



Nova Análise do Modelo Processo Estocástico Neural

Camilo Pinto Rodrigues

JUIZ DE FORA

JULHO, 2014

Nova Análise do Modelo Processo Estocástico Neural

CAMILO PINTO RODRIGUES

Universidade Federal de Juiz de Fora

Instituto de Ciências Exatas

Departamento de Ciência da Computação

Bacharelado em Ciência da Computação

Orientador: Luciana Conceição Dias Campos

JUIZ DE FORA

JULHO, 2014

NOVA ANÁLISE DO MODELO PROCESSO ESTOCÁSTICO NEURAL

Camilo Pinto Rodrigues

MONOGRAFIA SUBMETIDA AO CORPO DOCENTE DO INSTITUTO DE CIÊNCIAS EXATAS DA UNIVERSIDADE FEDERAL DE JUIZ DE FORA, COMO PARTE INTEGRANTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE BACHAREL EM CIÊNCIA DA COMPUTAÇÃO.

Aprovada por:

Luciana Conceição Dias Campos
D. Sc. em Engenharia Elétrica, PUC - Rio

Heder Soares Bernadino
D.Sc. em Modelagem Computacional, LNCC

Bruno Henriques Dias
D.Sc. em Engenharia Elétrica, PUC - Rio

JUIZ DE FORA
18 DE JULHO, 2014

*A Marília de Melo Costa, pelo apoio crucial
nas horas mais difíceis. Nada teria acontecido
sem você.*

Resumo

O modelo PEN (Processo Estocástico Neural) é um modelo de processo estocástico não linear que utiliza redes neurais, podendo ser usado para analisar problemas de fenômenos de comportamento periódico. Um problema tratado pelo PEN foi o de tratamento das afluências mensais, importantes para o planejamento do sistema hidrotérmico do Brasil. Fixando a quantidade de dados na autoregressão do modelo, utilizando diferentes configurações de redes neurais e com o uso de novos dados históricos sobre a série de afluências mensais, foram feitos novos testes com o PEN. Comparamos esses novos testes com o modelo PAR(p), que é o modelo usado atualmente pelo Operador Nacional do Sistema Elétrico (ONS) para gerar as séries sintéticas da série de afluências mensais, e também o comparamos com os próprios resultados mais antigos do modelo PEN.

Palavras-chave: Processos Estocásticos, Séries Temporais, Redes Neurais

Abstract

The PEN model (Neural Stochastic Process - PEN in portuguese) is a non-linear stochastic process which uses neural networks, and can be used to analyze periodical behaved phenomena. One problem treated by PEN is the treating of monthly inflows, important to the planning of Brazil's hidrothermal system. Making constant the quantity of data in the model's autoregression, using different neural networks configurations and with new historical data about the monthly inflows series, we did new tests with PEN, compared then with PAR(p), which is the model used nowadays by Nacional Operator of the Electric System (ONS in portuguese) to generate the synthetic series of the monthly inflows series, and we compared the new tests with older PEN's tests.

Keywords: Estochastic Processes, Time Series, Neural Networks.

Agradecimentos

A minha mãe, Maristela, pelo apoio incondicional.

A minha tia Ana Cláudia, por acreditar sempre!

Aos meus irmãos Raissa e Mateus, por me proporcionarem grande motivação.

A minha orientadora Luciana, por ter tido mais do que paciência comigo e com as minhas limitações.

“Todas as coisas no começo são pequenas” (Omnium Rerum Principia Parva Sunt).

Cicero, Século 1 A.C.

Sumário

Lista de Figuras	8
Lista de Tabelas	9
Lista de Abreviações	10
1 Introdução	11
2 Fundamentos Teóricos	14
2.1 Redes Neurais Artificiais Feedforward	14
2.2 Processos Estocásticos	19
2.2.1 Média	20
2.2.2 Variância	21
2.2.3 Autocovariância	21
2.2.4 Erro Médio Absoluto Percentual	21
2.2.5 Processo Estacionário	21
2.3 Séries Temporais	22
2.3.1 Série Temporal	22
2.3.2 Ruído Branco	23
2.3.3 Modelo Autoregressivo de Ordem p	23
2.4 Modelo Auto-regressivo Periódico de Ordem P - PAR(p)	24
2.5 Testes de Aderência	26
2.5.1 Teste de Kolmogorov-Smirnov	26
2.5.2 Teste T	26
2.5.3 Teste de Levene	27
3 Modelo PEN e aplicação em ENA	28
3.1 Descrição do modelo PEN	28
3.2 Aplicação do modelo PEN em ENA	30
4 Estudo de Caso	32
4.1 Considerações	32
4.2 Experimentos com o modelo PEN	33
4.2.1 Subsistema Sudeste	33
4.2.2 Subsistema Sul	41
4.2.3 Subsistema Nordeste	48
4.2.4 Subsistema Norte	56
4.2.5 Comparação com os testes de aderência do PEN original e PAR(p)	64
4.3 Gráficos PEN	67
4.3.1 Subsistema Sudeste	68
4.3.2 Subsistema Sul	70
4.3.3 Subsistema Nordeste	72
4.3.4 Subsistema Norte	74
5 Conclusão	77
5.1 Trabalhos Futuros	77

Lista de Figuras

2.1	Uma rede neural artificial. (16)	14
2.2	Um neurônio biológico. (17)	15
2.3	O neurônio artificial de McCulloch-Pitts. (7)	15
2.4	Uma rede neural artificial feedforward. (24)	18
2.5	Modelo de treinamento supervisionado. (21)	18
2.6	Modelo de treinamento não supervisionado. (21)	19
4.1	Análise dos cenários gerados no PAR(p). (6)	66
4.2	Envoltória teste, LP	68
4.3	Envoltória validação, LP	68
4.4	Envoltória teste, TP	69
4.5	Envoltória validação, TP	69
4.6	Envoltória teste, PP	69
4.7	Envoltória validação, PP	69
4.8	Envoltória teste, TL	70
4.9	Envoltória validação, TL	70
4.10	Envoltória teste, LP	70
4.11	Envoltória validação, LP	70
4.12	Envoltória teste, TP	71
4.13	Envoltória validação, TP	71
4.14	Envoltória teste, PP	71
4.15	Envoltória validação, PP	71
4.16	Envoltória teste, TL	72
4.17	Envoltória validação, TL	72
4.18	Envoltória teste, LP	72
4.19	Envoltória validação, LP	72
4.20	Envoltória teste, TP	73
4.21	Envoltória validação, TP	73
4.22	Envoltória teste, PP	73
4.23	Envoltória validação, PP	73
4.24	Envoltória teste, TL	74
4.25	Envoltória validação, TL	74
4.26	Envoltória teste, LP	74
4.27	Envoltória validação, LP	74
4.28	Envoltória teste, TP	75
4.29	Envoltória validação, TP	75
4.30	Envoltória teste, PP	75
4.31	Envoltória validação, PP	75
4.32	Envoltória teste, TL	76
4.33	Envoltória validação, TL	76

Lista de Tabelas

2.1	Tipos de processos estocásticos	20
4.1	Caso LP, Subsistema Sudeste	33
4.2	Caso TP, Subsistema Sudeste	35
4.3	Caso PP, Subsistema Sudeste	37
4.4	Caso TL, Subsistema Sudeste	39
4.5	Caso LP, Subsistema Sul	41
4.6	Caso TP, Subsistema Sul	43
4.7	Caso PP, Subsistema Sul	45
4.8	Caso TL, Subsistema Sul	47
4.9	Caso LP, Subsistema Nordeste	49
4.10	Caso TP, Subsistema Nordeste	51
4.11	Caso PP, Subsistema Nordeste	52
4.12	Caso TL, Subsistema Nordeste	54
4.13	Caso LP, Subsistema Norte	56
4.14	Caso TP, Subsistema Norte	58
4.15	Caso PP, Subsistema Norte	60
4.16	Caso TL, Subsistema Norte	62
4.17	Testes de aderência, subsistema sudeste.	64
4.18	Testes de aderência no subsistema sul	64
4.19	Testes de aderência no subsistema nordeste	65
4.20	Testes de aderência no subsistema norte	65
4.21	Testes de aderência, modelo PEN original. (6)	66

Lista de Abreviações

DCC	Departamento de Ciência da Computação
UFJF	Universidade Federal de Juiz de Fora
RNA	Rede Neural Artificial
CE	Componente Estocástica
PEN	Processo Estocástico Neural
PAR(p)	Modelo Auto-regressivo Periódico de Ordem p

1 Introdução

O planejamento energético de um país é de importância capital para garantir suprimentos adequados a todos os seus cidadãos. No caso particular do Brasil, a nossa matriz energética é em grande parte dependente de hidroelétricas e em menor parte, das termelétricas, mais caras e poluentes (22). O fato do Brasil apresentar uma dependência tão grande de hidroelétricas faz com que o país seja dependente do seu clima (isto é, das suas chuvas) para a geração de energia. A grosso modo, conseguir prever com um bom grau de precisão o quanto de chuva cada mês terá em períodos futuros é crucial. Seguindo essa previsão, o país pode saber que proporção de energia elétrica gerada por hidrelétricas usará em relação à térmica. Para se entender a importância de saber o quanto de chuva cada reservatório conterá, pode ser considerada a seguinte pergunta: o país deve usar a água nas hidrelétricas para gerar energia agora ou usar termelétricas e só depois usar o que há de água? O custo de um erro em responder essa pergunta pode ser oneroso. Caso o país use toda a água disponível agora para gerar eletricidade (pensando que no futuro haverá chuva o bastante para reabastecer) e não ocorram chuvas o suficiente no futuro, então as termelétricas deverão ser acionadas, a um custo maior e com risco de não conseguir suprir a demanda energética. Se o país usar termelétrica agora e poupar água (pensando que não vai haver chuvas o bastante) e chuvas substanciais aparecerem, as hidroelétricas serão obrigadas a abrir as comportas e liberar o excesso de água, o que significa literalmente desperdiçar energia. O Operador Nacional do Sistema Elétrico (ONS) (18) é quem precisa tomar a decisão de qual a proporção de energia hidrelétrica/termelétrica a ser usada. O ONS é o órgão responsável pela coordenação e controle da operação das instalações de geração, transmissão de energia elétrica no Sistema Interligado Nacional (SIN) e opera sob a fiscalização e regulação da Agência Nacional de Energia Elétrica (Aneel) (2).

O Sistema Interligado Nacional (SIN) é composto de empresas de 4 subsistemas equivalentes: sul, sudeste e centro-oeste, nordeste e norte (18). Para facilitar, toda a água de cada uma dessas regiões é denominada Energia Natural Afluentes (ENA), que é basicamente a energia equivalente nessas regiões, ou a estimativa de energia que pode ser

gerada. A ENA é uma das variáveis com maior incerteza para poder ser feita a previsão do preço da energia elétrica no Brasil (14) e, por isso, precisa-se saber quanto de ENA é o ideal em cada região. O planejamento do SIN é tremendamente complexo; isso motivou a adoção de simplificações como a adoção de diversos subproblemas com horizontes de tempo distintos. Nesse trabalho é utilizado o modelo do processo estocástico neural (PEN), que optou por se concentrar no período médio prazo (1-5 anos) (6). Atualmente, os cenários de ENA são gerados pelo modelo PAR(p), que são utilizados por um *software* chamado de NEWAVE, peça principal no planejamento de médio prazo do ONS, desde o seu primeiro uso nos anos 1980 (12). Esse modelo PAR(p), descrito em detalhes em Macieira (11) e comentado brevemente no nosso estudo, foi usado como fonte de comparação.

Quando o modelo PEN foi criado, a série histórica de ENA disponibilizada pelo ONS tinha um total de 75 anos (1931-2005). O PEN foi projetado para se obter as características de uma série temporal, e o faz através do uso de redes neurais artificiais. Essas redes neurais precisam ter dados suficientes para que sejam bem treinadas e tenham boa capacidade de generalização. Assim sendo, 75 anos é uma quantidade pequena para o treinamento de uma rede neural. Por isso uma das escolhas feitas ao rodar simulações utilizando o PEN foi dividir os dados da série histórica em um conjunto de treinamento e um conjunto de validação. Nesse trabalho houve acesso a mais dados em relação ao trabalho original do PEN, totalizando 80 anos (1931-2010); Utilizamos uma divisão dos dados em três partes, sendo treinamento, validação e teste (respectivamente 70 anos, 5 anos e 5 anos). Outros parâmetros do PEN foram modificados, como será explicado oportunamente. Com isso dito, é feita a tentativa de se responder uma pergunta simples: como os novos resultados gerados são comparados com os resultados originais do PEN e com os resultados do PAR(p)? Vale lembrar que para o modelo PAR(p), foi usada a implementação baseada em Macieira (11) para gerar esses resultados, do mesmo modo que foi feito por Campos (6).

A proposta desse estudo é continuar a exploração do PEN. Através do uso de novos dados e de uma nova análise, o modelo PEN é mais explorado, mostrando a influência de diversos parâmetros na geração das séries sintéticas, utilizando de modificações em alguns desses parâmetros e de restrições em outros. Foi feita uma comparação de como

o modelo se comporta em relação ao modelo PAR(p) e mostramos como ele evoluiu em relação ao modelo original.

Uma questão que pode surgir é por que explorar mais esse modelo PEN, com novos dados e parâmetros? A resposta passa pela motivação de buscar modelos alternativos para o setor elétrico. De acordo com Campos (6), desde a época da crise energética no Brasil em 2001-2002 (o “apagão”), em que o Governo brasileiro foi obrigado a tomar medidas de racionamento de energia, ficou muito claro a necessidade de explorar novos modelos para o setor elétrico.

A estrutura do trabalho é a que segue. No capítulo 2 alguns conceitos teóricos importantes para o entendimento do PEN são explicados, tais como redes neurais, que são usadas para compor o PEN, seguidos por processos estocásticos e séries temporais. É citado brevemente um modelo precursor do PAR(p) que é o AR(p) (modelo autorregressivo de ordem p) e finalizando com uma explicação sucinta a respeito do modelo PAR(p). No capítulo 3 há a apresentação do modelo PEN, primeiro descrevendo-o e em seguida mostrando como ele é aplicado na geração de séries sintéticas de energia natural afluyente (ENA). No capítulo 4 os resultados das simulações são mostrados, mostrando o efeito da variação de alguns parâmetros do PEN nos cenários gerados por ele, comparando-os com os resultados do PAR(p). Finalmente, no capítulo 5 foram feitas as conclusões e algumas considerações sobre trabalhos futuros.

2 Fundamentos Teóricos

2.1 Redes Neurais Artificiais Feedforward

Segundo Princípio (20) , redes neurais artificiais (RNAs) são máquinas de aprendizado distribuídas e adaptativas, compostas de unidades processadoras (UP). Cada UP recebe conexões de outras UPs (podendo formar várias camadas de UPs) ou de si própria. Essas interconexões definem a topologia de uma RNA. Os sinais que fluem dessas conexões são dimensionados por parâmetros denominados pesos sinápticos, w_1, \dots, w_n . As UPs somam todas essas contribuições para produzir uma saída. As saídas de uma UP ou são as saídas do sistema ou são enviadas a outra UP. Como mostra a figura 2.1, nós temos diversas entradas x_1, \dots, x_n , que são somadas, alteradas com um *bias* b e uma não linearidade $\varphi(\cdot)$ e produzem uma saída y .

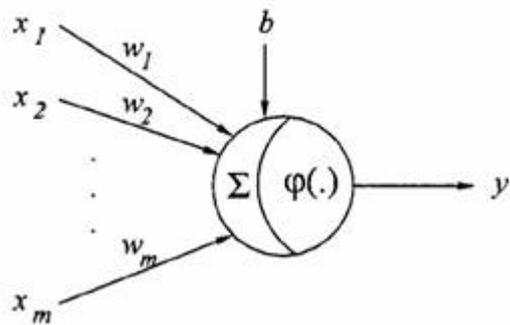


Figura 2.1: Uma rede neural artificial. (16)

As UPs são também chamadas de neurônios artificiais. A motivação para os neurônios artificiais são os neurônios biológicos - os principais componentes do sistema nervoso. São células excitáveis eletricamente, que processam e transmitem informações através de sinais químicos e elétricos. Como modelo simplificado, um neurônio típico possui um corpo celular, dendritos e um axônio, ilustrado na Figura 2.2.

