CADERNOS DO IME – Série Estatística

Universidade do Estado do Rio de Janeiro - UERJ

ISSN impresso 1413-9022 / ISSN on-line 2317-4536 - v.37, p.29 - 39, 2014

DOI: 10.12957/cadest.2014.18302

STATIONARITY OF ANNUAL MAXIMUM DAILY STREAMFLOW TIME SERIES IN SOUTH-EAST BRAZILIAN RIVERS

Jorge Machado Damázio
UERJ - Universidade do Estado do Rio de Janeiro
CEPEL - Centro de Pesquisas de Energia Elétrica
damazio@ime.uerj.br

Fernanda da Serra Costa
UERJ - Universidade do Estado do Rio de Janeiro
CEPEL - Centro de Pesquisas de Energia Elétrica
f.scosta@globo.com

Abstract

The paper presents a statistical analysis of annual maxima daily streamflow between 1931 and 2013 in South-East Brazil focused in detecting and modelling non-stationarity aspects. Flood protection for the large valleys in South-East Brazil is provided by multiple purpose reservoir systems built during 20th century, which design and operation plans has been done assuming stationarity of historical flood time series. Land cover changes and rapidly-increasing level of atmosphere greenhouse gases of the last century may be affecting flood regimes in these valleys so that it can be that nonstationary modelling should be applied to re-asses dam safety and flood control operation rules at the existent reservoir system. Six annual maximum daily streamflow time series are analysed. The time series were plotted together with fitted smooth loess functions and non-parametric statistical tests are performed to check the significance of apparent trends shown by the plots. Non-stationarity is modelled by fitting univariate extreme value distribution functions which location varies linearly with time. Stationarity and non-stationarity modelling are compared with the likelihood ratio statistic. In four of the six analyzed time series non-stationarity modelling outperformed stationarity modelling.

***Keywords**: Stationarity; Extreme Value Distributions; Flood Frequency Analysis; Maximum Likelihood Method.*

1. Introduction

Although it is recognized that natural streamflow process is inherently a non-stationary process, hydrologists performing flood frequency analysis applied to water resources management and planning problems usually assume stationarity of historical flood time series whenever the drainage system and the land cover in the studied watershed remain without obvious substantial changes. In many watersheds which drainage systems and land covers are known to have been impacted by minor natural and man-induced events, estimation of future flood frequency conditions can be improved with procedures for detecting and modelling non-stationarity in historical time series. Climate change has put another potential source of non-stationarity increasing the interest on detection and modelling non-stationarity in flood time series.

The large valleys in South-East Brazil are among the economic heartland of the country. Protection of its flood-prone areas is provided by multiple purpose reservoir systems built during 20th century, which design and operation plans has been done assuming stationarity of flood time series. Land cover changes and rapidly-increasing level of atmosphere greenhouse gases of the last century may be affecting flood regimes in these valleys so that it can be that non-stationary modelling should be applied to re-assess dam safety and flood control operation rules at the existent reservoir system.

The paper present a statistical analysis of annual maxima daily streamflow between 1931 and 2013 in South-East Brazil focused in detecting and modelling non-stationarity aspects. Section 2 presents a description of the six analyzed annual maximum daily streamflow time series, it´s time plots together with smooth fitted curves and results of non-parametric statistical tests for assessing the significance of trends found in the series. Section 3 discusses the representation of location non-stationarity in the analyzed series by fitting univariate extreme value distribution functions which locations vary linearly with time. Section 4 gives conclusions.

2. Descriptive Statistics, Time Plots and Non-Parametric Tests for Trend

The Brazilian national electrical system operator (ONS) maintains a database with records of natural streamflow since 1931 for the main hydropower plants in the country, which operation is on its responsibility. These natural streamflow values are obtained from field measured streamflow values which latter on are corrected to consider known upstream abstractions from the river flow and filling/draining operation of reservoirs. No correction is provided for considering other anthropogenic activities in the hydrograph basin, like land use changes, neither for considering changes in the precipitation regimes. The stationarity assumption is considered in the stochastic daily streamflow model used for designing flood control operation rules for the hydropower plants reservoirs (Costa *et al*., 2014).

