

3

Processo Estocástico Neural

Como mencionado anteriormente, o objetivo deste trabalho é a criação de um modelo de processo estocástico que possa ser aplicado em séries temporais de comportamento estocástico e também em séries temporais de comportamento periódico das suas propriedades probabilísticas, como por exemplo, a média, a variância, entre outras. Um exemplo são as séries hidrológicas sazonais de intervalos de tempo mensais.

Decidiu-se construir um processo estocástico baseado em redes neurais artificiais (RNAs), denominado Processo Estocástico Neural (PEN). O uso de RNAs faz com que o PEN seja um modelo não-linear capaz de capturar as características da série temporal, sem necessitar de fazer suposições a priori sobre o comportamento da série ou realizar algum tipo de decomposição na mesma, como deixá-la estacionária ou retirar algumas características como ciclo ou tendência. Para isso é preciso que as entradas das RNAs do PEN apresentem uma memória de curto prazo, contendo termos passados da série temporal a ser simulada. Os parâmetros do modelo PEN correspondem aos pesos sinápticos das RNAs e para simular realizações estocásticas (séries sintéticas) é preciso adicionar um valor aleatório às saídas das RNAs. Esses valores aleatórios são obtidos através das distribuições de probabilidade dos resíduos das RNAs do PEN.

Para trabalhar com séries temporais periódicas, os parâmetros do modelo PEN devem se ajustar não só ao intervalo de tempo da série, mas também ao seu período. Por essa razão, o PEN é modelado com uma componente estocástica (CE) para cada período da série. Por exemplo, no caso de período mensal, o PEN é composto por 12 CEs - uma para cada mês; já no caso de período semestral o PEN é composto por 2 CEs e no caso da série não ser periódica o PEN é básico, contendo apenas uma CE. Cada componente estocástica (CE) do PEN é formada por uma RNA e por uma distribuição de probabilidade para gerar valores aleatórios na geração de cenários, como ilustra a Figura 3.1.

Quando o PEN é formado por mais de uma CE, ocorre um encadeamento

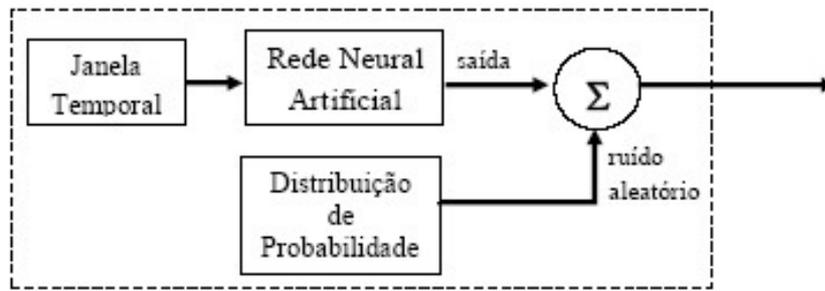


Figura 3.1: Componente Estocástica do Processo Estocástico Neural

entre elas, onde o valor da série fornecido pela CE de um período faz parte da janela temporal de entradas da RNA da CE do período seguinte, como ilustra a Figura 3.2.

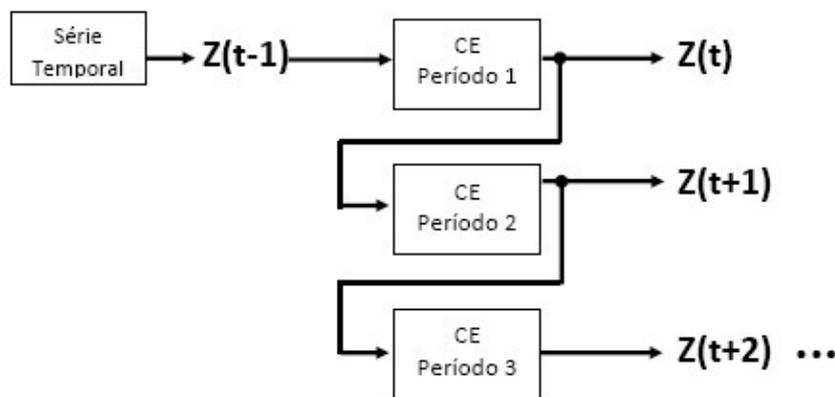


Figura 3.2: Encadeamento entre as CEs do Processo Estocástico Neural

Devido à sua estrutura, o PEN é classificado como um modelo estocástico periódico não-linear auto-regressivo genérico.

3.1 Descrição do Processo Estocástico Neural

Seja $Z(t)$ uma série temporal com período sazonal s e n observações simultâneas em todos os períodos. O índice de tempo t é descrito pela equação 3-1:

$$t = (r - 1) \cdot s + m \quad (3-1)$$

onde

- $r = 1 \dots n$ é o número da observação de cada período da série;
- $m = 1 \dots s$ corresponde a um período da série;

- $s \in \mathbb{N}$ é o total de períodos da série;
- $n \cdot s$ é o tamanho da série observada.

Por exemplo, no caso de uma série mensal, r corresponde ao ano, m corresponde ao mês, $s = 12$ e n corresponde ao total de anos que a série foi observada.

Ao modelar essa série, tem-se que o PEN é composto por s componentes estocásticas (CEs), uma para cada período m da série. Isso permite que o modelo possa ser descrito por cada CE m .

No estudo de caso desse trabalho utilizou-se na construção do PEN redes neurais *feedforward* MLP (*Multilayer Perceptron*) com uma camada oculta. Porém, isso não é uma restrição do modelo, isto é, o PEN suporta mais camadas ocultas e até outro tipo de topologia de RNA.

Conforme comentou-se, para que as RNAs *feedforward* se comportem como um modelo de processamento temporal, é preciso que elas apresentem habilidades de memória de curto prazo, que é feita através de técnicas de “janelamento” [24]. Essas técnicas consistem em introduzir memória nos neurônios da primeira camada escondida, fornecendo aos neurônios valores passados da série temporal analisada. Por isso o PEN é classificado como um modelo auto-regressivo, uma vez que as variáveis independentes ou entradas do modelo utilizam valores atrasados da variável dependente modelada.

