Multiregime SETAR models
with regime switching by means of the aggregation function value

Danuša Szőkeová

Department of Mathematic and Descriptive Geometry
Faculty of Civil Engineering, Slovak University of Technology
Bratislava, Slovakia

Silvia Kohnová

Department of Land and Water Resources Management
Faculty of Civil Engineering, Slovak University of Technology
Bratislava, Slovakia

SETAR models

Self exciting threshold autoregressive (SETAR) models are regime switching models based on the idea there are a finite set of various regimes in the stochastic process.

Regimes are represented by linear AR(p) models and in any time point the process may switch with the past data value (lagged value) relative to the thresholds from one member of the regime set to another one.

Consider an univariate time series $\{y_t\}$ which is observed at time points t=1,2,…,n.

We can generalized two–regime SETAR(c;p,d) model, r=2, in the next way

A) switching by means of the lagged value

$$y_t = (\Phi_1(L)y_t + \varepsilon_t^1)\ \mathrm{I}[y_{t-d} \le c] + (\Phi_2(L)y_t + \varepsilon_t^2)\ \mathrm{I}[y_{t-d} > c]\ ,$$

where

- c is the threshold (a real number, $|c|<\infty$),
- $\Phi_i(L) = \phi_{0,i} + \phi_{1,i}L + \ldots + \phi_{p,i}L^p$, i=1,2, are autoregressive polynoms of order p (L is the backshift operator),
- $\mathrm{I}[A]$ is the indicator function ($\mathrm{I}[A] = 1$ if A is true, $\mathrm{I}[A] = 0$ in another case),
- d is the delay parameter, $d \ge 1$ (a positive integer),
- $\varepsilon_t^i \approx$ i.i.d $N(0,\sigma_i^2)$, i=1,2, σ_i^2 is the regime variance.

B) switching by means of the aggregation function value

$$y_t = (\Phi_1(L)y_t + \varepsilon_t^1)\, \mathrm{I}[F(y_{t-d},...,y_{t-d+k}) \le c] + (\Phi_2(L)y_t + \varepsilon_t^2)\, \mathrm{I}[F(y_{t-d},...,y_{t-d+k}) > c] \,,$$

where

- c is the threshold (a real number, $|c|<\infty$),
- $\Phi_i(L) = \phi_{0,i} + \phi_{1,i}L + \ldots + \phi_{p,i}L^p$, i=1,2, are autoregressive polynomials of order p,
- $\mathrm{I}[A]$ is the indicator function ,
- $F(y_{t-d},...,y_{t-d+k})$ is the value of a k-dimensional aggregation function, $F:R^k \to R$, d is the delay parameter, $1 \le k \le d$, , (d,k – positive integers),
- $\varepsilon_t^i \approx$ i.i.d.$N(0,\sigma_i^2)$, i=1,2, σ_i^2 is the regime variance.

In modified SETAR (c;p,d,k) models we have used the next functions:

1. Mean$(x_1,...,x_m)$, m=2,3,
2. Max$(x_1,...,x_m)$, m=2,3,
3. Min$(x_1,...,x_m)$, m=2,3.

Model estimation

In SETAR model building we expect that the number of regime r is small, observation set is $\{y_1,y_2,\ldots,y_n\}$.

The proposed modeling procedure consists of the following steps:

- select maximal order p of AR models,
- select the threshold set $\boldsymbol{c}=(c_0,c_1,\ldots,c_r)$, $-\infty<c_0<c_1<\ldots<c_r<\infty$,
- for the set $\boldsymbol{c}$ divide the time series to groups according to threshold values and estimate sets of autoregressive parametres $\Phi_1,\ldots,\Phi_{\mathbf{r}}$ and standard deviations $\sigma_1,\ldots,\sigma_r$,
- fit linear AR(p_i) models, i=1,..,r,
- select delay parameter d (for example by using Akaike information criterion),
- in the case of switching by means of the aggregation function value select parameter k,
- fit multiregime SETAR model,
- evaluate model via diagnostic tests and forecasts,
- refine the model if necessary.

Application SETAR models in the river discharge modeling

We propose monthly average discharge analysis of five Slovak rivers in the basin Hron.

All computing were made in system Mathematica 7.

In such hydrologic time series modeling it is usual to proceed by specifying the systematic function which is the sum of an appropriate trend and cyclical functions.

1.	***Čierny Hron***	**2.283+0.0016 t– 1.089 Cos[(πt)/6]+ 1.076 Sin[(πt)/6]– 1.059 Sin[(πt)/3]**
2.	***Hron BB***	**22.090+ 0.0061 t- 8.369 Cos[(πt)/6]+ 9.545 Sin[(πt)/6]- 9.139 Sin[(πt)/3]**
3.	***Hron Brehy***	**42.650– 0.0193 t+ 4.218 Cos[(7πt)/45]– 10.272 Cos[(πt)/6]+ 8.768 Sin[(7πt)/45]+ 22.291 Sin[(πt)/6] –14.305 Sin[(πt)/3]**
4.	***Štiavnický p.***	**1.051– 0.0013 t– 0.476 Cos[(πt)/6]+ 0.326 Sin[(πt)/6]–0.536 Sin[(πt)/3]**
5.	***Vajskovský p.***	**1.165+ 0.0006 t– 0.627 Cos[(πt)/6]+ 0.209 Cos[(πt)/3]+ 0.124 Cos[(πt)/2]+ 0.357 Sin[(πt)/6]-0.648 Sin[(πt)/3]+0.229 Sin[(πt)/2]**

Table 1 Systematic functions of the monthly average discharge

The data from 1981 to 2000 have been divided to in-of-sample part (1981 to 1995) and out-of-sample part (1996 to 2000).

No.	River	Monthly average discharge characteristics									
		In-of-sample (1981-1995)				Out-of sample (1996-2000)		Systematic component		Systematic component residuals	
		min [m^3/s]	max [m^3/s]	μ_{is}	σ_{is}	μ_{os}	σ_{os}	μ_s	σ_s	μ_r	σ_r
1.	***Čierny Hron***	0.6	12.7	2.43	2.15	2.42	2.09	2.43	1.32	0.	1.69
2.	***Hron BB***	7.3	94.8	22.64	15.59	23.22	16.80	22.64	11.09	0.	10.96
3.	***Hron Brehy***	10.9	171.1	40.91	30.49	40.18	30.25	40.91	21.36	0.	21.80
4.	***Štiavnický p.***	0.2	3.9	0.94	0.77	0.90	0.78	0.94	0.56	0.	0.53
5.	***Vajskovský p.***	0.3	4.6	1.22	0.90	1.29	0.95	1.22	0.73	0.	0.53

Table 2 Statistic characterics of the monthly average discharge and systematic function values

