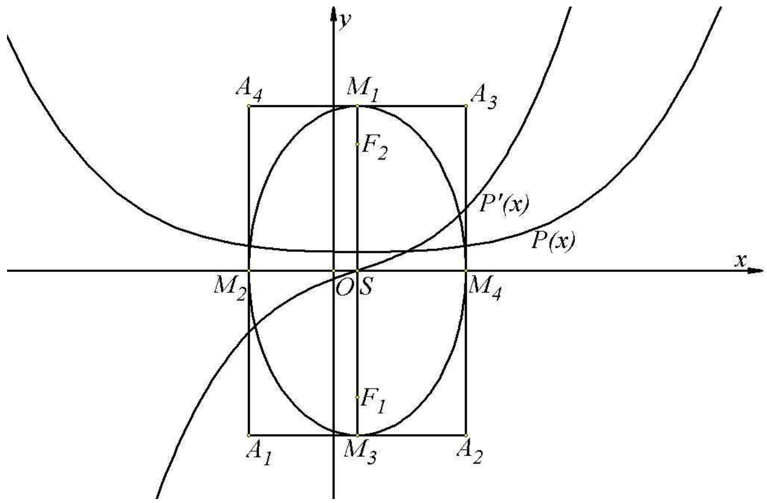


Fig. 9

Besides the ellipse that is mentioned in theorem 3, infinitely many ellipses could be inscribed in each parallelogram. A question arises about the possibility that some of the ellipses also realize geometric relations between polynomials and their corresponding derivatives. It turns out indeed, that some of these ellipses realize geometric relations between some kinds of polynomials with multiple roots in the vertices of a given parallelogram and the roots of their derivatives.

Theorem 4. *If a polynomial* $P(z)$ *of degree* $n = 2.(k_1 + k_2)$ *has* k_1*-multiple roots in the vertices* A_1 *and* A_3 *, also* k_2*-multiple roots in the vertices* A_2 *and* A_4 *of the parallelogram* $A_1A_2A_3A_4$*, then the derivative* $P'(z)$ *of* $P(z)$ *has roots in the foci and the center of the ellipse* k*, which is tangent to the lines* A_3A_4*,* A_4A_1*,* A_1A_2 *and* A_2A_3 *at the points* P_1*,* P_2*,* P_3 *and* P_4*,respectively, verifying the equalities*

$$\overline{A_3P_1} : \overline{A_4P_1} = -k_1 : k_2 , \ \overline{A_4P_2} : \overline{A_1P_2} = -k_2 : k_1 ,$$

$$\overline{A_1P_3} : \overline{A_2P_3} = -k_1 : k_2 , \ \overline{A_2P_4} : \overline{A_3P_4} = -k_2 : k_1 .$$

A proof of theorem 4 is presented in [5].

Corollary 2. *If a polynomial* $P(z)$ *of degree* $n=4.m$ *has a* m*-multiple root in the vertex* A_j $(j=1,2,3,4)$ *of the parallelogram* $A_1A_2A_3A_4$*, then the derivative* $P'(z)$ *of* $P(z)$ *has roots in the foci and the center of the ellipse* k*, which is tangent to the segments* A_3A_4*,* A_4A_1*,* A_1A_2 *and* A_2A_3 *at their midpoints.*

It is clear, that when $m=1$ corollary 2 implies theorem 1. If $P(z)=P(x)$ is a polynomial with real coefficients of a real variable, then some geometric interpretations of the corollary could be presented. Fig. 10 and Fig. 11 present polynomials $P(x)$ with double roots in the vertices of rhombus $A_1A_2A_3A_4$. In the case of Fig. 10 the derivative $P'(x)$ has real roots in the foci F_1, F_2 and the center O of the ellipse k, which is tangent at the mid-points of the sides A_3A_4, A_4A_1, A_1A_2 and A_2A_3. For this reason the graphical representation of $P'(x)$ passes through the points F_1, F_2 and the center of k. In the case of Fig. 11 the derivative $P'(x)$ has no real root in the foci F_1 and F_2 of the ellipse k. For this reason the graphical representation of $P'(x)$ passes through the center of k only.