A RNA de cada CE do PEN vai ter um número de termos passados da série na sua entrada, chamado de ordem da RNA. Neste trabalho, utiliza-se termos passados consecutivos da série, porém, o PEN permite o uso de termos passados não consecutivos. A ordem da RNA da CE do período m é representada por p_m . Assim, para obter um valor para a série no instante de tempo t , o PEN acessa a componente CE do período m correspondente e a sua RNA recebe os p_m termos auto-regressivos da série na camada de entrada. Ou seja, a fim de fornecer $Z(t)$, a RNA recebe como entrada os termos: $Z(t-1), Z(t-2), \dots, Z(t-p_m)$. Além disso, para reforçar o aprendizado do comportamento periódico da série é acrescentado na entrada o valor da série do período anterior, correspondente a $Z(t-s)$. A estrutura de uma RNA de ordem p_m é representada na Figura 3.3. Os *bias* e a descrição dos pesos foram omitidos de forma a simplificar a figura. A Figura 3.4 apresenta, em detalhes, um neurônio pertencente à camada oculta dessa RNA de ordem p_m , cuja saída é dada pela equação 3-2.

$$y_i = \varphi \left(\omega_{i,0} \cdot Z(t-s) + \left(\sum_{j=1}^{p_m} \omega_{i,j} \cdot Z(t-j) \right) + \theta_i \right) \quad (3-2)$$

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* desse neurônio.

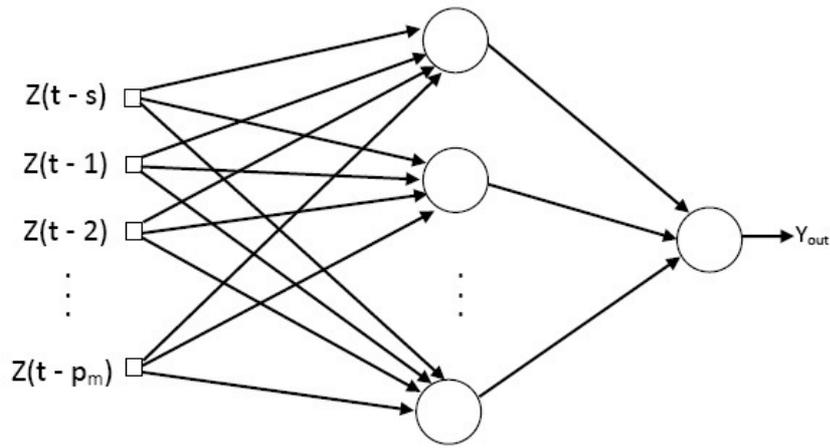


Figura 3.3: Rede Neural do PEN de ordem p_m

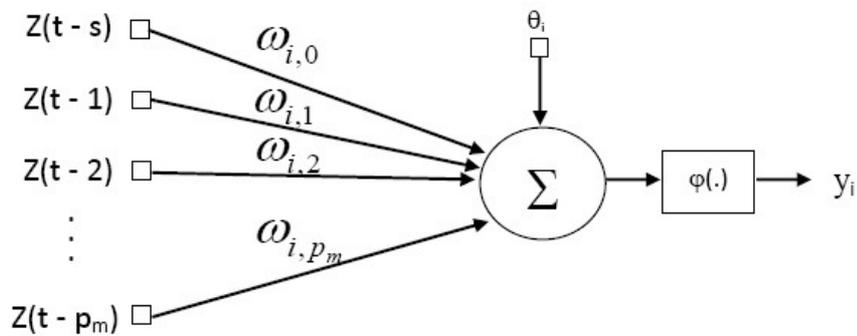


Figura 3.4: Neurônio da camada oculta de uma rede do PEN de ordem p_m

Considerando que a RNA de ordem p_m contém l_m neurônios na camada oculta, o neurônio da camada de saída é apresentado na Figura 3.5, sendo sua saída calculada pela equação 3-3.

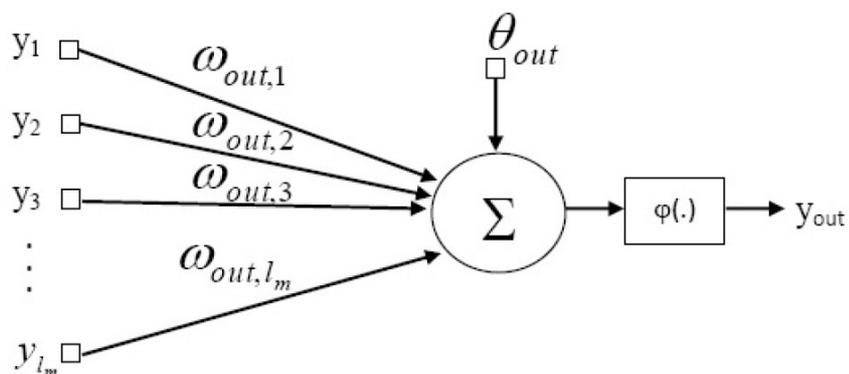


Figura 3.5: Neurônio de saída de uma rede do PEN com l_m neurônios na camada oculta

$$y_{out} = \varphi_{out} \left(\sum_{i=1}^{l_m} \omega_{out,i} \cdot y_i + \theta_{out} \right) \quad (3-3)$$

onde φ_{out} é a função de ativação do neurônio da camada de saída, representado por out , $\omega_{out,i}$ é o peso sináptico da conexão entre a entrada i (que corresponde à saída do neurônio i da camada oculta) e o neurônio out e θ_{out} é o *bias* desse neurônio.

