Definition of the Streamflow Scenario Tree for Long-Term Operation Planning Studies of Hydrothermal Power Generation System

Débora D.J. Penna[1]; *Jorge Machado Damázio*[1,2]; *Maria Elvira P. Maceira*[1,2]

[1] Electric Power Research Center – CEPEL, CP 68007 Rio de Janeiro, Brazil.
[2] Rio de Janeiro State University, São Francisco Xavier 525, Rio de Janeiro, Brazil.

Introduction

The long-term operation planning model of the Brazilian interconnected system defines for each month of the planning horizon (5 to 10 years) the strategy of optimal allocation of hydro and thermal resources. This strategy is defined by a policy of optimal operation represented by an expected cost-to-go function estimated by stochastic dual dynamic programming .

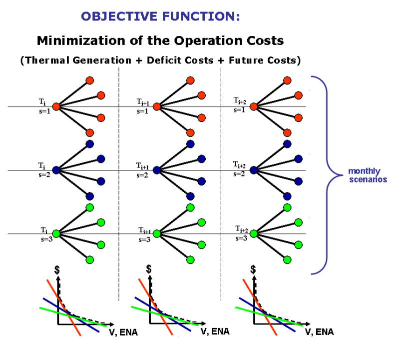

Figure 1- Scheme of long-term operation planning

The streamflow stochasticity is represented by a large number of hydrological scenarios synthetically generated. The scenario generator model uses the periodic autoregressive modelling approach referred as **PAR(p)** model.

The set of all possible realizations of the streamflow stochastic process throughout the planning horizon forms a scenario tree. This tree represents the entire probabilistic universe on which the optimal operation strategies are calculated. As the scenario tree of the long-term operation planning problem has a high cardinality, in practice it is impossible to visit the complete tree due to computational effort. Therefore, only a portion of the tree (sub-tree) is covered.

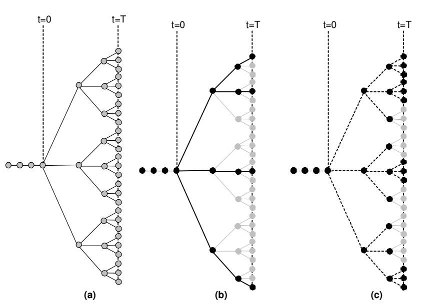

Figure 2- Scheme of Streamflow Scenario Tree
(a) Complete (b) Sub-tree Forward (c) Sub-tree Backward

This work describes a method, so called **Selective Sampling**, for defining the scenario sub-tree in order to make more robust the results from this operation policy with regard to variations in the sample of inflow scenarios.

Two proposals for definition of the sub-tree are analyzed:

- substitution of the simple random sampling (**SRS**) by latin hypercube sampling (**LHC**) in the multivariate streamflow scenario generation model
- substitution of the SRS by selective sampling (**SS**) in the multivariate streamflow scenario generation model

Methodology

The PAR(p) model describes the standardized streamflow at time period t, as a linear combination of the standardized streamflows in periods t-1, t-2, …, $t\text{-}p_t$ summed to a random noise.

$$\frac{z_t-\mu_t}{\sigma_t}=\phi_1^t\left(\frac{z_{t-1}-\mu_{t-1}}{\sigma_{t-1}}\right)+\ldots+\phi_{p_t}^t\left(\frac{z_{t-p_t}-\mu_{t-p_t}}{\sigma_{t-p_t}}\right)+a_t$$

In **Current Option**, the sample of random noises is selected using the Monte-Carlo method with classical SRS from a multivariate lognormal distribution. The forward sample noise is obtained by conditioned draws from backward sample and the scenarios are equiprobables.

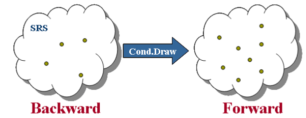

Figure 3- Current Option

The **Selective Sampling** method consists in application of the clustering techniques in the process of generation streamflow scenarios for forward and backward simulations. In this case, the clustering techniques are used to obtain samples of multivariate noises used by the PAR(p). The chosen method is the non-hierarchical K-Means method.

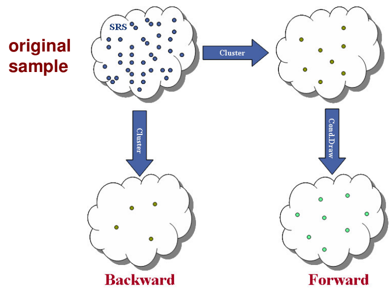

Figure 4- Scheme of Selective Sampling

The probability of a streamflow scenario is equal to the probability of the vector of noises from which it was created. The hydrological scenarios for the forward simulation will be obtained by conditioned draws from probability distributions resulting from the clustering procedure, and therefore they will always be treated as equiprobables.

