STAHY - Workshop Advances in statistical hydrology

Taormina May 23-25, 2010

PREVIVAZ – Improving Weekly Streamflow Time Series Forecasts With the Current Hydrologic State of the River Basin

S. A. Souza (CEPEL), F. S. Costa (CEPEL/UERJ), J. M. Damázio (CEPEL/UERJ), M. E. P. Maceira (CEPEL/UERJ), L. N. R. Xavier (CEPEL)

Department of Energy Optimization and Environment - Electric Power Research Center - CEPEL

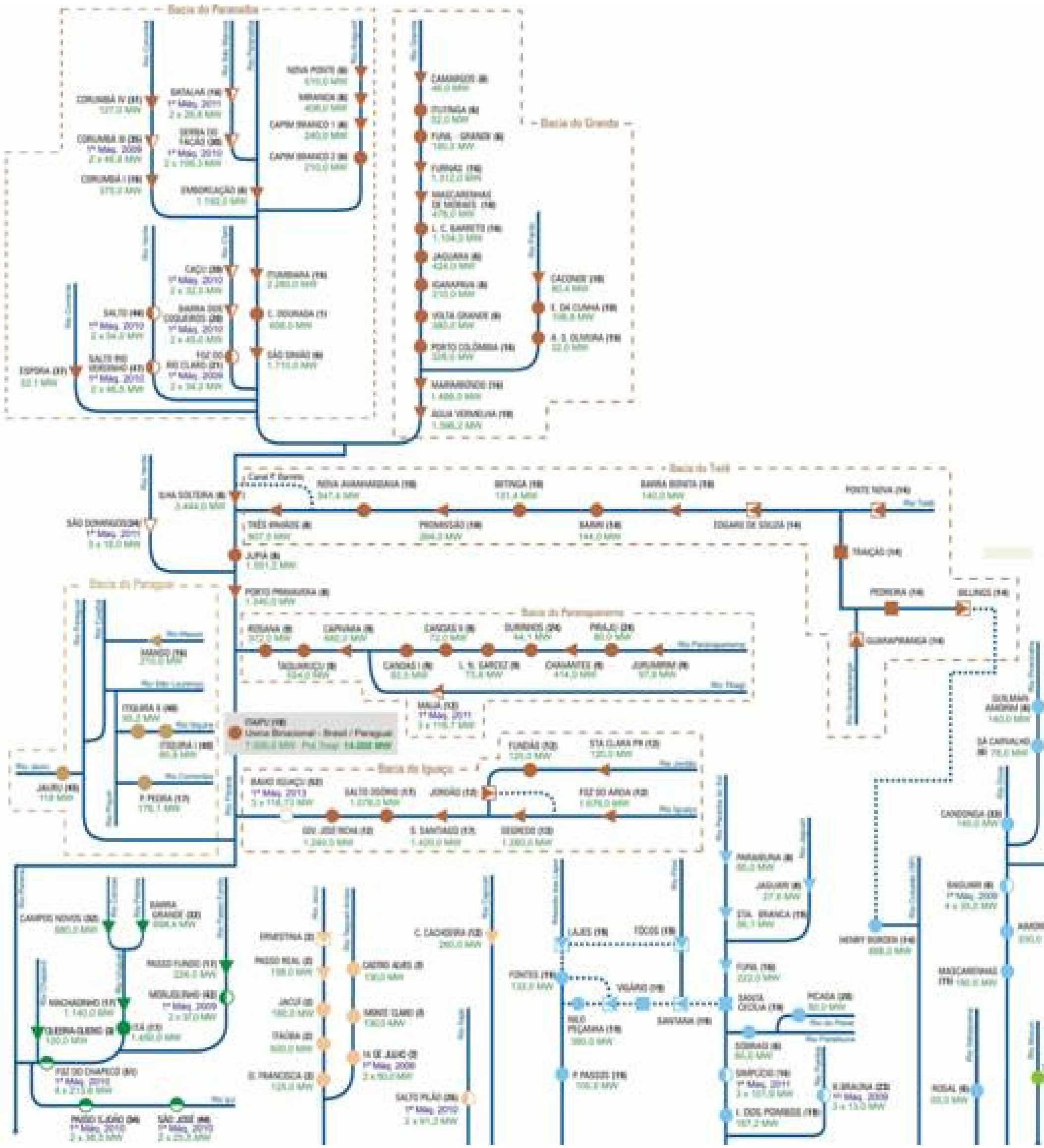
OPERATION PLANNING OF BRAZILIAN HYDRO THERMAL SYSTEM

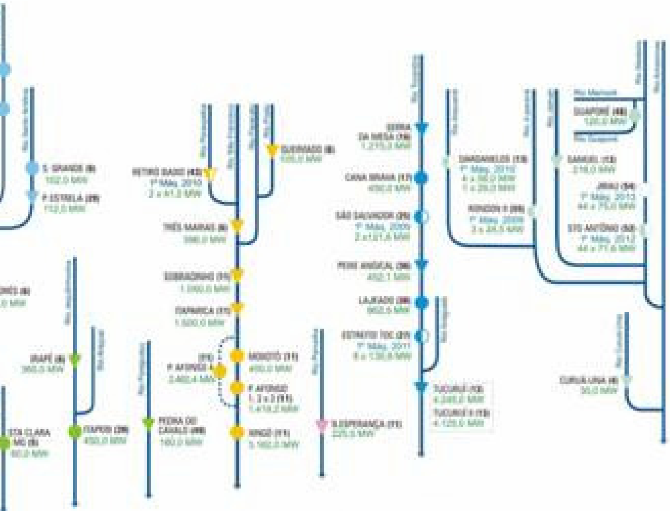
CEPEL 35 anos
Sistema Eletrobrás

- Hydraulic based (> 80% Total Power Generation)
  - 12 large watersheds (cascades)
  - Many large reservoirs **(> 1 km^3)**
    - Inter-annual streamflow regulation
    - Other uses (e.g.: Flood control)

