

Minimization of the conflict between electric power generation and flood control of the Brazilian reservoirs operation planning

I. Raupp¹, F. Costa^{1,2} & J. Damázio^{1,2}

¹*Department of Energy Optimization and Environment,
Electric Power Research Center – CEPEL, Brazil*

²*Rio de Janeiro State University – UERJ, Brazil*

Abstract

The Brazilian Power System is a hydrothermal system characterized by large reservoirs arranged in complex cascades over several river basins. The Brazilian Water Resource Policy determines that water resource management has to provide multiple use of water. Thus, the System Operation Planning should consider the use of part of the hydro plants reservoirs' volumes as flood control volumes for attenuation of possible floods, avoiding damage to downstream areas. The use of reservoir volumes for electricity generation and flood control is conflicting since the first use attempts to keep the reservoirs as full as possible and the second one attempts to keep part of it empty at the same period. As a result, some years, when the end of the rainy season inflows are low, the reservoirs may not end the season completely filled. To address this problem, this paper presents a methodology that changes the flood control volumes before the end of the rainy season according to hydrological information. The methodology proposed was tested in a Brazilian river basin to check its applicability and effectiveness.

Keywords: flood control, reservoir's operation, operation planning.

1 Introduction

The Interconnected Brazilian System of electricity production and transmission (SIN) is a large-scale hydrothermal system, with strong dominance of hydroelectric power plants with multiple owners.



One of the drawbacks of hydro-based electric generation system is the time fluctuations of natural streamflows. In order to provide protection against streamflow fluctuations, the Brazilian electric energy generation system operates with complementary thermal plants, whose operation involves additional fuel expenses.

The coordination of the operation of the reservoirs with the complementary thermal plants permits the best use of the natural streamflows, avoiding waste of water and excessive fuel expenditures. This coordination is done in the framework of the Interconnected System Operation Planning Studies performed at the Brazilian Electrical System National Operator – ONS.

In this planning, the main goal is to minimize the expected value of the operation cost (thermal generation expenses and penalties for not attempting the demand) over the planning horizon. However, the Brazilian Water Resource Policy determines that the water resource management has to provide for the multiple uses of water, in that way, it is necessary to take into account the multiple uses of water in the hydropower plant's reservoirs, for example, the flood control.

The use of the hydroelectric plants' reservoirs for flood control imposes constraints to the operation of reservoirs for electricity generation, for example: on the maximum outflows (to prevent flooding in downstream valley) and on the maximum levels (to prevent flooding upstream of the reservoir). To preserve these restrictions it is necessary to reserve part of the reservoirs' storage capacity for flood control. This situation can create a conflict, once the majority of hydroelectric plants were not designed to consider the flood control and each use (flood control and energy generation) operates the reservoir's volumes differently. The electricity generation demands to keep the volumes filled (to face the dry season) while the flood control needs to keep part of them empty (named *waiting volumes*), in the same period. This justifies the need of an optimized and careful determination of the *waiting volumes*.

The Interconnected System Operation Planning tries to minimize this conflict through the Annual Flood Plan. In this planning, performed by ONS, annually before the beginning of the rainy season, the *waiting volumes* is defined for the reservoirs located in the river basins where there are flood control operational restrictions. The definition of these volumes is associated with a risk (recurrence time) of not protection the downstream valley.

The methodology in use nowadays [1] adopts some resources for a better *waiting volumes* definition, such as: (i) seasonal *waiting volumes* allocation to increase the frequency of years in which the reservoirs are filled on the beginning of the dry season, (ii) use of stochastic scenarios of possible floods that might occur in the river basin, on the calculation of the *waiting volumes*, (iii) System Reservoir Analysis: for the protection of a flood control point (where there are flood restrictions), it is not consider only the reservoir immediately upstream, but all reservoirs in the river basin upstream the flood control point, and (iv) consideration of macro-climate information for the generation of possible floods scenarios [2].



Despite those resources, some years, the reservoirs may not reach the end of the rainy season fully filled, because the inflows are not able to fill the volumes allocated in this period for flood control.

In that way, this paper presents a methodology that aims to increase the probability that the reservoirs used for flood control and energy generation are going to be full at the end of the rainy season.

