

MODELLING OF BURNT FOREST AREAS IN BULGARIA BY MODIFIED METHOD OF THE TIME SERIES AND REGRESSION ANALYSIS

Mitko M. Petrov. Tatiana S. Ilkova

Institute of Biophysics and Biomedical Engineering – Bulgarian Academy of Sciences.

105 Acad. George Bonchev Str. 1113 Sofia, Bulgaria

Abstract

In this study a models of burnt forest areas and numbers of fires in Bulgaria has been developed with a modified method of the time series analysis (TSA). The method of regression analysis also has been used for modeling of burnt forest area in dependence from the following climatic indexes: rainfall, maximum twenty-four-hour period rainfall, maximum height of the snow-cover, and overall year temperature. The basic tendencies in the evolution of the burnt forest areas for the period from years 1988 to 2010 on all investigated indexes has done.

Key words: modelling, times series analysis method, regression analysis, burnt forest areas

1. INTRODUCTION

In recent years vastly increase of the fire activity in the last decade in the forest and agricultural region is observed.

In [12] the fire history in 23-years period from 1988 to 2010 year is investigated. The results for the different years show in the second part of the investigated period the fire activity is vastly extended.

The previous fire regimes (to 1988) were such that in Bulgaria a perfect and powerful system for the fire control has not necessary. The statistic shows the activity rural population and old system for the fire control successfully put out the fires in the forest areas before 1990. At creating of a prognosis systems of the danger from fires it is importantly at first to be understood what is the process of burning of the fires and what are the probable parameters that influence under their behavior. Important measurements can exert influence [7]:

- The prevalent atmosphere conditions and short-term effects of the rainfall together determine the index of burning for the plant fuelling materials. The index can be integrated with physical characteristics of the fuel the probability from beginning of a fire and the difficulty at its crushing for concrete condition to be fixed.
- Fuels – structure and distribution, accessibility, (fade/marasmus), quantity, continuity, type of the fuels, size and form of the fragments, all that has a significance for the characteristic of the burning and contributes for accelerating of the mechanisms of carrying such as fire spots, received from carrying of fuel fraction by wind.
- Topography – influences under the level of the fire distribution. For given meteorological conditions the fire moving up on the slope the level of distribution can double increased in comparison to the such of a flat area for each 10 degrees incidence on the slope.

The analysis of the time series shows the availability of periodical and long-term changes of some indices for burnt forest areas, which may be used in water quality management. On analyzing the time series the physical, chemical, biological and biochemical characteristics of the processes are not directly taken into consideration; an attempt however, was made for a description in the time of burnt forest areas in definite points as an integral result of forest ecosystem functioning. The time series analysis for forest quality is a part of the retrospective modelling of forest ecosystem functioning carried out with deterministic methods. These models use information from environmental monitoring [1-3, 5, 6, 11, 12, 14], and their realizations.

The major goal of the present investigation is to develop a model for the analysis and prognosis of burnt forest areas in Bulgaria, based on time series analysis method and regression analysis.

2. MATERIALS AND METHODS

For modelling of burnt forest areas, a method of TSA is a determined component (trend) – $\mathbf{x}_T$ describing the regularity of the development of the examined phenomenon, periodical component ($\mathbf{x}_p$) and stochastic variable (ε_t) [1, 2]:

$$\mathbf{x} = \mathbf{x}_T + \mathbf{x}_p + \varepsilon_t \quad (1)$$

where $\mathbf{x}$ is the vector of the investigated indexes of the pollution.

The determined component (trend) – $\mathbf{x}_T$ is a polynomial of 1st to 3rd degrees and the periodical component – $\mathbf{x}_p$ is described by the order of Fourier.

The main trend shows the main tendencies in the alteration of the studied indices, and it is a straight line:

$$\mathbf{x}_T^M = A_0 + A_1 t \quad (2)$$

In contrast to the conventional method of TSA the determined component as is a polynomial to high degrees:

$$\mathbf{x}_T = \sum_{j=0}^r a_j t^j \quad (3)$$

where a_j – coefficients of polynomial, $j = 0, 1, 2, \dots, r$; r – degree of the polynomial, $r \leq 5$.