- Planning problem
  - Load supply with **minimum expected total cost** (Thermal Units operation + deficits)
  - Energetic Modeling System
    - Optimization models (planning, dispatch)
    - **Hydrological models (streamflow forecasting & scenarios generation**

OPERATION PLANNING OF BRAZILIAN HYDRO THERMAL SYSTEM





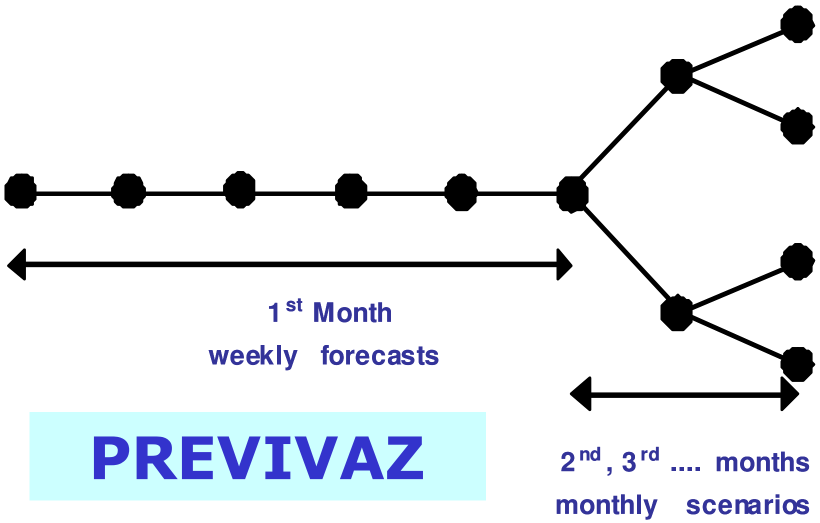
OPERATION PLANNING OF BRAZILIAN HYDRO THERMAL SYSTEM

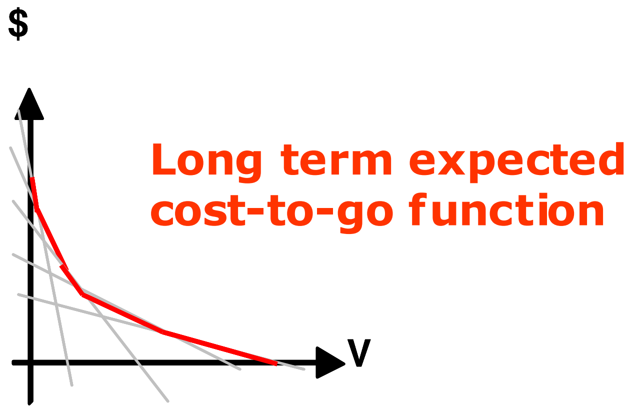
CEPEL 35 anos
Sistema Eletrobrás

Inflow representation	Hydrological models		Planning Horizon
Monthly inflow scenarios	GEVAZP	NEWAVE	5 years
Weekly inflow forecasts Monthly inflow scenarios	PREVIVAZ GEVAZP	DECOMP	2 months
Daily inflow forecasts	PREVIVAZH	DESSEM	2 weeks

OPERATION PLANNING OF BRAZILIAN HYDRO THERMAL SYSTEM

- (Short-term) Weekly Operation Scheduling
  - Operational constraints
  - Initial reservoirs' storage
  - Last weeks observed inflows
  - Future inflows: Weekly forecasts + monthly scenarios





PREVIVAZ MODEL

- Stochastic univariate model
- Weekly streamflow forecasts (up to 6 weeks ahead)
- Used in short-term Planning Operation (DECOMP model)

PREVIVAZ model consider **different modeling strategies** based on:

- Linear ARMA(p,q) models, periodic/non-periodic
- Different **estimation methods** (moments, regression)
- Series transformation (Box & Cox)
- Best model: select by RMSE (cross-validation)

PREVIVAZ MODEL

- Models evaluated by PREVIVAZ

Forecasting Model	Characteristic	Estimation Method
Constant	*Annual average*	*Moments*
Seasonal	*Weekly average*	
AR(p) $(1 \leq p \leq 4)$	*Stationary correlation structure*	
ARMA(p,1) $(1 \leq p \leq 3)$		
PAR(p) $(1 \leq p \leq 4)$	*Seasonal correlation structure (weekly, monthly, quarterly, semi-annual)*	*Moments and regression*
PARMA(p,1) $(1 \leq p \leq 3)$		