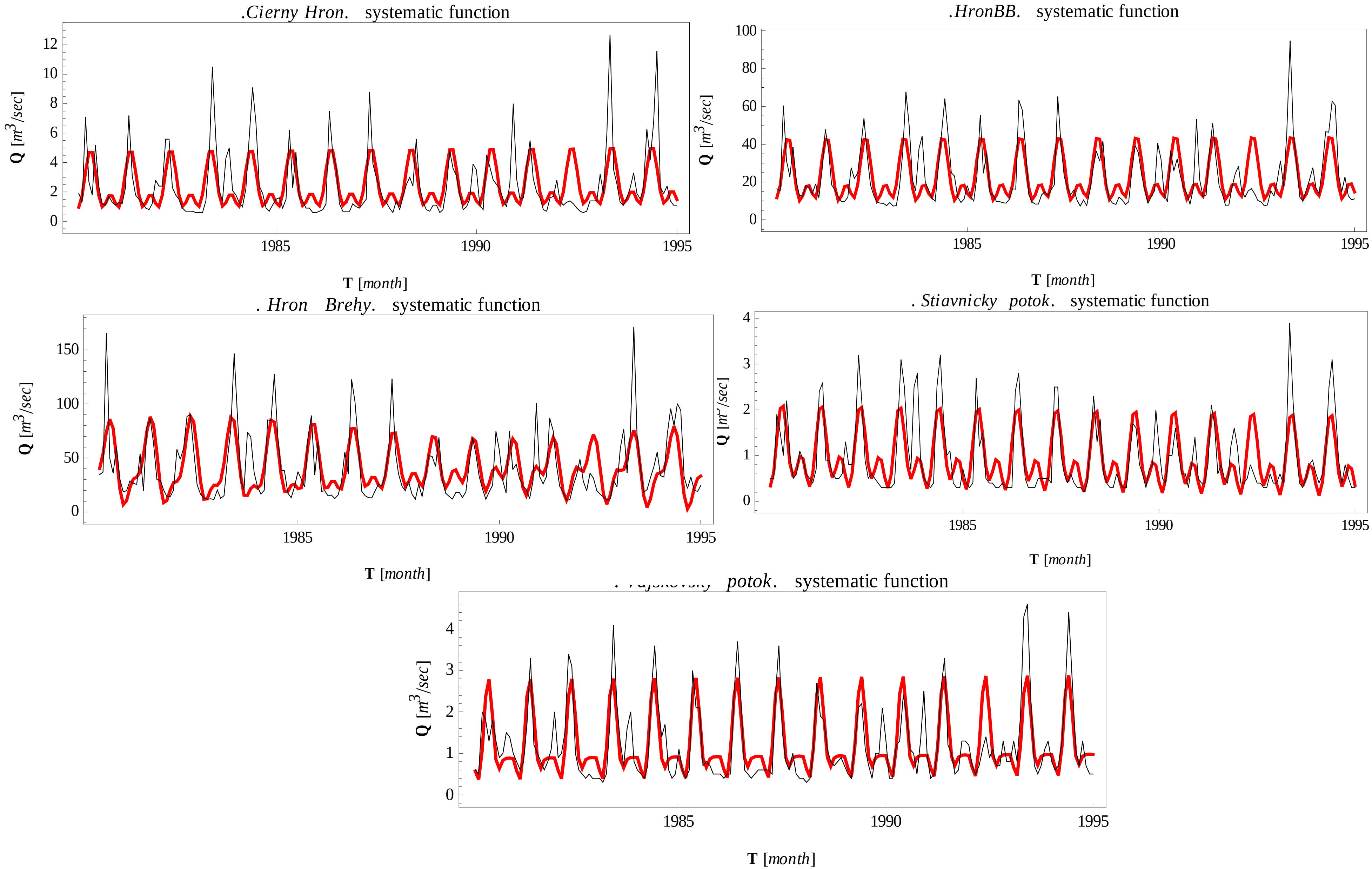

Figure 1 Monthly average discharge time plot (black color) and the systematic function graph (red color), period 1981–1995

Instead of the systematic function we have simulated observed time series by

1. the single regime linear AR(p) model,
2. two regime SETAR($c;p_1,p_2,d$) model with switching by means of the lagged value y_{t-d} ,
3. two regime SETAR($c;p_1,p_2,d,k$) model with switching by means of the aggregation function value $F(y_{t-d},\ldots,y_{t-d+k})$.

Čierny Hron **estimated models :**

AR(4):

$$\hat{y}_t = 1.763 + 0.515\, y_{t-1} - 0.158\, y_{t-2} + 0.0435\, y_{t-3} - 0.130\, y_{t-4} + \varepsilon_t$$

SETAR(4.3;4,4,12):

$$\hat{y}_t = \begin{cases} 1.136 + 0.269\, y_{t-1} + 0.002\, y_{t-2} - 0.017\, y_{t-3} - 0.008\, y_{t-4} + \varepsilon_t^1, & \text{if } y_{t-12} \le 4.3,\ \sigma_1 = 0.34, \\ 6.384 + 0.287\, y_{t-1} - 0.339\, y_{t-2} + 0.406\, y_{t-3} - 0.463\, y_{t-4} + \varepsilon_t^2, & \text{if } y_{t-12} > 4.3,\ \sigma_2 = 0.94, \end{cases}$$

SETARF(4.3;4,4,13,1):

$$\hat{y}_t = \begin{cases} 1.136 + 0.269\, y_{t-1} + 0.002\, y_{t-2} - 0.017\, y_{t-3} - 0.008\, y_{t-4}, & \text{if } \mathrm{Max}(y_{t-13}, y_{t-12}) \le 4.3,\ \sigma_1 = 0.34, \\ 6.384 + 0.287\, y_{t-1} - 0.339\, y_{t-2} + 0.406\, y_{t-3} - 0.463\, y_{t-4}, & \text{if } \mathrm{Max}(y_{t-13}, y_{t-12}) > 4.3,\ \sigma_2 = 0.94. \end{cases}$$

N.	River	Single regime AR(p), p=4			Two regimes SETAR($c;p_1,p_2,d$) the lagged value switching							Two regime SETARF($c;p_1,p_2,d,k$) the function val. switching				
		μ	σ	σ_r	c	p_1	p_2	d	μ	σ	σ_r	F	k	μ	σ	σ_r
1.	***Čierny Hron***	2.42	1.08	2.47	4.30	4	4	12	2.21	1.75	2.22	max	1	2.48	1.88	2.25
2.	***Hron BB***	22.58	8.09	17.56	32.80	4	4	12	21.66	12.44	15.52	max	1	24.34	12.75	14.14
3.	***Hron Brehy***	40.07	27.61	31.98	64.84	4	3	12	40.07	27.61	29.53	mean	2	38.04	24.74	29.53
4.	***Štiavnický p.***	0.93	0.38	0.87	1.47	4	4	12	0.93	0.75	0.73	mean	1	0.88	0.75	0.78
5.	***Vajskovský p.***	1.21	0.40	0.95	1.62	4	3	12	1.12	0.68	0.73	max	1	1.25	0.76	0.91