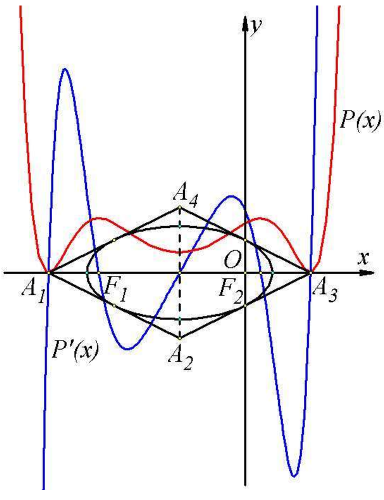


Fig. 10

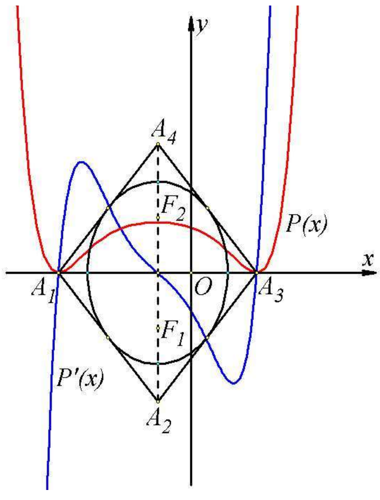


Fig. 11

Fig. 12 and Fig. 13 present polynomials $P(x)$ with 3-multiple roots in the vertices of the rectangle $A_1A_2A_3A_4$. In the case of Fig. 12, the derivative $P'(x)$ has real roots in the foci F_1, F_2 and the center of the ellipse k, which is tangent at the mid-points of the sides A_3A_4, A_4A_1, A_1A_2 and A_2A_3. For this reason the graphical representation of $P'(x)$ passes through the points F_1, F_2 and the center of k. In the case of Fig. 13, the derivative $P'(x)$ has no real root in the foci F_1 and F_2 of the ellipse k. For this reason the graphical representation of $P'(x)$ passes through the center of k only.

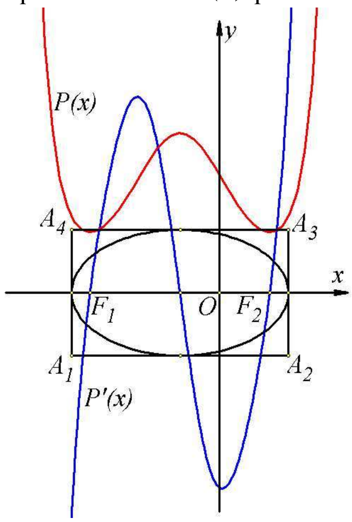


Fig. 12

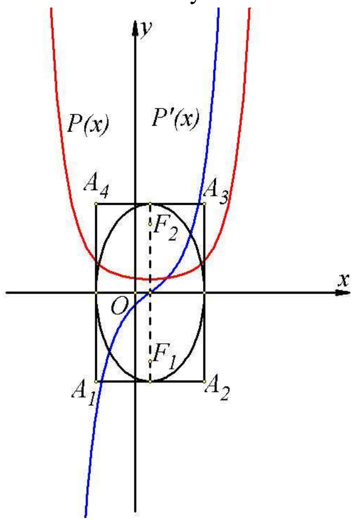


Fig. 13

The geometric relation presented in theorem 4, contains the foci of a suitable in-ellipse of a parallelogram. On the other hand infinitely many ellipses could be inscribed in each convex quadrilateral. Thus, a question arises for the existence of a relation between some of those ellipses and the derivatives of the polynomials with roots in the vertices of a given quadrilateral only. Such a relation is presented in the following assertion.

Theorem 5. *Let* k_j $(j=1,2,3,4)$ *be positive integers and* k *be an in-ellipse of the convex quadrilateral* $A_1A_2A_3A_4$, *such that the tangent points* P_1, P_2, P_3 *and* P_4 *of* k *with the segments* A_1A_2, A_2A_3, A_3A_4 *and* A_4A_1, *respectively, verify the equalities:*

$$\overline{A_1P_1}:\overline{A_2P_1}=-k_1:k_2,\ \overline{A_2P_2}:\overline{A_3P_2}=-k_2:k_3,$$
$$\overline{A_3P_3}:\overline{A_4P_3}=-k_3:k_4,\ \overline{A_4P_4}:\overline{A_1P_4}=-k_4:k_1.$$

If a polynomial $P(z)$ *of degree* $n=k_1+k_2+k_3+k_4$ *and complex variable, with complex coefficients has a* k_j *multiple root in the vertices* A_j $(j=1,2,3,4)$ *of* $A_1A_2A_3A_4$, *then the derivative of* $P(z)$ *has roots in the point* $P_0=A_1A_3\cap A_2A_4$ *and in the foci of the ellipse* k.

A proof of theorem 5 is presented in [6]. The polynomial $P'(z)$ has a $k_j - 1$ multiple root in the vertex A_j $(j = 1,2,3,4)$ [3], while the other three roots are described by theorem 5. In such a way we obtain a full geometric picture of the roots of $P'(z)$. Of course, theorem 5 is a generalization of theorem 4.