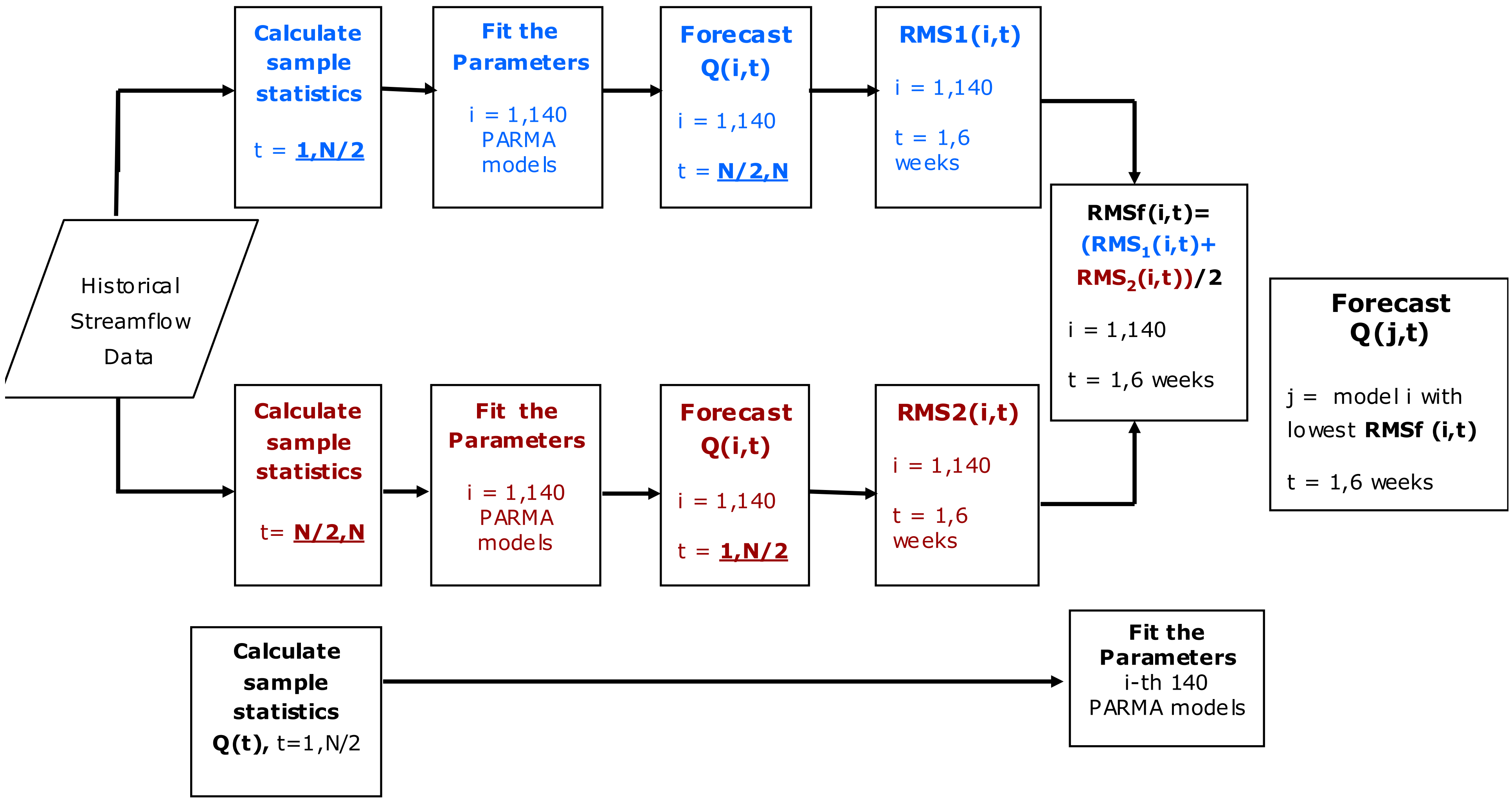
=> Up to 140 models (including B&C transformation)

PREVIVAZ: estimation procedure

1. Evaluate the discharge series statistics for week t;
2. Taking into account these series statistics, estimate M different linear PARMA models (up to 140);
3. For the i-th PARMA model, forecast Qi(t) on the period spanning the first half of historical record and calculate the mean RMS error over this period, $RMS_1 i$; repeat this for the M alternative PARMA models;
4. The same as 3, but for the period spanning the second half of historical record;
5. Evaluate each PARMA Model forecast, i=1,..,M, taking into account their mean RMS error, RMSi (average of $RMS_1 i$ and $RMS_2 i$);
6. Rank the M alternative models with respect to their average RMSi errors; chose the best one (k = 1);
7. With the model chosen in Step 6, forecast the discharge Q*(t) to week t and calculate the forecast confidence intervals;
8. Repeat the procedure describe in Steps 1 to 7 for the next weeks (t+1 to t+5). The last weeks (t,…,t-4) forecast will be considered as trend to the model’s choice for the next weeks(t+1,..,t+5)

PREVIVAZ MODEL

PREVIVAZ: estimation procedure



PREVIVAZ MODEL

- Problem: some **great errors** caused by
  - Unlikely large/small forecasts obtained on calibration period (inducing the rejection of an alternative model)
    - Weak time correlation structure
    - Last observed values statistics different from the historical record ones (short record?)
  - Unlikely large streamflow forecasts after inversion (from B&C transformed values)

=> Unlikely Q(t) variations

How could these unlikely forecasts be eliminated on estimation procedure?

PREVIVAZ MODEL

How could these unlikely forecasts be rejected on estimation procedure?