Conforme ilustra a Figura 3.1, a saída de uma CE corresponde ao somatório da saída da RNA com um valor aleatório, proveniente da distribuição de probabilidade dos resíduos da RNA, obtidos da diferença entre a saída desejada e a saída da RNA (y_{out}). Assim, a série temporal $Z(t)$, cujo índice de tempo t é descrito pela equação 3-1, é simulada através da equação 3-4.

$$Z(t) = y_{out} + \alpha(t) \quad (3-4)$$

onde $\alpha(t)$ é o valor aleatório proveniente da distribuição de probabilidade dos resíduos da RNA da CE do período m correspondente, descrito posteriormente nesse capítulo.

Concatenando as equações 3-2, 3-3 e 3-4, chega-se à equação 3-5, a qual corresponde à descrição matemática da componente estocástica (CE) do período m do Processo Estocástico Neural (PEN).

$$Z(t) = \varphi_{out} \left(\sum_{i=1}^{l_m} \omega_{out,i} \cdot \varphi_i \left[\omega_{i,0} Z(t-s) + \left(\sum_{j=1}^{p_m} \omega_{i,j} Z(t-j) \right) + \theta_i \right] + \theta_{out} \right) + \alpha(t) \quad (3-5)$$

Os termos da série sintética $Z(t), Z(t+1), Z(t+2), \dots$ são simulados serialmente pelo PEN. Como foi comentado anteriormente, ocorre um encadeamento entre as CEs do PEN, onde a série gerada pela CE do período m serve de entrada para a RNA da CE do período $m+1$ e pode servir de entrada para a RNA da CE dos períodos $m+2, m+3, \dots$, ou seja, cada termo $Z(t)$ gerado corresponde a um termo passado da série a ser usado como entrada na geração dos termos posteriores ($Z(t+1), Z(t+2), \dots$). A Figura 3.6 ilustra, de forma genérica, o encadeamento das CEs do PEN, considerando um dado instante de tempo t .

Antes de estar apto a gerar séries sintéticas, é preciso ajustar o modelo PEN com a série temporal histórica a ser simulada. Esse ajuste deve seguir três etapas importantes:

1. Determinação da estrutura do modelo
2. Estimação dos seus parâmetros
3. Validação dos resíduos gerados

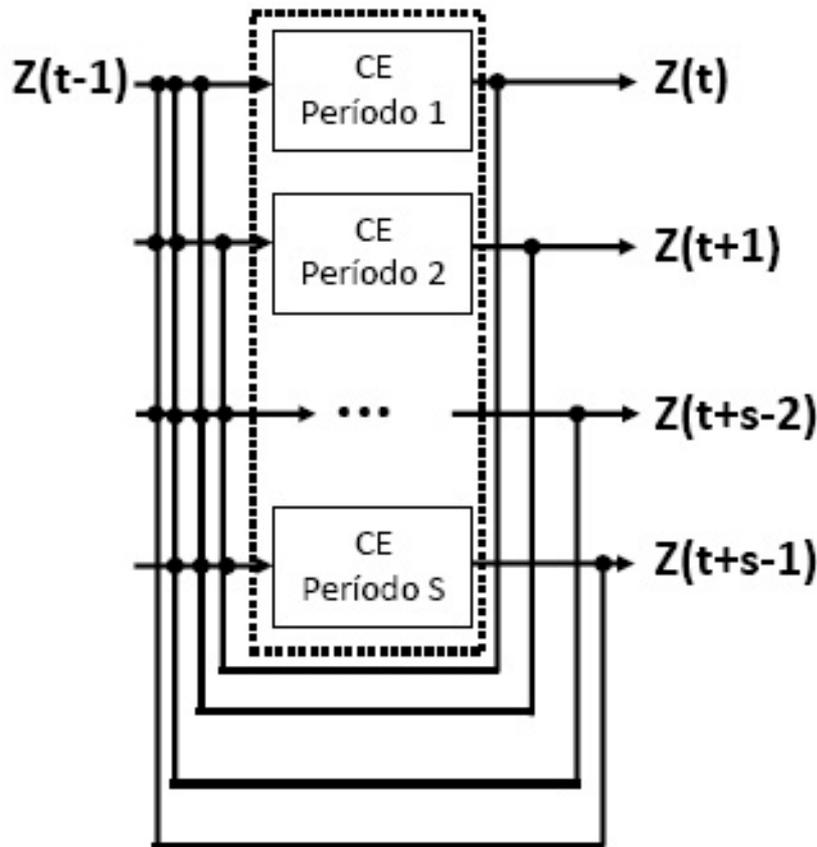


Figura 3.6: Encadeamento entre as entradas/saídas das Componentes Estocásticas do Processo Estocástico Neural

3.2

Determinação da estrutura do Processo Estocástico Neural

Nessa etapa define-se a arquitetura das redes neurais artificiais (RNAs) que fazem parte do modelo Processo Estocástico Neural (PEN). A definição da arquitetura de uma RNA consiste em determinar: o número de entradas, o número de camadas e de neurônios por camada, padrão de conexão entre os neurônios e a função de ativação dos neurônios.

Existem na literatura técnicas que tentam encontrar a arquitetura ideal através de um algoritmo genético, ou através de técnicas da diminuição de neurônios redundantes ou do compartilhamento de pesos, entre outras, porém, nenhuma conseguiu resultados convenientes para todas as aplicações [55]. De maneira geral, a escolha da arquitetura das RNAs é determinada principalmente pela experimentação sobre ajustamento e pela sensibilidade do projetista [1].

A escolha do número de entradas influencia diretamente o desempenho de uma RNA. Ao se utilizar uma quantidade muito grande de entradas em uma

RNA, maior será o número de parâmetros a serem ajustados, o que aumenta a complexidade do modelo, aumenta o tempo computacional do treinamento e, devido a inclusão de entradas pouco relevantes, diminui o desempenho do modelo [55]. Entretanto, uma quantidade muito pequena de entradas pode não conter informações suficientes para um bom aprendizado da RNA.