In contrast to the classical method for analyzing temporary series, where the Fourier series are used, the present research proposes the uses of the periodical functions from the type:

$$\mathbf{x}_p = \sum_{k=0}^p b_k \sin\left(\frac{2\pi}{c_k} t + d_k\right) \quad (4)$$

where: p – number of the periodical functions; b_k , c_k and d_k – coefficients in the periodical functions, $k=0, \dots, p$.

The number of the periodical functions p in (4) and the polynomial degree (2) is determined on basis the statistical criteria experimental Fisher function – F_E and experimental correlation coefficient – R^2_E .

Then the model (1) for analysis and prognosis has the following form ($\varepsilon_t=0$):

$$\mathbf{x} = \sum_{j=0}^r a_j t^j + \sum_{k=0}^p b_k \sin\left(\frac{2\pi}{c_k} t + d_k\right) \quad (5)$$

An algorithm and program is developed of the modified TSA.

2.1. The main algorithm of the Modified TSA

The generalized algorithm scheme is:

1. BEGINNING

2. Statistical treatment and analysis of main indexes for the burnt forest areas. Computing of statistical the experimental data. For this aim an algorithm and program developed in [9, 10]. Output results from subroutine program are: the real, centred or normalized experimental data. The centred the data are

computing is: $\tilde{x}_{i,j} = x_{i,j} - \bar{x}_j$, where: $\bar{x}_j = \frac{1}{n} \sum_{i=1}^n x_{i,j}$; n – number of time partitions ($i=1, \dots, n$); m –

number of investigation parameters of the indexes, ($j=1, \dots, m$). Subprogram also makes the following: check-ups for presence of trend, test for stationary, computing of autocorrelation, correlation matrix, and significance of correlation coefficients.

3. The computing time in year [8, 10]: $t = Year + (Month - 0.5)/12$ if necessary.
4. The coefficients A_0 , A_1 are determined in the main trend (2), and the coefficients a_j , $j=0, \dots, r$ in trend (3). The procedure is carried out with the help of the subroutines RLINE and RCURV from IMSL library of Compaq Visual FORTRAN 90 Pro.

5. The experimental values of the periodical function: $\mathbf{x}_p = \mathbf{x} - \mathbf{x}_T$ are calculated.
6. The coefficients b_k , c_k and d_k in periodical functions (4) and their numbers are determined. Comparison according to F and R^2 . The procedure is carried out with the help of the subroutine for nonlinear regression RNLIN from IMSL library of Compaq Visual FORTRAN 90 Pro.
7. If necessary the models (5) additional are determined by the programs for SIMPLEX method for optimization. The optimization program for direct search of the minimum of the criteria. The minimization criterion represents the Sum of the Squares of Weighed Residues:
$$SSWR = \sum_{i=1}^n \frac{(x_i^m - x_i)^2}{\max(x_i, x_i^m)^2}.$$
8. Computing adequacy of the models: the correlation coefficient (R_E^2), the experimental value of the Fisher's function (F_E) and the relative error (S_L). The S_L is calculated by [4, 13]:
$$S_L = \sqrt{\frac{1}{(n-\nu)} \sum_{i=1}^n \left(1 - \frac{x_i^m}{x_i^e}\right)^2},$$
 where x_i^e are the experimental values, x_i^m are the values calculated with the help of the model; ν – degrees of freedom; n – number of the experimental data points.
9. Comparison with the theoretical correlation coefficient R_T^2 and Fisher's function F_T is made. IF $R_E^2 > R_T^2$ and $F_E < F_T$, then the model is adequate. The lower the values of the relative error S_L are the higher the validity of the model for the investigated indexes.
10. Write: R_E^2 , F_E , and S_L .
11. Write: Time, $\mathbf{x}_M^T$ from (2), $\mathbf{x}_T$ from (3), model ($\mathbf{x}$) for analysis from (5).
12. **END**

The main algorithm is shown in Fig. 1. The coefficients A_0 , A_1 are determined in the main trend (2) with RLINE the procedure of a linear regression. The coefficients a_j , $j=0, \dots, r$ in trend (3) with RCURV is a procedure of the polynomial regression. RNLIN is a procedure for nonlinear regression. **SIMPLEX** is a procedure for optimization. SIMPLEX serves for specifying of the model coefficients, if it is necessary.