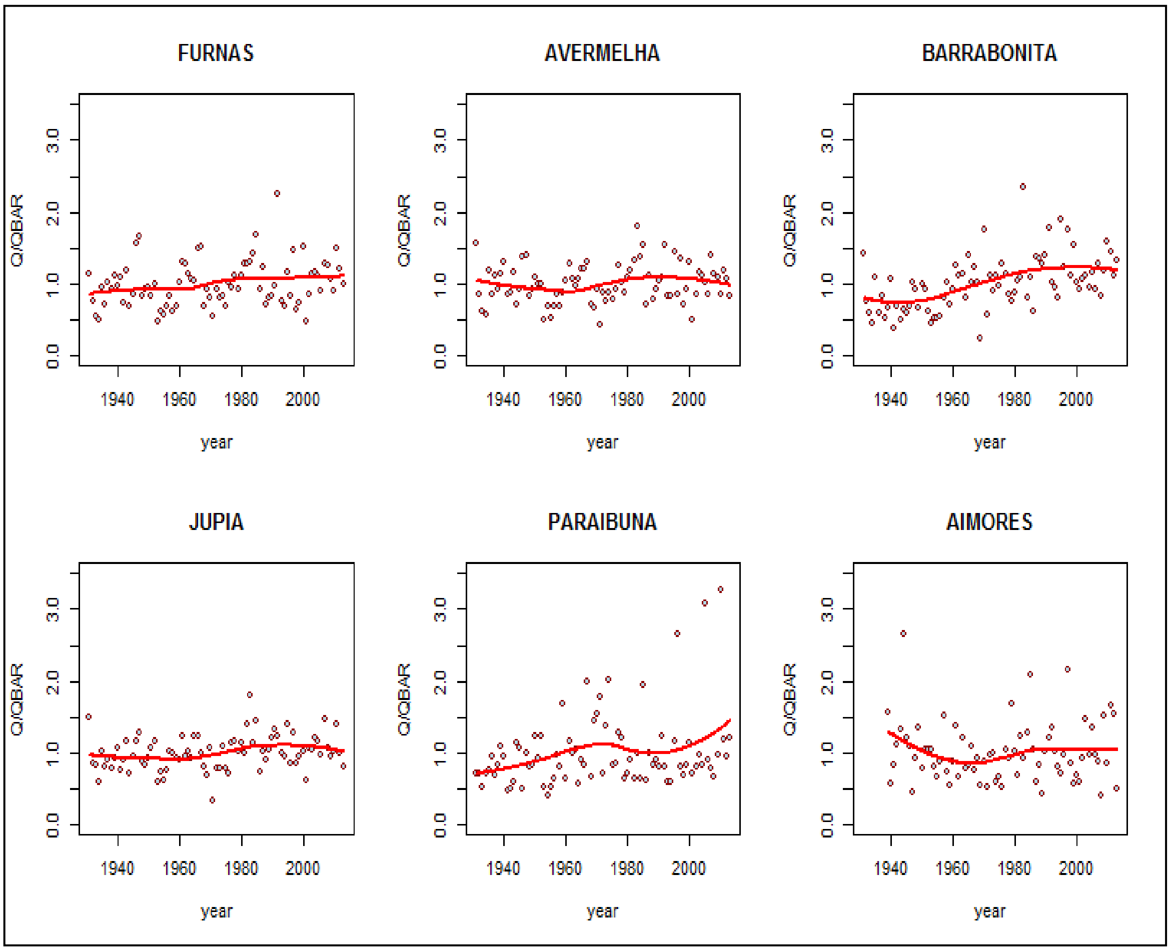
For this study it was chosen streamflow time series from six hydropower plants located in Brazilian South-East valleys (Table 1). Out of the chosen six hydro plants, four are located in the Paraná River valley: Furnas and Água Vermelha, located in Grande River, one of the Paraná River forming tributaries; Barra Bonita, located in Tietê River, the first left main tributary, and Jupiá, located in Paraná River channel just downstream of Tietê mouth. The other two chosen hydro plants are located in neighbor valleys: Paraibuna, located in the valley of Paraíba do Sul River and Aimorés, located in the valley of Doce River. Stream flow at all these hydro plants are influenced by the same macroclimatic events and share the same hydrologic regime with a much defined rainy season going from October to April, when major floods occur, and a dry season going from May to September, when streamflow is mainly base flow and no major flood can occur. For each natural daily streamflow time series, “annual” maximum daily streamflow values were obtained considering the maximum of each rainy season. With this method, the “annual” maximum daily streamflow values can be assumed independent. Table 1 also presents the record span and main descriptive statistics of each annual maximum daily streamflow series. Mean values in Table 1 vary highly reflecting the sites drainage area variability and extreme precipitation variability. In order make the univariate analysis more comparable, the values in each time series were divided by the correspondent mean.