The **LHC Sampling** aims a partial control of the sampling process, thus resulting in a reduction of the variance of the estimators. It will be also used to generate the noises of the PAR(p) model of hydrologic scenarios in the backward and forward steps. The forward and backward samples of noises are generated independently.

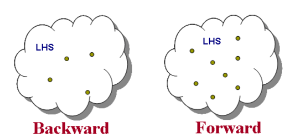

Figure 5 - Scheme of LHC Sampling

Results

The alternatives proposed to build the scenario tree were tested in a real case study, considering a system configuration formed by more than 100 hydro plants displayed in several cascades (aggregate) and the generation of 10 years ahead monthly streamflows, adopting 200 forward scenarios and 20 openings (backward scenarios). The several reservoirs are aggregated in 4 equivalent energy reservoirs.

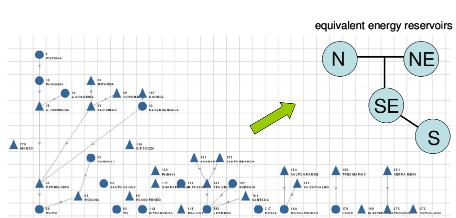

Figure 6 – System Configuration

It was used an original sample with 2000 noise vectors for the clustering process. It is important to emphasize that the computational time spent on the scenario tree construction grows considerably as the size of original sample of noise vectors used in the clustering process increases.

It was defined as relative deviation of a statistic $\hat\theta$ (mean, standard deviation etc.) the difference between the estimate and its expected value divided by the standard deviation of the estimate.

$$\frac{\hat\theta-\mu_\theta}{\sigma_\theta}\sim t_n$$

The adopted significance level for statistical tests is 5% and as population parameters are considered the historical values (non-conditional tests). In Figure 7 are presented the relative deviations of mean and standard deviation.

It could be observed that the current option presents substantial variability in the mean and standard deviation. This variability is clearly reduced when the clustering technique is applied in forward and backward samples. The LHC method also reduces the variability of mean and standard deviation.

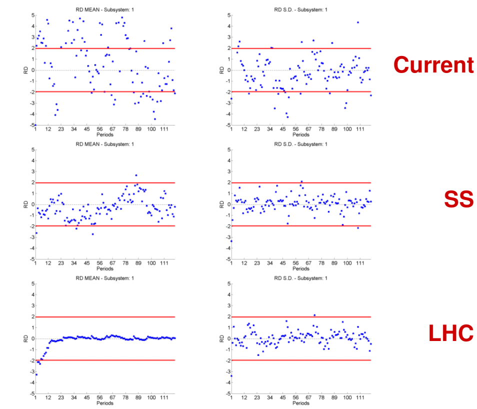

Figure 7 - Validation of the Generated Scenarios

The average values of cross correlation for forward sample are shown in Figure 8. The hydrological scenarios obtained with LHC method have a smaller cross correlation. The cross correlation values obtained using selective sampling is closer than historical values in comparison with other methods.

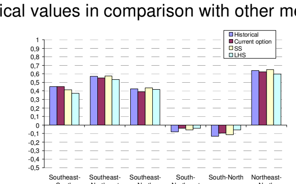

Figure 8 – Cross-correlation

In Figure 9, it is shown the expected value of the total operation cost (ZSUP), its estimated value (ZINF) and confidence interval of 95% for the last iteration of the convergence process. It was done a sensitivity analysis with respect to the sampling of scenarios. It was used different samples for forward and backward steps.

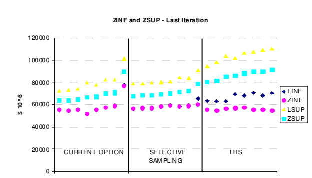

Figure 9 – ZINF and ZSUP (varying sample)

The variations of ZINF and ZSUP values for different samples are smaller in selective sampling than the in current option. This indicates that the use of clustering techniques in the scenario construction also provides greater robustness. It can also be seen that the results using the LHC method does not vary significantly with different noise samples, but can be observed a slight increase in the ZSUP values when compared with the same values obtained with other methods.

Conclusions

The scenarios representation of the historical experience with respect to the monthly means and standard deviations of energy inflows was very good for proposed methods and the variability of the mean and standard deviation was reduced. The cross-correlation was well preserved when using the selective sampling method, but it there was a small degradation with the LHC method. With respect to the sensitivity analysis, we can conclude that the application of LHC and SS make the scenario sample more robust in terms of variations in values of ZSUP and ZINF due to variations in sample scenarios. ZSUP values have increased when the LHC method was used.

E-mail for contact: newave@cepel.br
Site: www.cepel.br