STOCHASTIC MODELLING OF THE NATURAL RIVER FLOWS

Elena Bojilova

National Institute of Meteorology and Hydrology – Bulgarian Academy of Sciences, Bulgaria, Sofia, Bulgaria, e-mail: Elena.Bojilova@meteo.bg

Abstract: In design and operation of water resource systems the stochastic modelling plays a very important role. In the research disaggregation models are selected and applied. The objects of disaggregation are natural river flows for chosen Bulgarian rivers. As a result synthetic flow sequences are generated.
Disaggregation models are basically divided into three main groups: temporal, spatial and temporal-spatial. The focus of this paper is application of temporal disaggregation models for disaggregation of annual flow totals to seasonal or monthly flows. Two basic models are applied: the Original model of Mejia and Rousselle and the Corrected extended Lin model, one-stage disaggregation. Data from five rivers in different areas of Bulgaria for the aims of the study are selected and executed.
Our conclusion is that both stochastic models are suitable for disaggregation of the natural river flows with snowmelt conditions together with rainfall in the catchment areas. They are applicable for the specific climatic and hydrological conditions of Bulgaria. Furthermore, the best result is obtained from the Corrected extended Lin model, one-stage disaggregation for single-site approach.
Keywords: stochastic and disaggregation modelling, natural river flow.

STOCHASTISCHE MODELLIERUNG DES NATÜRLICHEN ABFLUSSES

Zusammenfassung: Die stochastische Modellierung spielt in der Auslegung und im Betrieb von Wasserbewirtschaftungssystemen eine sehr bedeutende Rolle. In der Forschung werden Disaggregationsmodelle ausgewählt und angewendet. Die Objekte der Disaggregation sind die natürlichen Abflüsse für ausgewählte Flüsse in Bulgarien. Als Ergebnis werden synthetische Abflussreihen erzeugt.

Disaggregationmodelle werden grundsätzlich in drei Hauptgruppen unterteilt: zeitliche, räumliche und räumlich-zeitliche. Der Schwerpunkt dieses Beitrages liegt auf der Anwendung von zeitlichen Disaggregationmodellen zur Disaggregation von jährlichen Gesamtabflüssen in jahreszeitliche und monatliche Abflüsse. Zwei einfache Modelle finden Anwendung: das Originalmodell von Mejia und Rousselle und das korrigierte erweiterte Lin-Modell mit einstufiger Disaggregation. Daten von fünf Flüssen aus verschiedenen Teilen Bulgariens wurden für die Ziele der Studie ausgewählt und durchgerechnet.

Unsere Schlussfolgerung ist, dass beide stochastischen Modelle verwendbar sind für die Disaggregation des natürlichen Abflusses unter Schneeschmelzbedingungen verbunden mit Regen in den Einzugsgebieten. Sie sind für die spezifischen klimatischen und hydrologischen Bedingungen in Bulgarien verwendbar. Weiter wurden die besten Ergebnisse bei Verwendung des korrigierten erweiterten Lin-Modells mit einstufiger Disaggregation für Einzelmessstellenansatz erreicht
Stichworte: stochastische - und Disaggregationmodellierung, natürlicher Abfluss.

1. Introduction

In design and operation of water resource systems the stochastic modelling plays a very important role. In this study disaggregation models from the family of stochastic models are selected and applied to natural river flow. As a result river discharge synthetic sequences are generated. Disaggregation models are basically divided into three main groups: temporal, spatial and temporal-spatial.

The focus of this paper is application of temporal disaggregation models for the specific climatic and hydrological conditions in Bulgaria. The applicability of disaggregation models for natural river flow is studied. Two disaggregation models in single-site disaggregation schemes are selected and executed. The selected stations with natural river flow can be considered as examples of drainage basins with relatively large areas (from 5 000 km^2 to 20 000 km^2) and low mean altitude of the river basin (between 250 *m* and 600 *m* above sea level).

2. Additional remarks

According to their mathematical expressions in Salas, disaggregation models can be classified into two types: basic and extended (1985). The basic model proposed by Valencia and Schaake (1973) provided for the disaggregation of annual flow totals into seasonal flows. Mejia and Rousselle (1976) added a term to the Valencia and Schaake model to created the so-called "extended model" (M-R model), and provided the link with the past at seasonal level. In the M-R model an additional term is included to preserve the seasonal covariances between the first season of current year and the last season of the past year (Hakem 1991).