Table 1: Annual Maximum Daily Streamflow Time Series Descriptive Statistics

Hydropower	valley	period	years	mean (m^3/s)	cv	skewness
Furnas	Paraná/Grande	1931-2013	83	3,312	0.32	0.95
Água Vermelha	Paraná/Grande	1931-2013	83	6,174	0.27	0.33
Barra Bonita	Paraná/Tietê	1931-2013	83	1,814	0.38	0.75
Jupiá	Paraná	1931-2013	83	16,880	0.24	0.40
Paraibuna	Doce	1931-2013	83	325	0.52	2.39
Aimorés	Paraíba do Sul	1939-2013	75	3,724	0.41	1.39

The six time series are plotted in Figure 1 together with fitted smooth loess curves showing graphically the trends suggested by the records. For every time series, its annual maxima daily streamflow shows a tendency to rise over the record time span with the exception of the Aimorés time series. It can also be noted that the rise period in the Paraná valley time series seems to be concentrated between early fifties and late eighties. This behavior is consistent with the increase of urbanization in these valleys after the Second War which stabilized around the nineties.

Figure 1: Annual Maximum Daily Streamflow Time Series and Trend Loess Curve



The time series were tested for trend using non-parametric procedures provided by Spearman´s rho and Mann-Kendall tests. Table 2 shows that at the significance level of 0.05 (5%) the two procedures yielded the same conclusion in every series, which is to asses as non-significant the trends in the Água Vermelha and Aimorés time series otherwise asses as significant the trends in the other four time series.

Table 2: Non-parametric tests for trend

Hydropower	valley	Spearman´s rho		Mann-Kendall	
		statistics	pvalue	statistics	pvalue
Furnas	Paraná/Grande	0.232	3.526e-02	533	3.642e-02
Água Vermelha	Paraná/Grande	0.128	2.494e-01	315	2.169e-01
Barra Bonita	Paraná/Tietê	0.559	3.940e-08	1308	2.746e-07
Jupiá	Paraná	0.278	1.098e-02	636	1.251e-02
Paraibuna	Doce	0.231	3.592e-02	537	3.501e-02
Aimorés	Paraíba do Sul	-0.023	8.417e-01	-61	7.837e-01

3. Stationary and non-stationary modeling using extreme value distributions

Gumbel´s theory of extreme value statistics was used for modeling the distribution of the annual daily streamflow maxima by considering the family of extreme value distributions referred to as Generalized Extreme Value (GEV) distributions, which cumulative probability function (CDF) are:

$$F(x) = exp\left\{-\left[1 - k\left(\frac{x-\mu}{\sigma}\right)^{-1/k}\right]\right\} \qquad for\ k \neq 0 \qquad (1)$$

where ∞<μ<∞ is a location parameter, $\sigma > 0$ is a scale parameter and k a shape parameter. For $k > 0, x > \mu - \sigma/k$; for $k < 0, x < \mu + \sigma/(-k)$. The case k=0 is obtained taking the limit as k→0 in (1) to find the Extreme Value type 1 (EV1) or Gumbel distribution valid for ∞<x<∞ given by:

$$F(x) = exp\left\{-e^{-\left(\frac{x-\mu}{\sigma}\right)}\right\} \quad for\ k = 0 \qquad (2)$$

Modeling non-stationarity in the annual maximum daily streamflow time series was done by representing trends in the location parameter μ with a linear function of the year t as:

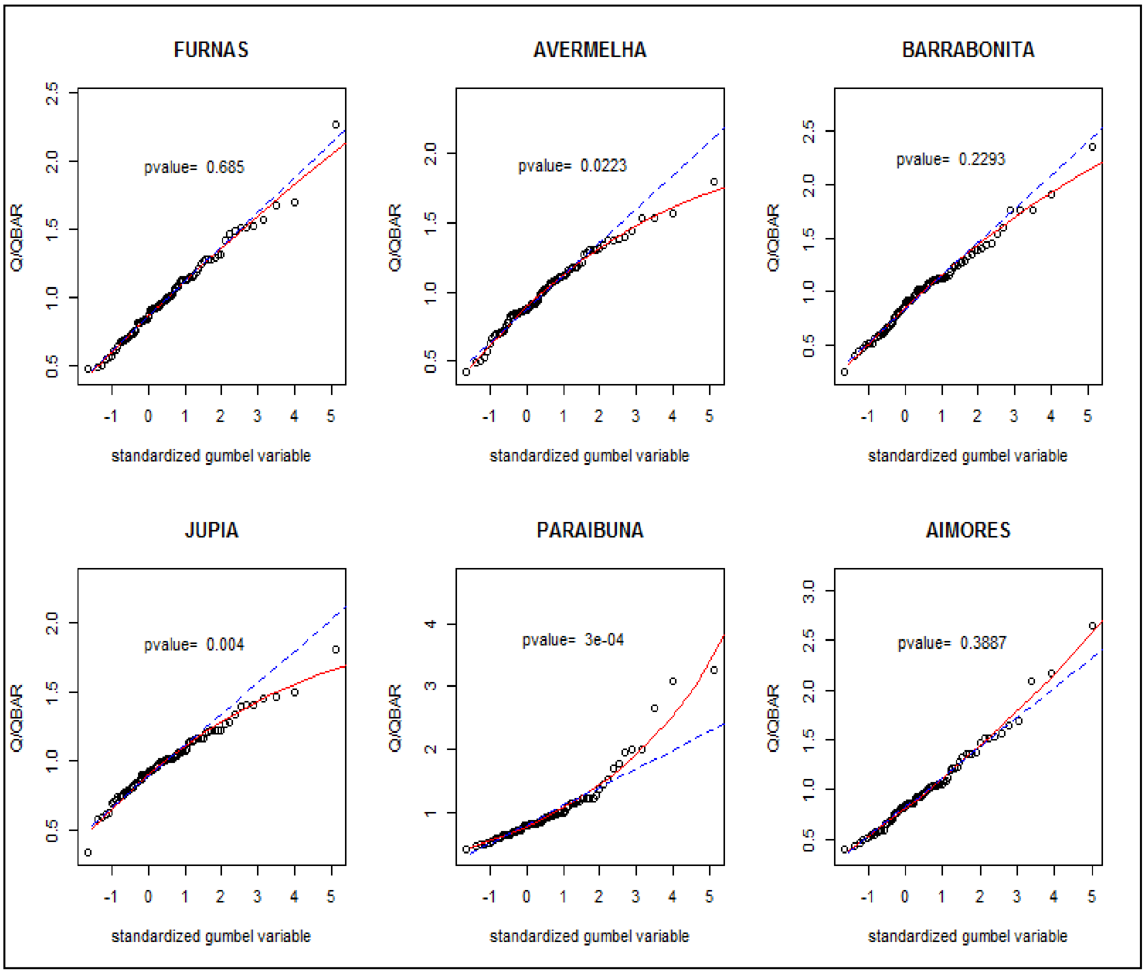
$$\mu(t) = \mu_0 + t\ .\ \mu_1/100 \qquad (3)$$

where μ_1 is location change rate in percent/years.

Stationary and non-stationary models were fitted to the annual maximum daily streamflow time series by maximum likelihood with the aid of the R package extRemes, version 2.0-5 (Gilleland and Katz, 2011).

For stationary models, both Gumbel and GEV distributions were considered and likelihood-ratio statistics were calculated to assess the significance of the estimated GEV shape parameters. Figure 2 present for each series the comparison of the goodness of fit of both distribution and the p-levels attained by likelihood-ratio statistics.

Figure 2- Goodness of fit of fitted stationary GEV (solid red curve) and Gumbel (dashed blue straight line) distributions and likelihood-ratio statistics p-level.



For non-stationary models, GEV distributions was considered only for series which estimated shape parameters in the stationary model proved significant at the 0.05 level (Água Vermelha, Jupiá and Paraibuna) and Gumbel distributions was considered for the other series. Two non-stationary models were fitted: Model 1 assumes that equation (3) is valid for all t whereas Model 2 assumes equation (3) is valid only between 1955 and 1990, with constant location parameter for years up to 1955 and after 1989.

The significance of the estimated location change rates were evaluated by calculating the p-values of likelihood-ratio statistics calculated with the non-stationary and stationary maximum likelihood fits and the negatives of the likelihood attained.

Figures 3 and 4 show for each series the profile likelihood curves for non-stationary modelling (figure 3 for Model 1 and figure 4 for Model 2) obtained varying location change rate and fixing other parameters at maximum likelihood estimate together with the acceptance interval for the stationary null hypothesis at the 0.05 significance level.