- Approach: **to limit the Q(t) variations between two consecutive weeks:**
  - Analyzing the Q(t)/Q(t-1) *cumulative distribution function* (historical record)
  - Defining probabilities of **minimum/maximum acceptable** Q(t) variation
  - With the last observed Q(t-1) values, defining a **range (bounds) of acceptable** Q(t) forecasts
  - Once this range has been defined, taking it into account on model selection procedure

PREVIVAZ MODEL

CEPEL 35 anos
Sistema Eletrobrás

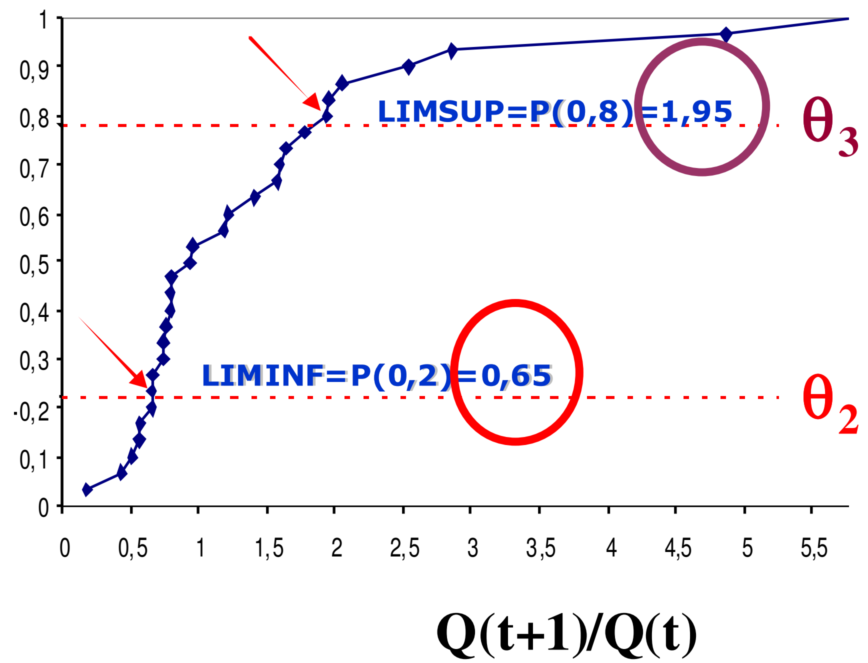
How the forecast bounds can be defined?

1) Build the Q(t+1)/Q(t) *cdf* (from historical record), having defined:

i) The **period** (week, month etc) where the *cdf* will be estimated;

ii) The θ_1 ranges of **Q(t) magnitude** to which the *cdf* will be estimated

2) Establish the minimum/ maximum **probabilities of non-exceedance** $\{\theta_2, \theta_3\}$



Thus, with the **Q(t-1) value** no week t-1:

Lower bound = Q(t-1)*(**value** corresponding to the θ_2 percentile)

Upper bound = Q(t-1)*(**value** corresponding to the θ_3 percentile)

How the forecast bounds can be defined?

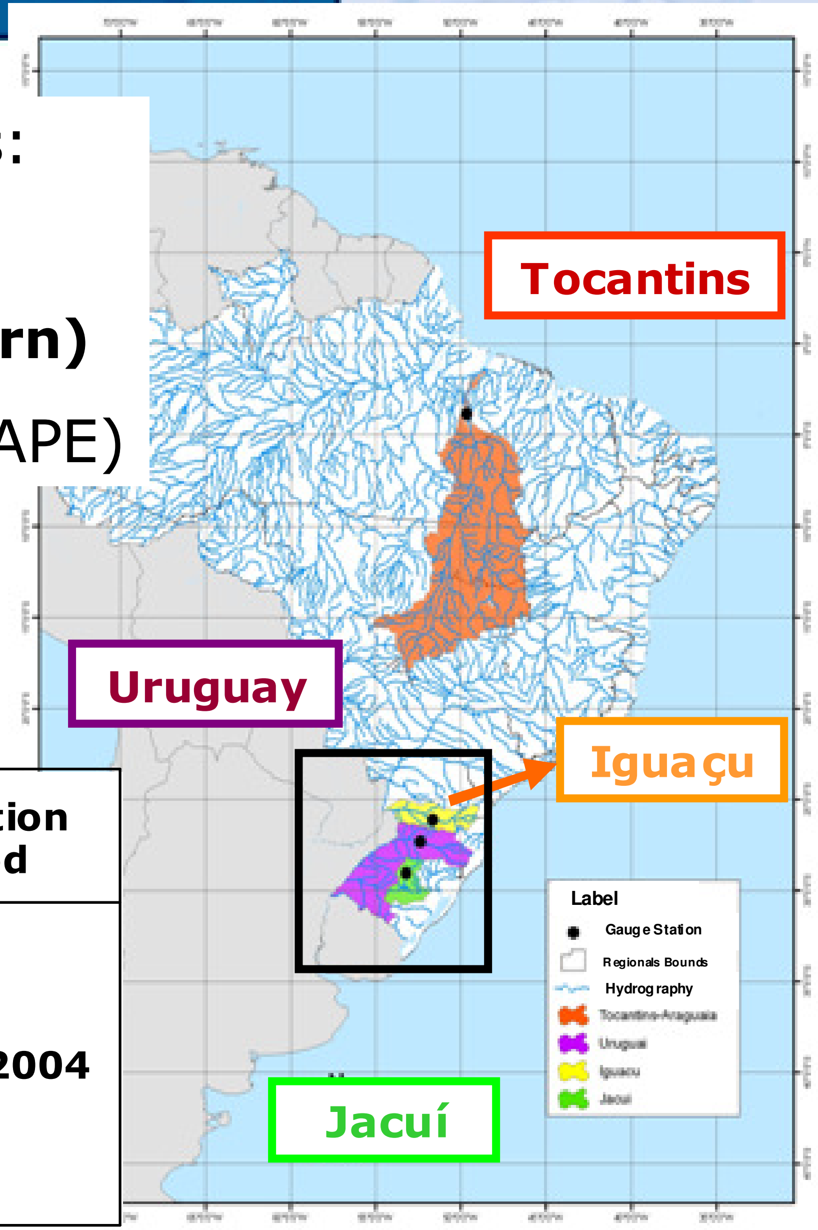
1. Define (i) the period on which the cdf will be estimated, (ii) the q1 streamflow magnitude ranges, (iii) the minimum/maximum allowed variation probabilities {q2,q3};
2. With the procedure described in Item 2.1 calculate the i=1,…,M alternative PARMA models to the week t and rank (k=1,…M) them following their average RMSi errors;
3. Estimate the q1 variation ratio (Q(t)/Q(t-1)) cdf's from the historical inflow time series;
4. From these q1 cdf's and the {q2,q3} values, calculate the q1 discharge variation bounds {LBk(t),UBk(t)};
5. For the k=1,…,M alternative models, calculate Qk(t) and verify if it lies in the correspondent (according to the magnitude of Qk(t)) bounds [LBk(t),UBk(t)]; if not, Qk(t) will be replaced by the lower or upper bound, and the their RMSk error, re-evaluated; a new ranking (k'=1,…,M) is then performed;
6. Taking into account the whole time series, and the ranking done in Step 5, evaluate if the best model (k'=1) forecast Q1(t) lies in [LB1(t),UB1(t)];
7. If the Step 6 verification does not hold, take the second best model (k'=2) and perform again the Step 5 verification; if this second verification does not hold neither, take the third best model (k'=3) forecast Q3(t) and so on, until finding the best model (k* = k') whose forecast respect the bounds defined in Step 4 and that will be retained thereafter;
8. Repeat the procedure described in Steps 1 to 6 to the next 5 weeks (t+1,…,t+5).