Disaggregation modelling is a process by which time series are generated from a time series already available. If the set of independent series is available, the corresponding shorter–interval series can be obtained. Typically, the independent time series has been previously generated and after these series are disaggregated into the sub-series. The independent input time series or so-called "key" series can be different: annual, seasonal or monthly flows. Disaggregation can be done in one stage or in several stages. For example, annual to semi-annual, semi-annual to quarterly, and quarterly to monthly flows - a three-stage disaggregation approach (Todini 1980).

The summary of main temporal disaggregation models is presented (Table 1). The main advantage of any disaggregation model is the preservation of statistical properties at more that one time interval. Additionally, it is highly desirable to preserve certain statistics at more than one level, i.e. monthly, annual. These statistical properties considered as important ones are the first two moments, the probability distribution of the series and some covariances (or equivalently correlations). The second major advantage of disaggregation is that it is a technique, which allows a more flexible approach for generation of synthetic data (Lin 1990). In recent study (Koutsoyiannis 1999) the adjusting procedure has been developed and proposed in case of not full preservation of additivity property.

All disaggregation models may be reduces to a form, which is termed the linear dependence model. A typical application of these models would be to generate a series of monthly river flows ($Y(t)$) from a given series of annual flow totals ($X(t)$). In Table 2 the mathematical description of the used in the research models is presented. The two used models are: the Original Mejia and Rousselle model (Case 1) and Modified Mejia and Rousselle or so-call "extended" Lin model (Case 2). The models reviewed in this study are "single-site" temporal models (Salas et al., 1985).

Annual flow volumes, modelled by the *AR(1)* model, are generated and disaggregated to monthly values. For to overcome the difficulty originating from the non-normality of the historical data series, the annual and monthly totals are transformed and standardised before modelling. The three-parameter lognormal transformation and the square root transformation are applied and the normality of the transformed data examined. The model results are back transformed and their means, standard deviations and coefficient of skewness (or asymmetry) are compared to the corresponding descriptors of the historical data. The sums of the monthly-generated stream flow totals are compared to the observed annual flow totals for to check preservation of additivity. Plotting the original annual totals and the sum of disaggregated time series is carried out to check the additivity.

Table 1. Summary table of temporal disaggregation models (*Source: Bojilova 2003)

Authors & year	Model name	Index	Advantage of the model	Disadvantage of the model
Valencia and Schaake, 1973	basic model	V-S	- statistics at both annual and seasonal levels are preserved; -basic clean form.	- link with the past only at annual level; - large number of parameters.
Mejia and Rousselle, 1976	original extended model	M-R	- link with the past at seasonal level.	- the model does not preserve the statistics, which was desired to preserve.
Lane, 1979	condensed		- decreasing the number of parameters.	- it fails to preserve the additivity.
Hoshi and Burges, 1979		H-B	- seasonal correlations and statistical moments are preserved; - successfully maintains correlations between the season that join successive water year; -introduced scheme for 3-PLN Distribution.	- for disaggregation scheme of 3-PLN distribution - distortion in the additivity property occurs.
Todini, 1980	modification of V-S and M-R		- suggested scheme for preservation of the skewness.	
Stedinger and Vogel, 1984		S-V	- reproduce the covariance between current upper and lower level flows as well as the covariance of the lower level with themselves; - in addition, reproduce reasonable lag one covariance matrix of lower level flow vectors.	- in particular case, the lag one covariance matrix was both practically and statistically different from the true population values.
Stedinger-Pei-Cohn, 1985	modification of condensed	SPC	- preserve the additivity and correlation among the seasonal generated value.	- it does not preserve exactly additivity.
Grygier and Stedinger, 1988	combination of spatial and temporal scheme	G-S	- preserves the additivity and correlation among the seasonal generated flows.	- it preserves the at-site lag-1 correlations in each month, and not the lag-1 cross-correlations.
Lin, 1990	modification of M-R	Lin	- the corrected parameter estimate was proposed for two-stage disaggregation; - the parameter estimation equations are mathematically consistent; -the model can preserve the important moments and the additivity.	

