Inflows, wind and price coherent scenario simulation

VEIGA, Álvaro, PUC-Rio
SOUTO, Mário, PUC-Rio
PEREIRA, Guilherme, PUC-Rio and Technische Universität München

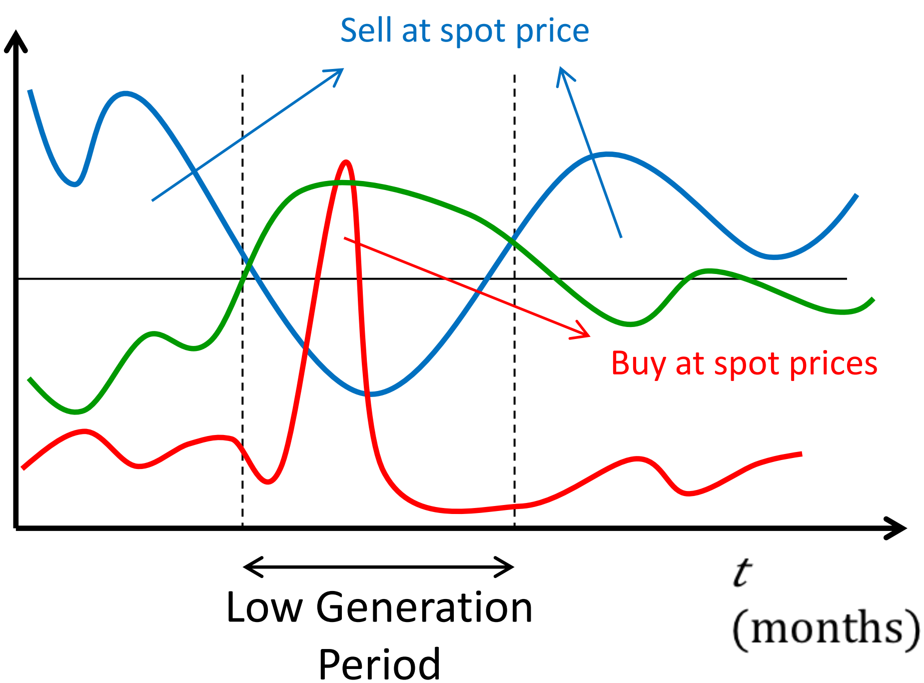
Agenda

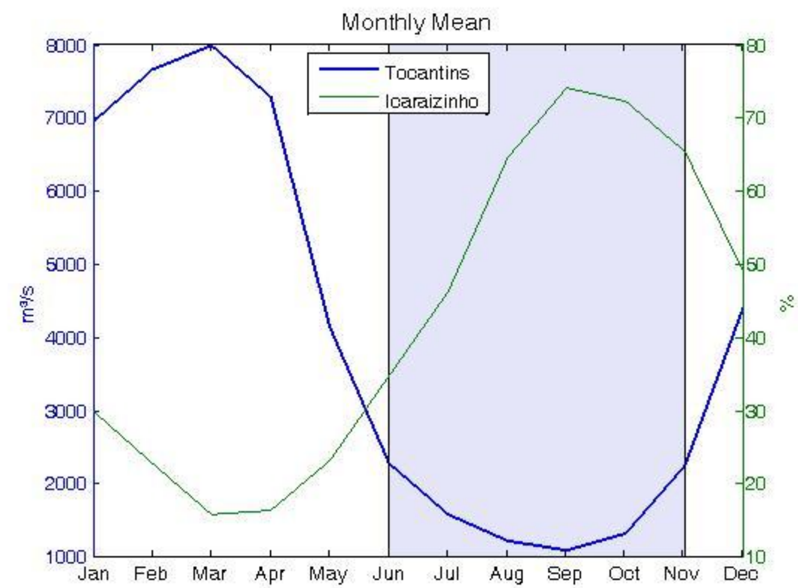
- Motivation
  - Coherent simulation of inflows, wind and prices
  - Connection variables are the ENA's
    - ENA'S → price DETERMINISTIC
    - ENAS → WIND/inflows STOCHASTIC
- Basic concepts
  - Main features of the series
    - Seasonality in the mean variance and covariances
    - Valid values → transformation
  - Variable selection in large dimension context: Lasso
- The model: VARX with incomplete rank seasonal VC matrix
- LASSO Maximum likelihood Estimation method
- Simulating wind, inflows and price
  - Bootstrap
  - Copulas

Motivation

- Renewable in Brazil (2014)
  - Solar: insignificant
  - Wind: 167 Farms , more ~400 licenses for future.
  - Run-of-the-river (Small Hydros): 463
- Clean energy but….intermittent = Problems
  - Expansion planning
  - Dispatch
  - Commercialization
- Barrier to expansion: Financial RISK
- Goal: to form portfolios of small hydros and wind farms with lower risk
  - Scenario generation methods for hydros, wind and price
  - Stochastic optimization with risk constrain → win-win contracts

Stochastic behavior of renewable power plants, with seasonal and complimentary behavior:



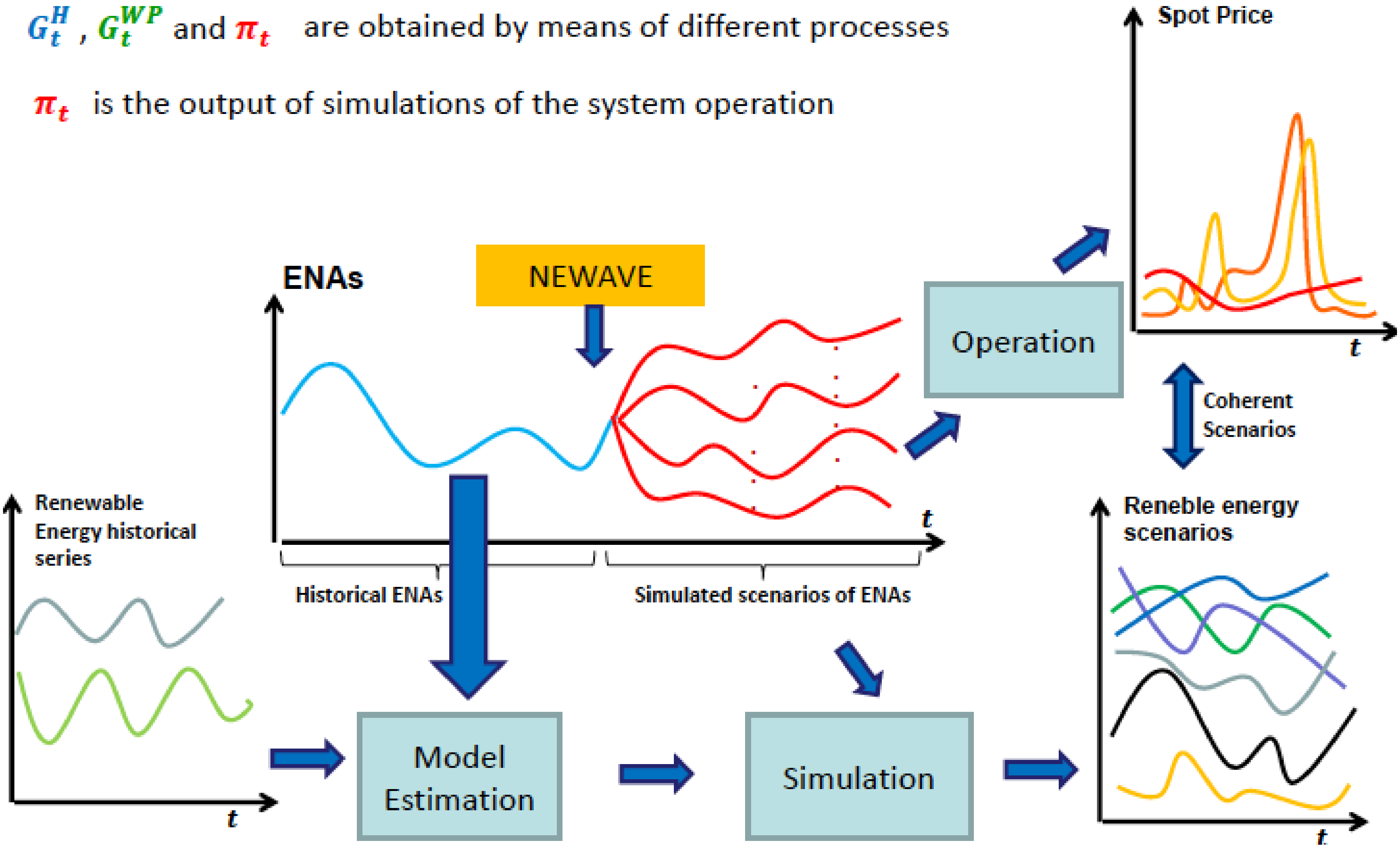


Thus, we need coherent scenarios of G_t^H, G_t^{WP} and π_t

Structure of the model

G_t^H , G_t^{WP} and π_t are obtained by means of different processes

π_t is the output of simulations of the system operation



Data for inflows and wind

- **Inflows: public access data bases:** 70 years of monthly data
  - Relationship to production is approximately linear
  - Positive values , no zeros values.
- **Wind Public data bases:** AMA system (EPE, www.epe.gov.br)
  - Potential evaluation: http://www.epe.gov.br/mercado/Documents/Série%20Estudos%20de%20Energia/20130925_1.pdf
  - 63 stations (>= 9 months in 03/2013, 28 months in 08/2014), 10 min.
    - North East: 30 (78m to 100m), Bahia (81m to 120m), South (78m to 121m)
    - Pression, direction, speed, temperature, humidity
- **Wind Private data bases**
  - Usually 2 to 4 years 10 min interval
  - Aggregated effective production is published monthly by ONS
    - Operational data unavailable → not all turbines are operational all the time
  - Transformation **wind → production** is non-linear: temperature, topology, configuration, wind direction….
  - Solution:
    - Wind data construction from satellite data/physical models --> rough approx.
    - Wind data extension from near by measured wind + satellite.
  - Values between 0 and nominal capacity: Capacity factor between 0 and 1
  - Zeros can appear in high frequency (from daily to 10min)