Figure 3- Location Change Rate Estimation for Model 1. Profile likelihood graph (solid black curve), maximum likelihood estimate (dashed black vertical line) and stationarity acceptance interval at the 5% significance level (gray vertical interval).

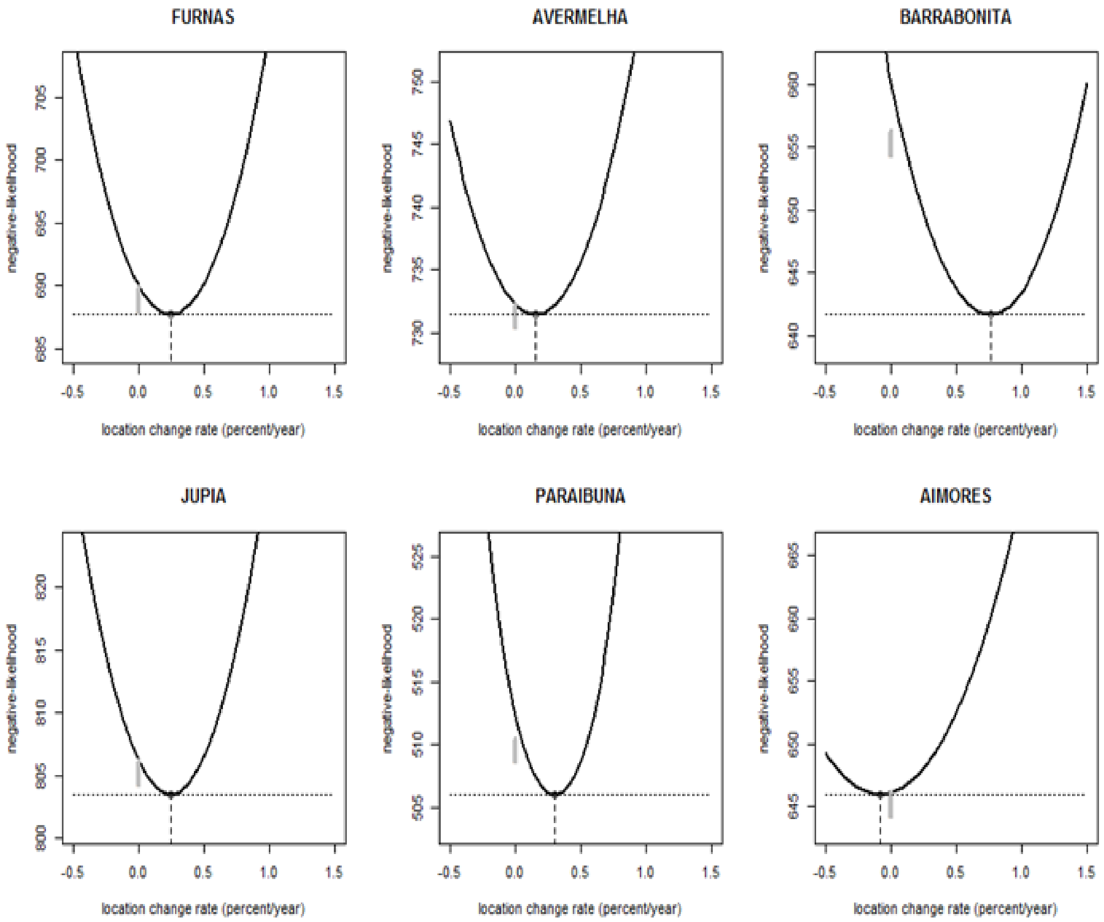


Figure 4- Location Change Rate Estimation for Model 2. Profile likelihood graph (solid black curve), maximum likelihood estimate (dashed black vertical line) and stationarity acceptance interval at the 5% significance level (gray vertical interval).