Table 3 Statistic characteristics of river discharge models, period 1981–1995

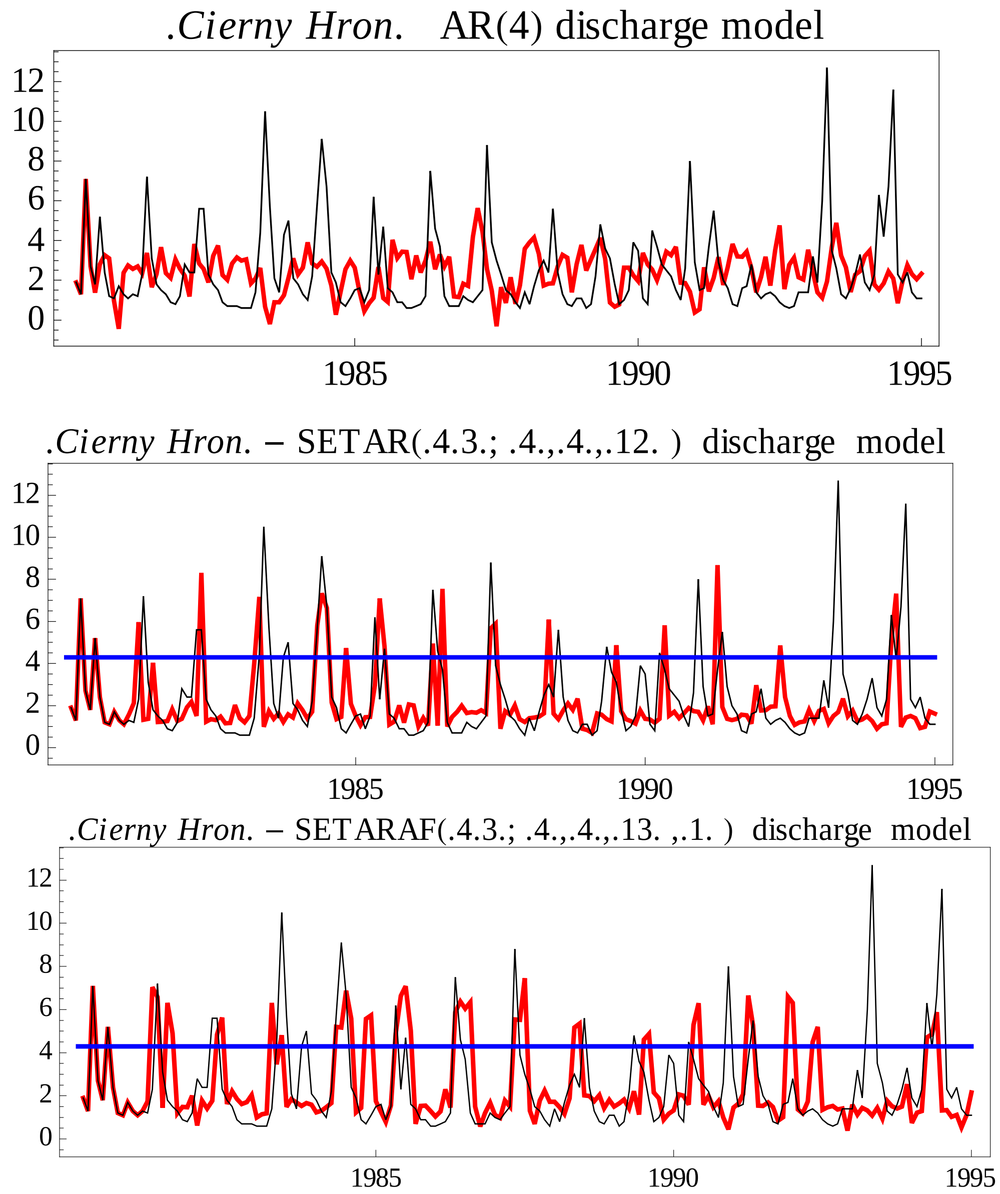

Figure 2 Time plots of Cierny Hron monthly average discharge models (threshold value marked by blue line), period 1981–1995

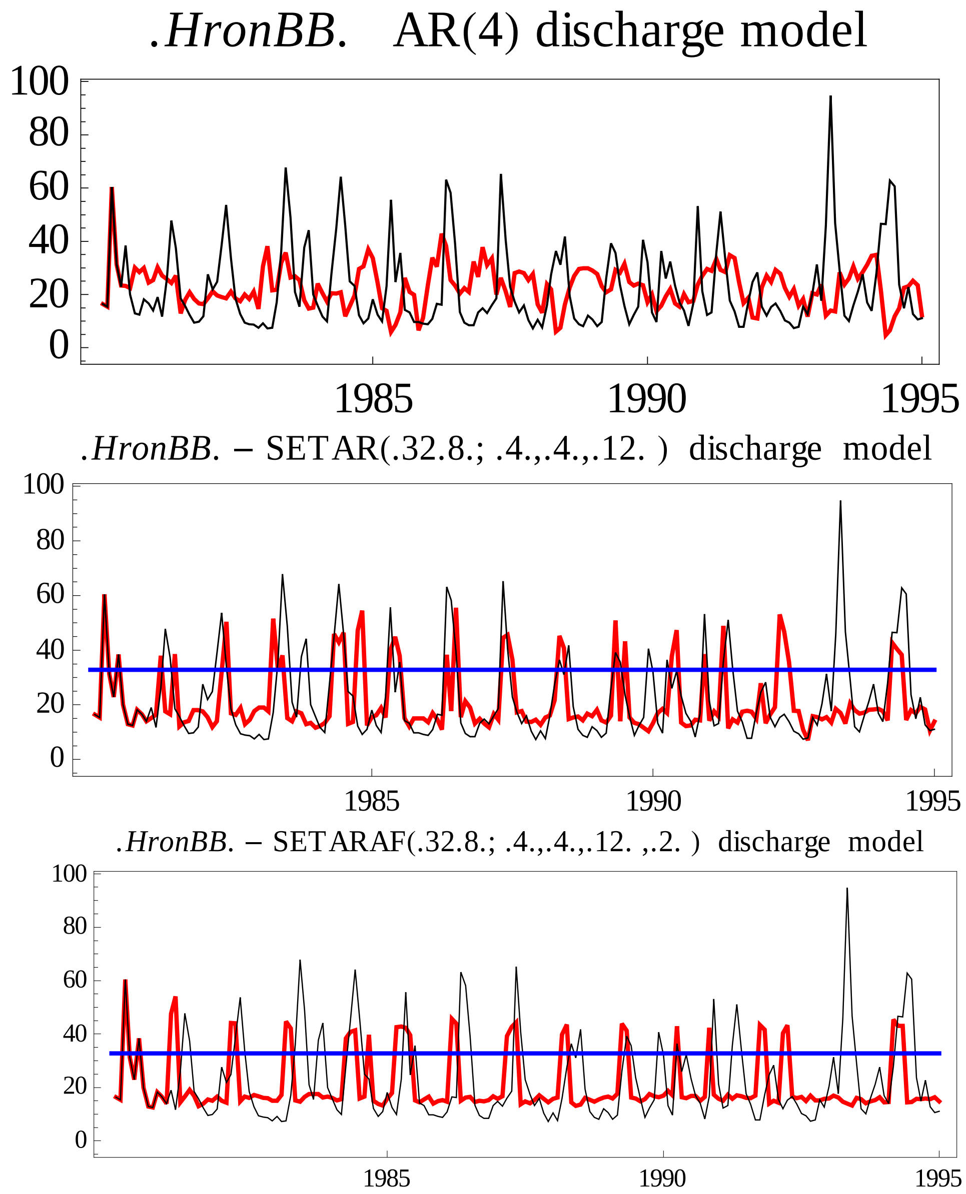

Figure 3 Time plots of Hron, Banska Bystrica monthly average discharge models, period 1981–1995