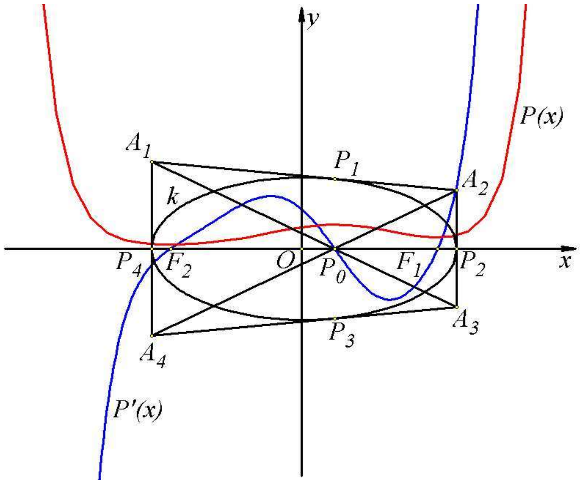


Fig. 14

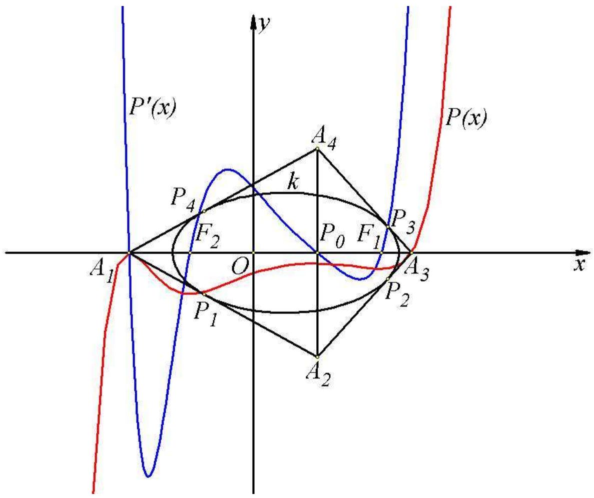


Fig. 15

If $P(z)=P(x)$ is a polynomial with real coefficients, of real variable, we could present some geometric interpretations of the proved theorem. Fig. 14 presents a polynomial $P(x)$ with roots in the vertices of a trapeze, which is symmetric with respect to the abscise axis. It has 3-multiple complex conjugated roots in the points A_1, A_4 and double complex conjugated roots in the points A_2 and A_3. Fig. 15 presents a polynomial $P(x)$ with roots in the vertices of a deltoid, which is symmetric with respect to the abscise axis. It has a double real root in the point A_1, a simple real root in the point A_3 and a 3-multiple complex conjugated roots in the points A_2 and A_4.

The geometric relation described in theorem 5, contains the foci of a suitable in-ellipse of the quadrilateral under consideration (if such an ellipse exists). On the other hand, each convex pentagon $A_1A_2A_3A_4A_5$ contains exactly two ellipses – the one is inscribed in $A_1A_2A_3A_4A_5$, while the other one is inscribed in the star-like pentagon $A_1A_3A_5A_2A_4$. In such a way a question arises for the existence of a relation between some kinds of convex pentagons and their in-ellipses in comparison with the derivatives of the polynomials with roots in the vertices of the pentagons under consideration only. It turns out that such a relation exists, which is similar to the one for quadrilaterals. It is expressed by the following assertion.

Theorem 6. *Let* k_j $(j=1,2,3,4,5)$ *be positive integers and* k' *be an in-ellipse of the convex pentagon* $A_1A_2A_3A_4A_5$, *such that the tangent points* P_1, P_2, P_3, P_4 *and* P_5 *of* k' *with the segments* A_1A_2, A_2A_3, A_3A_4, A_4A_5 *and* A_5A_1, *respectively, verify the equalities:*

$$\overline{A_1P_1}:\overline{A_2P_1}=-k_1:k_2,\ \overline{A_2P_2}:\overline{A_3P_2}=-k_2:k_3,\ \overline{A_3P_3}:\overline{A_4P_3}=-k_3:k_4,$$
$$\overline{A_4P_4}:\overline{A_5P_4}=-k_4:k_5\ \ \overline{A_5P_5}:\overline{A_1P_5}=-k_5:k_1.$$

If the polynomial $P(z)$ *of degree* $n=k_1+k_2+k_3+k_4+k_5$ *has a* k_j *multiple root in the vertices* A_j $(j=1,2,3,4,5)$ *of* $A_1A_2A_3A_4A_5$, *then the derivative of* $P(z)$ *has roots in the foci of the ellipse* k' *and the foci of the ellipse* k'', *which is inscribed in the star-like pentagon* $A_1A_3A_5A_2A_4$.

A proof of theorem 6 is presented in [7]. The polynomial $P'(z)$ has a k_j-1 multiple root in the vertex A_j $(j=1,2,3,4,5)$ [3], while the other four roots are described by theorem 6. In such a way we obtain a full geometric picture of the roots of $P'(z)$.

If $P(z)=P(x)$ is a polynomial with real coefficients of a real variable we could present some geometric interpretations of theorem 6. Fig. 16 presents a polynomial $P(x)$ with roots in the vertices of a pentagon, which is symmetric with respect to the abscise axis. It has double complex conjugated roots in the points A_1 and A_5, 3-multiple complex conjugated roots in the points A_2 and A_4 and a real root in the point A_3.