Table 2. Mathematical description of the models

Model	Mathematical form of the models*
Mejia and Rousselle model (Case 1)	$Y(t) = AX(t) + BV(t) + CY(t-1)$
Lin model one-stage (Case 2)	$X(t) = GX(t-1) + HU(t)$ $Y(t) = A[GX(t-1) + HU(t)] + BV(t) + CY(t-1)$

* $Y(t)$ is a nm - dimensional zero-mean vector of normally distributed monthly flows;
m is a number of months in the year;
n is a number of sites, in the single-site scheme $n = 1$;
$X(t)$ is a n - dimensional zero-mean vector of normally distributed annual river flows;
t is an index corresponding to the year;
A and *B* are $nm \times N$, $nm \times nm$ coefficient matrices, respectively;
$V(t)$ is a $nm \times N$ - dimensional matrix of independent standard normal deviates;
$Y(t-1)$ is a nm - dimensional zero-mean vector of normally distributed monthly flows for the year $t-1$;
C is a additional parameter matrix $nm \times nm$;
$X(t-1)$ is a n - dimensional zero-mean vector of normally distributed annual flows for the year $t-1$;
$U(t)$ is a n - dimensional vector of independent standard normal deviates;
G and *H* are $n \times n$ parameter matrices.

The means and standard deviations were checked using the average percentage change (*APCh*) – equation (1) and root mean square (*RMS*) - equation (2)

$$APCh = mean\left(\frac{Generated - Original\,value}{Original\,value}\right)*100\% \tag{1}$$

$$RMS = \sqrt{\frac{\sum(Generated - Original\,value)^2}{Number\,of\,values}} \quad . \tag{2}$$

The preservation of skewness was checked according to the type of transformation using the average percentage change - equation (3)

$$APCh = mean\left(\frac{Generated - Original\,value}{Mean\,of\,original\,value}\right)*100\% \tag{3}$$

and the *RMS*. The preservation of the covariance matrices S_{xy}; S_{yy} and S_{yy1} is evaluated using the *APCh* for single-site disaggregation (Bojilova 1999 and 2003). The work has been done (Koutsoyiannis, 2001 and 1999) in direction to preserve the coefficient of skewness in the process of time series disaggregation. The preservation of the third order moments (skewness) is of importance in application of disaggregation models to Mediterranean ephemeral and intermittent rivers (Koutsoyiannis et al., 1996).

3. Application of disaggregation models

Five river stations located in different valley regions of Bulgaria were chosen. The mean altitude of the chosen stations varied between 250 m to 600 m above sea level. The selected stations represent different climatic and hydrologic conditions of the country. They are included in the National Hydrological Network, see Table 3 for more comprehensive summary. Database for this study was received from the project - General master plans for water resources management (2000). Time series with natural river discharge representing the three main hydrological zones of Bulgaria were executed (Bojilova 2003). The selected stations with natural flow are characterized with drainage area in the range between 5 000 km^2 to 20 000 km^2. The analyses in the research were made for the 1961-1998 period.

Table 3. Information for the chosen river stations (*Source: “General master plans for water resources management”, 2000)

River	Number of the station	Location	Beginning of observation	Area	Average mean altitude
			year	*(km^2)*	*(m)*
Tundja River	74850*	Elxovo town	1949	5551	475
Kamchia River	43800	Grosdevo village	1938	4857	352
Cherni Lom River	31550	Shirokovo village	1948	1383	287
Maritza River	73750	Harmanly town	1936	19 693	603
Provadiiska River	42850	Sindel town	1922	1856	220

*For the project, data are prepared as follow:
74850 Str.G. Gerassimov and E.K. Bojilova;
43800, 31550 and 42850 Str.G. Gerassimov and N. Todorova;
73750 Str.G. Gerassimov and K. Krumova;
All mentioned authors are from Department of Hydrology, NIMH-BAS.

For application of disaggregation models five river stations were selected. The extended Mejia and Rousselle model and the Lin model (one-stage disaggregation approach) were applied using ten realisation of each model. The historical flow discharge data for the selected stations were used. The length of the data series for disaggregation aims was thirty-one years. Two flow sequences – annual and monthly – were organised first. For discussed models the annual flow totals $X(t)$ were disaggregated into monthly flows $Y(t)$ of twelve months. Single-site disaggregation was executed.

The following preliminary conclusions may be drawn:
- For all models used, most of disaggregated monthly time series before back transformation had means close to zero and standard deviations close to unit.
- For the two models, the means and standard deviations of the disaggregated time series after back transformation were close to the corresponding historical values.

In Figure 1 and Figure 2 the preservation of the first and second order statistical moments (mean and standard deviation) for the Cherni Lom River is presented.