Overall, this methodology doesn't modify the calculation of *waiting volumes*, but proposes to change the *waiting volumes*, in a specific date in the final period of the rainy season, according to the average monthly streamflow observed in the initial months of the season. This methodology is called "*Change of the waiting volumes curve at the end of the rainy season according to the hydrological information on the beginning of the season*" [3].

In this same paper, the methodology is implemented and tested in the Paraná River's basin, located in the Brazil's southeast region, to verify the applicability and effectiveness of the proposal [3].

2 Current methodology to calculate the waiting volumes

The methodology to calculate the *waiting volumes* to be allocated in the reservoirs of the SIN is based on the *controllability conditions theory* [4, 5]. This theory is an extension of the *critical path method* for systems with multiple reservoirs and multiple flood control points.

The *critical path method* works as a water balance between the amount of water that arrives at the reservoir (inflow), the maximum outflow that can be thrown, according to the flood control point constrain (outflow constrains), and the *waiting volumes* allocated in the time step before. This method uses a backward algorithm in time, i.e., the calculation is done from the last day of the rainy season (T) until the first day, as eqn. 1.

$$VV(t-1, s) = \max[0, (Q_{IN}(t-1, s) - Q_R) \times \Delta t + WV(t, s)], \quad t = T, \dots, 1 \quad (1)$$

where $WV(t, s)$ – *waiting volume* at day “t” at rainy season “s”; $Q_{IN}(t, s)$ – daily average inflow at day “t” at rainy season “s”; Q_R – outflow restriction (maximum outflow permitted); Δt – number of seconds at a day.

The *waiting volumes* calculated for all days of the rainy season of a year is called *critical path* for this year. The *critical path* ensures the protection of the reservoir downstream areas for one daily inflows sequence. However, as the *waiting volumes* were defined before the beginning of the rainy season, it is not possible to know what daily inflows sequence will occur during this season. This problem, due to randomness of the inflows is stochastic. To solve this problem, it is considered a set of daily inflows scenarios possibly to occur in the rainy season. To generate these scenarios is used a stochastic model that generates the synthetic daily inflows scenarios [6]. For each scenario is calculated its corresponding *critical path*.

If the objective is to provide 100% protection, given all these scenarios, so it must be set the envelopment of the *critical paths*. This envelopment, called *envelope curve*, is calculated using eqn (2), providing a limit curve, that indicates



for each day of the rainy season the necessary *waiting volume* to ensure that for any of the scenarios considered, the flood control point, associated with the reservoir, will be protected.

$$ENV(t) = \max[WV(t, s); s = 1, \dots, n]; t = 1, \dots, T \quad (2)$$

where: $ENV(t)$ – *envelope curve* at the day “ t ”; n – number of scenarios considered on the *envelope curve* calculation.

If the protection of the downstream area is associated to a risk, it does not need to ensure that the outflows do not exceed the maximum outflow restriction for all the scenarios considered. The number of scenarios that will not be protected is function of the risk (recurrence time) and is defined by the ratio between the total number of scenarios and the recurrence time (in years).

After determining the number of scenarios that will not be protected, it is necessary to chose a criterion to dispose the scenarios that won't be considered in the *envelope curve* calculation. In this paper, we adopt a criterion associated with the maximization of the *waiting volumes* recovery probability. This criterion disregards the scenarios that require greater inflows to ensure that the reservoir reaches full at the end of the rainy season. By this criterion, the scenarios to be disregard will be the ones that present the major values for the maximum tangential (eqn (3)):

$$TAN_{MAX}(s) = \max \left[\frac{WV(t, s)}{(T + 1) - t}; t = 1, \dots, T \right]; s = 1, \dots, n \quad (3)$$

where: $TAN_{MAX}(s)$ – maximum tangential for the scenario “ s ”.

3 Anticipating the recovery of the *waiting volumes*

Imagine, for simplicity, the situation of a single reservoir, which the *envelope curve* is shown in Figure 1 (the continuous curve), located in a river basin where the rainy season goes from November 1st to April 30th. Suppose that the inflows of the past months of the rainy season are below their average long-term inflow value. Thus, we may question: Would be possible, for example, to anticipate the *waiting volumes* recovery at the beginning of the penultimate month of the rainy season (March), replacing the *envelope curve* represented by the continuous curve by the dashed one (Figure 1)? This may increase the guarantee that the reservoir will be full at the end of the rainy season.