As UPs podem se conectar umas às outras, formando uma rede neural. A estrutura que permite essas conexões é denominada sinapse, que permite que um neurônio

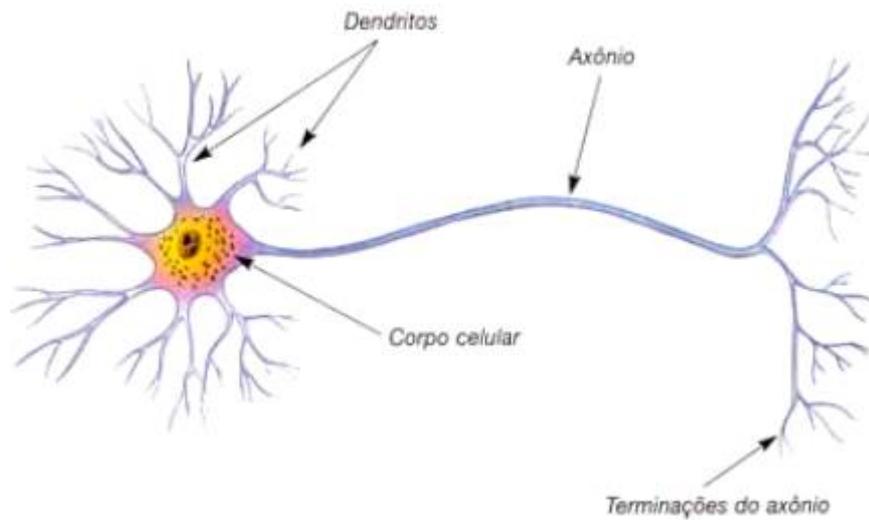


Figura 2.2: Um neurônio biológico. (17)

passa a outro um sinal (elétrico ou químico, como dissemos anteriormente) (23). As conexões formadas podem ser neurônio-neurônio, ou neurônio com outros tipos de células (por exemplo, musculares). O grande número de neurônios no cérebro humano (na ordem dos 100 bilhões) e a capacidade de conexão que a sinapse fornece geram estruturas muito complexas e grande processamento paralelo. O neurônio artificial tenta ao máximo emular o neurônio biológico. No nosso estudo, o modelo de neurônio artificial é modelo clássico de McCulloch-Pitts (13), como podemos ver na Figura 2.3 (7).

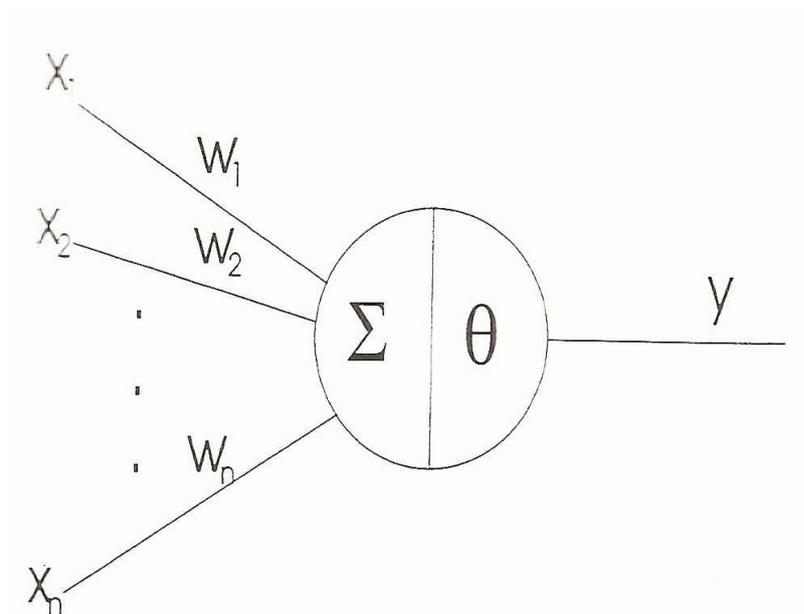


Figura 2.3: O neurônio artificial de McCulloch-Pitts. (7)

Usando da analogia com o neurônio biológico, o neurônio artificial tem diversas entradas, nomeadas X_1, \dots, X_n , que correspondem aos dendritos. Para essas entradas, uma única saída é produzida, nomeada Y , que corresponde ao axônio. Para cada entrada X_i , há um peso sináptico W_i correspondente. Esses pesos simulam as sinapses, e eles podem ter valores positivos ou negativos, dependendo do tipo de estímulo. Matematicamente, o efeito de um estímulo de uma sinapse i é dado pelo produto $X_i \cdot W_i$. O corpo celular é simulado pelo somatório 2.1.

$$\sum_{i=1}^n X_i \cdot W_i \quad (2.1)$$

No neurônio biológico, dizemos que o neurônio disparou se a soma dos estímulos que ele recebeu ultrapassou o seu limiar de excitação. O somatório que representa o corpo celular determina se o neurônio disparou ou não, dependendo se a soma ultrapassou ou não o limiar θ . No modelo de McCulloch-Pitts, a saída do neurônio é binária, e depende apenas da equação 2.2.

$$\sum_{i=1}^n X_i W_i > \theta \quad (2.2)$$

Sendo a saída 1 quando a equação for verdadeira e 0 quando falsa. De acordo com Carvalho (7), diversos modelos foram propostos a partir do modelo original de McCulloch-Pitts e diversas funções de ativação foram criadas, ampliando o conjunto de valores produzidos na saída. Quatro funções de ativação de interesse são as funções linear, rampa, degrau e sigmóide.

1. Função linear

$$Y = \alpha X \quad (2.3)$$

em que Y é a saída, X a entrada e α um número real aleatório.

2. Função Rampa

$$Y = \begin{cases} \gamma, & \text{se } x \geq \gamma \\ x, & \text{se } x \leq |\gamma| \\ -\gamma, & \text{se } x \leq -\gamma \end{cases} \quad (2.4)$$

A função rampa é utilizada quando queremos limitar a saída da função linear em um intervalo $[-\gamma, \gamma]$, fazendo com que a saída seja constante em determinados intervalos. Essa função é geralmente usada para simplificar uma função não linear.

3. Função Degrau

$$Y = \begin{cases} \gamma, & \text{se } x \geq 0 \\ -\gamma, & \text{se } x < 0 \end{cases} \quad (2.5)$$

Similar a função rampa, com a diferença que a saída são dois valores rígidos. Podemos adaptar essa função caso seja desejável que a saída seja 0 ou 1.

4. Função Sigmóide

$$Y = \frac{1}{1 + e^{-\frac{x}{T}}} \quad (2.6)$$

onde ϵ é um valor real e T define a suavidade da curva. Na verdade, essa equação particular é apenas um tipo de sigmóide, denominada função logística. As características comuns de sigmóides são semilinearidade, são limitadas e monotônicas. É uma das funções mais utilizadas na construção de redes neurais artificiais.

Diversos tipos de RNAs podem ser construídos para a resolução de problemas, e a forma como são construídas limita o tipo de problemas que podem ser resolvidos. Dentre as características que controlamos em um projeto de RNA, destacamos a escolha da função de ativação, o número de camadas de neurônios, como os neurônios se conectam uns aos outros (topologia) e como se dá o aprendizado da RNA. No nosso caso específico, estamos interessados em redes neurais não recorrentes, também chamadas redes neurais feedforward. Redes neurais artificiais feedforward (RNAFF), também chamadas de sem memória, são aquelas em que, dadas as camadas $1, \dots, i, \dots, n$ da RNA, não há conexões entre os neurônios de uma camada com neurônios de camadas anteriores ou com outra camada que não seja a sucessiva, isto é: para os neurônios da camada i , existe apenas conexão com os neurônios da camada $i+1$ (8). A Figura 2.4 mostra uma RNAFF. Essa rede possui, na sua primeira camada, a entrada dos dados; uma ou mais camadas que ficam entre as camadas de entrada e saída, chamadas de camadas escondidas. Essas camadas transmitem suas saídas para as camadas seguintes, até chegar à saída.

Para que uma rede neural produza uma saída, o conjunto de dados de entrada deve passar por um processo de treinamento. Esse treinamento nada mais é que um ajuste dos pesos p_1, \dots, p_n da Figura 2.1. O ajuste é feito de modo que esses pesos corrigidos possam fazer os valores de entrada convergirem para determinados valores, que representam o conhecimento adquirido pela RNA. Como a RNA aprende a partir de

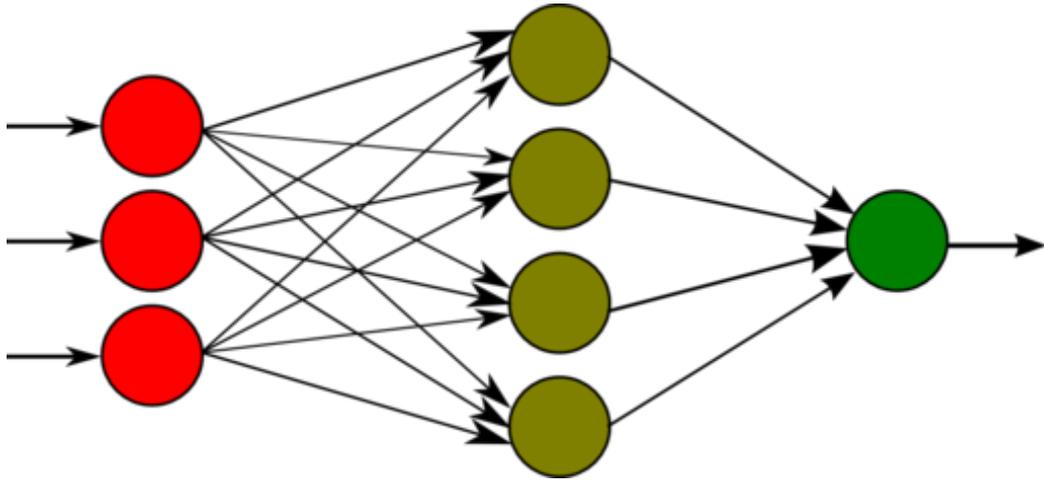


Figura 2.4: Uma rede neural artificial feedforward. (24)

padrões de dados, vale ressaltar que a qualidade dos dados de entrada é muito importante para o bom aprendizado da RNA, ou seja, a escolha de boas variáveis de entrada X_1, \dots, X_n , é fundamental.

Basicamente, há dois processos de treinamento, o supervisionado e o não supervisionado.

1. Treinamento supervisionado Um modelo simples de aprendizado supervisionado é dado na Figura 2.5 . O conjunto de dados que representa a entrada, x , é dado a

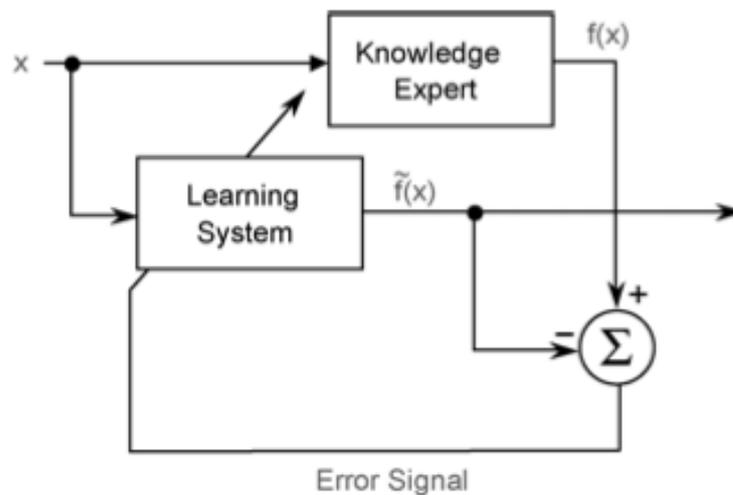


Figura 2.5: Modelo de treinamento supervisionado. (21)

um expert (*Knowledge Expert*, na figura 2.5). Esse expert produz a saída desejada, que chamamos $f(x)$. A RNA, que representa o sistema de aprendizado, recebe o

mesmo conjunto x de entrada, e produz uma outra saída, chamada $\tilde{f}(x)$. O erro entre a saída produzida pela RNA e pelo expert é então calculado, e utilizado para modificar a saída produzida pela RNA, adaptando os pesos desta, para tornar a saída mais próxima da saída produzida pelo expert. Esse processo é repetido até que a saída seja considerada próxima o suficiente da saída do expert. Esse expert pode ser uma função, um conjunto de regras, um conjunto de saídas medidas, um ser humano e assim por diante.

2. Treinamento não supervisionado O treinamento não supervisionado (Figura 2.6) é parecido com o supervisionado, mas não há nenhum expert que diga qual a saída desejada (21).

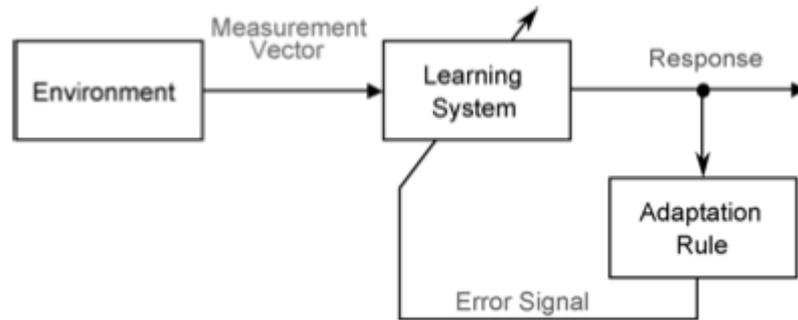


Figura 2.6: Modelo de treinamento não supervisionado. (21)

Esse modelo consiste de um ambiente (*environment*), representado por um vetor (*measurement vector*), que a RNA (*learning system*) recebe como entrada. A RNA produz então uma saída. Baseada nessa saída e na regra de adaptação escolhida, os pesos da RNA são ajustados até se obter o desempenho desejado.

2.2 Processos Estocásticos

Um processo estocástico é um conjunto de variáveis aleatórias definidas em um espaço de probabilidades comum (3). Podemos representar um processo estocástico como

$$\{X(t) : t \in T\} \quad (2.7)$$

onde $X(t)$ é uma variável aleatória denominada *estado do processo* no instante t e T é o *espaço paramétrico* do processo estocástico. Observemos também que a natureza de T pode ser contínua ou discreta, definindo também a classificação do processo estocástico. Se T for um intervalo de valores reais, o processo estocástico é denominado *processo contínuo*; se T for um conjunto de elementos contável, o processo estocástico é denominado *processo discreto*. Vamos utilizar a nomenclatura sugerida por Barros (3), que significa que para processos discretos usaremos a notação X_t ao invés de $X(t)$. Notando que o estado do processo $X(t)$ também pode ser discreto ou contínuo, e combinando esse conhecimento com a natureza do espaço paramétrico, temos 4 tipos de processos estocásticos possíveis, de acordo com as naturezas de $X(t)$ e T , que podem ser, de acordo com a tabela 2.1:

Tabela 2.1: Tipos de processos estocásticos

$X(t)$	T
Discreto	Discreto
Discreto	Contínuo
Contínuo	Discreto
Contínuo	Contínuo

Para descrevermos um processo estocástico de forma simples nós utilizamos os *momentos* das variáveis aleatórias, em especial os três primeiros, que são a média, variância e autocovariância do processo. Considerando um processo estocástico discreto, temos as seguintes definições, de acordo com (3):

2.2.1 Média

Média de um processo estocástico $X(t)$ é dada por

$$\mu = E(X_t) \tag{2.8}$$

Que não necessariamente é uma constante, e sim uma função do instante de tempo.