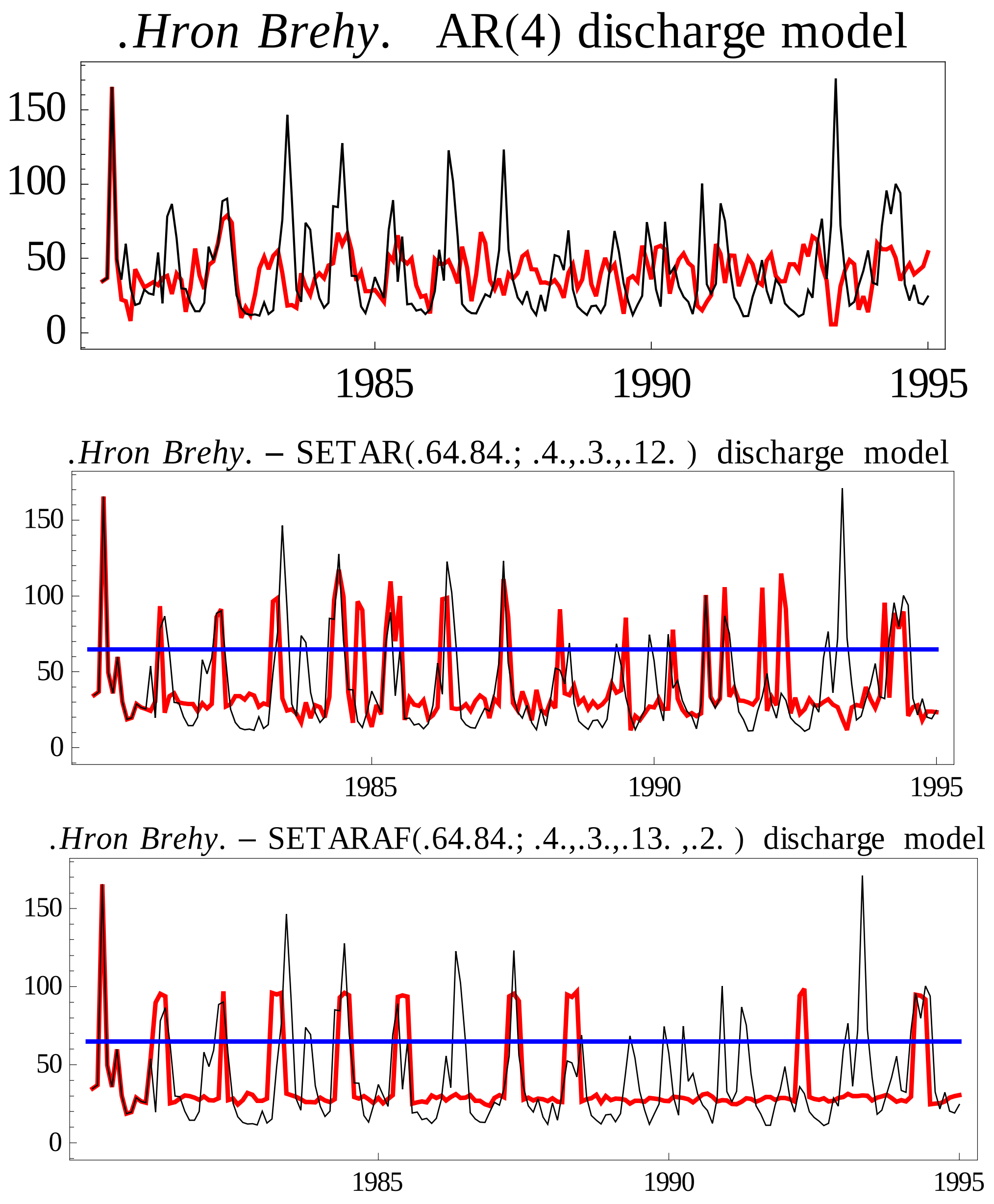

Figure 4 Time plots of Hron, Brehy, monthly average discharge models, period 1981–1995