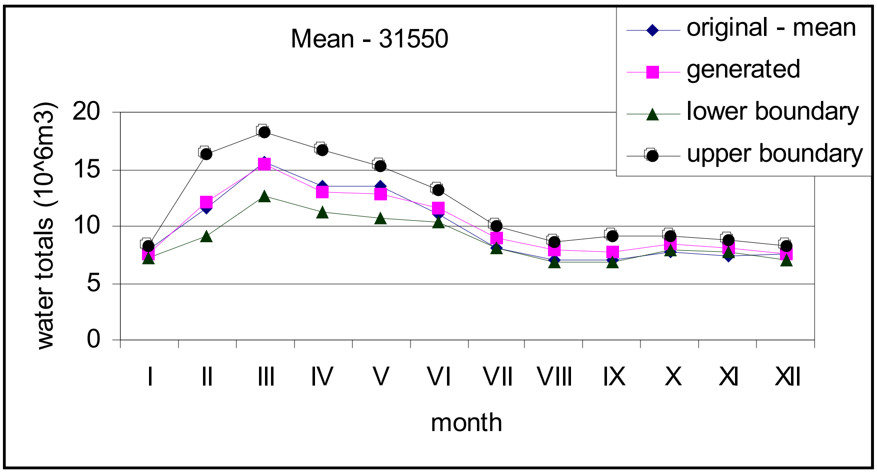

Figure 1. Preservation of mean, Cherni Lom River, 1961-1990

Goodness-of-fit of the different disaggregation models to the water year totals - additivity was tested. In Figure 3 the preservation of additivity is shown for Tundja River near Elxovo city. The results presented are the averages from ten realisations. In the figure the original values for the study period, lower and upper boundaries of generated flow values from ten runs are presented.