Outra decisão importante, que influencia o desempenho de uma RNA, é sobre o número de camadas ocultas, o número de neurônios em cada camada e a função de ativação dos neurônios da RNA [1].

No modelo PEN, a cada instante de tempo t definido pela equação 3-1, a janela temporal de entrada da RNA da componente estocástica (CE) do período m é composta pelos $t - j$ termos passados da série, onde $j = 1 \dots p_m$, mais o termo $t - s$ do período passado. A utilização do termo $t - s$ como entrada da RNA do período m correspondente, reforça à rede o aprendizado do período m . Por consequência, o número de entradas das RNAs é dado pela sua ordem + 1, ou seja, é igual a $p_m + 1$. Logo, a identificação do número de entradas consiste em determinar a ordem p_m , para cada período m da série a ser simulada pelo PEN.

A ordem de um modelo não-linear, como as RNAs, pode ser escolhida:

- através de informações conhecidas a priori;
- através de critérios que determinam a ordem, como por exemplo o critério PMI (*Partial Mutual Information*) [35], que determina os *lags* relevantes e não necessariamente consecutivos que forneçam a maior quantidade de informação ao sistema a partir de um número mínimo de entradas ou *lags*;
- de forma empírica, via tentativa e erro.

Como mencionado, no estudo de caso desse trabalho decidiu-se utilizar redes neurais *feedforward* MLP (*MultiLayer Perceptron* [58]). Como foi visto na Seção 2.4, essa topologia de RNA é não recorrente, portanto, não apresenta conexão de realimentação e nem conexão entre neurônios da mesma camada ou entre camadas não consecutivas. Segundo Haykin [28], uma única camada oculta é suficiente para uma RNA MLP computar aproximações de funções contínuas. Esse resultado só é válido para classes restritas de funções de ativação, por exemplo, funções sigmóides. Com base nisso, decidiu-se utilizar nesta aplicação uma única camada oculta e funções sigmóides como função de ativação dos neurônios da RNA. Porém, vale lembrar que o modelo PEN suporta RNAs com mais camadas ocultas e até outro tipo de topologia.

No PEN, cada RNA MLP apresenta um neurônio na camada de saída e o número de neurônios na camada oculta é determinado de forma empírica, isto

é, testa-se diferentes arquiteturas de RNAs, variando o número de neurônios na camada oculta.

Portanto, é possível definir diferentes estruturas de modelo PEN, determinando o número de neurônios na camada oculta e a ordem da RNA de cada CE. Seja l_m o número de neurônios da camada oculta e p_m a ordem da RNA da CE do período m . Tem-se que $m = 1, 2, \dots, s$, sendo s o número de períodos da série e, conseqüentemente, o número de CEs utilizadas no PEN. Seja p um vetor com a ordem de cada CE do PEN, logo, $p = (p_1, p_2, \dots, p_s)$, e seja l um vetor com o número de neurônios na camada escondida da RNA de cada CE do PEN, logo, $l = (l_1, l_2, \dots, l_s)$. Assim, o modelo de processo estocástico neural é referenciado por $\text{PEN}(p, l)$. A quantidade de parâmetros do modelo é a soma do número de parâmetros (número de pesos sinápticos, incluindo os *bias*, da RNA) de cada CE do PEN, que é dado por:

$$\sum_{m=1}^s p_m^{l_m} + 2 \cdot l_m + 1$$

A definição do modelo $\text{PEN}(p, l)$ consiste na identificação dos termos p e l , que podem ser determinados através de estudos preliminares sobre a série ou por tentativa e erro.

3.3

Estimação dos parâmetros do Processo Estocástico Neural

Nesta etapa faz-se a estimativa dos valores dos parâmetros do modelo processo estocástico neural (PEN), que correspondem aos pesos sinápticos das redes neurais artificiais (RNAs) das suas componentes estocásticas (CEs). Conforme descrito na Seção 2.4, as RNAs possuem uma regra de treinamento que ajusta estes pesos através de padrões de dados que lhe são apresentados. Existem diversos tipos de algoritmos de treinamento de redes neurais, os quais diferem entre si pela forma como esses pesos são modificados [28].

No modelo PEN, os pesos das RNAs precisam ser ajustados por um algoritmo de treinamento do tipo supervisionado. Como foi comentado na Seção 2.4, nesse tipo de algoritmo de treinamento os padrões de dados utilizados são formados pelo conjunto de entradas e pelo conjunto de saídas desejadas.

O algoritmo de treinamento é executado por um dado número de épocas, estipulado a priori, e em cada época de treinamento ocorre o ajuste dos pesos sinápticos da RNA. No modelo PEN, o ajuste dos parâmetros (pesos sinápticos de cada RNA) é realizado de maneira independente por período

m , $m = 1, \dots, s$. Isto é, em cada época as RNAs do PEN são treinadas individualmente, como ilustra a Figura 3.7. Isso permite que a etapa de estimativa dos parâmetros seja paralelizável, contudo, isso não foi aplicado no estudo de caso dessa tese.

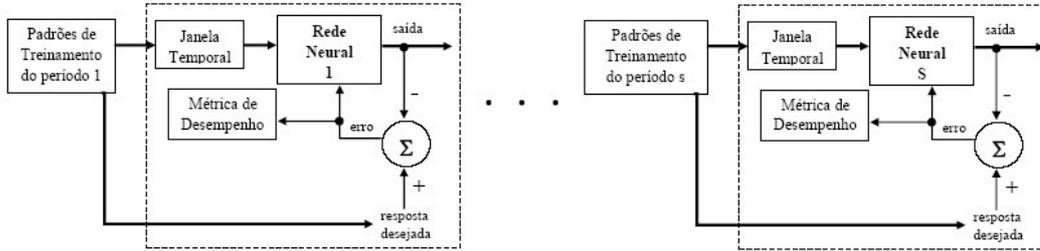


Figura 3.7: Treinamento das Redes Neurais do Processo Estocástico Neural

Para cada RNA é criado um conjunto de padrões de treinamento, com os valores das entradas e da saída desejada retirados da série histórica. Como foi comentado na Seção 2.4, para que o algoritmo de treinamento da RNA possa convergir de forma eficiente, os dados de entrada devem ser normalizados, dentro dos limites estabelecidos pela função de ativação dos neurônios.