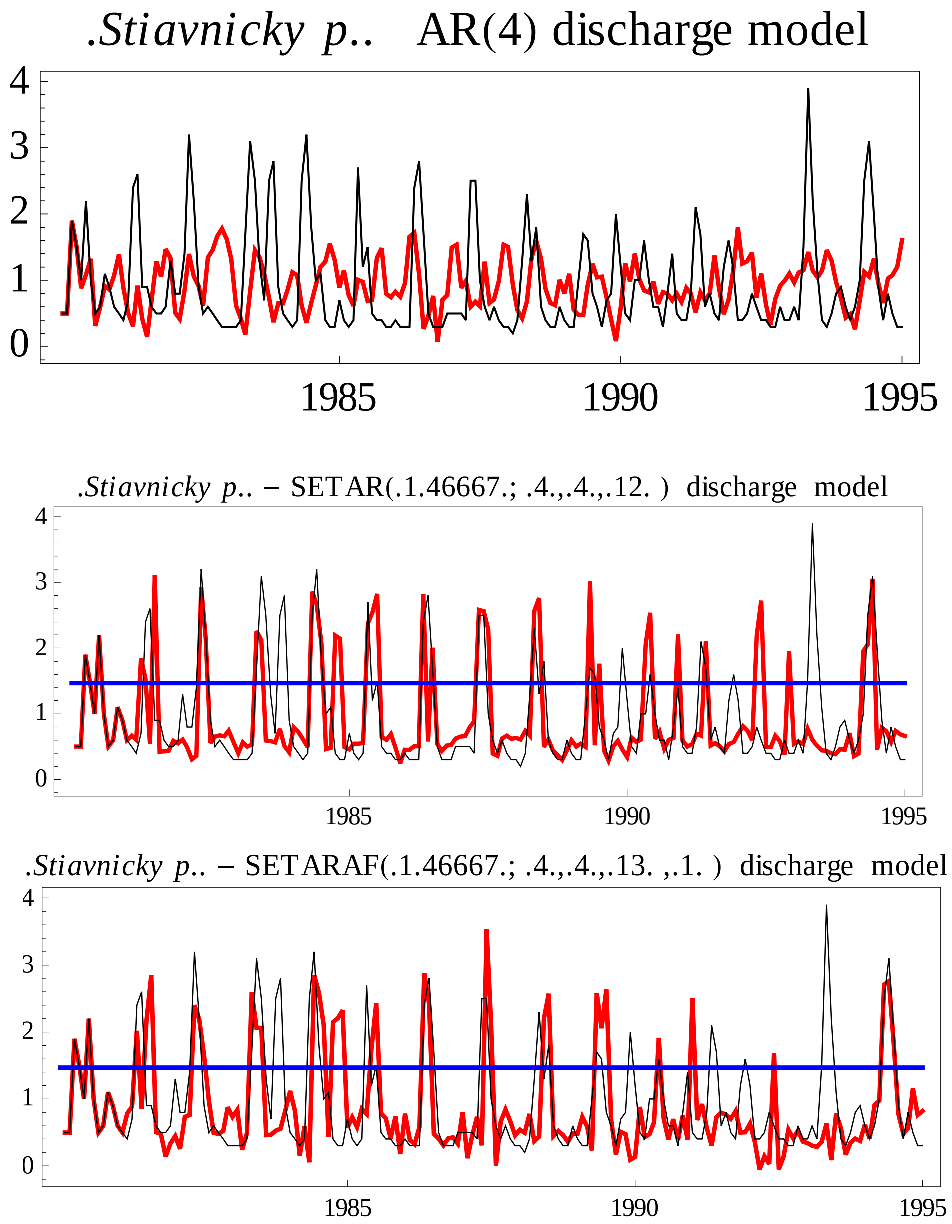

Figure 5 Time plots of Stiavnicky potok monthly average discharge models, period 1981–1995

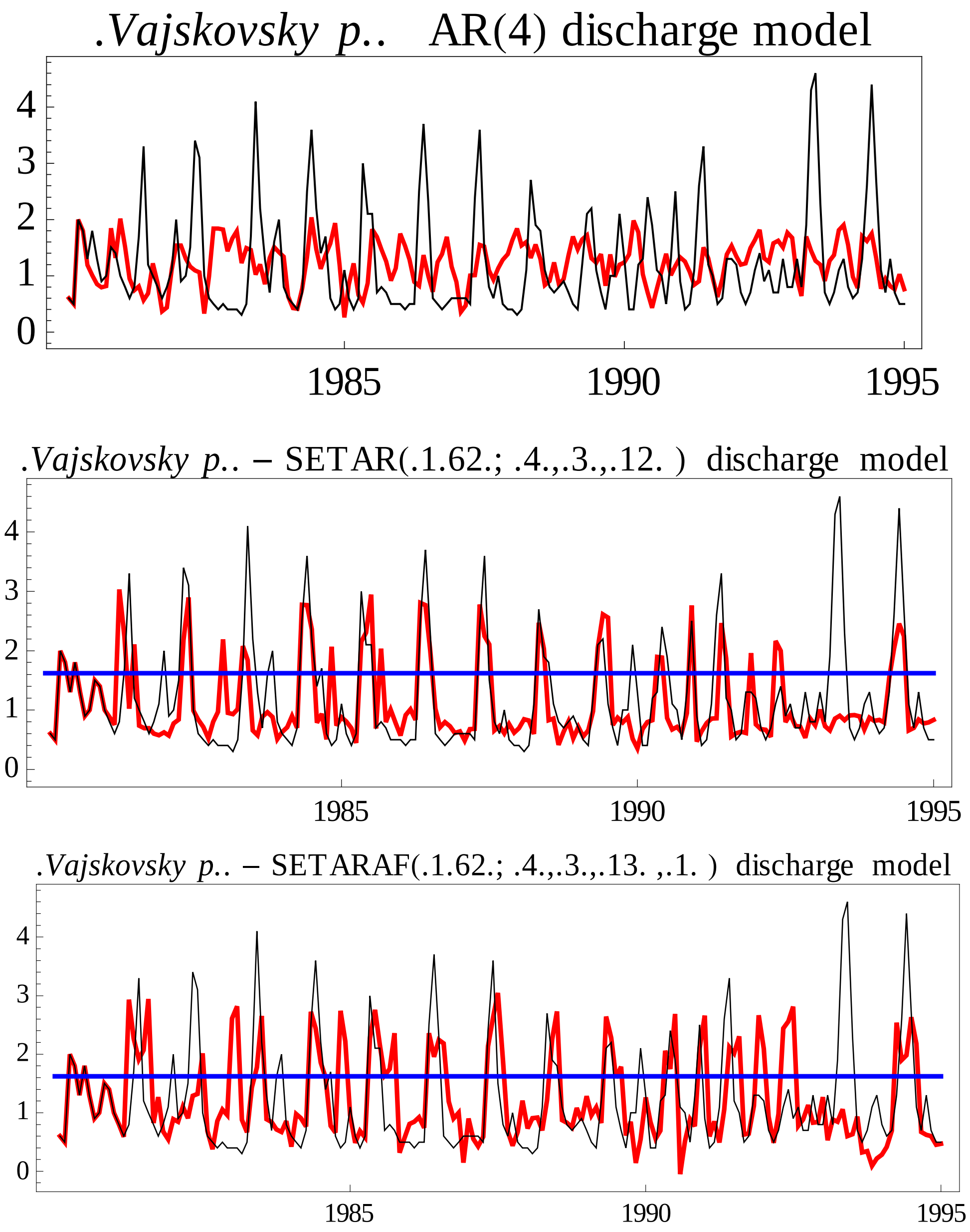

Figure 6 Time plots of Vajskovsky potok monthly average discharge models, period 1981–1995

Forecasting with alternative models

To evaluate the adequancy of fitted models we have computed the point forecasts by all alternative models and compare forecasts with out-of-sample.

We focuse on the usefulness of two-regime SETAR models for out-of-sample forecasting relative to single regime linear AR models and systematic function values.

N.	River	Out-of-sample forecasts with							
		Systematic function		AR(p)		SETAR($c;p_1,p_2,d$)		SETARF($c;p_1,p_2,d,k$)	
		μ	σ	μ	σ	μ	σ	μ	σ
1.	***Čierny Hron***	2.62	1.33	2.42	1.02	2.21	1.75	2.48	1.88
2.	***Hron BB***	23.38	11.15	22.43	7.64	22.72	12.62	24.00	13.86
3.	***Hron Brehy***	38.80	25.64	40.67	11.92	40.62	27.61	38.86	24.56
4.	***Štiavnický p.***	0.79	0.56	0.93	0.38	1.02	0.84	1.03	0.83
5.	***Vajskovský p.***	1.29	0.73	1.22	0.38	1.16	0.69	1.33	0.76

Table 4 Statistic characteristics of out-of-sample forecasts with alternative models, period 1996–2000