2.2.2 Variância

A variância de um processo estocástico $X(t)$ é dada por

$$\sigma^2(t) = Var(X_t) = E\{(X_t - \mu(t))^2\} \quad (2.9)$$

2.2.3 Autocovariância

A autocovariância de um processo estocástico $X(t)$ entre os instantes de tempo t_1 e t_2 é dada por

$$\gamma(t_1, t_2) = Cov\{X_{t_1}, X_{t_2}\} = E\{(X_{t_1} - \mu(t_1))(X_{t_2} - \mu(t_2))\} \quad (2.10)$$

O que significa que a autocovariância de um processo entre os instantes de tempo t_1 e t_2 é apenas a covariância para instantes de tempos diferentes. Se $t_1 = t_2$, temos a definição de variância. Essas características de processos estocásticos estão intimamente ligadas a noção de processo estacionário (3).

2.2.4 Erro Médio Absoluto Percentual

Consideremos Y_t uma observação no tempo t e F_t como sendo a previsão de Y_t . Definindo erro de precisão e_t como sendo $Y_t - F_t$. Temos que o Erro Médio Absoluto Percentual (MAPE) seria simplesmente a média percentual de e_t , dada por (9),

$$MAPE = \frac{1}{t} \sum_{i=1}^t \left| \frac{Y_i - F_i}{F_i} \right| \quad (2.11)$$

onde t seria o número de observações .

2.2.5 Processo Estacionário

Para a definição de processo estacionário vamos utilizar a definição de Papoulis (19). Um processo estocástico $X(t)$ é denominado *estritamente estacionário* se suas propriedades estatísticas são invariantes em relação a um deslocamento da origem dos t . Isso significa que $X(t)$ e $X(t+c)$ tem as mesmas características para qualquer c . Dois processos $X(t)$ e $Y(t)$ são chamados *conjuntamente estacionários* se as propriedades estatísticas conjuntas

de $X(t)$ e $Y(t)$ são as mesmas que propriedades estatísticas conjuntas de $X(t+c)$ e $Y(t+c)$. Um processo complexo $Z(t) = X(t) + jY(t)$ é estacionário se os processos $X(t)$ e $Y(t)$ forem conjuntamente estacionários. Dessas definições podemos concluir que a densidade de n -ésima ordem de um processo estritamente estacionário deve ser tal que $f(x_1, \dots, x_n; t_1, \dots, t_n) = f(x_1, \dots, x_n; t_1+c, \dots, t_n+c)$ para todo c . Podemos também chegar a conclusões particulares:

- Para $n=1$, a média e a variância são constantes para todo t ;
- Para $n=2$, a distribuição conjunta de $X(t_1)$ e $X(t_2)$ só depende do *lag* t_2-t_1 , assim como a autocovariância.

A autocovariância pode ser escrita como:

$$\gamma(k) = E\{(X(t) - \mu)(X(t+k) - \mu)\} \quad (2.12)$$

A função $\gamma(k)$ é chamada autocovariância de lag k .

Um processo estocástico $X(t)$ é denominado *fracamente estacionário* se sua média é constante e sua autocovariância depende apenas do lag. Um processo fracamente estacionário tem condições bem mais fracas que o processo estritamente estacionário, pois as condições dependem apenas dos dois primeiros momentos, em contrapartida a todos os momentos. O nosso interesse em um processo estocástico deriva da possibilidade de utilizá-lo para analisar séries temporais de forma a capturar suas características a fim de reproduzir séries sintéticas igualmente prováveis.

2.3 Séries Temporais

2.3.1 Série Temporal

Segundo Box (5), Uma série temporal é um conjunto de observações geradas sequencialmente no tempo. Se o conjunto é contínuo, a série temporal é dita contínua. Se o conjunto é discreto, a série temporal é dita discreta. Assim, observações de uma série temporal discreta feitas nos tempos τ_1, \dots, τ_n podem ser denotadas por $z(\tau_1), \dots, z(\tau_n)$. Consideraremos aqui apenas séries temporais discretas onde as observações são feitas em

um intervalo fixo h , isto é, se o ponto de tempo inicial for τ_0 e tivermos N observações, a série temporal será $\tau_0, \tau_0 + h, \tau_0 + 2h, \dots, \tau_0 + Nh$. Para simplicidade, denotaremos essa série como $z_1, z_2, \dots, z_t, \dots, z_n$, onde z_t pode ser considerado como a observação no instante t . Também podemos entender uma série temporal como sendo a realização de um processo estocástico. Séries temporais discretas podem aparecer de duas maneiras:

1. Tirando amostras de uma série temporal contínua. Imagine que há um forno industrial em que a entrada de ar é contínua e a saída de CO_2 também. Se definirmos intervalos discretos, por exemplo a cada 10 segundos, e fizermos a medição dessas entradas e saídas nesses intervalos, teremos uma série temporal discreta.
2. Através de acumulação de uma variável em um período de tempo. Exemplo comum seria a chuva que normalmente é acumulada em períodos como um dia ou mês (como em “no mês de março a quantidade de chuva costuma ser superior a 100mm”)

2.3.2 Ruído Branco

Um processo estocástico discreto é chamado de *ruído branco* se ele é uma sequência de variáveis aleatórias independentes e igualmente distribuídas, escritas como a_1, \dots, a_t, \dots , que nós assumimos como tendo média zero e variância σ_a^2 (5). Independência significa que os a_t são não correlacionados, a função de autocovariância é definida apenas por

$$\gamma_k = \begin{cases} \sigma_a^2, & \text{se } k = 0 \\ 0, & \text{se } k \neq 0 \end{cases} \quad (2.13)$$

2.3.3 Modelo Autoregressivo de Ordem p

AR(p) é definido como

$$Z_t = c + \sum_{i=1}^p \varphi_i Z_{t-i} + \varepsilon_t \quad (2.14)$$

onde c é uma constante, $\varphi_1, \dots, \varphi_p$ são os parâmetros de peso do modelo e ε_t é ruído branco. O processo definido por 2.14 é chamado de processo *autoregressivo de ordem p*, ou simplesmente AR(p). Se definirmos um operador de *backshift* B como sendo $BZ_t = Z_{t-1}$,

podemos reescrever a equação 2.14 como sendo

$$Z_t = c + \sum_{i=1}^p \varphi_i B^i Z_t + \varepsilon_t \quad (2.15)$$

Trazendo o somatório da equação 2.15 para o lado esquerdo e usando notação polinomial, temos

$$\phi(B)X_t = c + \varepsilon_t \quad (2.16)$$

2.4 Modelo Auto-regressivo Periódico de Ordem P - PAR(p)

Existem algumas séries temporais com intervalos menores que um ano, como por exemplo algumas séries hidrológicas mensais. Esse tipo de série apresenta comportamento periódico de algumas de suas propriedades, como por exemplo a média, a variância e a estrutura de auto-correlação (11). Para estudarmos esse tipo de séries, podemos fazer uso de modelos auto-regressivos, mas modificados de modo que seus parâmetros apresentem comportamento periódico. Chamaremos esse novo modelo modificado de modelo auto-regressivo periódico, PAR(p), onde p é o número de termos auto-regressivos do termo, também chamado de ordem do modelo. Normalmente p é um vetor $p = (p_1, p_2, \dots, p_{12})$ (até p_{12} para períodos mensais) e de acordo com Macieira (11), podemos descrever o modelo PAR(p) com a equação:

$$\left(\frac{Z_t - \mu_m}{\sigma_m} \right) = \phi_1^m \left(\frac{Z_{t-1} - \mu_{m-1}}{\sigma_{m-1}} \right) + \dots + \phi_{p_m}^m \left(\frac{Z_{t-p_m} - \mu_{m-p_m}}{\sigma_{m-p_m}} \right) \quad (2.17)$$

onde temos

Z_t é uma série sazonal de período s

s é o número de períodos (s = 12 para séries mensais)

t é o índice do tempo, $t = 1, 2, \dots, s$, função do ano T ($T = 1, 2, \dots, N$) e do período m

($m = 1, 2, \dots, s$)

N é o número de anos

μ_m é a média sazonal de período s

σ_m é desvio-padrão sazonal de período s

ϕ_i^m é o i -ésimo coeficiente auto-regressivo do período m

p_m é a ordem do operador auto-regressivo do período m

ε_t série de ruídos independentes (ruído branco) com média zero e variância $\sigma_a^{2(m)}$

Considerando $\rho^m(k)$ a correlação entre z_t e z_{t-k} , com t correspondendo ao período m :

$$\rho^m(k) = E \left[\left(\frac{Z_t - \mu_m}{\sigma_m} \right) \left(\frac{Z_{t-k} - \mu_{m-k}}{\sigma_{m-k}} \right) \right] \quad (2.18)$$

as diversas equações dos períodos $1, \dots, s, \dots$ das funções de autocorrelação $\rho^m(k)$ descrevem a estrutura de dependência temporal da série.

O PAR(p) funciona construindo um modelo AR(p) para cada mês. Podemos descrever o funcionamento do PAR(p) em seis etapas:

1. Leitura dos dados históricos
2. Cálculo da média e desvio padrão de cada mês
3. Cálculo da função de autocorrelação para cada mês, que é feita usando a equação 2.12
4. Cálculo da função de autocorrelação parcial para cada mês. Para isso, é necessária a utilização de um sistema de equações denominado de Yule-Walker, descrito em (1)
5. Identificação da ordem ρ_m do modelo AR de cada mês. Isso é feito resolvendo o sistema de Yule-Walker sucessivas vezes, com diferentes valores de ρ_m .
6. Cálculo dos resíduos históricos, média e desvio padrão

2.5 Testes de Aderência

2.5.1 Teste de Kolmogorov-Smirnov

O teste de Kolmogorov-Smirnov (teste K-S) é um teste não paramétrico de igualdade entre uma distribuição de probabilidade contínua e unidimensional (10). Esse teste pode ser usado para comparar uma amostra com um distribuição de probabilidade de referência, ou comparar duas amostras. Supondo que observamos X_1, \dots, X_n de uma função de distribuição contínua $F(x)$. Queremos testar a hipótese nula $F(x) = F_0(x)$ para todo x , contra a hipótese $F(x) \neq F_0(x)$ para algum x , onde $F_0(x)$ é uma função de probabilidade conhecida. Seja $\hat{F}_n(x)$ a função de distribuição empírica definida por

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^n I[X_i \leq x] \quad (2.19)$$

Assim, para uma amostra, temos o teste K-S

$$M = \max_x |\hat{F}_n(x) - F_0(x)| \quad (2.20)$$

que queremos rejeitar se M for muito grande(4). Quando comparamos duas amostras, temos

$$M = \max_x |\hat{F}_n(x) - \hat{G}_n(x)| \quad (2.21)$$

onde G seria a outra amostra.

2.5.2 Teste T

O teste T é comumente recomendado para testar a igualdade de médias em duas populações normalmente distribuídas (15). Existem diversas variações do teste T. A que usamos e é definida aqui é o teste T para variâncias diferentes. Consideremos duas amostras aleatórias de tamanho n_1 e n_2 , ou seja, $X_{11}, X_{12}, \dots, X_{1n_1}$ e $X_{21}, X_{22}, \dots, X_{2n_2}$. Essas amostras provém das populações 1 e 2, que tem médias μ_1 e μ_2 e variâncias σ_1^2 e σ_2^2 . Temos as seguintes hipóteses:

A estatística t do teste é dada por:

$$t = \frac{|\mu_1 - \mu_2|}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \quad (2.22)$$

Para o cálculo dos graus de liberdade, temos (15)

$$v = \frac{\left(\frac{1}{n_1} + \frac{u}{n_2}\right)^2}{\frac{1}{n_1^2(n_1-1)} + \frac{u^2}{n_2^2(n_2-1)}} \quad (2.23)$$

onde $u = \frac{\sigma_2^2}{\sigma_1^2}$

2.5.3 Teste de Levene

O teste de Levene foi projetado para se checar a igualdade de variâncias (25). Tomando $k \geq 2$ amostras aleatórias independentes X_{i1}, \dots, X_{in_i} ($i = 1, \dots, k$), onde cada i é uma coleção de n_i variáveis aleatórias independentes e identicamente distribuídas, sendo G_i a distribuição, μ_i a média e σ_i^2 a variância, todos desconhecidos. Cada observação i é substituída pelo desvio absoluto da variável em relação à média do grupo. Temos a hipótese H_0

$$H_0 : \sigma_1^2 = \dots = \sigma_k^2, \quad i = 1, \dots, k$$

Que é testada contra a hipótese H_1 :

$$H_1 : \sigma_i \neq \sigma_j, \quad \text{para algum } i \neq j, \quad j = 1, \dots, k$$

Os desvios absolutos das variáveis X_{ij} com relação as médias dos grupos \bar{X}_i (onde $\bar{X}_i = n_i^{-1} \sum_{j=1}^{n_i} X_{ij}$) são dados por $Z_{ij} = |X_{ij} - \bar{X}_i|$, $j = 1, \dots, n_i$, $i = 1, \dots, k$. Definindo a estatística W_0 como sendo

$$W_0 = \left(\frac{n-k}{k-1}\right) \frac{\sum_{i=1}^k n_i (\bar{Z}_i - \bar{Z}_{..})^2}{\sum_{i=1}^k \sum_{j=1}^{n_i} (Z_{ij} - \bar{Z}_i)^2} \quad (2.24)$$

Onde temos $\bar{Z}_i = n_i^{-1} \sum_{j=1}^{n_i} Z_{ij}$, $\bar{Z}_{..} = n^{-1} \sum_{i=1}^k n_i \bar{Z}_i$ e $n = \sum_{i=1}^k n_i$. O teste de Levene é rejeitar H_0 se $W_0 > F_{(k-1, n-k), (1-\alpha)}$, onde $F_{(k-1, n-k), (1-\alpha)}$ representa o quantil de ordem $1 - \alpha$ da distribuição $F_{(k-1, n-k)}$ e α é o nível de significância do teste (25).

3 Modelo PEN e aplicação em ENA

3.1 Descrição do modelo PEN

Processo Estocástico Neural (PEN) é um processo estocástico baseado em redes neurais artificiais (RNA), o que o torna um modelo não-linear capaz de assimilar características de uma série temporal não-linear, estocástica e até periódica (6). Apesar de o PEN não precisar decompor a série temporal ou fazer suposições a respeito dela, o PEN precisa que suas entradas contenham termos passados da série temporal a ser simulada.

Os parâmetros do PEN são os pesos sinápticos das RNAs. Para construirmos uma série sintética dessa série temporal (como uma simulação), precisamos acrescentar um valor aleatório à saída da RNA. Esse valor aleatório é obtido através das distribuições de probabilidades dos resíduos das RNAs do PEN.

A aplicação utilizada no trabalho de Campos (6) e neste trabalho é nas séries hidrológicas sazonais de intervalos mensais, particularmente no contexto do planejamento hidrotérmico de energia no Brasil. A próxima seção dá mais detalhes sobre essa aplicação. Ao lidar com essa aplicação, o PEN se ajusta ao intervalo e ao período da série temporal. Isso significa, nesse caso particular, que o PEN será composto de 12 componentes estocásticas (CE), uma para cada mês, sendo que o mês corresponde ao período da série temporal. Cada CE é uma RNA e uma distribuição de probabilidade para gerar valores aleatórios, e além disso, as CEs são encadeadas, ou seja, a saída de uma CE serve de entrada para a CE seguinte. Assim, se Z_t for uma série temporal com período s e n observações no total, temos a seguinte equação do tempo t :

$$t = (r - 1)s + m \quad (3.1)$$

onde,

r número de observações de cada período da série, $1, \dots, n$

s total de períodos da série

m correspondente a um período da série, 1, ..., s

e no nosso caso particular, $s = 12$, $r = 1, \dots$, ano, m é o mês e n o total de anos observados, que no estudo de caso desse trabalho é 80.

O PEN utilizado usa uma RNA feedforward com uma única camada oculta. Outra definição importante é a de ordem da RNA, que é simplesmente o número de termos anteriores (no nosso caso, meses) passados como entrada a um neurônio da camada escondida da RNA, e é representada por ρ_m .

Para obtermos Z_t , a RNA recebe $Z_{t-1}, Z_{t-2}, \dots, Z_{t-\rho_m}$ e ainda Z_{t-s} , para reforço do aprendizado do caráter periódico da série temporal. A saída de um neurônio da camada oculta é dada por

$$y_i = \varphi \left(\omega_{i,0} Z_{t-s} + \left(\sum_{j=1}^{\rho_m} \omega_{i,j} Z_{t-j} \right) + \theta_i \right) \quad (3.2)$$

onde φ é a função de ativação do neurônio i , $\omega_{i,j}$ é o peso sináptico da conexão entre a entrada j e o neurônio i e θ_i é o bias do neurônio i .