Como o treinamento é supervisionado, durante a fase de treinamento a resposta do neurônio da camada de saída (representada por y_{out} na equação 3-3) no instante de tempo t é comparada com a resposta desejada vinda do padrão de dados. A diferença desses valores corresponde ao erro usado no ajuste dos pesos sinápticos, pelo algoritmo de treinamento, e no cálculo da medida de desempenho do treinamento.

Uma medida de desempenho muito utilizada para validar modelos de séries temporais é o erro médio percentual absoluto (MAPE) [63]. Decidiu-se então utilizar o MAPE como a métrica de desempenho do PEN. No treinamento da RNA do período m o MAPE é calculado através da equação 3-6.

$$MAPE = \frac{1}{N} \cdot \sum_{k=1}^N \left| \frac{Z(k) - Y(k)}{Z(k)} \right| \cdot 100 \quad (3-6)$$

onde N corresponde ao total de padrões e $Z(k)$ é o valor da k -ésima saída desejada do padrão de treinamento do período m . $Y(k)$ é a saída desnormalizada da RNA do período m para o k -ésimo padrão de entrada.

Também conforme descrito na Seção 2.4.1, após a passagem de uma época de treinamento, onde os pesos sinápticos foram ajustados, costuma-se fazer uma propagação com os dados de entrada do conjunto de validação, composto por dados diferentes dos utilizados no conjunto de treinamento.

Porém, diferente do treinamento, a etapa de validação no PEN não é feita individualmente por RNA. O objetivo da etapa de validação é gerar 1 cenário de $x \cdot s$ elementos, montado com as saídas seriais das RNAs de cada período, como ilustra a Figura 3.8. Dessa forma, essa etapa envolve todas as RNAs do PEN. O cenário de validação é composto pelas saídas das RNAs e é disparado com os últimos padrões de dados de entrada do conjunto de treinamento. A construção do cenário é feita de forma seqüencial, através do encadeamento entre as redes, da seguinte forma: a RNA $m = 1$ monta sua janela temporal e gera a sua saída, que consiste no primeiro elemento do cenário de validação a ser usado na janela temporal da RNA $m + 1$, e assim por diante. A Figura 3.9 ilustra, de forma simplificada, como é feito o encadeamento das RNAs para construir o cenário de validação.