To answer this question, it is necessary to verify if the information of the hydrology (inflows) of the months that precedes the end of the rainy season can predict the flooding behavior in the final months of the rainy season. If this occurs, it is also necessary to define how to incorporate this information to the procedures and methods of flood control, preserving its risks (recurrence time). Following this, the methodology that will be presented in the next items will be divided in two parts: (i) verification of the capability to preview the flood behavior at the end of the rainy season and (ii) the methodology to incorporate

this information in the current *envelope curves* calculation. The correspondent application will be presented after each part.

The methodology was applied in Paraná River's basin, located in the Brazil's southeast region. In this river basin there are 32 hydro plants and 15 flood control points. The most downstream hydro plant is Jupiá, which has a flood control point immediately downstream, to avoid the flood damage in the valley, the flow in this point must be below 16,000 m³/s. For simplification, this reservoir's system was considered as a single reservoir until Jupiá hydro plant and the inflows as the natural flows until this plant. The historical daily inflows serie used is formed by 55 rainy seasons (from November 1st to April 30th), which will be called *years*, for simplification. Typically, in this river basin, the concern with the *waiting volumes* recovery starts in early March, when the inflows of the previous months are below the average long-term inflow.

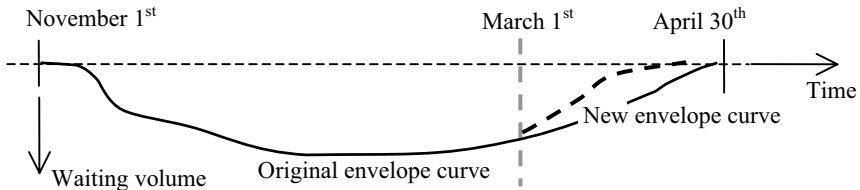


Figure 1: Example of an *Envelope curve* recovered on March 1st.

3.1 Verification of the capability to preview the flood behavior at the end of the rainy season

Suppose a river basin in which the rainy season covers the period from November 1st to April 30th and the concern with the *waiting volumes* recovery starts in early March. On that way, to check if it is possible to anticipate the *waiting volumes* recovery from March 1st, it would be natural to consider the monthly inflows from November to February as the hydrological information that can anticipate the flood behavior in March and April.

Assuming that, the hydrological information of the rainy season previous months can be represented by the average of different sets of monthly inflows: November-February, December-February, January-February and February. The flood behavior at the final months of the rainy season (March and April) can be represented by the maximum streamflows, for 1 day, 7 days and 15 days, and also by the amount of *waiting volumes* required. All these variables can be calculated by the historical daily streamflows series.

To identify which combination of average monthly streamflows has more influence on the flood behavior at the period March-April, we can calculate the correlation matrix between the average monthly streamflows and the maximum streamflows of different durations (1 day, 7 days and 15 days). The combination that shows the highest correlations is selected to represent the hydrological behavior in the beginning of the rainy season.

After selecting this variable, it is necessary to classify the hydrology (for example: High, Average and Low). It is important to consider the followings criteria: (i) The procedure to classify has to be related with the flood behavior in March-April (for example, maximum streamflow and *waiting volumes*); (ii) The parsimony, which implies that the decision to increase the number of classes is only justified if it increases significantly the flood behavior explanation in March and April; and (iii) the number of years of the historical inflows in each class has to be sufficient to estimate the parameters of the synthetic daily streamflows series generation model.

The method adopted to classify the hydrology was the Classification and Regression Tree [7] which is a technique for exploratory data analysis used when the data structure is unknown and when the amount of information is large. The Classification/Regression Tree divides the observation space into subsets whose elements belong to the same class. The structure of a Tree is formed by an initial node, called “root”, each one of the decision points is called “node” and the terminal nodes are called “leaves”. Each “node” is associated with a decision criterion for the sub division. The number of “leaves” represents the number of classes in which the observation space was divided. In Figure 2 is shown an example of Classification/Regression Tree, where the observation space was divided into 3 classes.

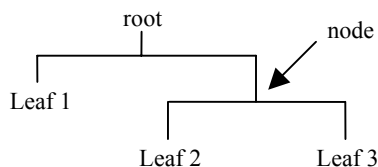


Figure 2: Example of a tree divided into 3 classes.