N.	River	Statistic forecasts errors							
		Systematic function		AR(p)		SETAR($c;p_1,p_2,d$)		SETARF($c;p_1,p_2,d,k$)	
		RMSPE	MAPE	RMSPE	MAPE	RMSPE	MAPE	RMSPE	MAPE
1.	***Čierny Hron***	2.79	0.73	2.01	0.87	2.94	0.66	2.88	0.64
2.	***Hron BB***	27.78	0.54	27.82	0.75	27.80	0.54	27.75	0.58
3.	***Hron Brehy***	49.38	0.62	49.36	0.79	49.19	0.87	49.44	0.82
4.	***Štiavnický p.***	1.21	0.57	2.01	1.07	1.64	0.73	1.46	0.69
5.	***Vajskovský p.***	1.39	0.50	2.00	0.80	1.39	0.51	1.49	0.57

Table 5 RMSPE and MAPE errors of out-of-sample forecasts with alternative models, period 1996–2000

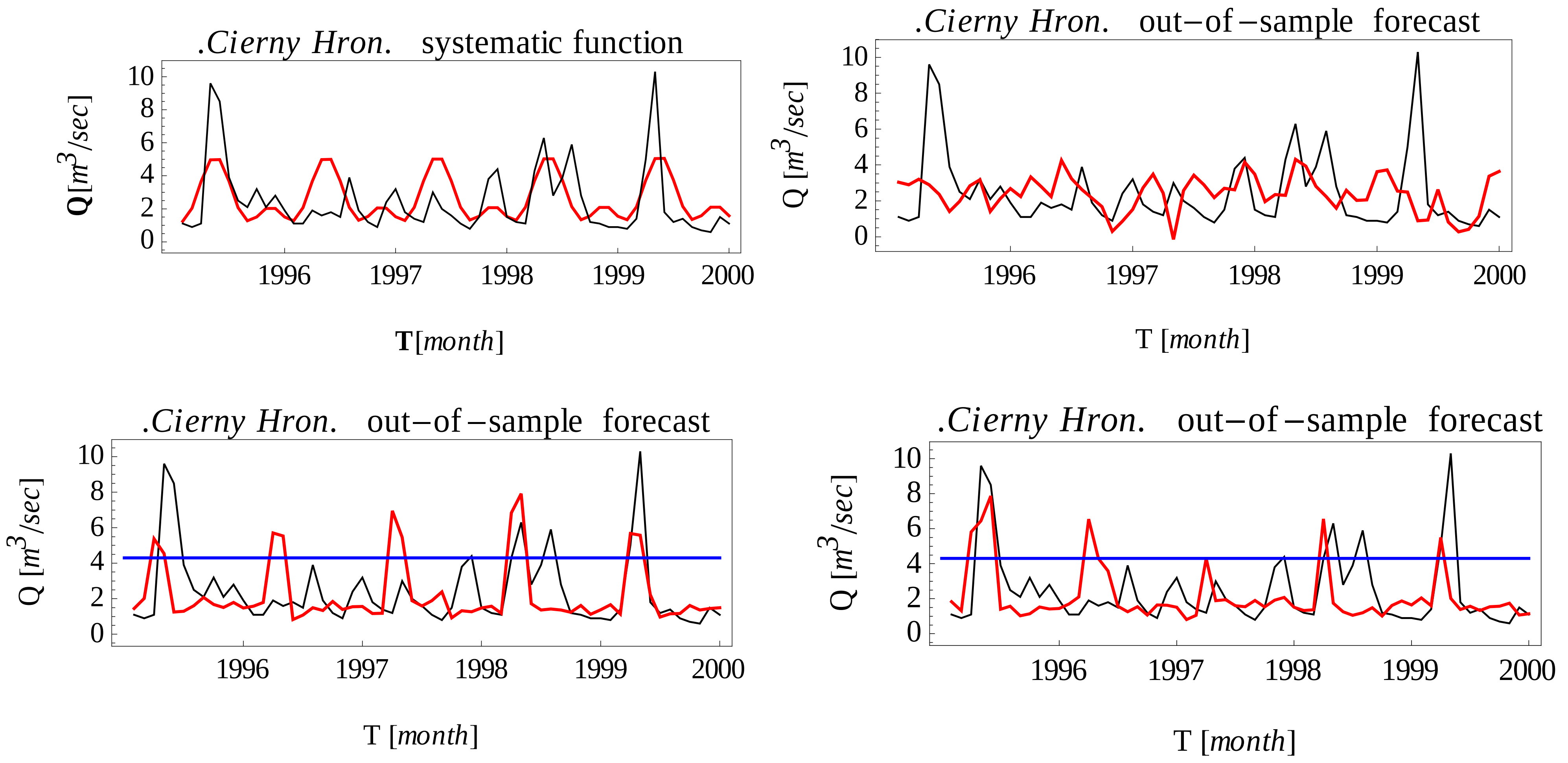

Figure 7 Forecasts time plot with alternative discharge models, Cierny Hron, period 1996-2000