CASE STUDY

Gauge Stations in 4 Brazilian Basins:

- Tocantins (Northern)
- Jacuí, Uruguay and **Iguaçu (Southern)**

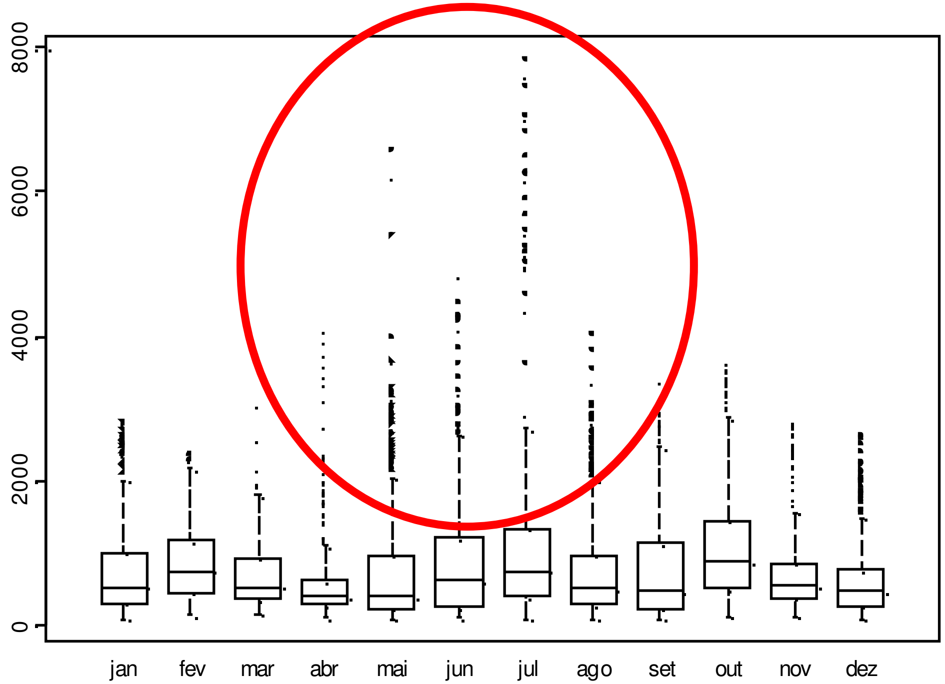
Different goodness-of-fit criteria (MAPE)