The growth of the tree is based on a binary partitioning algorithm on which each node divides the data into classes, having as a criterion the minimization of the residual deviations sum. In case of the Regression Tree (when the explanatory variable and the response variable are numeric), the criterion is the minimization of the residual variance sum (D) (eqn. (4)) and each division is made to maximize the reduction of D .

$$D = \sum_j \sum_i (y_{j,i} - \mu_j)^2 \quad (4)$$

where: $y_{j,i}$ - response variable “i” of the class “j”; μ_j - average response variable of the elements of the class “j”.

The complexity of a tree is described by the number of nodes. A tree can be more complex than necessary to describe the data. One way to balance the complexity of the tree and its ability to adjust (describe) the data is by pruning successively the less important branches. The pruning algorithm has a function similar to the residual variance (eqn (5)). The pruning algorithm obtains the sub-

tree that minimizes $D_k(T')$ over all sub-trees. The larger K , the lower is the number of nodes.

$$D_K(T') = D(T') + K \cdot \text{size}(T') \quad (5)$$

where: $D_k(T')$ - corrected residual variance of the sub-tree (T'); K - cost-complexity parameter (parsimony lost); $\text{size}(T')$ - number of terminal nodes of T' .

In this paper was used the statistical software S-Plus2000 [8], which has the functions *Tree* and *Prune.tree*. Three specific functions were created to obtain the results presented in the next item.

3.1.1 Application in Paraná River Basin

The procedures proposed in section 3.1 were applied in Paraná River's basin. On that way, to check if it is possible to anticipate the *waiting volumes* recovery from March 1st, it was considered the monthly inflows of November to February as the hydrological information that can anticipate the flood behavior in March and April. So it was calculated the correlation matrix between the average monthly streamflows and the maximum streamflows of different durations (Table 1). Initially the variables selected to represent the hydrological behavior were the average monthly streamflow of January-February and February, because they presented the highest correlations with the maximum flow rates.

Table 1: Correlation matrix.

Average Monthly flows	Maximum flows in March-April		
	1 day	7 days	15 days
November-February	0.49	0.49	0.52
December-February	0.51	0.51	0.54
January-February	0.58	0.58	0.61
February	0.60	0.58	0.59

Considering the maximum daily streamflow in March-April and February average monthly streamflow, the Regression Tree methodology described in Section 3.1, was used to build the tree shown in Figure 3. On the left part of Figure 3, it is possible to observe that dividing the maximum daily streamflows in two classes, depending on the February monthly streamflows, the residual variance is reduced in 35% (5 to 3.25) and dividing into three classes, the reduction of the residual variance is 50% (5 to 2.5). Increasing the number of classes for four obtains the same residual variance obtained by three classes. Following the criterion defined in Section 3.1, it was decided to classify the February monthly streamflows in three classes: High, Average and Low. The limits for this classification obtained through the tree model (on the right part of Figure 3) were:

High - February monthly flow exceeding 13,479.5 m³/s;

Average - February monthly flow between 7,625.0 m³/s and 13,479.5 m³/s;

Low - February monthly flow less than 7,625.0 m³/s.



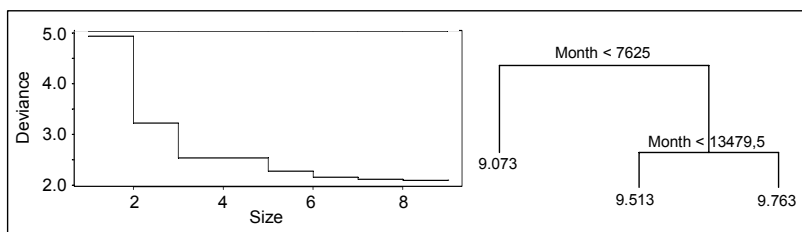


Figure 3: Regression Tree model classification.

The average daily inflow logarithmic in March-April, for each class, is shown in the leaves of the tree model (Figure 3).

Table 2: February monthly flow upper limits and the frequency of maximum daily flows greater than 16,000 m³/s in March-April.