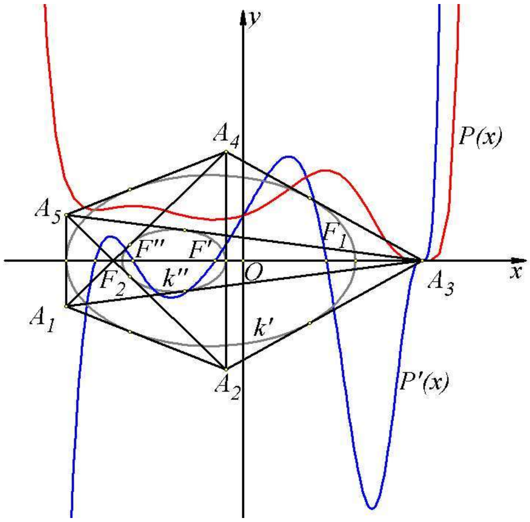


Fig. 16

Theorems 1 and 2 formulated before concern polynomials of third and fourth degree, located in the vertices of a triangle and a parallelogram, respectively. They could be generalized considering a special class of polygons, which includes the triangle and the parallelogram. This class consists of polygons which are simultaneously inscribed and exscribed for concentric homothetic ellipses. It is determined in the following way: Let $A_1A_2\ldots A_n$ $(n \geq 3)$ be an arbitrary convex polygon for which B_1, B_2, …, B_{n-1}, B_n are the mid-points of the sides A_1A_2, A_2A_3, …, $A_{n-1}A_n$, A_nA_1, respectively. If $A_1A_2\ldots A_n$ has an in-ellipse k, such that the tangent points of k with the sides of $A_1A_2\ldots A_n$ are B_j $(j = 1,2,\ldots,n)$, we will call $A_1A_2\ldots A_n$ to be *meditangential polygon of* k, while we will call k to be *meditangential ellipse for* $A_1A_2\ldots A_n$. The following theorem is true for the *meditangential polygons.*

Theorem 7. *If a polynomial* $P(z)$ *of degree* n *has roots in the vertices of a meditangential n-polygon* $A_1A_2\ldots A_n$, *then the roots of the derivative* $P'(z)$ *of* $P(z)$ *are in the foci of all meditangential ellipses, generated by the vertices of* $A_1A_2\ldots A_n$, *and in the center of* $A_1A_2\ldots A_n$, *when* n *is even.*

A proof of theorem 7 is presented in [8]. This theorem implies directly the following:

Corollary 3. *If the roots of the polynomial* $P(z)$ *are in the vertices of a meditangential polygon, then the roots of its derivative* $P'(z)$ *are located on a straight line.*

Corollary 4. *If a polynomial* $P(z)$ *of degree* $n.p$ *has a* p *multiple root in the vertex* A_j $(j=1,2,\ldots,n)$ *of a meditangential polygon* $A_1A_2\ldots A_n$*, then the derivative* $P'(z)$ *of* $P(z)$ *has roots in the foci of all meditangential ellipses, generated by the vertices of* $A_1A_2\ldots A_n$ *and in the center of* $A_1A_2\ldots A_n$*, when* n *is even.*

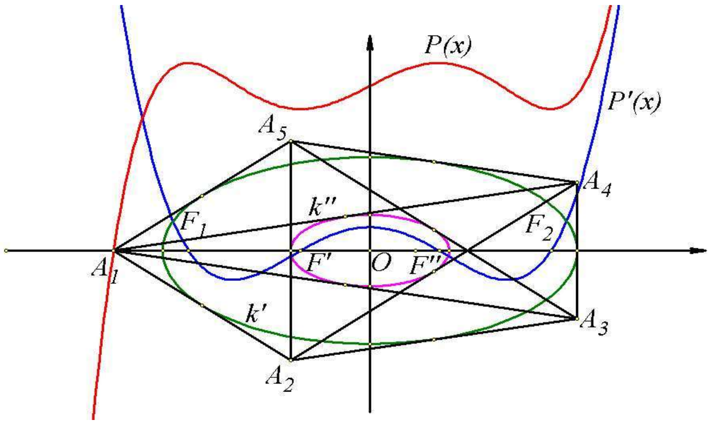


Fig. 17

If $P(z)=P(x)$ is a polynomial with real coefficients of real variable we could present some geometric interpretations of theorem 7. Fig. 17 presents a polynomial $P(x)$ with roots in the vertices of a *meditangential pentagon*, which is symmetric with respect to the abscise axis. Fig. 18 presents a polynomial $P(x)$ with roots in the vertices of a *meditangential pentagon*, symmetric with respect to the abscise axis, which has a six-multiple real root in the point A_1, six multiple complex conjugated roots in the pairs of points (A_2,A_5) and (A_3,A_4). This case demonstrates corollary 4.

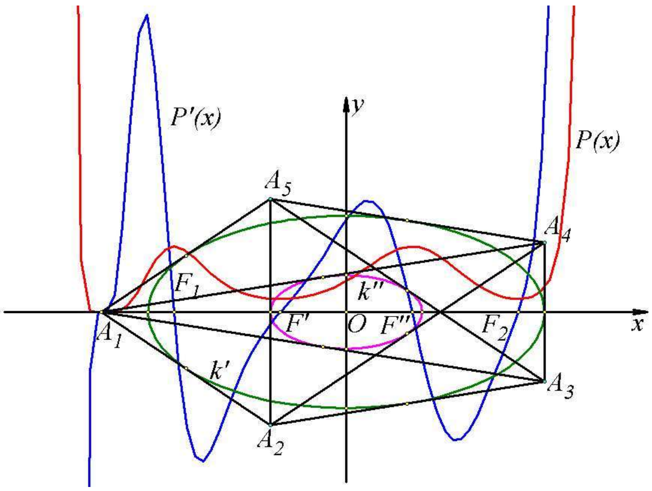


Fig. 18

REFERENCES

[1] V. Nenkov, *A relation between non-real roots and a real ellipse*, Mathematics Plus, 2, Sofia (2010), 61–63, ISSN 0861-8321.