  → **We use the monthly wind production as the variable to model**
- **Satellite data + Climate modelling :** ERA-INTERIM , NNRP

LAMPS

Data trasformation

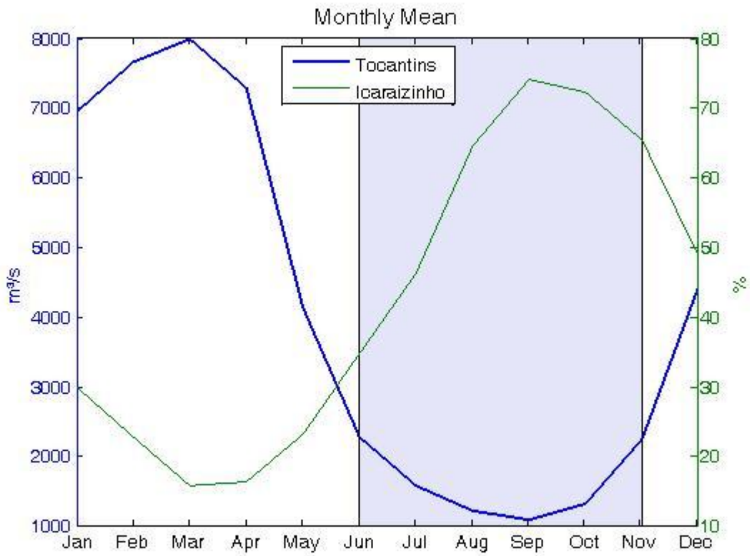
- Wind energy production (WEP) and small hydro production (SHP) are non negative variables.
  - Wind
    - 0 < WEP < Max → Capaticy factor (CFW) = WEP/Max → 0< CF < 1
  - Run-of-the-river
    - 0 < SHP < Max → Capaticy factor (C) = SHP/Max → 0< CF < 1
- Scenarios should respect this constrains
  - Transformation:

$$X_t = ln\left(\frac{FC_t - Min_{m(t)}}{Max_{m(t)} - FC_t}\right).$$

  - m(t): month of instant t
  - $Min_{m(t)}$ and $Max_{m(t)}$ are defined heuristically for each month in order to get a approximately symmetric distribution for X_t.

LAMPS

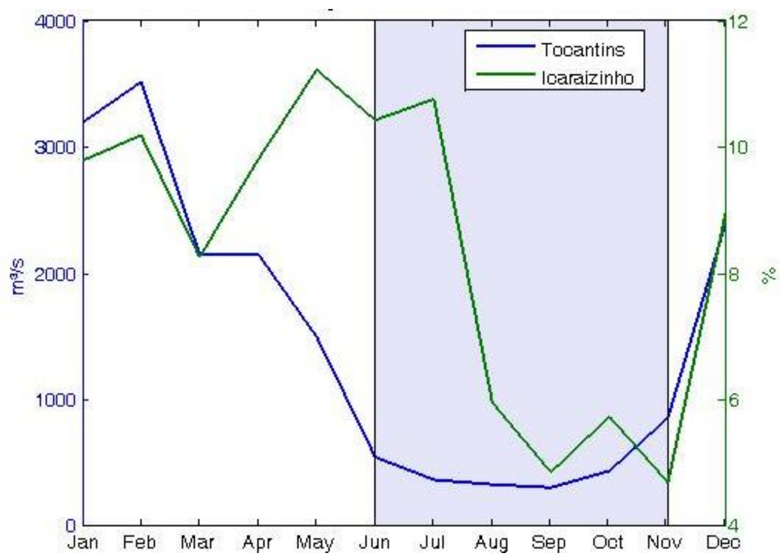
Seasonal mean, variance and covariance

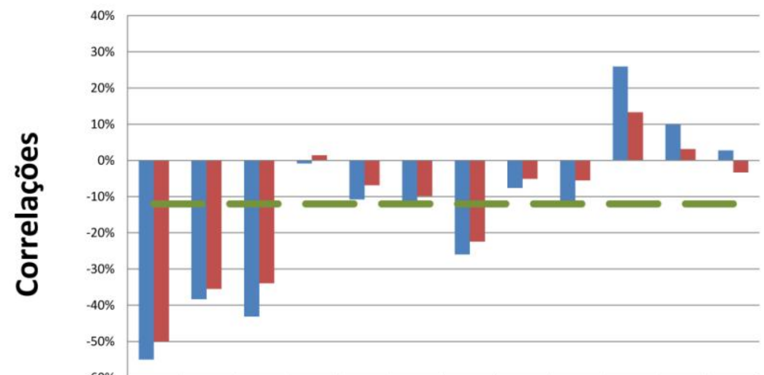


← MEAN

← VARIANCE

← CORRELATION





	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Não-Heurístico	-54,97%	-38,32%	-43,11%	-0,82%	-10,78%	-12,25%	-25,96%	-7,58%	-11,04%	25,97%	9,91%	2,77%
Heurístico	-50,00%	-35,46%	-33,94%	1,42%	-6,85%	-9,82%	-22,45%	-5,08%	-5,50%	13,34%	3,14%	-3,30%

LAMPS

Stochastic model: VARX model with periodic variance

$$Y_t = c + \sum_{i=1}^{p} \Phi_i Y_{t-i} + \sum_{j=0}^{q} \Theta_j X_{t-j} + A_{m(t)}\varepsilon_t \;\; \forall\, t = 1, \ldots, T$$

$$\varepsilon_t \sim \mathcal{N}(0, I) \qquad \Sigma_{m(t)} = A_{m(t)} A_{m(t)}^\top \qquad m(t) \in \{1, 2, \ldots, 12\}$$

$Y_t \in \mathbb{R}^k$, k is the number of renewable power plants.

$X_t \in \mathbb{R}^r$, r is the number of exogenous variables (ENAs).

$c \in \mathbb{R}^k$, vector of intercepts

Φ_i , Matrix (k x k) coefficients of autoregressive components

Θ_j , Matrix (k x r) coefficients of exogenous variables

Σ , Matrix (k x k) covariance matrix

LAMPS

Stochastic model: VARX model with periodic variance

Defining:

$$\beta = vec([c \;\; \Phi_1 \;\; \Phi_2 \;\; ... \;\; \Phi_{12} \;\; \Theta_1 \;\;\; ... \;\; \Theta_q])$$

$$W_{t-1} = [1^\top \; Y_{t-1}^\top \; Y_{t-2}^\top \; \cdots \; Y_{t-12}^\top \; X_t^\top \; X_{t+1-q}^\top] \otimes I_{k \times k}$$

We obtain:

$$Y_t = W_{t-1}\beta + A_{m(t)}\varepsilon_t$$

$$\varepsilon_t \sim \mathcal{N}(0, I)$$

$$m(t) \in \{1, 2, \ldots, 12\}$$

Large Dimension Issues

- **More parameters than data**
  - Data for 31 years = 372 months
  - With dimension 50:
    - 372 x 50 = 1860
    - 30450 coefficients
    - 2500 x 12 =30000 variânces-covariances

→We need a sparse model → Classical variable selection is unfeasible→ LASSO

- **Highly correlated variables → $\Sigma_{m(t)}$ t=1,..,12 not full rank**
  - Inflows in the same basin
  - Wind in close to each other

→ Spectral regularization is needed

Estimation algorithm

Estimation via Maximum likelihood method:

$$\ell(\psi) = \frac{Tk}{2}\ln(2\pi) + \frac{1}{2}\sum_{t=13}^{T}\ln(|\Sigma_{m(t)}^{-1}|) - \frac{1}{2}\sum_{t=13}^{T}(Y_t - W_{t-1}\beta)^\top \Sigma_{m(t)}^{-1}(Y_t - W_{t-1}\beta)$$

$$\psi = \{\beta, \{\Sigma_m\}_{m=1}^{12}\}$$

From first order conditions we obtain the following system:

$$\begin{cases} \hat{\Sigma}_m = \displaystyle\sum_{t=13|m(t)=m}^{T} \frac{(Y_t - W_{t-1}\beta)(Y_t - W_{t-1}\beta)^\top}{n} \quad \forall m = 1, 2, \ldots, 12 \\ \hat{\beta} = \left[\displaystyle\sum_{t=13}^{T} W_{t-1}^\top \Sigma_{m(t)}^{-1} W_{t-1}\right]^{-1} \left[\displaystyle\sum_{t=13}^{T} W_{t-1}^\top \Sigma_{m(t)}^{-1} Y_t\right] \end{cases}$$