2.2. Algorithm for statistical treatment data

The block scheme of the algorithm [9] is shown in Fig. 2, where: **SPLINE** is a procedure for extrapolation of the experimental data in the time; **STATIONARY** is a procedure for check-up of the *stationary* of the experimental observations; **TREND** is a procedure for test for presence of a *trend* (check-up for normally distribution of the observations); **CORR** is a procedure for calculating of the auto- and correlation functions; **STUDNT** is a procedure for determining of the correlation coefficients. All programs are developed on Compaq Visual FORTRAN 90.

3. MODELLING OF THE BURNT FOREST AREAS AND NUMBER OF FIRES

The burnt forest areas in Bulgaria is carried out on the base on the data obtained by the Global Fire Monitoring Centre (GFMC) for International Forest Fire News for the period from years 1988 to 2010 [7]. The value of the measured indexes are shown in Table 1.

In Table 1 y_1 is burnt forest area and y_2 is number of fires, and climatic indexes x_1 –rainfall, x_2 – maximum twenty-four-hour period rainfall, x_3 –maximum height of the snow-cover, x_4 –overall year temperature.

3.1. Modelling of the burnt forest areas and number of fires with modified TSA

We examined the burnt forest area (y_1) and number of the fires (y_2) in Bulgaria for period from 23 years with modified TSA. The model (5) has the type:

$$\mathbf{y} = \mathbf{a}_0 + \sum_{j=1}^5 \mathbf{a}_j t^j + \boldsymbol{\theta}_1 \sin\left(\frac{2\pi}{\boldsymbol{\theta}_2} t + \boldsymbol{\theta}_3\right) + \boldsymbol{\theta}_4 \sin\left(\frac{2\pi}{\boldsymbol{\theta}_5} t + \boldsymbol{\theta}_6\right) \quad (6)$$

where: $\mathbf{y} = \mathbf{y}[y_1 \ y_2]^T$, $\mathbf{a}_0 = \mathbf{a}_0[a_{0,1} \ a_{0,2}]^T$, $\mathbf{a}_j = \mathbf{a}_j[a_{1,1} \ a_{1,2} \dots a_{5,1} \ a_{5,2}]^T$, $\boldsymbol{\theta}_j = \boldsymbol{\theta}_j[\theta_{1,1} \ \theta_{1,2} \dots \theta_{6,1} \ \theta_{6,2}]^T$.