[2] S. Grozdev, V. Nenkov. *Polynomials with multiple roots in the vertices of a triangle*, Mathematics and Informatics, 4, Sofia (2018), 352–359, ISSN 1310-2230.

[3] G. Genov, S. Mihovski, T. Molov, *Algebra with Number theory.* Nauka I Izkustvo, Sofia (1991).

[4] S. Grozdev, V. Nenkov, *Polynomials of forth degree with roots in the vertices of a parallelogram*, Mathematics and Informatics, 3, Sofia (2018), 283–293, ISSN 1310-2230.

[5] S. Grozdev, V. Nenkov. *Polynomials with multiple roots in the vertices of a parallelogram*, Mathematics and Informatics, 4, Sofia (2019), 435–443, ISSN 1310-2230.

[6] S. Grozdev, V. Nenkov. *About the polynomials with roots in the vertices of a class convex quadrilaterals*, Mathematics and Informatics 3, Sofia (2020), 324–329, ISSN 1310-2230.

[7] S. Grozdev, V. Nenkov. *About the polynomials with roots in the vertices of a class pentagons*, Mathematics Plus, 3, Sofia (2020), 45–55, ISSN 0861-8321.

[8] S. Grozdev, V. Nenkov. *Polynomials with roots in the vertices of meditangential polygons*, Mathematics and Informatics, 5, Sofia (2019), 544–549, ISSN 1310-2230.

WORKING WITH MATHEMATICALLY GIFTED STUDENTS IN PRIMARY EDUCATION – PART ONE

UDC: 373.3.011.3-056.317:51

Risto Malcheski[1)], Aleksa Malcheski[2)], Katerina Anevska[3)], Metodi Glavche[4)]

Abstract. The mathematics content enables differentiation of instruction as early as initial education. For this purpose, it is essential to discover and identify mathematically gifted students. After discovering and identifying the mathematically gifted students, we need to organize instruction that will result in faster advancement. For this goal, in this paper we have provided an integral teaching program for working with mathematically gifted students in Grade VI (Grade V) in the nine-year (eight-year) long primary education. This teaching program is an upgrade of the corresponding teaching programs presented in [3] – [5] intended for students in the initial education, i.e. students in Grades II – V in the nine-year long education, i.e. students in Grades I – IV in the nine-year long education. Alongside, almost the complete teaching program is supplemented by a system of tasks for particular topics (solved and unsolved), given in [10] – [17].

1. INTRODUCTION

Work with mathematically gifted students is usually carried out as a part of mathematical sections which help the students prepare for the upcoming mathematics competitions. Practice shows that the educational process does not give as much attention as needed to identify the mathematically gifted students or to work with them. Our analysis will not focus as much on identifying these students, because this matter is thoroughly treated in the existing literature, such as [1], [2], and [18].

During the last several years we have seen efforts to organize systematic and continuous work with mathematically gifted students. Nevertheless, these efforts were accompanied by numerous contradictions, such as the absence of adequate programs, as well as lack of necessary literature for realizing these programs. Having this into consideration, we made an effort to develop integral programs for work with mathematically gifted students in the initial education, i.e. students in grades III-V (I-IV) in the nine-year long (eight-year) primary education in the papers [3], [4] and [5]. These papers also recommend adequate literature that covers the suggested programs in a significant part. The literature has been enriched with new collections of tasks in recent years.

2010 *Mathematics Subject Classification.* Primary: 97-XX
Key words and phrases. mathematically gifted students, primary education

This paper is a continuation of the papers [3], [4] and [5], and its goal is to offer adequate integral programs for work with mathematically gifted students aged 11-12. Taking into consideration the experiences in the Republic of Macedonia and some other countries, we are going to make an effort to offer a system of tasks for one of the topics in the program. This system of tasks can be used to discover and identify mathematically gifted students as well as work with them (done in [3], [4], and [5] for the students in the initial education).

2. PROGRAM FOR WORKING WITH MATHEMATICALLY GIFTED STUDENTS AGED 11-12

In this part we are going to present an integral program for work with mathematically gifted students aged 11-12, i.e. students in grade VI (grade V) in the nine-year long (eight) primary education. This teaching program should be realized continuously, not only when the students are practicing for mathematics competitions.

Goals of the program for students aged 11-12:

- The student understands and adopts operations on sets and uses them when solving tasks,
- The student distinguishes finite and infinite sets,
- The student understands the decimal number system and completes arithmetic operations with natural numbers,
- The student understands the commutative and associative laws of addition and multiplication, and the distributive law and can apply these laws when calculating the numerical expressions,
- The student is able to solve word problems arithmetically,
- The student is able to solve equations with one variable and can apply them when solving word problems,
- The student adopts divisibility rule of 2, 3, 4, 5, 8, and 9 and can use this knowledge when solving appropriate tasks,
- The student adopts the terms greatest common divisor and least common multiple and can use them when solving tasks,
- The student adopts the terms prime and composite number and can solve tasks with prime and composite numbers,
- The student can compare and reduce fractions and does operations with fractions,
- The student transforms fractions into decimal numbers and percentages and can do operations with decimal numbers,
- The student distinguishes the basic terms (point, line, plane, distance) from the derivative terms (semi-line, line segment, angle, polygon, circle, etc.),
- The student adds line segments graphically and arithmetically,
- The student adds angles graphically and arithmetically,