Class	Upper limit of February monthly streamflow (m ³ /s)	N° of daily streamflows greater than 16,000m ³ /s in March-April		Total
		Yes	No	
High	∞	10	7	17
Average	13,479.5	6	25	31
Low	7,625.0	0	7	7

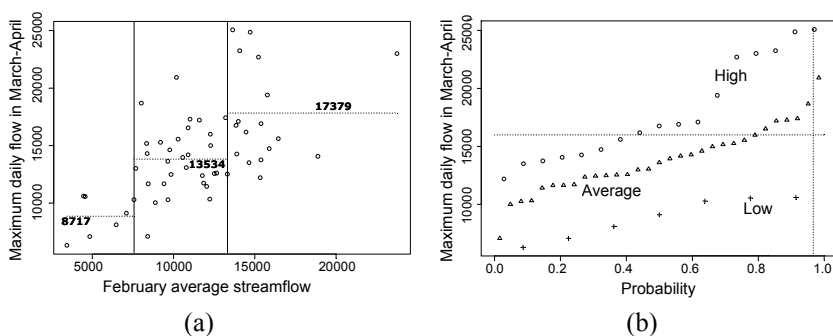


Figure 4: (a) Dispersion of the maximum daily flows in March-April versus the February monthly average flows. (b) Cumulative probabilities distributions of maximum flows in March-April for each class.

Table 2 shows the upper limits for the February monthly flows for each class, and the number of occurrence of the maximum daily flows greater than the 16,000 m³/s in March-April. We can note that the occurrences of daily inflows greater than 16,000 m³/s, in each class, are very distinct. Figure 4(a) shows the maximum 1-day flow in March-April versus a February monthly average flow graphic.

Figure 4(b) shows the cumulative probability distribution of maximum flow for each class. There is a huge difference between them, which means that it is a good classification. In order to verify if the average monthly streamflows during

January-February resulted in a better classification, the analysis presented in this section were repeated, replacing the February average monthly streamflows by January-February average monthly streamflows. However, the average flow in February resulted in a better classification, because the difference in the frequency of streamflows greater than 16,000 m³/s in the three classes was higher and the distribution of the observations in each class was more uniform. So the February average streamflow was considered as the variable to indicate the behavior of the necessity of *waiting volumes* in March-April, according to the classification: High (flow exceeding 13,479.5 m³/s), Average (flow between 13,479.5 m³/s and 7,625 m³/s) and Low (flow less than 7,625 m³/s).

Figure 5(a) shows the annual maximum *waiting volumes* in March-April versus the February monthly flow and in Figure 5(b) is shown the cumulated probability distributions of maximum *waiting volume* in March-April for the three classes. It can be observed that in the Low class it was not necessary *waiting volumes* and the cumulative distributions for the three classes are very distinct. In Figure 5(a), the horizontal lines represent the average annual maximum *waiting volumes* in March-April for each class and the vertical lines, the limits of the classes.

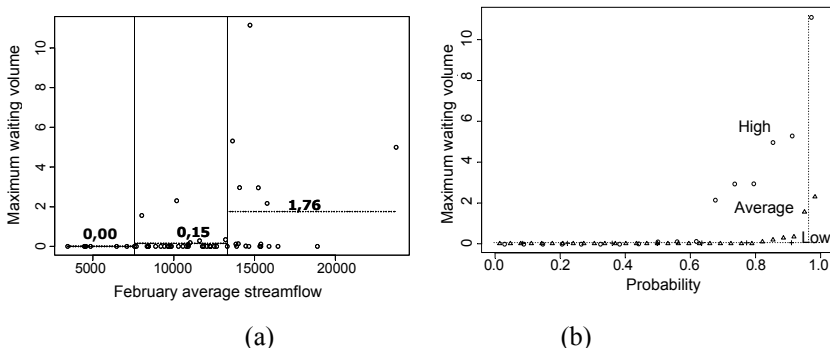


Figure 5: (a) Annual maximum *waiting volumes* in March-April versus February monthly inflows. (b) Maximum *waiting volumes* cumulative probabilities distributions per class.

3.2 Methodology to incorporate the recently past hydrological information in the current envelope curves calculation

Suppose the same river basin described in section 3.1 and that is desirable to review the *envelope curves* on March 1st, using as a criterion the classification of the monthly flow occurred in February, preserving the risk established in the beginning of the study. For this, it is necessary to have one *envelope curve* for the period from November 1st to February 28th, and three *envelope curves* for the period from March 1st to April 30th, each one corresponding to a class of February monthly streamflow (High, Average and Low), as illustrated in Figure 6. It is emphasized that the review of the *waiting volumes* on March 1st not necessarily indicates that every year is going to have a reduction on these

volumes. Some years, when the observed February monthly streamflow is classified as High, this review may indicate the need of higher *waiting volumes* during March-April than the ones calculated earlier.