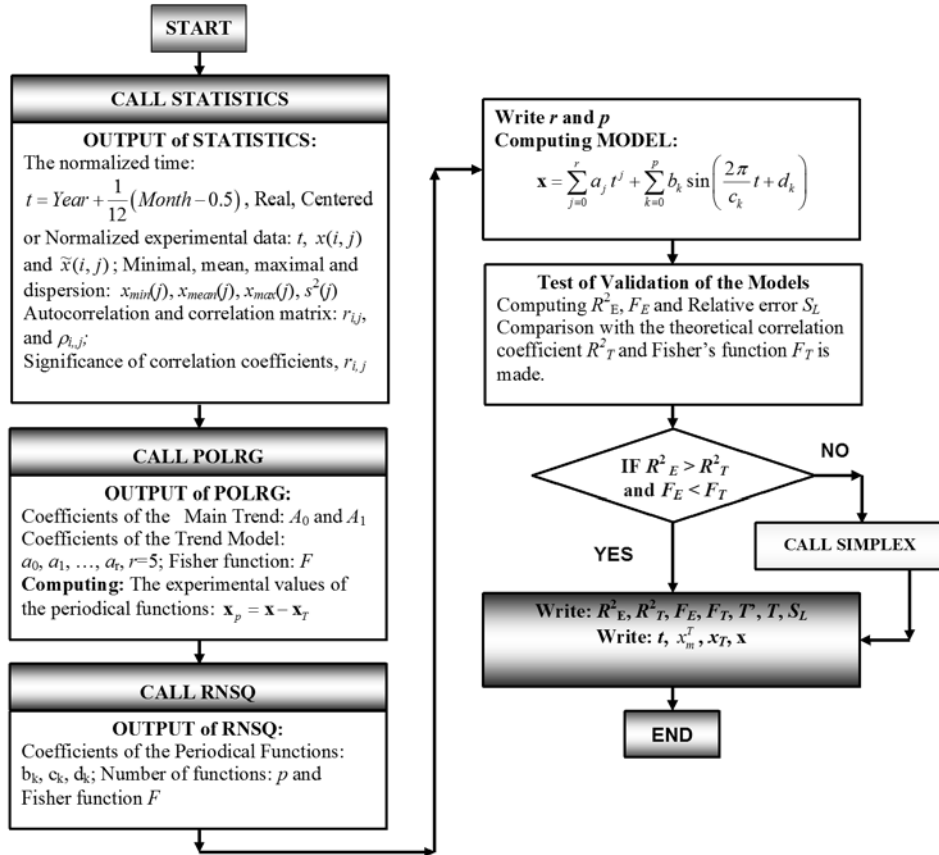


Figure 1. Main algorithm of modified TSA

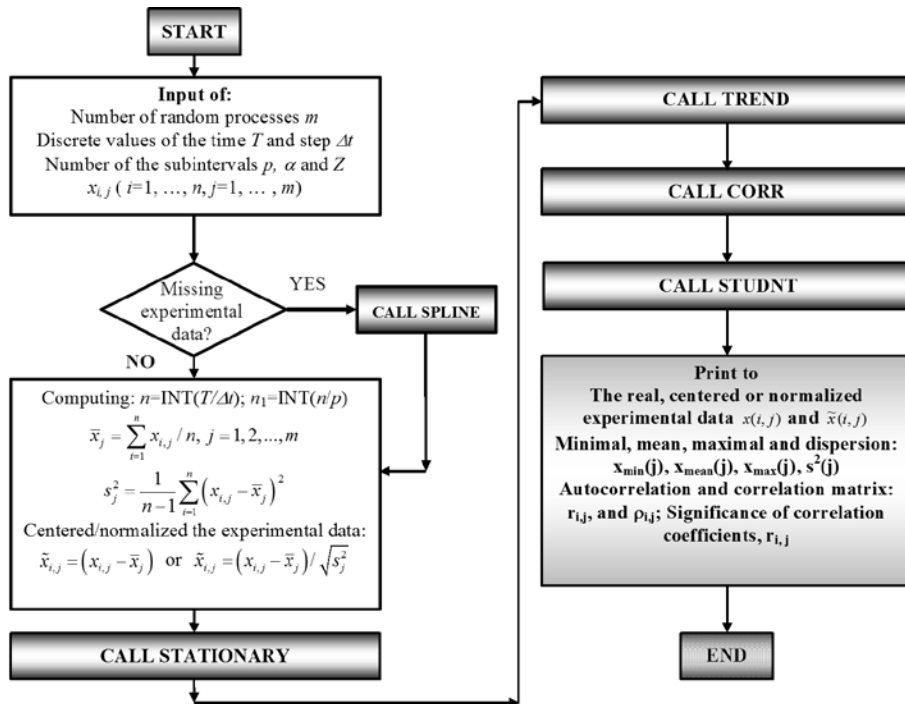


Figure 2. Block scheme of the algorithm for sub-programme Statistics

The parameters in main trends (2), trend functions (3), periodical functions (4) and full models (6) for the burnt forest area (y_1) and number of fires (y_2) with modified TSA are shown in Table 2.

Table 1. Fire statistic for Bulgaria for the period from 1988 to 2010 year

Year	y_1	y_2	x_1	x_2	x_3	x_4	Year	y_1	y_2	x_1	x_2	x_3	x_4
1988	173	67	645	184.9	60	11.2	2000	4154	602	456	138.0	33	11.7
1989	199	81	673	183.5	55	10.7	2001	1147	1196	475	92.4	50	11.6
1990	168	110	452	134.5	43	12.3	2002	3708	667	528	263.0	31	13.0
1991	126	47	863	164.4	52	11.5	2003	527	114	697	135.4	64	11.2
1992	98	51	949	104.5	56	11.2	2004	1933	246	599	122.2	44	11.0
1993	1151	100	563	143.4	32	11.1	2005	472	200	662	110.6	55	11.3
1994	340	57	342	199.8	53	12.3	2006	3060	578	678	157.6	61	12.1
1995	301	81	784	156.6	49	11.4	2007	4198	320	633	268.5	54	12.1
1996	462	101	579	197.5	60	11.9	2008	3431	1710	377	160.0	65	12.4
1997	223	63	546	164.0	44	12.4	2009	1463	825	549	100.5	62	12.3
1998	1012	208	459	135.5	38	11.8	2010	5910	402	743	158.0	55	11.9
1999	471	73	641	182.0	45	10.6	2011	–	–	–	–	–	–