LAMPS

Estimation algorithm: Fixed-point method

Algorithm Fixed-Point Method

iter $\leftarrow 0$

$\hat{\Sigma}_m^0 \leftarrow I_{k\times k} \quad \forall m = 1, 2, ..., 12$

Obtain $\hat{\beta}^0$

while $||\hat{\beta}^{iter} - \hat{\beta}^{iter-1}||_2 > tolerance$ **do**

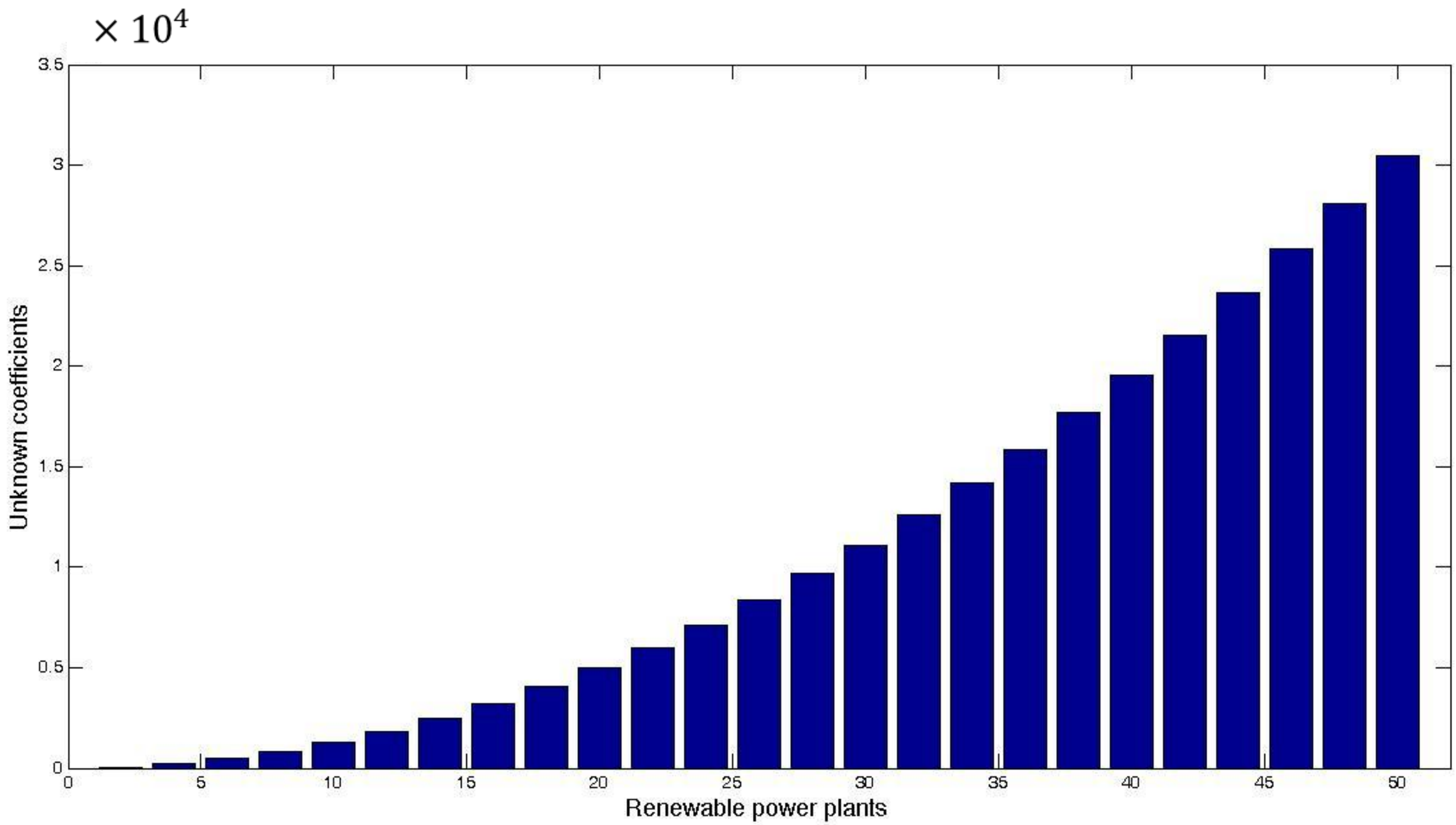
  iter $\leftarrow$ iter $+ 1$

  Obtain $\hat{\Sigma}_m^{iter}$

  Obtain $\hat{\beta}^{iter}$

end while

High-Dimensional degenerated model



For a multivariate model with only 6 plants the model is already over fitted:

β : 486 x 1

Σ: 6 x 6 (one for each month)

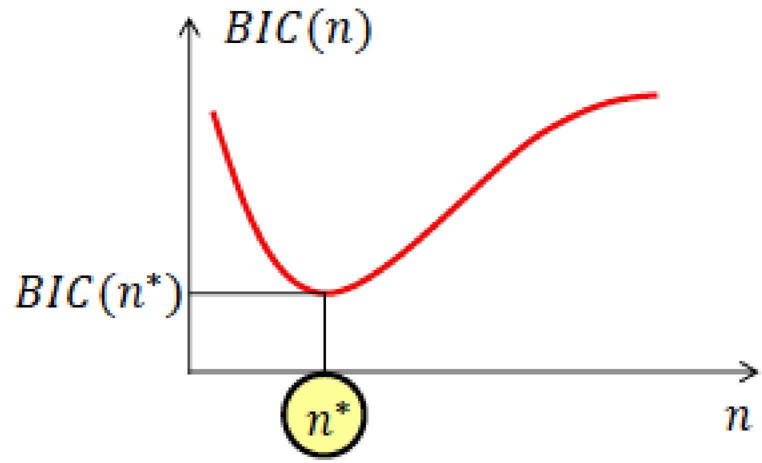
*Typical number of observations is 360 (monthly observations for 30 years)

LAMPS

Regularization

Objective: to find a sparse or regularized estimator that accounts for the tradeoff between “fitness” (likelihood) and the “degrees of freedom to adjust” (number of parameters in the model)

- Bayesian Information Criterion (BIC)
- $BIC_n = n \cdot [ln(T) - \ln(2\pi)] - 2\ell_n$

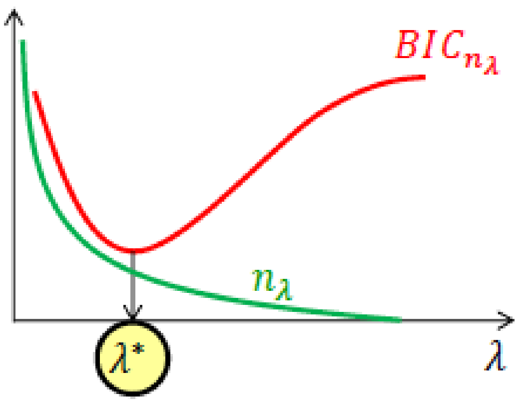


SCHWARZ, G. The annals of statistics. Estimating the dimension of a model, journal, v.6, n.2, p. 461–464, 1978.

Regularization by LASSO

Least Absolute Shrinkage and Selection Operator optimization

- $BIC_n = n \cdot [ln(T) - \ln(2\pi)] - 2\ell_n^0$
- $\ell_n^0 = \max_\psi\{\ell(\psi) | s.t.: \|\beta\|_0 \leq n\}$, where $\|\beta\|_0 = \#\{i \in N | \beta_i \neq 0\}$
- Which is a NP-Hard problem
- But we can use a norm-1 relaxation: LASSO
- $\psi_\lambda^* \in \arg\left(\max_\psi\{\ell(\psi) - \lambda \cdot \|\beta\|_1\}\right)$

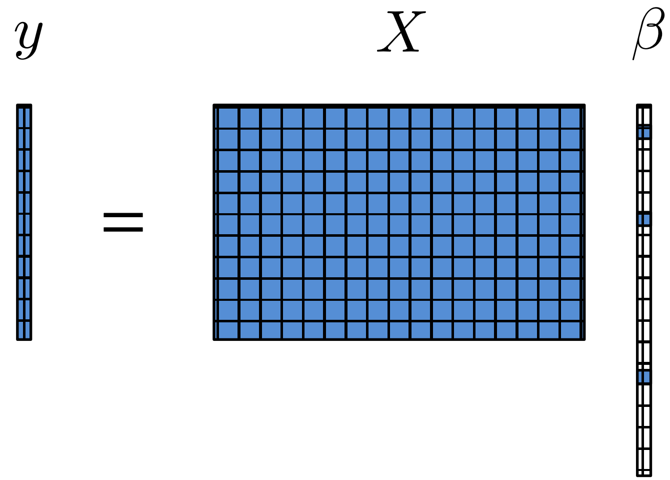


- Thus, for each λ, there is a $n_\lambda = \#\left\{i \in N \middle| (\beta_\lambda^*)_i \neq 0\right\}$ and BIC_{n_λ}

DONOHO, D. L.; ELAD, M. Proceedings of the National Academy of Sciences. Optimally sparse representation in general (nonorthogonal) dictionaries via l1 minimization, journal, v.100, n.5, p. 2197–2202, 2003.

A short digression on LASSO

Indetermined systems solution → assume the solution is sparse.