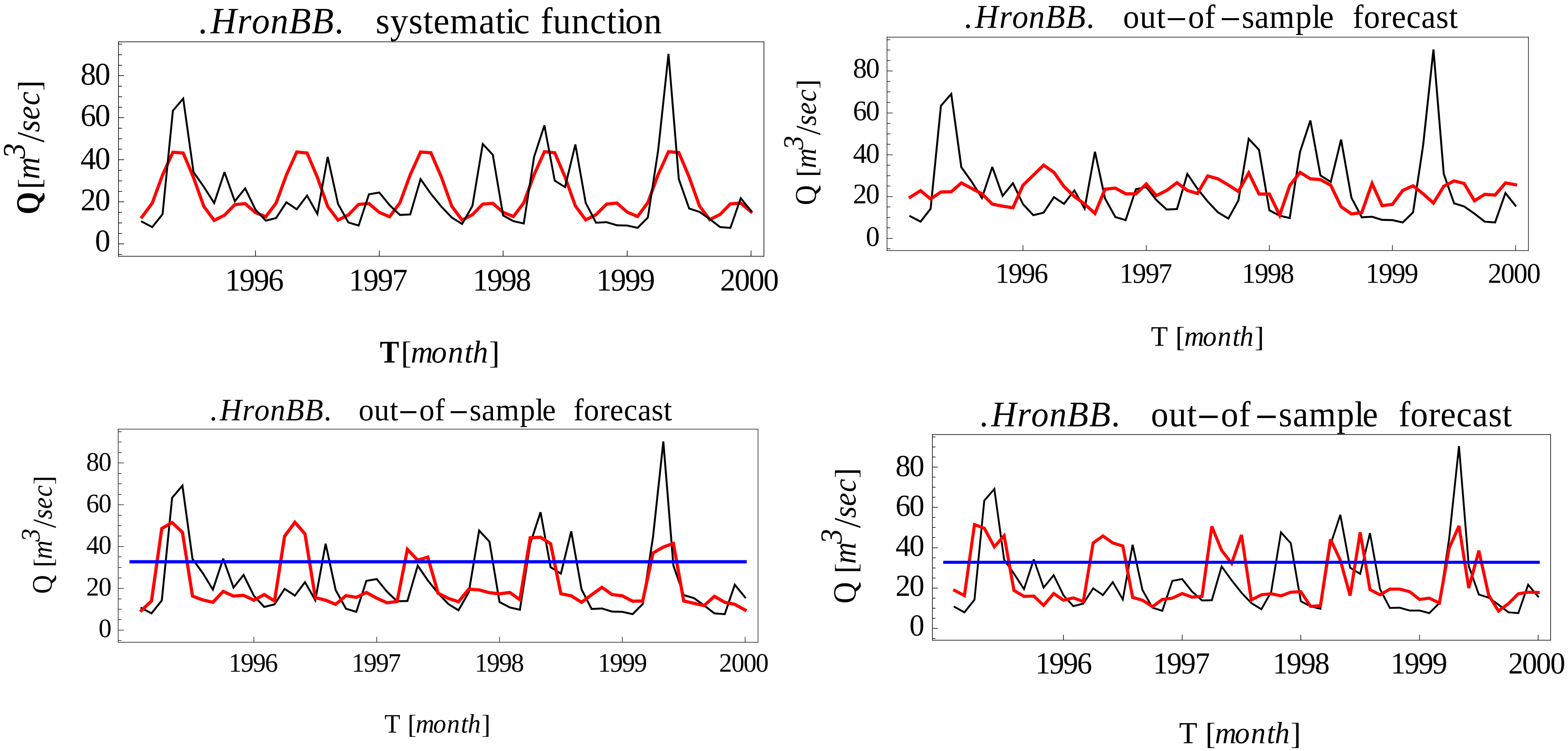

Figure 8 Forecasts time plot with alternative discharge models, Hron, Banska Bystrica, period 1996-2000

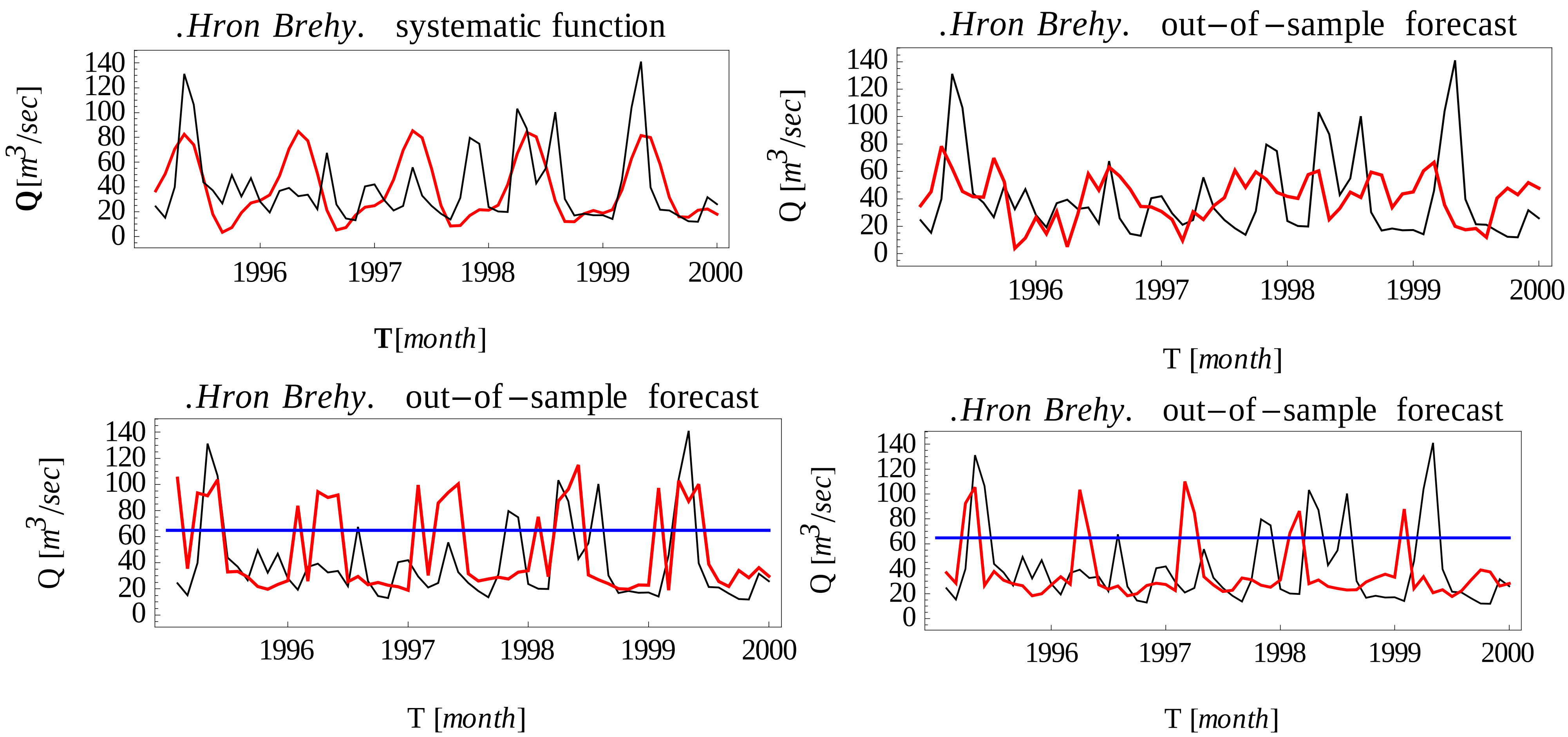

Figure 9 Forecasts time plot with alternative discharge models, Hron, Brehy, period 1996-2000