Table 2. Parameters in the models

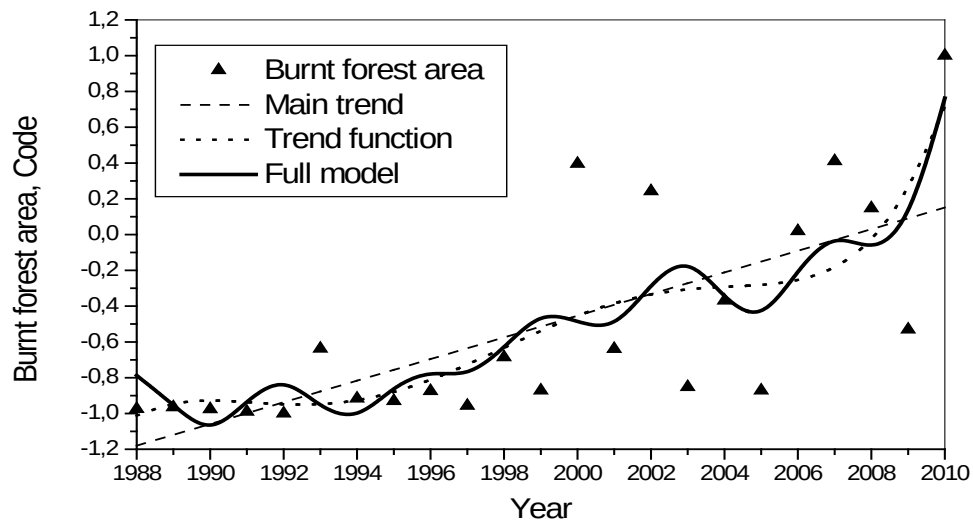
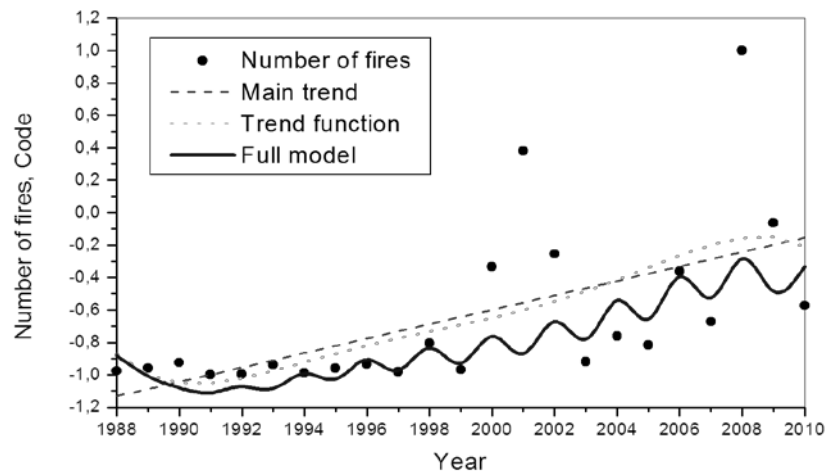
Burnt forest area, $y_1=f(\text{Year})$			Number of fires, $y_2=f(\text{Year})$		
Parameters in main trend					
$A_{0,2}=-1.240252$		$A_{1,2}=0.06050824$		$A_{0,1}=-0.1177$	
				$A_{1,1}=0.0445$	
Parameters inn trend functions					
$a_{0,2}= -1.1930$		$a_{1,2}= 0.2525$		$a_{0,1}=-0.6401$	
				$a_{1,1}=-0.2952$	
$a_{2,2}= -0.0816$		$a_{3,2}= 0.0107$		$a_{2,1}= 0.0706$	
				$a_{3,1}=-6.7293\times10^{-3}$	
$a_{4,2}= -56.928\times10^{-5}$		$a_{5,2}= 10.5556\times10^{-6}$		$a_{4,1}= 0.2977\times10^{-3}$	
				$a_{5,1}=-0.4902\times10^{-5}$	
Parameters in periodical functions					
$\theta_{1,2}=0.0758$		$\theta_{2,2}=0.0051$		$\theta_{3,2}=0.1234$	
				$\theta_{1,1}=0.1049$	
$\theta_{4,2}=0.1520$		$\theta_{5,2}=0.0974$		$\theta_{2,1}=6.8734\times10^{-3}$	
				$\theta_{3,1}=0.1869$	
		$\theta_{6,2}=0.0775$		$\theta_{4,1}=0.2302$	
				$\theta_{5,1}=0.1431$	
				$\theta_{6,1}=0.0899$	

The results from the modelling with modified TSA of the burnt forest area (y_1) and number of fires (y_2) and are shown in Fig. 3 and Fig. 4.