On can solve the problem

$$(P_0): \quad min \; ||\beta||_0 \quad s.a. \quad y = X\beta$$

$$||\beta||_0 = \#\{j : \beta_j \neq 0\}$$

(P_0) Is a NP-hard problem!

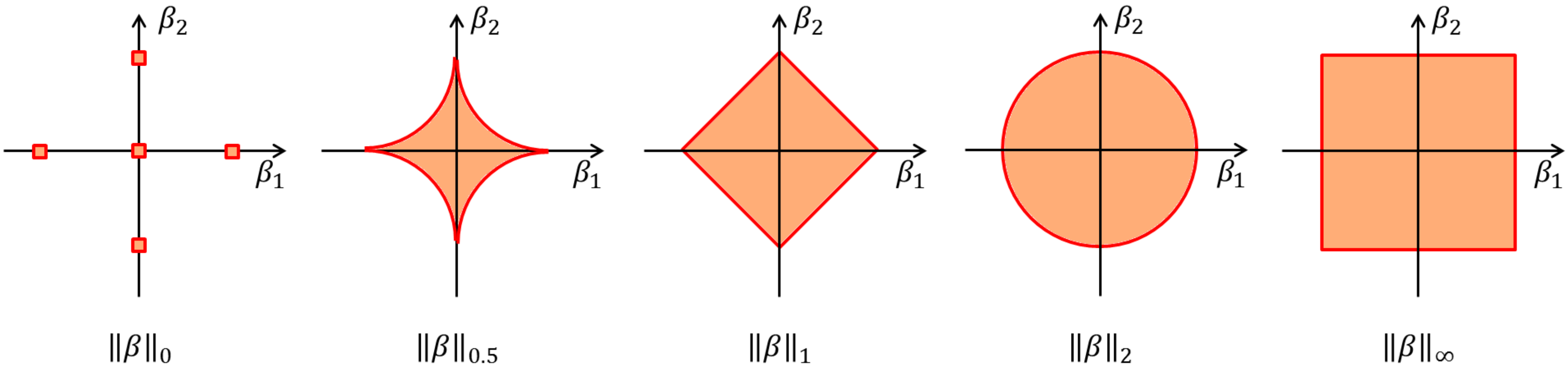
One idea is to relax the problem to the closest reasonable approximation

$$(P_0): \quad min\ ||\beta||_0 \quad s.a. \quad y = X\beta$$

$\Rightarrow$

$$(P_1): \quad min\ ||\beta||_1 \quad s.a. \quad y = X\beta$$

The norm ℓ_1 is the closest convex norm.



LAMPS

If the system has non exact solution, or some noise,

$$
\begin{aligned}
(P_{\delta,0}): \quad & min\ ||\beta||_0 \\
& s.a. \\
& ||y - X\beta||_2 \leq \delta
\end{aligned}
$$

A possible convex relaxation is

$$
\begin{aligned}
(P_{\delta,1}): \quad & min\ ||\beta||_1 \\
& s.a. \\
& ||y - X\beta||_2 \leq \delta
\end{aligned}
$$

$(P_{\delta,1})$ Is known as the *Basis Pursuit in* signal processing.

LAMPS

One can rewrite the problem under the Lagrangian form $(P_{\delta,1})$

$$LASSO: \quad \underset{(\beta)}{min}\{\ ||y - X\beta||_2^2 + \lambda||\beta||_1\}$$

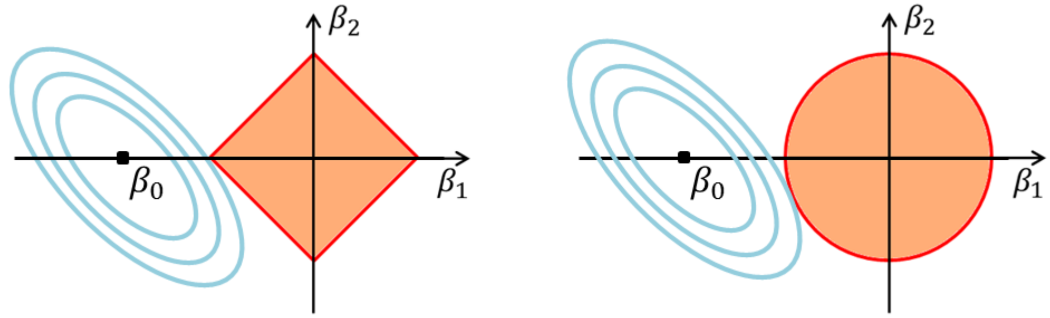
LASSO: least absolute shrinkage and selection operator

The dual form is easier to interpret

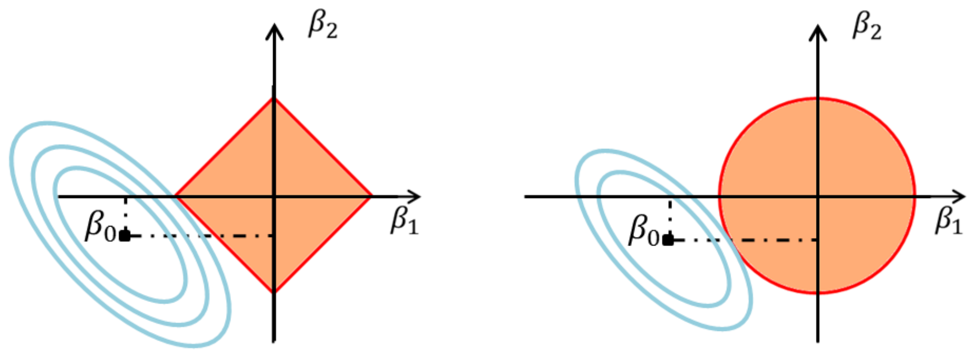
$$\hat{\beta}^{lasso} \in argmin||y - X\beta||_2^2$$
$$s.a.$$
$$\sum_{j=1}^{p} |\beta_j| \leq s$$

LAMPS

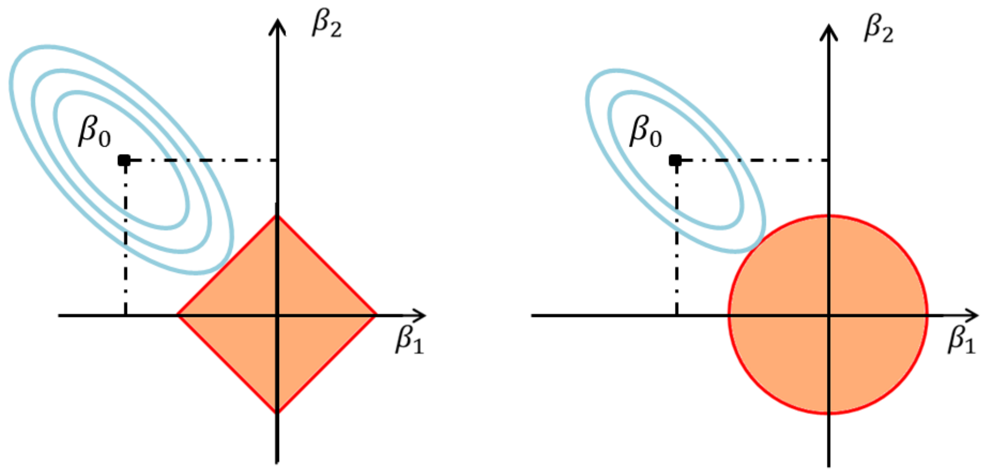
Caso 1:



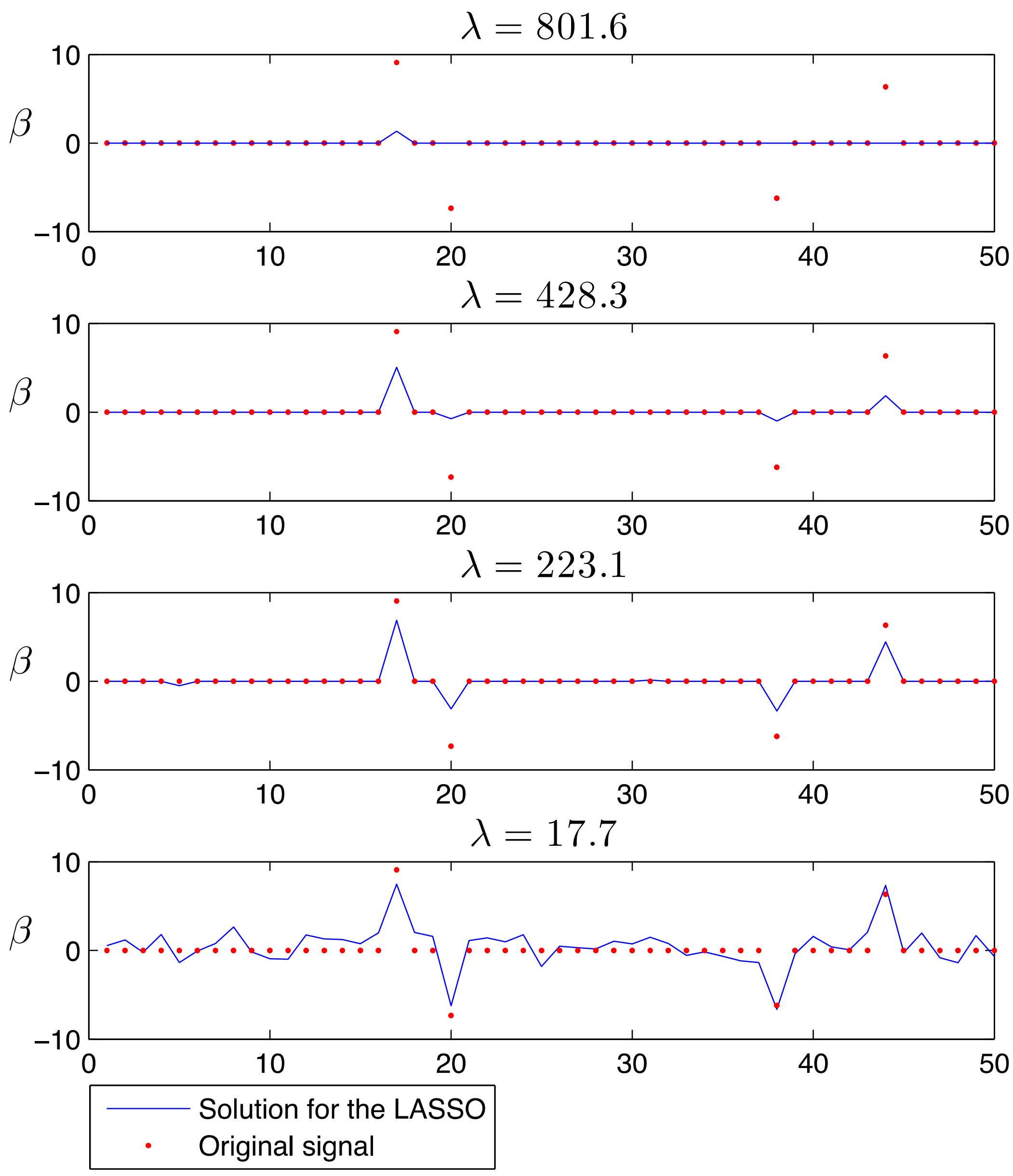
Caso 2:



Caso 3:



The parameter λ controls the sparcity of the model (50 variables).

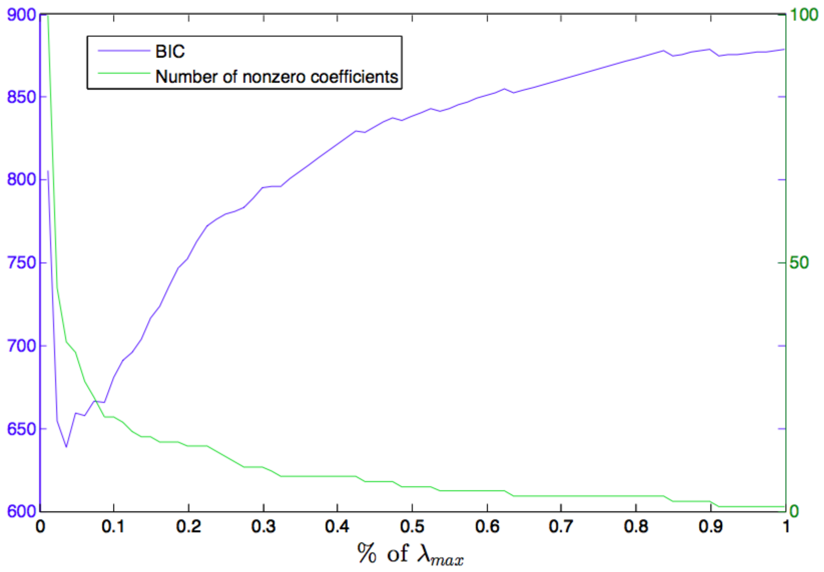


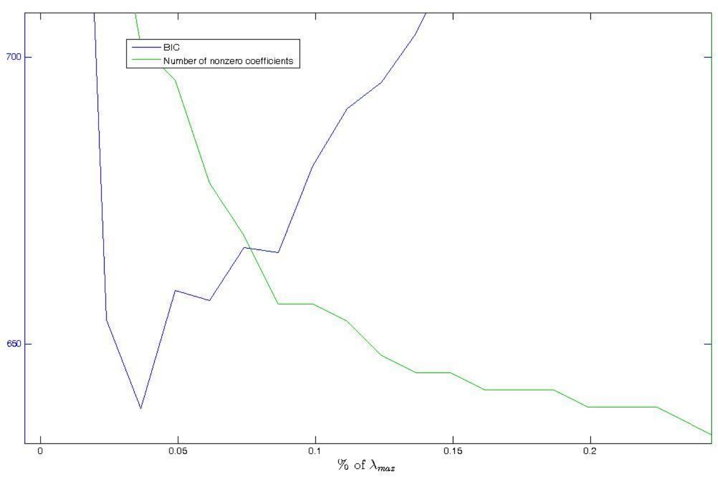
Shrinkage parameter l must be estimated by grid search between 0 and λ_{max}

$$\lambda_{max} = ||X^\top y||_\infty$$

Better value for λ depends on the goal of the model

- Forecasting: Cross-Validation
- True model: BIC (consistent asymptotically)





Optimum at 0.04*λ_{max} → 34 variables

LAMPS

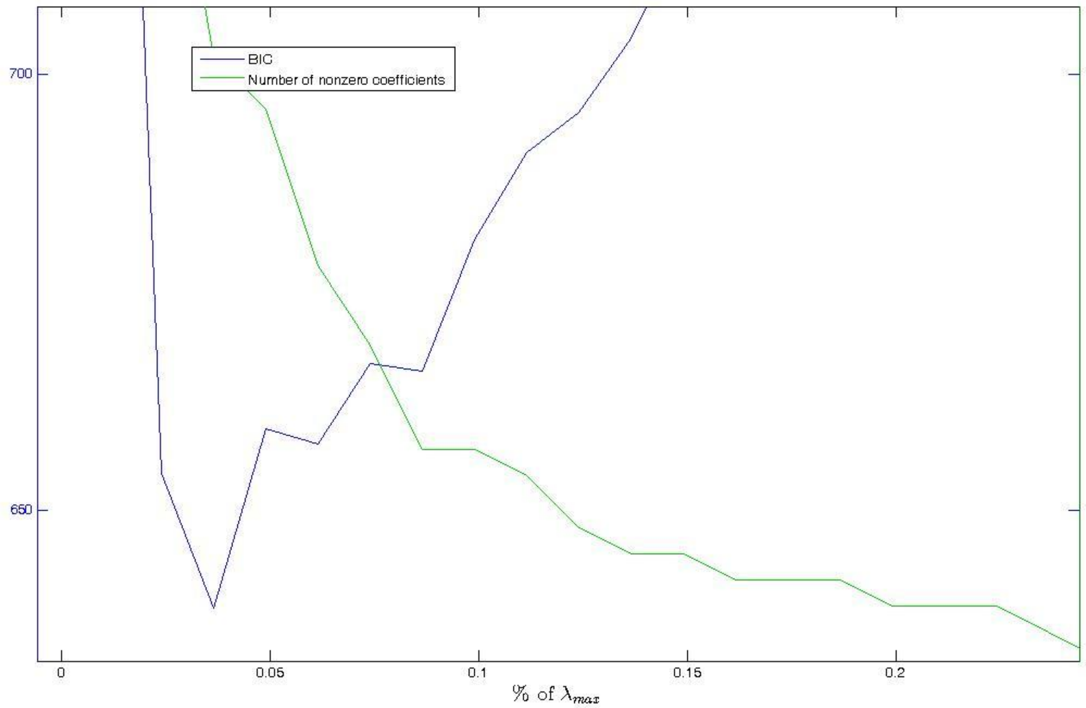
O “melhor” valor de λ não é conhecido *a priori.* Sendo assim, é necessário avaliar um grid de λ‘s no intervalo [0, λ_{max}].

$$\lambda_{max} = ||X^\top y||_\infty$$

O Critério para escolher o λ depende do obtivo do modelo:

- Previsão: Cross-Validation
- Recuperar sinal/modelo original: BIC (assintóticamente consistente)

4% do λ_{max}
Selecionando 34 variáveis



AMPS

End of digression on LASSO

Maximum likelihood estimation via LASSO

LASSO: $\max_{\{\beta,\Sigma_{m(k)},k=1,12\}}\{l^*(\beta,\Sigma_{m(k),k=1,12}) - \lambda\|\beta\|_1\}$

Tibshirani, Robert. "Regression shrinkage and selection via the lasso." *Journal of the Royal Statistical Society. Series B (Methodological)* (1996): 267-288.