O único neurônio da camada de saída, que recebe a saída da camada oculta (que contém l_m neurônios), tem equação

$$y_{out} = \varphi_{out} \left(\sum_{i=1}^{l_m} \omega_{out,i} \cdot y_i + \theta_{out} \right) \quad (3.3)$$

onde φ_{out} é a função de ativação do neurônio da camada de saída, $\omega_{out,i}$ é o peso sináptico da conexão entre a entrada i (que é um neurônio da camada escondida) e o neurônio da camada de saída e θ_{out} é o bias desse neurônio. Como comentamos anteriormente, a saída de cada CE do PEN é a saída de uma RNA mais um valor aleatório, como na equação

$$Z_t = y_{out} + \alpha(t) \quad (3.4)$$

onde y_{out} é a saída da CE e $\alpha(t)$ o valor aleatório gerado da distribuição de probabilidades da RNA do CE, no período m .

Por fim, através da combinação das equações 3.2, 3.3 e 3.4, chegamos a equação da saída da CE de período m do PEN, dada por 3.5:

$$Z_t = \varphi_{out} \left(\sum_{i=1}^{l_m} \omega_{out,i} \cdot \varphi \left[\omega_{i,0} Z_{t-s} + \left(\sum_{j=1}^{p_m} \omega_{i,j} Z_{t-j} \right) + \theta_i \right] + \theta_{out} \right) + \alpha(t) \quad (3.5)$$

3.2 Aplicação do modelo PEN em ENA

As séries de aflúncias têm natureza não estacionária, por causa dos períodos de seca e cheia durante o ano. Isso somado a sazonalidade e incerteza nas aflúncias dos rios brasileiros, mostra a importância de um modelo como o PEN, que é capaz de capturar essa não-linearidade. Como já comentamos, o SIN é dividido em 4 regiões (subsistemas) interligadas: Sul, Sudeste/Centro-Oeste, Nordeste e Norte. Por isso, no planejamento de médio prazo (que o modelo PEN está focado) ocorre a agregação dos reservatórios das usinas em reservatórios equivalentes de energia, um para cada subsistema (6). Há também a agregação das aflúncias às usinas em ENAs, que como também falamos, energia natural afluyente é uma estimativa da energia que pode ser gerada.

Vamos utilizar o PEN para gerar cenários de ENA. Para avaliar as características das séries geradas, precisamos verificar a aderência dessas séries sintéticas às séries históricas de ENA, aplicando alguns testes de hipóteses relevantes (6). Nossa comparação é feita utilizando esses mesmos testes com as séries sintéticas geradas pela implementação de Campos (6) do PAR(p), baseada na descrição de Macieira (11).

Para cada subsistema do SIN há um sistema do PEN. Esses sistemas são compostos de 12 componentes estocástica (CE), que correspondem aos 12 meses de um ano, isso por que as séries de ENA tem período de 12 meses. Em seguida, cada sistema PEN precisa ser ajustado a sua série de ENA correspondente, e isso é feito em 3 etapas (6), descritas sucintamente aqui:

Determinação da estrutura do modelo: Considerando Z_t o valor de ENA para o tempo t , temos

$$t = (r - 1)s + m$$

onde m corresponde ao mês ($m=1, \dots, s$), $s=12$ e r o ano. Cada CE do PEN recebe valores de ENA. Um desses valores é o do instantes $t-s$, que equivale a um período

inteiro anterior, para reforçar o caráter periódico. Os outros valores são os ρ_m meses anteriores ao mês m , que são os termos auto-regressivos da série. ρ_m foi definida como ordem da série. Esse período, no modelo PEN original assumiu diversos valores: 3, 6, 9 e 11. No estudo de caso desse trabalho, o ρ_m ficou fixado em 11. Outra informação importante que é definida é o número l_m de neurônios na camada escondida, que pode variar de 1 a 20 para cada ordem p_m .

Estimação dos parâmetros: Esses parâmetros são os pesos sinápticos da RNA. O Conjunto de dados históricos usados para realizar esse ajuste foi dividido em dois conjuntos, correspondentes a treinamento e validação. A etapa de validação gera cenários de um número arbitrário de anos com as saídas da RNA. Após o fim do treinamento, a rede com a melhor configuração é escolhida para ser a RNA da CE. Para definir essa melhor configuração o MAPE (erro médio percentual absoluto) (9) é utilizado. No PEN original diversos MAPEs correspondentes a variados horizontes de tempo e aos conjuntos de treinamento e validação. No estudo de caso desse trabalho foram utilizamos apenas o MAPE mensal de validação e de teste.

Validação dos resíduos gerados: Após as duas etapas anteriores serem cumpridas, ocorre o ajuste da série de resíduos obtidos pela diferença entre a saída da RNA e os valores do conjunto de treinamento, que são ajustadas a uma distribuição de probabilidade arbitrária. Em seguida o teste de Kolmogorov-Smirnov é aplicado, para se verificar a aderência entre a série de resíduos e a distribuição de probabilidades. A distribuição com o melhor ajuste é escolhida para compor a distribuição de probabilidade que pertence à CE do PEN, para o mês testado.

4 Estudo de Caso

4.1 Considerações

No estudo de caso desse trabalho, foi usada uma versão mais recente do Matlab (2013a), que a utilizada por (6). O código fonte usado foi revisado e simplificado, utilizando funções mais modernas no Matlab. Fixamos o parâmetro *ordem* na entrada dos dados no PEN em 11, o que significa dizer que estamos utilizando sempre 11 meses anteriores para gerar uma saída. O fato da ordem estar fixa pode ser responsável pela piora de alguns resultados, como será visto adiante nas comparações. Conforme foi comentado anteriormente, a massa de dados utilizada conta com 80 anos, o que são cinco anos a mais que a massa usada no trabalho original do modelo PEN de (6). Isso fez com que a massa de dados pudesse ser dividida em três conjuntos: 70 anos para treinamento, cinco anos para validação e cinco anos para testes. Esses fatos também causaram uma modificação nos resultados esperados. Quanto às funções de ativação de entrada e saída utilizadas, 4 casos são definidos:

Caso PP: A função de ativação da camada escondida é linear, e a função de ativação do neurônio da camada de saída é também linear. Essa combinação foi escolhida de modo a tentar "linearizar" o PEN, de modo a aproximá-lo do PAR(p). Talvez esse caso seja melhor visto como uma regressão linear, em vez de rede neural;

Caso LP: A função de ativação da camada escondida é uma sigmóide, e a função de ativação do neurônio da camada de saída é linear;

Caso TP: A função de ativação da camada escondida é uma tangente hiperbólica, e a função de ativação do neurônio da camada de saída é linear.

Caso TL: A função de ativação da camada escondida é uma tangente hiperbólica, e a função de ativação do neurônio da camada de saída é uma sigmóide. Essa foi a combinação utilizada no trabalho de Campos (6), que como já comentamos, é o

trabalho que criou o PEN.

Para as comparações serem feitas, o MAPE foi utilizado. Para cada tabela, é apresentado o número do mês (1 a 12), o critério escolhido (MAPE de validação ou de teste) e o número de neurônios encontrados na camada escondida que proporcionou esse melhor ajuste (isto é, com menor MAPE).

4.2 Experimentos com o modelo PEN

São descritos agora os resultados obtidos em nossas simulações, divididas por subsistema. São 4 tabelas, uma para cada subsistema, contendo cada combinação de funções de ativação. Cada tabela é dividida em doze meses e em três colunas: critério, valor e número de neurônios (denotado por # neurônios). A coluna critério mostra os dois tipos de MAPE calculados para a rede neural, teste e validação. A coluna valor mostra o percentual calculado (por exemplo, 5.4503%). Todos esses MAPEs obtidos são os melhores de cada mês, obtidos depois de diversas inicializações (variando de 1 a 10) no MATLAB e com diversos números diferentes de neurônios na camada escondida. A coluna # neurônios mostra o número de neurônios na camada escondida que foram utilizados para se chegar a esses MAPEs. A rede neural foi treinada com o algoritmo de LevenbergMarquardt

4.2.1 Subsistema Sudeste

A tabela 4.1 mostra o caso LP no subsistema sudeste. O que podemos perceber na tabela 4.1 é que em 8 dos 12 meses, o MAPE de validação foi menor que o de teste. O melhor mês é novembro.

Tabela 4.1: Caso LP, Subsistema Sudeste

Mês	1	
critério	valor	# neurônios
MAPE Teste	5.4503	7
MAPE Validação	11.7784	6
Mês	2	

Tabela 4.1 – *Continuando da página anterior*

critério	valor	# neurônios
MAPE Teste	4.5075	5
MAPE Validação	6.8020	12
Mês	3	
critério	valor	# neurônios
MAPE Teste	8.2905	18
MAPE Validação	6.0646	8
Mês	4	
critério	valor	# neurônios
MAPE Teste	5.6841	18
MAPE Validação	5.1597	5
Mês	5	
critério	valor	# neurônios
MAPE Teste	4.9608	4
MAPE Validação	2.4361	1
Mês	6	
critério	valor	# neurônios
MAPE Teste	2.0670	12
MAPE Validação	3.9646	11
Mês	7	
critério	valor	# neurônios
MAPE Teste	4.8947	20
MAPE Validação	2.3252	15
Mês	8	
critério	valor	# neurônios
MAPE Teste	5.9102	10
MAPE Validação	2.6574	2
Mês	9	

Tabela 4.1 – *Continuando da página anterior*

critério	valor	# neurônios
MAPE Teste	9.5020	14
MAPE Validação	4.5278	1
Mês	10	
critério	valor	# neurônios
MAPE Teste	14.4039	14
MAPE Validação	11.8040	12
Mês	11	
critério	valor	# neurônios
MAPE Teste	2.7211	3
MAPE Validação	2.3526	8
Mês	12	
critério	valor	# neurônios
MAPE Teste	11.0977	16
MAPE Validação	8.9534	19

Na tabela 4.2 temos o caso TP no subsistema sudeste. Podemos apontar que em 6 dos 12 meses o MAPE de validação é menor que o de teste. O melhor mês é novembro.

Tabela 4.2: Caso TP, Subsistema Sudeste

Mês	1	
critério	valor	# neurônios
MAPE Teste	4.4324	9
MAPE Validação	10.8755	4
Mês	2	
critério	valor	# neurônios
MAPE Teste	3.8199	12
MAPE Validação	6.9340	16
Mês	3	

Tabela 4.2 – *Continuando da página anterior*

critério	valor	# neurônios
MAPE Teste	6.0109	16
MAPE Validação	5.8016	5
Mês	4	
critério	valor	# neurônios
MAPE Teste	2.3138	4
MAPE Validação	5.7981	8
Mês	5	
critério	valor	# neurônios
MAPE Teste	3.3740	5
MAPE Validação	4.6424	2
Mês	6	
critério	valor	# neurônios
MAPE Teste	3.7147	4
MAPE Validação	3.9711	16
Mês	7	
critério	valor	# neurônios
MAPE Teste	3.6642	20
MAPE Validação	2.2094	3
Mês	8	
critério	valor	# neurônios
MAPE Teste	5.3746	2
MAPE Validação	2.3095	1
Mês	9	
critério	valor	# neurônios
MAPE Teste	7.9069	18
MAPE Validação	4.3963	8
Mês	10	

Tabela 4.2 – *Continuando da página anterior*

critério	valor	# neurônios
MAPE Teste	14.5389	1
MAPE Validação	11.7815	20
Mês	11	
critério	valor	# neurônios
MAPE Teste	3.2642	7
MAPE Validação	1.4548	9
Mês	12	
critério	valor	# neurônios
MAPE Teste	8.8860	13
MAPE Validação	10.3074	18

Na tabela 4.3 temos o caso PP no subsistema sudeste. Podemos apontar que em 6 dos meses o MAPE de validação é menor que o de teste, idêntico ao caso TP anterior. O pior mês é outubro. Esse é um resultado melhor que os dois casos anteriores. O melhor mês é novembro.

Tabela 4.3: Caso PP, Subsistema Sudeste

Mês	1	
critério	valor	# neurônios
MAPE Teste	4.9527	13
MAPE Validação	8.8202	13
Mês	2	
critério	valor	# neurônios
MAPE Teste	2.3558	3
MAPE Validação	7.7699	11
Mês	3	
critério	valor	# neurônios
MAPE Teste	4.9638	7

Tabela 4.3 – *Continuando da página anterior*

MAPE Validação	6.8461	7
Mês	4	
critério	valor	# neurônios
MAPE Teste	7.5723	20
MAPE Validação	4.6983	10
Mês	5	
critério	valor	# neurônios
MAPE Teste	4.5370	9
MAPE Validação	4.2737	6
Mês	6	
critério	valor	# neurônios
MAPE Teste	3.8497	15
MAPE Validação	4.5649	19
Mês	7	
critério	valor	# neurônios
MAPE Teste	5.1051	7
MAPE Validação	2.0375	2
Mês	8	
critério	valor	# neurônios
MAPE Teste	4.6251	8
MAPE Validação	2.0971	1
Mês	9	
critério	valor	# neurônios
MAPE Teste	12.4072	20
MAPE Validação	4.9460	2
Mês	10	
critério	valor	# neurônios
MAPE Teste	13.3155	12

Tabela 4.3 – *Continuando da página anterior*

MAPE Validação	9.4250	19
Mês	11	
critério	valor	# neurônios
MAPE Teste	2.6430	3
MAPE Validação	2.8517	3
Mês	12	
critério	valor	# neurônios
MAPE Teste	7.8466	15
MAPE Validação	10.8685	1

Na tabela 4.4 temos o caso TL no subsistema sudeste. Podemos apontar que em 7 dos 12 meses o MAPE de validação é menor que o de teste. O pior mês é dezembro.

Tabela 4.4: Caso TL, Subsistema Sudeste

Mês	1	
critério	valor	# neurônios
MAPE Teste	6.6490	11
MAPE Validação	10.6384	9
Mês	2	
critério	valor	# neurônios
MAPE Teste	3.1076	16
MAPE Validação	6.8034	14
Mês	3	
critério	valor	# neurônios
MAPE Teste	6.2856	5
MAPE Validação	6.5704	17
Mês	4	
critério	valor	# neurônios
MAPE Teste	6.9665	8

Tabela 4.4 – *Continuando da página anterior*

MAPE Validação	6.1060	7
Mês	5	
critério	valor	# neurônios
MAPE Teste	3.8928	1
MAPE Validação	3.7912	10
Mês	6	
critério	valor	# neurônios
MAPE Teste	3.1247	1
MAPE Validação	4.6494	11
Mês	7	
critério	valor	# neurônios
MAPE Teste	6.6390	9
MAPE Validação	2.0422	5
Mês	8	
critério	valor	# neurônios
MAPE Teste	3.6812	5
MAPE Validação	2.5843	9
Mês	9	
critério	valor	# neurônios
MAPE Teste	13.2181	17
MAPE Validação	5.5558	7
Mês	10	
critério	valor	# neurônios
MAPE Teste	13.6400	5
MAPE Validação	7.1640	18
Mês	11	
critério	valor	# neurônios
MAPE Teste	3.0562	15

Tabela 4.4 – *Continuando da página anterior*

MAPE Validação	4.0814	9
Mês	12	
critério	valor	# neurônios
MAPE Teste	11.6212	20
MAPE Validação	11.5040	19

Analisando as tabelas do subsistema sudeste, podemos verificar que o melhor caso é o caso TP.

4.2.2 Subsistema Sul

Na tabela 4.5 temos o caso LP no subsistema sul. Podemos apontar que em 6 dos 12 meses o MAPE de validação é menor que o de teste. O pior mês é outubro. O melhor mês é novembro.