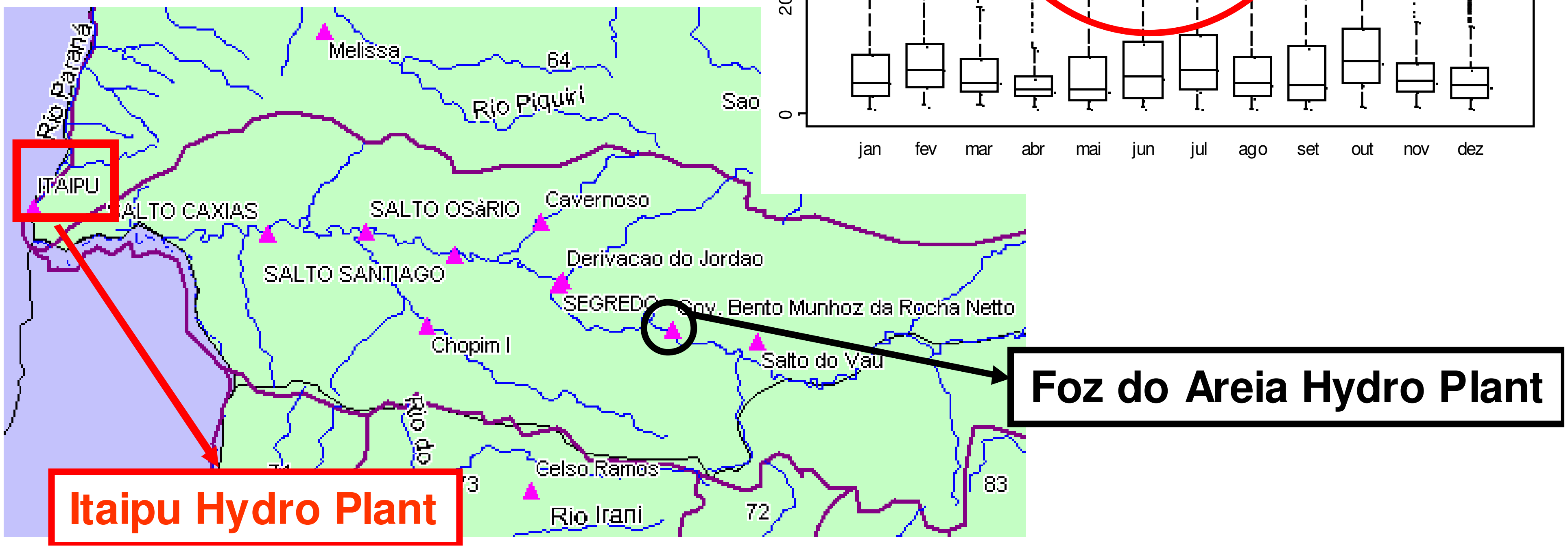


Watershed	Gauge station	Estimation period	Validation period
Iguaçu	**Foz do Areia**	**1969 – 1994**	**1995 – 2004**
Uruguay	**Itá**	**1973 – 1994**	
Jacuí	**Passo Real**	**1940 – 1994**	
Tocantins	**Tucuruí**	**1970 – 1994**	

CASE STUDY – Foz do Areia

Mean	**780 m³/s**
S.Deviation	**710 m³/s**
Maximum	**7830 m³/s**
Minimum	**68 m³/s**





CASE STUDY – Foz do Areia

- Foz do Areia gauge station
  - 1969 – 1994 : parameter estimation
  - 1995 – 2004 : validation

Bounds

- *Cdf* estimation: weekly, monthly, quarterly and semi-annual;
- 4 Q(t) magnitude ranges: {0;25%;50%;75%;100%};
- Minima/maxima probabilities: {10%;90%}, {20%;80%} and {30%;60%}

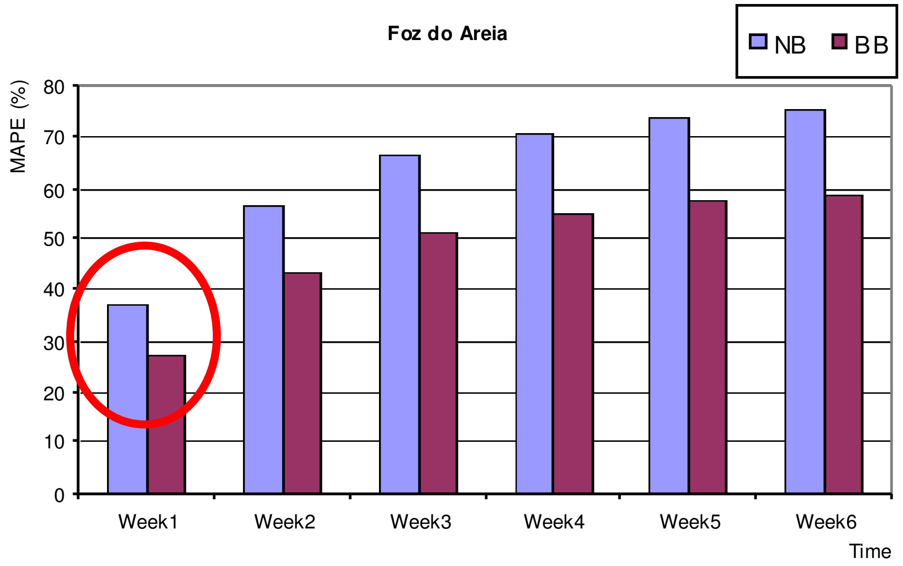
=> **48 bounds parameterizations to be evaluated**

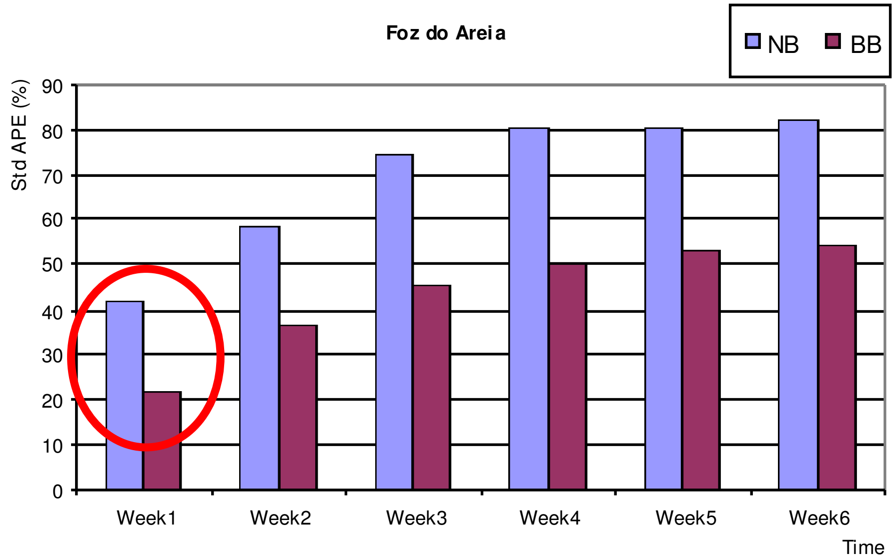
CASE STUDY – Foz do Areia

Goodness of fit criterion: Mean Absolute Percentage Error

Best bounds parameterization leads to:

- Average error reduction: **17.4 %** (6 weeks)
- Great reduction of Error standard deviation : **43.1%** (6 weeks)





Reduction : 26.4 % (1st Week)

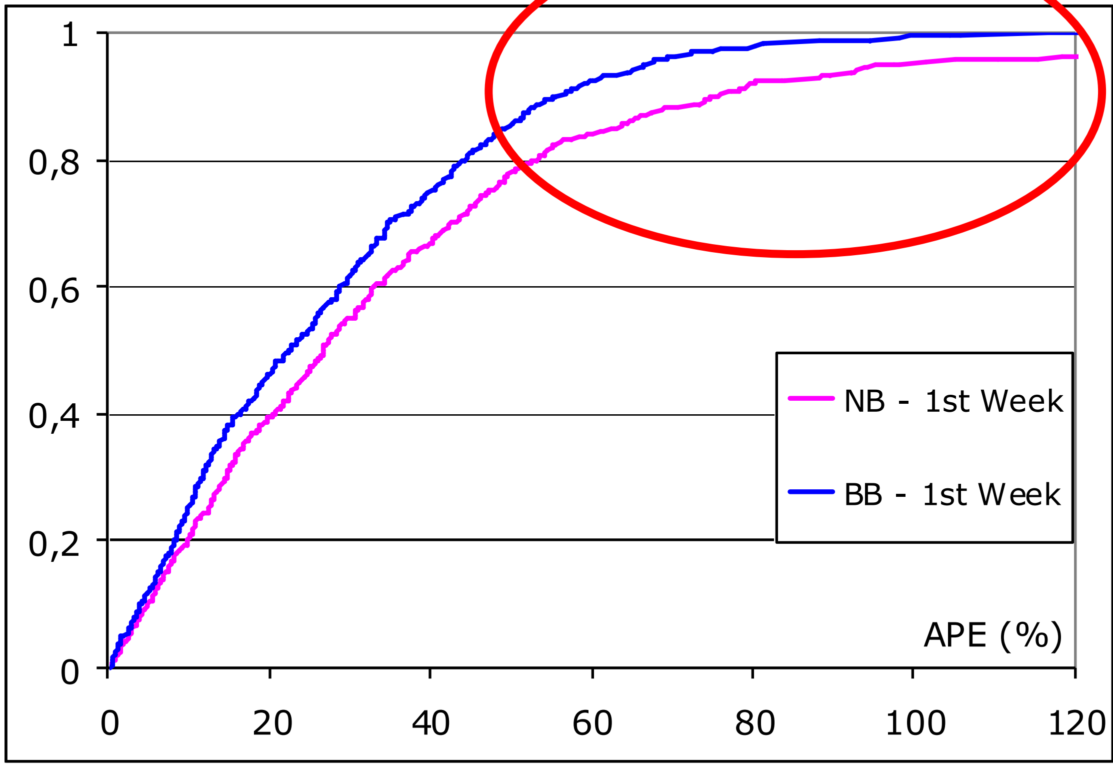
Reduction : 50.2 % (1st Week)

CASE STUDY – Foz do Areia

Goodness of fit criterion: Mean Absolut Percentage Error

Best bounds parameterization leads to:

- Remarkable reduction of large errors (over the whole period)



Week	Q25%	Q50%	Q75%	Q95 %
1	13,85	9,98	16,12	**33,91**
2	6,65	10,08	12,26	**31,42**
3	4,17	4,79	10,57	**31,26**
4	-4,05	1,98	14,56	**32,78**
5	0,76	-2,32	14,55	**28,75**
6	11,03	5,80	11,40	**28,27**