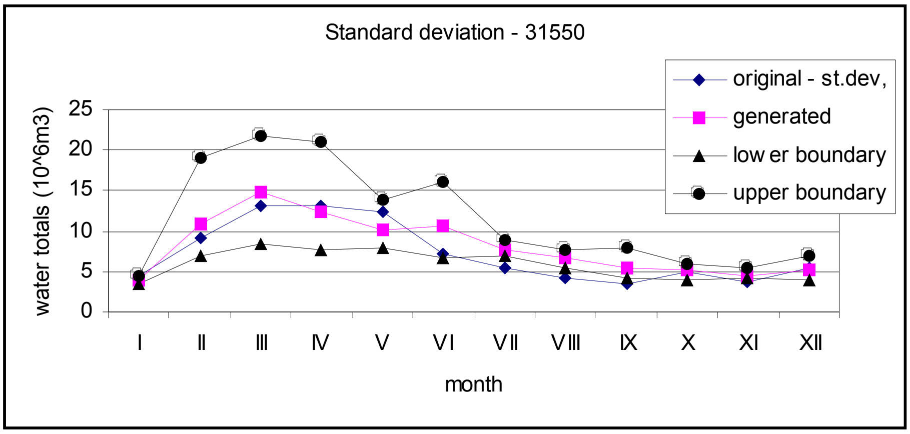

Figure 2. Preservation of standard deviation, Cherni Lom River, 1961-1990

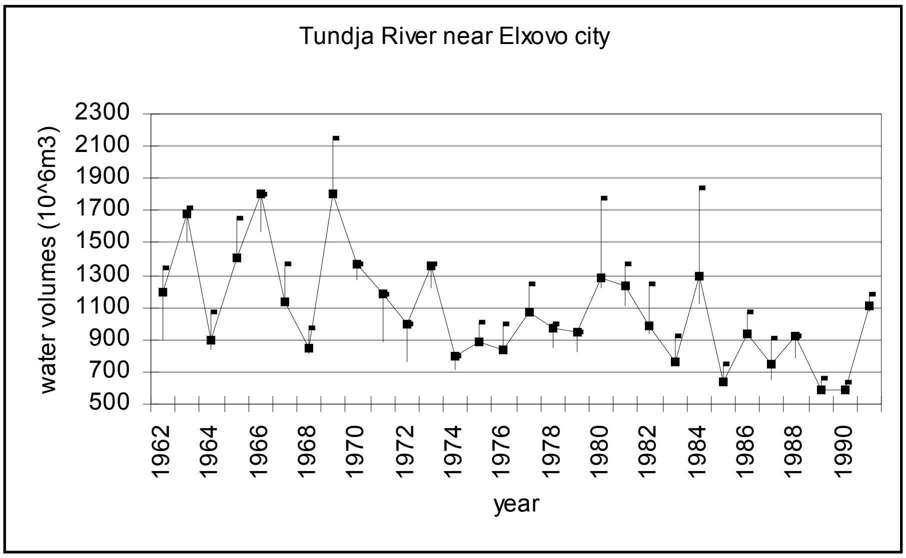

Figure 3. Additivity, Case 2, Tundja River, 1961-1990

4. Comparison between the two approaches

4.1. Case 1 - The original Mejia and Rousselle model

The original M-R model preserves the first two statistical moments and the additivity quite well, but fails to preserve the correlation matrices, possibly due to unreliable parameter estimation.

4.2. Case 2 - The corrected extended Lin model, one-stage disaggregation

Table 4 presents the results from one-stage single-site disaggregation – Lin model, Case 2 (average results from 10 realization of the model) using *APCh* and *RMS*. The Lin model preserves the first two statistical moments very well. The additivity is preserved very well in the single-site approach for all executed stations. The correlation matrices are preserved very well.

Table 4. Results from one-stage single-site disaggregation – Lin model, Case 2 (average results from 10 realization of the model)

River station number		$\bar{x}$		σ_x		C_s	
		from	to	from	to	from	to
74850	*APCh, %*	0.65	8.14	4.06	12.70	-27.67	7.35
	RMS	5.24	14.59	17.23	36.84	1.13	1.55
43800	*APCh, %*	0.85	8.51	3.63	23.77	-21.80	5.48
	RMS	7.24	16.52	22.46	43.28	1.93	2.55
31550	*APCh, %*	2.35	5.02	17.62	21.40	8.87	38.04
	RMS	0.96	1.90	2.53	4.13	1.06	1.46
42850	*APCh, %*	2.17	4.23	9.08	18.57	-9.46	13.05
	RMS	1.01	1.49	1.85	4.21	0.97	1.09
73750	*APCh, %*	-2.14	11.67	6.92	34.38	-31.01	-5.39
	RMS	16	63.58	68.51	113.2	0.67	1.16

* $\bar{x}$ - mean; σ_x - standard deviation; C_s - coefficient of skewness.

5. Data generation

For the chosen five stations synthetic data generation for a new realisation of thirty years was carried out. The multi-variate lag-one auto-regressive model or so-called *AR(1)* model was applied for generation of the annual flows - the "*key*" series for disaggregation model. The 1961-1998 period is used in the process of generation. The annual series or $X(t)$ were generated using *AR(1)* model. With *AR(1)* model, sixty years of data for the chosen five stations were generated. The first generated thirty years, or what it is called the "warm-up" length, were neglected. The next thirty years were used to disaggregate the monthly river flows. The annual generated series, transformed and normalised, were the input to disaggregation models.

Generation was performed using Lin model, single-site approach, Case 2, Figures 4 and 5. The disaggregated series had zero mean and unit variance and were in normalised form, so that to obtain the generated monthly flows back transformations are needed. Using the results for *APCh* and the *RMS*, the preservation of mean, standard deviation and skweness can be accepted.