Thus

$$\hat{\beta}_\lambda \in argmax\{-\frac{1}{2}\sum_{t=13}^{T}(Y_t - W_{t-1}\beta)^\top \Sigma_{m(t)}^{-1}(Y_t - W_{t-1}\beta) - \lambda||\beta||_1\}$$

$$\hat{\Sigma}_m = \sum_{t=13|m(t)=m} \frac{(Y_t - W_{t-1}\beta)(Y_t - W_{t-1}\beta)^\top}{n} \quad \forall m = 1, 2, \ldots, 12$$

Estimation algorithm via LASSO

- The covariance matrix is not in general full rank → Computing Σ_m^{-1}
  - Compute the eigenvalues of Σ_m
  - Eliminate those negatives (it happens….) and too small
  - Construct a new diagonal matrix Λ^{-1} with the inverse of theremaining eigenvalues and zeros elsewhere.
  - Compute

    $$\Sigma_{m(t)}^{-1} = P\Lambda^{-1}P^\top$$

- The maximum value for l worth looking at can be show (Souto, 2013) to be

$$\lambda^* = \left\| 2 \left| \sum_{t=13}^{T} W_{t-1}^\top \Sigma_{m(t)}^{-1} Y_t \right| \right\|_\infty$$

Estimation algorithm

Algorithm Fixed-Point & ℓ^1-regularization

$\hat{\Sigma}_m^0 \leftarrow I_{k\times k} \quad \forall m = 1, 2, ..., 12$

for $\lambda = 0$ to λ^* **do**

  iter $\leftarrow 0$

  $\hat{\beta}^0 = argmax\{\ell^*(\beta, \{\hat{\Sigma}_m^0\}_{m=1}^{12})\}$

  while $||\hat{\beta}^{iter} - \hat{\beta}^{iter-1}||_2 > tolerance$ **do**

    iter $\leftarrow$ iter $+ 1$

    Obtain $\hat{\Sigma}_m^{iter}$

    Calculate the pseudo-inverse $\Sigma_m^{-1} \quad \forall m = 1, ..., 12$

    $\hat{\beta}^{iter} = argmax\{\ell^*(\beta, \{\hat{\Sigma}_m^{iter}\}_{m=1}^{12})\}$

  end while

  Obtain $BIC(\lambda)$ using $\hat{\beta}^{iter}$ and $\{\hat{\Sigma}_m^{iter}\}_{m=1}^{12}$

end for

LAMPS

Outline of the scenario generation scheme

- Consider the output of the dispatch tool for t=1,T
  - S paths for ENA's and PLDs
  - Configuration of the system can change: $C_1, \ldots, C_K$
  - Reconstruct historical ENAs and estimate the model for each configuration $C_1, \ldots, C_{K.}$
- For each path s=1,…,S
  - Generate a path by bootstrapping monthly wise standardized residuals of the model and reinserting them in the model recursively, changing parameters when configuration changes

$$\boldsymbol{Y}_t^{(r,s)} = \hat{\boldsymbol{c}}^{(d(t))} + \sum_{i=1}^{12} \widehat{\boldsymbol{\Phi}}_i^{(d(t))} \boldsymbol{Y}_{t-1}^{(r,s)} + \sum_{j=1}^{2} \widehat{\boldsymbol{\theta}}_j^{(d(t))} \boldsymbol{x}_{t-j-1}^{s} + \widehat{\boldsymbol{A}}_{m(t)}^{(d(t))} \boldsymbol{\varepsilon}_{\boldsymbol{t}}^{(r,s)}.$$

- The scenarios for the price PLD come ready from the dispatch tool.

LAMPS

Case study: Data

Case study data:

16 Wind Farms (% of)

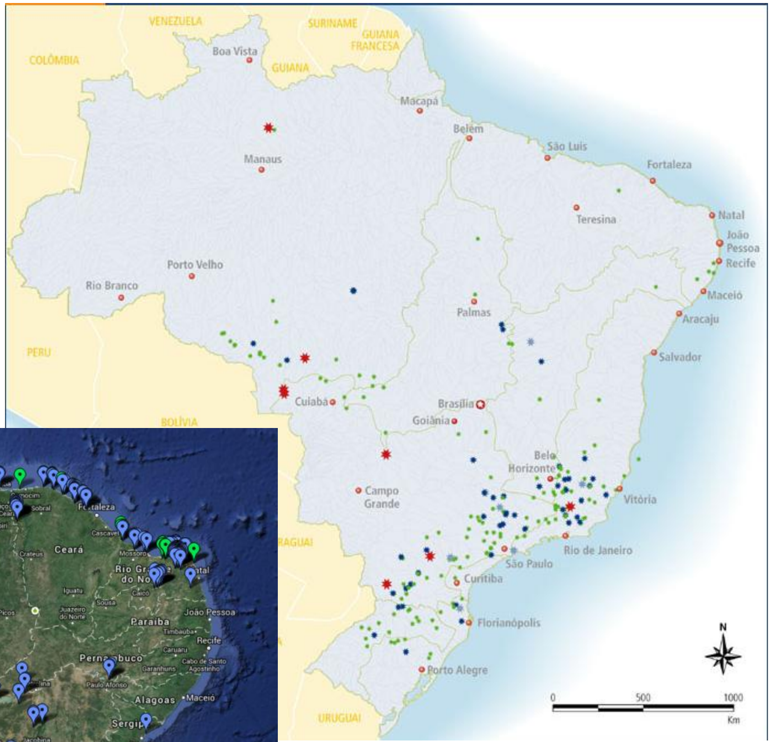
34 Hydro's (m^3/s)

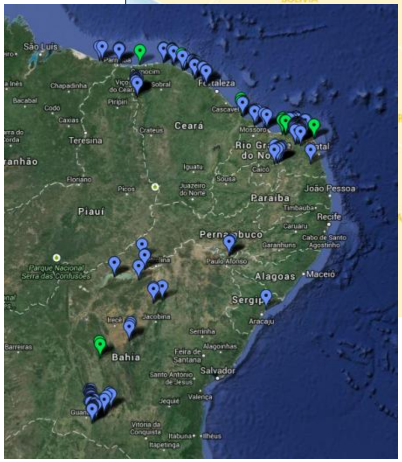
Monthly Data from January 1981 to December 2011: totalizing 372 observations.

Thus we have:

β: 30450 x 1

Σ: 50 x 50 (one for each month)





All series are log-transformed to avoid the simulation of negative results

LAMPS

Case study: Results

Result:

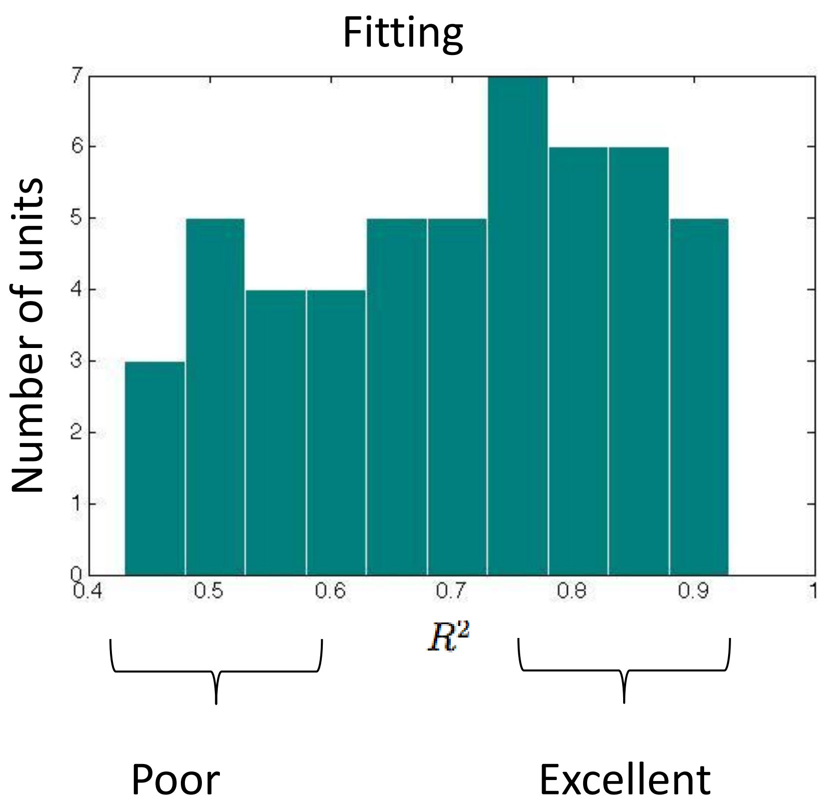
By BIC, 97,15% of β coefficients were set to zero.

Natural Water Inflow	MAPE	R^2
Salto Verdinho	14.99	0.69
Vigario	7.18	0.64
Itaguaçu	16.52	0.75
Pereira Passos	18.37	0.61
Teles Pires	9.07	0.88
Santana	10.83	0.62
Ferreira Gomes	13.18	0.85
Ilha dos Pombos	36.11	0.59
Santa Cecília	15.33	0.57
Belo Monte	14.95	0.85
Dardanelos	28.61	0.91
Salto	15.26	0.66
Santo Antonio do Jari	15.96	0.83
Tocos	50.86	0.53
Olho D'Água	16.21	0.58
Jupia	22.76	0.72
Coaracy Nunes	20.72	0.85

Natural Water Inflow	MAPE	R^2
Manso	17.59	0.72
Ponte de Pedra	9.26	0.77
Samuel	65.65	0.82
Santa Isabel	17.15	0.92
Balbina	19.5	0.68
Estreito Tocantins	21.2	0.67
Lajeado	26.2	0.48
Tucurui	12.06	0.9
Jirau	25.03	0.91
Foz do Rio Claro	18.42	0.79
Guilman Amorim	30.88	0.54
Itaipu	18.54	0.45
Itiquira	11.74	0.75
Peixe Angical	35.42	0.43
Porto Estrela	52.27	0.49
Barra dos Coqueiros	24.74	0.65
Cacu	18.58	0.79