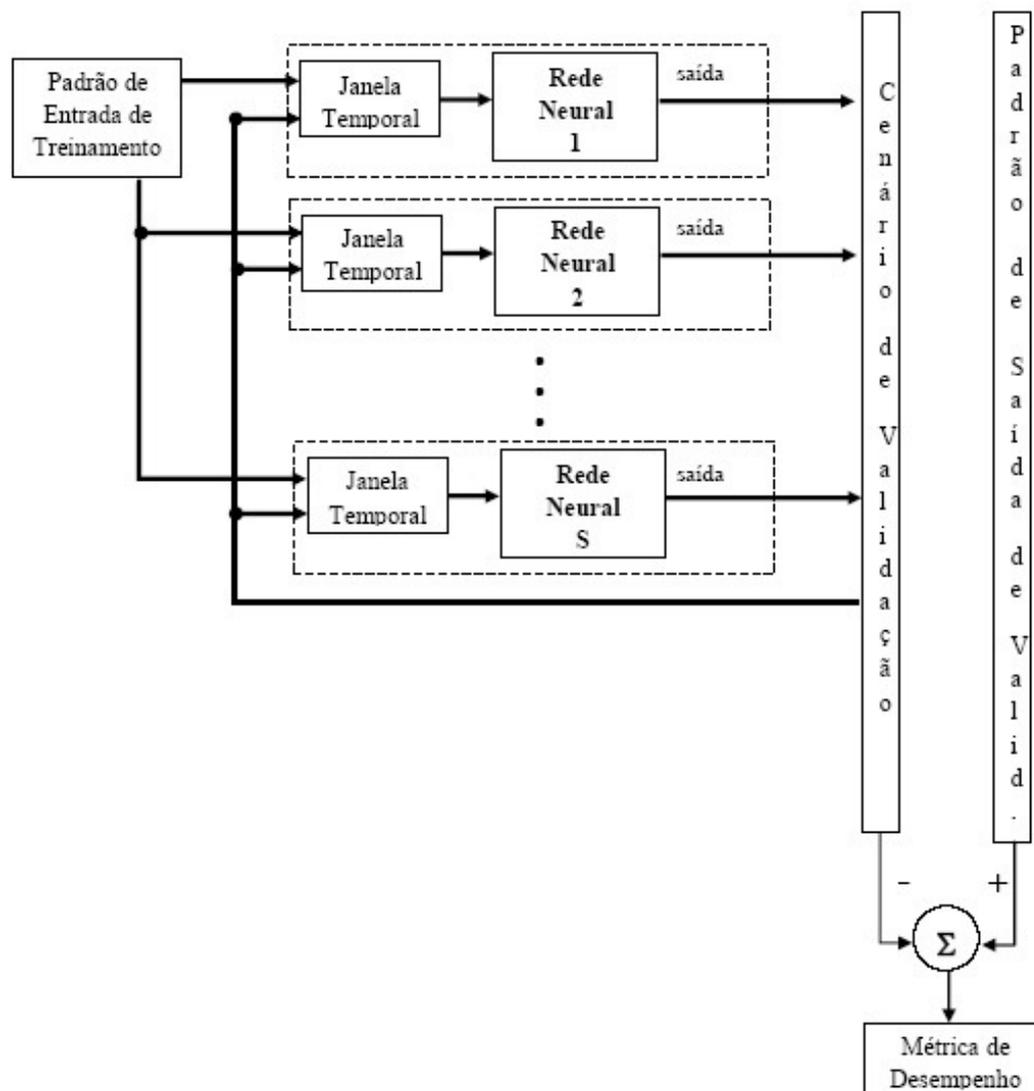


Figura 3.8: Validação das Redes Neurais do Processo Estocástico Neural

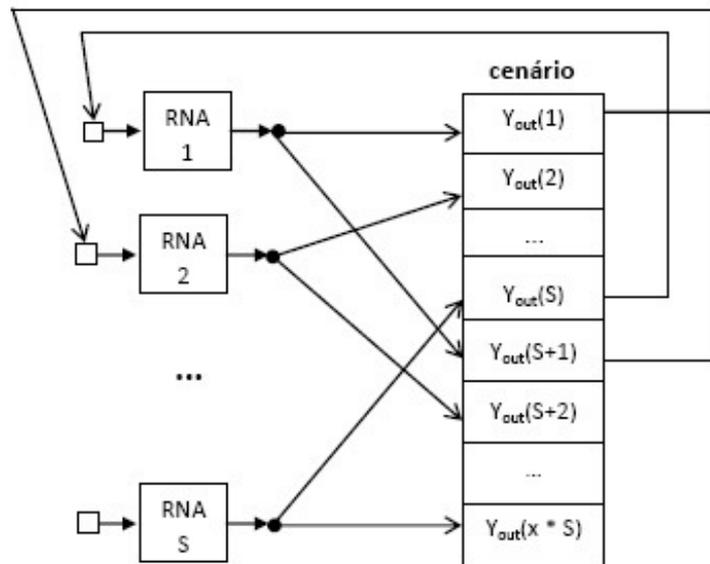


Figura 3.9: Construção do Cenário de Validação

É calculada uma métrica de desempenho, igual a utilizada no treinamento, mas referente aos dados do conjunto de validação. Para calcular a métrica de desempenho de validação, compara-se os dados do cenário com os dados que compõem o padrão de saída desejada do conjunto de validação. Dessa forma, é possível obter 2 tipos de métrica de desempenho de validação:

1. **Por cenário:** calcula-se a métrica percorrendo seqüencialmente todo o cenário.
2. **Por período:** compara-se os x valores do período m presentes no cenário com os x padrões de saída do conjunto de validação do período m .

Através do índice i , definido pela equação 3-7, é possível percorrer todo o cenário, representado aqui por C .

$$i = (k - 1) \cdot s + m \quad (3-7)$$

onde $k = 1 \dots x$ é o k -ésimo elemento de cada período no cenário, $m = 1 \dots s$ é o período e $s \in \mathbb{N}$ é o total de períodos considerados.

Logo, $C(i)$ corresponde ao i -ésimo elemento do cenário. A equação 3-8 apresenta o MAPE calculado por cenário.

$$MAPE = \frac{1}{x \cdot s} \cdot \sum_{i=1}^{x \cdot s} \left| \frac{Z(i) - C(i)}{Z(i)} \right| \quad (3-8)$$

onde $x \cdot s$ corresponde ao total de elementos do cenário e i é o índice definido pela equação 3-7, $Z(i)$ é a k -ésima saída desejada do padrão de validação do período m , sendo k e m definidos por i .

Já a equação 3-9, a seguir, apresenta o MAPE calculado individualmente por período.

$$MAPE = \frac{1}{x} \cdot \sum_{k=1}^x \left| \frac{Z(k) - C((k-1) \cdot s + m)}{Z(k)} \right| \quad (3-9)$$

onde x corresponde ao total de elementos do período m no cenário, k é o índice do elemento, s é o número de períodos da série, $Z(k)$ é a k -ésima saída desejada do padrão de validação do período m , C é o cenário e m é fixo.

Portanto, os treinamentos das RNAs do PEN ocorrem em paralelo, intercalando cada época de treinamento com uma época de validação que envolve todas as redes. Como há duas formas de calcular o MAPE de validação, é possível obter dois tipos de função de custo para avaliar a interrupção do treinamento:

1. **Em conjunto:** Utiliza-se o MAPE do cenário calculado na equação 3-8. Caso essa métrica comece a subir, interrompe o treinamento de todas as s RNAs do PEN na mesma época. A Figura 3.10 ilustra essa situação.
2. **Separado por rede neural:** Cada RNA m do PEN utiliza o seu MAPE calculado na equação 3-9. Caso o MAPE do período m comece a subir, a RNA m interrompe o seu treinamento individualmente. Assim as s RNAs do PEN podem interromper os seus treinamentos em épocas distintas. A Figura 3.11 ilustra essa situação.

3.4

Validação dos resíduos gerados

Durante a fase de treinamento da rede neural artificial (RNA) é calculado um conjunto de diferenças entre a saída dada pela rede e a saída desejada do padrão de treinamento. Ao término do treinamento, o conjunto de diferenças obtido pela RNA do período m corresponde à sua série de resíduos de estimativa.

Nessa etapa, busca-se ajustar uma distribuição de probabilidade teórica que possua uma boa aderência com a série de resíduos da RNA treinada para o período m ($m = 1, \dots, s$), pois através da distribuição de probabilidade teórica se obtém uma descrição aproximada das características desses resíduos.