- The student adopts the measurement units for mass, length, time, temperature, area and volume,
- The student operates with named numbers and transforms single-named into multi-named numbers,
- The student calculates the perimeter and area of a rectangle and a square, as well as the area of more complex shapes composed of rectangles and squares,
- The student calculates the area of a cube and a cuboid, as well as more complex shapes composed of cubes and cuboids,
- The student solves basic logical tasks,
- The student understands the basic combining principles and combining configurations,
- The student solves elementary tasks with coloring, covering and dissecting figures into simpler figures,
- The student uses the invariant method at an elementary level,
- The student develops qualities of thinking, such as elasticity, pattern-making, depth, rationalization, and critical thinking,
- Efforts will be made for the student to adopt the scientific methods informally: observation, comparison, experiment, analysis, and synthesis,
- Efforts will be made for the student to adopt the types of conclusion-making informally: induction, deduction, and analogy, while presenting suitable examples from which the students will learn that analogously based conclusion is not always correct.

The following content needs to be learned in order to achieve the previously mentioned goals:

Topic I. *Sets*: definition and notion for a set, ways of presenting sets, a subset, equal sets, set operations (union, intersection, difference, and Cartesian product), and tasks with Venn diagrams.

Topic II. *Natural numbers*: the set of natural numbers, extended set of natural numbers, notion for a finite and an infinite set, decimal number system, operations with natural numbers (addition, multiplication, subtraction, and division) and their properties, numerical expression, order of arithmetical operations, order of natural numbers, inequalities, degrees, extended form of a natural number in a decimal entry, solving equations with one variable, solving numerical rebus puzzles with addition, subtraction, multiplication and division, and sequences of numbers which satisfy a certain property and magical figures.

Topic III. *Number theory*: divisibility in the set of natural numbers, division with remainders, general divisibility properties, divisibility properties of 2, 3, 4, 5, 9 and 11, prime and composite numbers, decomposing a composite number into prime factors, common divisors and common multiples of two or more natural numbers, greatest common divisor and least common multiple, elementary Diophantine equations.

Topic IV. *Fractions*: Decimal numbers: fractions, reading and writing fractions, representing fractions on a number line, expanding and reducing fractions, comparing fractions, arithmetic mean of natural numbers, addition and subtraction of fractions, decimal fraction and a decimal number, comparing decimal numbers, operations with decimal numbers (addition, subtraction, multiplication and division), transforming fractions into decimal numbers and vice versa, finite and infinite decimal number, periodic decimal number and decimal rounding, and solving equations with one variable with decimal numbers.

Topic V. ***Solving word problems*****:** tasks with numbers and numerals, tasks with measure numbers, percentage tasks, and money tasks.

Topic VI. Geometry: point and line, mutual position of two lines, the distance between two points, semi-line, and line segment, length of a line segment, basic and derivative terms, graphic and arithmetic line segments addition, length of a broken line and perimeter of a polygon, and a circle, common position of a circle and a point, circle and a line, and of two circles, the notion of an angle, adjacent, linear and vertical angles, measuring angles, types of angles according to size, graphic and arithmetic angles addition, complementary and supplementary angles, bisection of a line segment and bisection of an angle, distance from a point to a line, convex and concave polygons, types of polygons, triangle (axial symmetry of a triangle, triangle inequality), elementary constructive tasks, area and perimeter of a square and a rectangle and more complex shapes consist of squares and rectangles, area of a cube, a cuboid and more complex shapes built of cubes and cuboids.

Topic VII. ***Logic and combinatorics*****:** elementary logical tasks, classical logical tasks (finding the culprit, liar, etc.), Dirichlet's principle (intuitive use), method of invariants (elementary level), counting and recounting by using the principles of sum, difference, and product (intuitive use), coloring, covering and dissection, comparison of weight and pouring liquids, elementary games and strategies.

3. AN EXAMPLE OF A SYSTEM OF TASKS FOR WORKING WITH MATHEMATICALLY GIFTED STUDENTS AGED 11-12

Adequate instructional materials, i.e. textbooks, accompanied by collections of tasks have to be developed to realize the suggested program for work with mathematically gifted students aged 11-12. The textbook and the collection of tasks must contain adequate tasks that will be used to discover and identify mathematically gifted students. We are going to present an example of a system of tasks for this age group. The tasks cover the topic Sets, part of the collections of tasks [10] – [17].

Task 1. Let
$A=\{a\,|\,2 \text{ is devisor of } a, a\in\mathbb{N}\}$ and $B=\{b\,|\,3 \text{ is devisor of } b, b\in\mathbb{N}\}$. and. Determine the set $A\cap B$.

Task 2. Let

$$A=\{a \mid a\in\mathbb{N}, a\leq 7\} \text{ and } B=\{b \mid b\in\mathbb{N}, 4\leq b<9\}.$$

Determine the elements of the set C if

$$C=\{c \mid c\in\mathbb{N}, c=a-b, a\in A, b\in B\}.$$

Task 3. Let

$A=\{a \mid a\in\mathbb{N}, a<7\}$, $B=\{b \mid b\in\mathbb{N}, 5<b<6\}$ and $C=\{c \mid c\in\mathbb{N}, 5<c<8\}$. Determine the set $(B\cap C)\setminus A$.