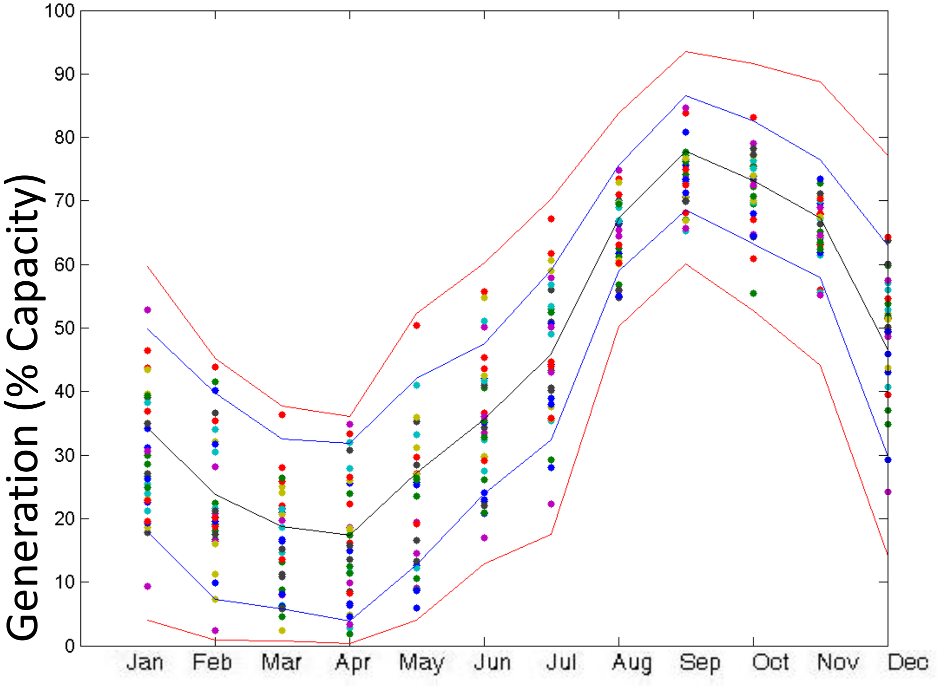
Case study: Results

Wind Farms	MAPE	R^2
Alegria 1	21.59	0.75
Alegria 2	18.33	0.63
Bons Ventos	21.49	0.76
Canoa Quebrada	40.64	0.81
Cerro Chato	11.08	0.51
Cerro Chato 2	11.16	0.51
Icaraizinho	26.67	0.88
Mangue Seco 1	14.36	0.74
Mangue Seco 2	9.01	0.51
Mangue Seco 3	8.39	0.55
Mangue Seco 4	14.28	0.79
Praia do Morgado	21.46	0.93
Praia Formosa	20.47	0.86
Rio do Fogo	18.38	0.68
Sangradouro	16.59	0.64
Volta do Rio	22.61	0.76

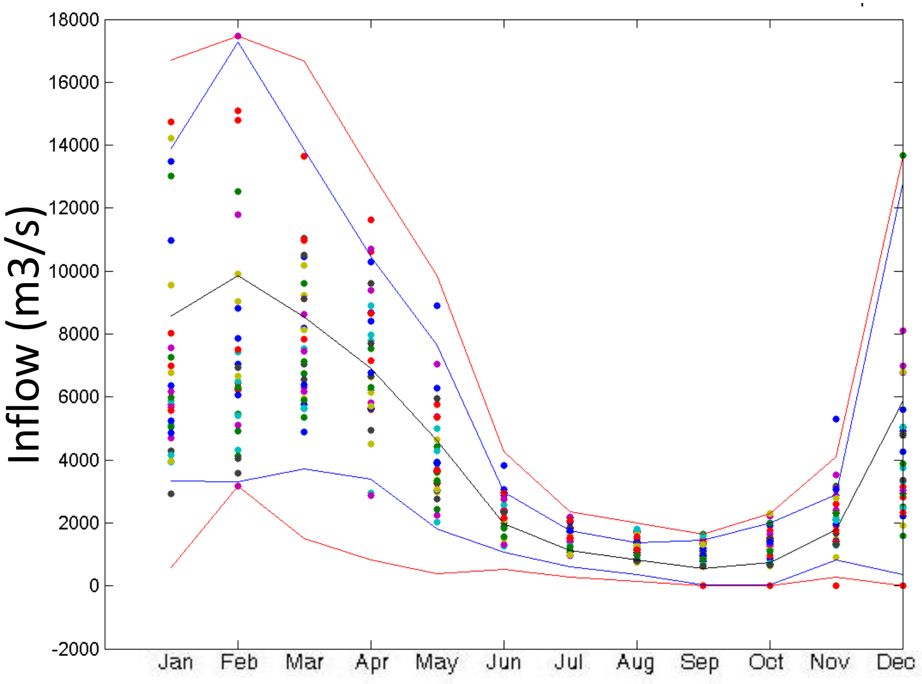


Case study: Simulation results 1

Icaraizinho



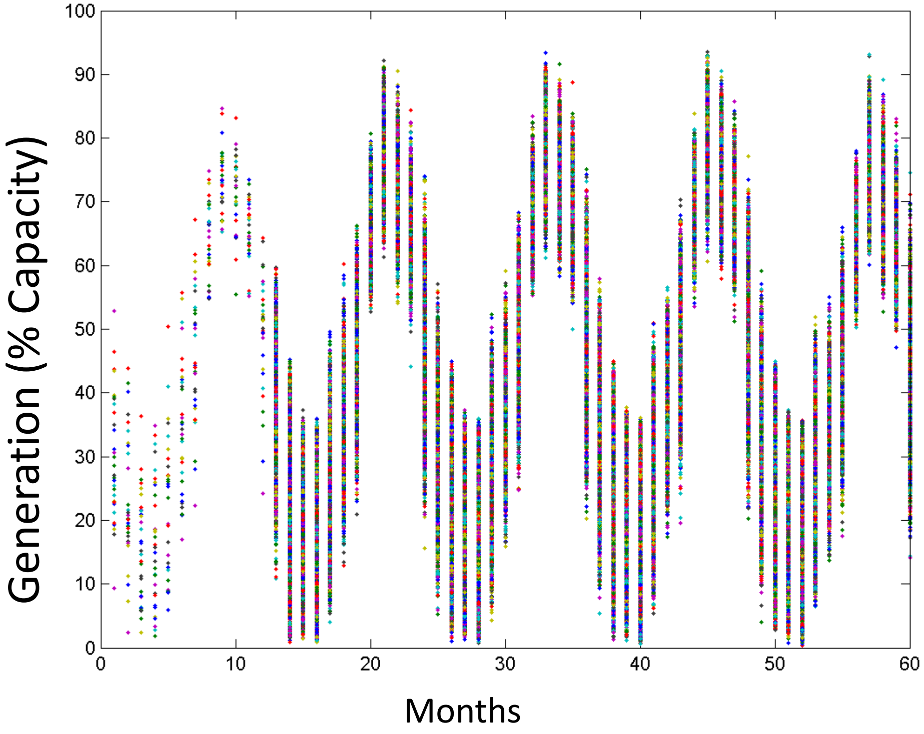
Tocantins



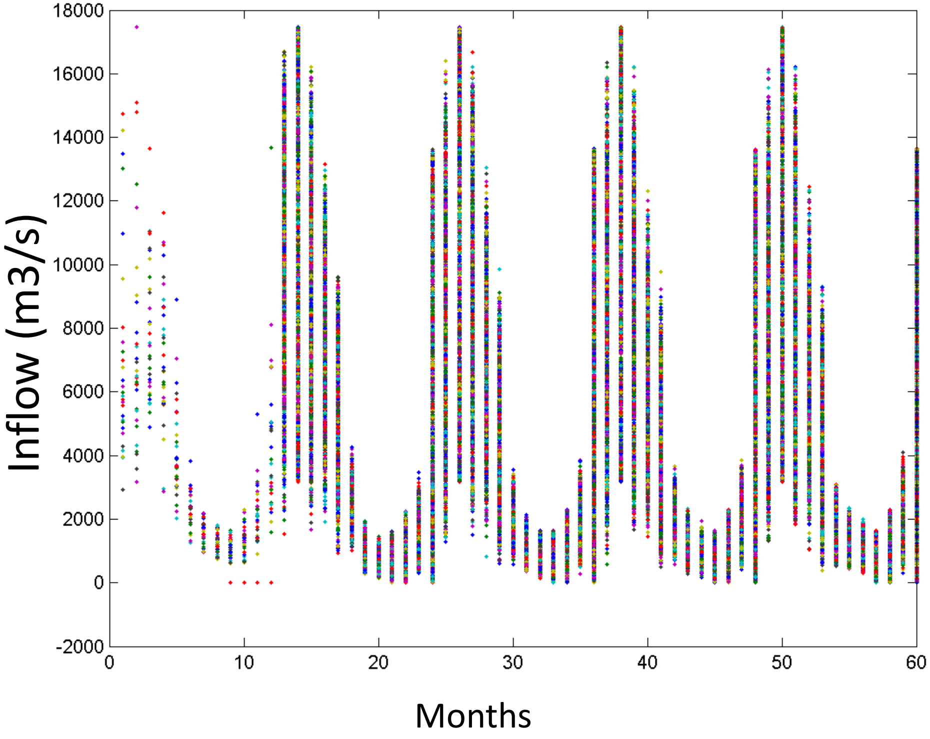
LAMPS

Case study: Simulation results 2

Icaraizinho



Tocantins



LAMPS

Generating scenarios by copulas – on going work

- The conjoint distribution is as important as the marginals for portfolio evaulation
- Multivariate copulas in large dimension is a difficult subject
- Vine pair-copulas
  - Bivariate copulas are building blocks for higher-dimensional Distributions
  - The dependency structure is determined by the bivariate copulas and a nested set of trees.
- Estimation
  - Graph theory to determine the dependency structure of the data
  - statistical inference (maximum-likelihood, Bayesian approach ...) to fit bivariate copulas
- A number of algorithms are available
  - R-Vine : Maximum Spanning Tree
  - D-Vine: "Traveling Salesman Problem"
- The R-Vine was applied to the standardized residuals
- New simulated residuasl were then re-injected in the model