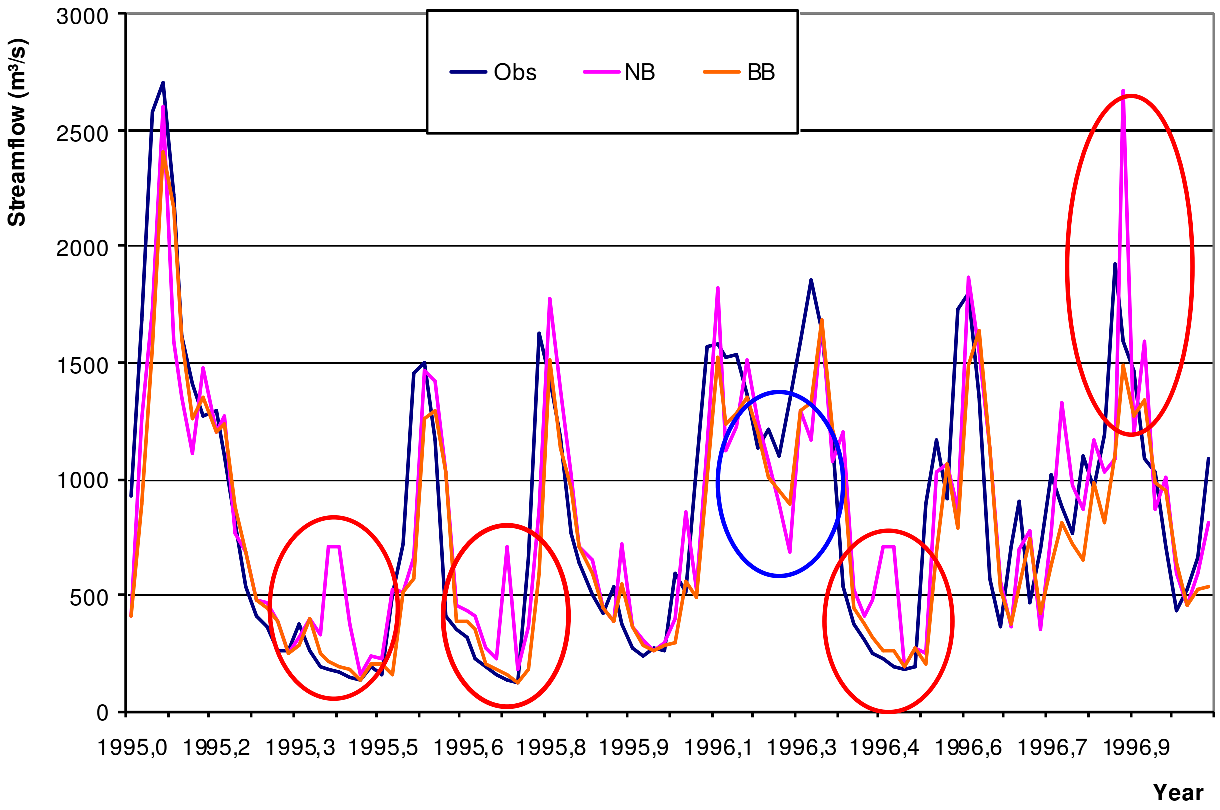
CASE STUDY – Foz do Areia

CEPEL
Sistema Eletrobrás

Goodness of fit criterion: Mean Absolut Percentage Error

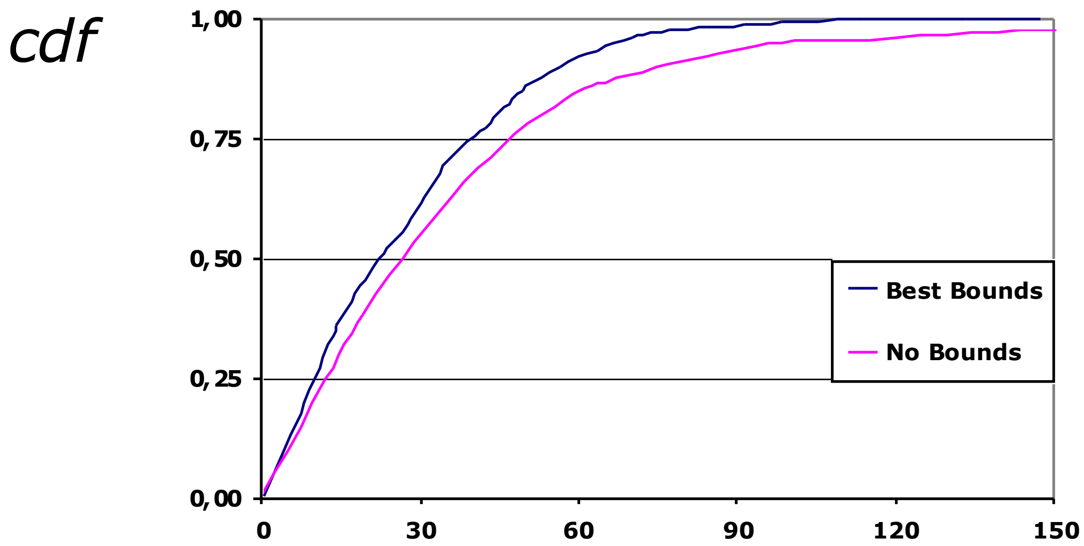
Best bounds parameterization leads to:

- Reduction of errors on recession periods



CASE STUDY – Foz do Areia

1st Week Statistics



1st Week	Mean	St. Dev.	Q25%	Q50%	Q75%	Q95%
No Bounds	37,16	42,07	12,43	26,79	47,32	98,26
Best Bounds	27,35	20,94	10,71	24,12	39,69	64,95
Reduction %	**26,39**	**50,23**	**13,85**	**9,98**	**16,12**	**33,91**

CONCLUDING REMARKS

- Considering the bounds on model selection led to smaller mean forecast errors over the validation period;

- A remarkable reduction of large errors has been obtained;

- The error reduction were observed to the 6 weeks;

- Similar results were obtained with the other metrics (RMSE, Nash-Sutcliffe index) considered in our study;

- Similar results were obtained to the other basins, especially to the other 2 basins in Southern Brazil (Uruguay and Jacuí);

- Other results (not presented) showed that the application of this approach to other watersheds whose hydrological features are similar to the ones considered herein leads to a similar error reduction pattern.

ONGOING WORK AND FUTURE DEVELOPMENTS

- Development of an automatic procedure aiming to define the “best bounds parameterization” to each series, taking into account:
    - Antecedent (wet/dry) conditions
        - Precipitation ; Moisture (estimated)
    - Seasonal hydrological response patterns
- Integrate daily forecasting (conceptual & stochastic) model results in 1st week PREVIVAZ modeling;
    - with forecasted “some-days ahead” precipitation
- Investigate the constraining of last weeks (5th/6th) PREVIVAZ modeling by a monthly forecasting model results.
    - Scenarios generation in 5th/6th weeks ahead?

Thank you !!!

Contact : **previvaz@cepel.br**