Table 3. Statistical information of the models. Theoretical Fisher function and theoretical coefficient R_T^2 are

$$F_T(5, 22)=4.5579 \text{ and } R_T^2(21) = 0.347 .$$

Statistical information						
Burnt forest area, $y_1=f(t)$				Number of fires, $y_2=f(t)$		
Model	R^2	F_E	S_L	R^2	F_E	S_L
Main trend	0.7041	2.0171	3.9586	0.5919	2.8541	1.3911
Trend functions	0.7563	1.7484	1.5745	0.6016	2.6735	1.7994
Full model	0.7847	1.7012	1.2052	0.6356	3.9221	1.0619

Figure 3. Modelling of the burnt forest area (y_1)Figure 4. Modelling of the number of the fires (y_2)

3.2. Modelling of the burnt forest areas with regression analysis

Except with modified TSA the burnt forest areas $y_1 = f(x_1, x_2, x_3, x_4)$ are examined from the climatic indexes x_1, x_2, x_3 and x_4 . After a preliminary analysis (comparison on statistical indexes R^2 and Fisher function) from the full model of the regression analysis [5]

$$y = b_0 + \sum_{i=1}^m b_i x_i + \sum_{i=1}^m b_{ii} x_i^2 + \sum_{i=1}^{m-1} \sum_{j=i+1}^m b_{ij} x_{ij}^2 \quad (7)$$

the numbers (b_{ij}) that read interactions drop out:

$$y = b_0 + \sum_{i=1}^m b_i x_i + \sum_{i=1}^m b_{ii} x_i^2 \quad (8)$$

The burnt forest areas and climatic indexes (x_i) are coded in the interval from -1 to +1. The calculated parameters in the model (8) have the following values:

$$b_0 = -0.4507722,$$

$$b_1 = 0.0876752, b_2 = 0.2895196, b_3 = -0.1249314, b_4 = 0.2504738,$$

$$b_{11} = -0.2821069, b_{22} = 0.2530676, b_{33} = 0.4773276, b_{44} = -0.4832127,$$

with correlation coefficient $R^2 = 0.583$, standard error of estimate, $SERR = 0.594$, and Fisher function, $F_E = 0.901$. These results show the regression model (8) is adequate and can be used for investigation of the influence of the included in its climatic indexes on y_1 the burnt forest areas. Fig. 5 shows the experimental and simulated data from the regression model results.