Comparison Historical x Simulated

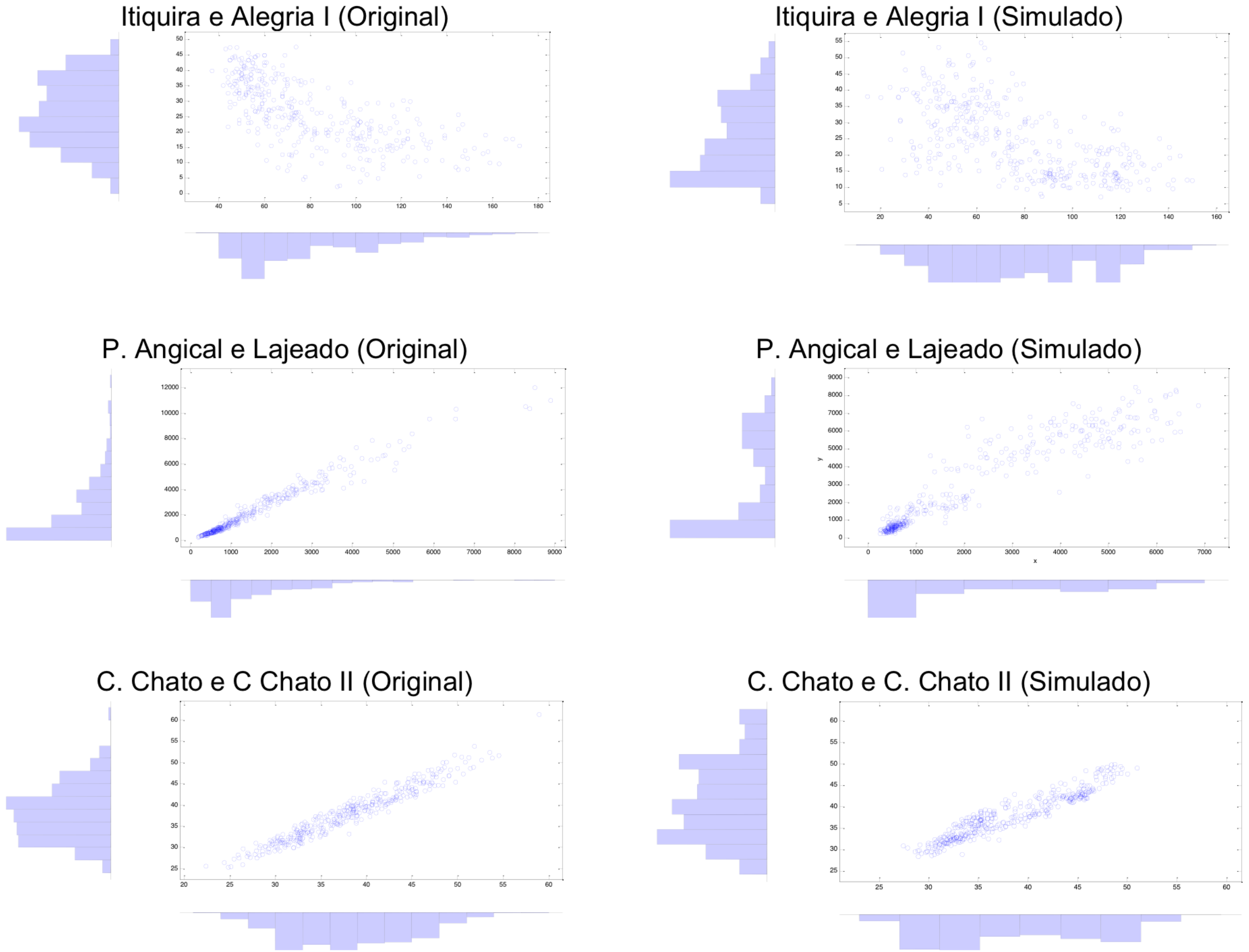
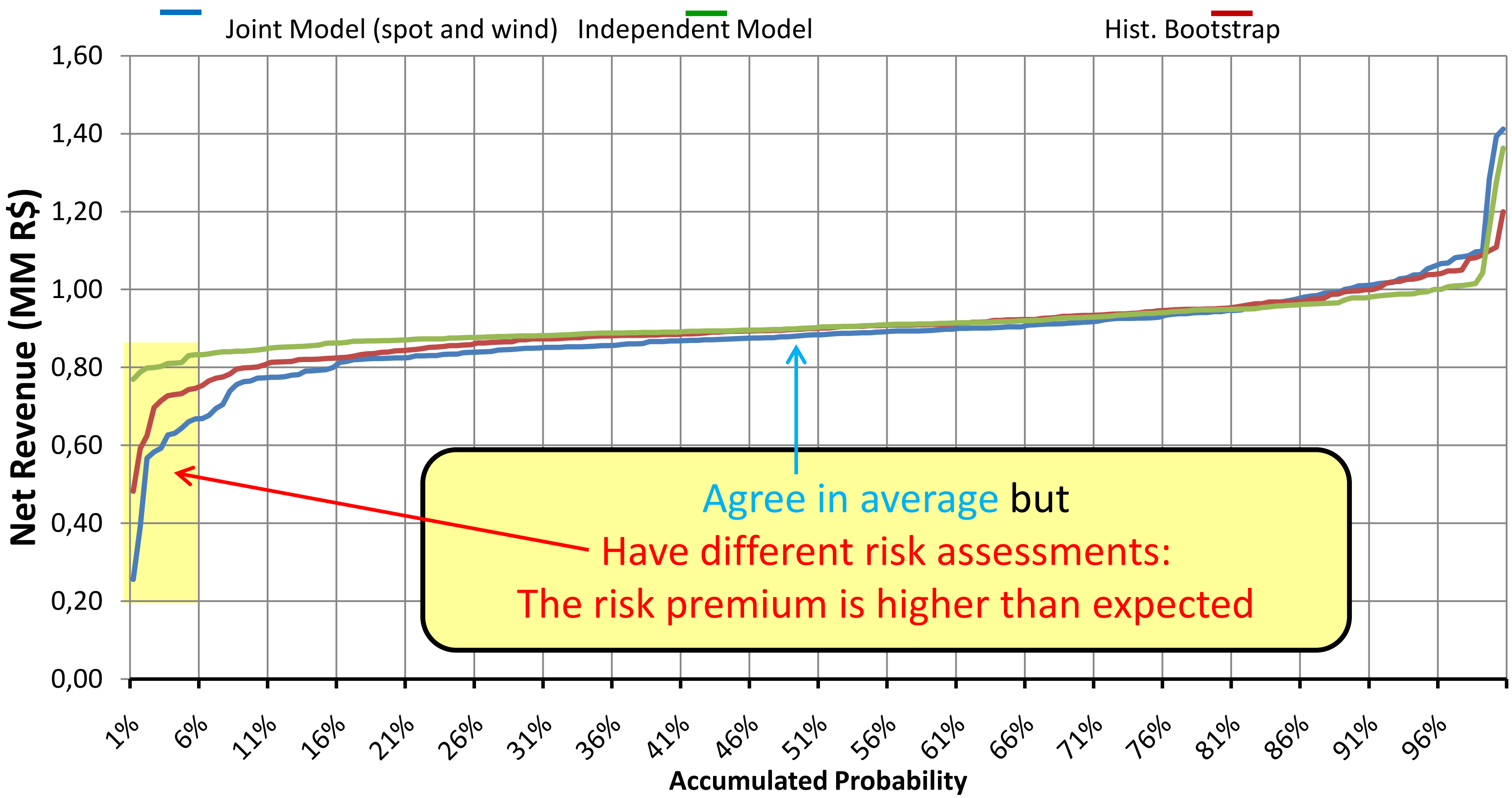


Figura 4 – Histograma e gráfico de dispersão das séries originais e das séries simuladas

LAMPS

Risk Assessment: Free Environment (ACL)

- **Annual net revenue impact:**
  - → **Wind power producer FEC = 1 avgMW**
  - → **Selling a 1-year forward contract P = 100 R$/MWh, Q = 1 avgMW.**



LAMPS

Conclusions

- We were able to make the joint estimation and simulation for 50 plants
- BIC was shown to be a good criterion for setting the regularization parameter
- Simulations of renewables were obtained and they are coherent with operation od the system and spot prices
- Taking into account the dependence between price and renewables can be seen in risk assessment
- Copulas show to be a promising way to correctly describe dependency patterns
- The method is integrated to a computational tool for optimization of electric energy contracts portfolio with risk constrains