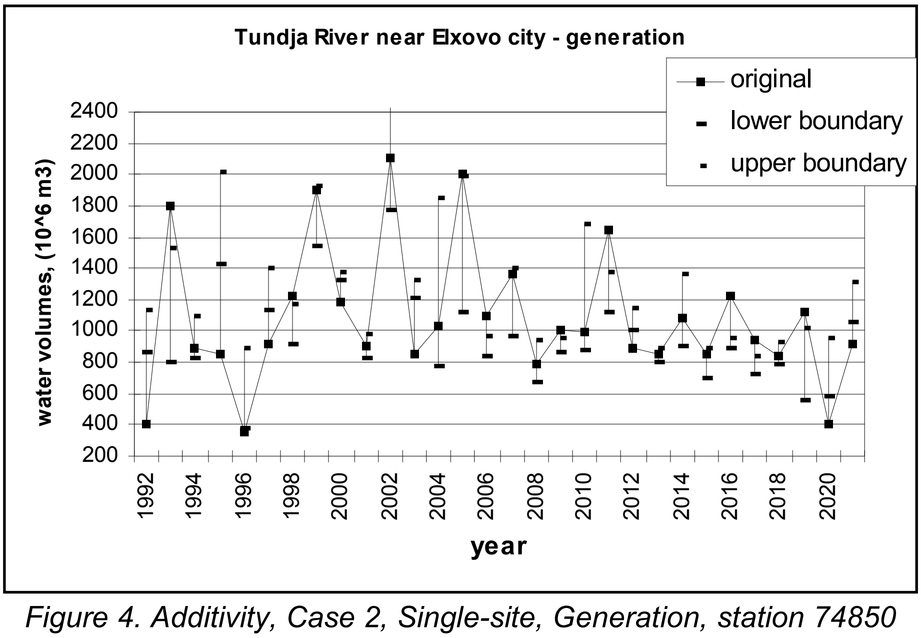

Figure 4. Additivity, Case 2, Single-site, Generation, station 74850

The disaggregated series had zero mean and unit variance and were in normalised form, so that to obtain the generated monthly flows back transformations are needed. Using the results for *APCh* and the *RMS*, the preservation of mean, standard deviation and skweness can be accepted.

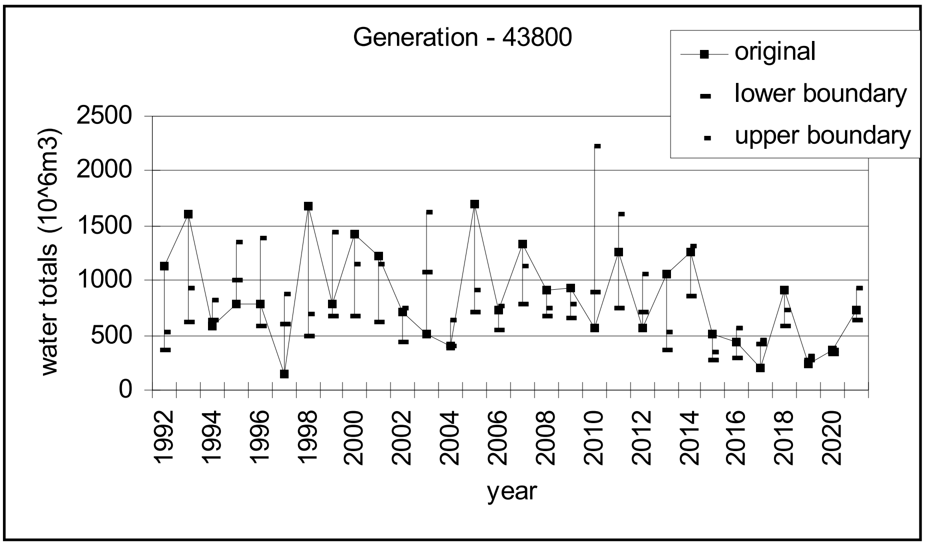

Figure 5. Additivity, Case 2, Single-site, Generation, Kamchia River

6. Conclusions

In the study, the applicability of the temporal disaggregation models was tested for selected stations with natural river flow in Bulgaria. Two models for a single-site disaggregation scheme were applied. The following conclusions can be drawn.

- The corrected parameter estimation in the Lin model succeeded in preservation of the additivity and the covariance matrices as expected. The corrected Lin model (Case 2) is an improvement compared to the extended Mejia and Rousselle model.
- The Lin model preserves the first two statistical moments and the covariance matrices rather well. However, the preservation of the second-order statistical moment (variance and covariance) is poorer than that of the first-order moment (mean), and the third-order (skewness) is even worse.
- Both models are suitable for disaggregation of the natural flows with snowmelt conditions together with rainfall in the catchment areas. They are applicable for the specific climatic and hydrological conditions of Bulgaria.
- The models are suitable for the specific conditions in the valley regions in the country for natural river flow with average mean altitude of the drainage basins between 200 *m* to 500 *m*. Furthermore, the best result was obtained from the corrected extended model, one-stage disaggregation single-site approach.

The further studies should be carried out to test applicability of the models to multi-site approach. Another interesting possibility for the future research is to apply the disaggregation models to the non-perennial and ephemeral rivers in single- and multi-site scheme. Also, will be very interesting to execute the new techniques capable to preserve the skewness in case of disaggregation modelling of the Mediterranean rivers.

7. Acknowledgment

This paper is partly based on the Ph.D. thesis of the author (Bojilova, 2003). The author wishes to express her deep sense of gratitude and appreciations to Prof. M.J. Hall and Prof. Str.G. Gerasimov for their supervision and invaluable advices during her research and the thesis works.