Task 4. Determine the elements of the set B if

$$A\cup B\cup C=\{a,b,c,d,e,m,n,n,p,q\},\ A\setminus B=\{e,m\},$$
$$A\cap C=\varnothing,\ C\setminus B=\{p,c\}.$$

Task 5. Let $A=\{1,2,3\}$ and $B=\{3,4,5\}$. Determine the sets X if $A\cup X=A$ and $B\cap X=A\setminus(A\setminus B)$. How many solutions are there?

Task 6. Let $S=\{1,2,3,4,5,6,7,8,9,10\}$. Are there exist sets A and B such that $A\cup B=S$, $A\cap B=\varnothing$ and the sum of the elements of the set A equals the sum of the elements of the set B?

Task 7. Let $A=\{1,2,3,4,5,6,8\}$ and $B=\{2,4,5,6,8\}$. Determine the set S for which the following equations are true

$$A\cap S=\{3,4\} \text{ and } B\cup S=\{2,3,4,5,6,7,8,9\}.$$

Task 8. Let

$$A=\{a,b,c,d,e\}, B=\{a,d,f\}, C=\{b,e,f,g\}, D=\{a,f,g,h\}.$$

Determine the set S given that

$$S\subseteq A,\ S\cap(B\cup D)=\varnothing,\ (A\cap C)\setminus S=\varnothing,\ \{c\}\setminus S=\{c\}.$$

Task 9. Let $A=\{5,2x+2\}$ and $B=\{2x+1,y-3\}$, for x and y natural numbers.Determine all the values of x and y such that B is the subset of A.

Task 10. Let $A=\{x \mid x=2k, k\in\mathbb{N}\}$ and $B=\{x \mid x=2k-1, k\in\mathbb{N}\}$. Using symbols determine the relationship between the sets

a) A and $\mathbb{N}$, b) B and $\mathbb{N}$,

c) A and B, d) $A\cup B$ and $\mathbb{N}$.

Task 11. The elements 1, 2, 3, 4, 5, 6 7, 8 are the only elements that form the sets A,B and C. Determine the sets A,B and C given that the following conditions are fulfilled:

a) none of the elements belong to all three sets,

b) the elements 3 and 8 belong only to set B,

c) each of the elements 1, 4, and 7 belong to only one of the sets A,B and C,

d) the set A contains only two elements,

e) the element 2 belongs exactly to two of the sets A,B and C,

f) $1\notin B, 2\notin A, 4\in A, 6\in A, 6\in B, 7\notin B$.

Task 12. Determine the sets A and B which fulfill the conditions: $A\cup B=\{2,3,4,5,6,7\}$, $A\cap B=\{3,4,5,6\}$, $7\notin A\setminus B$ and $2\notin B\setminus A$.

Task 13. Determine the sets A, B and C which fulfill the conditions:

$$A\cup B\cup C=\{1,2,3,4,5,6,7,8,9\},$$
$$A\cup B=\{1,2,3,6,7,8\},$$
$$A\cup C=\{1,2,3,4,5,7,9\},$$
$$A\cap B=\{1,2\}, \qquad A\cap C=\{3,7\}.$$

Task 14. Let $S=\{x\,|\,x=3a+5, x<20\}$.

a) For which values of the number a from the set $\mathbb{N}_0$ the set S will not be empty?

b) For which values of the number a from the set $\mathbb{N}$ the set S will be empty?

Task 15. Let

$A=\{0,2,3,5,9\}$, $B=\{1,2,7,8,9\}$, $C=\{2,4,5,6,7\}$ and $D=\{2,4,5,6,7,9\}$. Represent the set D in terms of the sets A,B,C and the set operations.

Task 16. Let:

$$A=\{x\,|\,x\in\mathbb{N}, x\leq 10\},\ B=\{x\,|\,x\in\mathbb{N}, 5\leq x<15\} \text{ and}$$
$$C=\{x\,|\,x\in\mathbb{N}, x\leq 12 \text{ and } x \text{ is an even number}\}.$$

Prove that for the given sets, the following is true

$$A\setminus(B\cup C)=(A\setminus B)\cap(A\setminus C).$$

Task 17. Determine the sets A,B,C such that the following hold true:

$$A\cup B=\{2,3,4,5,6,7,8\}, \qquad A\cap B=\{2\},$$
$$B\cup C=\{1,2,4,6,8\}, \qquad B\cap C=\{2,4,8\},$$
$$C\cup A=\{1,2,3,4,5,7,8\}, \qquad C\cap A=\{2\}.$$

Task 18. Let $A=\{x\,|\, x\in\mathbb{N} \text{ and } x<6\}$ and $B=\{4,5,6,7,8\}$. Determine the set X if $X\subset(A\cup B)$, $X\cap A=A\setminus B$, $X\cap B=B\setminus A$.