Tabela 4.5: Caso LP, Subsistema Sul

Mês	1	
critério	valor	# neurônios
MAPE Teste	5.8830	3
MAPE Validação	10.6991	11
Mês	2	
critério	valor	# neurônios
MAPE Teste	5.1323	15
MAPE Validação	7.5409	6
Mês	3	
critério	valor	# neurônios
MAPE Teste	9.3370	1
MAPE Validação	5.8206	11
Mês	4	
critério	valor	# neurônios

Tabela 4.5 – *Continuando da página anterior*

MAPE Teste	6.0869	13
MAPE Validação	6.5837	14
Mês	5	
critério	valor	# neurônios
MAPE Teste	4.4535	7
MAPE Validação	3.1437	14
Mês	6	
critério	valor	# neurônios
MAPE Teste	3.3626	1
MAPE Validação	4.3309	11
Mês	7	
critério	valor	# neurônios
MAPE Teste	4.6594	19
MAPE Validação	2.1162	1
Mês	8	
critério	valor	# neurônios
MAPE Teste	4.9897	12
MAPE Validação	2.2997	2
Mês	9	
critério	valor	# neurônios
MAPE Teste	9.4709	20
MAPE Validação	3.7813	9
Mês	10	
critério	valor	# neurônios
MAPE Teste	13.3263	20
MAPE Validação	8.2004	17
Mês	11	
critério	valor	# neurônios

Tabela 4.5 – *Continuando da página anterior*

MAPE Teste	2.9835	10
MAPE Validação	3.2259	6
Mês	12	
critério	valor	# neurônios
MAPE Teste	9.7938	16
MAPE Validação	10.0059	14

Na tabela 4.6 temos o caso TP no subsistema sul. Podemos apontar que em 7 dos 12 meses o MAPE de validação foi menor que o de teste. O pior mês é outubro. O melhor mês é novembro.

Tabela 4.6: Caso TP, Subsistema Sul

Mês	1	
critério	valor	# neurônios
MAPE Teste	7.0649	12
MAPE Validação	8,5501	19
Mês	2	
critério	valor	# neurônios
MAPE Teste	4.0614	9
MAPE Validação	7.1396	16
Mês	3	
critério	valor	# neurônios
MAPE Teste	5.9610	3
MAPE Validação	7.3770	20
Mês	4	
critério	valor	# neurônios
MAPE Teste	7.3584	10
MAPE Validação	6.6984	11
Mês	5	

Tabela 4.6 – *Continuando da página anterior*

critério	valor	# neurônios
MAPE Teste	2.2244	13
MAPE Validação	3.9421	14
Mês	6	
critério	valor	# neurônios
MAPE Teste	4.1933	2
MAPE Validação	4.1471	7
Mês	7	
critério	valor	# neurônios
MAPE Teste	5.5409	11
MAPE Validação	2.3516	10
Mês	8	
critério	valor	# neurônios
MAPE Teste	5.8855	10
MAPE Validação	1.8186	3
Mês	9	
critério	valor	# neurônios
MAPE Teste	12.7881	14
MAPE Validação	5.1049	14
Mês	10	
critério	valor	# neurônios
MAPE Teste	12.9990	10
MAPE Validação	11.4393	5
Mês	11	
critério	valor	# neurônios
MAPE Teste	2.1204	4
MAPE Validação	2.5562	8
Mês	12	

Tabela 4.6 – *Continuando da página anterior*

critério	valor	# neurônios
MAPE Teste	9.3322	2
MAPE Validação	7.8301	10

Na tabela 4.7 temos o caso PP no subsistema sul. Podemos apontar que em 8 dos 12 meses o MAPE de validação foi menor que o de teste. O pior mês é novamente outubro. O melhor mês é novembro.

Tabela 4.7: Caso PP, Subsistema Sul

Mês	1	
critério	valor	# neurônios
MAPE Teste	6.2967	8
MAPE Validação	9.2120	4
Mês	2	
critério	valor	# neurônios
MAPE Teste	4.8473	2
MAPE Validação	5.9920	18
Mês	3	
critério	valor	# neurônios
MAPE Teste	6.1564	2
MAPE Validação	6.8175	12
Mês	4	
critério	valor	# neurônios
MAPE Teste	6.7414	17
MAPE Validação	5.2422	18
Mês	5	
critério	valor	# neurônios
MAPE Teste	6.4043	8
MAPE Validação	4.9802	4

Tabela 4.7 – *Continuando da página anterior*

Mês	6	
critério	valor	# neurônios
MAPE Teste	3.4284	12
MAPE Validação	4.5894	3
Mês	7	
critério	valor	# neurônios
MAPE Teste	6.7807	11
MAPE Validação	1.2720	10
Mês	8	
critério	valor	# neurônios
MAPE Teste	4.9392	19
MAPE Validação	2.4544	6
Mês	9	
critério	valor	# neurônios
MAPE Teste	13.7136	20
MAPE Validação	4.4745	7
Mês	10	
critério	valor	# neurônios
MAPE Teste	14.0340	16
MAPE Validação	10.3235	17
Mês	11	
critério	valor	# neurônios
MAPE Teste	3.0998	4
MAPE Validação	2.9788	7
Mês	12	
critério	valor	# neurônios
MAPE Teste	7.5997	5
MAPE Validação	6.3162	4

Na tabela 4.8 temos o caso TL no subsistema sul. Podemos apontar que em 7

dos 12 meses o MAPE de validação foi menor que o de teste. O pior mês é outubro. O melhor mês é novembro.

Tabela 4.8: Caso TL, Subsistema Sul

Mês	1	
critério	valor	# neurônios
MAPE Teste	9.2030	7
MAPE Validação	10.6134	8
Mês	2	
critério	valor	# neurônios
MAPE Teste	5.1656	2
MAPE Validação	6.4742	7
Mês	3	
critério	valor	# neurônios
MAPE Teste	5.6390	13
MAPE Validação	6.1835	16
Mês	4	
critério	valor	# neurônios
MAPE Teste	5.4567	3
MAPE Validação	5.3596	19
Mês	5	
critério	valor	# neurônios
MAPE Teste	5.4813	4
MAPE Validação	4.5837	1
Mês	6	
critério	valor	# neurônios
MAPE Teste	3.9585	3
MAPE Validação	3.8176	8
Mês	7	
critério	valor	# neurônios

Tabela 4.8 – *Continuando da página anterior*

MAPE Teste	5.8700	18
MAPE Validação	1.8729	1
Mês	8	
critério	valor	# neurônios
MAPE Teste	5.8924	15
MAPE Validação	2.1919	2
Mês	9	
critério	valor	# neurônios
MAPE Teste	12.1239	10
MAPE Validação	3.3482	8
Mês	10	
critério	valor	# neurônios
MAPE Teste	15.1559	5
MAPE Validação	10.9831	3
Mês	11	
critério	valor	# neurônios
MAPE Teste	3.1382	4
MAPE Validação	1.6436	15
Mês	12	
critério	valor	# neurônios
MAPE Teste	7.1094	9
MAPE Validação	9.0052	12

Analisando as tabelas do subsistema sul, podemos verificar que o melhor caso é o caso TL.

4.2.3 Subsistema Nordeste

Na tabela 4.9 temos o caso LP no subsistema nordeste. Podemos apontar que em 8 dos 12 meses o MAPE de validação foi menor que o que teste. O pior mês é outubro. O

melhor mês é junho.

Tabela 4.9: Caso LP, Subsistema Nordeste

Mês	1	
critério	valor	# neurônios
MAPE Teste	6.6318	20
MAPE Validação	9.7828	5
Mês	2	
critério	valor	# neurônios
MAPE Teste	4.8578	8
MAPE Validação	6.9107	11
Mês	3	
critério	valor	# neurônios
MAPE Teste	8.9434	5
MAPE Validação	7.7205	14
Mês	4	
critério	valor	# neurônios
MAPE Teste	6.1532	16
MAPE Validação	5.9034	8
Mês	5	
critério	valor	# neurônios
MAPE Teste	5.4500	11
MAPE Validação	4.7808	3
Mês	6	
critério	valor	# neurônios
MAPE Teste	2.6982	1
MAPE Validação	3.9014	17
Mês	7	
critério	valor	# neurônios
MAPE Teste	6.4449	3

Tabela 4.9 – *Continuando da página anterior*

MAPE Validação	2.2167	5
Mês	8	
critério	valor	# neurônios
MAPE Teste	5.4034	2
MAPE Validação	2.8669	3
Mês	9	
critério	valor	# neurônios
MAPE Teste	12.1025	8
MAPE Validação	2.8046	17
Mês	10	
critério	valor	# neurônios
MAPE Teste	10.9916	17
MAPE Validação	10.6912	10
Mês	11	
critério	valor	# neurônios
MAPE Teste	4.1144	13
MAPE Validação	2.2953	8
Mês	12	
critério	valor	# neurônios
MAPE Teste	10.9364	11
MAPE Validação	8.5209	16

Na tabela 4.10 temos o caso TP no subsistema nordeste. Podemos apontar que em 8 dos 12 meses o MAPE de validação foi menor que o de teste. O pior mês continua sendo outubro. Esse resultado é ligeiramente melhor que o do caso anterior, caso LP. Apesar de o número de MAPEs de validação menores ser igual ao do caso anterior, a distribuição (i. e., os meses) é diferente. De todos os casos vistos até agora, esse caso TP foi o primeiro a apresentar MAPEs de validação menores nos primeiros meses do ano (janeiro, março e abril). O melhor mês é novembro.

Tabela 4.10: Caso TP, Subsistema Nordeste

Mês	1	
critério	valor	# neurônios
MAPE Teste	10.1253	17
MAPE Validação	7.5579	7
Mês	2	
critério	valor	# neurônios
MAPE Teste	4.2239	7
MAPE Validação	7.1895	3
Mês	3	
critério	valor	# neurônios
MAPE Teste	9.0599	18
MAPE Validação	5.8158	8
Mês	4	
critério	valor	# neurônios
MAPE Teste	7.6571	17
MAPE Validação	5.3212	3
Mês	5	
critério	valor	# neurônios
MAPE Teste	3.8356	10
MAPE Validação	4.2161	1
Mês	6	
critério	valor	# neurônios
MAPE Teste	2.5817	15
MAPE Validação	3.2632	9
Mês	7	
critério	valor	# neurônios
MAPE Teste	6.5256	18

Tabela 4.10 – *Continuando da página anterior*

MAPE Validação	2.3590	7
Mês	8	
critério	valor	# neurônios
MAPE Teste	5.1960	7
MAPE Validação	2.0119	5
Mês	9	
critério	valor	# neurônios
MAPE Teste	13.5983	18
MAPE Validação	4.6065	1
Mês	10	
critério	valor	# neurônios
MAPE Teste	11.1771	3
MAPE Validação	7.9289	17
Mês	11	
critério	valor	# neurônios
MAPE Teste	2.1705	14
MAPE Validação	3.7812	19
Mês	12	
critério	valor	# neurônios
MAPE Teste	7.8359	17
MAPE Validação	7.6891	12

Na tabela 4.11 temos o caso PP no subsistema nordeste. Podemos apontar que em 7 dos 12 meses o MAPE de validação foi menor que o de teste. O pior mês é janeiro. O melhor mês é novembro.

Tabela 4.11: Caso PP, Subsistema Nordeste

Mês	1	
critério	valor	# neurônios

Tabela 4.11 – *Continuando da página anterior*

MAPE Teste	8.6407	10
MAPE Validação	9.7797	14
Mês	2	
critério	valor	# neurônios
MAPE Teste	4.2594	7
MAPE Validação	6.1104	7
Mês	3	
critério	valor	# neurônios
MAPE Teste	6.5518	10
MAPE Validação	3.7463	14
Mês	4	
critério	valor	# neurônios
MAPE Teste	5.8003	15
MAPE Validação	5.5756	13
Mês	5	
critério	valor	# neurônios
MAPE Teste	2.5209	2
MAPE Validação	4.4158	3
Mês	6	
critério	valor	# neurônios
MAPE Teste	3.9399	16
MAPE Validação	2.3739	14
Mês	7	
critério	valor	# neurônios
MAPE Teste	4.9147	19
MAPE Validação	1.7056	3
Mês	8	
critério	valor	# neurônios

Tabela 4.11 – *Continuando da página anterior*

MAPE Teste	4.5062	18
MAPE Validação	1.7615	1
Mês	9	
critério	valor	# neurônios
MAPE Teste	5.4603	11
MAPE Validação	5.1249	18
Mês	10	
critério	valor	# neurônios
MAPE Teste	18.1062	1
MAPE Validação	6.4588	12
Mês	11	
critério	valor	# neurônios
MAPE Teste	2.2318	20
MAPE Validação	2.1525	15
Mês	12	
critério	valor	# neurônios
MAPE Teste	7.7071	15
MAPE Validação	9.5064	10

Na tabela 4.12 temos o caso TL no subsistema nordeste. Podemos apontar que em 7 dos 12 meses o MAPE de validação foi menor que o de teste. O pior mês é janeiro. O melhor mês é novembro.

Tabela 4.12: Caso TL, Subsistema Nordeste

Mês	1	
critério	valor	# neurônios
MAPE Teste	4.1457	17
MAPE Validação	10.0605	7
Mês	2	

Tabela 4.12 – *Continuando da página anterior*

critério	valor	# neurônios
MAPE Teste	3.8715	8
MAPE Validação	7.1501	8
Mês	3	
critério	valor	# neurônios
MAPE Teste	6.9816	13
MAPE Validação	5.0021	9
Mês	4	
critério	valor	# neurônios
MAPE Teste	6.3386	16
MAPE Validação	6.1389	20
Mês	5	
critério	valor	# neurônios
MAPE Teste	4.0561	5
MAPE Validação	4.2703	7
Mês	6	
critério	valor	# neurônios
MAPE Teste	2.5630	4
MAPE Validação	4.1017	11
Mês	7	
critério	valor	# neurônios
MAPE Teste	4.6442	16
MAPE Validação	2.5292	1
Mês	8	
critério	valor	# neurônios
MAPE Teste	5.5058	14
MAPE Validação	2.3376	11
Mês	9	

Tabela 4.12 – *Continuando da página anterior*

critério	valor	# neurônios
MAPE Teste	11.1546	6
MAPE Validação	5.8273	1
Mês	10	
critério	valor	# neurônios
MAPE Teste	13.4576	3
MAPE Validação	9.1778	12
Mês	11	
critério	valor	# neurônios
MAPE Teste	2.0617	10
MAPE Validação	3.4053	1
Mês	12	
critério	valor	# neurônios
MAPE Teste	10.9349	5
MAPE Validação	7.6711	11

Analisando todas as tabelas do subsistema nordeste, temos que o melhor caso é o caso PP.

4.2.4 Subsistema Norte

Na tabela 4.13 temos o caso LP no subsistema norte. Podemos apontar que em 7 dos 12 meses o MAPE de validação foi menor que o de teste. O pior mês é outubro. O melhor mês é junho.

Tabela 4.13: Caso LP, Subsistema Norte

Mês	1	
critério	valor	# neurônios
MAPE Teste	7.0464	7
MAPE Validação	6.1563	9

Tabela 4.13 – *Continuando da página anterior*

Mês	2	
critério	valor	# neurônios
MAPE Teste	4.2669	10
MAPE Validação	7.4281	8
Mês	3	
critério	valor	# neurônios
MAPE Teste	5.3890	20
MAPE Validação	6.8658	9
Mês	4	
critério	valor	# neurônios
MAPE Teste	7.0689	6
MAPE Validação	7.6046	9
Mês	5	
critério	valor	# neurônios
MAPE Teste	4.2642	6
MAPE Validação	3.5766	14
Mês	6	
critério	valor	# neurônios
MAPE Teste	1.7641	5
MAPE Validação	4.5000	4
Mês	7	
critério	valor	# neurônios
MAPE Teste	6.2444	4
MAPE Validação	2.3860	3
Mês	8	
critério	valor	# neurônios
MAPE Teste	5.8073	17
MAPE Validação	2.6878	17

Tabela 4.13 – *Continuando da página anterior*

Mês	9	
critério	valor	# neurônios
MAPE Teste	11.2179	16
MAPE Validação	4.8013	11
Mês	10	
critério	valor	# neurônios
MAPE Teste	13.3120	6
MAPE Validação	8.3155	14
Mês	11	
critério	valor	# neurônios
MAPE Teste	3.0374	7
MAPE Validação	3.5698	8
Mês	12	
critério	valor	# neurônios
MAPE Teste	11.0647	12
MAPE Validação	5.0160	4

Na tabela 4.14 temos o caso TP no subsistema norte. Podemos apontar que em 7 dos 12 meses o MAPE de validação foi menor que o de teste. O pior mês é outubro novamente. O melhor mês é novembro.