The author want to acknowledge that data evaluated in the study are received from "*General master plans for water resources management*" (2000) - project with Bulgarian Ministry of Water and Environment.

8. References

Bojilova, E. (2003): *Applicability of disaggregation models to the conditions of river runoff in Bulgaria*. Ph.D. thesis, 212 p, Sofia, Bulgaria.

Bojilova, E. (1999): *Generation of Monthly and Seasonal Streamflow Data Using Disaggregation Models*. 28th IAHR Congress, Topic C2: 187, 8 p, Graz, Austria.

General master plans for water resources management. (2000) Project with Bulgarian Ministry of Water and Environment, Sofia, Bulgaria.

Hakem, J. (1991): *Application of disaggregation models to rainfall in The Syrian Arab Republic*. MSc Thesis H.H. 97, IHE-Delft, The Netherlands.

Koutsoyiannis, D. (2001): *Coupling stochastic models of different time scales*. Water Resources Research, 37(2), 379-392.

Koutsoyiannis, D. (1999): *Optimal decomposition of covariance matrices for multivariate stochastic models in hydrology*. Water Res. Res., 35(4), 1219-1229.

Koutsoyiannis, D., Manetas, A. (1996): *Simple disaggregation by accurate adjusting procedure*. Water Res. Research, 32(7), 2105-2117.

Lin, Gwo-Fong. (1990): *Parameters estimation for seasonal to subseasonal disaggregation*. Journal of Hydrology, 120, 65-77.

Mejia, J.M., Rouselle, J. (1976): *Disaggregation models in Hydrology Revisited*. Water Res. Research, Vol. 12 (2), 185-186.

Salas, J.D., Delleur, J.W., Yevjevich, V.M. and Lane, W.L. (1985): *Applied modelling of hydrologic time series*. Water Res. Pub, 2nd printing, Littleton, Colorado, USA.

Todini, E. (1980): *The preservation of skewness in linear disaggregation schemes*. J of Hydrology, 47, 199-214.

Valencia, D.R., Schaake, J.C. (1973): *Disaggregation Processes in Stochastic Hydrology*. Water Resource Research, 9 (3), pp. 580-585.

STOCHASTIC MODELLING OF THE NATURAL RIVER FLOWS

Elena Bojilova[1]
[1]National Institute of Meteorology and Hydrology – Bulgarian Academy of Sciences, Bulgaria, Tsharigradsko chaussee 66, 1784 Sofia, Bulgaria,
E-mail: Elena.Bojilova@meteo.bg

Extended abstract: In design and operation of water resource systems the stochastic modelling plays a very important role. In the presented research disaggregation models are selected and applied. The objects of disaggregation are stations with natural river flows for chosen Bulgarian rivers – Cherni Lom, Kamchia, Provadiiska, Tundja and Maritza Rivers. As a result synthetic flow sequences are generated.

Disaggregation models are basically divided into three main groups: temporal, spatial and temporal-spatial. The focus of the paper is application of temporal disaggregation models for disaggregation of annual flow totals to monthly flows. Two basic models are applied: the Original model of Mejia and Rousselle and the Corrected extended Lin model, one-stage disaggregation. Data from five rivers in different areas of Bulgaria for the aims of the study are selected and executed. The comparison between the applied models is made. The Lin model, single-site is selected for generation of river flow totals.

The following conclusions can be drawn:

- Corrected parameter estimation in the Lin model succeeded in preservation of the additivity and the covariance matrices as expected. The corrected Lin model is an improvement compared to the extended Mejia and Rousselle model.
- The Lin model preserves the first two statistical moments and the covariance matrices rather well. However, the preservation of the second-order statistical moment is poorer than that of the first-order moment.
- Both models are suitable for disaggregation of the natural flows with snowmelt conditions together with rainfall in the catchment areas. They are applicable for the specific climatic and hydrological conditions of Bulgaria.
- The models are suitable for the specific conditions in the valley regions in the country for natural river flow with average mean altitude of the drainage basins between 200 *m* to 500 *m*. Furthermore, the best result is obtained from the corrected extended model, one-stage disaggregation single-site approach.

Further studies should be carried out to test applicability of the models to multi-site approach; and to apply disaggregation models to non-perennial and ephemeral rivers in single- and multi-site scheme.

Keywords: stochastic and disaggregation modelling, natural river flow.