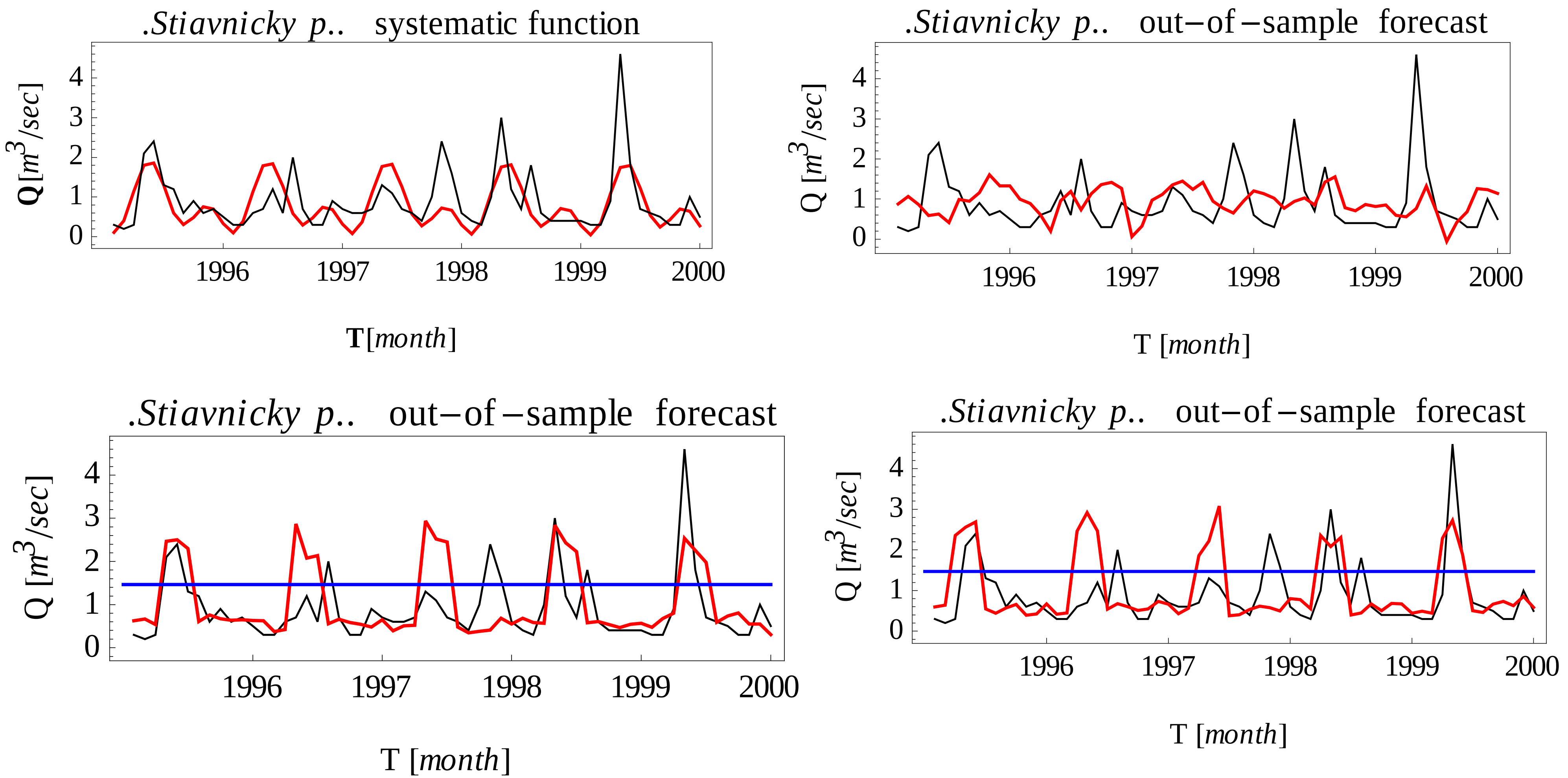

Figure 10 Forecasts time plot with alternative discharge models, Stiavnicky potok, period 1996-2000

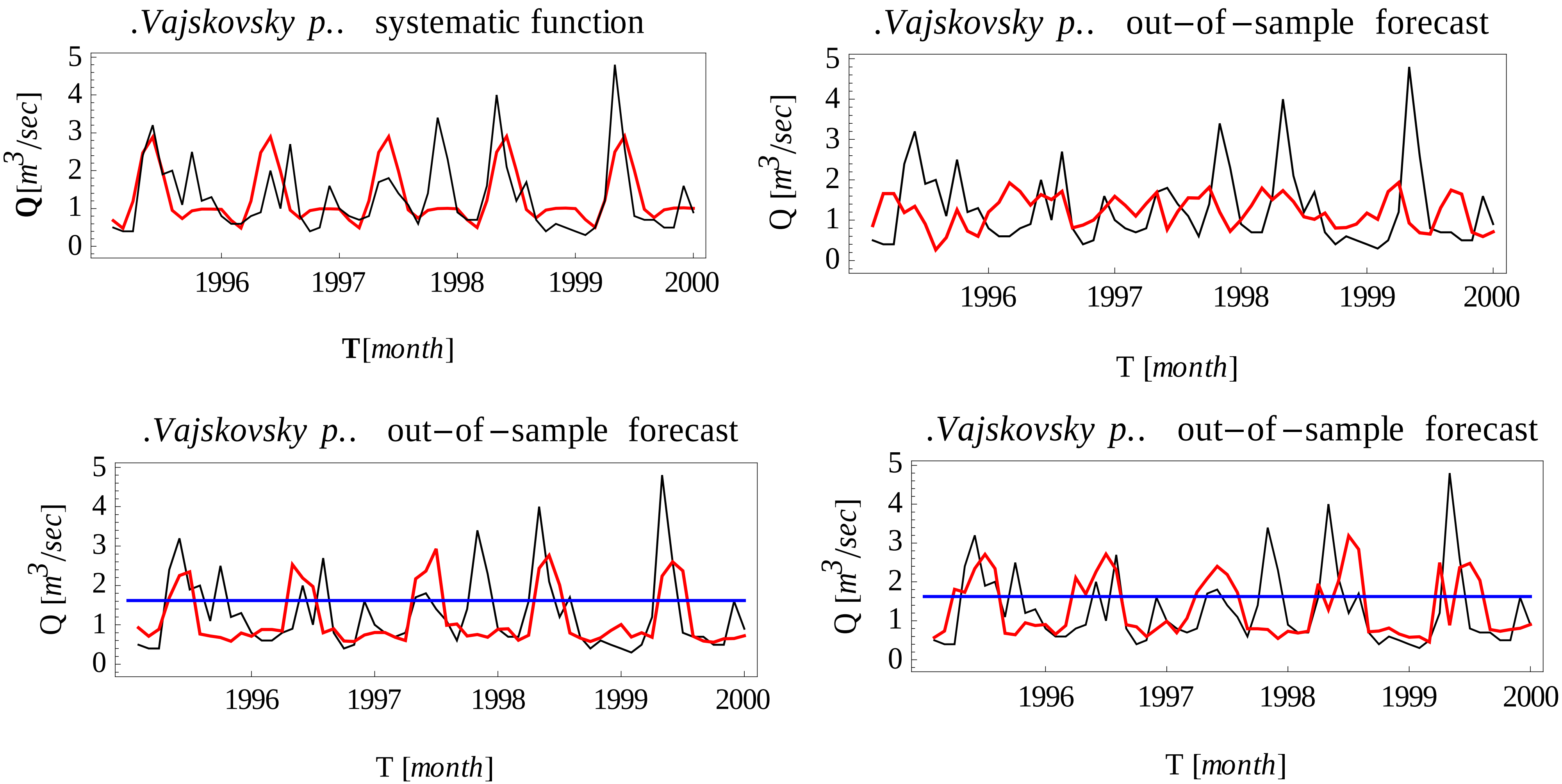

Figure 11 Forecasts time plot with alternative discharge models, Vajskovsky potok, period 1996-2000

Conclusion

We study two regime nonlinear SETAR models of univariate time series and try to improve the regime switching thrue aggregation function value with more lagged values in input.

We were used the monthly average streamflow time series to illustrate the modeling procedure and the forecasts estimation.

We were compared results of two regime SETAR modeling with single regime linear AR models and systematic function values.

We can summarize that in the case of monthly average discharge time series results of two regime SETAR modeling are comparative with the systematic function values for descriptive and forecast purposes.

We can suppose that next study and experiments with multiregime nonlinear models not only SETAR type but also another types as STAR (smooth transition autoregressive) and MSW (Markow switching) will be succesfull and contribute to better simulation of hydrologic data.

Thank you for your attention