Tabela 4.14: Caso TP, Subsistema Norte

Mês	1	
critério	valor	# neurônios
MAPE Teste	7.9689	12
MAPE Validação	10.1249	1
Mês	2	
critério	valor	# neurônios
MAPE Teste	2.8603	5

Tabela 4.14 – *Continuando da página anterior*

MAPE Validação	4.8958	7
Mês	3	
critério	valor	# neurônios
MAPE Teste	6.6600	16
MAPE Validação	6.4532	6
Mês	4	
critério	valor	# neurônios
MAPE Teste	5.0851	2
MAPE Validação	6.4331	8
Mês	5	
critério	valor	# neurônios
MAPE Teste	4.2682	8
MAPE Validação	4.6953	7
Mês	6	
critério	valor	# neurônios
MAPE Teste	2.8927	9
MAPE Validação	4.1291	10
Mês	7	
critério	valor	# neurônios
MAPE Teste	6.0158	9
MAPE Validação	2.6718	13
Mês	8	
critério	valor	# neurônios
MAPE Teste	6.3799	12
MAPE Validação	2.1809	7
Mês	9	
critério	valor	# neurônios
MAPE Teste	13.2491	16

Tabela 4.14 – *Continuando da página anterior*

MAPE Validação	4.6983	20
Mês	10	
critério	valor	# neurônios
MAPE Teste	16.7179	17
MAPE Validação	11.4713	11
Mês	11	
critério	valor	# neurônios
MAPE Teste	3.8938	4
MAPE Validação	2.9925	14
Mês	12	
critério	valor	# neurônios
MAPE Teste	8.9133	18
MAPE Validação	5.1806	12

Na tabela 4.15 temos o caso PP no subsistema norte. Podemos apontar que em 9 dos 12 meses o MAPE de validação foi menor que o de teste. O melhor mês é julho.

Tabela 4.15: Caso PP, Subsistema Norte

Mês	1	
critério	valor	# neurônios
MAPE Teste	7.9153	17
MAPE Validação	7.3871	8
Mês	2	
critério	valor	# neurônios
MAPE Teste	6.0801	19
MAPE Validação	6.7280	17
Mês	3	
critério	valor	# neurônios
MAPE Teste	5.3420	20

Tabela 4.15 – *Continuando da página anterior*

MAPE Validação	5.2724	8
Mês	4	
critério	valor	# neurônios
MAPE Teste	7.3759	4
MAPE Validação	5.1186	3
Mês	5	
critério	valor	# neurônios
MAPE Teste	3.3294	10
MAPE Validação	4.1598	6
Mês	6	
critério	valor	# neurônios
MAPE Teste	2.6987	1
MAPE Validação	4.6496	8
Mês	7	
critério	valor	# neurônios
MAPE Teste	3.8084	13
MAPE Validação	1.9835	3
Mês	8	
critério	valor	# neurônios
MAPE Teste	6.0499	5
MAPE Validação	2.0088	1
Mês	9	
critério	valor	# neurônios
MAPE Teste	10.1402	19
MAPE Validação	6.0079	10
Mês	10	
critério	valor	# neurônios
MAPE Teste	15.3282	12

Tabela 4.15 – *Continuando da página anterior*

MAPE Validação	6.0195	7
Mês	11	
critério	valor	# neurônios
MAPE Teste	3.2858	6
MAPE Validação	2.7364	3
Mês	12	
critério	valor	# neurônios
MAPE Teste	10.0214	5
MAPE Validação	9.2073	13

Na tabela 4.16 temos o caso TL no subsistema norte. Podemos apontar que em 6 dos 12 meses o MAPE de validação foi menor que o de teste. O melhor mês é novembro.

Tabela 4.16: Caso TL, Subsistema Norte

Mês	1	
critério	valor	# neurônios
MAPE Teste	7.6277	8
MAPE Validação	10.9842	13
Mês	2	
critério	valor	# neurônios
MAPE Teste	4.7214	6
MAPE Validação	6.4578	8
Mês	3	
critério	valor	# neurônios
MAPE Teste	6.5855	7
MAPE Validação	3.9478	17
Mês	4	
critério	valor	# neurônios
MAPE Teste	4.8341	13

Tabela 4.16 – *Continuando da página anterior*

MAPE Validação	5.2608	11
Mês	5	
critério	valor	# neurônios
MAPE Teste	3.9277	2
MAPE Validação	3.1200	7
Mês	6	
critério	valor	# neurônios
MAPE Teste	1.6638	1
MAPE Validação	4.5109	5
Mês	7	
critério	valor	# neurônios
MAPE Teste	6.4882	13
MAPE Validação	2.7681	1
Mês	8	
critério	valor	# neurônios
MAPE Teste	6.8076	12
MAPE Validação	2.5065	1
Mês	9	
critério	valor	# neurônios
MAPE Teste	13.9495	16
MAPE Validação	4.0698	8
Mês	10	
critério	valor	# neurônios
MAPE Teste	11.4159	6
MAPE Validação	11.5668	4
Mês	11	
critério	valor	# neurônios
MAPE Teste	2.5055	9

Tabela 4.16 – *Continuando da página anterior*

MAPE Validação	3.5088	12
Mês	12	
critério	valor	# neurônios
MAPE Teste	8.7806	10
MAPE Validação	9.5966	7

Analisando todas as tabelas do subsistema norte, temos que o melhor caso é o caso LP.

4.2.5 Comparação com os testes de aderência do PEN original e PAR(p)

As tabelas a seguir mostram os resultados dos cenários gerados usando o nosso modelo PEN, o PEN original de (6) e o PAR(p). Em negrito e sombreado temos os melhores resultados de cada subsistema. A Tabela 4.17 mostra a análise de cenários do PEN no subsistema sudeste. O melhor caso é TL.

Tabela 4.17: Testes de aderência, subsistema sudeste.

	Histórico Teste			Histórico Validação		
	Hipótese Teste %			Hipótese Teste %		
Caso	t	Levene	K-S	t	Levene	K-S
LP	72	80	65	69	92	55
TP	84	80	89	94	90	85
PP	89	94	89	92	97	74
TL	90	97	85	94	100	87

A Tabela 4.18 mostra a análise de cenários do PEN no subsistema sul. O melhor caso é PP.

Tabela 4.18: Testes de aderência no subsistema sul

Tabela 4.18 – *Continuando da página anterior*

	Histórico Teste			Histórico Validação		
	Hipótese Teste %			Hipótese Teste %		
Caso	t	Levene	K-S	t	Levene	K-S
LP	75	80	84	100	92	95
TP	92	80	90	79	99	62
PP	90	89	94	90	100	70
TL	85	90	87	84	85	65

A Tabela 4.19 mostra a análise de cenários do PEN no subsistema nordeste. O melhor caso é LP.

Tabela 4.19: Testes de aderência no subsistema nordeste

	Histórico Teste			Histórico Validação		
	Hipótese Teste %			Hipótese Teste %		
Caso	t	Levene	K-S	t	Levene	K-S
LP	92	94	92	94	99	84
TP	85	90	85	80	95	77
PP	72	87	72	87	92	70
TL	75	72	90	90	90	77

A Tabela 4.20 mostra a análise de cenários do PEN no subsistema norte. O melhor caso é LP.

Tabela 4.20: Testes de aderência no subsistema norte

	Histórico Teste			Histórico Validação		
	Hipótese Teste %			Hipótese Teste %		
Caso	t	Levene	K-S	t	Levene	K-S
LP	90	97	90	99	99	85
TP	72	80	74	90	97	75

Tabela 4.20 – *Continuando da página anterior*

PP	85	85	89	95	90	79
TL	80	87	85	100	94	94

A Figura 4.1 mostra os resultados dos testes realizados com o PAR(p), de acordo com as especificações de (6), para cada subsistema, sendo: (1) sudeste, (2) sul, (3) nordeste e (4) norte.

Subsistema	Teste	Histórico Teste	Histórico Estimativa
1	T	100%	90%
	Levene	100%	99%
	K-S	100%	94%
2	T	97%	100%
	Levene	100%	100%
	K-S	100%	92%
3	T	47%	89%
	Levene	100%	99%
	K-S	50%	95%
4	T	95%	99%
	Levene	99%	99%
	K-S	75%	99%

Figura 4.1: Análise dos cenários gerados no PAR(p). (6)

A Tabela 4.21 é um compêndio de todos os melhores resultados obtidos por (6), para cada subsistema e para os históricos teste e estimativa.

Tabela 4.21: Testes de aderência, modelo PEN original.

(6)

	Histórico Teste			Histórico Estimativa		
	Hipótese Teste %			Hipótese Teste %		
Subsistema	t	Levene	K-S	t	Levene	K-S
Sudeste	97	100	97	94	95	94

Tabela 4.21 – *Continuando da página anterior*

Sul	95	99	99	75	90	90
Nordeste	69	64	54	97	99	94
Norte	74	99	67	85	90	64

Para comparar os resultados do subsistema sudeste, basta examinarmos as Tabelas 4.17 e 4.21 e a Figura 4.1. Podemos ver que no histórico de teste, o PAR(p) é claramente o vencedor, com 100% em todos os testes. Já no histórico estimativa (que comparamos com a nossa validação), vemos que os nossos resultados estão competitivos, sendo melhores nos testes t e Levene que o PAR(p), e tendo resultado próximos ao PEN original de (6).

Tomando as Tabelas 4.18 e 4.21 e a Figura 4.1, podemos comparar os resultados do subsistema sul. Aqui o PAR(p) é claramente vencedor em todos os testes. No histórico teste o PEN original foi melhor; no histórico validação os nossos resultados perdem apenas no teste de Kolmogorov-Smirnov.

Se considerarmos as Tabelas 4.19 e 4.21 e a Figura 4.1, podemos comparar os resultados do subsistema nordeste. No histórico teste os nossos resultados foram vencedores em todos os testes. No histórico validação o PEN original foi o vencedor, embora tanto os nossos resultados quanto o PAR(p) tenham resultados bons, competitivos.

Finalmente, para comparar os resultados do subsistema norte nós tomamos as Tabelas 4.20 e 4.21 e a Figura 4.1. No histórico teste o PAR(p) é o vencedor, embora perca no teste K-S. Os nossos resultados estão competitivos, superando o PEN original. No histórico validação, o PAR(p) vence novamente, embora os nossos resultados percam apenas no teste K-S.

4.3 Gráficos PEN

Nesta seção vamos fazer uma inspeção visual dos nossos casos. Todos os gráficos são compostos da envoltória do cenário (A maior parte em preto), a média histórica (a curva contínua) e a média dos cenários (a curva contínua marcada por cruces, onde cada cruz é um mês). No eixo horizontal nós temos o tempo variando de 0 a 60 meses. No eixo

vertical temos os valores de ENA. A seguir vamos mostrar cada um dos 32 gráficos.

4.3.1 Subsistema Sudeste

Os gráficos 4.2 e 4.3 são correspondentes às envoltórias de teste e validação do caso LP, subsistema sudeste. No gráfico 4.2, podemos notar que apesar de a média dos cenários seguir a forma geral da média histórica, o ajuste da média dos cenários é bem fraco, com poucos meses (cruzes) em cima da da média histórica. Há problemas em particular em situações de máximo, em que a média dos cenários tende a ultrapassar a média histórica (seria o mês de fevereiro). Após o primeiro ano, a média dos cenários está sempre maior que a média histórica. No gráfico 4.3, podemos notar que a curva da média dos cenários acompanha a média histórica, mas o ajuste é também ruim. O primeiro ano apresenta os melhores ajustes (as cruzeiros estão bem próximas da curva da média histórica), mas nos anos seguintes há diversos problemas. Há problemas nos picos históricos, em que a média de cenário fica abaixo. O ajuste é particularmente fora nos meses de março e dezembro.

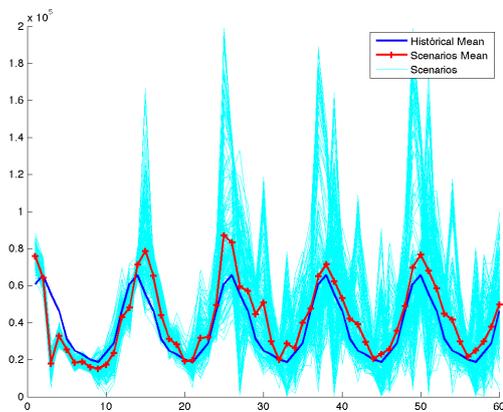


Figura 4.2: Envoltória teste, LP

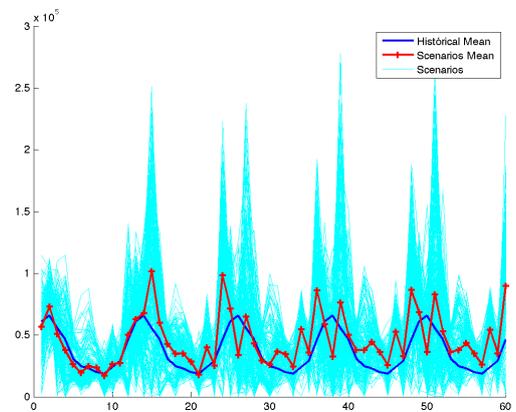


Figura 4.3: Envoltória validação, LP

Os gráficos 4.4 e 4.5 são correspondentes às envoltórias de teste e validação do caso TP, subsistema sudeste. No gráfico 4.4, podemos notar que o ajuste está muito melhor que no caso anterior, embora ainda haja erros. Novamente percebemos problemas nos picos, e um dos meses de janeiro está bem desajustado. No gráfico 4.5, podemos notar que o ajuste continua parecido em relação ao gráfico de testes. Temos problemas em quase todos os meses de fevereiro (que correspondem aos picos).

Os gráficos 4.6 e 4.7 são correspondentes às envoltórias de teste e validação

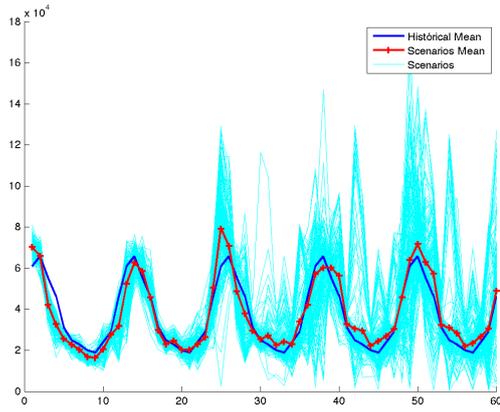


Figura 4.4: Envoltória teste, TP

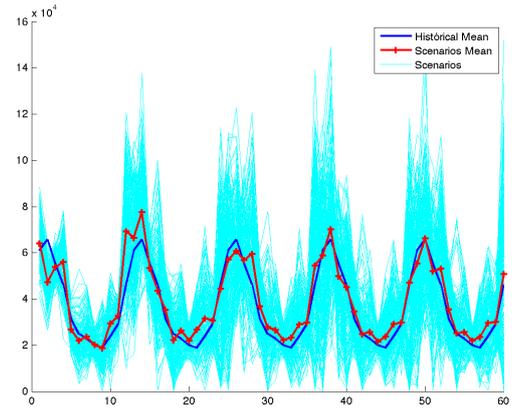


Figura 4.5: Envoltória validação, TP

do caso PP, subsistema sudeste. No gráfico 4.6, podemos notar que, com exceção do primeiro ano o ajuste está muito bom. Há erros nos dois últimos picos, e quase o primeiro ano todo está abaixo da média histórica. No gráfico 4.7, podemos notar que embora o ajuste do primeiro ano tenha melhorado bastante em relação ao gráfico anterior, o curva como um todo está ligeiramente pior, embora mais simétrica. Percebemos também um erro pronunciado nos meses de junho.