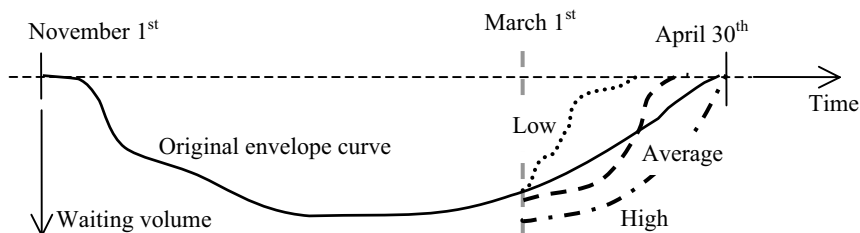


Figure 6: Example of *envelope curves* for the review of the waiting volumes.

Suppose available a daily inflows historical series with NH rainy seasons (years). The procedure to calculate their *envelope curves* is described above:

- a. Generation of N synthetic daily flows series conditioned by the three classes;
 - Division of the NH years into three sets according to their February average monthly flow: Low (NHL), Average (NHA), and High (NHH);
 - For each class, estimation of the parameters for the stochastic model to generate synthetic daily inflows series;
 - Generation of NH, NA and NL synthetic series from each set of parameters. NH, NA and NL are the number of synthetic daily inflows series, whose sum is equal to N. These sets of series are classified as High, Average and Low respectively. Where:

$$NH = (N \cdot NHH) / NH \quad NA = (N \cdot NHA) / NH \quad NL = (N \cdot NHL) / NH \quad (6)$$

- b. Calculation of the *critical path* (eqn. (1)) of the N synthetic series for the period from November 1st to April 30th;
- c. Definition of the Recurrence time (TR) and determination of the number of series that will be not protected (NP);
- d. Selection of the NP synthetics series that will be not protected, eqn. (3);
- e. Calculation of the *envelope curve* of the N-NP *critical paths*, eqn. (2);
- f. Determination, in the set of NP series, whose are not protected in March-April and how many of them belong to each class;
- g. Calculation of the *envelope curve* for March-April for each set of synthetic series (NH, NA and NL), not protecting the synthetic series identified on item f.

3.2.1 Application in Paraná River Basin

The methodology proposed in section 3.2 was applied in Paraná River's basin. For the classification of the historical daily inflows serie, the 55 years were divided as shown in Table 3, column 2. We adopted N equal to 12,000, and the number of synthetic series generated for each class is presented in column 3.

The recurrence time adopted was 30 years. Thus, the number of series that should be not protected was 400 (NP). On the set of 400 unprotected synthetic

series, 229 were not protected in March-April. On the fourth column of Table 3 is shown how many of these 229 series belong to each class.

In Figure 7 are presented the *envelope curves* for the whole rainy season (November-February), that will be called *original envelope curve*, and the three *envelope curves* for each class in March-April, that were calculated maintaining their original recurrence time in that period (Table 3, column 5). In Figure 7, the y-axis is the difference between the reservoir storage capacity and the *waiting volume* (in %). In this figure, we can observe the increase of the *waiting volumes* necessity when the February monthly flow is classified as high, but in the last three weeks it is equal to the original curve. The *envelope curve* for the Average class anticipates in one week the *waiting volumes* recovery and allocates less volume from week 20 to 24. The *envelope curve* for the Low class anticipates in 3 weeks the recovery and has a huge difference in the *waiting volumes* allocation.

Table 3: Historical and synthetic information per class.