Task 19. Determine the sets A, B and C, given that

$$A\cup B\cup C=\{1,2,3,4,5,6,7,8\},\ A\cap B\cap C=\{3\},\ C\setminus(A\cup B)=\{5,8\},$$
$$A\cap B=\{1,3\},\ B\setminus C=\{1,2\},\ A\setminus C=\{1,4\},\ B\cap C=\{3,6\}.$$

Task 20. Determine the sets A, B and C, if

$$A\cup B\cup C=\{n\,|\,n\in\mathbb{N} \text{ and } n \text{ is a one digit number}\},$$
$$A\cap B\cap C=\{2\},\ C\setminus(A\cup B)=\{1,3,5\},\ (B\cap C)\setminus A=\{6,7\},$$
$$(A\cap C)\setminus B=\varnothing,\ A\cap(B\cup C)=\{2,4,5\}.$$

Task 21. Determine the sets A and B if

$$A\cup B=\{1,2,3,4,5,6,7,8,9,10,11,12\},\ A\cap B=\{8,9,10,11,12\}$$

and the sum of all the elements of A equals the sum of all the elements of B.

Task 22. Let $A=\{1,2,3,4\}, B=\{0,1,2\}, C=\{5,6\}$. Check if the following equations are correct:

$$(A\cap B)\times C=(A\times C)\cap(B\times C) \text{ and } (A\setminus B)\times C=(A\times C)\setminus(B\times C).$$

Task 23. Each of the members of one team plays football or tennis. How many members are there in the team if 18 play football and tennis, 23 play football and 21 tennis?

Task 24. An international convention was the place where a group of people from Spain, Finland, Mongolia and Vietnam, a total of 21 people met. Five of them speak Spanish, 14 Finnish, 14 Mongolian and 10 Vietnamese. One Spaniard speaks only his native tongue. Two people from Finland speak Mongolian and Vietnamese, but not Spanish. Of the remaining people, none speak more than two languages, and eight people speak Finnish and Mongolian. How many people from Spain speak Vietnamese?

Task 25. A sports club has 60 members, of which 39 play football, 28 handball and 16 play both football and handball. Are there members of the club who neither play football nor handball?

Task 26. A class has 31 students. Each of the students can either swim or drive a bicycle, and 11 students can both swim and ride a bicycle. How many students can swim and how many can ride a bicycle if there are twice more students who can swim than students who can ride a bicycle?

Task 27. A class has 33 students. If 21 play basketball, 18 football and 6 neither basketball nor football, how many students play both football and basketball?

Task 28. In Gorjan's school in sixth grade, the students learn three foreign languages: 2 French, English and Russian, 9 only French and English, 13 French and English, 12 Russian and English, 29 English, 6 learn only French and 7 learn only Russian.

a) How many students in total are there in sixth grade?

b) How many students learn the Russian language?

c) How many students learn the French language?

Task 29. Let A be the set of natural numbers less than 2018, divisible by 4, B is a set of natural numbers less than 2018 which can be divided by 6 and C is a set of natural numbers less than 2018 which can be divided by 15. Determine the number of elements of the set $A\setminus(A\cap B\cap C)$.

Task 30. Let A be a set of natural numbers less than 500 and divisible by 2, B be a set of natural numbers less than 500 divisible by 3, C be a set of natural numbers less than 500 divisible by 4 and D be a set of natural numbers less than 500 divisible by 5. Determine the set $A\cap B\cap C\cap D$.

4. CONCLUSIONS

One of the forms for differentiation of instruction in primary education is the work with mathematically gifted students. This work is usually carried out within the preparation of the students for participation in many mathematical competitions. This practice, in our opinion, is neither a successful differentiation of instruction nor well-organized work with mathematically gifted students. The work with the mathematically gifted students should be organized in the following way:

- The acknowledgement and identification of the mathematically gifted students will be carried out in the upper grades as well, because when this is done only during the initial years of elementary education, most often than not, the gifted students characterized with width, critical thinking and depth, remain off the radar of the teacher,
- A special program will be developed for every age group, such as the program in this paper and it will be realized during the entire school year. Content which is not included in the regular program will be learned during extracurricular instruction, where material from the regular program will be upgraded, and
- For realizing the program with work with the gifted students adequate didactic material will be developed (a complete textbook or mathematical manuals for the content not included in the regular program and collections of tasks that cover the entire program for work with mathematically gifted students).

CONFLICT OF INTEREST

No conflict of interest was declared by the authors.

AUTHOR'S CONTRIBUTIONS

All authors contributed equally and significantly to writing this paper. All authors read and approved the final manuscript.

References

[1] Đorđević, B.: *Školski oblici podrške u razvoju nadarenih učenika*, Nastava i vaspitanje, Beograd, 1990

[2] Gallagher, J. J.: *Teaching the gifted children*, Boston, Allyn & Bacon, 1985

[3] Glavche, M., Anevska, K., Malcheski, R.: Working with mathematically gifted students in the initial education – part one, Teacher, Faculty of education, Bitola

[4] Glavche, M., Anevska, K., Malcheski, R.: Working with mathematically gifted students in the initial education – part two, Conference Procededings of the Second International Scientific Conference: Gifted and talented creators of the progress, Faculty of Education – Bitola