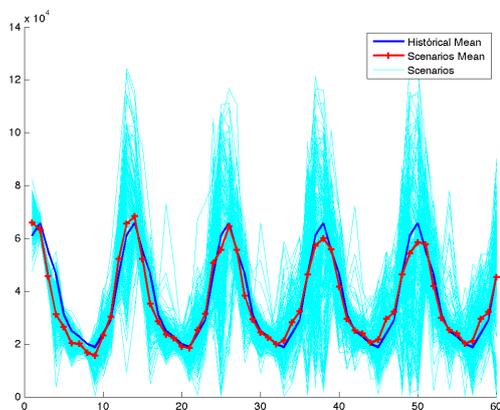


Figura 4.6: Envoltória teste, PP

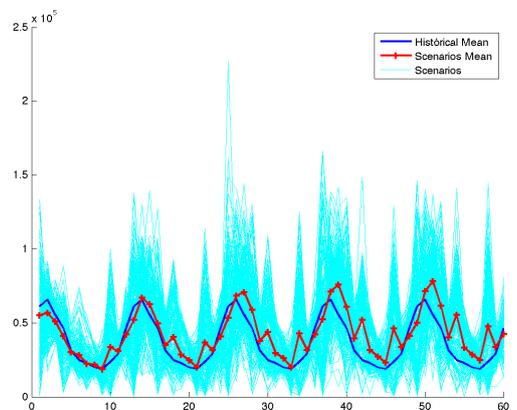


Figura 4.7: Envoltória validação, PP

Os gráficos 4.8 e 4.9 são correspondentes às envoltórias de teste e validação do caso TL, subsistema sudeste. No gráfico 4.8, podemos notar que a curva de estimativa de cenários acompanha o histórico, com algumas falhas. O primeiro janeiro está completamente fora. Há problemas nos picos também, e há razoável simetria. No gráfico 4.9, podemos notar que o ajuste do primeiro ano melhorou em relação ao gráfico anterior, particularmente o primeiro mês de janeiro. Os picos estão ainda mais desajustados e os

meses de janeiro subsequentes também. É interessante perceber que apesar do caso TL ganhar nos teste de aderência, visualmente ele está pior que o caso PP. Se observarmos novamente a Tabela 4.17, vemos que a diferença de aderência entre os casos TL e PP não é grande - por isso o modelo escolhido para a ser o PP e não o TL.

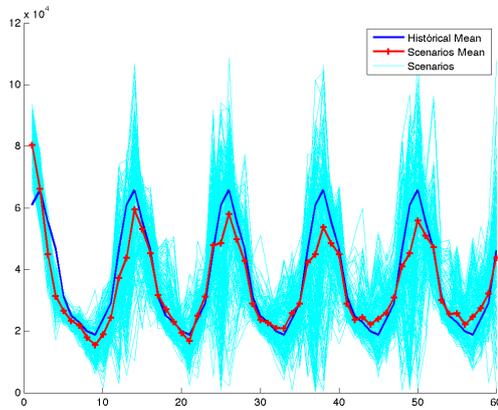


Figura 4.8: Envoltória teste, TL

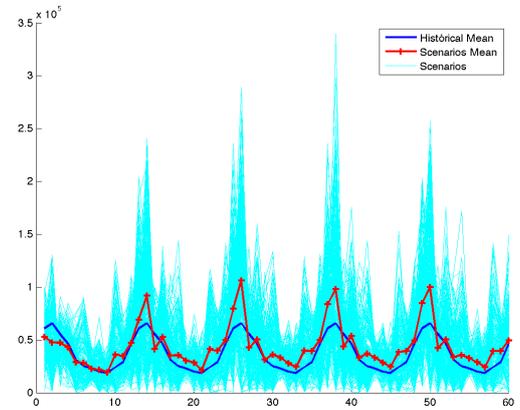


Figura 4.9: Envoltória validação, TL

4.3.2 Substema Sul

Os gráficos 4.10 e 4.11 são correspondentes às envoltórias de teste e validação do caso LP, substema sul. No gráfico 4.10, temos problemas nos picos, mas a curva da média acompanha bem a curva de média do histórico. No gráfico 4.11, temos que os problemas nos picos continuam, e o mês de janeiro também é problemático agora.

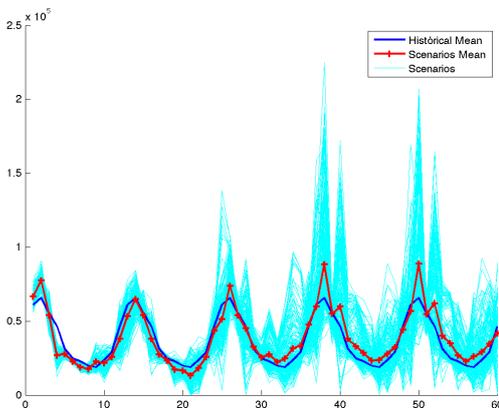


Figura 4.10: Envoltória teste, LP

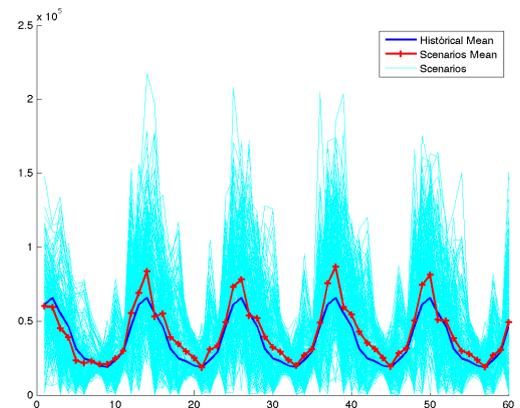


Figura 4.11: Envoltória validação, LP

Os gráficos 4.12 e 4.13 são correspondentes às envoltórias de teste e validação do caso TP, substema sul. No gráfico 4.12, Vemos que o ajuste da curva de médias

de cenário é muito bom, o melhor até agora. Chama atenção o fato de o primeiro mês de janeiro estar muito desajustado, e os picos também, embora de maneira bem mais discreta. No gráfico 4.13, temos a curva está muito desajustada. Curioso notar que o primeiro mês de janeiro está correto. Os meses janeiro, fevereiro, março e maio estão muito desajustados.

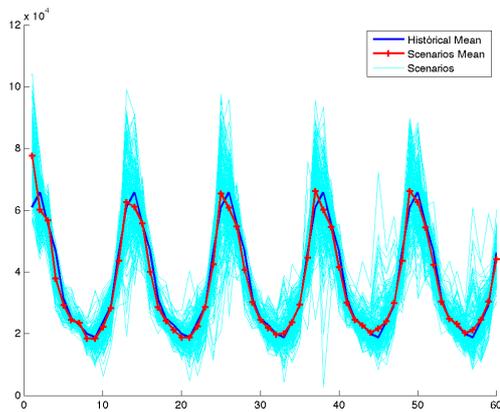


Figura 4.12: Envoltória teste, TP

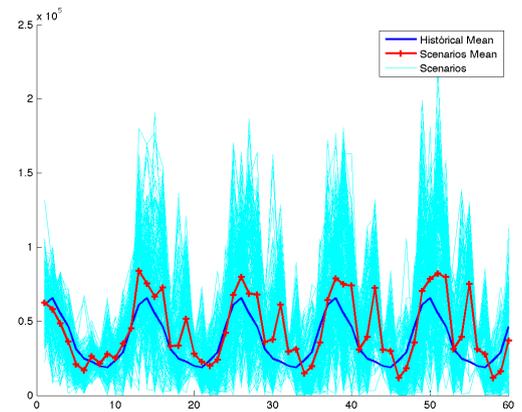


Figura 4.13: Envoltória validação, TP

Os gráficos 4.14 e 4.15 são correspondentes às envoltórias de teste e validação do caso PP, subsistema sul. No gráfico 4.14, podemos perceber que o ajuste também é muito bom, com numerosos meses corretos. Vale notar que o primeiro mês de janeiro está desajustado, e o segundo fevereiro está abaixo da média histórica. No gráfico 4.15, temos a curva está novamente muito desajustada. Os meses de fevereiro, março e abril estão particularmente ruins. Interessante notar que o ajuste do primeiro ano é bem razoável.

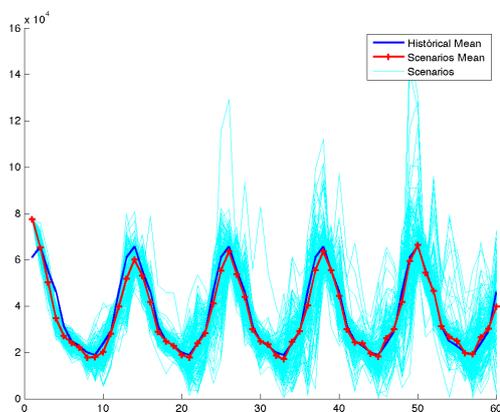


Figura 4.14: Envoltória teste, PP

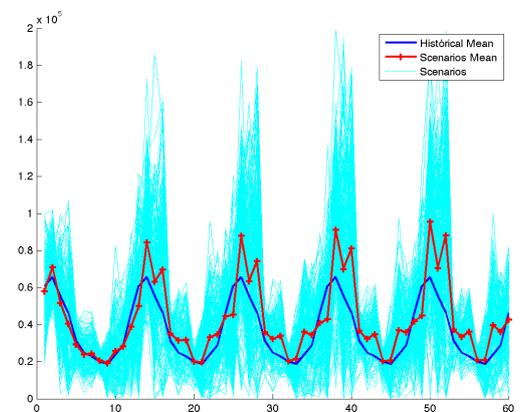


Figura 4.15: Envoltória validação, PP

Os gráficos 4.16 e 4.17 são correspondentes às envoltórias de teste e validação do

caso TL, subsistema sul. No gráfico 4.16, no primeiro ano apenas os meses de novembro e dezembro estão ajustados. Temos problemas nos picos, e nos meses de junho, julho, agosto e setembro. No gráfico 4.17, temos, a partir do terceiro ano, grande desajuste nos meses de fevereiro, março e abril. A curva como um todo está bem desajustada.

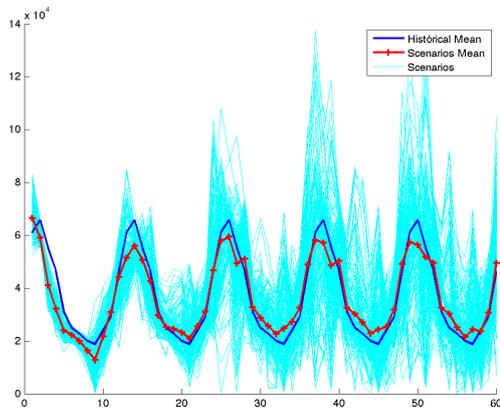


Figura 4.16: Envoltória teste, TL

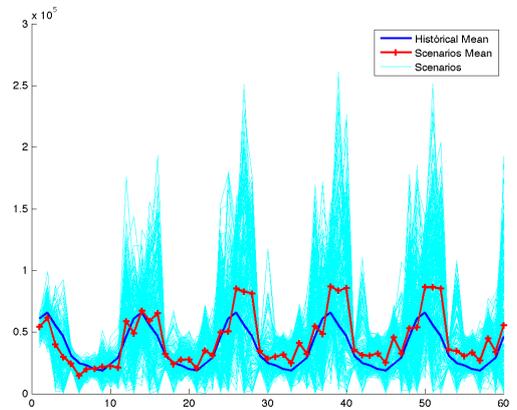


Figura 4.17: Envoltória validação, TL

4.3.3 Subsistema Nordeste

Os gráficos 4.18 e 4.19 são correspondentes às envoltórias de teste e avaliação do caso LP, subsistema nordeste. No gráfico 4.18 temos, com exceção do primeiro ano, um ajuste muito bom. Temos também problemas novamente nos picos e no primeiro janeiro. No gráfico 4.19 temos que o primeiro ano melhorou, mas o ajuste do gráfico como um todo piorou. Temos ainda os problemas nos picos, embora o problema do primeiro janeiro tenha sumido.

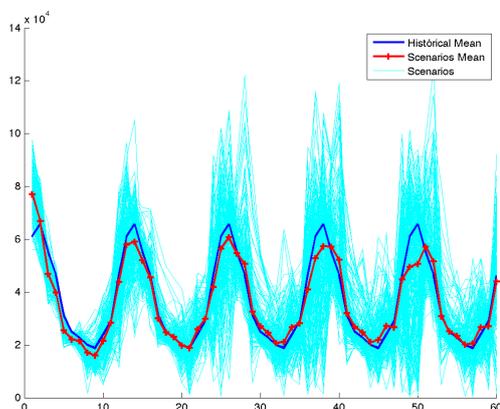


Figura 4.18: Envoltória teste, LP

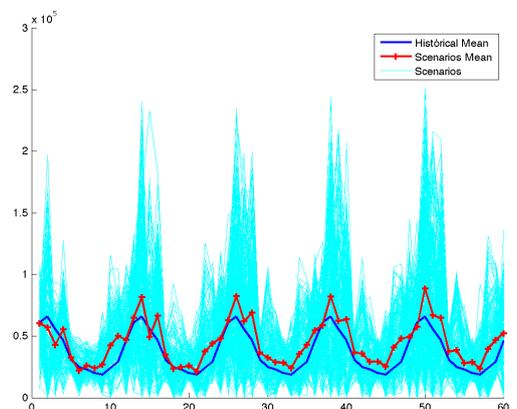


Figura 4.19: Envoltória validação, LP

Os gráficos 4.20 e 4.21 são correspondentes às envoltórias de teste e avaliação do caso TP, subsistema nordeste. No gráfico 4.20, temos a curva da média dos cenários com um bom ajuste, embora o mês de janeiro esteja desajustado, e alguns outros pontos tem desajustes menores. No gráfico 4.21, esse gráfico está explicitamente desajustado. O que é digno de nota é que o maior desajuste ocorre nos meses de abril, e todos os meses de maio e setembro estão corretos.

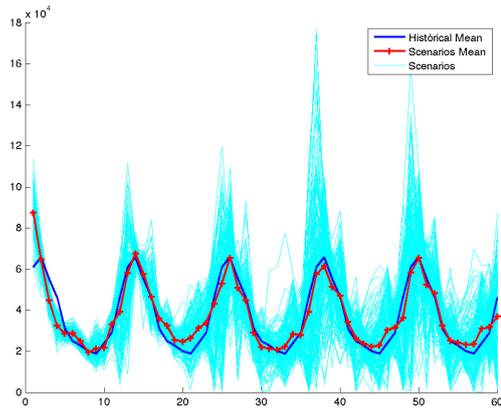


Figura 4.20: Envoltória teste, TP

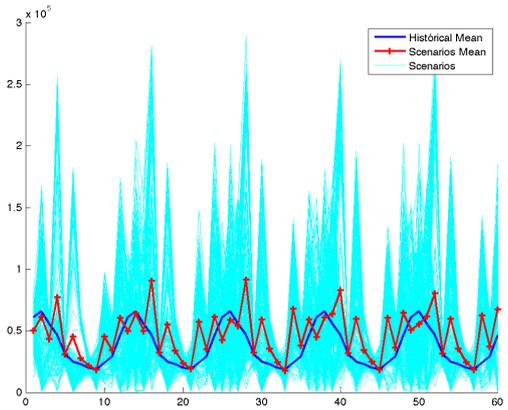


Figura 4.21: Envoltória validação, TP

Os gráficos 4.22 e 4.23 são correspondentes às envoltórias de teste e avaliação do caso TP, subsistema nordeste. No gráfico 4.22, podemos perceber que o gráfico está desajustado nos meses próximos aos picos (janeiro, fevereiro e março). No gráfico 4.23, o que percebemos é que o principal problema ocorre nos meses de abril e maio.