Class	Historical serie	Synthetic serie	Number of series not protected in March-April	Recurrence time in March-April
Low	NHL = 7 (12.7%)	NL = 1527	0	∞ years
Average	NHA = 31 (56.4%)	NA = 6763	4	3000 years
High	NHH = 17 (30.9%)	NH = 3710	225	53.3 years
Total	NH = 55	N = 12000	229	52 years

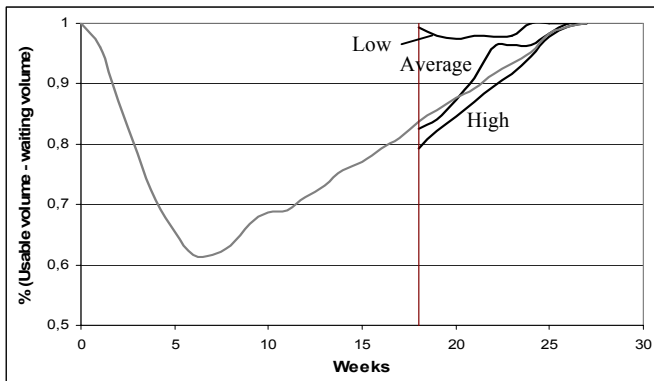


Figure 7: *Envelope curves* for the period November to April.

4 Conclusions

In this paper was shown that, in Paraná River's basin, the most relevant information to indicate the flood behavior in March-April is the average monthly flow of February. Adopting the Regression Tree methodology and the historical inflows series of 55 years (rainy seasons), the February monthly flow was classified as High, Average and Low. In the period March-April none of the

years, classified as Low, required *waiting volumes*, 6 of 31 years, classified as Average, needed *waiting volumes*, and 10 of 17, classified as High required *waiting volumes*.

The methodology proposed to change the *envelope curve* in March-April considers that the 12,000 synthetic daily inflows series is formed by three sets of synthetic series, each one conditioned on a class of the February monthly flow. The number of series in each set is proportional to the number of historical inflows series in its class. Then, the *envelope curve* of each class is calculated separately considering just the years that belong to each class. The *original envelope curve* (from November to February) is calculated considering all the synthetic series disregarding the ones that will be not protected. The *envelope curves* obtained for the period March-April for each class indicates: (i) Increased necessity of *waiting volumes* when the February monthly flow is classified as High; (ii) Decreased necessity of *waiting volumes* when the February monthly flow is classified as Low and anticipation in 3 weeks the recover; and (iii) for the Average class the recovery is anticipated in one week and the necessity of *waiting volumes* is decreased in the majority of the period.

Based on the results, this methodology seems to be promising to increase the probability to get the end of the rainy season with the reservoirs full, minimizing the conflict of the reservoirs uses for flood control and hydroelectric generation. Others improvements are being incorporated in this methodology as considering the inflows forecasting and an alternative form to calculate the *critical path*.

References

- [1] Costa, F.S., Damázio, J. M., Neves, F. P., et al, Sistema SPEC – Sistema para Estudos de Prevenção de Cheias em Sistemas Hidroelétricos. *XV Water Resources Brazilian Symposium*, Brazil, 1999. (In Portuguese)
- [2] Costa, F.S., Damázio, J.M., Ghirardi, A.O., Incorporação de Tendências Macro-climáticas na Operação de Controle de Cheias. *CEPEL technical report*, Brazil, 1996. (In Portuguese)
- [3] Raupp, I. P., Minimização do Conflito entra a Geração de Energia Elétrica e o Controle de Cheias no Planejamento da Operação de Reservatórios, *Msc Thesis*, COPPE-UFRJ, Brazil, 2008. (In Portuguese)
- [4] Damázio, J.M., Condições de Controlabilidade de Sistemas de Reservatórios para Controle de Cheias e seu Uso na Operação de Sistemas com Múltiplos Usos, *Dsc Thesis*, COPPE-UFRJ, Brazil, 1988. (In Portuguese)
- [5] Damázio, J., M., Marien, I., Costa, F.S., Building Flood Control Rule Curves for Multipurpose Multireservoir System Using Controllability Condition, *Water Resources Research*, Vol.30, n.04, Pp.1135-1144, 1994.
- [6] Kelman, J., Damázio, J.M., Costa, J.P., Geração de Séries Sintéticas de Vazões Diárias – Modelo Diana, *Revista Brasileira de Engenharia*, Vol.1, Nº 2, Pp. 5-22, 1983. (In Portuguese)
- [7] Braun, J., Maindonald, J., Data Analysis and Graphics Using R, *Cambridge Series in Statistical and Probabilistic Mathematics*, Cambridge Press, 2003.
- [8] Mathsoft, S-Plus 2000, Guide to Statistics Volume I, *Data Analysis Products Division*, MathSoft, Seattle, WA, 1999.