A natureza específica de uma distribuição teórica é determinada pelos parâmetros da distribuição. Por exemplo, uma distribuição de probabilidade

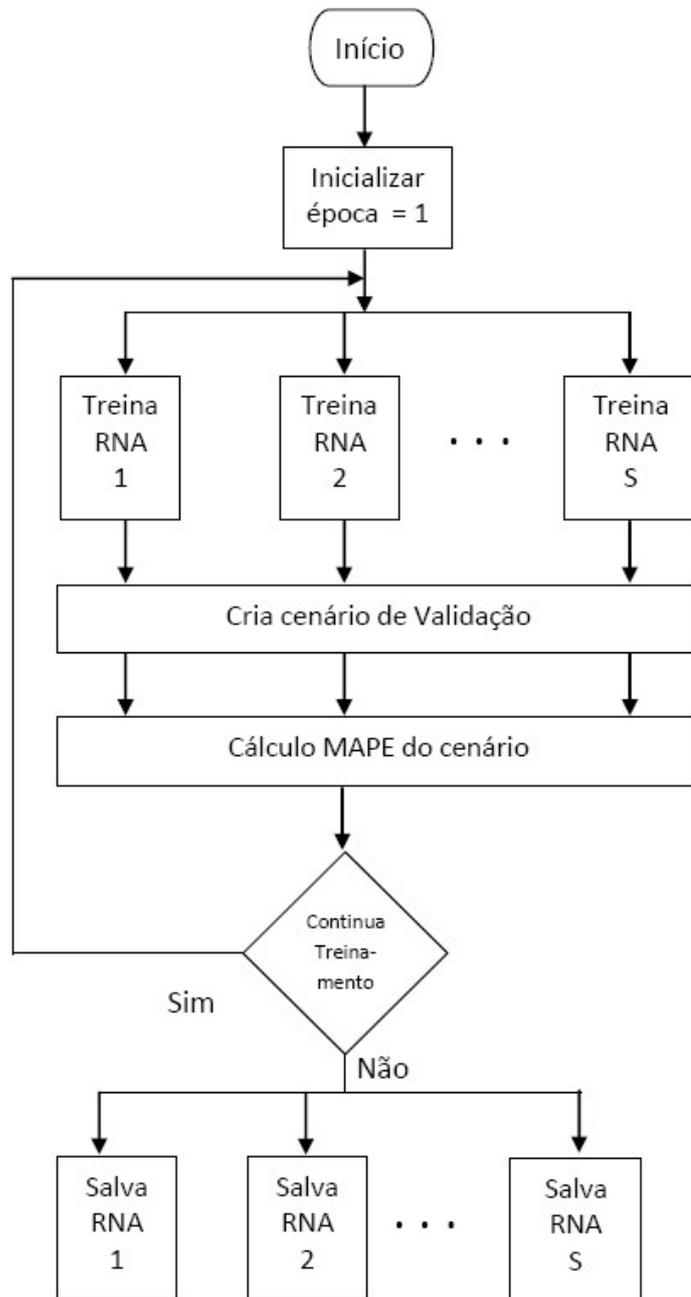


Figura 3.10: Interrupção do Treinamento Coletiva

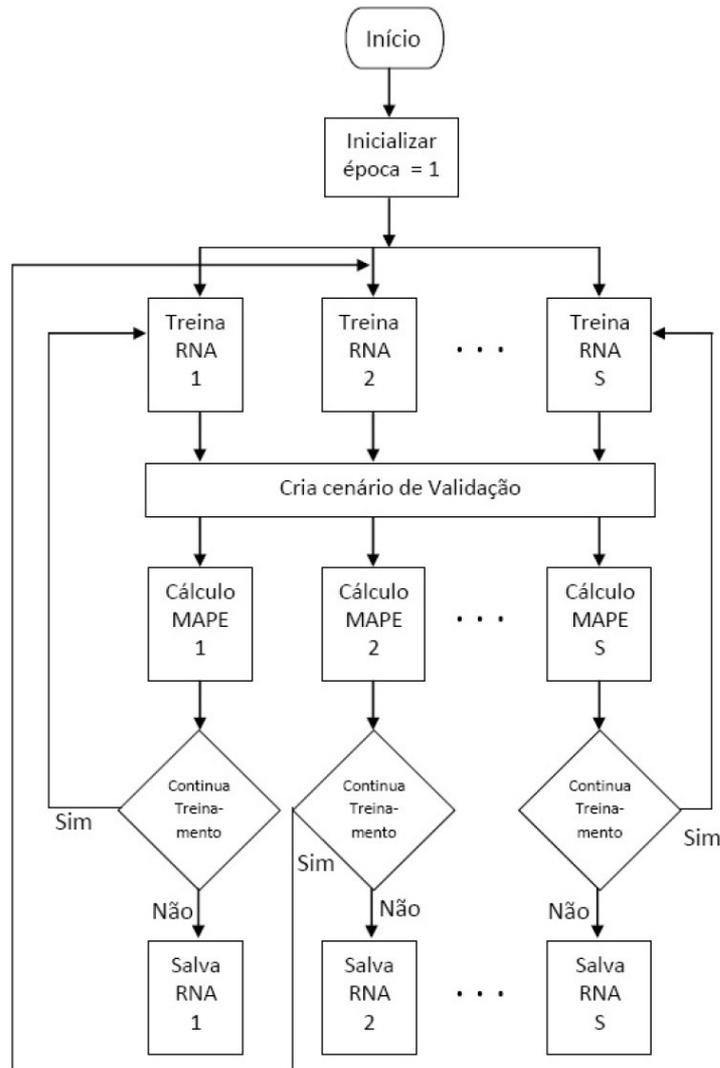


Figura 3.11: Interrupção do Treinamento Individual por RNA

normal é determinada pelos momentos média e variância [7]. Utiliza-se o teste de Kolmogorov-Smirnov (vide Seção 2.3.2), que é o teste indicado para se verificar a aderência da distribuição de probabilidade teórica à série de resíduos de um período $m = 1, \dots, s$. Escolhe-se a distribuição de probabilidade que conseguir o menor erro de ajuste através do teste de Kolmogorov-Smirnov. Essa corresponderá à distribuição de probabilidade da componente estocástica (CE) do período m do Processo Estocástico Neural (PEN), como ilustra a Figura 3.1.