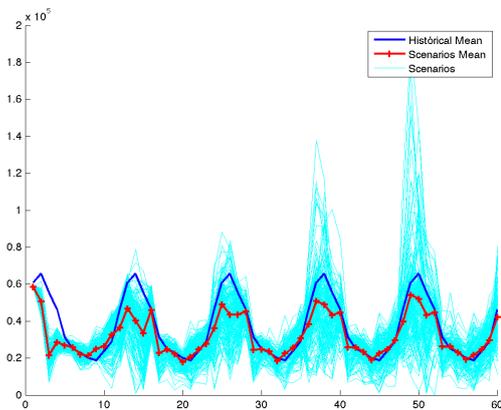


Figura 4.22: Envoltória teste, PP

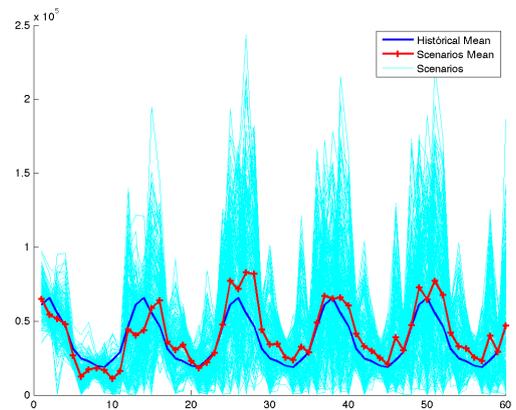


Figura 4.23: Envoltória validação, PP

Os gráficos 4.24 e 4.25 são correspondentes às envoltórias de teste e avaliação do caso TL, subsistema nordeste. No gráfico 4.24, a curva da média dos cenários está

acompanhando muito bem a média histórica. Vale notar que o primeiro janeiro está desajustado, e os picos também não estão. No gráfico 4.25, podemos ver que o ajuste da curva das médias dos cenários ainda está bom, com exceção dos meses de fevereiro e março.

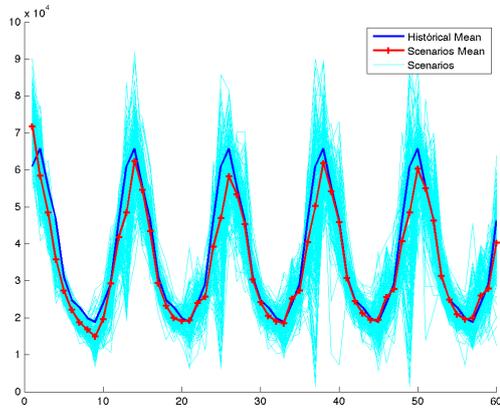


Figura 4.24: Envoltória teste, TL

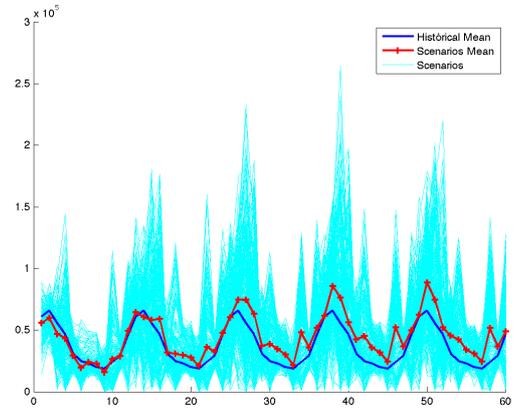


Figura 4.25: Envoltória validação, TL

4.3.4 Subsistema Norte

Os gráficos 4.26 e 4.27 são correspondentes às envoltórias de teste e avaliação do caso LP, subsistema norte. No gráfico 4.26, temos um bom ajuste da curva das médias dos cenários, com exceção do primeiro mês de janeiro e dos picos. No gráfico 4.27, continuamos com um bom ajuste, embora alguns meses de janeiro estejam com problemas. Os picos melhoraram.

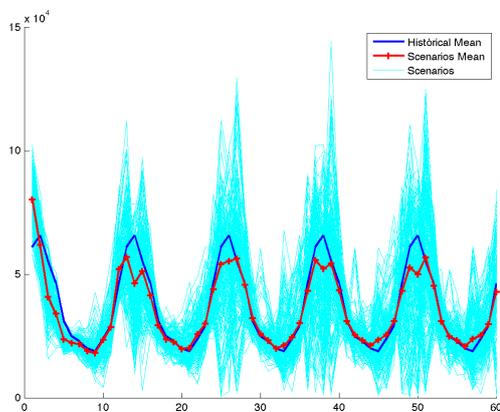


Figura 4.26: Envoltória teste, LP

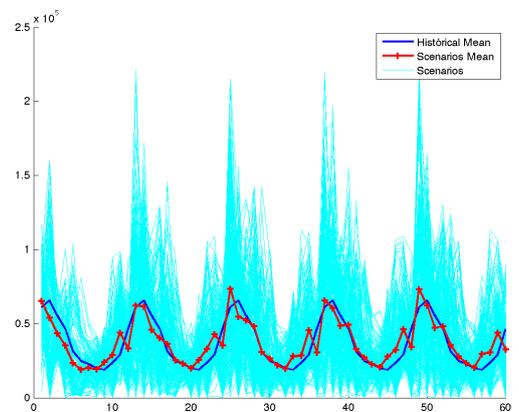


Figura 4.27: Envoltória validação, LP

Os gráficos 4.28 e 4.29 são correspondentes às envoltórias de teste e avaliação do

caso TP, subsistema norte. No gráfico 4.28, temos um ajuste razoável, mas novamente notamos que o primeiro mês de janeiro está bem desajustado. No gráfico 4.29, podemos perceber que a partir do terceiro ano, os meses de fevereiro e abril vão progressivamente ficando piores, assim como os vales também.

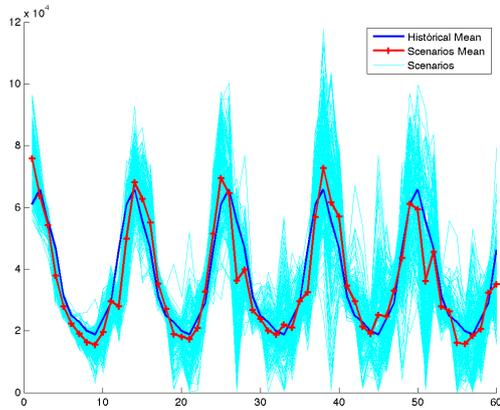


Figura 4.28: Envoltória teste, TP

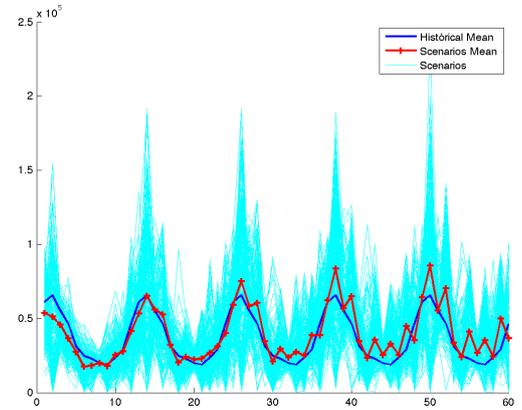


Figura 4.29: Envoltória validação, TP

Os gráficos 4.30 e 4.31 são correspondentes às envoltórias de teste e avaliação do caso PP, subsistema norte. No gráfico 4.30, notamos um ajuste muito bom, um dos melhores de todos. Destacamos que no primeiro ano esse ajuste não é tão bom, e novamente o mês de janeiro está desajustado. No gráfico 4.31, podemos ver que o ajuste da curva das médias dos cenários ainda é razoável, mas dessa vez temos problemas nos picos.

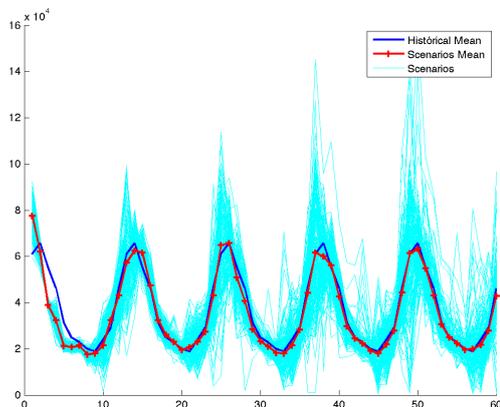


Figura 4.30: Envoltória teste, PP

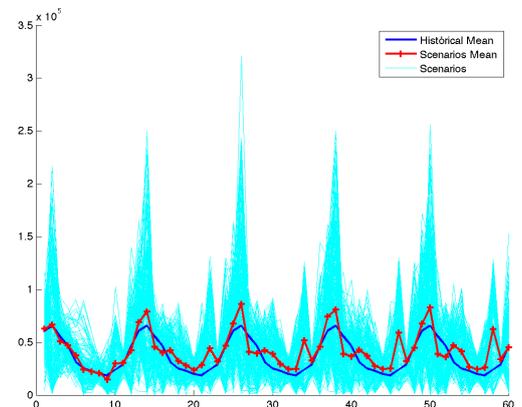


Figura 4.31: Envoltória validação, PP

Os gráficos 4.32 e 4.33 são correspondentes às envoltórias de teste e avaliação do caso TL, subsistema norte. No gráfico 4.32, podemos apontar o primeiro ano como o

pior, e os meses de janeiro e fevereiro como os que tem o pior ajuste. No gráfico 4.33, percebemos um ajuste bem razoável, com exceção dos meses de fevereiro, correspondentes aos picos.

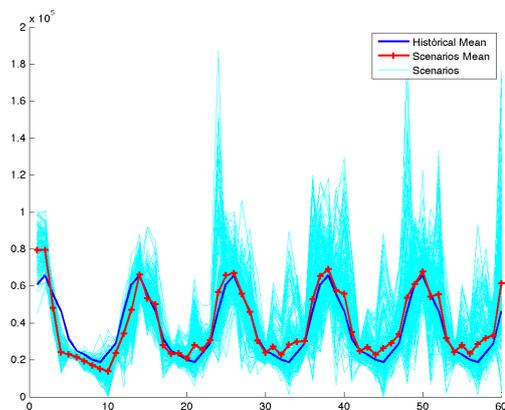


Figura 4.32: Envoltória teste, TL

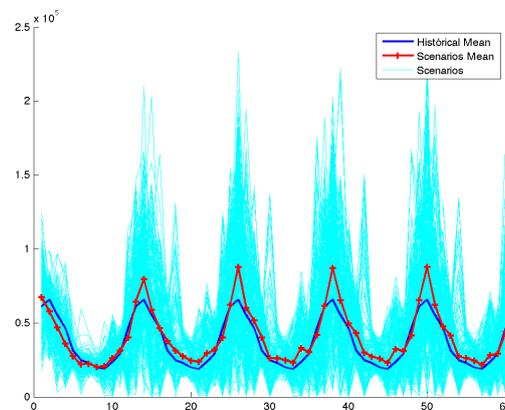


Figura 4.33: Envoltória validação, TL

5 Conclusão

Apesar das restrições da construção do modelo PEN neste trabalho (como limitar a ordem em 11, em vez de variar em 3, 6, 9 ou 11 no PEN original de (6)), foi obtida uma aderência boa nos testes t , Levene e K-S, competindo com o PEN original e mesmo com o PAR(p). Isso mostra que o fato de haver mais dados históricos na construção desse trabalho balanceou, de alguma forma, as restrições impostas. Nos resultados obtidos, o caso LP foi o que mais obteve resultados bons, possuindo o maior número de MAPEs menores nos subsistemas norte e nordeste. Vale destacar que o melhor caso do subsistema sul foi o caso PP, que corresponde a uma tentativa de linearizar o PEN, o que foi pensado que daria sempre resultados ruins, pois linearizar o PEN é tentar desfazer a característica não linear que ele possui, que é uma de suas vantagens. Mesmo assim, se fossemos forçados a escolher apenas um caso para tratar todos os subsistemas, esse seria o caso LP, pelo fato de ter sido melhor em dois subsistemas. Na inspeção visual dos cenários, notamos que houve diversos casos (principalmente na validação) em que os picos, correspondentes ao mês de fevereiro, e outros meses diversos, estavam bem distantes da curva de histórico. Especulando o motivo desses problemas como sendo a escolha da função de ruído, que superou os limites aceitáveis e por isso saiu da escala. Apesar disso, os gráficos tenderam a fazer um bom acompanhamento do histórico.

5.1 Trabalhos Futuros

Em relação a trabalhos futuros, há diversos estudos de caso a se fazer com o modelo PEN. Um deles seria aplicar o PEN em outra série temporal mensal. Também podem ser feitos mais tratamentos da série de ENA, inclusive variando a autoregressão, que foi fixa em 11 nesse trabalho. Um outro trabalho interessante seria a aplicação da teoria de *wavelets*, visando melhorar a aderência da série sintética com a série histórica.

Referências Bibliográficas

- [1]
- [2] Abril 2014.
- [3] Barros, M. **Processos Estocásticos**. Papel Virtual Editora, Julho 2004.
- [4] Bendre, S. Nonparametric statistics. **Unpublished manuscript, University of Hyderabad**, 2010.
- [5] Box, G. E.; Jenkins, G. M. ; Reinsel, G. C. **Time series analysis: forecasting and control**. Wiley.com, 2008.
- [6] Campos, L. C. D.; Vellasco, M. M. B. R. ; Lazo, J. G. L. **Modelo Estocástico Periódico baseado em Redes Neurais**. 2010. Tese de Doutorado - PUC-Rio.
- [7] de Leon F de Carvalho, A. P.; de Pádua Braga, A. ; Ludermir, T. B. **Fundamentos de Redes Neurais Artificiais**. 11a Escola de Computação, 1998.
- [8] Hagan, M.; Menhaj, M. Training feedforward networks with the marquardt algorithm. **IEEE Transactions on Neural Networks**, 1994.
- [9] Hyndman, R. J.; Koehler, A. B. Another look at measures of forecast accuracy. **International journal of forecasting**, v.22, n.4, p. 679–688, 2006.
- [10] Lilliefors, H. W. On the kolmogorov-smirnov test for normality with mean and variance unknown. **Journal of the American Statistical Association**, v.62, n.318, p. 399–402, 1967.
- [11] Maceira, M. E. P.; Penna, D. D. J. ; Damázio, J. M. Geração de cenários sintéticos de energia e vazão para o planejamento da operação energética. **XVI Simpósio Brasileiro de Recursos Hídricos, João Pessoa, Paraíba, Brasil**, p. 3–12, 2005.
- [12] Maceira, M. E. P.; Duarte, V.; Penna, D.; Moraes, L. ; Melo, A. Ten years of application of stochastic dual dynamic programming in official and agent studies in brazil–description of the newave program. **16th PSCC, Glasgow, Scotland**, p. 14–18, 2008.
- [13] McCulloch, W. S.; Pitts, W. A logical calculus of the ideas immanent in nervous activity. **The bulletin of mathematical biophysics**, v.5, n.4, p. 115–133, 1943.
- [14] de Medeiros, L. **Previsão do preço spot no mercado de energia elétrica**. 2004. Tese de Doutorado - Tesis para optar al grado de doctor. Pontifícia Universidade Católica do Rio de Janeiro. 2004.[Links].
- [15] Moser, B. K.; Stevens, G. R. Homogeneity of variance in the two-sample means test. **The American Statistician**, v.46, n.1, p. 19–21, 1992.
- [16] Maio 2014.
- [17] Abril 2014.

- [18] Janeiro 2014.
- [19] Papoulis, A. **Probability, Random Variables and Stochastic Processes**. McGraw-Hill, 1991.
- [20] Principe, J. C.; Euliano, N. R. ; Lefebvre, W. C. **Neural and Adaptive Systems: Fundamentals Through Simulations**. 1. ed., Wiley, 1999.
- [21] Priddy, K. L.; Keller, P. E. **Artificial neural networks: an introduction**. SPIE Publications, 2005.
- [22] Rosa, L. P. Geração hidrelétrica, termelétrica e nuclear. **Estudos Avançados**, v.21, n.59, p. 39, 2007.
- [23] Schacter, D. L.; Gilbert, D. T. ; Wegner, D. M. **Psychology**. 2nd. ed., Worth Publishers, 2011.
- [24] Junho 2014.
- [25] de Almeida, A.; Elian, S. ; Nobre, J. Modificações e alternativas aos testes de levene e de brown e forsythe para igualdade de variâncias e médias. **Revista Colombiana de Estadística**, v.31, n.2, p. 241–260, 2008.