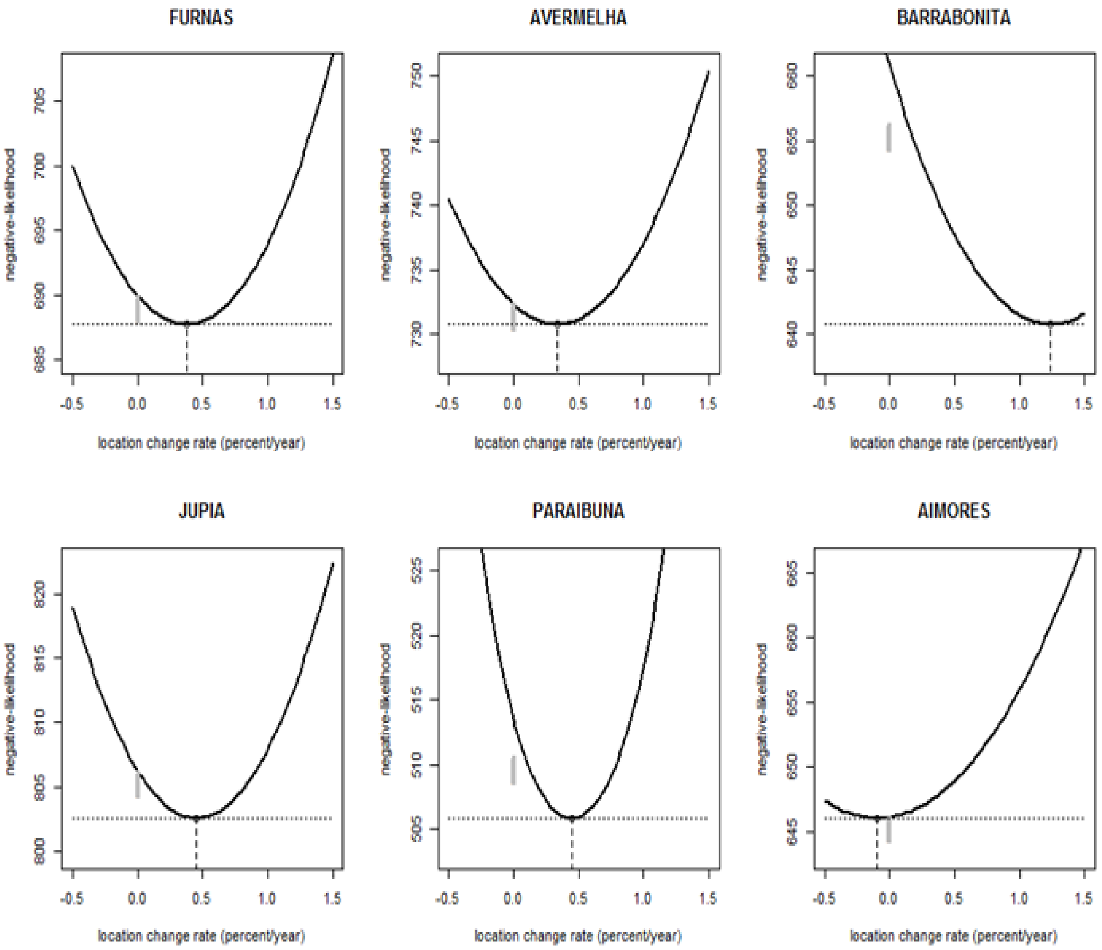


Table 3 presents the estimates of the location change rate and the p-values of its likelihood-ratio statistics for the two non-stationary models. Note that for both non-stationary models only the location change rates of Água Vermelha and Aimorés time series are non-significant at the 0.05 significance level, which is the same conclusion obtained with the above reported application of non-parametric trend tests. It can also be noted that the two non-stationary models are statistically undistinguishable except for Jupiá, where Model 2 is preferred.

Figures 5 and 6 show plots of the analyzed time series together with curves of estimates of quantiles of probabilities 0.03, 0.50 and 0.97, obtained with stationary models for Água Vermelha and Aimorés and non-stationary models (Figure 3 with Models 1 and figure 4 with Models 2) for the others time series.

Table 3: ME estimates of location change rate in non-stationary models and p-values

Hydropower	Model 1			Model 2		
	location change rate	p-value	Negative Log-Likelihood	location change rate	p-value	Negative Log-Likelihood
Furnas	0.243	3.659e-02	14.967	0.375	3.858e-02	15.011
Água Vermelha	0.158	1.878e-01	7.016	0.331	8.274e-02	6.378
Barra Bonita	0.760	6.499e-08	18.896	1.240	2.677e-08	18.034
Jupiá	0.242	2.345e-02	-1.856	0.455	8.515e-03	-5.318
Paraibuna	0.303	2.779e-03	25.998	0.445	2.303e-03	25.826
Aimorés	-0.082	6.167e-01	29.289	-0.094	7.025e-01	29.342

Figure 5- Annual Maximum daily streamflow series and curves of (0.03, 0.5, 0.97) quantiles estimates for stationary models at Água Vermelha and Aimorés and for non-stationary Models 1 at the other sites