Conforme especificado na equação 3-4, para a criação de um cenário, a série sintética a ser gerada no instante de tempo t (definido pela equação 3-1), depende de uma parcela estocástica, dada por um valor aleatório $\alpha(t)$. Esse valor $\alpha(t)$ é obtido através de um sorteio da distribuição de probabilidade da

CE do período m correspondente.

Com o intuito de reduzir a variância das amostras geradas por cada distribuição do PEN, decidiu-se utilizar o método de amostragem Quase Monte Carlo com a seqüência de Sobol (apresentado na Seção 2.2). Dessa forma, os valores aleatórios gerados pela distribuição de probabilidade de uma CE abrangem, o mais uniformemente possível, o espaço amostral dessa distribuição. Assim, será possível obter valores aleatórios mais equidistantes, o que evita a geração de cenários muito divergentes.

Uma vez ajustados os parâmetros da RNA e determinada a distribuição de probabilidade de cada CE do modelo PEN, é possível gerar diferentes cenários da série temporal analisada.

3.5

Geração de cenários através do processo estocástico neural

O processo estocástico neural (PEN) deve capturar o comportamento estocástico da série histórica, sendo capaz de gerar, artificialmente, séries sintéticas igualmente prováveis à série histórica.

A equação 3-5 é adequada para a geração de séries sintéticas, pois possibilita um esquema de geração seqüencial, onde a cada instante de tempo a série é obtida como uma função de seus valores passados (termos auto-regressivos) e de um termo aleatório. No início do processo de geração de séries sintéticas é preciso fornecer os primeiros valores dos termos auto-regressivos. Esses valores iniciais são tirados diretamente da série histórica original.

O PEN é composto por s componentes estocásticas (CEs), uma para cada período m da série. Como foi comentado anteriormente, quando a série é periódica, $s > 1$, na geração de cenários ocorre o encadeamento entre as CEs: o valor da série sintética gerada pela CE do período m consiste em uma das entradas que compõem a janela temporal da RNA das CEs dos períodos $m + 1$, $m + 2$, etc. A Figura 3.12 ilustra o diagrama de blocos do PEN na geração de cenários. As setas grossas é que fazem parte do fluxo de informações na geração de cenários.

Para cada cenário de séries sintéticas a ser gerado, defini-se o número de anos n do cenário e executa-se seqüencialmente, a cada instante de tempo t ($t = 1, 2, \dots, n \cdot s$), a equação 3-5.

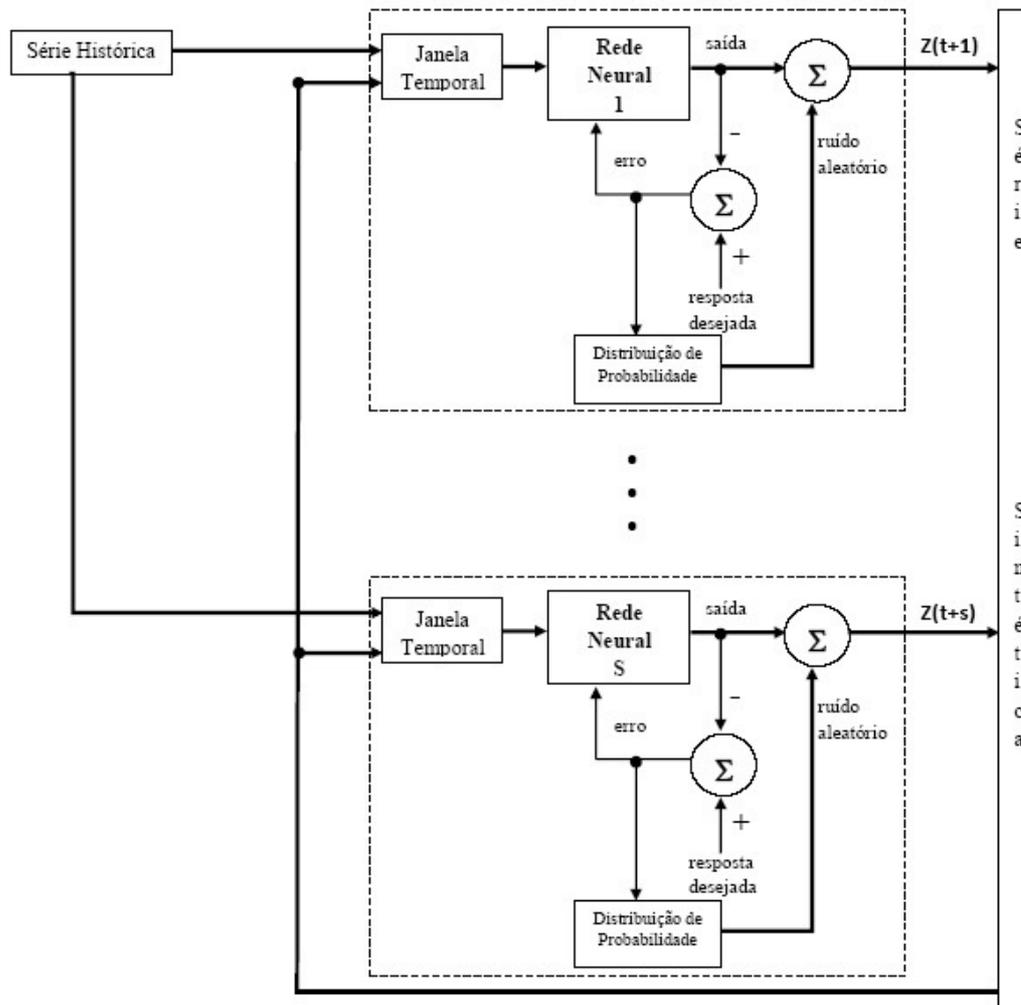


Figura 3.12: Geração de cenários no PEN

3.6 Resumo

Nesse capítulo foi apresentada a descrição genérica do modelo proposto nessa tese, o Processo Estocástico Neural. O próximo capítulo apresenta a modelagem do Processo Estocástico Neural para ser utilizado na geração de cenários de afluências mensais.