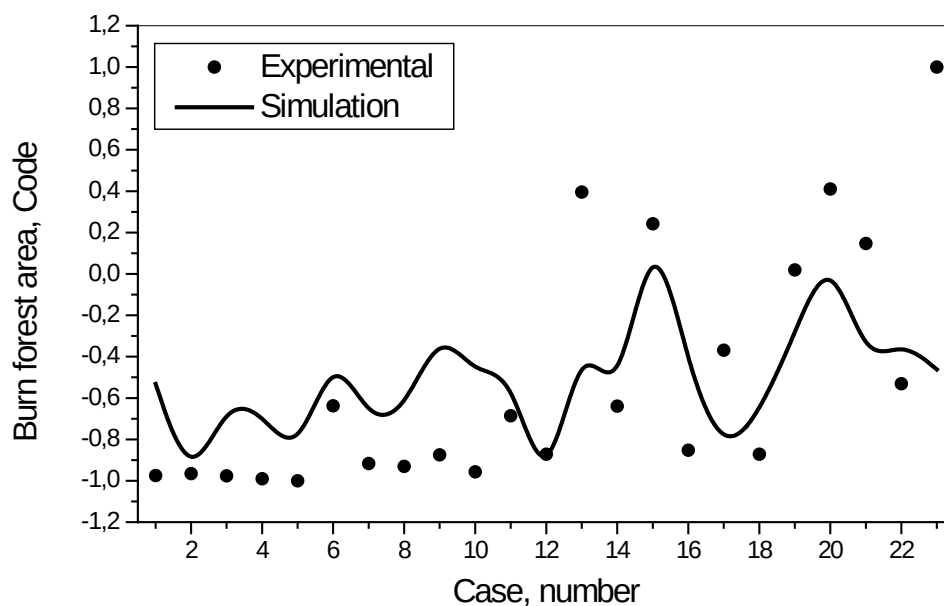


Figure 5. Experimental and simulation results with regression model

The calculated correlation coefficients are in the interval from 0.7041 to 0.7847 and they are larger than the theoretical correlation coefficient. The calculated Fisher functions are in the interval from 1.7012 to 2.0171 and they are smaller than the theoretical function. The relative errors S_L varies from are from 1.2052 to 3.9586 which proves the adequacy of the models for main trend, trend function and full model (Table 3).

The results for modelling of the main trend and trend function (Fig. 3 and Fig. 4) show increase of the burnt forest areas also of the number of fires.

4. CONCLUSIONS

From the modelling of the burnt forest area and number of fires the following conclusions can be drawn:

1. A method, algorithm and program for statistical treatment and analysis of experimental observation for the burnt forest areas in Bulgaria are developed. The statistical data treatment is used at modelling of the burnt forest areas in Bulgaria. The program can be used at the statistical data treatment in similar problems.
2. A model for analysis of the burnt forest areas, based on a modified time series analysis is developed. This approach can be used at modelling and prognosis of different regions in the world. The proposed modified approach for time series analysis can be used on solving different problems connected with ecosystem management. A models for analysis of the burnt forest areas is developed, which adequately expresses the dynamics of the pollution for a defined period.
3. A model for analysis of the burnt forest areas, based on a regression analysis is developed. This approach can be used at modelling of different regions in the world for their modelling of burnt forest and field fire.

ACKNOWLEDGMENTS

The authors are thankful for the support provided by the project DFNI-I-01/0006 "Simulating the behavior of forest and field fires", funded by the National Science Fund. Bulgarian Ministry of Education and Science.

REFERENCES

1. Box G., D. Djenkins (1974) The Time Series Analysis Method. Mir, Moskva, pp 366 (in Russian).
2. Box G., N. Draper (1987) Empirical Model Building and Response Surface. J. Wiley and Sons. N. Y., pp 460.
3. Crabree R., L. Cluckre L.. C. Forste. C. Crocket (1986) A Comparison of Two River Quality Models. *Water Research*. Vol. 20(3), pp. 53-61.
4. Cressie N. (1991) Statistics for Spatial Data. J. Wiley and Sons. N.Y., pp 900.
5. Diadovski I., T. Hristova. M. Petrov (2002) A Model for Assessment and Prognosis of the River Water Quality. *Chem. Biochem. Eng. Q.*, Vol. 16(2), pp. 75-80.
6. http://en.wikipedia.org/wiki/Time_series
7. <http://www.itl.nist.gov/div898/handbook/pmc/section4/pmc4.htm>
8. Ilkova T., M. Petrov (2013) An Application of Regression Analysis for Modelling of the Forest Fire, *Journal of International Scientific Publications: Materials, Methods & Technology*, Vol. 7, Part 1, pp. 444-455
9. Ikova T., M. Petrov (2006) Statistical Treatment and Analysis of Main Indexes for the Water Quality of the Struma river. *Ecological Engineering and Environment Protection*. Vol. 3-4, pp. 10-18, (in Bulgarian).
10. Ikova T., M. Petrov (2007) Modeling and Prognosis of the Strouma River Pollution. *Proceedings of the Int. Symposium BioPS'07*. Sofia. 6-7 November 2007, pp. II.15-II.24.
11. Isaaks E., Srivastava R. (1989) Applied Geostatistics. Oxford University Press. N.Y., p. 561.
12. Project Management Ltd. (2003) Project for development of forest area. Technical report for fires - N: PE-PO33964, Bulgaria
13. Sachs L. (1992) Angewandte Statistik. Anwendung statistischer methoden. Springer Verlag. Berlin. pp 450.
14. Stegmann H.-J. Ehrig. P. Liem (1978) Anwendung Einer Zeitreihenanalyse in der Wassergutewirtschaft. *Wasser und Boden*. Vol. 3, pp. 50-54.