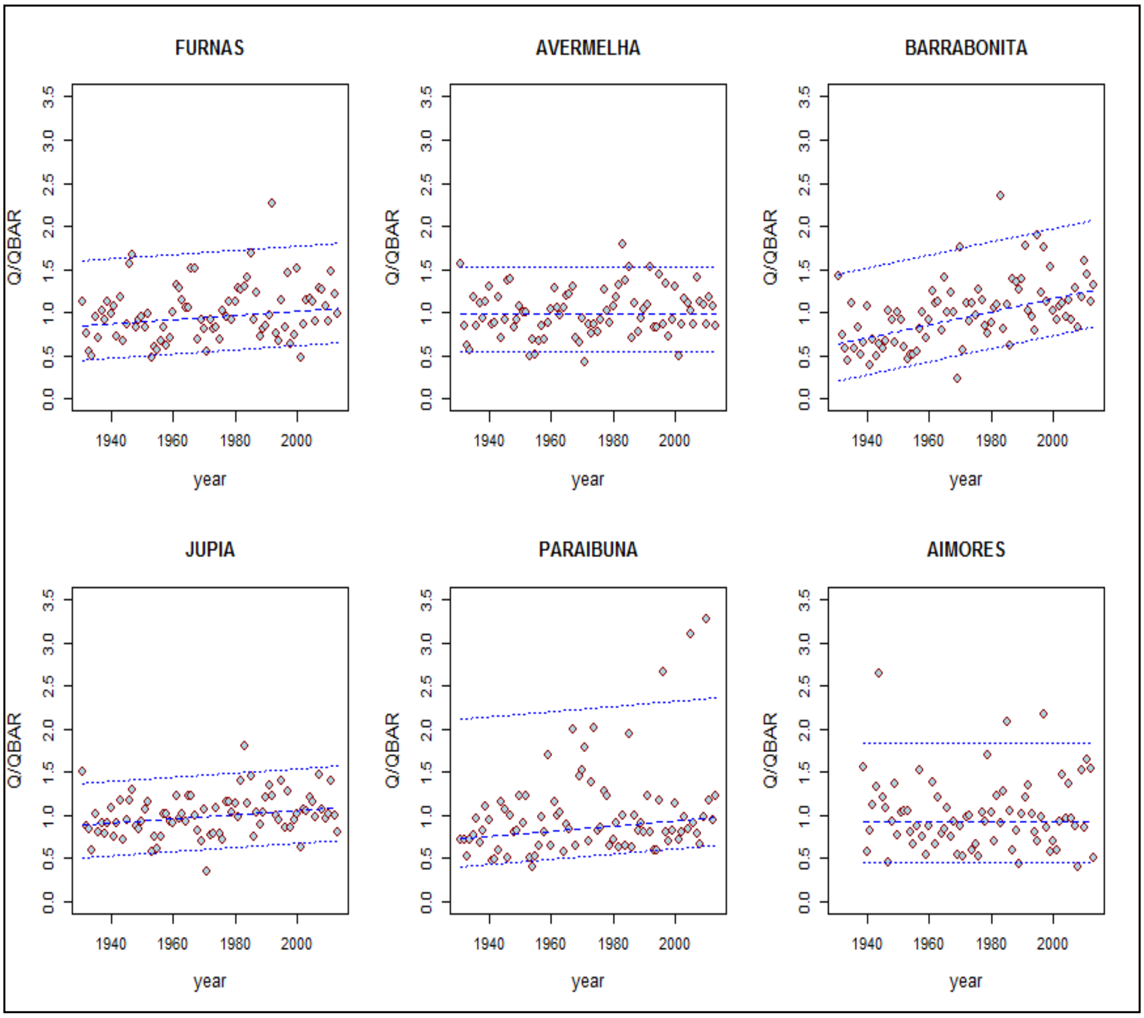
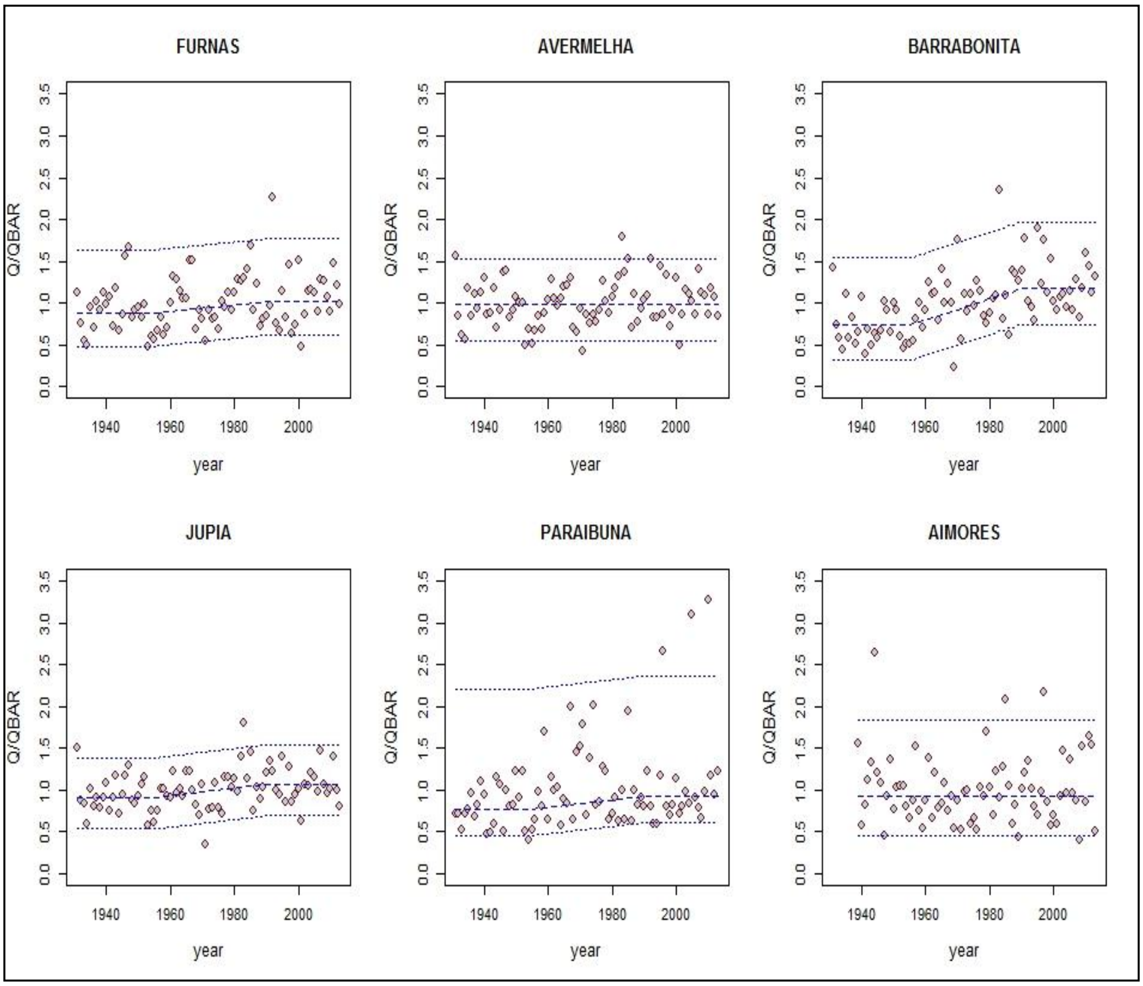


Figure 6- Annual Maximum daily streamflow series and curves of (0.03, 0.5, 0.97) quantiles estimates for stationary models at Água Vermelha and Aimorés and for non-stationary Models 2 at the other sites



4. Conclusions

The results of non-parametric trend tests identified significant positive trends in 4 out of 6 analyzed maximum daily flow time series at South-East Brazilian valleys. Two extreme value non-stationary modeling options imposing a linear equation for the location parameter either for all the period or for only in a limited portion of it (1955 to 1990) attained statistically similar goodness of fit, although the practical consequences of either model in re-assessing dam safety and flood control operation rules at the existent reservoir system in these valleys can be very different.

References

GILLELAND E; KATZ R. W (2011). New software to analyze how extremes change over time. **Eos**,11January, 92, (2), pp. 13-14

COSTA, F. S.; RAUPP, I. P.; DAMAZIO, J. M.; OLIVEIRA, P.D;.GUILHON, L.G (2014). The Methodologies for the Planning of the Flood Control using Hydropower Reservoirs in Brazil, **In.** 6^{th} International Conference on Flood Management (ICFM6), 16-18 September 2014, São Paulo – Brazil

ONS (2015); http://www.ons.org.br/operacao/vazoes_naturais.aspx; visited in 10/10/2014.